AMM 2025

CONFERENCE PROCEEDINGS

The 29th Annual Meeting in Mathematics

การประชุมวิชาการทางคณิตศาสตร์ ครั้งที่ 29

Lead a better future with mathematics

21 – 23 พฤษภาคม 2568

ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)

พิมพ์ครั้งแรก กรกฎาคม พุทธศักราช 2568
ลิขสิทธิ์ของภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

ข้อมูลทางบรรณานุกรมของหอสมุดแห่งชาติ

มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะวิทยาศาสตร์. ภาควิชาคณิตศาสตร์.
Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025).-- กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ, 2568.
788 หน้า.

1. คณิตศาสตร์--การประชุม. I. ชื่อเรื่อง.

510
ISBN : 978-616-296-315-5

จัดทำโดย

ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
114 สุขุมวิท 23
แขวงคลองเตยเหนือ เขตวัฒนา
กรุงเทพมหานคร 10110

คณะผู้จัดทำ

บรรณาธิการ	รองศาสตราจารย์ ดร.วรานุช แขมมณี
	อาจารย์กาญจนา พานิชการ
บรรณาธิการฝ่ายจัดทำ	ผู้ช่วยศาสตราจารย์ ดร.ธันวา ธีระกาญจน์
	ผู้ช่วยศาสตราจารย์ ดร.วิศรุต โพธิ์อ้น
ผู้ออกแบบปก	ผู้ช่วยศาสตราจารย์ ดร.สุชาดา พงษ์ประเสริฐ

หมายเหตุ

ความคิดเห็นและทัศนะที่ปรากฏในบทคัดย่อและบทความวิจัยเป็นของผู้เขียนแต่เพียงผู้เดียว ไม่จำเป็นต้องสะท้อนถึงความคิดเห็นหรือทัศนะของบรรณาธิการ คณะผู้จัดทำ ภาควิชาคณิตศาสตร์ หรือบุคลากรของภาควิชาคณิตศาสตร์แต่อย่างใด

AMM
2025

Conference Proceedings

การประชุมวิชาการทางคณิตศาสตร์ ครั้งที่ 29 ประจำปี 2568
The Annual Meeting in Mathematics 2025

21–23 พฤษภาคม 2568

ศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย
ร่วมกับ
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ

สารจากอธิการบดีมหาวิทยาลัยศรีนครินทรวิโรฒ

มหาวิทยาลัยศรีนครินทรวิโรฒ มีวิสัยทัศน์ในการพัฒนาคุณภาพการศึกษาให้ก้าวหน้าในระดับสากล มุ่งผลิตบัณฑิตที่พร้อมทั้งด้านวิชาการและเป็นผู้ที่สร้างคุณค่าแก่สังคม ด้วยแนวคิด “Learning University for Society” ภายใต้สโลแกน LOVES SWU Growth for All โดย LOVES มาจาก

L: Learning มุ่งสู่ความเป็นเลิศทางวิชาการ มีแพลตฟอร์มเพื่อการเรียนรู้ตลอดชีวิต ร่วมมือกับภาครัฐและภาคเอกชนทั้งในประเทศและต่างประเทศเพื่อยกระดับการเรียนรู้ รวมถึงเป็นศูนย์กลางทางวิชาการเพื่อสังคม

O: Opportunity โอกาสสำหรับทุกคน ส่งเสริมและสนับสนุนนิสิตด้วยการมอบทุนการศึกษาให้สวัสดิการต่าง ๆ ผลักดันนิสิตสู่การเป็นพลเมืองที่ดีของโลก

V: Value Added ยกระดับคุณค่าขององค์กร มุ่งสู่การเป็นพื้นที่ของการเรียนรู้แห่งอนาคต ส่งเสริมการนำองค์ความรู้มาสร้างนวัตกรรมอย่างมืออาชีพ พัฒนางานวิจัยที่ตอบโจทย์รัฐบาลและทิศทางของโลก

E: Environment สนับสนุนแนวคิดมหาวิทยาลัยเป็นพื้นที่ของทุกคน เพื่อทุกคน ปรับภูมิทัศน์ให้ตอบโจทย์การใช้งานและรองรับการใช้พื้นที่สำหรับผู้คนหลากหลาย เพิ่มศักยภาพการเป็นมหาวิทยาลัยสีเขียวที่คำนึงถึงสุขภาวะของทุกคน และ

S: Social Engagement ยกระดับความร่วมมือเพื่อสังคม ภายใต้แนวคิด มศว มหาวิทยาลัยเพื่อสังคม

มหาวิทยาลัยศรีนครินทรวิโรฒ โดยภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ รู้สึกเป็นเกียรติที่ได้รับความไว้วางใจให้เป็นเจ้าภาพจัดการประชุมวิชาการทางคณิตศาสตร์ ครั้งที่ 29 ประจำปี 2568 (The 29th Annual Meeting in Mathematics หรือ AMM 2025) ภายใต้หัวข้อ “Lead a Better Future with Mathematics” การประชุมครั้งนี้เป็นเวทีสำคัญสำหรับการนำเสนอผลงานวิจัยด้านคณิตศาสตร์ สถิติ คณิตศาสตร์ประยุกต์ สถิติประยุกต์ วิทยาการข้อมูล คณิตศาสตรศึกษา และสาขาที่เกี่ยวข้อง โดยเปิดโอกาสให้นักวิจัยทั้งในประเทศและต่างประเทศเผยแพร่งานวิจัย แลกเปลี่ยนองค์ความรู้ และร่วมกันผลักดันการนำคณิตศาสตร์ไปประยุกต์ใช้ในการพัฒนาคุณภาพชีวิต ซึ่งการจัดงาน AMM 2025 ไม่เพียงส่งเสริมคุณภาพการศึกษาและการพัฒนางานวิจัยเท่านั้น แต่ยังช่วยต่อยอดองค์ความรู้ให้สามารถนำไปใช้ประโยชน์ต่อองค์กร สังคม และประเทศชาติ ทั้งนี้ยังสอดคล้องกับวิสัยทัศน์ของมหาวิทยาลัยศรีนครินทรวิโรฒในการขับเคลื่อนองค์ความรู้ เพื่อการพัฒนาอย่างยั่งยืน

มหาวิทยาลัยศรีนครินทรวิโรฒหวังเป็นอย่างยิ่งว่า AMM 2025 จะเป็นอีกก้าวสำคัญในการสร้างเครือข่ายความร่วมมือทางวิชาการที่มีคุณภาพ พร้อมทั้งส่งเสริมการประยุกต์ใช้คณิตศาสตร์ และองค์ความรู้ทางวิทยาศาสตร์ เพื่อประโยชน์สูงสุดแก่สังคมและประเทศชาติในอนาคต ในนามของมหาวิทยาลัยศรีนครินทรวิโรฒ ขอขอบคุณวิทยากรและผู้เข้าร่วมเสนอผลงานวิจัยทุกท่านที่ให้เกียรติมาร่วมบรรยายและถ่ายทอดความรู้แก่ผู้ร่วมงาน รวมถึงผู้เข้าร่วมประชุมทุกท่านที่มีส่วนร่วมในการแลกเปลี่ยนความรู้ทางวิชาการ ซึ่งเป็นการสร้างสรรค์โอกาสใหม่ ๆ ในการพัฒนาความรู้และแนวคิดด้านคณิตศาสตร์และวิทยาศาสตร์ ขอบคุณคณะกรรมการและผู้เกี่ยวข้องทุกฝ่ายที่มีบทบาทสำคัญในการจัดงาน AMM 2025 ให้เป็นไปอย่างราบรื่นและมีประสิทธิภาพ จนเกิดเป็นกิจกรรมที่ทรงคุณค่า และเป็นประโยชน์อย่างยิ่งต่อการพัฒนางานวิจัยและการศึกษาทางคณิตศาสตร์และสาขาวิชาที่เกี่ยวข้อง



(รองศาสตราจารย์ ดร.ชลวิทย์ เจียรจิตต์)
อธิการบดีมหาวิทยาลัยศรีนครินทรวิโรฒ

สารจากนายกสมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์

ในนามของสมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์ ดิฉันขอแสดงความยินดีเป็นอย่างยิ่งที่การจัดประชุมวิชาการคณิตศาสตร์ครั้งที่ 29 ประจำปี 2568 (AMM 2025) ภายใต้หัวข้อ “Lead a Better Future with Mathematics” สามารถดำเนินการไปได้ด้วยดี โดยความร่วมแรงร่วมใจของบุคลากรภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒเป็นสำคัญ การประชุมวิชาการประจำปี หรือ AMM ของชาวคณิตศาสตร์จากทั่วประเทศ ที่ศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย (CEPMART) ภายใต้สมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์ และมหาวิทยาลัยต่าง ๆ ทั่วประเทศ มีส่วนร่วมกันจัดขึ้นนั้น มีความมุ่งหมายหลักเพื่อให้เป็นเวทีในการนำเสนอผลงานวิจัยใหม่ ๆ ทางคณิตศาสตร์ สถิติ และคณิตศาสตรศึกษา รวมถึงงานวิจัยในสาขาอื่น ๆ ที่เกี่ยวข้อง อีกทั้งยังเป็นเวทีในการแลกเปลี่ยนเรียนรู้ความก้าวหน้าทางวิชาการและการวิจัยด้านคณิตศาสตร์แขนงต่าง ๆ โดยมีกำหนดจัดประชุมเป็นประจำทุกปี

ในการประชุมครั้งนี้ มีการบรรยายพิเศษ และการเสวนาทางวิชาการโดยวิทยากรผู้ทรงคุณวุฒิจากทั้งในและต่างประเทศ สมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์ หวังเป็นอย่างยิ่งว่าคณาจารย์ นักวิจัย นิสิต นักศึกษา และผู้ที่สนใจคณิตศาสตร์ ตลอดจนศาสตร์ที่เกี่ยวข้องที่ได้เข้าร่วมประชุมวิชาการในครั้งนี้ ได้รับประโยชน์และนำประสบการณ์ที่ได้รับไปพัฒนาและต่อยอดองค์ความรู้ อีกทั้งนำไปถ่ายทอดให้แพร่หลายต่อไปด้วย

สมาคมคณิตศาสตร์แห่งประเทศไทยในพระบรมราชูปถัมภ์ ขอขอบคุณภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ที่ให้เกียรติเป็นเจ้าภาพการจัดการประชุม AMM ครั้งนี้ และขอขอบคุณทุกท่านที่มีส่วนร่วมในการจัดประชุมด้วยความวิริยะอุตสาหะยิ่ง

สุดท้ายนี้ ขอให้การประชุมวิชาการครั้งนี้ เป็นดั่งสะพานเชื่อมไปสู่ความร่วมมือระหว่างนักคณิตศาสตร์ทุกแขนงจากสถาบันและองค์กรต่าง ๆ ทั่วประเทศ เพื่อนำไปสู่การประยุกต์ใช้ความรู้ ความสามารถไปพัฒนานวัตกรรมอันทันสมัยต่าง ๆ ทำให้ประเทศของเราก้าวไปสู่อนาคตที่ดีกว่าด้วยคณิตศาสตร์อย่างรุ่งเรืองและยั่งยืนสืบไป



(ศาสตราจารย์กิตติคุณ ดร.พัฒนี อุดมกะวานิช)
นายกสมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์

สารจากผู้อำนวยการศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย

การจัดประชุมวิชาการทางคณิตศาสตร์ประจำปี (Annual Meeting in Mathematics) เป็นการประชุมที่สำคัญของประชาคมชาวคณิตศาสตร์ในประเทศไทย ที่จะได้มีโอกาสมาพบปะกันเพื่อฟังการบรรยายจากผู้เชี่ยวชาญสาขาต่าง ๆ ทางคณิตศาสตร์และสาขาที่เกี่ยวข้อง เพื่อให้เกิดการตระหนักรู้ในความก้าวหน้าและวิทยาการใหม่ ๆ รวมไปถึงการนำคณิตศาสตร์ไปใช้ในด้านต่าง ๆ และยังเป็นเวทีให้นักวิจัย ทั้งรุ่นเก่าและรุ่นใหม่ได้นำเสนอผลงาน เพื่อแลกเปลี่ยนเรียนรู้กับผู้ที่สนใจที่อยู่ต่างสถาบันกัน อันจะนำไปสู่ความร่วมมือทางด้านวิชาการและงานวิจัยต่อไปในอนาคต

สถาบันที่เปิดสอนหลักสูตรคณิตศาสตร์ในประเทศไทยได้ร่วมมือหมุนเวียนกันเป็นเจ้าภาพจนถึงครั้งนี้ โดยภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ได้รับเป็นเจ้าภาพจัดการประชุมวิชาการทางคณิตศาสตร์ ครั้งที่ 29 ประจำปี 2568 ระหว่างวันที่ 21–23 พฤษภาคม 2568 ในหัวข้อ “Lead a Better Future with Mathematics” ทั้งนี้ การจัดประชุมวิชาการทางคณิตศาสตร์ จะเป็นเวทีสำหรับนักวิจัยทั่วประเทศได้เผยแพร่งานวิจัย และแลกเปลี่ยนองค์ความรู้ เพื่อจะนำไปสู่การประยุกต์ใช้คณิตศาสตร์ในการยกระดับคุณภาพชีวิตให้ดีขึ้นในอนาคต

ในนามของผู้อำนวยการศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย (CEPMART) ขอขอบคุณคณะทำงานภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒเป็นอย่างยิ่ง ที่เป็นเจ้าภาพการจัดการประชุมวิชาการทางคณิตศาสตร์ครั้งนี้ และขออวยพรให้การจัดงานครั้งนี้สำเร็จราบรื่นด้วยดีทุกประการ ขอให้คณะทำงานทุกท่านมีสุขภาพแข็งแรงทั้งกายและใจ มีกำลังใจในการทำงาน ก้าวผ่านปัญหาและอุปสรรคต่าง ๆ อย่างราบรื่นด้วยดี

ศจี เพียรสกุล

(รองศาสตราจารย์ ดร.ศจี เพียรสกุล)
ผู้อำนวยการศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย

สารจากคณบดีคณะวิทยาศาสตร์

คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ มุ่งมั่นสู่การเป็นสถาบันชั้นนำระดับประเทศ ที่ขับเคลื่อนอนาคตด้วยวิทยาศาสตร์และเทคโนโลยี โดยมีพันธกิจสำคัญ 3 ด้าน ได้แก่ ด้านที่ 1 ส่งเสริมการเรียนรู้วิทยาศาสตร์และเทคโนโลยีผ่านสภาพแวดล้อมการเรียนรู้แบบดิจิทัลที่สร้างสรรค์ ด้านที่ 2 พัฒนางานวิจัยแบบบูรณาการผ่านระบบนิเวศการวิจัยที่มุ่งเน้นนวัตกรรม และด้านที่ 3 แสดงศักยภาพความเชี่ยวชาญด้านวิทยาศาสตร์และเทคโนโลยีผ่านการบริการวิชาการแก่ชุมชนและสังคม ในการประชุมวิชาการทางคณิตศาสตร์ระดับชาติ ครั้งที่ 29 ประจำปี 2568 หรือ AMM 2025 ภายใต้หัวข้อ “Lead a Better Future with Mathematics” ซึ่ง ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ได้รับเกียรติเป็นเจ้าภาพ มีเป้าหมายสำคัญในการเป็นเวทีวิชาการสำหรับการเผยแพร่งานวิจัยในสาขาคณิตศาสตร์ สถิติ และคณิตศาสตรศึกษา ตลอดจนเปิดโอกาสให้คณาจารย์ นักวิชาการ นักวิจัย และนิสิต นักศึกษา จากหลากหลายสถาบันได้แลกเปลี่ยนความรู้และประสบการณ์ในการทำวิจัยและการสอน ซึ่งจะนำไปสู่การประยุกต์ใช้คณิตศาสตร์เพื่อยกระดับคุณภาพชีวิตในอนาคต

การประชุมครั้งนี้ถือเป็นกิจกรรมทางวิชาการที่มีบทบาทสำคัญในการส่งเสริมและสนับสนุน การพัฒนาคุณภาพการศึกษาและงานวิจัย รวมถึงการสร้างเครือข่ายความร่วมมือเพื่อผลิตผลงานวิจัยแบบบูรณาการ นอกจากนี้ยังช่วยผลักดันนักวิจัยรุ่นใหม่ให้เติบโตในสายงานวิจัย รวมถึงส่งเสริมการเรียนรู้อย่างต่อเนื่องของนักวิจัยและคณาจารย์ เพื่อพัฒนาองค์ความรู้ที่มีคุณภาพและนำไปประยุกต์ใช้ให้เกิดประโยชน์ต่อองค์กร สังคม และประเทศชาติ อันสอดคล้องกับวิสัยทัศน์และพันธกิจของคณะวิทยาศาสตร์

คณะวิทยาศาสตร์หวังเป็นอย่างยิ่งว่าการประชุมวิชาการครั้งนี้ จะเป็นอีกก้าวหนึ่งที่เปิดโอกาสให้คณาจารย์ นักวิจัย นิสิต นักศึกษา และผู้สนใจทุกท่าน ได้แลกเปลี่ยนเรียนรู้ร่วมกัน เสริมสร้างเครือข่ายวิชาการ และสามารถนำองค์ความรู้ที่ได้ไปประยุกต์ใช้ให้เกิดประโยชน์อย่างแท้จริงต่อสังคมและประเทศชาติในอนาคต



(ศาสตราจารย์ ดร.ปรินทร์ ชัยวิสุทธางกูร)
รองอธิการบดีฝ่ายวิจัยและนวัตกรรม มหาวิทยาลัยศรีนครินทรวิโรฒ
รักษาการแทนคณบดีคณะวิทยาศาสตร์

คำนำ

การประชุมวิชาการทางคณิตศาสตร์ระดับชาติในประเทศไทย เริ่มจัดขึ้นตั้งแต่ปี พ.ศ. 2538 โดยได้รับการริเริ่มจากศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย (Center for Promotion of Mathematical Research of Thailand หรือ CEPMART) ภายใต้สมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์ การประชุมนี้จัดขึ้นเป็นประจำทุกปี โดยหมุนเวียนให้ภาควิชาคณิตศาสตร์ของมหาวิทยาลัยสมาชิก CEPMART ทำหน้าที่เป็นเจ้าภาพ ในปี พ.ศ. 2568 ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ได้รับเกียรติให้เป็นเจ้าภาพจัด การประชุมวิชาการทางคณิตศาสตร์ระดับชาติ ครั้งที่ 29 ภายใต้หัวข้อ “Lead a Better Future with Mathematics” จัดขึ้นระหว่างวันที่ 21–23 พฤษภาคม พ.ศ. 2568 การประชุมครั้งนี้มีการนำเสนอผลงานวิจัยในหลากหลายสาขา ได้แก่ คณิตศาสตร์ สถิติ วิทยาการข้อมูล คณิตศาสตรศึกษา และสาขาอื่น ๆ ที่เกี่ยวข้อง โดยมุ่งเน้นการส่งเสริมการเผยแพร่องค์ความรู้และผลงานวิจัย ตลอดจนเป็นเวทีแลกเปลี่ยนความรู้และประสบการณ์ระหว่างคณาจารย์ นักวิชาการ นักวิจัย นิสิต นักศึกษา และผู้สนใจ ซึ่งจะนำไปสู่การพัฒนาการเรียนการสอนและการวิจัยสาขาต่าง ๆ

คณะกรรมการจัดการประชุม AMM 2025 หวังเป็นอย่างยิ่งว่าการประชุมครั้งนี้จะก่อให้เกิดความร่วมมือทางวิชาการระหว่างสถาบันการศึกษาต่าง ๆ รวมถึงเสริมสร้างเครือข่ายความร่วมมือเพื่อพัฒนาวงการคณิตศาสตร์ของประเทศไทยให้ก้าวหน้าต่อไป

คณะกรรมการจัดการประชุม AMM 2025 ขอขอบคุณสมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์ ศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ และทุกหน่วยงานที่ให้การสนับสนุนการจัดประชุมในครั้งนี้ ตลอดจนวิทยากรการบรรยายพิเศษและเสวนาทางวิชาการ ผู้ทรงคุณวุฒิ ผู้นำเสนอผลงาน และผู้เข้าร่วมประชุมทุกท่านที่มีส่วนร่วมทำให้การประชุมสำเร็จลุล่วงไปด้วยดี หากมีข้อบกพร่องประการใด ภาควิชาคณิตศาสตร์ขอน้อมรับคำติชมเพื่อนำไปปรับปรุงและพัฒนาในโอกาสต่อไป

(ผู้ช่วยศาสตราจารย์ ดร.ณหทัย ฤกษ์ฤทัยรัตน์)
หัวหน้าภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ

สารบัญ

สารจากอธิการบดีมหาวิทยาลัยศรีนครินทรวิโรฒ 1

สารจากนายกสมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์ 3

สารจากผู้อำนวยการศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย 4

สารจากคณบดีคณะวิทยาศาสตร์ 5

คำนำ 6

1 กำหนดการประชุม 13

2 กำหนดการนำเสนอ 18

3 บทคัดย่อการบรรยายพิเศษ 43

4 บทความวิจัย 51

AL: Algebra

AL44 Some Conditions of the Relation between BN-Algebra and Group
Teerapan Jodnok, Pimpattra Duangnoi, Paweena Tontan, Amphon Liangprom, Kanokphitcha Chansaengkrachang and Juree Suwannasri* . 52

AL49 A Note on Endomorphism Categories
Korn Kruaykitanon and Phichet Chaoha* 58

AL50 Splitness of the Characteristic Polynomial of 2×2 Normal Matrices over a Field of Characteristic 2
Sarawut Therdchanthuk and Keng Wiboonton* 76

DC: Data Science, Computer Science, and Artificial Intelligence

DC26 ระบบการแนะนำการซื้อขายตราสารอนุพันธ์: โดยใช้รูปแบบราคาทะลุผ่านกราฟสามเหลี่ยมและตัวแบบไลต์เกรเดียนต์บูสติ้งแมชชีน
*ณัฐนนท์ เลี้ยงประไพพันธิ์, พิชายุตม์ ตันติพัฒน์พงศ์ และ โชติรส สุรพลชัย** 85

DC80 Forecasting Pollen Concentration from Past Meteorological Factors Using Deep Learning
Suppakrit Sookhanonsawat, Paponsak Sookserisarp*, Akarapark Kongsiriwong* and Sittichoke Som-am* 93

DC81 AI-Assisted Reference Retrieval For Mathematical Theorem Proving: A Step Toward Reliable Proof Generation
Krong Krongyuth, Ekawat Chaowicharat and Pannapa Changpetch* . . 104

DC82 แอปพลิเคชันเว็บสำหรับการตรวจสอบและประเมินแผลเบาหวานด้วยการเรียนรู้เชิงลึก
วรัท เมฆาสุวรรณดำรง, ธัชธรรม์ กษิดิ์เดชพงศ์*, ปุณยวีร์ อริยกิตติวงศ์*, สิทธิโชค โสมอ่ำ และ เจิมพิภัช เจริญปัญญาศักดิ์* 119

DC83 แอปพลิเคชันสำหรับการให้ข้อมูลสมุนไพรและระบุสมุนไพร จากรูปภาพใบสมุนไพร โดยใช้เทคนิคการเรียนรู้ของเครื่อง
ตฤน จันทร์อร่าม, ณัฐธีร์ ศรุติเจริญวงศ์*, วิชานันท์ ชานันโท*, สิทธิโชค โสมอ่ำ, วรพันธ์ หอมสมบูรณ์ และ เจิมพิภัช เจริญปัญญาศักดิ์* 132

GC: Graph Theory and Combinatorics

GC09 Properties of the Double Domination Numbers of Some Glued Paths
Punvarot Rattakorn and Piyashat Sripratak* 164

GC10 Equitable Colorings of Cartesian Products of Fans and Tripartite Graphs
Annob Jobpan and Kittikorn Nakprasit* 184

GC12 Simple Necessary Conditions for Achieving Byzantine Agreement over (2,3)-Uniform Hypergraphs
Supphachai Sonwai and Wutichai Chongchitmate* 201

GC58 A Distinct Length Path Decomposition of Graphs
*Sittichoke Songsa-ard and Atthakorn Sakda** 212

ME: Mathematics Education

ME43 Integrating Active Learning with The Learning Pit Model to Foster Mathematical Resilience
Gritchanat Sangsupakorn and Ratchanikorn Chonchaiya* 224

ME64 Developing an Understanding of “a half” in 3- to 4-Year-Old Children Lacking Cardinality Knowledge Through Play-Based Learning Activities
Sinchai Jansem and Amika Wachirawin* 237

ME66 การศึกษาสภาพปัญหาทักษะการสอนด้านเทคโนโลยีดิจิทัลของครู สำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของวิชาชีพครู ในโรงเรียนจังหวัดสุรินทร์
ประฎิชญา สร้อยจิตร, สุพัตรา วะยะลุน, อรุณธิดา สีเชียงหา และ วิภาวดี มูลไชยสุข* 249

ME67 การพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4
เวธน์วศิล พิรุณกาญจน์, อัคเรศ เตี๋ยวสกุล*, กาญจนา เจริญช่วย, กรวิกา ก้องกุล, อลงกรณ์ แซ่ตั้ง, เกษราภรณ์ แซ่ตั้ง และ ปรีดาภรณ์ กาญจนสำราญวงศ์* 260

ME72 การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ ที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนและเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 วิทยาลัยเทคนิคมีนบุรี
อภิสิทธิ์ กินรีวงศ์ และ ภาสุดา ภาคาผล* 272

ME74 ผลการจัดการเรียนรู้แบบ Concrete – Pictorial – Abstract ร่วมกับเทคนิคเพื่อนคู่คิด ที่ส่งผลต่อความสามารถในการให้เหตุผลทางคณิตศาสตร์ และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน
อนุชา เสาวคนธ์, ธารีรัตน์ ธนัตถ์พาณิชย์ และ วราภรณ์ จาตนิล* 286

ME75 ผลการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ที่มีต่อความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 4
เริงฤทธิ์ รัตนชวางกูร และ รัชนิกร ชลไชยะ* 299

NM: Numerical Analysis, Mathematical Modeling, and Operations Research

NM17 Numerical Simulation of Contaminant in Transient Groundwater System
Visarut Huayshelake and Khamron Mekchay* 321

NM18 A Large Neighborhood Search for Two-picker Routing Problem with Picker Blocking Consideration
Panupong Daengpradap and Panupong Vichitkunakorn* 334

NM21 Analysis of a Conservative Finite Difference Method for Burgers-type Equation
Keerati Buchatip and Nattapol Ploymaklam* 344

NM22 Determining Characteristics of Solution for Vehicle Routing Problem with Time Windows
Teerapong Chaochanglek and Boonyarit Intiyot* 357

NM23 Two-Step-Change Pivot Rule for the Simplex Method
Rawit Chaiareekij and Krung Sinapiromsaran* 368

NM76 Mathematical Model of Helicobacter with Liver Fluke as Reservoir
Tachin Kaewmala, Apichat Suratanee, Kitiporn Plaimas and Ratinan Boonklurb* 389

NM77 ตัวแบบคณิตศาสตร์การแพร่กระจายของเชื้อเอชไอวีและโรคเอดส์ที่มีผลจากประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พ และการรักษาด้วยวิธียาต้านไวรัส ART
ฐาปณี ทองคำ, เกตุกนก หนูดี, อัญชุลี ณ ตะกั่วทุ่ง และ กันยากร อ่อนรักษ์* 404

NM78 Influenza Dynamics: A Mathematical Modeling Approach
Wirawan Chinviriyasit and Settapat Chinviriyasit* 428

NT: Number Theory

NT05 พหุนามเกาส์เซียนเพลล์ควอเทอร์เนียนและพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน
ธีรภัทร ไกรช่วย, พณิชพล รามมาก, อดิศักดิ์ เด็นเพ็ชรหน่อง และ ภัทราวรรณ สิงควานนท์* 436

NT07 สมบัติบางประการของพหุนามไบเกาส์เซียนเพลล์และพหุนามไบเกาส์เซียนเพลล์ลูคัส
อรประภา เถี้ยมแก้ว, สุตาภัทร ชุมสมบูรณ์, ศรัณยา เฮงสวัสดิ์ และ ภัทราวรรณ สิงควานนท์* 448

PS: Probability Theory, Statistics, and Mathematical Finance

PS29 Optimal Parameters for an Investing Strategy based on Relative Strength Index (RSI) Indicator
Kamonwan Konghom , Weenakorn Ieosanurak and Watcharin Klongdee* 462

PS31 The Mixed Poisson Distribution with a New Two-Parameter Weighted Exponential Distribution for Overdispersed Count Data and its Application in Motor Insurance
Jiramet Hengcharoensuk and Adisak Moumeesri* 488

PS32 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณราคาปาล์มน้ำมันประเทศไทยปี พ.ศ. 2558 – 2566
ทรงกรด มัคอาน อัลฟาตีฮะห์ แมเราะ และ ปิยธิดา บุญสนอง* 508

PS33 การเปรียบเทียบประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวนเมื่อมีค่านอกเกณฑ์ในชุดข้อมูล
*วนิดา พงษ์ศักดิ์ชาติ** 527

PS34 แอพพลิเคชั่นช่วยเลือกแผนประกันสุขภาพและโรคร้ายแรง
ศุภณัฐ แหยมดี, ตรีทศ กิจวรรณ, อมรรัตน์ อ้วนไตร และ อัชฌา เฉลยจิตร์* 540

PS89 Logistic Regression for Mitigating Subpopulation Effects in Familial DNA Search
Akaraphon Jantaraphum and Monchai Kooakachai* 572

PS91 Analysis of Liquidity Provider Returns in Automated Market Makers Under Independent Black-Scholes Dynamics
Nuttanicha Jirakupt and Parkpoom Phetpradap* 591

PS95 SL-vine Copula Model
Tanabodee Monggon, Songkiat Sumetkijakan, Tippawan Santiwipanont and Monchai Kooakachai* 610

PS97 Portfolio Selection by GJR-GARCH Model and Gaussian Mixture Model
Kornprapa Chotsirikunnawat and Dawud Thongtha* 626

PS98 การจำแนกระยะการเป็นโรคไตเรื้อรังโดยตัวแบบการถดถอยลอจิสติกพหุ
จริยาวัฒน์ ลี้พันธ์ และ วิภาวรรณ เล้าอรุณ* 639

PS99 A Comparative Study of Odds Ratio Estimators under a Fixed Effect Model in Meta-Analysis Studies
Phimpatta Ariyawutthichai, Chukiat Viwatwongkasem and Jutatip Sillabutra* 657

TG: Topology and Geometry

TG13 The Properties of Function for Searching the Bound of Nieuwland Constant in Rupert's Problem
Patid Srikirinth and Wacharin Wichiramala* 686

TG14 รูปเรขาคณิตของผลลัพธ์ของผลคูณเชิงเวกเตอร์
พัฒนพรรณ เหลืองปฐมธนา และ เป็นหญิง โรจนกุล* 703

TG15 Generalized Final Topologies and Hereditary Classes
Paphatphong Painee and Pongdate Montagantirud* 713

TG16 ความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลม
เวธน์วศิล พิรุณกาญจน์, อัคเรศ เตี่ยวสกุล*, กรวิกา ก้องกุล, อลงกรณ์ แซ่ตั้ง และ เกษราภรณ์ แซ่ตั้ง* 725

TG57 Counting Umbilical Points on Generalised Fermat Surfaces
Phemmatad Tansoontorn and Nithi Rungtanapirom* 734

TG85 Geometric Data Analysis for Classifying Patterns of Points Derived from Partial Circles
Asama Jampeepan and Supanut Chaidee* 747

TG87 การสร้างแบบจำลองแผนผังสำหรับตลาดนัด มศว
ปริยาภัทร นิ่มเจริญ, ณัฐณิชา ใจฝัน, ธัญสินี มะนะนะโส, ธันยพร นาคหรั่ง และ อนัญญา เมฆเกิดชู* 762

บรรณาธิการ **783**

ผู้สนับสนุนการประชุม **788**

AMM
2025

กำหนดการประชุม

กำหนดการประชุมวิชาการทางคณิตศาสตร์ ครั้งที่ 29 ประจำปี 2568
วันที่ 21–23 พฤษภาคม 2568
โรงแรม ดิ เอมเมอรัลด์ – กรุงเทพฯ

วันพุธที่ 21 พฤษภาคม 2568	
8.00 - 9.00 น.	**ลงทะเบียน** ณ ห้อง Grand Ballroom
9.00 - 9.30 น.	**พิธีเปิดการประชุม** **กล่าวต้อนรับ** โดย ศาสตราจารย์กิตติคุณ ดร.พัฒนี อุดมกะวานิช นายกสมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์ **กล่าวรายงาน** โดย ศาสตราจารย์ ดร.ปรินทร์ ชัยวิสุทธางกูร คณบดีคณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ **กล่าวเปิดการประชุม** โดย รองศาสตราจารย์ ดร.วัลนา ธเนศพงศ์ธรรม รองอธิการบดีฝ่ายวิชาการ มหาวิทยาลัยศรีนครินทรวิโรฒ (ห้อง Grand Ballroom)
9.30 - 10.30 น.	**การบรรยายพิเศษ (Keynote) เรื่อง Mathematics: The Key to a Better and More Sustainable Society** *ศาสตราจารย์ ดร.กฤษณะ เนียมมณี* ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย (ห้อง Grand Ballroom)
10.30 - 11.00 น.	พักรับประทานอาหารว่าง
11.00 - 12.00 น.	**การนำเสนอผลงาน** (ห้อง Grand Ballroom, Tabtim, Petai, Petchchompoo)
12.00 - 13.00 น.	พักรับประทานอาหารกลางวัน

13.00 - 14.00 น.	**การบรรยายพิเศษ เรื่อง Using Learning Theories to Transform Mathematics Learning** *Dr. Yeap Ban Har* Director of Curriculum & Teacher Development Pathlight School, Singapore and Anglo Singapore International School, Thailand (ห้อง Grand Ballroom 1)	
	การบรรยายพิเศษ เรื่อง Insurance Innovations for Social Resilience *คุณอิฎฐ์ อภิรักษ์ติวงศ์* ประธานเจ้าหน้าที่บริหารสายงานพัฒนาองค์กร บริษัท พรูเด็นเชียล ประกันชีวิต (ประเทศไทย) จำกัด (มหาชน) (ห้อง Grand Ballroom 2)	
14.15 - 15.00 น.	**การนำเสนอผลงาน** (ห้อง Grand Ballroom 1, Grand Ballroom 2, Tabtim, Petai)	**การประชุม CEPMART** (ห้อง Petchchompoo)
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	
15.30 - 17.00 น.	**การนำเสนอผลงาน** (ห้อง Grand Ballroom 1, Grand Ballroom 2, Tabtim, Petai)	**การประชุม CEPMART** (ห้อง Petchchompoo)

วันพฤหัสบดีที่ 22 พฤษภาคม 2568	
9.30 - 10.30 น.	**การเสวนาทางวิชาการ เรื่อง Mathematics Through the Lens of Life: Honoring the Legacy of Prof. Dr. Narong Punnim** *รองศาสตราจารย์ ดร.จิณดิษฐ์ ละออปักษิณ* สาขาวิชาการศึกษาคณิตศาสตร์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย *รองศาสตราจารย์ ดร.จริยา อุ่ยยะเสถียร* ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย *รองศาสตราจารย์ ดร.วรานุช แขมมณี* *ผู้ช่วยศาสตราจารย์ ดร.ธันวา ธีระกาญจน์* *อาจารย์ ดร.เสริมศรี ไทยแท้* ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ (ห้อง Grand Ballroom)
10.30 - 11.00 น.	พักรับประทานอาหารว่าง
11.00 - 12.00 น.	**การบรรยายพิเศษ เรื่อง A Recent Development in Probabilistic Cluster Validation and its Software Packages with Healthcare Applications** *ผู้ช่วยศาสตราจารย์ ดร.ณรรฐคุณ วิรุฬห์ศรี* ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี (ห้อง Grand Ballroom 1)
	การบรรยายพิเศษ เรื่อง From Data to Impact: Transforming Public Services with Data Science *ดร.สุนทรีย์ ส่งเสริม* รองผู้อำนวยการสถาบันข้อมูลขนาดใหญ่ (องค์การมหาชน) (ห้อง Grand Ballroom 2)
12.00 - 13.00 น.	พักรับประทานอาหารกลางวัน
13.00 - 15.00 น.	**การนำเสนอผลงาน** (ห้อง Grand Ballroom 1, Grand Ballroom 2, Petai, Petchchompoo)
15.00 - 15.30 น.	พักรับประทานอาหารว่าง

15.30 - 16.45 น.	**การนำเสนอผลงาน** (ห้อง Grand Ballroom 1, Grand Ballroom 2, Petai, Petchchompoo)

วันศุกร์ที่ 23 พฤษภาคม 2568	
9.30 - 10.30 น.	**การเสวนาทางวิชาการ เรื่อง Lead a Better Future with Mathematics: Insights from Industry Experts** *ดร.ศรัณย์ อาฮูยา* Chief Data Scientist, Abacus Digital *คุณเทพรัตน์ วงศ์เจริญวนกิจ* ผู้อำนวยการกลุ่มงานลงทุน กองบริหารการลงทุน สำนักงานประกันสังคม กระทรวงแรงงาน *ดร.พศิน มนูรังษี* Staff Research Scientist, Google Research *ผู้ช่วยศาสตราจารย์ ดร.อุดมศักดิ์ รักวงษ์วาน* (พิธีกร) ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ (ห้อง Grand Ballroom)
10.30 - 11.00 น.	พักรับประทานอาหารว่าง
11.00 - 12.00 น.	**พิธีปิดการประชุมและพิธีมอบธงการเป็นเจ้าภาพ** **กล่าวปิดการประชุม** โดย รองศาสตราจารย์ ดร.ศจี เพียรสกุล ผู้อำนวยการศูนย์ส่งเสริมการวิจัยคณิตศาสตร์แห่งประเทศไทย (ห้อง Grand Ballroom)
12.00 - 13.00 น.	พักรับประทานอาหารกลางวัน

AMM
2025

กำหนดการนำเสนอ

ตารางเวลาการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568				
8.00 - 9.00 น.	ลงทะเบียน			
9.00 - 9.30 น.	พิธีเปิด			
9.30 - 10.30 น.	การบรรยายพิเศษ (Keynote)			
10.30 - 11.00 น.	พักรับประทานอาหารว่าง			
	GB	**Tabtim**	**Petai**	**Petchchompoo**
11.00 - 11.15 น.	AN01	NT05	GC09	TG13
11.15 - 11.30 น.	AL02	NT06	GC10	TG14
11.30 - 11.45 น.	PS03	NT07	GC11	TG15
11.45 - 12.00 น.	ME04	NT08	GC12	TG16
12.00 - 13.00 น.	พักรับประทานอาหารกลางวัน			
13.00 - 14.00 น.	การบรรยายพิเศษ			
	GB 1	**GB 2**	**Tabtim**	**Petai**
14.15 - 14.30 น.	NM17	DC26	AN35	AL44
14.30 - 14.45 น.	NM18	DC27	AN36	AL45
14.45 - 15.00 น.	NM19	DC28	AN37	AL46
15.00 - 15.30 น.	พักรับประทานอาหารว่าง			
15.30 - 15.45 น.	NM20	PS29	AN38	AL47
15.45 - 16.00 น.	NM21	PS30	AN39	AL48
16.00 - 16.15 น.	NM22	PS31	ME40	AL49
16.15 - 16.30 น.	NM23	PS32	ME41	AL50
16.30 - 16.45 น.	NM24	PS33	ME42	AL51
16.45 - 17.00 น.	NM25	PS34	ME43	

วันพฤหัสบดีที่ 22 พฤษภาคม 2568				
9.30 - 10.30 น.	การเสวนาทางวิชาการ			
10.30 - 11.00 น.	พักรับประทานอาหารว่าง			
11.00 - 12.00 น.	การบรรยายพิเศษ			
12.00 - 13.00 น.	พักรับประทานอาหารกลางวัน			
	GB 1	**GB 2**	**Petai**	**Petchchompoo**
13.00 - 13.15 น.	AN52	ME64	NM76	PS88
13.15 - 13.30 น.	AN53	ME65	NM77	PS89
13.30 - 13.45 น.	AN54	ME66	NM78	PS90
13.45 - 14.00 น.	AN55	ME67	NM79	PS91
14.00 - 14.15 น.	พักการนำเสนอ			
14.15 - 14.30 น.	TG56	AL68	DC80	NT92
14.30 - 14.45 น.	TG57	AL69	DC81	NT93
14.45 - 15.00 น.	GC58	AL70	DC82	NT94
15.00 - 15.30 น.	พักรับประทานอาหารว่าง			
15.30 - 15.45 น.	GC59	ME71	DC83	PS95
15.45 - 16.00 น.	GC60	ME72	TG84	PS96
16.00 - 16.15 น.	GC61	ME73	TG85	PS97
16.15 - 16.30 น.	GC62	ME74	TG86	PS98
16.30 - 16.45 น.	GC63	ME75	TG87	PS99

กำหนดการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568 ช่วงเช้า

ห้อง Grand Ballroom

วันพุธที่ 21 พฤษภาคม 2568 ช่วงเช้า ห้อง Grand Ballroom		
9.00 - 9.30 น.	พิธีเปิด	
9.30 - 10.30 น.	**การบรรยายพิเศษ (Keynote)** Mathematics: The Key to a Better and More Sustainable Society *ศาสตราจารย์ ดร.กฤษณะ เนียมมณี*	
10.30 - 11.00 น.	พักรับประทานอาหารว่าง	
Chair: ผู้ช่วยศาสตราจารย์ ดร.ณหทัย ฤกษ์ฤทัยรัตน์		
11.00 - 11.15 น.	AN01	Common Fixed Point Theorems for $M_{f,g}$-Contractions with Applications to Nonlinear Systems of Matrix Equations *Chalermpon Bunpog**
11.15 - 11.30 น.	AL02	Unit-regular Elements of Variant Semigroups of Transformation Preserving Double Direction Equivalences *Chanpen Phokaew*, Kittisak Tinpun, Juwainee Thotthing and Fateemah Do*
11.30 - 11.45 น.	PS03	Estimating the Minimum Initial Capital Required for an Insurance Company's Discrete Time Surplus Process *Soontorn Boonta*, Kanoktip Kotsamran and Prakrong Voraka*
11.45 - 12.00 น.	ME04	An Analysis of the Mathematisation Process for Approaching the Concept of Ratio in a Problem-Solving Based Mathematics Textbook *Jensamut Saengpun**

กำหนดการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568 ช่วงเช้า

ห้อง Tabtim

วันพุธที่ 21 พฤษภาคม 2568 ช่วงเช้า ห้อง Tabtim		
Chair: รองศาสตราจารย์ ดร.วิภาวี ตั้งใจ		
11.00 - 11.15 น.	NT05	พหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน *ธีรภัทร ไกรช่วย*, พณิชพล รามมาก, อดิศักดิ์ เด็นเพ็ชรหน่อง และ ภัทราวรรณ สิงควานนท์*
11.15 - 11.30 น.	NT06	Closed-Form Expressions of Catalan-Like Series *Nattapat Bunmak* and Detchat Samart*
11.30 - 11.45 น.	NT07	สมบัติบางประการของพหุนามไบเกาส์เซียนเพลล์และพหุนามไบเกาส์เซียนเพลล์ลูคัส *อรประภา เถี้ยมแก้ว*, สุตาภัทร ชุมสมบูรณ์, ศรัณยา เฮงสวัสดิ์ และ ภัทราวรรณ สิงควานนท์*
11.45 - 12.00 น.	NT08	On the Exponential Diophantine Equation $2568^x - 2025^y = z^2$ *Jaruwat Nakwimol* and Chalongchai Klanarong*

กำหนดการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568 ช่วงเช้า

ห้อง Petai

วันพุธที่ 21 พฤษภาคม 2568 ช่วงเช้า ห้อง Petai		
Chair: อาจารย์ ดร.อุทุมพร มาโต		
11.00 - 11.15 น.	GC09	Properties of the Double Domination Numbers of Some Glued Paths *Punvarot Rattakorn* and Piyashat Sripratak*
11.15 - 11.30 น.	GC10	Equitable Colorings of Cartesian Products of Fans and Tripartite Graphs *Annob Jobpan* and Kittikorn Nakprasit*
11.30 - 11.45 น.	GC11	3-PGDDs with block size four and two groups of size n and $n-2$ *Apiwat Peereeyaphat*, Dinesh G. Sarvate and Chariya Uiyyasathian*
11.45 - 12.00 น.	GC12	Simple Necessary Conditions for Achieving Byzantine Agreement over (2,3)-Uniform Hypergraphs *Supphachai Sonwai* and Wutichai Chongchitmate*

กำหนดการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568 ช่วงเช้า

ห้อง Petchchompoo

วันพุธที่ 21 พฤษภาคม 2568 ช่วงเช้า ห้อง Petchchompoo		
Chair: รองศาสตราจารย์ ดร.ระศิเมษ เมืองช้าง		
11.00 - 11.15 น.	TG13	Solving Rupert's Problem Using Grid Search *Patid Srikirinth* and Wacharin Wichiramala*
11.15 - 11.30 น.	TG14	รูปเรขาคณิตของผลลัพธ์ของผลคูณเชิงเวกเตอร์ *พัฒนพรรณ เหลืองปฐมธนา* และ เป็นหญิง โรจนกุล*
11.30 - 11.45 น.	TG15	Generalized Final Topologies and Hereditary Classes *Paphatphong Painee* and Pongdate Montagantirud*
11.45 - 12.00 น.	TG16	ความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลม *เวธน์วศิล พิรุณกาญจน์**, อัคเรศ เตี๋ยวสกุล*, กรวิกา ก้องกุล, อลงกรณ์ แซ่ตั้ง และ เกษราภรณ์ แซ่ตั้ง*

กำหนดการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568 ช่วงบ่าย

ห้อง Grand Ballroom 1

วันพุธที่ 21 พฤษภาคม 2568 ช่วงบ่าย ห้อง Grand Ballroom 1		
13.00 - 14.00 น.	**การบรรยายพิเศษ** Using Learning Theories to Transform Mathematics Learning *Dr. Yeap Ban Har*	
Chair: ผู้ช่วยศาสตราจารย์ปัญญวัฒน์ หาอาษา		
14.15 - 14.30 น.	NM17	Numerical Simulation of Contaminant in Transient Groundwater System *Visarut Huayshelake* and Khamron Mekchay*
14.30 - 14.45 น.	NM18	A Large Neighborhood Search for Two-Picker Routing Problem with Picker Blocking Consideration *Panupong Daengpradap* and Panupong Vichitkunakorn*
14.45 - 15.00 น.	NM19	The Optimal Slouch Angle for Running in the Rain *Kantee Leenanunath*, Phuwadej Tantisittiporn*, Sorawich Chanchotsathien* and Thanatkrit Kaewtem*
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	
15.30 - 15.45 น.	NM20	แบบจำลองปัญหาเส้นทางพาหนะที่มีความจุคันรถในการรับและส่งพัสดุในกรณีมีการตีกลับสินค้า *ภัทรพล จำปาหอม*, วชิรวิชญ์ ชาภูมี*, ศิวนาท ทนงจิตร* และ มนสิการ จันทร์สร้าง*
15.45 - 16.00 น.	NM21	Analysis of a Conservative Finite Difference Method for Burgers-type Equation *Keerati Buchatip* and Nattapol Ploymaklam*
16.00 - 16.15 น.	NM22	Determining Characteristics of Solution for Vehicle Routing Problem with Time Windows *Teerapong Chaochanglek* and Boonyarit Intiyot*

16.15 - 16.30 น.	NM23	Two-Step-Change Pivot Rule for the Simplex Method *Rawit Chaiareekij* and Krung Sinapiromsaran*
16.30 - 16.45 น.	NM24	แบบจำลองพยากรณ์การระบาดของเพลี้ยแป้งในสวนกระท้อน *วรพิชชา ไกรสิทธิพาณิชย์* และ กมลชนก อินทพาส**
16.45 - 17.00 น.	NM25	การเปรียบเทียบการคำนวณหาค่าลำดับฟีโบนัชชี ด้วยวิธีอิเทอเรทีฟกับวิธีเมทริกซ์ผ่านการเขียนอัลกอริทึม *พิชชาพร พาลีพจน์*, ณิชกานต์ ตันสุวรรณ* และ อมรรัตน์ อ้วนไตร*

กำหนดการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568 ช่วงบ่าย

ห้อง Grand Ballroom 2

วันพุธที่ 21 พฤษภาคม 2568 ช่วงบ่าย ห้อง Grand Ballroom 2		
13.00 - 14.00 น.	**การบรรยายพิเศษ** Insurance Innovations for Social Resilience *อิฎฐ์ อภิรักษ์ติวงศ์*	
Chair: อาจารย์ ดร.จิดาภา ธาดาจรัสศิริ		
14.15 - 14.30 น.	DC26	ระบบการแนะนำการซื้อขายตราสารอนุพันธ์โดยใช้รูปแบบราคาทะลุผ่าน กราฟสามเหลี่ยมและตัวแบบไลต์เกรเดียนต์บูสติ้งแมชชีน *ณัฐนนท์ เลี้ยงประไพพันธิ์, พิชายุตม์ ตันติพัฒน์พงศ์* *และ โชติรส สุรพลชัย**
14.30 - 14.45 น.	DC27	Supervised Machine Learning Classification Models for Academic Performance Prediction *Rungruttikarn Moungmai**
14.45 - 15.00 น.	DC28	การระบุตำแหน่งเส้นประสาทจากภาพอัลตราซาวด์บริเวณเหนือหัวเข่าด้วยโมเดลทางปัญญาประดิษฐ์ *วรรณกร ลาภพิรุณรัตน์, นิภาพร ประดับเวทย์*,* *ทิพวรรณ อ้อยหวาน* และ ณัชพล ประเทศ*
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	
Chair: อาจารย์ ดร.อัจฉริยะ วัธนวิสูตร		
15.30 - 15.45 น.	PS29	Optimal Parameters for an Investing Strategy Based on Relative Strength Index (RSI) Indicator *Kamonwan Konghom*, Weenakorn Ieosanurak and Watcharin Klongdee*
15.45 - 16.00 น.	PS30	A Comparison of the ARIMA–GARCH and ARIMA–EGARCH Hybrid Models for Forecasting Ethereum's Daily Minimum Prices *Natchapat Chaipeththirakul* and Benjawan Rodjanadid*

16.00 - 16.15 น.	PS31	The Mixed Poisson Distribution with a New Two-Parameter Weighted Exponential Distribution for Overdispersed Count Data and Its Application in Motor Insurance *Jiramet Hengcharoensuk* and Adisak Moumeesri*
16.15 - 16.30 น.	PS32	การวิเคราะห์การถดถอยเชิงเส้นพหุคูณราคาปาล์มน้ำมันประเทศไทยปี พ.ศ. 2558 – 2566 *ทรงกรด มัคอาน*, อัลฟาตีฮะห์ แมเราะ และ ปิยธิดา บุญสนอง*
16.30 - 16.45 น.	PS33	การเปรียบเทียบประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวนเมื่อมีค่านอกเกณฑ์ในชุดข้อมูล *วนิดา พงษ์ศักดิ์ชาติ**
16.45 - 17.00 น.	PS34	แอปพลิเคชันช่วยเลือกแผนประกันสุขภาพและโรคร้ายแรง *ศุภณัฐ แหยมดี, ตรีทศ กิจวรรณ*, อมรรัตน์ อ้วนไตร และ อัชฌา เฉลยจิตร์*

กำหนดการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568 ช่วงบ่าย

ห้อง Tabtim

วันพุธที่ 21 พฤษภาคม 2568 ช่วงบ่าย ห้อง Tabtim		
Chair: อาจารย์ ดร.นพดล วิชิตสงคราม		
14.15 - 14.30 น.	AN35	The Delay-Dependent Stability Criteria on Robust H_∞ Performance for Delayed Fuzzy Systems *Sirada Pinjai*, Mingkwan Gunjina and Jenjira Thipcha*
14.30 - 14.45 น.	AN36	Fixed Point Theorems for Wardowski-Type Mappings Based on Three-Point Conditions *Kanokwan Sawangsup* and Wutiphol Sintunavarat*
14.45 - 15.00 น.	AN37	Comparison of High-Order Compact Finite Difference Schemes for the Benjamin-Bona-Mahony-KdV Equation with Different Nonlinear Approximation Approaches *Apipoom Polwang**
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	
15.30 - 15.45 น.	AN38	A Chaos-Driven Approach to Augmenting and Enriching Symbolic Music Datasets *Kanatsanun Sub-udom*, Pichate Kunakornvong and Ratthaprom Promkam*
15.45 - 16.00 น.	AN39	On the Relationship Between Kannan Contractions and Banach Contractions in Quasi-Metric Spaces *Wasin Supakul* and Wutiphol Sintunavarat*
Chair: ผู้ช่วยศาสตราจารย์ ดร.เอนก จันทรจรูญ		
16.00 - 16.15 น.	ME40	การพัฒนารูปแบบการจัดกิจกรรมการเรียนรู้แบบบูรณาการสะเต็มศึกษาร่วมกับภูมิปัญญาท้องถิ่น เพื่อส่งเสริมทักษะชีวิตและทักษะกระบวนการทางคณิตศาสตร์ เรื่อง อนุพันธ์ของฟังก์ชันพีชคณิต รายวิชาเพิ่มเติม คณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 6 *สุทธดา เหลืองห่อ**

16.15 - 16.30 น.	ME41	Developing an Interactive Platform for Teaching AI Fundamentals Through Coding *Sittichoke Som-am**
16.30 - 16.45 น.	ME42	Integrating ChatGPT in the Calculus Classroom: Student and Teacher Perceptions *Thanatkrit Kaewtem**
16.45 - 17.00 น.	ME43	Integrating Active Learning with the Learning Pit Model to Foster Mathematical Resilience *Gritchanat Sangsupakorn* and Ratchanikorn Chonchaiya*

กำหนดการนำเสนอ

วันพุธที่ 21 พฤษภาคม 2568 ช่วงบ่าย

ห้อง Petai

วันพุธที่ 21 พฤษภาคม 2568 ช่วงบ่าย ห้อง Petai		
Chair: ผู้ช่วยศาสตราจารย์ ดร.นิสรา สิรสุนทร		
14.15 - 14.30 น.	AL44	Some Conditions of the Relation between BN-Algebra and Group *Teerapan Jodnok*, Pimpattra Duangnoi, Paweena Tontan, Amphon Liangprom, Kanokphitcha Chansaengkrachang and Juree Suwannasri*
14.30 - 14.45 น.	AL45	Power Ternary Semigroups on Ternary Semihypergroups *Anak Nongmanee*, Kanruethai Jeenkaew and Sorasak Leeratanavalee*
14.45 - 15.00 น.	AL46	A Note on Hybrid Quasi-Ideals in Ternary Semigroups via Rough Approximations *Rukchart Prasertpong* and Aiyared Iampan*
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	
Chair: ผู้ช่วยศาสตราจารย์ ดร.สุชาดา พงษ์ประเสริฐ		
15.30 - 15.45 น.	AL47	A Classification of Maximal Subsemigroups of Finite Semigroups of Transformations with Invariant Set *Azizah Baka* and Yanisa Chaiya*
15.45 - 16.00 น.	AL48	Implicative Negatively Partially Ordered Ternary Semigroups *Kansada Nakwan*, Panuwat Luangchaisri and Thawhat Changphas*
16.00 - 16.15 น.	AL49	A Note on Endomorphism Categories *Korn Kruaykitanon* and Phichet Chaoha*

16.15 - 16.30 น.	AL50	Splitness of the Characteristic Polynomial of 2×2 Normal Matrices over a Field of Characteristic 2 *Sarawut Therdchanthuk* and Keng Wiboonton*
16.30 - 16.45 น.	AL51	Generalized Quaternions Sylvester Matrix Equation with Applications to Color Image Deblurring *Janthip Jaiprasert*, and Pattrawut Chansangiam*

กำหนดการนำเสนอ

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงเช้า

ห้อง Grand Ballroom 1-2

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงเช้า ห้อง Grand Ballroom	
9.30 - 10.30 น.	**การเสวนาทางวิชาการ** Mathematics Through the Lens of Life: Honoring the Legacy of Prof. Dr. Narong Punnim *รองศาสตราจารย์ ดร.จิณดิษฐ์ ละออปักษิณ* *รองศาสตราจารย์ ดร.จริยา อุ่ยยะเสถียร* *รองศาสตราจารย์ ดร.วรานุช แขมมณี* *ผู้ช่วยศาสตราจารย์ ดร.ธันวา ธีระกาญจน์* *อาจารย์ ดร.เสริมศรี ไทยแท้*
10.30 - 11.00 น.	พักรับประทานอาหารว่าง
11.00 - 12.00 น.	**ห้อง Grand Ballroom 1**
	การบรรยายพิเศษ A Recent Development in Probabilistic Cluster Validation and its Software Packages with Healthcare Applications *ผู้ช่วยศาสตราจารย์ ดร.ณรรฐคุณ วิรุฬห์ศรี*
	ห้อง Grand Ballroom 2
	การบรรยายพิเศษ From Data to Impact: Transforming Public Services with Data Science *ดร.สุนทรีย์ ส่งเสริม*

กำหนดการนำเสนอ

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงบ่าย

ห้อง Grand Ballroom 1

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงบ่าย ห้อง Grand Ballroom 1		
Chair: ผู้ช่วยศาสตราจารย์ ดร.เรืองวรินท์ อินทรวงษ์ สราญรักษ์สกุล		
13.00 - 13.15 น.	AN52	Henstock-Fresnel Integrals in $\mathbb{R}$ *Suppakorn Sungsillert* and Varayu Boonpongkrong*
13.15 - 13.30 น.	AN53	Some Fixed Point Properties in Composite Generalized Hilbert Spaces and Their Applications in Quantum Mechanics *Perawit Boonsomchua* Atchariyapol Chanhorm* and Kanokwan Sitthithekingkiat*
13.30 - 13.45 น.	AN54	Convergence Analysis with Self-Comparative Rate Assessment of Iterative Methods *Akarate Singta* and Wutiphol Sintunavarat*
13.45 - 14.00 น.	AN55	A Novel Pexiderized Functional Equation of Davison Type *Pratchaya Singjanusong* and Wutiphol Sintunavarat*
14.00 - 14.15 น.	พักการนำเสนอ	
Chair: ผู้ช่วยศาสตราจารย์ ดร.ธันวา ธีระกาญจน์		
14.15 - 14.30 น.	TG56	Greybody Factors of 4D Einstein-Gauss-Bonnet Metric *Kunlapat Sansuk*, Petarpa Boonserm and Sattha Phalungsongsathit*
14.30 - 14.45 น.	TG57	Counting Umbilical Points on Generalised Fermat Surfaces *Phemmatad Tansoontorn* and Nithi Rungtanapirom*
14.45 - 15.00 น.	GC58	A Distinct Length Path Decomposition of Graph *Sittichoke Songsa-ard and Atthakorn Sakda**
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	

Chair: ผู้ช่วยศาสตราจารย์ ดร.ศุภโชค อิสริยปาลกุล		
15.30 - 15.45 น.	GC59	Connectedness of the Graph from (n^2-1)-Puzzle with 1 Fixed Cell *Piyashat Sripratak**
15.45 - 16.00 น.	GC60	The $(3,3)$-Colorability of Planar Graphs without Cycles of Length $3, 4$, and 6 *Pongpat Sittitrai and Wannapol Pimpasalee**
16.00 - 16.15 น.	GC61	On Structural and Invariant Properties for Annihilator Graphs of P_n *Chollawat Pookpienlert*, Nuttawoot Nupo and Yanisa Chaiya*
16.15 - 16.30 น.	GC62	Domination Parameters of Bishop Graphs Defined on Honeycomb Chessboards *Siripong Sirisuk, Sakollawat Sookjam*, Graiwit Wanthong and Tonnam Ninrat*
16.30 - 16.45 น.	GC63	Numbers of Locally Strong Homomorphisms from Path Graphs to Path Graphs and from Path Graphs to Cycle Graphs *Yananan Mati*, Natthapong Wongkam, Sayan Panma and Penying Rochanakul*

กำหนดการนำเสนอ

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงบ่าย

ห้อง Grand Ballroom 2

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงบ่าย ห้อง Grand Ballroom 2		
Chair: ผู้ช่วยศาสตราจารย์ ดร.สายัณห์ โสธะโร		
13.00 - 13.15 น.	ME64	Developing an Understanding of “a half” in 3- to 4-Year-Old Children Lacking Cardinality Knowledge Through Play-Based Learning Activities *Sinchai Jansem* and Amika Wachirawin*
13.15 - 13.30 น.	ME65	ผลการฝึกสติต่อความเครียด สมาธิ และผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนมหิดลวิทยานุสรณ์ *นงลักษณ์ อาภาสัตย์* และ อรวิสาส์ กิจพิทักษ์**
13.30 - 13.45 น.	ME66	การศึกษาสภาพปัญหาทักษะการสอนด้านเทคโนโลยีดิจิทัลของครูสำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของวิชาชีพครูในโรงเรียนจังหวัดสุรินทร์ *ประฎิชญา สร้อยจิตร, สุพัตรา วะยะลุน, อรุณธิดา สีเชียงหา* และ วิภาวดี มูลไชยสุข*
13.45 - 14.00 น.	ME67	การพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็นของนักเรียนชั้นมัธยมศึกษาปีที่ 4 *เวธน์วศิล พิรุณกาญจน์*, อัคเรศ เตี๋ยวสกุล*, กาญจนา เจริญช่วย, กรวิกา ก้องกุล, อลงกรณ์ แซ่ตั้ง, เกษราภรณ์ แซ่ตั้ง และ ปรีดาภรณ์ กาญจนสำราญวงศ์*
14.00 - 14.15 น.	พักการนำเสนอ	
Chair: อาจารย์ ดร.เสริมศรี ไทยแท้		
14.15 - 14.30 น.	AL68	On Semicomplete Leibniz Algebras *Chaleamrach Malangpoo*, Suchada Pongprasert and Nissara Sirasuntorn*

14.30 - 14.45 น.	AL69	On Almost Inner Derivations of Leibniz Algebras *Pattraporn Bumrungkit*, Kittipong Navisej*,* *Wannisa Phanfai* and Suchada Pongprasert*
14.45 - 15.00 น.	AL70	Magic Card Trick with Two Errors *Sasithorn Insakul*, Nattanan Kongjan,* *Thitarie Rungratgasame and Neeracha Supakosol*
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	
Chair: ผู้ช่วยศาสตราจารย์ ดร.ขวัญ เพียซ้าย		
15.30 - 15.45 น.	ME71	ผลการเรียนรู้วิชาคณิตศาสตร์ เรื่อง การแก้โจทย์ปัญหาการนำเสนอข้อมูล โดยใช้การจัดการเรียนรู้แบบสืบเสาะหาความรู้ (5E) ร่วมกับแนวคิดโพลยา สำหรับนักเรียนชั้นประถมศึกษาปีที่ 5 โรงเรียนสามเสนนอก (ประชาราษฎร์อนุกูล) *พรลิตา ประกอบดี* ณัฏฐกานต์ แก้วกำพล* และ สุภาภรณ์ สดวกดี*
15.45 - 16.00 น.	ME72	การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ ที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนและเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 วิทยาลัยเทคนิคมีนบุรี *อภิสิทธิ์ กินรีวงศ์* และ ภาสุดา ภาคาผล*
16.00 - 16.15 น.	ME73	การวิเคราะห์การสนทนาเชิงวิภาษวิธีในหนังสือเรียนคณิตศาสตร์ที่สอนผ่านการแก้ปัญหาในบทเรียนเรื่องความยาวชั้นประถมศึกษาปีที่ 2 *วีรภัทร ดากลาง* และ เจนสมุทร แสงพันธ์*
16.15 - 16.30 น.	ME74	ผลการจัดการเรียนรู้แบบ Concrete – Pictorial – Abstract ร่วมกับเทคนิคเพื่อนคู่คิด ที่ส่งผลต่อความสามารถในการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน *อนุชา เสาวคนธ์*, ธารีรัตน์ ธนัตถ์พาณิชย์ และ วราภรณ์ จาตนิล*
16.30 - 16.45 น.	ME75	ผลการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ที่มีต่อความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 4 *เริงฤทธิ์ รัตนชวางกูร* และ รัชนิกร ชลไชยะ*

กำหนดการนำเสนอ

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงบ่าย

ห้อง Petai

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงบ่าย ห้อง Petai		
Chair: อาจารย์ ดร.ธีรศักดิ์ ฉลาดการณ์		
13.00 - 13.15 น.	NM76	Mathematical Model of *Helicobacter* with Liver Fluke as Reservoir *Tachin Kaewmala*, Apichat Suratanee, Kitiporn Plaimas and Ratinan Boonklurb*
13.15 - 13.30 น.	NM77	ตัวแบบคณิตศาสตร์การแพร่กระจายของเชื้อเอชไอวีและโรคเอดส์ที่มีผลจากประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พและการรักษาด้วยวิธียาต้านไวรัส ART *ฐาปณี ทองคำ*, เกตุกนก หนูดี, อัญชุลี ณ ตะกั่วทุ่ง และ กันยากร อ่อนรักษ์*
13.30 - 13.45 น.	NM78	Influenza Dynamics: A Mathematical Modeling Approach *Wirawan Chinviriyasit* and Settapat Chinviriyasit*
13.45 - 14.00 น.	NM79	Subspace Distances in Parametric Nonlinear Model Reduction: A Comparative Approach with the KPP-Fisher Equation *Suvarin Phungngam*, Norapon Sukuntee and Saifon Chaturantabut*
14.00 - 14.15 น.	พักการนำเสนอ	
Chair: ผู้ช่วยศาสตราจารย์ ดร.อิทธิเทพ นวาระสุจิตร		
14.15 - 14.30 น.	DC80	Forecasting Pollen Concentration from Past Meteorological Factors Using Deep Learning *Suppakrit Sookhanonsawat*, Paponsak Sookserisarp*, Akarapark Kongsiriwong* and Sittichoke Som-am*

14.30 - 14.45 น.	DC81	AI-Assisted Reference Retrieval for Mathematical Theorem Proving: A Step Toward Reliable Proof Generation *Krong Krongyuth*, Ekawat Chaowicharat and Pannapa Changpetch*
14.45 - 15.00 น.	DC82	แอปพลิเคชันเว็บสำหรับการตรวจสอบและประเมินแผลเบาหวานด้วยการเรียนรู้เชิงลึก *วรัท เมฆาสุวรรณดำรง*, ธัชธรรม์ กษิดิ์เดชพงศ์*, ปุณยวีร์ อริยกิตติวงศ์*, สิทธิโชค โสมอ่ำ และ เจิมพิภัช เจริญปัญญาศักดิ์*
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	
15.30 - 15.45 น.	DC83	แอปพลิเคชันสำหรับการให้ข้อมูลสมุนไพรและระบุสมุนไพรจากรูปภาพใบสมุนไพรโดยใช้เทคนิคการเรียนรู้ของเครื่อง *ตฤน จันทร์อร่าม*, ณัฐธีร์ ศรุติเจริญวงศ์*, วิชานันท์ ชานันโท*, สิทธิโชค โสมอ่ำ, วรพันธ์ หอมสมบูรณ์ และ เจิมพิภัช เจริญปัญญาศักดิ์*
Chair: ผู้ช่วยศาสตราจารย์ ดร.วิศรุต โพธิ์อ้น		
15.45 - 16.00 น.	TG84	Properties of the Region Generated by Slabs Intersection of a Set of Three Circles *Phapaengmuang Sukkasem* and Supanut Chaidee*
16.00 - 16.15 น.	TG85	Geometric Data Analysis for Classifying Patterns of Points Derived from Partial Circles *Asama Jampeepan* and Supanut Chaidee*
16.15 - 16.30 น.	TG86	Infinite Families of Hexagon Pairs Admitting Angle Permutations *Pakkawan Kusuwan*, Poombadee Panawas, Pasagorn Pirom*and Thunwa Theerakarn*
16.30 - 16.45 น.	TG87	การสร้างแบบจำลองแผนผังสำหรับตลาดนัด มศว *ปริยาภัทร นิ่มเจริญ*, ณัฐณิชา ใจฝัน, ธัญสินี มะนะโส, ธันยพร นาคหรั่ง และ อนัญญา เมฆเกิดชู*

กำหนดการนำเสนอ

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงบ่าย

ห้อง Petchchompoo

วันพฤหัสบดีที่ 22 พฤษภาคม 2568 ช่วงบ่าย ห้อง Petchchompoo		
Chair: ผู้ช่วยศาสตราจารย์ ดร.สุกัญญา อินทรภักดิ์		
13.00 - 13.15 น.	PS88	Expectation and Variance of Ruin Time in the n-Player Gambler's Ruin Problem with One Player Having Twice the Capital *Parkpoom Phetpradap* and Natkamon Sripanitan*
13.15 - 13.30 น.	PS89	Logistic Regression for Mitigating Subpopulation Effects in Familial DNA Search *Akaraphon Jantaraphum* and Monchai Kooakachai*
13.30 - 13.45 น.	PS90	Modeling Relegation Risk in the Premier League with Logistic Regression *Kittiyakorn Naiwikun* and Monchai Kooakachai*
13.45 - 14.00 น.	PS91	Analysis of Liquidity Provider Returns in Automated Market Makers under Independent Black-Scholes Dynamics *Nuttanicha Jirakupt* and Parkpoom Phetpradap*
14.00 - 14.15 น.	พักการนำเสนอ	
Chair: อาจารย์ ดร.นพดล วิชิตสงคราม		
14.15 - 14.30 น.	NT92	Base-$b(\mathcal{C})$ Representations and Irreducible Polynomials with Integer Coefficients *Phitthayathon Phetnun, Mayurachat Janthawee*, Narakorn Rompurk Kanasri and Pattaraporn Boonying*
14.30 - 14.45 น.	NT93	On the Exponential Diophantine Equations $3^x - (5p)^y = z^2$ and $(5p)^x - 3^y = z^2$ *Wariam Chuayjan*, Thapani Kraiwiradechachai, Theeradach Kaewong, and Sutthiwat Thongnak*

14.45 - 15.00 น.	NT94	On the Exponential Diophantine Equation $3^x + (2p)^y = z^2$ *Thapani Kraiwiradechachai*, Wariam Chuayjan, Theeradach Kaewong and Sutthiwat Thongnak*
15.00 - 15.30 น.	พักรับประทานอาหารว่าง	
Chair: อาจารย์ ดร.อังคนา เกาะแก้ว		
15.30 - 15.45 น.	PS95	SL-vine Copula Model *Tanabodee Monggon*, Songkiat Sumetkijakan, Tippawan Santiwipanont and Monchai Kooakachai*
15.45 - 16.00 น.	PS96	Comparing Survival Prediction Models with and without Interactions for Probability of Defaults *Chaitawat Seedang* and Kannigar Hirunkasi*
16.00 - 16.15 น.	PS97	Portfolio Selection by GJR-GARCH Model and Gaussian Mixture Model *Kornprapa Chotsirikunnawat* and Dawud Thongtha*
16.15 - 16.30 น.	PS98	การจำแนกระยะการเป็นโรคไตเรื้อรังโดยตัวแบบการถดถอยลอจิสติกพหุ *จริยาวัฒน์ ลี้พันธ์* และ วิภาวรรณ เล้าอรุณ*
16.30 - 16.45 น.	PS99	A Comparative Study of Odds Ratio Estimators under a Fixed Effect Model in Meta-Analysis Studies *Phimpatta Ariyawutthichai*, Chukiat Viwatwongkasem and Jutatip Sillabutra*

กำหนดการนำเสนอ

วันศุกร์ที่ 23 พฤษภาคม 2568 ช่วงเช้า
ห้อง Grand Ballroom

วันศุกร์ที่ 23 พฤษภาคม 2568 ช่วงเช้า ห้อง Grand Ballroom	
9.30 - 10.30 น.	**การเสวนาทางวิชาการ** Lead a Better Future with Mathematics: Insights from Industry Experts *ดร.ศรัณย์ อาฮูยา* *เทพรัตน์ วงศ์เจริญวนกิจ* *ดร.พศิน มนูรังษี* *ผู้ช่วยศาสตราจารย์ ดร.อุดมศักดิ์ รักวงษ์วาน* (พิธีกร)
10.30 - 11.00 น.	พักรับประทานอาหารว่าง
11.00 - 12.00 น.	พิธีปิดและพิธีมอบธงการเป็นเจ้าภาพ

AMM
2025

บทคัดย่อการบรรยายพิเศษ

Mathematics: The Key to a Better and More Sustainable Society

Professor Dr. Kritsana Neammanee

Keynote Speaker

Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University

Abstract

In this talk, I will present examples of the roles that Thai mathematicians play in various aspects of development and solving important problems. These contributions are crucial for fostering a better and more sustainable society, in alignment with the key issues addressed by the World Economic Forum.

Using Learning Theories to Transform Mathematics Learning

Dr. Yeap Ban Har

Invited Speaker

Director of Curriculum & Teacher Development

Pathlight School, Singapore

Anglo Singapore International School, Thailand

Abstract

Over the years, the speaker has travelled to more than forty countries to advise schools and education systems on using learning theories to help teachers help students learn mathematics well. This lecture outlines the conditions for transforming mathematics teaching in schools.

Insurance Innovations for Social Resilience

อิฎฐ์ อภิรักษ์ติวงศ์

Invited Speaker

ประธานเจ้าหน้าที่บริหารสายงานพัฒนาองค์กร

บริษัท พรูเด็นเชียล ประกันชีวิต (ประเทศไทย) จำกัด (มหาชน)

Abstract

By leveraging their expertise in risk assessment, pricing, and data analysis, actuaries help insurers develop sustainable solutions that not only protect policyholders but also contribute to broader social resilience efforts.

Mathematics Through the Lens of Life: Honoring the Legacy of Professor Dr. Narong Punnim

Associate Professor Dr. Jinnadit Laorpaksin[1]
Associate Professor Dr. Chariya Uiyyasathian[2]
Associate Professor Dr. Varanoot Khemmani[3]
Assistant Professor Dr. Thunwa Theerakarn[3]
Dr. Sermsri Thaithae[3]

[1]Mathematics Education Program, Faculty of Education, Chulalongkorn University
[2]Department of Mathematics and Computer Science,
Faculty of Science, Chulalongkorn University
[3]Department of Mathematics, Faculty of Science, Srinakharinwirot University

Abstract

Professor Dr. Narong Punnim was a visionary educator, mathematician, and mentor who made a lasting impact on Thailand's mathematical community. As Thailand's leading researcher in graph theory and design theory, he published over 50 papers and mentored many students. Beyond research, he played a key role in advancing mathematics education in Thailand by helping to establish foundational curricula for the Mathematics Olympiad program under the Promotion of Academic Olympiad and Development of Science Education Foundation (POSN) and for science schools nationwide.

In this panel discussion, students and colleagues will take the audience on a journey through mathematics. Through personal reflections and shared experiences, the panel will explore how Professor Dr. Narong Punnim inspired generations of thinkers and helped shape the mathematical landscape of Thailand. More than a tribute, this session invites the audience to see mathematics not just as a subject, but as a way of life–deeply human, profoundly impactful, and endlessly inspiring.

A Recent Development in Probabilistic Cluster Validation and its Software Packages with Healthcare Applications

Assistant Professor Dr. Nathakhun Wiroonsiri

Invited Speaker

Department of Mathematics, Faculty of Science,
King Mongkut's University of Technology Thonburi

Abstract

Cluster analysis is one of the most important unsupervised machine learning tools with applications in several areas such as business, healthcare, finance, social science, etc. It is used to separate sample points into groups which share some similar characteristics. The most known algorithms such as K-means, Fuzzy C-means, hierarchical, and EM clustering require a validation step to detect the final number of clusters. Countless numbers of cluster validity indices (CVIs) were introduced to seek the optimal number of clusters, yet no work has been discussed on sub-optimal options. In some situations, more than one choice can be chosen and the one suitable for a specific application is not always the optimal one. To fill this gap, we develop correlation-based WI and WPI for hard and soft clustering, respectively. We also introduce a new connection between the Bayesian framework and CVIs which we name BCVI. It is devoted to searching the final number of clusters when experience matters. Potential applications to healthcare problems such as MRI brain tumor detection will also be discussed. Our CRAN packages called 'UniversalCVI' and 'BayesCVI' for computing these indices are available to install.

From Data to Impact: Transforming Public Services with Data Science

Dr. Soontaree Songserm

Invited Speaker

Senior Executive Vice President

Big Data Institute (Public Organization)

Abstract

In the era of data abundance, government agencies possess vast and underutilized datasets that, when analyzed effectively, can significantly enhance public service delivery. This talk explores how both descriptive and predictive analytics are being used to drive data-informed decision-making in the public sector. Showcasing real-world use cases developed by Big Data Institute (Public Organization), this includes impactful projects such as Travellink, a dashboard that synthesizes tourism data to support policy and business planning; Healthlink, which integrates patient data for health analytics including budget prediction for dialysis, causal impact analysis, and disease progression modeling; Envilink, which uses data aggregation and image processing to calculate baseline estimation to address environmental issues like air pollution and carbon emissions; and Youthlink, a platform that connects fragmented early childhood data to forecast IQ, EQ, and SDQ scores, helping shape developmental policy. Each use case highlights the mathematical and data science methods behind the insights–from statistical modeling to machine learning and causal inference. Together, they demonstrate how unlocking the value of government-held data can lead to macro-level impact, smarter governance, and measurable improvements in citizens' quality of life.

Lead a Better Future with Mathematics: Insights from Industry Experts

Dr. Sarun Ahuja[1]
Tepparat Wongcharoenwanaki[2]
Dr. Pasin Manurangsi[3]
Assistant Professor Dr. Udomsak Rakwongwan[4] **(Moderator)**

[1]Chief Data Scientist, Abacus Digital

[2]Director of Investment Unit, Investment Management Division
Social Security Office, Ministry of Labour

[3]Staff Research Scientist, Google Research

[4]Department of Mathematics, Faculty of Science, Kasetsart University

Abstract

This panel brings together three distinguished industry professionals whose careers demonstrate the power and versatility of mathematics in driving innovation, informing policy, and solving real-world problems. From financial technology and cyber privacy to national social security systems, the panelists will share how their mathematical backgrounds have enabled them to lead in diverse sectors. Aimed at students and faculty alike, this session will explore career pathways, the societal impact of mathematics, and essential skills for the future. Join us for a lively and inspiring discussion on how mathematics can lead us toward a better future.

AMM
2025

บทความวิจัย

Some Conditions of the Relation between BN-algebra and Group

Teerapan Jodnok[1,*], Pimpattra Duangnoi[1], Paweena Tontan[1],
Amphon Liangprom[1], Kanokphitcha Chansaengkrachang[1]
and Juree Suwannasri[1]

[1]Department of Mathematics, Faculty of Science and Technology,
Surindra Rajabhat University, Surin 32000, Thailand

Abstract

Algebraic structures were widely studied in fields of mathematics consisting of a non-empty set with a binary operation $*$ and a constant 0 and axioms of its. Particularly, B-algebra was an important structure and was introduced important properties dramatically which was continued a novel structure as BN-algebra.

In this article, we intend to prove that associative property, commutative property and 0-commutative property have relations mutually on BN-algebra. Moreover, we introduce additional identities of BN-algebra. Particularly, we can show that a BN-algebra X is a group under the condition $(x * y) * z = (z * y) * x$.

Keywords: B-algebra, BN-algebra, associative, commutative, 0-commutative
2020 MSC: 06F35, 06F15

1 Introduction

An algebra $(X, *, 0)$ of type $(2, 0)$ is mathematical system that consists of a non-empty set with a binary operation $*$ and a constant 0 and axioms of its [2, 3, 6, 9]. J. Neggers and H. S. Kim [9] introduced the idea and the definition about B-algebra which defined a B-algebra as an algebra $(X, *, 0)$ of type $(2, 0)$ (i.e., a non-empty set with a binary operation $*$ and a constant 0) satisfying the following axioms: $x*x = 0, x*0 = x$ and $(x*y)*z = x*[z*(0*y)]$ for any $x, y, z \in X$. Next, this notion was used to study some axiomatization of commutative B-algebras by A.Walendziak [3] which some systems of axioms defining a B-algebra were given with proof of the independence of the axioms. And they introduced a simplified axiomatization of commutative B-algebras. Recently, the notion of B-algebra was continued as BN-algebra by H. S. Kim [4] which was defined as an algebra $(X, *, 0)$ of type $(2, 0)$ consisting of a non-empty set, a binary operation $*$ and a constant 0 with satisfying the axioms in definition 2.1. Moreover, they proved that a BN-algebra is 0-commutative (see in definition 2.4), and an algebra X is a BN-algebra if and only if it is a 0-commutative BF-algebra. And they showed some relations between BN-algebras and several algebras. In this paper, we define some conditions on an algebra $(X, *, 0)$ to prove some properties on a BN-algebra and show relations between several properties on a BN-algebra. In addition, we will show that a BN-algebra is group under condition(P) (see in definition 3.7).

*Speaker and corresponding author (teerapan.jo@srru.ac.th)

2 Preliminaries

In this section, we will use both some definitions and theorems for novel notion to introduce some new identities on a BN-algebra. Moreover, we will introduce some important concepts on B-algebra and BN-algebra to use as tools and methods for proving our results in the next section.

Definition 2.1. ([4]) A **BN-algebra** is an algebra $(X, *, 0)$ of type $(2, 0)$ satisfying the following axioms:
(BN1) $x * x = 0$,
(BN2) $x * 0 = x$ and
(BN3) $(x * y) * z = (0 * z) * (y * x)$
for any $x, y, z \in X$.

Example 2.2. ([4]) Let the set $X = \{0, 1, 2\}$. If we define a binary operation $*$ on X as the following table:

$*$	0	1	2
0	0	1	2
1	1	0	1
2	2	1	0

then $(X, *, 0)$ is a BN-algebra.

Theorem 2.3. *([4]) If $(X, *, 0)$ is a BN-algebra, then*

1. $0 * (0 * x) = x$,
2. $y * x = (0 * x) * (0 * y)$,
3. $(0 * x) * y = (0 * y) * x$,
4. *if* $x * y = 0$ *then* $y * x = 0$,
5. *if* $0 * x = 0 * y$ *then* $x = y$,
6. $(x * z) * (y * z) = (z * y) * (z * x)$

for any $x, y, z \in X$.

Definition 2.4. ([9]) An algebra $(X, *, 0)$ is said to be **0-commutative** if $x * (0 * y) = y * (0 * x)$ for any $x, y \in X$.

Theorem 2.5 ([9]). *If $(X, *, 0)$ is 0-commutative B-algebra, then $(0 * x) * (0 * y) = y * x$ for any $x, y \in X$.*

Theorem 2.6 ([9]). *If $(X, *, 0)$ is B-algebra, then $0 * (x * y) = y * x$ for any $x, y \in X$.*

Corollary 2.7 ([4]). *If $(X, *, 0)$ is a 0-commutative B-algebra, then it is a BN-algebra.*

The converse of Corollary 2.7 does not hold in general. For example, Let $X = \{0, 1, 2\}$ be a set with the table given in Example 2.2. Then it is easy to show that $(X, *, 0)$ is a BN-algebra, but not a B-algebra, since $(1 * 1) * 2 = 2 \neq 1 * [2 * (0 * 1)] = 0$.

Every abelian group was shown that it was a BN-algebra which presented the following theorem.

Theorem 2.8 ([9]). *Let $(X, \circ, 0)$ be an abelian group. If we define $x * y = x \circ y^{-1}$ for any $x, y \in X$, then $(X, *, 0)$ is a BN-algebra.*

3 Main Results

In this section, we will introduce our main results which focus on some conditions for some properties on BN-algebra. Moreover, we will introduce some conditions to show important identities consisting of additional properties on BN-algebra and the relation between BN-algebra and group. First step, we will present additional properties on BN-algebra satisfying the following definitions and theorems.

Definition 3.1. An algebra $(X, *, 0)$ is said to be **commutative** if $x * y = y * x$ for any $x, y \in X$.

Definition 3.2. An algebra $(X, *, 0)$ is said to be **associative** if $x * (y * z) = (x * y) * z$ for any $x, y, z \in X$.

Theorem 3.3. *If $(X, *, 0)$ is an associative BN-algebra, then*

1. $0 * x = x$,
2. $x * (y * z) = (z * y) * x$

for any $x, y, z \in X$.

Proof. Let $(X, *, 0)$ is an associative BN-algebra and $x, y, z \in X$.
From (BN2) and associative property, we obtain

$$0 * x = (0 * 0) * x = 0 * (0 * x) = x.$$

And from (BN2) and $0 * x = x$, we obtain

$$x * (y * z) = (0 * x) * (y * z) = (z * y) * x.$$

This completes the proof. □

Theorem 3.4. *If $(X, *, 0)$ is an associative and 0-commutative BN-algebra, then it is commutative.*

Proof. Let $(X, *, 0)$ is an associative and 0-commutative BN-algebra.
For any $x, y \in X$, from (BN2), we obtain

$$\begin{aligned} x * y &= (x * 0) * y \\ &= x * (0 * y) \\ &= y * (0 * x) \\ &= (y * 0) * x \\ &= y * x. \end{aligned}$$

Therefore, $(X, *, 0)$ is commutative. □

Theorem 3.5. *If $(X, *, 0)$ is an associative BN-algebra, then it is 0-commutative.*

Proof. Let $(X, *, 0)$ is an associative BN-algebra and $x, y \in X$.
From Theorem 3.3 and (BN3), we obtain

$$\begin{aligned} x * (0 * y) &= (0 * x) * (0 * y) \\ &= (y * 0) * x \\ &= y * (0 * x) \end{aligned}$$

Therefore, $(X, *, 0)$ is 0-commutative. □

Corollary 3.6. *If $(X, *, 0)$ is an associative BN-algebra, then it is commutative.*

Proof. From Theorem 3.4 and Theorem 3.5, then X is commutative. □

Next step, we will define a condition as condition (P) to use as tool for introducing some properties and show the relation between BN-algebra and group satisfying the following definitions and theorems.

Definition 3.7. An algebra $(X, *, 0)$ is said to have a **condition (P)** if $(x * y) * z = (z * y) * x$ for any $x, y, z \in X$.

Example 3.8. Let the BN-algebra $X = \{0, a\}$ and a binary operation $*$ on X as the following table:

$*$	0	a
0	0	a
a	a	0

then $(X, *, 0)$ has a condition(P).

Theorem 3.9. *If $(X, *, 0)$ is a BN-algebra with the condition (P), then it is commutative.*

Proof. Let $(X, *, 0)$ is a BN-algebra with the condition (P) and $x, y \in X$.
From (BN2), we obtain

$$\begin{aligned} x * y &= (x * 0) * y \\ &= (y * 0) * x \\ &= y * x \end{aligned}$$

Therefore, $(X, *, 0)$ is commutative. □

Theorem 3.10. *Let $(X, *, 0)$ is a BN-algebra, then it has the condition (P) if and only if it is associative.*

Proof. For any $x, y, z \in X$, by Theorem 3.9, we obtain

$$\begin{aligned} (x * y) * z &= (z * y) * x \\ &= x * (z * y) \\ &= x * (y * z). \end{aligned}$$

Therefore, $(X, *, 0)$ is an associative.
Conversely, assume that $(X, *, 0)$ is associative.
From Theorem 3.3 and (BN3), we obtain

$$\begin{aligned}(x * y) * z &= x * (y * z) \\ &= (0 * x) * (y * z) \\ &= (z * y) * x.\end{aligned}$$

This completes the proof. □

Theorem 3.11. *If $(X, *, 0)$ is a BN-algebra with the condition (P), then it is a group.*

Proof. Let $(X, *, 0)$ is a BN-algebra with the condition (P). From Theorem 3.10, X is an associative. Next, we will consider identity properties of group. From (BN2) and Theorem 3.3, 0 is an identity element. Finally, from (BN1), x is an inverse element of x on X. Therefore, $(X, *, 0)$ is a group. □

4 Discussion

In this article, we focus on basic definition and theorem on BN-algebra which these properties lead to these novel additional identities of BN-algebra. Particularly, we introduce some condition as a condition(P): $(x * y) * z = (z * y) * x$ in order to show the relation between BN-algebra and group.

Acknowledgment. The authors are highly grateful to the referees for their valuable comments and suggestions. This article was supported by the Department of Mathematics, Faculty of Science and Technology, Surindra Rajabhat University.

References

[1] A. Walendziak, A note on normal subalgebras in B-algebras, Sci. Math. Japonicae, Online e-2005 (2005) 49–53.

[2] A. Walendziak, On BF-algebras, Math Slovaca 57 (2007) 119–128.

[3] A. Walendziak, Some axiomatizations of B-algebras, Math Slovaca 56 (2006) 301–306.

[4] C.B. Kim, On BN-algebras, KYUNGPOOK Math. J. 53 (2013) 175–184.

[5] C. B. Kim and H. S. Kim, On BM-algebras, Sci. Math. Japonicae 63 (2006) 421–427.

[6] C. B. Kim and H. S. Kim, On BG-algebras, Demo. Math. 41 (2008) 497–505.

[7] H. S. Kim, Y. H. Kim and J. Neggers, Coxeters and pre-Coxeter algebras in Smarandache setting, Honam Math. J. 26 (2004) 251–272.

[8] J. Neggers and H. S. Kim, A fundamental theorem of B-homomorphism for B-algebras, Intern. Math. J. 2 (2002) 207–214.

[9] J. Neggers and H. S. Kim, On B-algebras, Mate. Vesnik 54 (2002) 21–29.

[10] J. Neggers and H. S. Kim, On β-algebras, Math. Slovaca 52 (2002) 517–530.

[11] J. R. Cho and H. S. Kim, On B-algebras and quasigroups, Quasigroups and related systems 8 (2001) 1–6.

[12] X. Zhang and R. Ye, BZ-algebra and group, J. Math. Phy. Sci. 29 (1995) 223–233.

[13] Y. B. Jun, E. H. Roh and H. S. Kim, On BH-algebra, Sci. Math. 1 (1998) 347–354.

A Note on Endomorphism Categories

Korn Kruaykitanon[1,*] and Phichet Chaoha[1]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

For each category C, there is a category $\mathcal{M}(\mathsf{C})$ of endomorphisms in C. We proved that $\mathcal{M}(\mathsf{C})$ is (co)complete whenever C is. For suitable categories C, we constructed fixed point, n-periodic point, orbit, and n-orbit functors defined on $\mathcal{M}(\mathsf{C})$ and established adjunctions involving these functors. We proved that the construction of $\mathcal{M}(\mathsf{C})$ from C is functorial in that it defines an endofunctor $\mathcal{M}$ on the category CAT of locally small categories. Moreover, $\mathcal{M}$ defines a monad on CAT and preserves 2-categorical structure of CAT. Some results when C is the category of topological spaces are also discussed.

Keywords: Endomorphism category, Fixed Point Functor, Orbit Functor
2020 MSC: 18B99, 18A30, 18F60

1 Introduction

Let $f : X \to X$ and $g : Y \to Y$ be continuous self-maps on topological spaces X and Y. We say that f and g are **topological conjugate** if there exists a homeomorphism $\alpha : X \to Y$ such that $\alpha \circ f = g \circ \alpha$ (see [1]). In this case, the dynamical properties of f and g are identical (see [1]), and their fixed point spaces and orbit spaces are homeomorphic under the restriction of α (see Theorem 1.1 [2]).

One can construct the category $\mathcal{M}(\mathsf{Top})$ of endomorphisms of topological spaces whose objects are pairs (X, f), where X is a topological space and $f : X \to X$ is an endomorphism, i.e. continuous self-map, on X, and whose morphisms are continuous maps commuting a certain diagram comprised of the underlying endomorphisms. The notion of isomorphism in the category $\mathcal{M}(\mathsf{Top})$ is precisely topological conjugacy. The authors, interested in extending the results in [2], are thus prompted to study properties of the category $\mathcal{M}(\mathsf{Top})$.

It turns out that the construction of $\mathcal{M}(\mathsf{Top})$ can be generalized further to the category $\mathcal{M}(\mathsf{C})$ of endomorphisms in a general category C (see e.g. [3, 4]). In this paper, we investigate the categorical properties of the categories of endomorphisms as well as the functors defined on them. Along the way, we will obtain categorification of some results in [2].

Surprisingly, the construction of the categories of endomorphisms is functorial in that it defines an endofunctor $\mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$, which, together with appropriate natural transformations, forms a categorical algebraic structure known as a monad. In the end, we proved that the functor $\mathcal{M}$ preserves 2-categorical structure of CAT and hence preserves adjunctions.

*Speaker and corresponding author (korn.kruaykitanon@gmail.com)
To the loving memory of Suwit Kruaykitanon, the first author's beloved father.

Organization

This paper is organized as follows. In Section 2, we review the basics of category theory. In Section 3, we recall the construction of the category $\mathcal{M}(\mathsf{C})$ of endomorphisms in C and prove that $\mathcal{M}(\mathsf{C})$ is (co)complete whenever C is. In Section 4, we construct for suitable categories C, fixed point, and n-periodic point functors $\mathcal{M}(\mathsf{C}) \to \mathsf{C}$ and prove that they are right adjoint. In Section 5, we construct for suitable categories C, orbit and n-orbit functors $\mathcal{M}(\mathsf{C}) \to \mathsf{C}$ and prove that they are left adjoint. In Section 6, we construct an endomorphism functor $\mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$ and show that it defines a monad and is a strict 2-functor.

2 Recapitulation of Category Theory

We review the basics of category theory required to understand the result of this paper. For complete details, see standard references, e.g. [5] or [6].

A **category** C consists of objects, morphisms each with specified domain and codomain (write $f : X \to Y$, $X \xrightarrow{f} Y$ for 'f is a morphism in C with domain X and codomain Y' and $\mathsf{C}(X, Y)$ for the collection of morphisms $X \to Y$), and a composition law, specifying for each pair of composable morphisms $X \xrightarrow{f} Y \xrightarrow{g} Z$ a morphism $X \xrightarrow{g \circ f} Z$, such that the composition law is **associative**, i.e. $h \circ (g \circ f) = (h \circ g) \circ f$ whenever these compositions are defined, and **unital**, i.e. for any object X, there is an identity morphism $X \xrightarrow{1_X} X$, and $f \circ 1_X = f = 1_Y \circ f$ whenever these compositions are defined. We write $X \in \mathsf{C}$ for 'X is an object of C'. A category is said to be **small** if it has set-many morphisms, and is **locally small** if, between each pair of objects, there are only set-many morphisms. An **isomorphism** is a morphism f for which there is a morphism g such that $g \circ f$ and $f \circ g$ are identities. We write $X \cong Y$ if there is an isomorphism $X \to Y$.

For example, there is a category Top whose objects are topological spaces, whose morphisms are continuous maps, and whose composition law is the usual one. Since between any two topological spaces, there are set-many maps, hence only set-many continuous maps, the category Top is locally small. Since any set can be made into a topological space by putting the discrete topology, there are more than set-many objects in Top, hence more than set-many identity morphisms, so the category Top is not small. Isomorphisms in Top are precisely homeomorphisms.

An object $X \in \mathsf{C}$ is **initial** if for any object $Y \in \mathsf{C}$, there is a unique morphism $X \to Y$. Dually, an object $X \in \mathsf{C}$ is **terminal** if for any object $Y \in \mathsf{C}$, there is a unique morphism $Y \to X$. For example, the empty space is an initial object in Top, and any one-point space is a terminal object in Top.

Let C and D be categories. A **functor** $F : \mathsf{C} \to \mathsf{D}$ consists of a map sending each object in C to an object in D, and for each $X, Y \in \mathsf{C}$, a map $\mathsf{C}(X, Y) \to \mathsf{D}(F(X), F(Y))$, such that $F(1_X) = 1_{F(X)}$ and $F(g \circ f) = F(g) \circ F(f)$ whenever these are defined. For ease of reading, one often omits the parenthesis for a functor applied to objects and morphisms, i.e. write simply just FX and Ff instead of $F(X)$ and $F(f)$. Since an isomorphism is witnessed by a commutative diagram, any functor preserves isomorphisms (cf. Lemma 1.3.8 [5]).

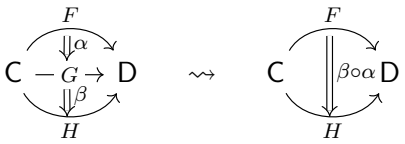
Let $F, G : \mathsf{C} \to \mathsf{D}$ be functors. A **natural transformation** $\eta : F \Rightarrow G$ is a family $\{FX \xrightarrow{\eta_X} GX\}_{X \in \mathsf{C}}$ of morphisms in D indexed by C such that for any $f \in \mathsf{C}(X, Y)$, $\eta_Y \circ Ff = Gf \circ \eta_X$,

i.e. the diagram

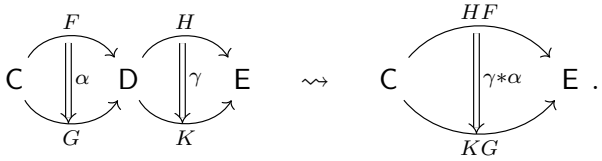
$$\begin{array}{ccc} FX & \xrightarrow{\eta_X} & GX \\ \downarrow{\scriptstyle Ff} & & \downarrow{\scriptstyle Gf} \\ FY & \xrightarrow{\eta_Y} & GY \end{array}$$

commutes. A **natural isomorphism** is a natural transformation η whose components η_X are isomorphisms for all X. We write $F \cong G$ if there exists a natural isomorphism $\eta : F \Rightarrow G$.

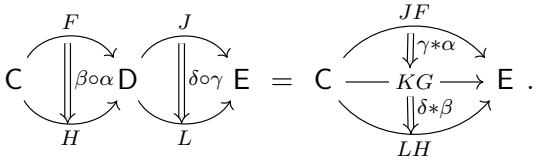
A **strict 2-category** is a category C together with an additional 2-dimensional structure (see e.g. Definition 1.7.8 [5], Definition 1.4.8 [6]). A prime example is the category CAT of locally small categories. (CAT itself is not locally small.) It has locally small categories as 0-cells (or objects), functors as 1-cells (or morphisms), and natural transformations as 2-cells. There are two composition laws for 2-cells: the **vertical composition**,



and the **horizontal composition**,



These two interact nicely via the interchange law:



A **strict 2-functor** is a functor that also preserves the 2-dimensional structure.

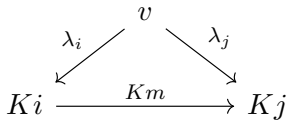
Let C and D be locally small categories and $F : \mathsf{C} \rightleftarrows \mathsf{D} : G$ be functors. An **adjunction** $F \dashv G$ consists of a family of bijections $\flat : \mathsf{D}(FX, Y) \rightleftarrows \mathsf{C}(X, GY) : \sharp$ for $X \in \mathsf{C}$ and $Y \in \mathsf{D}$, such that for any morphism $f : X \leftarrow X'$ and $g : Y \to Y'$, the diagram

$$\begin{array}{ccc} \mathsf{D}(FX, Y) & \underset{\sharp}{\overset{\flat}{\rightleftarrows}} & \mathsf{C}(X, GY) \\ {\scriptstyle g\circ -\circ Ff}\downarrow & & \downarrow{\scriptstyle Gg\circ -\circ f} \\ \mathsf{D}(FX', Y') & \underset{\sharp}{\overset{\flat}{\rightleftarrows}} & \mathsf{C}(X', GY') \end{array}$$

commutes. Since an adjunction is witnessed by a certain commutative 2-diagrams, it is preserved by any strict 2-functor $\mathsf{CAT} \to \mathsf{CAT}$ (cf. Proposition 4.4.6 [5]).

Let J be a small category. A **diagram** in C of shape J is a functor $K : \mathsf{J} \to \mathsf{C}$. A **cone over** the diagram K consists of an object v, called a **vertex**, and morphisms $\{v \xrightarrow{\lambda_i} Ki\}_{i \in \mathsf{C}}$ so that

for any morphism $m : i \to j$ in J, $\lambda_j = Km \circ \lambda_i$, i.e. the diagram



commutes. (In fact, the reason for naming it a cone is the shape of the diagram above.) A **limit** of the diagram K is a cone $\{v \xrightarrow{\lambda_i} Ki\}_{i\in\mathsf{J}}$ such that for any cone $\{w \xrightarrow{\mu_i} Ki\}_{i\in\mathsf{J}}$, there is a unique morphism, called the **factorization**, $f : w \to v$ such that $\lambda_i \circ f = \mu_i$ for all $i \in \mathsf{J}$. It can be shown that such a limit is unique up to a unique canonical isomorphism. A **limit object** is the vertex of a limit cone. A limit of discrete shape is called a **product**. A limit of shape $\bullet \rightrightarrows \bullet$ is called an **equalizer**. A category is **complete** if every diagram has a limit.

For example, the category Top is complete. A limit of the diagram consisting of two objects X and Y and no morphisms between them is nothing but the product space $X \times Y$ together with the canonical projection maps. An equalizer of the diagram $X \underset{g}{\overset{f}{\rightrightarrows}} Y$ is given by the subspace $\{x \in X : f(x) = g(x)\}$ of X together with the inclusion map into X.

One can **dualize** the definition of cones over a diagram and limits by reversing the direction of morphisms and obtain the definition of **cones under** a diagram and **colimits**. A colimit of discrete shape is called a **coproduct**. A colimit of shape $\bullet \rightrightarrows \bullet$ is called a **coequalizer**. A category is **cocomplete** if every diagram has a colimit.

An important result in category theory is the **duality principle** which states that if a theorem about categories is true, then its dual version, obtained by appropriately reversing the direction of morphisms in the statement, also holds for all categories as well. The key is that the proofs of dual theorems can be extracted from the proofs of original theorems by carefully reversing the direction of morphisms that appear in the proof. For this reason, one often omits statements and proofs of dual theorems. For example, uniqueness up to isomorphism of limits immediately dualizes to uniqueness up to isomorphism of colimits.

A functor $F : \mathsf{D} \to \mathsf{C}$ is said to **strictly create limits** if for any diagram K in D, any limit cone over the diagram FK in C has a unique **lift** along F, i.e. a cone over K whose image under F is the starting limit cone, and such a unique lift is a limit cone of K. In this case, any limit that exists in C is lifted to a limit in D, so the completeness of C implies that of D.

Let C be a category. A **monad** on C consists of an endofunctor $T : \mathsf{C} \to \mathsf{C}$ and natural transformations $\eta : 1_{\mathsf{C}} \Rightarrow T$ and $\mu : T^2 \Rightarrow T$ so that $\mu \circ \mu T = \mu \circ T\mu$ and $\mu \circ T\eta = 1_T = \mu \circ \eta T$ as natural transformations. A monad (T, η, μ) is, roughly speaking, an abstract algebraic structure on C carried out in the background by T whose properties are witnessed by η and μ. For example, given a topological space X, the power set $P(X)$ with the inclusions as morphisms is a category on which the closure operator $\mathrm{Cl}_X : P(X) \to P(X)$ defines a monad.

3 The Category of Endomorphisms

3.1 Construction of the Category of Endomorphisms

Construction 3.1 (See also [3,4])**.** Let C be a category. The category $\mathcal{M}(\mathsf{C})$ of endomorphisms in C (or the endomorphism category of C) is constructed as follows.

- An object of $\mathcal{M}(\mathsf{C})$ is a pair (X, f), where X is an object of C and f is a morphism $X \to X$ in C.
- A morphism $(X, f) \xrightarrow{\alpha} (Y, g)$ in $\mathcal{M}(\mathsf{C})$ is a morphism $X \xrightarrow{\alpha} Y$ in C such that $\alpha \circ f = g \circ \alpha$, i.e. the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \alpha} \\ Y & \xrightarrow{g} & Y \end{array}$$

  in C commutes.
- The composition law of morphisms $\mathcal{M}(\mathsf{C})$ is given by the composition law in C.

It can be shown that if C is locally small, then so is $\mathcal{M}(\mathsf{C})$.

Remark 3.2. An isomorphism $\alpha : (X, f) \to (Y, g)$ in $\mathcal{M}(\mathsf{C})$ is precisely an isomorphism $\alpha : X \to Y$ in C such that $\alpha \circ f = g \circ \alpha$. In particular, an isomorphism in $\mathcal{M}(\mathsf{Top})$ precisely encodes topological conjugacy.

Proposition 3.3. *The identity morphism functor $I : \mathsf{C} \to \mathcal{M}(\mathsf{C})$, given on objects by $X \mapsto (X, 1_X)$ and identity on morphisms, is fully faithful and injective on objects. Hence, the category C is embedded into $\mathcal{M}(\mathsf{C})$ via I.*

Proof. It follows from the observation that any morphism $f : X \to Y$ is automatically a morphism $(X, 1_X) \to (Y, 1_Y)$ since $f \circ 1_X = f = 1_Y \circ f$. □

3.2 Completeness and Cocompleteness

Theorem 3.4. *The forgetful functor $U : \mathcal{M}(\mathsf{C}) \to \mathsf{C}$ strictly creates limits and colimits. Hence, if C is complete (or cocomplete, respectively), then $\mathcal{M}(\mathsf{C})$ is complete (or cocomplete, respectively).*

Proof. By duality principle, we will prove only results for limits. Let $K : \mathsf{J} \to \mathcal{M}(\mathsf{C})$ be a diagram. Say $Ki = (X_i, f_i) \in \mathcal{M}(\mathsf{C})$ for $i \in \mathsf{J}$. Let $\{X \xrightarrow{\pi_i} X_i\}_{i \in \mathsf{J}}$ be a limit cone over the diagram $UK : \mathsf{J} \to \mathsf{C}$.

- uniqueness of the lift along U of the limit cone : By the description of U, the lift must be of the form $\{(X, f) \xrightarrow{\pi_i} (X_i, f_i)\}_{i \in \mathsf{J}}$. It remains to argue why $f : X \to X$ is uniquely determined.

  Note that $\{X \xrightarrow{f_i \circ \pi_i} X_i\}$ is a cone over the diagram UK: for any morphism $m : i \to j$ in J, commutativity of the outer diagram

$$\begin{array}{ccc} & X & \\ {\scriptstyle \pi_i}\swarrow & & \searrow{\scriptstyle \pi_j} \\ X_i & \xrightarrow{Km} & X_j \\ \downarrow{\scriptstyle f_i} & & \downarrow{\scriptstyle f_j} \\ X_i & \xrightarrow{Km} & X_j \end{array}$$

follows from the top triangle (by the cone condition of $\{X \xrightarrow{\pi_i} X_i\}_{i\in\mathsf{J}}$ and the bottom square (by the morphism condition of $Km : (X_i, f_i) \to (X_j, f_j)$).

Now, $f : X \to X$ is the unique factorization of the cone $\{X \xrightarrow{f_i\circ\pi_i} X_i\}$ through the limit cone $\{X \xrightarrow{\pi_i} X_i\}_{i\in\mathsf{J}}$:

$$\begin{array}{ccc} X & \overset{f}{\dashrightarrow} & X \\ \downarrow{\scriptstyle \pi_i} & & \downarrow{\scriptstyle \pi_i} \\ X_i & \xrightarrow{f_i} & X_i \end{array} .$$

Hence, f is uniquely determined.

- existence of the lift along U of the limit cone : Construct f as in the uniqueness part. It is left to verify the cone condition of $\{(X, f) \xrightarrow{\pi_i} (X_i, f_i)\}_{i\in\mathsf{J}}$. This follows easily from cone condition of the cone $\{X \xrightarrow{\pi} X_i\}_{i\in\mathsf{J}}$.

- the lift is a limit cone : Let us show that $\{(X, f) \xrightarrow{\pi_i} (X_i, f_i)\}_{i\in\mathsf{J}}$ is a limit cone of the diagram $K : \mathsf{J} \to \mathcal{M}(\mathsf{C})$. Suppose a cone $\{(Y, g) \xrightarrow{\lambda_i} (X_i, f_i)\}_{i\in\mathsf{J}}$ over K is given. We have to show existence and uniqueness of the factorization $\xi : Y \to X$ of cones over K:

  - uniqueness : Such a factorization ξ is, in particular, a factorization of the cone $\{Y \xrightarrow{\lambda_i} X_i\}_{i\in\mathsf{J}}$ through the limit cone $\{X \xrightarrow{\pi_i} X_i\}_{i\in\mathsf{J}}$ of UK, so it is unique.
  - existence : Let $\xi : Y \to X$ be the factorization of the cone $\{Y \xrightarrow{\lambda_i} X_i\}_{i\in\mathsf{J}}$ through the limit cone $\{X \xrightarrow{\pi_i} X_i\}_{i\in\mathsf{J}}$ of UK. To see that ξ is a factorization of cones over K, it remains to argue why $\xi : (Y, g) \to (X, f)$ is a morphism in $\mathcal{M}(\mathsf{C})$, i.e. $\xi \circ g = f \circ \xi$. To see this, compute for each $i \in \mathsf{J}$,

$$\pi_i \circ \xi \circ g = \lambda_i \circ g = f_i \circ \lambda_i = \pi_i \circ f \circ \xi.$$

    Since $\{X \xrightarrow{\pi_i} X_i\}_{i\in\mathsf{J}}$ is a limit cone, we must have $\xi \circ g = f \circ \xi$ as desired.

In summary, U strictly creates limits. □

Example 3.5. If X is an initial (or terminal, respectively) object in C, then $(X, 1_X)$ is also initial (or terminal, respectively) in $\mathcal{M}(\mathsf{C})$.

Example 3.6. Since Top is complete and cocomplete (see e.g. Proposition 3.5.2 [5]), the category $\mathcal{M}(\mathsf{Top})$ is complete and cocomplete.

For example, the categorical product of the family $\{(X_i, f_i)\}_{i\in\mathsf{J}}$ of objects in $\mathcal{M}(\mathsf{Top})$ is given by $(\prod_{i\in\mathsf{J}} X_i, \prod_{i\in\mathsf{J}} f_i)$, where

- $\prod_{i\in\mathsf{J}} X_i$ is the product space of $\{X_i\}_{i\in\mathsf{J}}$,
- the map $\prod_{i\in\mathsf{J}} f_i : \prod_{i\in\mathsf{J}} X_i \to \prod_{i\in\mathsf{J}} X_i$ is given by $(x_i)_{i\in\mathsf{J}} \mapsto (f_i(x_i))_{i\in\mathsf{J}}$,

together with the canonical projection morphisms $\pi_j : (\prod_{i\in\mathsf{J}} X_i, \prod_{i\in\mathsf{J}} f_i) \to (X_j, f_j)$.

Dually, the categorical coproduct of the family $\{(X_i, f_i)\}_{i\in\mathsf{J}}$ of objects in $\mathcal{M}(\mathsf{Top})$ is given by $(\coprod_{i\in\mathsf{J}} X_i, \coprod_{i\in\mathsf{J}} f_i)$ together with the canonical inclusion morphisms $\iota_j : (X_j, f_j) \to (\coprod_{i\in\mathsf{J}} X_i, \coprod_{i\in\mathsf{J}} f_i)$.

4 The Fixed Point and Period Point Functors

Let X be a set and $f : X \to X$ be a map. The **fixed point set** $X^f := \{x \in X : f(x) = x\}$ of f, together with the canonical inclusion map, is the equalizer of the diagram (in the category of sets) $X \underset{1_X}{\overset{f}{\rightrightarrows}} X$. This shows that the fixed point set can be generalized to general categories as the equalizers of endomorphisms and identity morphisms.

In this section, suppose that the category C has all equalizers.

4.1 The Fixed Point Functor

Construction 4.1. Construct the **fixed point functor** $\wedge : \mathcal{M}(\mathsf{C}) \to \mathsf{C}$ by specifying its action

- on objects : for each pair $(X, f) \in \mathcal{M}$, define $\wedge(X, f)$ to be the equalizer object X^f in an arbitrary but fixed[†] equalizer diagram

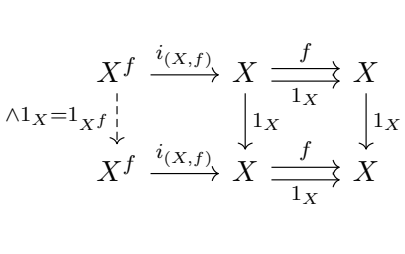
$$X^f \xrightarrow{i_{(X,f)}} X \underset{1_X}{\overset{f}{\rightrightarrows}} X \ ;$$

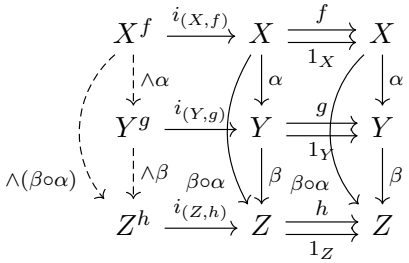
- on morphisms : a morphism $(X, f) \xrightarrow{\alpha} (Y, g)$ is mapped to the unique morphism $\wedge\alpha : X^f \to Y^g$ between the equalizers commuting the diagram

$$\begin{array}{ccccc} X^f & \xrightarrow{i_{(X,f)}} & X & \underset{1_X}{\overset{f}{\rightrightarrows}} & X \\ \downarrow{\scriptstyle \wedge\alpha} & & \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \alpha} \\ Y^g & \xrightarrow{i_{(Y,g)}} & Y & \underset{1_Y}{\overset{g}{\rightrightarrows}} & Y \end{array}$$

(obtained from the universal property of the equalizer Y^g)

Functoriality comes from the uniqueness of factorization through equalizers and the following diagram chasings:





Remark 4.2. Functoriality of the fixed point functor is a special case of functoriality of the limit functor $\lim_{\mathsf{J}} : [\mathsf{J}, \mathsf{C}] \to \mathsf{C}$ when $\mathsf{J} = \bullet \rightrightarrows \bullet$ (cf. Proposition 3.1.7 [5]).

Corollary 4.3. *If $(X, f) \cong (Y, g)$ as objects in $\mathcal{M}(\mathsf{C})$, then $X^f \cong Y^g$ as objects in C.*

Proof. It follows from the basic theorem in category theory that every functor preserves isomorphisms. In particular, $\wedge : \mathcal{M}(\mathsf{C}) \to \mathsf{C}$ preserves isomorphisms. □

[†]A diagram could possibly have more than one equalizers. Nevertheless, they are unique up to a unique canonical isomorphism. In order to define a functor, we need to choose a distinguished equalizer for each diagram.

Remark 4.4. When $\mathsf{C} = \mathsf{Top}$, Corollary 4.3 is precisely Theorem 1.1(1) [2].

Theorem 4.5. *There is an adjunction* $\mathsf{C} \underset{\wedge}{\overset{I}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathcal{M}(\mathsf{C})$ *. In particular, the functor* $\wedge$ *is right adjoint.*

Proof. For each $X \in \mathsf{C}$ and $(Y, g) \in \mathcal{M}(\mathsf{C})$, for any given $\alpha : (X, 1_X) \to (Y, g)$, there is a unique morphism $\alpha^\flat : X \to Y^g$ such that $i_{(Y,g)} \circ \alpha^\flat = \alpha$.

$$\begin{array}{ccccc} Y^g & \xrightarrow{i_{(Y,g)}} & Y & \underset{1_Y}{\overset{g}{\rightrightarrows}} & Y \\ {\scriptstyle \alpha^\flat}\uparrow & \nearrow_{\alpha} & & & \\ X & & & & \end{array} .$$

We get a bijection

$$\begin{aligned} \flat : \mathcal{M}(\mathsf{C})((X, 1_X), (Y, g)) &\leftrightarrows \mathsf{C}(X, Y^g) : \sharp, \\ \alpha &\longmapsto \alpha^\flat, \\ i_{(Y,g)} \circ \beta &\longmapsfrom \beta. \end{aligned}$$

It is left to show naturality: for any $X \xleftarrow{\phi} X'$ and $(Y, g) \xrightarrow{\psi} (Y', g')$, we need to show that the diagram

$$\begin{array}{ccc} \mathcal{M}(\mathsf{C})((X, 1_X), (Y, g)) & \xleftarrow[\sharp]{} & \mathsf{C}(X, Y^g) \\ \downarrow{\scriptstyle \psi \circ - \circ I\phi} & & \downarrow{\scriptstyle \wedge\psi \circ - \circ \phi} \\ \mathcal{M}(\mathsf{C})((X', 1_{X'}), (Y', g')) & \xleftarrow[\sharp]{} & \mathsf{C}(X', Y'^{g'}) \end{array}$$

commutes. For any $\beta \in \mathsf{C}(X, Y^g)$, we have

$$\psi \circ \beta^\sharp \circ I\phi = \psi \circ i_{(Y,g)} \circ \beta \circ \phi = i_{(Y,g)} \circ \wedge\psi \circ \beta \circ \phi = (\wedge\psi \circ \beta \circ \phi)^\sharp.$$

This completes the proof. □

Corollary 4.6. *For any* $(X, f), (Y, g) \in \mathcal{M}(\mathsf{C})$, $(X \times Y)^{f \times g} \cong X^f \times Y^g$ *as objects in* C.

Proof. This is a consequence of the classical result in category theory that a right adjoint functor preserves limits. In particular, $\wedge$ preserves binary products. □

Example 4.7. For any $(X, f), (Y, g) \in \mathcal{M}(\mathsf{Top})$, $(X \times Y)^{f \times g} \cong X^f \times Y^g$ as spaces.

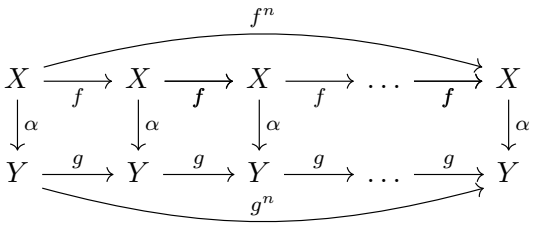
4.2 The Period Point Functor

Let $n \geq 2$. An **n-periodic point** of a map $f : X \to X$ is a point $x \in X$ such that $f^n(x) = x$. Notice that an n-periodic point of f is precisely the fixed point of the map f^n obtained from iterating f n times.

Construction 4.8. The n**-iterator functor** is defined as

$$\begin{aligned} \square^n : \mathcal{M}(\mathsf{C}) &\longrightarrow \mathcal{M}(\mathsf{C}), \\ ((X, f) \xrightarrow{\alpha} (Y, g)) &\longmapsto ((X, f^n) \xrightarrow{\alpha} (Y, g^n)). \end{aligned}$$

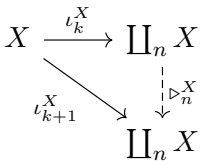
The functor $\square^n$ is well-defined on morphisms by gluing diagrams:



Construction 4.9. The n**-periodic point functor** is the composite functor $\wedge^n := \wedge \circ \square^n : \mathcal{M}(\mathsf{C}) \to \mathcal{M}(\mathsf{C})$.

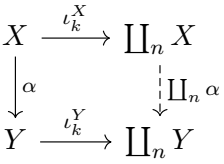
Construction 4.10. Suppose that C has binary coproducts. Fix $X \in \mathsf{C}$.

Construct the n-th copower $\coprod_n X \in \mathsf{C}$ with inclusion morphisms $\{X \xrightarrow{\iota_k^X} \coprod_n X\}_k$. Let $\triangleright_n^X : \coprod_n X \to \coprod_n X$ be the unique morphism commuting the diagram for all k,



where the indices of inclusion morphisms is considered modulo n. Thus, we have an object $\left(\coprod_n X, \triangleright_n^X\right) \in \mathcal{M}(\mathsf{C})$.

For any morphism $\alpha : X \to Y$ in C, let $\coprod_n \alpha : \coprod_n X \to \coprod_n Y$ be the unique morphism commuting the diagram for all k,



In fact, we will show that $\coprod_n \alpha : \left(\coprod_n X, \triangleright_n^X\right) \to \left(\coprod_n Y, \triangleright_n^Y\right)$ is a morphism in $\mathcal{M}(\mathsf{C})$.

Now, construct a functor

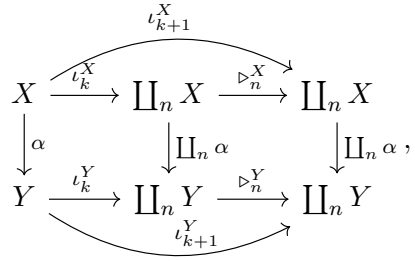
$$\begin{aligned} L_n : \mathsf{C} &\longrightarrow \mathcal{M}(\mathsf{C}), \\ (X \xrightarrow{\alpha} Y) &\longmapsto \left(\left(\coprod_n X, \triangleright_n^X \right) \xrightarrow{\coprod_n \alpha} \left(\coprod_n Y, \triangleright_n^Y \right) \right). \end{aligned}$$

Proposition 4.11. *$L_n : \mathsf{C} \to \mathcal{M}(\mathsf{C})$ constructed in Construction 4.10 is well-defined and indeed defines a functor.*

Proof. First, let us show that $\coprod_n \alpha : (\coprod_n X, \triangleright_n^X) \to (\coprod_n Y, \triangleright_n^Y)$ is a morphism in $\mathcal{M}(\mathsf{C})$, i.e. $\coprod_n \alpha \circ \triangleright_n^X = \triangleright_n^Y \circ \coprod_n \alpha$, so that L_n is well-defined. This is by computing for each k, that

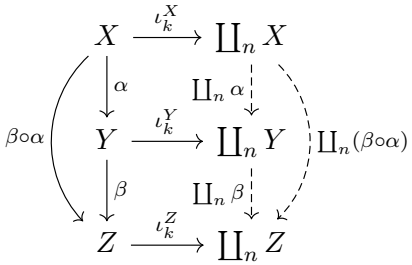
$$\coprod_n \alpha \circ \triangleright_n^X \circ \iota_k^X = \triangleright_n^Y \circ \coprod_n \alpha \circ \iota_k^X .$$

Indeed, by chasing the diagram below using the property of $\triangleleft_n$ and $\coprod_n \alpha$,



we have the desired equality.

Functoriality of L_n follows from the following diagram chasing



In summary, $L_n : \mathsf{C} \to \mathcal{M}(\mathsf{C})$ defines a functor. □

Theorem 4.12. *Suppose that* C *has binary coproducts. Then there is an adjunction*

$$\mathsf{C} \underset{\wedge^n}{\overset{L_n}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathcal{M}(\mathsf{C}) \ ,$$

where L_n *is constructed in Construction 4.10. In particular, the functor* $\wedge^n : \mathcal{M}(\mathsf{C}) \to \mathsf{C}$ *is right adjoint.*

Proof. Establish a one-to-one correspondence

$$\flat : \mathcal{M}(\mathsf{C}) \left(\left(\coprod_n X, \triangleright_n^X \right), (Y, g) \right) \longrightarrow \mathsf{C}(X, Y^{g^n}) : \sharp$$

as follows: let $Y^{g^n} \overset{i_{(Y,g^n)}}{\to} Y$ be the canonical monomorphism.

- For each $\alpha \in \mathcal{M}(\mathsf{C})((\coprod_n X, \triangleright_n^X), (Y, g))$, since

$$g^n \circ \alpha = \alpha \circ (\triangleleft_n^X)^n = \alpha = 1_Y \circ \alpha,$$

i.e. the diagram

$$\begin{array}{ccc} \coprod_n X & \xrightarrow{(\triangleright_n^X)^n = 1_{\coprod_n X}} & \coprod_n X \\ \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \alpha} \\ Y & \underset{1_Y}{\overset{g^n}{\rightrightarrows}} & Y \end{array}$$

commutes, there is a unique morphism $\alpha^\flat \in \mathsf{C}(X, Y^{g^n})$ such commuting the diagram

$$\begin{array}{ccc} X & \xrightarrow{\iota_0} & \coprod_n X \\ \downarrow{\scriptstyle \alpha^\flat} & & \downarrow{\scriptstyle \alpha} \\ Y^{g^n} & \xrightarrow{i_{(Y,g^n)}} & Y \end{array}$$

- For each $\beta \in \mathsf{C}(X, Y^{g^n})$, let $\beta^\sharp : \coprod_n X \to Y$ be the unique morphism (obtained from the universal property of coproduct) commuting the diagram

$$\text{for all } k: \quad \begin{array}{ccccc} X & \xrightarrow{\beta} & Y^{g^n} & \xrightarrow{i_{(Y,g^n)}} & Y \\ \downarrow{\scriptstyle \iota_k^X} & & & & \downarrow{\scriptstyle g^k} \\ \coprod_n X & & \xrightarrow{\beta^\sharp} & & Y \end{array}$$

This is well-defined for exponents of g modulo n since e equalizes g^n and 1_Y. To see that $\beta^\sharp \in \mathcal{M}(\mathsf{C})((\coprod_n X, \triangleright_n^X), (Y, g))$, consider the diagram

$$\text{for all } k: \quad \begin{array}{ccccccc} X & & \xrightarrow{\beta} & & Y^{g^n} & \xrightarrow{i_{(Y,g^n)}} & Y \\ & \searrow & & & & \nearrow & \\ & & X & \xrightarrow{\beta} & Y^{g^n} & \xrightarrow{i_{(Y,g^n)}} & Y \\ \downarrow{\scriptstyle \iota_{k-1}^X} & & \downarrow{\scriptstyle \iota_k^X} & & & & \downarrow{\scriptstyle g^k} \\ & & \coprod_n X & & \xrightarrow{\beta^\sharp} & & Y \\ & \nearrow{\scriptstyle \triangleright_n^X} & & & & \nwarrow{\scriptstyle g} & \\ \coprod_n X & & & & \xrightarrow{\beta^\sharp} & & Y \end{array} \qquad g^{k-1}$$

(upper) trivially (left) by def of $\triangleright_n^X$ (right) $g^k = g \circ g^{k-1}$ (inner) by def of $\beta^\sharp$ (k-th component) (outer) by def of $\beta^\sharp$ ($(k-1)$-th component). By the universal property of coproduct, we conclude that the lower diagram commutes.

Next, we will show that $\flat$ and $\sharp$ are mutually inverse.

- $(\beta^\sharp)^\flat = \beta$: By construction of $\beta^\sharp$, the diagram

$$\begin{array}{ccc} X & \xrightarrow{\iota_0} & \coprod_n X \\ \downarrow{\scriptstyle \beta} & & \downarrow{\scriptstyle \beta^\sharp} \\ Y^{g^n} & \xrightarrow{i_{(Y,g^n)}} & Y \end{array}$$

commutes. By uniqueness of factorization through Y^{g^n}, we have $(\beta^\sharp)^\flat = \beta$.

- $(\alpha^\flat)^\sharp = \alpha$: Consider

$$\text{for all } k: \quad \begin{array}{ccccc} X & \xrightarrow{\iota_0} & \coprod_n X & \xrightarrow{(\triangleright_n^X)^k} & \coprod_n X \\ \downarrow{\scriptstyle \alpha^\flat} & & \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \alpha} \\ Y^{g^n} & \xrightarrow{i_{(Y,g^n)}} & Y & \xrightarrow{g^k} & Y \end{array}$$

(left) def of $\alpha^\flat$ (right) morphism property of α (upper) definition of $\triangleright_n^X$. Hence, the outer diagram commutes. We have

$$\text{for all } k: \quad \begin{array}{ccccc} X & \xrightarrow{\alpha^\flat} & Y^{g^n} & \xrightarrow{i_{(Y,g^n)}} & Y \\ \downarrow{\scriptstyle \iota_k} & & & & \downarrow{\scriptstyle g^k} \\ \coprod_n X & & \xrightarrow{\quad\alpha\quad} & & Y \end{array}$$

By uniqueness of such morphism, $(\alpha^\flat)^\sharp = \alpha$.

Let us check naturality of $\flat$. For any $X \xleftarrow{\phi} X'$ and $(Y,g) \xrightarrow{\psi} (Y',g')$, we need to show that the diagram

$$\begin{array}{ccc} \mathcal{M}(\mathsf{C})((\coprod_n X, \triangleright_n^X), (Y,g)) & \xrightarrow{\flat} & \mathsf{C}(X, Y^{g^n}) \\ \downarrow{\scriptstyle \psi \circ - \circ \coprod_n \phi} & & \downarrow{\scriptstyle \wedge^n \psi \circ - \circ \phi} \\ \mathcal{M}(\mathsf{C})((\coprod_n X', \triangleright_n^X), (Y',g')) & \xrightarrow{\flat} & \mathsf{C}(X', Y'^{g'^n}) \end{array}$$

commutes, i.e. for any $\alpha \in \mathcal{M}(\mathsf{C})((\coprod_n X, \triangleright_n^X), (Y,g))$,

$$(\psi \circ \alpha \circ \coprod_n \phi)^\flat = \wedge^n \psi \circ \alpha^\flat \circ \phi.$$

To achieve this, we show that $\wedge^n \psi \circ \alpha^\flat \circ \phi$ commutes the diagram

$$\begin{array}{ccc} X' & \xrightarrow{\iota_0} & \coprod_n X' \\ \downarrow{\scriptstyle \wedge^n \psi \circ \alpha^\flat \circ \phi} & & \downarrow{\scriptstyle \psi \circ \alpha \circ \coprod_n \phi} \\ Y'^{g'^n} & \xrightarrow{i_{(Y,g^n)}} & Y' \end{array}$$

Indeed, the diagram is obtained from gluing diagrams:

$$\begin{array}{ccc} X' & \xrightarrow{\iota_0} & \coprod_n X' \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \coprod_n \phi} \\ X & \xrightarrow{\iota_0} & \coprod_n X \\ \downarrow{\scriptstyle \alpha^\flat} & & \downarrow{\scriptstyle \alpha} \\ Y^{g^n} & \xrightarrow{i_{(Y,g^n)}} & Y \\ \downarrow{\scriptstyle \wedge^n \psi} & & \downarrow{\scriptstyle \psi} \\ Y'^{g'n} & \xrightarrow{i_{(Y',g'^n)}} & Y' \end{array}$$

(upper) def of $\coprod_n \phi$ (middle) def of $\alpha^\flat$ (lower) def of $\wedge^n \psi$. In total, $\flat$ is natural.

This completes the proof. □

5 The Orbit and n-Orbit Functors

Given a set X and a map $f : X \to X$, the **orbit set** X_f of f is obtained by identifying points of X under the equivalence relation on X generated by the relations $x \sim f(x)$ for $x \in X$. This construction can be viewed as the coequalizer of the diagram (in the category of sets) $X \underset{1_X}{\overset{f}{\rightrightarrows}} X$. Hence, it is a dual notion to the fixed point set of f.

For $n \geq 2$, we also have the **n-orbit set** X_{f^n} of f which can be obtained from forming the orbit set of $f^n : X \to X$. Thus, the concept of n-orbit sets is dual to that of n-periodic point sets.

The results from Section 4 on fixed point and n-periodic point functors dualize to those of the orbit and n-orbit functors. We will only provide sketches of proofs.

In this section, we assume that the category C has all coequalizers.

5.1 The Orbit Functor

Construction 5.1. Construct the **orbit functor** $\vee : \mathcal{M}(\mathsf{C}) \to \mathsf{C}$ by specifying its action

- on objects : for each pair $(X, f) \in \mathcal{M}$, define $\vee(X, f)$ to be the coequalizer object X_f in an arbitrary but fixed[‡] coequalizer diagram

$$X \underset{1_X}{\overset{f}{\rightrightarrows}} X \xrightarrow{q_{(X,f)}} X_f \ ;$$

- on morphisms : a morphism $(X, f) \xrightarrow{\alpha} (Y, g)$ is mapped to the unique morphism $\vee\alpha : X_f \to Y_g$ between the coequalizers commuting the diagram

$$\begin{array}{ccccc} X & \underset{1_X}{\overset{f}{\rightrightarrows}} & X & \xrightarrow{q_{(X,f)}} & X_f \\ \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \alpha} & & \vdots{\scriptstyle \vee\alpha} \\ Y & \underset{1_Y}{\overset{g}{\rightrightarrows}} & Y & \xrightarrow{q_{(Y,g)}} & Y_g \end{array}$$

(obtained from the universal property of the coequalizer X_f)

The proof of functoriality of $\vee$ is dual to that of $\wedge$ (see Construction 4.1).

Corollary 5.2. *If $(X, f) \cong (Y, g)$ as objects in $\mathcal{M}(\mathsf{C})$, then $X_f \cong Y_g$ as objects in C.*

Proof. It follows from the basic theorem in category theory that every functor preserves isomorphisms. In particular, $\vee : \mathcal{M}(\mathsf{C}) \to \mathsf{C}$ preserves isomorphisms. □

Remark 5.3. When $\mathsf{C} = \mathsf{Top}$, Corollary 5.2 is precisely Theorem 1.1(2) of [2].

The following theorem is dual to Theorem 4.5.

Theorem 5.4. *There is an adjunction* $\mathcal{M}(\mathsf{C}) \underset{I}{\overset{\vee}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathsf{C}$. *In particular, the functor $\vee$ is left adjoint.*

[‡]The remark concerning the fixed point functor also applies here.

Corollary 5.5. *For any* $(X, f), (Y, g) \in \mathcal{M}(\mathsf{C})$, $(X \sqcup Y)_{f \sqcup g} \cong X_f \sqcup Y_g$ *as objects in* C.

Proof. This is a consequence of the classical result in category theory that a left adjoint functor preserves colimits. In particular, $\vee$ preserves binary coproducts. □

Example 5.6. For any $(X, f), (Y, g) \in \mathcal{M}(\mathsf{Top})$, $(X \sqcup Y)_{f \sqcup g} \cong X_f \sqcup Y_g$ as objects in Top.

5.2 The n-Orbit Functor

Construction 5.7. The **n-orbit functor** is the composite functor $\vee^n := \vee \circ \square^n : \mathcal{M}(\mathsf{C}) \to \mathcal{M}(\mathsf{C})$.

Construction 5.8. Suppose that C has binary products. Fix $X \in \mathsf{C}$.

Construct the n-th power $\prod_n X \in \mathsf{C}$ with the projection morphisms $\{\pi_k : \prod_n X \to X\}_k$. Let $\lhd_n^X : \prod_n X \to \prod_n X$ be the unique morphism such that for all k, $\pi_k^X \circ \lhd_n^X := \pi_{k+1}^X$, where the indices of projection morphisms is modulo n.

This construction induces a functor

$$\begin{aligned} R_n : \mathsf{C} &\longrightarrow \mathcal{M}(\mathsf{C}), \\ (X \xrightarrow{\alpha} Y) &\longmapsto \left(\left(\prod_n X, \lhd_n^X \right) \xrightarrow{\prod_n \alpha} \left(\prod_n Y, \lhd_n^Y \right) \right). \end{aligned}$$

The proof that R_n is well-defined and is a functor is similar to that of L_n (cf. Proposition 4.11) and is omitted.

Remark 5.9. $L_1 = R_1 = I$ as functors $\mathsf{C} \to \mathcal{M}(\mathsf{C})$.

Theorem 5.10. *Suppose that* C *has binary products. Then there is an adjunction*

$$\mathcal{M}(\mathsf{C}) \underset{R_n}{\overset{\vee^n}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathsf{C} \ ,$$

where R_n *is constructed in Construction 5.8. In particular, the functor* $\vee^n : \mathcal{M}(\mathsf{C}) \to \mathsf{C}$ *is left adjoint.*

Proof. (sketch) Establish a one-to-one correspondence

$$\flat : \mathsf{C}(X_{f^n}, Y) \longrightarrow \mathcal{M}(\mathsf{C}) \left((X, f), \left(\prod_n Y, \lhd_n^Y \right) \right) : \sharp$$

as follows: let $X \overset{q_{(X,f)}}{\to} X_{f^n}$ be the canonical epimorphism.

- For each $\beta \in \mathcal{M}(\mathsf{C})((X, f), (\prod_n Y, \lhd_n^Y))$, $\beta^\sharp \in \mathsf{C}(X_{f^n}, Y)$ is the unique factorization

$$\begin{array}{ccccc} X & \underset{1_X}{\overset{f^n}{\rightrightarrows}} & X & \xrightarrow{q} & X_{f^n} \\ \downarrow{\scriptstyle \beta} & & \downarrow{\scriptstyle \beta} & & \vdots{\scriptstyle \beta^\sharp} \\ \prod_n Y & \xrightarrow{(\lhd_n^Y)^n = 1_{\prod_n Y}} & \prod_n Y & \xrightarrow{\pi_0} & Y \end{array}$$

- For each $\alpha \in \mathsf{C}(X_{f^n}, Y)$, $\alpha^\flat$ is the unique morphism commuting the diagram

$$\text{for all } k: \quad \begin{array}{ccccc} X & & \xrightarrow{\quad\alpha^\flat\quad} & & \prod_n Y \\ \downarrow{\scriptstyle f^k} & & & & \downarrow{\scriptstyle \pi_k^Y} \\ X & \xrightarrow{q} & X_{f^n} & \xrightarrow{\alpha} & Y \end{array}$$

  This is well-defined modulo n since q coequalizes f^n and 1_X. The proof of $\alpha^\flat \in \mathcal{M}(\mathsf{C})((X,f), (\prod_n Y, \lhd_n^Y))$ follows dual argument given in $L_n \dashv \wedge^n$.

The proof that $\flat$ and $\sharp$ are mutually inverse and that they are natural follows the dual argument given in $L_n \dashv \wedge^n$. □

6 Endomorphism Functor

In Section 3, we see that each (locally small) category C induces its (locally small) category $\mathcal{M}(\mathsf{C})$ of endomorphisms. In this section, we will prove that this construction is functorial in that it gives rise to the functor $\mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$, where CAT is the category of locally small categories. Moreover, the functor induces a monad on CAT and preserves the strict 2-categorical structure of CAT.

Construction 6.1. The **endomorphism functor** $\mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$ is defined as follows:

- on objects : a locally small category C is mapped to the category $\mathcal{M}(\mathsf{C})$ of endomorphisms on C, which can be shown to be locally small.
- on morphisms : a functor $F : \mathsf{C} \to \mathsf{D}$ is mapped to the functor $\mathcal{M}(F) : \mathcal{M}(\mathsf{C}) \to \mathcal{M}(\mathsf{D})$ given as follows:
  - on objects : $(X, f) \mapsto (F(X), F(f))$,
  - on morphisms : $((X,f) \xrightarrow{\alpha} (Y,g)) \mapsto ((F(X),F(f)) \xrightarrow{F\alpha} (F(Y),F(g)))$, which is well-defined by functoriality of F.

Functoriality of $\mathcal{M}$ includes

- preserving identity : For any C, $\mathcal{M}(1_\mathsf{C}) = 1_{\mathcal{M}(\mathsf{C})}$ as functors $\mathcal{M}(\mathsf{C}) \to \mathcal{M}(\mathsf{C})$;
- respecting composition : For any $\mathsf{C} \xrightarrow{F} \mathsf{D} \xrightarrow{G} \mathsf{E}$, $\mathcal{M}(G \circ F) = \mathcal{M}(G) \circ \mathcal{M}(F)$ as functors $\mathcal{M}(\mathsf{C}) \to \mathcal{M}(\mathsf{E})$.

These conditions are straightforward to verify.

6.1 $\mathcal{M}$ induces a monad on CAT

Construction 6.2. We construct a monad $(\mathcal{M}, \eta, \mu)$ on CAT as follows:

- the endofunctor $\mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$ is as in Construction 6.1,

- the unit $\eta : 1_{\mathsf{CAT}} \Rightarrow \mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$ has its components given by

$$\eta_{\mathsf{C}} : \mathsf{C} \longrightarrow \mathcal{M}(\mathsf{C}),$$
$$(X \xrightarrow{\alpha} Y) \longmapsto ((X, 1_X) \xrightarrow{\alpha} (Y, 1_Y)).$$

- the multiplication $\mu : \mathcal{M}\mathcal{M} \Rightarrow \mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$ has its components given by

$$\mu_{\mathsf{C}} : \mathcal{M}(\mathcal{M}(\mathsf{C})) \longrightarrow \mathcal{M}(\mathsf{C}),$$
$$(((X, f), \alpha) \xrightarrow{H} ((Y, g), \beta)) \longmapsto ((X, f \circ \alpha) \xrightarrow{H} (Y, g \circ \beta)).$$

Proposition 6.3. *$(\mathcal{M}, \eta, \mu)$ constructed in Construction 6.2 is indeed a monad on* CAT.

Proof. Functoriality of η_{C} and μ_{C} as well as naturality of η and μ are omitted.

Let us verify that the axioms of monads, i.e. commutativity of the following diagrams

$$\begin{array}{ccc} \mathcal{M}\mathcal{M}\mathcal{M} & \xRightarrow{\mathcal{M}\mu} & \mathcal{M}\mathcal{M} \\ \Downarrow{\scriptstyle \mu\mathcal{M}} & & \Downarrow{\scriptstyle \mu} \\ \mathcal{M}\mathcal{M} & \xRightarrow{\mu} & \mathcal{M} \end{array} \qquad \begin{array}{ccccc} \mathcal{M} & \xRightarrow{\mathcal{M}\eta} & \mathcal{M}\mathcal{M} & \xLeftarrow{\eta\mathcal{M}} & \mathcal{M} \\ & \searrow & \Downarrow{\scriptstyle \mu} & \swarrow & \\ & & \mathcal{M} & & \end{array}$$

commute.

Let C be a category. We need to show that the diagrams

$$\begin{array}{ccc} \mathcal{M}\mathcal{M}\mathcal{M}\mathsf{C} & \xrightarrow{\mathcal{M}\mu_{\mathsf{C}}} & \mathcal{M}\mathcal{M}\mathsf{C} \\ \downarrow{\scriptstyle \mu_{\mathcal{M}\mathsf{C}}} & & \downarrow{\scriptstyle \mu_{\mathsf{C}}} \\ \mathcal{M}\mathcal{M}\mathsf{C} & \xrightarrow{\mu_{\mathsf{C}}} & \mathcal{M}\mathsf{C} \end{array} \qquad \begin{array}{ccccc} \mathcal{M}\mathsf{C} & \xrightarrow{\mathcal{M}\eta_{\mathsf{C}}} & \mathcal{M}\mathcal{M}\mathsf{C} & \xleftarrow{\eta_{\mathcal{M}\mathsf{C}}} & \mathcal{M}\mathsf{C} \\ & \searrow & \downarrow{\scriptstyle \mu_{\mathsf{C}}} & \swarrow & \\ & & \mathcal{M}\mathsf{C} & & \end{array}$$

commute.

1. $\mu_{\mathsf{C}} \circ \mathcal{M}\mu_{\mathsf{C}} = \mu_{\mathsf{C}} \circ \mu_{\mathcal{M}\mathsf{C}}$ as functors $\mathcal{M}\mathcal{M}\mathcal{M}\mathsf{C} \to \mathcal{M}\mathsf{C}$:
   - on objects : For each $(((X, f), \alpha), H) \in \mathcal{M}\mathcal{M}\mathcal{M}\mathsf{C}$, we have
   $$(\mu_{\mathsf{C}} \circ \mathcal{M}\mu_{\mathsf{C}})((((X, f), \alpha), H)) = \mu_{\mathsf{C}}(((X, f \circ \alpha), H))$$
   $$= (X, f \circ \alpha \circ H) = \mu_{\mathsf{C}}(((X, f), \alpha \circ H)) = (\mu_{\mathsf{C}} \circ \mu_{\mathcal{M}\mathsf{C}})((((X, f), \alpha), H)).$$
   - on morphisms : For each $(((X, f), \alpha), H) \xrightarrow{\phi} (((Y, g), \beta), G)$, we have
   $$(\mu_{\mathsf{C}} \circ \mathcal{M}\mu_{\mathsf{C}})(\phi) = \mu_{\mathsf{C}}(\phi) = \phi = \mu_{\mathsf{C}}(\phi) = (\mu_{\mathsf{C}} \circ \mu_{\mathcal{M}\mathsf{C}})(\phi).$$
2. $\mu_{\mathsf{C}} \circ \mathcal{M}\eta_{\mathsf{C}} = 1_{\mathcal{M}\mathsf{C}}$ as functors $\mathcal{M}\mathsf{C} \to \mathcal{M}\mathsf{C}$:
   - on objects : For any $(X, f) \in \mathcal{M}\mathsf{C}$,
   $$(\mu_{\mathsf{C}} \circ \mathcal{M}\eta_{\mathsf{C}})((X, f)) = \mu_{\mathsf{C}}(((X, 1_X), f)) = (X, 1_X \circ f) = (X, f) = 1_{\mathcal{M}\mathsf{C}}((X, f)).$$
   - on morphisms : For any $(X, f) \xrightarrow{\alpha} (Y, g)$, we have
   $$(\mu_{\mathsf{C}} \circ \mathcal{M}\eta_{\mathsf{C}})(\alpha) = \mu_{\mathsf{C}}(\alpha) = \alpha = 1_{\mathcal{M}\mathsf{C}}(\alpha).$$

3. $\mu_{\mathsf{C}} \circ \eta_{\mathcal{M}\mathsf{C}} = 1_{\mathcal{M}\mathsf{C}}$ as functors $\mathcal{M}\mathsf{C} \to \mathcal{M}\mathsf{C}$:
   - on objects : For any $(X, f) \in \mathcal{M}\mathsf{C}$,
   $$(\mu_{\mathsf{C}} \circ \eta_{\mathcal{M}\mathsf{C}})((X, f)) = \mu_{\mathsf{C}}(((X, f), 1_X)) = (X, f \circ 1_X) = (X, f) = 1_{\mathcal{M}\mathsf{C}}((X, f)).$$
   - on morphisms : For any $(X, f) \xrightarrow{\alpha} (Y, g)$, we have
   $$(\mu_{\mathsf{C}} \circ \eta_{\mathcal{M}\mathsf{C}})(\alpha) = \mu_{\mathsf{C}}(\alpha) = \alpha = 1_{\mathcal{M}\mathsf{C}}(\alpha).$$

In total, $(\mathcal{M}, \eta, \mu)$ defines a monad on CAT. □

6.2 $\mathcal{M}$ is a strict 2-functor

Recall that the category CAT is endowed with a strict 2-category structure with categories as 0-cells, functors as 1-cells, and natural transformations as 2-cells. Denote by $\circ$ and $*$ vertical and horizontal 2-compositions, respectively.

Construction 6.4. On top of the functor $\mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$, assign to each natural transformation $\alpha : F \Rightarrow F' : \mathsf{C} \to \mathsf{D}$ a natural transformation $\mathcal{M}(\alpha) : \mathcal{M}(F) \Rightarrow \mathcal{M}(F') : \mathcal{M}(\mathsf{C}) \to \mathcal{M}(\mathsf{D})$ whose components are given by

$$\mathcal{M}(\alpha)_{(X,f)} := \alpha_X : (FX, Ff) \to (F'X, F'f),$$

which is a well-defined morphism by naturality of α at f. This assignment makes $\mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$ a strict 2-functor.

For the proof, it is left to verify that the assignment $\mathcal{M}$ on 2-cells respect 2-identities and 2-compositions.

- 2-identity : for any 1-cell $F : \mathsf{C} \to \mathsf{D}$, $\mathcal{M}(1_F) = 1_{\mathcal{M}(F)}$.
- horizontal 2-composition : whenever the 2-cell $\beta * \alpha$ is defined, $\mathcal{M}(\beta * \alpha) = \mathcal{M}(\beta) * \mathcal{M}(\alpha)$.
- vertical 2-composition : whenever the 2-cell $\beta \circ \alpha$ is defined, $\mathcal{M}(\beta \circ \alpha) = \mathcal{M}(\beta) \circ \mathcal{M}(\alpha)$.

These hold ultimately because each induced natural transformation $\mathcal{M}(\alpha)$ is essentially given by α.

Corollary 6.5. *$\mathcal{M}$ preserves adjunctions.*

Proof. Recall that any strict 2-functor $\mathsf{CAT} \to \mathsf{CAT}$ preserves adjunctions (cf. Proposition 4.4.6. [5]). In particular, $\mathcal{M} : \mathsf{CAT} \to \mathsf{CAT}$ preserves adjunctions. □

Example 6.6 (Cylinder-path adjunction)**.** Let $I = [0, 1]$ be the unit interval equipped with the standard topology. For each topological space X, one can construct the product space $I \times X$ and the space X^I of continuous maps $I \to X$ equipped with the compact-open topology. These constructions are functorial in that they define functors $I \times -, (-)^I : \mathsf{Top} \to \mathsf{Top}$. In fact, one can show that $I \times - \dashv (-)^I$.

Applying the functor $\mathcal{M}$, one obtain functors $C, P : \mathcal{M}(\mathsf{Top}) \to \mathcal{M}(\mathsf{Top})$ given on objects by

$$C(X, f) := (I \times X, 1_I \times f) \quad \text{and} \quad P(X, f) := (X^I, f \circ -).$$

Moreover, they form an adjunction $C \dashv P$.

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments.

References

[1] Layek, G. C. (2015). *An introduction to dynamical systems and chaos* (Vol. 1). West Bengal, India: Springer India.

[2] Chaoha, P., & Phuksawad, T. (2024). Homotopy Conjugate of Continuous Map. *Proceedings of Annual Pure and Applied Mathematics Conference 2024*, 90-100. https://math.sc.chula.ac.th/wordpress/2024/apam/apam-2024-2/

[3] Almkvist, G. (1974). The Grothendieck ring of the category of endomorphisms. *Journal of Algebra, 28*(3), 375-388.

[4] Song, K., Wu, L., & Zhang, Y. (2018). Endomorphism category of an abelian category. *Communications in Algebra, 46*(7), 3062-3070.

[5] Riehl, E. (2017). *Category theory in context*. Courier Dover Publications.

[6] Leinster, T. (2004). *Higher operads, higher categories* (No. 298). Cambridge University Press.

Splitness of the Characteristic Polynomial of 2×2 Normal Matrices over a Field of Characteristic 2

Sarawut Therdchanthuk[1,*] and Keng Wiboonton[1,†]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

The study of linear algebra over finite fields is different from the study of the one over the real field or the complex field. Some important results over either the real or the complex are no longer true when we work over finite fields. In [1], Basha and McDonald proved the spectral theorem for normal matrices over finite fields. However, the splitness of normal matrices over finite fields was not mentioned in [1]. Therefore, we investigate this issue in our work. In this paper, we provide results on the splitness of 2×2 normal matrices over a finite field of characteristic 2 and obtain the canonical forms of these matrices.

Keywords: finite field of characteristic 2, normal matrix, characteristic polynomial

2020 MSC: 15B33, 15A20, 11T06

1 Introduction

There are many theorems in linear algebra that agree in any field, but some of them do not, especially in finite fields. For instance, certain results on orthogonality of vectors and diagonalizability of matrices. Due to the characteristic of $\mathbb{F}_{q^2}$, the finite field with q^2 elements, there exists a self-orthogonal vector over $\mathbb{F}_{q^2}$, i.e., there is a vector that is perpendicular to itself. This leads to some challenges when we study Hermitian and normal matrices in this context. However, in [1], Basha and McDonald showed the following spectral theorem for normal matrices over $\mathbb{F}_{q^2}$.

Theorem 1.1 ([1], Theorem 7.7)**.** *If A is an $n \times n$ normal matrix over $\mathbb{F}_{q^2}$ and if A does not have any self-orthogonal eigenvectors, then A is unitarily diagonalizable.*

Nevertheless, in this theorem, there is no assumption about the splitness of the matrix A to guarantee the existence of the eigenvalues of A over $\mathbb{F}_{q^2}$. This makes us suspect that normal matrices over $\mathbb{F}_{q^2}$ may not be split over $\mathbb{F}_{q^2}$, that is, the characteristic polynomial of a normal matrix over $\mathbb{F}_{q^2}$ may not be factorized into a product of linear polynomials. So we try to find a normal matrix over $\mathbb{F}_{q^2}$ which is not split. For instance, working over $\mathbb{F}_4$, consider the matrix A as follows:

*Speaker (S.Therdchanthuk@hotmail.com)
†Corresponding author (kwiboonton@gmail.com)

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Notice that A is normal and the characteristic polynomial of A is

$$\chi_A(x) = (x-1)(x^3+x+1),$$

having only the eigenvalue 1 with the one-dimensional corresponding eigenspace spanned by the vector $(1,0,0,0)$. This implies that A is not diagonalizable over $\mathbb{F}_4$. Thus, from our perspective, to make the spectral theorem for normal matrices over a finite field completely valid, the splitness condition should be assumed. So the splitness of the characteristic polynomial of normal matrices over a given finite field is investigated in our study.

In this work, we focus on 2×2 normal matrices over a field of characteristic 2, and examine the factorization of the characteristic polynomial of these matrices. **In this work, F will always be a finite field of characteristic 2 and E a quadratic extension field of F, meaning that the degree of extension field E over F is 2.** The set of 2×2 matrices over a field F is denoted by $\mathrm{M}_2(F)$.

2 Preliminaries

In this section, we summarize appropriate definitions and results in linear algebra and field theory. These materials can be found in [2, 3].

Definition 2.1. Two matrices A and B are *similar*, denoted by $A \sim B$, if there exists an invertible matrix P for which $B = PAP^{-1}$. The equivalence classes associated with similarity are called *similarity classes*.

Theorem 2.2. *Let F be a field and let $p(x) \in F[x]$ be an irreducible polynomial. Then there exists a field K containing an isomorphic copy of F in which $p(x)$ has a root. Identifying F with this isomorphic copy shows that there exists an extension of F in which $p(x)$ has a root.*

Remark 2.3. The construction of a field K of Theorem 2.2 is as follows: Define the quotient $K := F[x]/(p(x))$ of the polynomial ring $F[x]$ by the ideal generated by $p(x)$. Then K is a field since $p(x)$ is an irreducible polynomial in the principal ideal domain $F[x]$ which implies that $(p(x))$ is maximal. Moreover, it is easy to see that K contains a root of the polynomial $p(x)$.

When we work on matrices over a field which is not algebraically closed, the following definition taken from [4] describes the diagonalization of a given matrix over some extension field.

Definition 2.4. A square matrix over a field F is called *potentially diagonalizable* if it is diagonalizable over some extension field of F.

According to [1, 5], we recall some definitions of quadratic extension fields of any field of characteristic 2.

Definition 2.5. The *F-conjugate* of an element $a \in E$, denoted by a^*, is the remaining Galois conjugate of a, which is the image of that element under the automorphisms of the field extension that fix the base field.

Definition 2.6. The *F-norm* of an element $a \in E$, denoted by $|a|$, is the product of all the Galois conjugates of a. It is clear by definition that $|a| = a^*a$ and $|a| \in F$.

Proposition 2.7. *Let F be a finite field of characteristic 2 and E a quadratic extension field of F. Then for any $a, b \in E$, the following statements are true.*

1. $(a+b)^* = a^* + b^*$.
2. $(ab)^* = a^*b^*$.
3. $(a^{-1})^* = (a^*)^{-1}$.
4. $a^* = a$ *if* $a \in F$.
5. $|ab| = |a||b|$.
6. $|a^{-1}| = |a|^{-1}$.
7. $|a| = a^2$ *if* $a \in F$.
8. $|ca| = c^2|a|$ *for all* $c \in F$.
9. $(a^*)^* = a$.

Remark 2.8. From the fact that $(a^*)^* = a$ for any $a \in E$ in the Proposition 2.7, we say that the map $(\,\cdot\,)^* : E \to E$ is an *involution*.

Definition 2.9. Let $A = [a_{ij}]$ be a $k \times n$ matrix with entries from E . Then the *conjugate transpose* of A, denoted by A^{H}, is the $n \times k$ matrix such that $[a_{ij}]^{\mathrm{H}} = [a^*_{ji}]$ for all i, j, where a^*_{ji} is the F-conjugate of a_{ij}. We will refer to A as *normal* provided $A^{\mathrm{H}}A = AA^{\mathrm{H}}$.

Proposition 2.10. *Let F be a finite field of characteristic 2 and E a quadratic extension field of F and $A, B \in \mathrm{M}_2(E)$. Then*

1. $(AB)^{\mathrm{H}} = B^{\mathrm{H}}A^{\mathrm{H}}$.
2. *if* $a \in E$, *then* $(aA)^{\mathrm{H}} = a^*A^{\mathrm{H}}$.

3 Main Results

3.1 Some auxiliary results

Before we start investigating the characteristic polynomial of normal matrices over a given quadratic extension of a finite field of characteristic 2, we study some facts about the roots of polynomials in a finite field of characteristic 2 and similarity properties of these matrices.

Proposition 3.1. *Let F be a field of characteristic 2 and $f(x) = x^2 + x + c \in F[x]$. Then α is a root of $f(x)$ if and only if $\alpha + 1$ is a root of $f(x)$.*

Proof. The relation

$$(\alpha+1)^2 + (\alpha+1) + c = \alpha^2 + 1 + \alpha + 1 + c = \alpha^2 + \alpha + c,$$

implies that α is a root of $f(x)$ if and only if $\alpha + 1$ is a root of $f(x)$. □

Remark 3.2. For the case $f(x) = x^2 + bx + c$ for some b, c in a field of characteristic 2 with $b \neq 0$, we can adjust the polynomial $f(x)$ so that

$$f(x) = b^2\left((b^{-1}x)^2 + (b^{-1}x) + b^{-2}c\right) = b^2 g(b^{-1}x)$$

where $g(x) = x^2 + x + b^{-2}c$. Now, if α is a root of $f(x)$, then $b^{-1}\alpha$ is a root of $g(x)$ and then by the Proposition 3.1 we have $b^{-1}\alpha + 1 = b^{-1}(\alpha + b)$ is a root of $g(x)$. Hence, $\alpha + b$ is a root of $f(x)$ distinct from α since $b \neq 0$. We can summarize the proposition as follows:

Proposition 3.3. *Let F be a field of characteristic 2 and $f(x) = x^2 + bx + c \in F[x]$ with $b \neq 0$. Then α is a root of $f(x)$ if and only if $\alpha + b$ is a root of $f(x)$.*

Proposition 3.4. *Let A be a square matrix over a field E. If A is a normal matrix, then so are $A - \alpha I$ and βA for all $\alpha \in E$ and $\beta \in E \setminus \{0\}$.*

Proof. For any $A \in \mathrm{M}_2(E)$ and $\alpha \in E$, we have

$$\begin{aligned}
(A - \alpha I)^{\mathrm{H}}(A - \alpha I) &= (A^{\mathrm{H}} - \alpha^* I)(A - \alpha I) \\
&= A^{\mathrm{H}}A - \alpha A^{\mathrm{H}} - \alpha^* A + \alpha^*\alpha I \\
&= AA^{\mathrm{H}} - \alpha A^{\mathrm{H}} - \alpha^* A + \alpha^*\alpha I \qquad \text{(since } A \text{ is normal)} \\
&= (A - \alpha I)(A^{\mathrm{H}} - \alpha^* I) \\
&= (A - \alpha I)(A - \alpha I)^{\mathrm{H}}
\end{aligned}$$

Therefore, $A - \alpha I$ is normal for all $\alpha \in E$.

Lastly, for any $\beta \in E \setminus \{0\}$, we have

$$\begin{aligned}
(\beta A)^{\mathrm{H}}(\beta A) &= \beta^*\beta A^{\mathrm{H}}A \\
&= \beta^*\beta AA^{\mathrm{H}} \qquad \text{(since } A \text{ is normal)} \\
&= (\beta A)(\beta A)^{\mathrm{H}}
\end{aligned}$$

Therefore, βA is normal for all $\beta \in E \setminus \{0\}$. □

Proposition 3.5. *Let A be a square matrix over a field E. If A is similar to B, then $\alpha A + \beta I$ is similar to $\alpha B + \beta I$ for all $\alpha, \beta \in E$.*

Proof. Suppose that $A \sim B$. Then $A = PBP^{-1}$ for some invertible matrix P. Then for any $\alpha, \beta \in E$, we have

$$\begin{aligned}
\alpha A + \beta I &= \alpha(PBP^{-1}) + \beta I \\
&= P(\alpha B)P^{-1} + P(\beta I)P^{-1} \\
&= P(\alpha B + \beta I)P^{-1}.
\end{aligned}$$

This follows that $\alpha A + \beta I \sim \alpha B + \beta I$. □

Corollary 3.6. *Let A be a square matrix over a field E. If A is diagonalizable over E, then so is $\alpha A + \beta I$ for all $\alpha, \beta \in E$.*

Proof. Suppose that A is diagonalizable over E. Then $A \sim D$ for some diagonal matrix D. Thus, for any $\alpha, \beta \in E$, we have $\alpha A + \beta I \sim \alpha D + \beta I$ by Proposition 3.5. Note that $\alpha D + \beta I$ is again a diagonal matrix since D is diagonal. It follows that $\alpha A + \beta I$ is diagonalizable over E. □

3.2 Normal matrices over the quadratic field extension of characteristic 2

Now, we are ready to investigate the characteristic polynomial of normal matrices over the quadratic extension of a field of characteristic 2.

Proposition 3.7. *Let F be a finite field of characteristic 2, E a quadratic extension field of F, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ a 2×2 matrix over E and $\chi_A(x)$ the characteristic polynomial of A (So $\chi_A(x) = x^2 - (a+d)x + (ad - bc)$).*

1. *If $a = d$, then there is an element α in some quadratic extension of E such that $\chi_A(x) = (x - \alpha)^2$ where α satisfies the equation $\alpha^2 = ad - bc$.*
2. *If $a \neq d$ and A is normal, then $\chi_A(x)$ splits as the product of distinct linear factors over E.*

Proof. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a matrix whose entries are in E. If $a = d$, then the characteristic polynomial of A becomes

$$\chi_A(x) = x^2 + (ad - bc).$$

The proof of the first part is obvious if $b = 0$ or $c = 0$ or $ad - bc$ is a square in E. Now assume b, c are nonzero elements in E and $ad - bc$ is not a square in E. Then $\chi_A(x)$ is irreducible in E and so it has a root in a quadratic extension $K := E[x]/(\chi_A(x))$ by Theorem 2.2, says $\alpha \in K$. Therefore $\alpha^2 = ad - bc$ and so $\chi_A(x) = (x - \alpha)^2$.

Now we are going to prove the second part. Suppose that A is a normal matrix such that $a \neq d$ (or equivalently, $a - d \neq 0$). Then by Proposition 3.4, $A - dI$ is normal and thus

$$B := (a-d)^{-1}(A - dI) = \begin{pmatrix} 1 & b(a-d)^{-1} \\ c(a-d)^{-1} & 0 \end{pmatrix} = \begin{pmatrix} 1 & bk \\ ck & 0 \end{pmatrix}$$

is also normal where $k = (a-d)^{-1}$. Note that $a - d \neq 0$ implies its inverse element exists, i.e. k exists and not equals to 0. From the normality of a matrix B, i.e. $B^{\mathrm{H}}B = BB^{\mathrm{H}}$, we have the following system of nonlinear equations:

$$1 + |bk| = 1 + |ck|, \tag{3.1}$$
$$ck = (bk)^*, \tag{3.2}$$
$$(ck)^* = bk, \tag{3.3}$$
$$|bk| = |ck|. \tag{3.4}$$

Obviously, the equations (3.1) and (3.4) are the same, and the equation (3.2) and (3.3) are identical since the F-conjugate is an involution. Moreover, equations (3.2) and (3.3) imply that (3.1) and (3.4) hold. Consider the characteristic polynomial of B. Simplifying $\chi_B(x)$ yields

$$\begin{aligned} \chi_B(x) &= \begin{vmatrix} x-1 & bk \\ ck & x \end{vmatrix} \\ &= x^2 - x - (bk)(ck) \\ &= x^2 - x - (bk)(bk)^* \qquad \text{by (3.2)} \\ &= x^2 - x - |bk|. \end{aligned}$$

Recall that $\chi_B(x)$ is a polynomial in F because the norm $|bk|$ belongs to F. Next, we examine whether the polynomial $\chi_B(x)$ has roots in the field E.

- If $\chi_B(x)$ is reducible in F, then $\chi_B(x)$ has a root in F. So we are done since $F \subseteq E$.
- If $\chi_B(x)$ is irreducible in F, then $\chi_B(x)$ has a root in a quadratic extension $K := F[x]/(\chi_B(x))$ by Theorem 2.2. Note that the degree of extension $[E : F]$ and $[K : F]$ are equal to 2. Therefore, $K \cong E$. Consequently, $\chi_B(x)$ has a root in E.

Both cases indicate that the characteristic polynomial $\chi_B(x)$ of B has a root in E, say β. Furthermore, by Proposition 3.1, $\chi_B(x)$ splits into distinct linear factors, namely, $\chi_B(x) = (x-\beta)(x-(\beta+1))$. Lastly, we will show that $\chi_A(x)$ also splits into distinct linear factors over E. By the definition of the matrix B, we find that

$$\begin{aligned}
\chi_B(x) &= \det(B - xI) \\
&= \det\big(k(A - dI) - xI\big) \\
&= \det\big(kA - (dk + x)I\big) \\
&= k^2 \det\big(A - k^{-1}(dk + x)I\big) \\
&= k^2 \chi_A\big(k^{-1}(dk + x)\big) \\
&= (a-d)^{-2} \chi_A\big((a-d)x + d\big).
\end{aligned}$$

Substituting x by β and $\beta+1$ shows that all the roots of $\chi_A(x)$ are $(a-d)\beta+d$ and $(a-d)(\beta+1)+d = (a-d)\beta+a$. The assumption $a \neq d$ asserts that the two roots are distinct. Therefore, $\chi_A(x)$ splits into distinct linear factors over E as desired. □

Remark 3.8. The first result in the above proposition can be shortened if the field E is perfect, i.e. if we can always pick an element $\alpha \in E$ such that $\alpha^2 = ad - bc$. This is the case if our fields are finite fields of characteristic 2.

Since Proposition 3.7 gives the splitness of the characteristic polynomial of normal matrices over the quadratic extension of a field of characteristic 2, this leads to diagonalizability of 2×2 normal matrices over the quadratic extension of a field of characteristic 2.

Theorem 3.9. *Let F be a finite field of characteristic 2, E a quadratic extension field of F and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ a 2×2 matrix over E.*

1. *If $a = d$ and at least one of the elements b or c are nonzero, then A is similar to a Jordan block of size 2.*
2. *If $a \neq d$, then A is potentially diagonalizable, in fact, A is diagonalizable over a quadratic extension of E, and in addition, if A is normal, then A is diagonalizable over E.*

Proof. If $a = d$, the Proposition 3.7 implies that there is an element α in some quadratic extension of E such that $\chi_A(x) = (x-\alpha)^2$ where $\alpha^2 = ad - bc$. Observe that the minimal polynomial of A, denoted by $m_A(x)$, is of the form $(x-\alpha)^2$ only. Otherwise, if $m_A(x) = x - \alpha$, then A is a diagonal matrix but we have a contradiction since $b \neq 0$ or $c \neq 0$ (by the assumption). Then A is not diagonalizable over a quadratic extension of E (and so not diagonalizable over

E). However, the algebraic multiplicity and geometric multiplicity of the eigenvalue α are 2 and 1 respectively, A is similar to a Jordan block of size 2. That is, A is similar to $\begin{pmatrix} \alpha & 0 \\ 1 & \alpha \end{pmatrix}$.

If $a \neq d$, then $\chi_A(x)$ splits as distinct linear factors over E by the Proposition 3.7 provided that A is normal. Now, consider just a general matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ over E with $a \neq d$. Recall that the characteristic polynomial of a matrix A is $\chi_A(x) = x^2 - (a+d)x + (ad - bc)$.

- If $\chi_A(x)$ has a root $\alpha \in E$, then A is diagonalizable over E since $\chi_A(x)$ splits as distinct linear factors by the Proposition 3.3.
- If $\chi_A(x)$ is irreducible, then $\chi_A(x)$ has a root in some quadratic extension of E. Hence $\chi_A(x)$ splits in $K := E[x]/(\chi_A(x))$ by the Proposition 3.3 (view $\chi_A(x)$ as a quadratic polynomial in K). So A is diagonalizable over K.

Both cases imply that A is diagonalizable over E or some quadratic extension of E, i.e. A is potentially diagonalizable. □

4 Discussion

We know from Theorem 3.9 that a normal matrix over a field of characteristic 2 is similar to either a diagonal matrix (or equivalently, diagonalizable) or a Jordan block of size 2. Now, we work out the details on these two canonical forms.

Let F be a finite field of characteristic 2 and E a quadratic extension field of F. Consider a normal matrix $A \in \mathrm{M}_2(E)$, say $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $a, b, c, d \in E$. Normality of a matrix A yields $AA^{\mathrm{H}} = A^{\mathrm{H}}A$ and so

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then we have the following system of nonlinear equations:

$$\begin{aligned} aa^* + bb^* &= a^*a + c^*c, \\ ac^* + bd^* &= a^*b + c^*d, \\ ca^* + db^* &= b^*a + d^*c, \\ cc^* + dd^* &= b^*b + d^*d. \end{aligned}$$

Since the map $(\,\cdot\,)^*$ is an involution, simplification of the system of the above nonlinear equations yields

$$|b| = |c|, \tag{4.1}$$
$$ac^* + bd^* = a^*b + c^*d. \tag{4.2}$$

First, note that $b = 0$ if and only if $c = 0$ by the equation (4.1). Therefore, if $b = 0$ or $c = 0$, then A is just a diagonal matrix. So, A is diagonalizable. Now assume that b and c are non-zero elements in E. We divide into two cases depending on the value of $a - d$.

Suppose that $a-d \neq 0$ (or equivalently, $a \neq d$). Then A is diagonalizable over E by Theorem 3.9. To explain the factorization of A, the characteristic polynomial of A splits as distinct linear factors over E according to Proposition 3.7. Moreover, following the proof of the proposition shows that $\lambda_1 := (a-d)\alpha + a$ and $\lambda_2 := (a-d)\alpha + d$ are eigenvalues of A where $k = (a-d)^{-1}$ and $\alpha \in E$ satisfies $\alpha^2 - \alpha - |bk| = 0$. Note that both eigenvalues belong to E. Next, we find an eigenvector corresponding to the eigenvalue λ_1, which is equivalent to finding a non-zero vector in the eigenspace $\ker(A - \lambda_1 I)$. To this end, we observe that

$$\begin{aligned}
(A - \lambda_1 I)\begin{pmatrix} bk \\ \alpha \end{pmatrix} &= \begin{pmatrix} -(a-d)\alpha & b \\ c & -(a-d)(\alpha+1) \end{pmatrix}\begin{pmatrix} bk \\ \alpha \end{pmatrix} \\
&= \begin{pmatrix} -(a-d)kb\alpha + b\alpha \\ cbk - (a-d)(\alpha+1)\alpha \end{pmatrix} \\
&= \begin{pmatrix} -b\alpha + b\alpha \\ cbk - (a-d)(bk)(bk)^* \end{pmatrix} \quad (\text{since } \alpha^2 - \alpha - |bk| = 0) \\
&= \begin{pmatrix} 0 \\ cc^*k^* - |b|k^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{by (4.1) and (4.2).}
\end{aligned}$$

Therefore, we conclude that $\left(\begin{smallmatrix} bk \\ \alpha \end{smallmatrix}\right)$ is an eigenvector corresponding to the eigenvalue $(a-d)\alpha + a$. A similar argument from above shows that an eigenvector corresponding to the eigenvalue $(a-d)\alpha + d$ is $\left(\begin{smallmatrix} bk \\ \alpha+1 \end{smallmatrix}\right)$. So we obtain an explicit form of the diagonalizability of a matrix A as the following:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} bk & bk \\ \alpha & \alpha+1 \end{pmatrix}\begin{pmatrix} (a-d)\alpha + a & 0 \\ 0 & (a-d)\alpha + d \end{pmatrix}\begin{pmatrix} bk & bk \\ \alpha & \alpha+1 \end{pmatrix}^{-1},$$

here $k = (a-d)^{-1}$ and $\alpha \in E$ satisfies $\alpha^2 - \alpha - |bk| = 0$. Notice that the matrix $\begin{pmatrix} bk & bk \\ \alpha & \alpha+1 \end{pmatrix}$ is invertible because it has determinant $bk \neq 0$, and the factorization is not unique, depending on the eigenvectors we choose.

Suppose that $a - d = 0$ (or equivalently, $a = d$). Then A is similar to a Jordan block of size 2 by Theorem 3.9. In addition, following Proposition 3.7, the characteristic polynomial of A is $\chi_A(x) = x^2 + (ad - bc) = (x-\beta)^2$ for some β in a quadratic extension K of E and $\beta^2 = ad - bc$. Note that the minimal polynomial in this case is exactly the characteristic polynomial. Next, we find an invertible matrix $P = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$ over K such that $A = PJP^{-1}$ (or equivalently, $AP = PJ$) where J is a Jordan block of size 2 whose diagonal entries are β, β. First, we define $B := A - dI = \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}$. Then the equation $AP = PJ$ is equivalent to $BP = P(J - dI)$. Some computation gives the following system of equations:

$$bx_2 = (\beta - d)x_1 + x_2, \tag{4.3}$$
$$cx_1 = (\beta - d)y_1 + y_2, \tag{4.4}$$
$$by_2 = (\beta - d)x_2, \tag{4.5}$$
$$cy_1 = (\beta - d)y_2. \tag{4.6}$$

Multiplying both sides of (4.3) by $\beta - d$ yields

$$\begin{aligned} b(\beta-d)x_2 &= (\beta-d)^2x_1 + (\beta-d)x_2 \\ b(\beta-d)x_2 &= bcx_1 + by_2 \qquad \text{by (4.5)} \\ b(\beta-d)x_2 &= b(\beta-d)y_1 \qquad \text{by (4.4)}. \end{aligned}$$

Note that $(\beta-d)^2 = \beta^2 - d^2 = (d^2-bc) - d^2 = -bc = bc$. To comply with the above equation, we choose $x_2 = y_1$. Next, multiplying both sides of (4.4) by $\beta - d$ yields

$$\begin{aligned} c(\beta-d)x_1 &= (\beta-d)^2y_1 + (\beta-d)y_2 \\ c(\beta-d)x_1 &= bcy_1 + cy_1 \qquad \text{by (4.6)} \\ c(\beta-d)x_1 &= (b+1)cy_1. \end{aligned}$$

Choosing $x_1 = b+1$ yields $x_2 = y_1 = \beta - d$ and thus we obtain $y_2 = c$. Therefore, from $B = P(J-dI)P^{-1}$, we have the following equation:

$$\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} = \begin{pmatrix} b+1 & \beta-d \\ \beta-d & c \end{pmatrix} \begin{pmatrix} \beta-d & 0 \\ 1 & \beta-d \end{pmatrix} \begin{pmatrix} b+1 & \beta-d \\ \beta-d & c \end{pmatrix}^{-1}.$$

Hence, by substituting $B = A - dI$, we obtain an explicit form of the Jordan canonical form of a matrix A as the following:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} b+1 & \beta-d \\ \beta-d & c \end{pmatrix} \begin{pmatrix} \beta & 0 \\ 1 & \beta \end{pmatrix} \begin{pmatrix} b+1 & \beta-d \\ \beta-d & c \end{pmatrix}^{-1}.$$

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments.

References

[1] A.I. Basha, J.J. McDonald, Orthogonality over finite fields, Linear and Multilinear Algebra 70 (2022) 7277–7289.

[2] D.S. Dummit, R.M. Foote, Abstract Algebra, Wiley, 3rd edition, 2003.

[3] S. Roman, Advanced Linear Algebra, Springer, New York, 2005.

[4] K. Conrad, Potential diagonalizability <https://kconrad.math.uconn.edu/blurbs/linmultialg/potdiagonalizable.pdf>, (accessed 07.03.25).

[5] A.I. Basha, Linear Algebra over Finite Fields, (Doctoral dissertation), Washington State University, 2020 https://rex.libraries.wsu.edu/esploro/outputs/doctoral/Linear-Algebra-over-Finite-Fields/99900581700601842/filesAndLinks?index=0.

ระบบการแนะนำการซื้อขายตราสารอนุพันธ์โดยใช้รูปแบบราคาทะลุผ่านกราฟสามเหลี่ยมและตัวแบบไลต์เกรเดียนต์บูสติงแมชชีน

ณัฐนนท์ เลี้ยงประไพพันธิ์[1], พิชายุตม์ ตันติพัฒน์พงศ์[1] และ โชติรส สุรพลชัย[1,*]

[1]ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ถนนพญาไท ปทุมวัน กรุงเทพฯ 10330

บทคัดย่อ

โครงงานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบการแนะนำการซื้อขายตราสารอนุพันธ์ที่อ้างอิงราคาดัชนี และแสดงผลลัพธ์โดยการนำปัญญาประดิษฐ์มาช่วยวิเคราะห์และตัดสินใจแทนนักลงทุน ซึ่งจะช่วยให้นักลงทุนประหยัดเวลาและช่วยลดความเสี่ยงจากการเลือกลงทุนด้วยตนเอง วิธีการวิจัยจะเริ่มจากการตรวจหารูปแบบราคาทะลุผ่านกราฟสามเหลี่ยม เพื่อใช้เป็นจุดเข้าซื้อ หลังจากนั้นจะคัดแยกคุณสมบัติของข้อมูล แล้วจึงกรองด้วยตัวแบบเชิงต้นไม้ไลต์เกรเดียนต์บูสติงแมชชีน เพื่อให้ได้จุดเข้าซื้อที่มีความปลอดภัยมากขึ้น โดยผลลัพธ์ของแบบจำลองจะสรุปโดยใช้กำไร อัตราการชนะ และผลขาดทุนสูงสุด หลังจากนั้นจะนำผลลัพธ์จากการทดลองไปจัดนำเสนอผ่านระบบ ผลวิจัยที่ได้แสดงให้เห็นว่า โมเดลสามารถช่วยลดการลงทุนที่จะทำให้ขาดทุนได้ โดยพิจารณาจากอัตราชนะที่คำนวณจากจำนวนครั้งของการซื้อขายที่กำไร และเมื่อเปรียบเทียบอัตราชนะหลังใช้โมเดลซึ่งเท่ากับ 59.09 เปอร์เซ็นต์ มีค่ามากกว่าอัตราการชนะก่อนใช้โมเดลซึ่งเท่ากับ 52.43 เปอร์เซ็นต์

คำสำคัญ: กราฟสามเหลี่ยม, ตราสารอนุพันธ์, ปัญญาประดิษฐ์, ไลต์เกรเดียนต์บูสติงแมชชีน
2020 MSC: 68T05, 68T20

1 บทนำ

การซื้อขายตราสารอนุพันธ์ในปัจจุบันมีการเปลี่ยนแปลงอย่างรวดเร็วและเต็มไปด้วยความซับซ้อน นักลงทุนจึงต้องวิเคราะห์ข้อมูลทางการเงินและข้อมูลในตลาดหุ้นอย่างละเอียดเพื่อการตัดสินใจที่ถูกต้องในการซื้อขายตราสารอนุพันธ์ เว็บไซต์การแนะนำการซื้อขายตราสารอนุพันธ์จึงมีบทบาทเพื่อรองรับความต้องการในการซื้อขายที่มีข้อมูลปริมาณมาก ๆ และช่วยลดความผิดพลาดของนักลงทุน โดยเว็บไซต์เหล่านั้นจะใช้ขั้นตอนวิธี (algorithm) ทำการ

*ผู้นำเสนอและผู้ประพันธ์บรรณกิจ (Chotiros.S@chula.ac.th)

วิเคราะห์และประมวลผลข้อมูล เพื่อช่วยตัดสินใจในการซื้อขายตราสารอนุพันธ์ให้แม่นยำและมีประสิทธิภาพที่สุด ซึ่งทำให้นักลงทุนมีโอกาสเลือกตราสารอนุพันธ์ในตลาดหุ้นมาลงทุนด้วยตนเองได้

นอกจากนี้การซื้อขายตราสารอนุพันธ์ยังมีจุดด้อยที่ต้องพิจารณา นั่นก็คือความซับซ้อนของระบบการแนะนำการซื้อขายตราสารอนุพันธ์ ซึ่งต้องใช้ทรัพยากรบุคคลและทรัพยากรทางเทคนิคจำนวนมาก เพื่อให้ระบบดังกล่าวสามารถทำงานได้อย่างถูกต้องและมีประสิทธิภาพที่สุด ดังนั้นการจัดการข้อมูลที่เปลี่ยนแปลงอย่างรวดเร็วในปัจจุบันต้องคำนึงถึงความเสี่ยงในการซื้อขายที่อาจเกิดขึ้นด้วย เช่น ระบบอาจเกิดข้อผิดพลาดหรือการตัดสินใจที่ไม่ถูกต้อง เป็นต้น และเพื่อรองรับและแก้ไขข้อจำกัดดังกล่าว การพัฒนาระบบการแนะนำการซื้อขายตราสารอนุพันธ์ยังต้องคำนึงถึงเทคโนโลยีขั้นสูง ไม่ว่าจะเป็นขั้นตอนวิธีการเรียนรู้ของเครื่อง (machine learning) เพื่อให้ปัญญาประดิษฐ์นำไปปรับขั้นตอนวิธีในการวิเคราะห์การตัดสินใจ การเรียกใช้ชุดข้อมูล รวมทั้งสถานการณ์ทางเศรษฐกิจในปัจจุบัน ปัจจัยทั้งหลายที่นำมาใช้ในการตัดสินใจลงทุนจึงเป็นสิ่งสำคัญอย่างยิ่ง ขั้นตอนวิธีเหล่านี้จะดึงข้อมูลในอดีตมาเรียนรู้และทดสอบ เพื่อรองรับกับสถานการณ์ที่ไม่คาดคิด [1,2]

ในโครงงานวิจัยนี้ผู้วิจัยจะพัฒนาขั้นตอนวิธีโดยใช้ข้อมูลราคา SET50 ที่ดาวน์โหลดจากโบรกเกอร์ Toptrader โดยเก็บข้อมูลตั้งแต่วันที่ 26 ตุลาคม พ.ศ. 2561 ถึงวันที่ 26 ตุลาคม พ.ศ.2566 โดยข้อมูลที่เก็บคือ วันที่ ราคาเปิดปิด ราคาสูงสุดต่ำสุด (8,970 แท่งกราฟ) แล้วตรวจหารูปแบบราคาทะลุผ่านกราฟสามเหลี่ยม [3,4] เพื่อใช้เป็นจุดเข้าซื้อ จากนั้นจะคัดแยกคุณสมบัติ (feature) แล้วกรองด้วยโมเดลเชิงต้นไม้ด้วยโมเดลไลต์เกรเดียนต์บูสติงแมช ชีน (LightGBM) [5-6] เพื่อสร้างโมเดลการทำนายราคาหุ้น หลังจากนั้นจะคาดคะเนกำไรหรือขาดทุนที่จะได้รับจากการลงทุนนั้น ๆ

2 ความรู้พื้นฐาน

2.1 ตราสารอนุพันธ์

ตราสารอนุพันธ์ (Derivative Warrant, DW) [8] คือสินทรัพย์ทางการเงินชนิดหนึ่งที่ตกลงซื้อขายกันโดยอ้างอิงจากราคาสินทรัพย์อื่น ที่เรียกว่า สินทรัพย์อ้างอิง และส่งมอบคืนในอนาคตตามราคาในสัญญาที่ตกลงกันไว้ระหว่างผู้ซื้อและผู้ขาย ทำให้มูลค่าของตราสารอนุพันธ์ขึ้นอยู่กับราคาของสินทรัพย์อ้างอิงในขณะนั้น นักลงทุนจะใช้ DW เพิ่มโอกาสให้ขาดทุนในการลงทุนน้อยลง หรืออาจเก็งกำไรในระยะสั้นที่จะทำกำไรได้บ้างเมื่อราคาหุ้นลดลงก็ตาม

ลักษณะเด่นของตราสารอนุพันธ์จะเป็นคู่สัญญาที่ทำสัญญาตกลงกันว่าจะซื้อหรือส่งมอบสิ่งใดสิ่งหนึ่งในอนาคตให้อีกฝั่งหนึ่ง ซึ่งเมื่อครบอายุของสัญญา มูลค่าของอนุพันธ์ก็จะหมดลงด้วย

2.2 รูปแบบราคาทะลุผ่านกราฟสามเหลี่ยม

รูปแบบราคาทะลุผ่านกราฟสามเหลี่ยม (Triangle Breakout Price Pattern) คือการวาดกราฟโดยที่เส้นแนวโน้มด้านบนเป็นราคาสูงสุด และเส้นแนวโน้มด้านล่างเป็นราคาต่ำสุดตัดกันเป็นรูปสามเหลี่ยม และสร้างเส้นแนวโน้มโดยการใช้แนวทางการถดถอยเชิงเส้น [4]

2.2.1 สามเหลี่ยมขาขึ้น

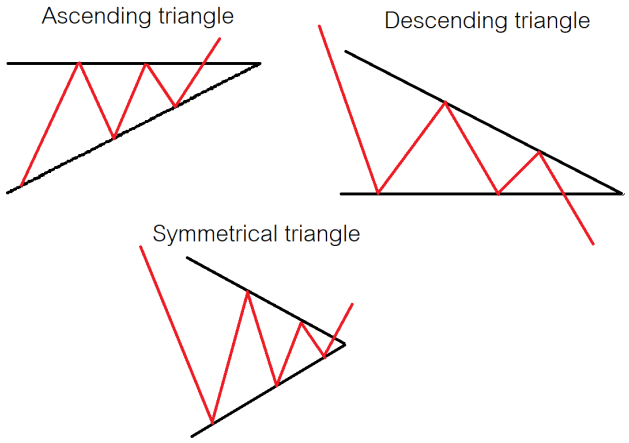
สามเหลี่ยมขาขึ้น (Ascending Triangle) จะเกิดขึ้นเมื่อราคาทะลุเส้นแนวโน้มแนวนอนด้านบนพร้อมกับปริมาณที่เพิ่มขึ้น จะได้รูปแบบที่มีแนวโน้มที่เชื่อว่าราคาหุ้นจะเพิ่มขึ้นในอนาคต โดยเส้นแนวโน้มด้านบนจะต้องเป็นแนวนอน ซึ่งแสดงถึงจุดสูงสุดที่เกือบจะเท่ากัน และก่อให้เกิดระดับแนวต้าน เส้นแนวโน้มด้านล่างเพิ่มขึ้นในแนวทแยงบ่งชี้ถึงจุดต่ำสุดที่สูงขึ้น

2.2.2 สามเหลี่ยมขาลง

สามเหลี่ยมขาลง (Descending Triangle) เป็นรูปแบบกลับหัวของสามเหลี่ยมขาขึ้น เส้นแนวโน้มด้านล่างควรเป็นแนวนอนโดยเชื่อมต่อจุดต่ำสุดที่เกือบจะเท่ากัน เส้นแนวโน้มด้านบนลดลงในแนวทแยงไปยังปลายทาง

2.2.3 สามเหลี่ยมสมมาตร

สามเหลี่ยมสมมาตร (Symmetrical Triangle) ประกอบด้วยเส้นแนวโน้มด้านบนที่ลดลงในแนวทแยง และเส้นแนวโน้มด้านล่างที่เพิ่มขึ้นในแนวทแยง เมื่อราคาเคลื่อนตัวไปยังปลายทาง มันจะทะลุเส้นแนวโน้มด้านบนอย่างหลีกเลี่ยงไม่ได้



รูปที่ 1 : รูปแบบราคาทะลุผ่านกราฟสามเหลี่ยม

2.3 โมเดลไลต์เกรเดียนต์บูสติงแมชชีน

โมเดลไลต์เกรเดียนต์บูสติงแมชชีน (Light Gradient-Boosting Machine Model, LightGBM) [7] จะเน้นใช้เทคนิคขั้นตอนวิธีการในการสร้างโครงสร้างต้นไม้ให้แตกกิ่งก้านของต้นไม้ไปตามใบ (leaf-wise) ซึ่งแตกต่างจากขั้นตอนวิธีอื่น ๆ ที่ต้นไม้จะเติบโตลงไปทีละระดับ (Level-wise) โดยจะทำการเลือกใบของต้นไม้ที่มีการสูญเสียข้อมูลมากสุดในการเจริญเติบโต โดยเลือกเทคนิคของ LightGBM มาใช้เป็นเฟรมเวิร์ค (framework) ดังนี้

2.3.1 เทคนิคการสุ่มด้านเดียวโดยใช้พื้นฐานจากเกรเดียนต์

การสุ่มด้านเดียวเป็นเทคนิคของ LightGBM โดยใช้พื้นฐานจากเกรเดียนต์ (Gradient-based One Side Sampling, GOSS) มาสร้างต้นไม้ตัดสินใจ เพื่อเพิ่มประสิทธิภาพของโมเดลและลดการใช้หน่วยความจำลง โดยเก็บตัวอย่างที่มีค่าเกรเดียนต์มาก และสุ่มกำจัดตัวอย่างที่มีค่าเกรเดียนต์น้อย

วิธี GOSS [5] ใช้การบูสบนชุดฝึก (train set) ที่มีตัวอย่าง n ตัว $\{ x_1, \dots , x_n \}$ โดยที่แต่ละตัวอย่าง x_i เป็นเวกเตอร์ที่มีมิติ s ในปริภูมิ X_s โดยในการวนลูปของการบูสแต่ละครั้งค่าเกรเดียนต์เชิงลบของฟังก์ชันสูญเสีย (Loss function) ที่เกี่ยวข้องกับผลลัพธ์ของโมเดลจะแสดงเป็น $\{ g_1, \dots , g_n \}$ ตัวอย่างใน train set จะได้รับการจัดอันดับจากมากไปน้อยตามค่าเกรเดียนต์สัมบูรณ์ และตัวอย่าง Top-A × 100% ที่มีค่าเกรเดียนต์มากที่สุดจะถูกเลือกเพื่อสร้างชุดย่อย A

สำหรับเซต A_c ที่เหลือ ประกอบด้วยตัวอย่าง (1- |A|) × 100% ที่มีค่าเกรเดียนต์น้อยกว่า จะนำไปตัดแบ่งตามสูตรความแปรปรวนโดยประมาณและเพิ่มเข้าไปในเวกเตอร์ตามลำดับต่อไป ทำให้ขนาดของตัวอย่างของชุดข้อมูล A เล็กลง

2.3.2 เทคนิคการมัดรวมคุณสมบัติพิเศษ

การประมวลผลชุดข้อมูลขนาดใหญ่อาจจะมีการตัดบางคุณสมบัติออกไป ทำให้เกิดการสูญเสียข้อมูลบางตัวได้ ดังนั้นเทคนิคการมัดรวมคุณสมบัติพิเศษ (Exclusive Feature Bundling, EFB) จะช่วยลดการสูญเสียข้อมูลถึงแม้จะลดจำนวนคุณสมบัติลงก็ตาม

2.4 คุณลักษณะเฉพาะที่นำมาใช้

คุณลักษณะเฉพาะหรือฟีเชอร์ (feature) สำหรับโมเดล จะเลือกใช้ฟีเชอร์ 4 ตัว คือ Trendline, Spearman_r, Dif_from_Trend และ Position_between_Minmax ดังนี้

2.4.1 Trendline

ฟีเชอร์ Trendline จะไว้คำนวณเส้นแนวโน้มของราคาในช่วงที่กำหนดว่าอยู่ในรูปแบบใด โดยใช้ค่าความชันของสมการถดถอยเชิงเส้น จากนั้นจะสร้างเส้นลากผ่านจุดยอดทั้งหมดที่อยู่ในช่วงแท่งเทียนก่อนหน้าในทุกคู่จุดยอดที่เป็นคู่จุดสูงสุดหรือคู่จุดต่ำสุดเพื่อให้เกิดกราฟสามเหลี่ยม

2.4.2 Spearman_r

ฟีเชอร์ Spearman_r จะนำมาคำนวณค่า spearman correlation ระหว่างข้อมูลราคาและตำแหน่งในช่วงเวลาที่กำหนดว่ามีความสัมพันธ์กันในลักษณะใด ค่าที่ได้จะมีค่าตั้งแต่ -1 ถึง 1 ถ้ามีค่าเป็น 1 จะเกาะเส้นกราฟราคาไป แต่ถ้ามีค่าเป็น -1 ค่าจะเกาะเส้นกราฟไปในทิศทางตรงกันข้าม

2.4.3 Dif_from_Trend

ฟีเชอร์ Dif_from_Trend จะนำมาคำนวณหาความต่างระหว่างราคาปัจจุบันเปรียบเทียบกับเส้นแนวโน้มในช่วงที่กำหนดว่ามีทิศทางสูงกว่าหรือต่ำกว่าค่าเฉลี่ย

2.4.4 Position_between_Minmax

ฟีเชอร์ Position_between_Minmax จะนำมาคำนวณว่าราคาปัจจุบันควรอยู่ในตำแหน่งใดของช่วงราคาซึ่งอยู่ระหว่างราคาต่ำสุดและราคาสูงสุดในช่วงที่กำหนด

โดยฟีเชอร์ทั้ง 4 ตัวจะถูกพิจารณาจากแท่งเทียนก่อนหน้าในชุดจำนวนแท่งเทียนที่กำหนด 6 ชุด คือ15 30 60 90 180 และ 360 แท่งก่อนหน้าตามลำดับ ดังนั้นฟีเชอร์ทั้งหมดที่เลือกใช้จะมี 24 ตัว

3 ผลการทดลอง

ข้อมูลที่นำมาใช้ในโมเดลจะเก็บรวบรวมตั้งแต่วันที่ 26 ตุลาคม พ.ศ. 2561 ถึงวันที่ 26 ตุลาคม พ.ศ. 2566 ซึ่งมีจำนวนข้อมูลทั้งหมด 8,970 แท่ง โดยพิจารณาจากข้อมูลที่เป็น test set ว่าควรจะมีทั้งช่วงที่มีแนวโน้มขาขึ้นและแนวโน้มขาลง ดังนั้นจึงแบ่งข้อมูลเป็น train set ร้อยละ 70 สำหรับการฝึกโมเดล ตั้งแต่วันที่ 26 ตุลาคม พ.ศ. 2561 เวลา 09.00 น. ถึงวันที่ 9 มิถุนายน พ.ศ. 2565 เวลา 09.00 น. ซึ่งมีจำนวนข้อมูล 6,279 แท่ง และที่เหลือเป็น test set ร้อยละ 30 ไว้สำหรับทดสอบประสิทธิภาพของโมเดล ตั้งแต่วันที่ 9 มิถุนายน พ.ศ. 2565 เวลา 10.00 น. ถึงวันที่ 26 ตุลาคม พ.ศ. 2566 เวลา 17.00 น. ซึ่งมีจำนวนข้อมูล 2,691 แท่ง

ในการเริ่มฝึกโมเดลจะมีการกำหนดสัญญาณการซื้อขายซึ่งเป็นการค่าเริ่มต้นของโมเดล ตัวอย่างเช่น เวลาสัญญาณซื้อ 2022-06-24 16:00 เวลาสัญญาณขาย 2022-06-28 09:00 ราคาเข้า 950.36 จุดตัดการขาดทุน 941.29 ราคาขาย 959.17 และกำไรที่คาดการณ์ 37.08 เป็นต้น

จากข้อมูล test set จะได้ผลการทดลองดังตารางที่ 1 ถึงตารางที่ 3 ซึ่งแสดงถึงผลการดำเนินการวิจัยในระบบการแนะนำการซื้อขายตราสารอนุพันธ์โดยใช้รูปแบบราคาทะลุผ่านกราฟสามเหลี่ยมและโมเดลไลต์เกรเดียนต์บูสติงแมชชีน [9] โดยแบ่งรูปแบบการซื้อเป็น 3 แบบคือ

1. ซื้อแบบขึ้น หมายถึง การซื้อขายในช่วงตลาดขาขึ้น
2. ซื้อแบบลง หมายถึง การซื้อขายในช่วงตลาดขาลง
3. ซื้อแบบปกติ หมายถึง การซื้อขายในตลาดทั้งช่วงขาขึ้นและช่วงขาลงรวมกัน

ตารางที่ 1 : ตารางจำนวนครั้งการซื้อขายก่อนและหลังใช้โมเดล

รูปแบบการซื้อ	**จำนวนครั้งการซื้อขายที่กำไรก่อนใช้โมเดล**	**จำนวนครั้งการซื้อขายที่ขาดทุนก่อนใช้โมเดล**	**จำนวนครั้งการซื้อขายที่กำไรหลังใช้โมเดล**	**จำนวนครั้งการซื้อขายที่ขาดทุนหลังใช้โมเดล**
ซื้อแบบขึ้น	27	24	11	12
ซื้อแบบลง	27	25	15	6
ซื้อแบบปกติ	54	49	26	18

จากตารางที่ 1 จะเห็นว่า ก่อนใช้โมเดลปัญญาประดิษฐ์มีการซื้อแบบปกติทั้งหมด 103 ครั้ง แต่หลังจากใช้โมเดลช่วยทำนายจะมีการซื้ออยู่ที่ 44 ครั้ง แสดงว่ามีจำนวนครั้งในการซื้อขายลดลงทั้งการซื้อขายที่กำไรและการซื้อขายที่ขาดทุน นั่นคือจากเดิมการซื้อขายที่กำไรทั้งหมด 54 ครั้ง ลดลงเหลือ 26 ครั้ง และจากเดิมการซื้อขายที่ขาดทุนทั้งหมด 49 ครั้ง ลดลงเหลือ 18 ครั้ง

จำนวนครั้งการซื้อขายในตารางที่ 1 จะนำมาคำนวณหาอัตราชนะ โดยคำนวณจากจำนวนครั้งการซื้อขายที่กำไร หารด้วยจำนวนครั้งที่ซื้อทั้งหมด (จำนวนครั้งการซื้อขายที่กำไร + จำนวนครั้งการซื้อขายที่ขาดทุน) ที่สอดคล้องกับรูปแบบการซื้อ (ซื้อแบบขึ้น ซื้อแบบลง และซื้อแบบปกติ) ดังในตารางที่ 2

ตารางที่ 2 : ตารางอัตราชนะของการซื้อขายก่อนและหลังใช้โมเดล

รูปแบบการซื้อ	**อัตราชนะก่อนใช้โมเดล (%)**	**อัตราชนะหลังใช้โมเดล (%)**
ซื้อแบบขึ้น	52.94	47.83
ซื้อแบบลง	51.92	71.43
ซื้อแบบปกติ	52.43	59.09

จากตารางที่ 2 จะเห็นว่าหลังใช้โมเดลช่วยทำนายในการเลือกซื้อแบบปกติมีอัตราชนะมากกว่าก่อนใช้โมเดลช่วยทำนาย จากก่อนใช้โมเดลมีอัตราชนะอยู่ที่ 52.43 เปอร์เซ็นต์ และหลังใช้โมเดลมีอัตราชนะเพิ่มขึ้นเป็น 59.09

เปอร์เซ็นต์ เมื่อแบ่งพิจารณาการซื้อแบบขึ้นและการซื้อแบบลง จะเห็นได้ว่าในการซื้อแบบขึ้นโมเดลไม่ได้ช่วยเพิ่มอัตราชนะ แต่ในการซื้อแบบลงโมเดลช่วยเพิ่มอัตราชนะอย่างมาก จากก่อนใช้โมเดลมีอัตราชนะอยู่ที่ 51.92 เปอร์เซ็นต์ และหลังใช้โมเดลมีอัตราชนะเพิ่มขึ้นเป็น 71.43 เปอร์เซ็นต์ จึงสรุปได้ว่าโมเดลไลต์เกรเดียนต์บูสติงแมชชีนสามารถช่วยทำนายอัตราชนะจาก 52.43 เปอร์เซ็นต์เป็น 59.09 เปอร์เซ็นต์ นับเป็น 6.66 เปอร์เซ็นต์ ซึ่งในการซื้อแบบลงทำได้ดีกว่าในการซื้อแบบขึ้น ทั้งนี้สาเหตุหนึ่งเนื่องมาจากข้อมูล test set ที่ใช้อยู่ในปี พ.ศ. 2566 ซึ่งสถานการณ์ของราคาตลาดอยู่ในช่วงขาลง

ตารางที่ 3 : ตารางผลขาดทุนสูงสุด

รูปแบบการซื้อ	**ผลขาดทุนสูงสุดก่อนใช้โมเดล (%)**	**ผลขาดทุนสูงสุดหลังใช้โมเดล (%)**
ซื้อแบบขึ้น	-72.5874	-9.5733
ซื้อแบบลง	-60.9739	-16.2367
ซื้อแบบปกติ	-72.5874	-16.8773

ในส่วนของกำไรจากการซื้อขายตราสารอนุพันธ์ จากการคำนวณปรากฏว่า ก่อนใช้โมเดลช่วยทำนายสามารถทำกำไรสุทธิได้ 100.3769 เปอร์เซ็นต์ และหลังใช้โมเดลช่วยทำนายสามารถทำกำไรสุทธิได้ 89.5352 เปอร์เซ็นต์ นั่นคือการใช้โมเดลไลต์เกรเดียนต์บูสติงแมชชีนนี้มาช่วยทำนายนั้น ไม่ได้ช่วยเพิ่มผลกำไร แต่จากตารางที่ 3 พิจารณาถึงผลขาดทุนสูงสุดนั้นปรากฏว่า ผลขาดทุนสูงสุดหลังใช้โมเดลนั้น น้อยกว่าผลขาดทุนสูงสุดก่อนใช้โมเดล จึงสามารถสรุปได้ว่า โมเดลไลต์เกรเดียนต์บูสติงแมชชีนนี้สามารถช่วยลดการขาดทุนให้ลดลงได้

4 ผลสรุป

จากผลวิจัยการใช้รูปแบบราคาทะลุผ่านกราฟสามเหลี่ยมและโมเดลไลต์เกรเดียนต์บูสติงแมชชีนจะได้ว่าโมเดลไลต์เกรเดียนต์บูสติงแมชชีนมีความสามารถในการทำนายการซื้อขายในตลาดขาลงด้วยความคลาด เคลื่อนน้อยกว่าการทำนายการซื้อขายในตลาดขาขึ้น และในการคำนวณหาอัตราชนะจะได้ว่า อัตราชนะในการซื้อแบบลงนั้นทำได้ดีกว่าการซื้อแบบขึ้น ส่วนในเรื่องของผลกำไรนั้น โมเดลไลต์เกรเดียนต์บูสติงแมชชีนทำนายได้ไม่ดีเท่าที่ควร ดังนั้นในอนาคตอาจจะทดลองใช้ปัญญาประดิษฐ์แบบการเรียนรู้ของเครื่องชนิดต่าง ๆ มาทดลองเพื่อหากำไรที่เพิ่มขึ้น หรือเพื่อกรองจุดเข้าซื้อ เป็นต้น

กิตติกรรมประกาศ ขอขอบคุณภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และกองทุนเพื่อการพัฒนาบุคลากรคณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

เอกสารอ้างอิง

[1] M. Alsulmi, "Reducing Manual Effort to Label Stock Market Data by Applying a Metaheuristic Search: A Case Study from the Saudi Stock Market", IEEE Access 9 (2021) 110493-110504.

[2] S. N. T. Nishitha, S. Bano, G. G. Reddy, P. Arja and G. L. Niharika, "Stock Price Prognosticator using Machine Learning Techniques", 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India (2020), pp. 1-7.

[3] K. Kamijo and T. Tanigawa, "Stock price pattern recognition-a recurrent neural network approach", 1990 IJCNN International Joint Conference on Neural Networks, San Diego, CA, USA (1990), pp. 215-221.

[4] A. Tripathi, J. Mathure, S. Deotarse, D. Rai and L. Gadhikar, "Linear Regression Approach For Stock Chart Pattern Recognition", 2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India (2023), pp. 1-6.

[5] Gupta, V., Kumar, E., “H3O-LGBM: hybrid Harris hawk optimization based light gradient boosting machine model for real-time trading”, Artif Intell Rev 56 (2023) 8697–8720.

[6] X. Zheng, J. Cai and G. Zhang, "Stock Trend Prediction Based on ARIMA-LightGBM Hybrid Model", 2022 3rd Information Communication Technologies Conference (ICTC), Nanjing, China (2022), pp. 227-231.

[7] Shreyanshisingh28, LightGBM (Light Gradient Boosting Machine) <https://www.geeksforgeeks.org/lightgbm-light-gradient-boosting-machine/>, 2023 (accessed 13.11.23).

[8] Dw คืออะไร? CALL กับ PUT ใน DW คืออะไร? นักลงทุนมือใหม่ควรรู้! <https://www.thaiwarrant.com/knowledge/articles-details-69 >, 2567 (สืบค้น 27 มี.ค. 67)

[9] ณัฐนนท์ เลี้ยงประไพพันธิ์, พิชายุตม์ ตันติพัฒน์พงศ์, ระบบการแนะนำการซื้อขายตราสาอนุพันธ์โดยใช้รูปแบบราคาทะลุผ่านกราฟสามเหลี่ยมและตัวแบบไลต์เกรเดียนต์บูสติ้งแมชชีน, (โครงงานวิทยาศาสตร์ปริญญาบัณฑิต), จุฬาลงกรณ์มหาวิทยาลัย, 2567.

Forecasting Pollen Concentration from Past Meteorological Factors Using Deep Learning

Suppakrit Sookhanonsawat[1,*,†], Paponsak Sookserisarp[1,*], Akarapark Kongsiriwong[1,*], Haruho Yamaguchi[2], Yuichiro Kawakatsu[3] and Sittichoke som-am[4]

[1]Mahidol Wittayanusorn School, Nakhon Pathom 73170, Thailand

[2]College of Informational Science and Engineering, Ritsumeikan University, Osaka, Japan

[3]Department of Electrical and Electronic Engineering, Ritsumeikan University, Shiga, Japan

[4]Department of Mathematics and Computer Science, Mahidol Wittayanusorn School, Nakhon Pathom 73170, Thailand

Abstract

Pollen is a major cause of allergic rhinitis, affecting approximately 12.3% to 39.8% of the global population, with prevalence reaching nearly 26% in the United States and almost 40% in Europe. In Japan, it is extremely high, with over 42.5%. Accurate pollen forecasts are essential for public health, yet existing forecasting systems are often costly and time-consuming. This study develops a deep learning-based model using Long Short-Term Memory (LSTM) networks to predict hourly pollen concentration based on past meteorological data. Pollen data from two cities in the Kyoto and Osaka prefectures, recorded from February to May from 2006 to 2021, were used for model training. We evaluated four different types of LSTM models, which are LSTM, Bi-LSTM, Stacked LSTM, and CNN-LSTM, each tested with four different hyperparameter configurations. The models were trained on hourly data collected from February to May from 2006 to 2021. The models utilized the past 24 hours of meteorological data to predict pollen concentration one hour ahead. The performance of each model was assessed with R-squared (R^2) scores and Root Mean Squared Error (RMSE). The optimal model was the LSTM model with ReLU activation, Adam optimizer, and MAE loss function, achieving an RMSE of 0.40235 and an R^2 of 0.72135. The findings demonstrate that deep learning approaches can achieve competitive accuracy in pollen concentration prediction, offering a viable alternative to traditional methods.

Keywords: Pollen allergy, LSTM, Meteorological data

2020 MSC: 68T07, 62M10

*Speaker (suppakrit.soo_g33@mwit.ac.th, paponsak.soo_g33@mwit.ac.th, akarapark.kon_g33@mwit.ac.th)
†Corresponding author (suppakrit.soo_g33@mwit.ac.th)

1 Introduction

Pollen is one of the main allergens causing allergic rhinitis, also known as hay fever. Allergic reactions to pollen, varying to each individual, can range from basic reactions, such as sneezing, nasal congestion, rhinorrhea, pruritus, headaches, and rashes, to life-threatening reactions like anaphylaxis. According to [1], 12.3% to 39.8% of the global population is allergic to pollen, with the demographic continuing to increase annually. In some regions, such as China, the United States, Europe, and Japan, the prevalence of pollen allergy is as high as 26%, 40%, and 42.5% of the population, respectively [2-4].

In Japan, millions experience seasonal allergic rhinitis from pollen, which is considered a public health concern [1]. The primary contributor to this phenomenon is the pollen of the Japanese cedar tree (*Cryptomeria japonica*), the most widespread tree species in Japan. With approximately 4.6 billion trees covering 18% of the country's land area, its pollen dominates the air during the dispersal season from February to April [4].

A nationwide survey found that 62.5% of affected individuals reported that their symptoms significantly interfered with daily activities [6]. The socioeconomic burden is also considerable; allergy-related industries, including pharmaceutical and preventive healthcare sectors, are estimated to generate seasonal revenues of 200 to 300 billion yen [7]. These findings underscore the wide-reaching impact of SAR on both individual well-being and the broader Japanese economy.

To mitigate the impacts of seasonal allergies, Japan employs both conventional and contemporary methods to forecast pollen. Traditional techniques, such as the Durham method, involve collecting pollen grains on coated slides and examining them under a microscope [8]. More recently, automated systems like the "Hanako-san" network and "Pollen Robo" have been implemented to provide real-time monitoring of pollen diffusion [9]. However, these systems often involve substantial infrastructure investment, making them costly and less accessible for widespread applications.

Considering these limitations, data-driven forecasting methods present a promising alternative. Recent advances in deep learning, particularly in time series modeling, offer new opportunities to predict pollen concentrations using only past meteorological data. Long Short-Term Memory (LSTM) networks, which are designed to capture temporal dependencies in sequential data, have shown significant potential in environmental and health-related prediction tasks [10], [11].

Therefore, this study aims to develop and compare LSTM-based deep learning models to forecast hourly pollen concentrations in the Kansai region of Japan using past meteorological variables. By focusing on data from Kyoto and Osaka prefectures between 2006 and 2021, this research seeks to provide an effective and scalable solution for pollen prediction, potentially improving public health responses to seasonal allergies.

2 Literature Review

Several studies have investigated the relationship between meteorological factors and pollen concentration using both statistical and machine learning approaches.

Shah and Takeuchi [12] examined the correlation between weather conditions and pollen counts of Japanese cedar and cypress in Tsu City, Japan, using data from 2009 to 2018. Pollen data were collected using Durham's sampler from February to May. Spearman's rank correlation was applied to assess the association between annual pollen counts and meteorological variables, followed by Multiple Linear Regression (MLR) for significant results. The study found that average sunshine duration in April (current year) and November (previous year), as well as average humidity in August (previous year), were positively correlated with the combined pollen count, with an R^2 of 0.980 and ANOVA $p < 0.000$. Conversely, Japanese cedar pollen showed a negative correlation with January's maximum temperature ($R^2 = 0.338$, $p < 0.045$), while cypress pollen was influenced positively by April sunshine duration and August maximum temperature, but negatively by October temperature ($R^2 = 0.872$, $p < 0.001$).

Murayama and Tonouchi [13] explored traditional and modern pollen observation methods in Japan. The Durham method, widely used across over one hundred observation stations, involves collecting pollen on Vaseline-coated slides for microscopic analysis. In contrast, the KH-3000 automatic counter utilizes laser-based identification to provide real-time data online. The study revealed that average summer temperature and solar radiation had a positive correlation with seasonal pollen levels, while warmer spring temperatures advanced the pollen season. Interestingly, in southern Japan, warmer winters delayed pollen release due to disruption in cedar tree dormancy. The study also linked global warming to increased pollen production, citing a 19% rise in solar radiation and an 11% increase in sunshine duration from 1987 to 2006.

Papadogiannaki et al. [14] applied Gradient Boosting Regression (GBR) to predict daily pollen concentrations of Oleaceae and Quercus families in Thessaloniki, Greece. Using meteorological and land surface data from ECMWF (2016–2021), the model incorporated variables such as water vapor, pressure, temperature, soil moisture, and vegetation cover. Pollen data were obtained from a Hirst-type trap at Aristotle University. The GBR model achieved high predictive accuracy with an Index of Agreement (IOA) of 0.86 for Oleaceae and 0.78 for Quercus, confirming the efficacy of ensemble-based machine learning in this domain.

Lops et al. [15] employed a Deep Convolutional Neural Network (DCNN) to estimate pollen concentrations based on weather data from 2009–2015, including temperature, humidity, rainfall, and wind. The model performed well in short-term forecasting, yielding IOA scores of 0.9 for 1-day forecasts and 0.81 for 7-day forecasts for weed pollen. Similar performance was observed for flower and grass pollen, although accuracy decreased for longer forecast horizons (e.g., grass pollen IOA dropped from 0.8 to 0.56 for 7-day forecasts).

Picornell et al. [16] compared five LSTM variants—standard LSTM, Stacked LSTM, Bidirectional LSTM, CNN-LSTM, and Autoencoder LSTM—to forecast monthly pollen integrals for Olea and Urticaceae in Málaga, Spain. The study used 30 years of meteorological data from the Spanish National Meteorological Agency and pollen data from a Hirst-type trap at the University of Málaga. Results showed that all LSTM models effectively captured interannual variability, with CNN-LSTM yielding the highest accuracy. The RMSE was generally under 12% of the annual pollen integral (APIn) for olive and under 6.5% for Urticaceae, except for autumn 2021, where performance declined slightly. Notably, Stacked LSTM did not outperform standard LSTM, while Bidirectional LSTM, CNN-LSTM, and Autoencoder LSTM offered superior results.

Collectively, these studies demonstrate the potential of both traditional statistical methods and advanced deep learning models in predicting pollen concentrations. While classical models offer interpretability, deep learning architectures such as LSTM and CNN-LSTM show promise in capturing temporal dynamics and nonlinear patterns in environmental data.

3 Methodology

The operational framework of this study is illustrated in Figure 1, outlining the complete process from data collection to model evaluation.

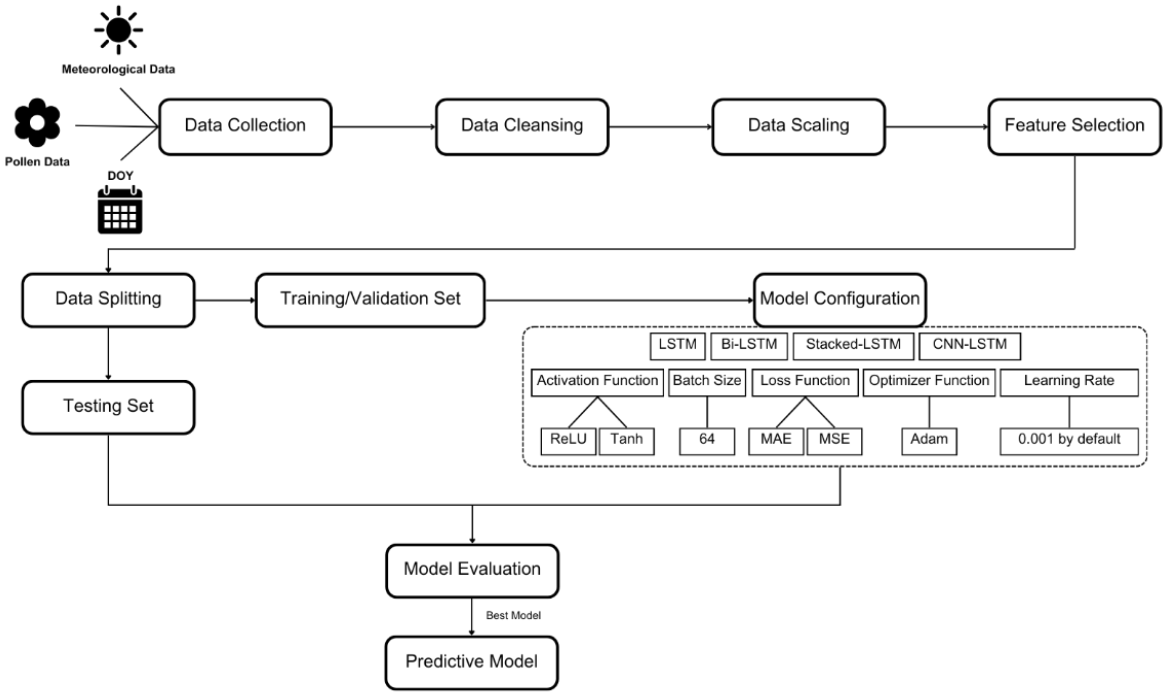


Figure 1: Operational framework

3.1 Data Collection

This study utilized hourly pollen concentration and meteorological data collected from February to June from 2006 to 2021. The selected observation sites are Kyoto City in Kyoto Prefecture and Toyonaka City in Osaka Prefecture due to the consistency and completeness of data during the entire observation period. Pollen concentration data were obtained from the Ministry of the Environment of Japan. In addition, the meteorological data were sourced from the Japan Meteorological Agency (JMA) and collected by their Automated Meteorological Data Acquisition System (AMeDAS) spread throughout Japan. The high-resolution temporal dataset allowed for fine-grained modeling of pollen dispersion patterns in response to short-term weather changes.

3.2 Data Preparation

3.2.1 Data Cleansing

The dataset for this study consisted of 46,084 records, each containing meteorological variables including temperature, precipitation, wind direction, and wind speed. After removing entries with missing values, 44,063 complete records remained, representing over 95% of the original dataset. Despite being taken into consideration, standard methods for handling missing data, such as imputation and interpolation, were not applied due to an adequate amount of clean data being available to provide efficient model training.

Although the potential impact of data removal on dataset diversity was acknowledged, the post-cleansing data still provided a wide range of values across the variables. Furthermore, imputing missing values in meteorological datasets, which can fluctuate unpredictably, risks introducing bias or unrealistic patterns that could reduce model performance. Therefore, removing incomplete entries was the most appropriate strategy in this situation.

3.2.2 Data Scaling

Given the differing magnitudes and units across meteorological variables, standard scaling was applied to normalize the data. Each feature was transformed to have a mean value of zero and a standard deviation value of one. In addition, wind direction, a circular variable, was encoded using sine and cosine transformations to retain its periodic properties, as recommended in prior time-series modeling studies [17].

3.2.3 Feature Selection

The selected features of meteorological datasets are temperature, precipitation, wind speed, and wind direction because they were the only meteorological factors provided. These factors are not the most effective in forecasting pollen concentrations [11]. However, the Japanese Meteorological Agency's datasets were the only source that were credible, accessible, and practical for the study's conditions at that time, making effective pollen forecasting methods with widely available data. Other alternatives have been presented but were not chosen due to uncertainty in credibility, consistency, and financial factors of the data.

3.2.4 Data Splitting

The dataset was chronologically split into three segments: 60% for training, 20% for validation, and 20% for testing. Temporal order was preserved to simulate real-world forecasting conditions and prevent data leakage. The training set was used to fit the models, the validation set to fine-tune hyperparameters, and the test set to evaluate the final performance on unseen data.

3.3 Model development

This study developed four deep learning models using Python and TensorFlow and Keras libraries: LSTM, Bidirectional LSTM (Bi-LSTM), Stacked LSTM, and CNN-LSTM. Each model was trained to predict the next hour's pollen concentration based on the previous 24 hours of meteorological data.

3.3.1 LSTM

The Long Short-Term Memory (LSTM) network is a type of Recurrent Neural Network (RNN) designed to retain long-range dependencies in sequential data. It is well-known for its ability to predict sequential, where past observations influence future predictions. It addresses vanishing gradient issues typical in conventional RNNs [17]. An LSTM unit comprises memory cells regulated by input, forget, and output gates, as well as candidate memory states (see Figure 2).

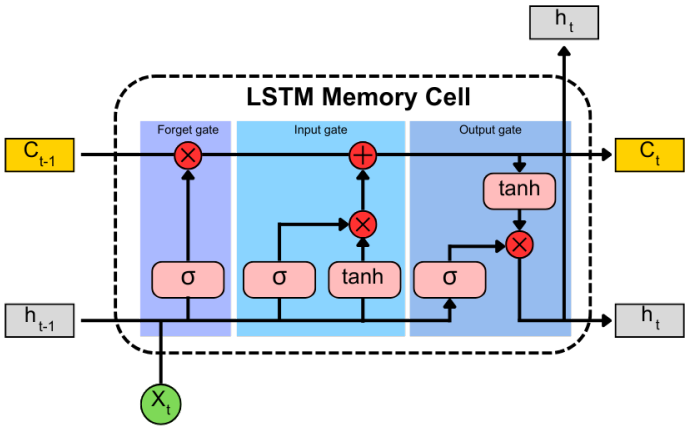


Figure 2: Architecture of an LSTM Unit

3.3.2 Bidirectional LSTM (Bi-LSTM)

The Bi-LSTM architecture processes input sequences in both forward and backward directions, enabling the model to learn from past and future contexts simultaneously [18]. This is particularly useful when predicting pollen patterns influenced by multiple interacting temporal dynamics.

3.3.3 Stacked LSTM

Stacked LSTM models involve multiple LSTM layers stacked on top of each other, enabling hierarchical representation learning. Deeper architecture can capture more complex temporal dependencies, although they may require more data and careful regularization to avoid overfitting.

3.3.4 CNN-LSTM

CNN-LSTM combines Convolutional Neural Networks (CNNs) for feature extraction with LSTM layers for temporal learning. CNN layers process local patterns in meteorological time windows before passing the sequence to LSTM units for long-term context modeling [16].

3.3.5 Model Specification

The models were implemented using the following hyperparameters:

- Activation Functions: ReLU, Tanh
- Batch Size: 64
- Loss Functions: Mean Absolute Error (MAE), Mean Squared Error (MSE)
- Optimizer: Adam
- Learning Rate: 0.001 (adaptive per optimizer dynamics)

3.4 Model Evaluation

Model performance was evaluated using the following metrics:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x_i})}$$

R-squared (R^2) measures how well the regression model approximates the real data points. It indicates the percentage of the response variable variation that is explained by a linear model.

$$R^2 = 1 - \frac{RSS}{TSS}$$

Root Mean Square Error (RMSE) is a measure of how much the predicted values from a regression model differ on average from the actual observed values. It is calculated as the square root of the average of the squared prediction errors.

4 Results

4.1 Model Evaluation Results

A total of 32 models were developed and evaluated, comprising four LSTM-based architectures (LSTM, Bi-LSTM, Stacked LSTM, and CNN-LSTM), each tested under four hyperparameter combinations across two datasets (Kyoto and Toyonaka). The models were evaluated using R-squared (R^2) and Root Mean Square Error (RMSE) on the test set.

Table 1 presents the R^2 values for each model. In general, models trained on Kyoto City data yielded higher R^2 values than those trained on Toyonaka City data. Among all configurations, the LSTM model with ReLU activation and MAE loss function trained on Kyoto data achieved the highest R^2 of 0.7595. This was closely followed by Bi-LSTM (0.7526), Stacked LSTM (0.7454), and CNN-LSTM (0.5628).

Table 2 displays RMSE values. In contrast to the R^2 results, RMSE values were consistently lower for models trained with Toyonaka City data. The lowest RMSE was observed in the Bi-LSTM model with ReLU activation and MAE loss trained on Toyonaka data, achieving an RMSE of 0.3662.

To better represent model performance across the Kansai region, the results from both cities were averaged, as shown in Table 3 (average R^2) and Table 4 (average RMSE). The use of average values is justified to generalize model effectiveness beyond localized behavior and evaluate adaptability across diverse microclimates.

Best overall average R^2: LSTM with ReLU and MAE at 0.7214.

Best overall average RMSE: LSTM with ReLU and MAE at 0.4024.

These results confirm that the LSTM model configured with ReLU activation, MAE loss function, and Adam optimizer consistently outperformed other variants across both metrics.

Table 1: R^2 of each model

Place	Activation Function	Loss Function	LSTM Type			
			LSTM	Bi-LSTM	Stacked LSTM	CNN-LSTM
Kyoto	ReLU	MSE	0.7322	0.7324	0.7421	0.5628
		MAE	0.7595	0.7526	0.7355	0.5328
	TanH	MSE	0.7536	0.7475	0.7421	0.5571
		MAE	0.7464	0.7351	0.7454	0.5229
Toyonaka	ReLU	MSE	0.6743	0.6685	0.6475	0.4230
		MAE	0.6832	0.6842	0.6291	0.4629
	TanH	MSE	0.6406	0.6734	0.6687	0.4572
		MAE	0.6711	0.6549	0.6584	0.4265

Table 2: RMSE of each model

Place	Activation Function	Loss Function	LSTM Type			
			LSTM	Bi-LSTM	Stacked LSTM	CNN-LSTM
Kyoto	ReLU	MSE	0.4622	0.4620	0.4535	0.5905
		MAE	0.4380	0.4442	0.4593	0.6105
	TanH	MSE	0.4433	0.4488	0.4535	0.5943
		MAE	0.4498	0.4596	0.4506	0.6169
Toyonaka	ReLU	MSE	0.3718	0.3751	0.3868	0.4949
		MAE	0.3667	0.3662	0.3968	0.4775
	TanH	MSE	0.3906	0.3723	0.3750	0.4800
		MAE	0.3736	0.3827	0.3808	0.4934

Table 3: Average R^2 of each model from Kyoto City and Toyonaka City

Activation Function	Loss Function	LSTM Type			
		LSTM	Bi-LSTM	Stacked LSTM	CNN-LSTM
ReLU	MSE	0.7033	0.7005	0.6948	0.4929
	MAE	0.7214	0.7184	0.6823	0.4979
TanH	MSE	0.6971	0.7105	0.7054	0.5072
	MAE	0.7088	0.6950	0.7019	0.4747

Table 4: Average RMSE of each model from Kyoto City and Toyonaka City

Activation Function	Loss Function	LSTM Type			
		LSTM	Bi-LSTM	Stacked LSTM	CNN-LSTM
ReLU	MSE	0.4170	0.4186	0.4202	0.5427
	MAE	0.4024	0.4052	0.4281	0.5440
TanH	MSE	0.4170	0.4106	0.4143	0.5372
	MAE	0.4117	0.4212	0.4157	0.5552

4.2 Real-world Accuracy Evaluation

To further assess model robustness, a real-world test was performed using meteorological data from February 1–16, 2022. The test measures the model's effectiveness in forecasting pollen concentrations for the next 24 hours and the next 168 hours. RMSEs of both scenarios were calculated using the predictions and the actual pollen concentrations of the same period. Two test scenarios were designed:

4.2.1 24-hour forecast vs actual pollen data

The RMSE for the 24-hour test was high at 0.5773 and deemed unreliable due to extremely low actual pollen concentration during this period, which led to a marginal prediction error being disproportionately penalized in RMSE.

4.2.2 16-day forecast vs actual over an extended horizon

The 16-day test also yielded a value of 0.0528, a deceptively accurate result, due to the same reason. As pollen concentrations were minimal during February, we recommend conducting real-world evaluations during peak pollen months (March to April) to gain meaningful assessments.

4.3 Discussion and Interpretation

The findings indicate that the basic LSTM model outperformed more complex variants (e.g., CNN-LSTM, Bi-LSTM, and Stacked LSTM), contradicting prior assumptions that deeper or more intricate models would yield higher accuracy [16], [18]. This may be due to the limited variability in the input features or the relatively stable seasonal pattern of pollen dispersion in the studied region.

The superior performance of the LSTM + ReLU + MAE configuration aligns with other environmental forecasting studies that recommend MAE over MSE for datasets with potential outliers [14]. Furthermore, ReLU's efficient gradient propagation likely contributed to better convergence and generalization.

While R^2 indicates how well the model captures the variance of pollen concentration, RMSE provides insight into the absolute error of predictions. Thus, using

both metrics offers a comprehensive view of performance, where R^2 captures the model's explanatory power and RMSE its practical prediction reliability.

5 Conclusion

This study aims to develop a deep learning model based on the Lon Short Term model to predict hourly pollen concentration using past meteorological data in the Kansai region of Japan. Datasets from Kyoto City and Toyonaka City between 2006 and 2021 were used, comprising pollen concentration, temperature, precipitation, wind speed, and wind direction. Four LSTM variants, which are LSTM, Bi-LSTM, Stacked LSTM, and CNN-LSTM, each with four hyperparameter configurations of MAE and MSE loss functions and Tanh and ReLU activation functions, were tested.

The results demonstrated that the standard LSTM model with ReLU activation, MAE loss function, and Adam optimizer provided the best overall performance, achieving an R^2 of 0.7214 and an RMSE of 0.4024 when averaged across both locations. These results indicate that the model is capable of accurately predicting pollen concentration one hour in advance based solely on past weather data.

Although more complex models such as Bi-LSTM, Stacked LSTM, and CNN-LSTM were explored, their performance did not surpass that of the standard LSTM. This suggests that simpler architecture may be more effective for environmental forecasting problems with well-defined temporal patterns. Additionally, real-world testing showed the need to evaluate model performance during peak pollen seasons for more significant assessments.

In conclusion, this research confirms the viability of LSTM deep learning models as an effective and scalable alternative to traditional pollen forecasting methods. With further validation and integration into public health systems, such models could support timely interventions for individuals affected by seasonal allergic rhinitis.

6 Future Work

Due to climate change, the model may not perform as effectively as expected. To address this, the researchers aim to apply more features, improve the effectiveness, and the reliability of the models in the future. The researchers also aim to develop an application integrating the best-performing model. The application will include predictions for multiple cities. A user feedback survey will be conducted for future improvements.

References

[1] China Meteorological Administration, "Increasing incidence of pollen allergy globally," CMA News, Apr. 2024. [Online]. Available: https://www.cma.gov.cn/en/outreach/news/202404/t20240422_6215001.html

[2] Asthma and Allergy Foundation of America, "Allergy Facts and Figures," 2022. [Online]. Available: https://aafa.org/wp-content/uploads/2022/08/aafa-allergy-facts-and-figures.pdf

[3] M. Sakashita, T. Hirota, Y. Harada, et al., “Prevalence of allergic rhinitis and sensitization to common aeroallergens in Japan: A population-based study,” Allergy, vol. 63, no. 3, pp. 320–328, 2008.

[4] H. Ohashi-Doi, et al., “Cryptomeria pollen dispersion and forest coverage in Japan,” Journal of Environmental Allergy Studies, vol. XX, no. X, pp. XXX–XXX, 2022.

[5] D. P. Skoner, “Allergic rhinitis: Definition, epidemiology, pathophysiology, detection, and diagnosis,” J. Allergy Clin. Immunol., vol. 108, no. 1, pp. S2–S8, 2001.

[6] K. Okuda, “Impact of cedar pollinosis on quality of life: A national survey,” Japanese Journal of Allergology, vol. XX, no. X, pp. XXX–XXX, 2003.

[7] M. Yamada, et al., “Economic impact of seasonal allergic rhinitis in Japan,” Allergol. Int., vol. XX, no. X, pp. XXX–XXX, 2014.

[8] R. Durham, “The Durham gravimetric pollen trap,” Rev. Palaeobot. Palynol., vol. XX, no. X, pp. XXX–XXX, 1946.

[9] Ministry of the Environment, Japan, “Real-time pollen monitoring using Hanako-san,” 2022. [Online]. Available: https://www.env.go.jp/en

[10] F. Chollet, “Deep learning with Python,” 2nd ed., Manning Publications, 2021.

[11] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997. [12] M. Shah and Y. Takeuchi, “Relationship between climate change and pollen counts of Japanese cedar and cypress trees in Tsu City,” J. Climate Allergy Stud., vol. 8, no. 2, pp. 45–58, 2020.

[13] H. Murayama and M. Tonouchi, “Meteorological factors and the timing of cedar pollen dispersal in Japan,” Aerobiologia, vol. 23, no. 4, pp. 277–285, 2007.

[14] E. Papadogiannaki, I. Birlidis, and K. Stefanopoulos, “Forecasting daily pollen concentrations using gradient boosting regression,” Environmental Modelling & Software, vol. 168, 2023.

[15] Y. Lops, A. M. Gadaleta, and N. Novielli, “Estimating airborne pollen concentration using deep convolutional neural networks,” Procedia Comput. Sci., vol. 170, pp. 549–556, 2020.

[16] A. Picornell, S. Pérez-Badia, and M. Belmonte, “Forecasting pollen integrals using advanced LSTM variants: A 30-year analysis from Málaga, Spain,” Environmental Research, vol. 241, 2024.

[17] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[18] M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks,” IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2673–2681, 1997.

AI-Assisted Reference Retrieval For Mathematical Theorem Proving: A Step Toward Reliable Proof Generation

Krong Krongyuth[1,*], Ekawat Chaowicharat[1], and Pannapa Changpetch[1]

[1]Department of Mathematics, Faculty of Science, Mahidol University,
272 Rama VI Road, Ratchathewi, Bangkok 10400, Thailand

Abstract

The ability to construct mathematical proofs is a fundamental yet challenging aspect of pure mathematics, requiring rigorous logical reasoning and a strong grasp of foundational concepts. Recent advances in artificial intelligence (AI) have explored enhancing mathematical reasoning, but AI models still struggle with reliability due to hallucinated proofs and evaluation difficulties. Previous research has relieved this issue using in-context learning, a technique that supplies relevant information to the LLM. This approach helps reduce hallucinations and enhances the model's reasoning capabilities. This research focuses on improving mathematical reference retrieval to assist in theorem proving by leveraging NaturalProofs, a corpus of mathematical statements and proofs, alongside various information retrieval techniques. We employ all-MiniLM-L6-v2, an open-source, widely used similarity search model, Gemini Embedding, the current state-of-the-art retriever, combined with traditional retrieval methods such as TF-IDF and BM25. Additionally, we fine-tune all-MiniLM-L6-v2 using a curated dataset from ProofWiki to optimize reference retrieval performance. Two experiments are conducted using different versions of the reference database: one is With-Theorems Experiment, and one is the Without-Theorem Experiment. Our results demonstrate that the **fine-tuned all-MiniLM-L6-v2 outperforms the state-of-the-art retriever in mAP and MRR in both experiments**. Moreover, we noticed the difference in performance in both experiments, showing promise that a lower data source density has a good effect on the retriever.

Keywords: Mathematical Reference Retrieval, Theorem Proving, Information Retrieval, Sentence Similarity, Artificial Intelligence

2020 MSC: 68V05, 68P20, 68T50, 68T01

1 Introduction

The rise of Artificial Intelligence (AI) has significantly transformed the way people learn and work. Large Language Models (LLMs) applications such as ChatGPT, Gemini, and Llama are widely used to alleviate the workload associated with complex tasks [8, 14, 18]. In mathematical research, a major focus is on enhancing AI's capabilities in mathematical reasoning, either by increasing model complexity or by introducing more challenging benchmarks [2, 3, 9–11].

*Speaker and corresponding author (krong.kro@student.mahidol.edu)

However, despite advancements in AI-driven proof generation, current models still struggle to produce reliable results. These challenges arise from the need for rigorous logical reasoning, a strong foundational understanding of mathematics, and the tendency of models to generate hallucinated proofs [1, 5]. Additionally, the difficulty of evaluating mathematical proofs further complicates efforts to improve AI's performance in this domain.

Since the introduce of **Retrieval-Augmented Generation (RAG)** propose the technique that grants capabilities of Information Retrieval to AI and providing promising result with decreasing of model hallucination within the domain of retrieved information [13].

The concept of RAG is highly effective in enhancing the capabilities of LLMs. By leveraging Information Retrieval, LLMs gain the capabilities to access vast external corpora (e.g., Wikipedia, specific PDF files) through a retriever. This approach improves question-answering performance by grounding responses in the retrieved corpus, ensuring more accurate and contextually relevant outputs [13].

This research aims to enhance the performance of reference retrievers in providing precise and effective premises for constructing mathematical proofs. To achieve this, we utilize **NaturalProofs**, an open-source corpus containing a rich collection of mathematical statements and their corresponding proofs written in natural mathematical language and sourced from various references [21]. Our approach leverages all-MiniLM-L6-v2, an efficient and wildly used model, Gemini Embedding, the current state-of-the-art retriever, and other models to accomplish our research objectives:

- all-MiniLM-L6-v2: A highly efficient and widely used similarity search model on Hugging Face, built on the concepts of Sentence-BERT and MiniLM [16, 20].
- Gemini-Embedding: A state-of-the-art embedding model leveraging the power of Gemini, Google's most capable large language model [**?**].
- TF-IDF: Well-known statistical model in text mining, developed for semantic search and information retrieval. It ranks documents based on Term Frequency (TF) and Inverse Document Frequency (IDF) scores [4].
- BM25: An enhanced version of TF-IDF that improves its scoring mechanism, preventing term frequency scores from increasing linearly and producing more realistic rankings [17].
- BERT: A widely known natural language model often used for masked token prediction and next-sentence prediction. In this research, we utilize BERT-Base (uncased) as a retriever to compare its performance with all-MiniLM-L6-v2 [6].

All models in this study are commonly used as retrievers in information retrieval tasks. We focus on developing an efficient retrieval system to retrieve the most relevant references and establish a solid infrastructure for future research.

2 Literature Review

Before delving deeper into the research, it is essential to first understand the foundational concepts and preliminary work in this field.

2.1 Retrieval Augmented Generation (RAG)

RAG is a framework that enhances the capabilities of LLMs by integrating information retrieval techniques [13]. Instead of relying solely on pre-trained knowledge, RAG retrieves relevant documents from an external knowledge base, such as Wikipedia or other data sources, to generate more informed and contextually accurate responses.

The effectiveness of RAG heavily depends on the quality of the retrieved data. High-quality data ensures that the retrieved documents are relevant, well-structured, and accurate, reducing the chances of hallucination and misinformation in generated responses.

In the context of mathematical theorem proving, good data quality is essential for maintaining logical consistency and precision. Well-structured theorem databases with correctly formatted proofs allow RAG to retrieve the most relevant mathematical references, improving the model's ability to generate valid proof steps and assist researchers or learners in solving complex mathematical problems efficiently.

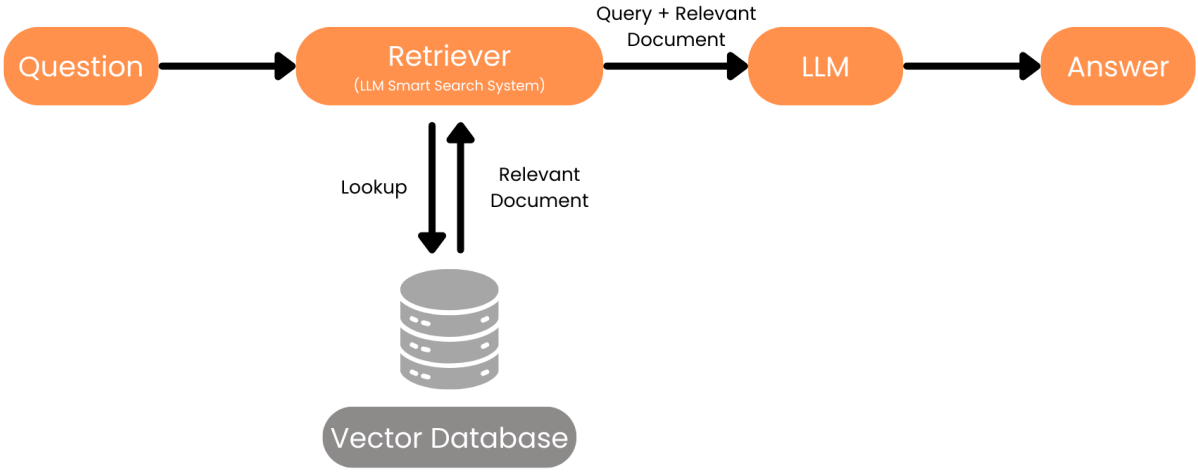


Figure 1: Retrieval Augmented Generation (RAG) framework

2.2 Mathematical reference retrieval

Mathematical reference retrieval refers to the process of identifying and retrieving relevant mathematical statements, such as theorems, definitions, and proofs, from a structured database or corpus. One of the most influential studies for this work is NaturalProofs. This research introduced a corpus of mathematical statements collected from various sources, including ProofWiki, Stacks Project, and mathematics textbooks [21]. Notably, their study closely aligns with ours, as they leveraged NaturalProofs to enhance mathematical reference retrieval performance. Given a mathematical claim, their approach retrieves a set of relevant references, such as theorems, lemmas, and definitions, that appear in its proof.

In contrast to their focus on retrieving references that contribute to the proof of a given claim, our approach emphasizes retrieving mathematical statements that explicitly appear within the given claim itself. This distinction ensures that our retrieval process directly captures the contextual and structural dependencies present in the mathematical expression.

2.3 Sentence similarity search

Sentence similarity search is a fundamental technique in information retrieval, designed to identify the most relevant documents within a large corpus. This field has received significant attention and has undergone extensive development. Concepts from text mining, such as Word2Vec, TF-IDF, and BM25, have led to numerous modern applications [4, 15, 17]. The introduction of Transformer architectures revolutionized natural language processing, enabling models like BERT and Sentence-BERT to enhance information retrieval for more complex textual data [6, 16, 19].

3 Methodology

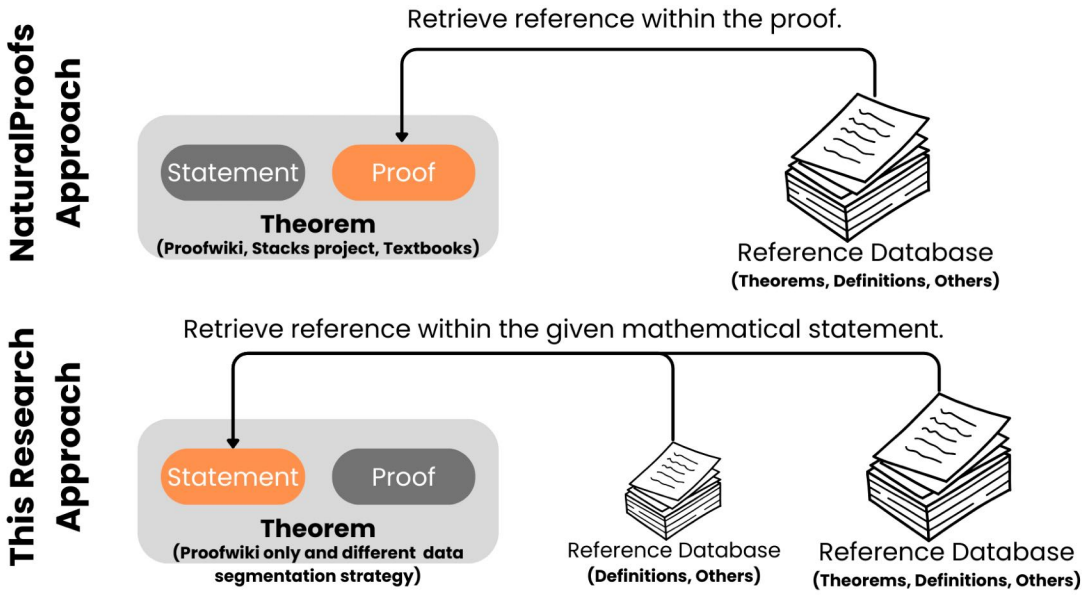


Figure 2: The difference between the NaturalProofs approach and This Research approach

This research aims to improve mathematical reference retrieval to support mathematicians and students in theorem proving by leveraging various information retrieval techniques. Unlike prior work Figure 2, which focuses on retrieving references that appear in a theorem's proof, our approach retrieves references for mathematical terms within a given theorem statement from the reference database a collection of theorems, definitions, and other mathematical resources compiled by NaturalProofs [21]. To achieve this, we implement sentence similarity search using **Facebook AI Similarity Search (FAISS)**, a vector database developed by Meta, designed for efficiently storing large-scale vector embeddings and performing high-speed similarity searches [7].

To evaluate our approach, we conduct two experiments using different sizes of the reference database. The details of these experiments will be discussed in the following sections.

3.1 Reference Database

We utilize the ProofWiki dataset from NaturalProofs to construct the **reference database**, which comprises **19,734 theorems, 12,420 definitions, and 1,006 other** mathematical resources.

reference database, in the context of information retrieval, refers to a Vector Database that stores and organizes documents as high-dimensional embeddings. Instead of relying on traditional keyword-based searches, a vector database enables efficient similarity searches by representing textual data in a continuous vector space.

However, through data exploration, we observed that over **96%** of theorems rely solely on definitions and other resources as references. Based on this insight, we designed two experimental settings: "With-Theorem Database" and "Without-Theorem Database," reducing the reference database size by **60%** to assess its impact on retrieval performance.

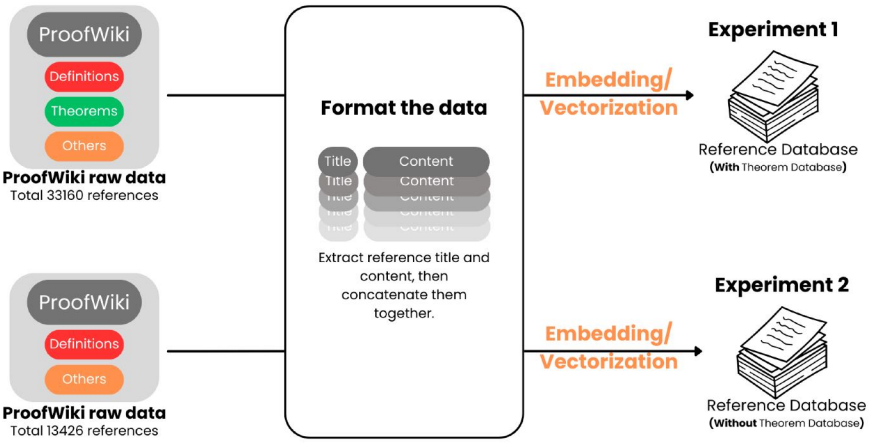


Figure 3: Process of formatting ProofWiki references into a reference database for retrieval.

Based on this finding, we created two versions of the reference database:

- **With-Theorem Database** Contains 33,160 entries (Theorems, Definitions, and Others), each formatted as "Title: Content." then vectorized and stored in the database Figure 3.

- **Without-Theorem Database** Contains 13,426 entries (Definitions and Others), each formatted as "Title: Content." then vectorized and stored in the database Figure 3.

The data structure in the database remains the same, but the amount of stored data varies, leading to distinct retrieval results for each approach, as shown in section 4.

3.2 Reference Retrieval

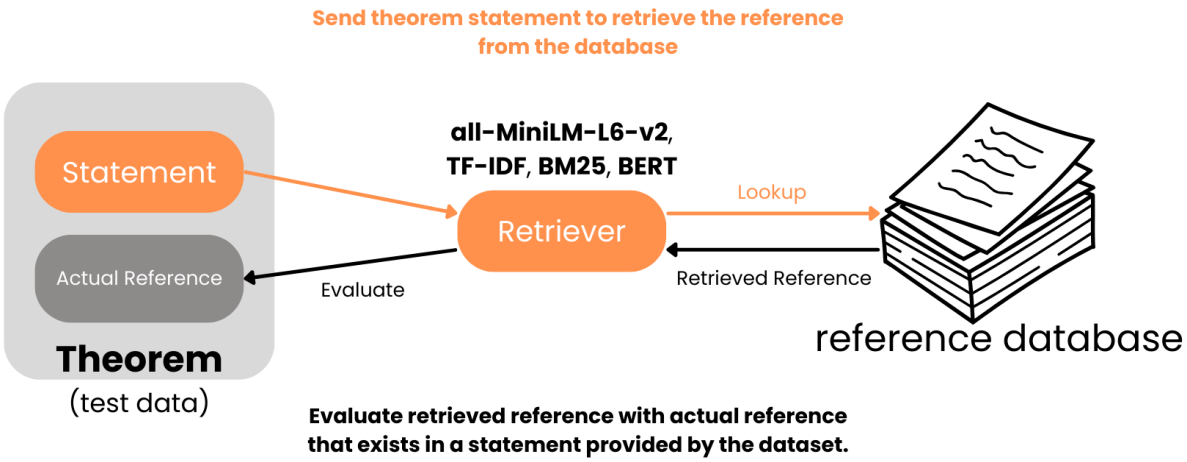


Figure 4: Diagram of retrieve and evaluate retrieved reference.

As mentioned previously, this project focuses on retrieving references within mathematical statements, specifically theorems, with the expectation of positively impacting proof construction. This is achieved through a retriever, a tool designed to identify the most relevant documents from a large collection referred to here as the reference database, as illustrated in Figure 4.

The all-MiniLM-L6-v2, Gemini Embedding and BERT retrievers utilize FAISS to store and index the reference database, significantly improving retrieval efficiency. However, due to differences in vectorization techniques, TF-IDF and BM25 require alternative approaches for database construction.

All retrieval methods rely on vector similarity to identify the most relevant references. Some methods, such as FAISS, incorporate additional optimizations to enhance search efficiency, ensuring faster and more accurate retrieval results.

3.3 Fine-Tuned all-MiniLM-L6-v2

Moreover, we develop a fine-tuned all-MiniLM-L6-v2 model specifically for this task by training it on the first **80%** of the theorems in the ProofWiki raw dataset, using CosineSimilarityLoss as the loss function.

Additionally, we fine-tuned the all-MiniLM-L6-v2 model separately using two distinct training sets, both derived from the first 80% of the theorems, to support different experimental settings.

training_set

	theorems_content	refs_content	score
0	Let $\left({S, \preceq}\right)$ be an [[Defi	Let $\left({S, \preccurlyeq}\right)$ be an [[Defir	1.0
1	There exists only one [[Definition:Strictly	Let n be a [[Definition:Number|number]] exp	1.0
2	Let $p, q \in \R_{\ne 0}$ be non-zero [[D	:$\map {\dfrac \d {\d x} } {\map {\tanh^{-1} } {\	0.0
4	Let $f: \R \to \R$ be the [[Definition:Absc	Let f be a [[Definition:Real Function|real func	1.0
6	The [[Definition:Disk|unit ball]] $\mathbb	Two [[Definition:Set|sets]] $S, T \subset \R^n$	1.0
12	There cannot exist a [[Definition:Sequen	Let $f: S \to T$ be a [[Definition:Mapping|mapp	1.0
15	Every [[Definition:Delta-Algebra|δ-	:$\vdash \paren {\paren {p \implies r} \land \pa	0.0
16	The [[Definition:Real Number|real numbe	Let $\mathbb F$ be a [[Definition:Field (Abstra	0.0
19	Of the 256 different types of [[Definitio	The [[Definition:Topological Subspace|induced	0.0
27	Let $\struct {G, \circ}$ be a [[Definition:C	Let $\struct {G, \circ}$ be a [[Definition:Group|	1.0
29	The [[Definition:Set Union|union]] of the [	The '''Cauchy criterion''' is the [[Definition:Cor	0.0
32	Let $\N$ be the [[Definition:Natural Num	$\mathbf A$ is in '''non-unity echelon form''' {	0.0

Figure 5: Illustration of sample training data from ProofWiki used in the Fine-Tuning process

Figure 5 illustrates a sample of the training data used in the fine-tuning process. As mentioned earlier, we utilize Cosine Similarity as the loss function, which requires a specific data format consisting of "theorems content", "reference content", and a corresponding "score" as the label. The model learns to adjust its parameters to transform "theorems content" and "reference content" into vectors that produce the given label score.

In our approach, we match a theorem's content with its corresponding reference content and assign a score of **1.0 (Maximum Score)**. However, without incorporating negative examples into the training set, the model may face challenges in learning effectively. To address this, we also pair a theorem's content with unrelated references and assign a score of **0 (Minimum Score)**, ensuring a balanced training process.

We fine-tune all-MiniLM-L6-v2 on Google Colab using the **sentence_transformers** package, allowing us to specify training parameters such as the loss function, number of epochs, and evaluation strategy.

A key distinction between the two training sets lies in their construction:

- For the "Without-Theorem Database" experiment, we filtered out any references that were theorems before training.
- For the "With-Theorem Database" experiment, no such filtering was necessary.

After completing the fine-tuning process, we obtain the Fine-Tuned model, which is now ready for evaluation. The experimental results can be found in section 4.

3.4 Evaluation

We evaluate the reference retrieval performance by comparing the retrieved references with the actual references provided in the dataset, as illustrated in Figure 4. In both experiments, we retrieve 3, 5, 10, and 50 documents from the reference database, as constructing a proof often requires multiple references.

To assess performance, we use well-established evaluation metrics in information retrieval, including Precision@K, Recall@K, and Mean Average Precision (mAP), which are standard for measuring overall retrieval effectiveness.

In addition, we add Reciprocal Rank (RR) and Mean Reciprocal Rank (MRR) to evaluate the performance of early retrieved reference from reference database which will detail as following:

- **Precision@K** The proportion of relevant documents among the top K retrieved documents. It measures how many of the retrieved results are actually relevant.

$$P@K = \frac{Number\ of\ relevant\ references\ among\ K}{K} \tag{3.1}$$

- **Recall@K** The proportion of relevant documents retrieved out of all relevant documents available. It assesses how well the system captures all possible correct references.

$$R@K = \frac{Number\ of\ relevant\ references\ among\ K}{Total\ number\ of\ reference} \tag{3.2}$$

- **Reciprocal Rank (RR)** The reciprocal of the rank of the first relevant document retrieved. It evaluates how early the first relevant result appears.

$$RR_i = \begin{cases} \frac{1}{rank_i} & \text{if } rank_i > 0 \text{ and } rank_i < K \\ 0 & \text{if } rank_i \text{ doesn't exist in the first } K \text{ elements.} \end{cases} \tag{3.3}$$

- **Mean Reciprocal Rank (MRR)** The average of the reciprocal ranks across multiple queries, providing a general measure of ranking quality.

$$MRR = \frac{1}{N} \sum_{i=1}^{N} RR_i \tag{3.4}$$

- **Mean Average Precision (mAP)** The mean of the Average Precision (AP) over all queries, considering both precision and recall at various ranks to assess overall ranking effectiveness.

$$mAP = \frac{1}{N} \sum_{k=1}^{k=N} AP_k \tag{3.5}$$

Each parameter corresponds to the following:

- **K** is the number of retrieved document.
- $\boldsymbol{rank_i}$ represents the position of the first relevant reference document retrieved, where $\boldsymbol{i}$ indicates the index of the theorems in the testing set.

- $\boldsymbol{N}$ is the total number of queries.
- $\boldsymbol{AP_k}$ The average precision of all relevant documents for each query.

We use the first 80% of the data for fine-tuning. To prevent data leakage, the remaining 20% is reserved as the test set. As a result, the test set consists of 2,746 rows and will use to evaluate both experiments.

4 Main Results

We assess the retrieval performance using a comprehensive set of evaluation metrics to ensure a thorough analysis. The experiments are conducted under two different settings: "With-Theorem Database" and "Without-Theorem Database". These settings allow us to compare the impact of including theorem references in the retrieval process. The results for each retrieval approach under these conditions are presented below.

The following table presents the overall results from both experiments. Values that are bold and underlined indicate the **highest scores**, bold numbers represent the **second highest scores**, and underlined numbers denote the third highest scores. Scores for P@3, P@5, P@10, P@50, R@3, R@5, R@10, and R@50 are reported as average values to show the overall score for each model.

With-Theorem Database										
	mAP	MRR	P@3	P@5	P@10	P@50	R@3	R@5	R@10	R@50
Fine-Tuned all-MiniLM-L6-v2	**12.46**	**0.31**	**12.43**	**10.22**	**7.23**	2.48	**13.29**	**17.89**	**24.63**	41.50
all-MiniLM-L6-v2	10.47	0.23	9.63	8.89	7.06	2.94	9.31	13.75	21.78	43.59
BERT-Base (uncased)	1.56	0.05	1.85	1.62	1.26	0.63	1.42	2.25	3.47	8.94
TF-IDF	10.64	0.23	9.49	8.40	7.06	**3.23**	9.10	13.42	21.81	**48.16**
BM25	2.87	0.09	2.86	2.64	2.30	1.28	2.28	3.48	5.95	16.96
Gemini Embedding	**11.01**	**0.24**	**10.16**	**9.11**	**7.17**	**2.95**	**10.18**	**14.77**	**22.68**	**44.15**

Table 1: The table presents the overall retrieval performance using the **With-Theorem Database** setting.

As shown in Table 1, the fine-tuned all-MiniLM-L6-v2 model outperforms all other approaches across eight evaluation metrics, achieving a **12.46%** mean Average Precision (mAP) and **0.31** Mean Reciprocal Rank (MRR). Notably, the fine-tuned model surpasses the Gemini Embedding model in both mAP and MRR, highlighting its strong in-field performance compared to a state-of-the-art baseline. Nevertheless, Gemini Embedding demonstrates consistent performance, ranking as the first runner-up across all metrics.

Overall, the retrieval scores in this experiment are relatively low across all models. However, TF-IDF, a traditional information retrieval method, delivers impressive results, ranking among the top three models in some metric. Notably, it performs comparably to all-MiniLM-L6-v2 and even outperforms the Fine-Tuned all-MiniLM-L6-v2 in P@50 and R@50, demonstrating its effectiveness in retrieving references at higher precision and recall thresholds.

Without-Theorem Database										
	mAP	MRR	P@3	P@5	P@10	P@50	R@3	R@5	R@10	R@50
Fine-Tuned all-MiniLM-L6-v2	**30.09**	**0.71**	**29.27**	**20.17**	11.60	3.17	**32.60**	**36.58**	40.58	50.85
all-MiniLM-L6-v2	25.73	0.53	25.01	18.81	12.52	4.02	23.97	29.35	38.64	59.69
BERT-Base (uncased)	5.24	0.16	6.49	4.76	3.16	1.08	5.45	6.66	9.00	15.44
TF-IDF	25.53	0.51	23.95	19.34	**13.13**	**4.46**	22.98	31.19	**40.83**	**65.18**
BM25	8.85	0.24	10.23	8.05	5.83	2.02	8.33	10.77	15.62	28.29
Gemini Embedding	**29.44**	**0.59**	**28.07**	**20.71**	**13.40**	**4.43**	**26.85**	**32.49**	**41.23**	**65.03**

Table 2: The table presents the overall retrieval performance using the **Without-Theorem Database** setting.

As shown in Table 2, the performance of each model in the "Without-Theorem Database" setting shows a significant difference from that in the "With-Theorem Database" setting in Table 1. The Fine-Tuned model continues to outperform others in more than **five metrics**, achieving **30.09% in mAP** and **0.71 in MRR**, both of which serve as strong indicators of overall retrieval performance and still outperformed the Gemini Embedding model. On the other hand, Gemini Embedding continues to demonstrate impressive results as a state-of-the-art retriever, ranking among the top two models across all metrics and outperforming the Fine-Tuned all-MiniLM-L6-v2 in P@5, P@10, P@50, R@10, and R@50.

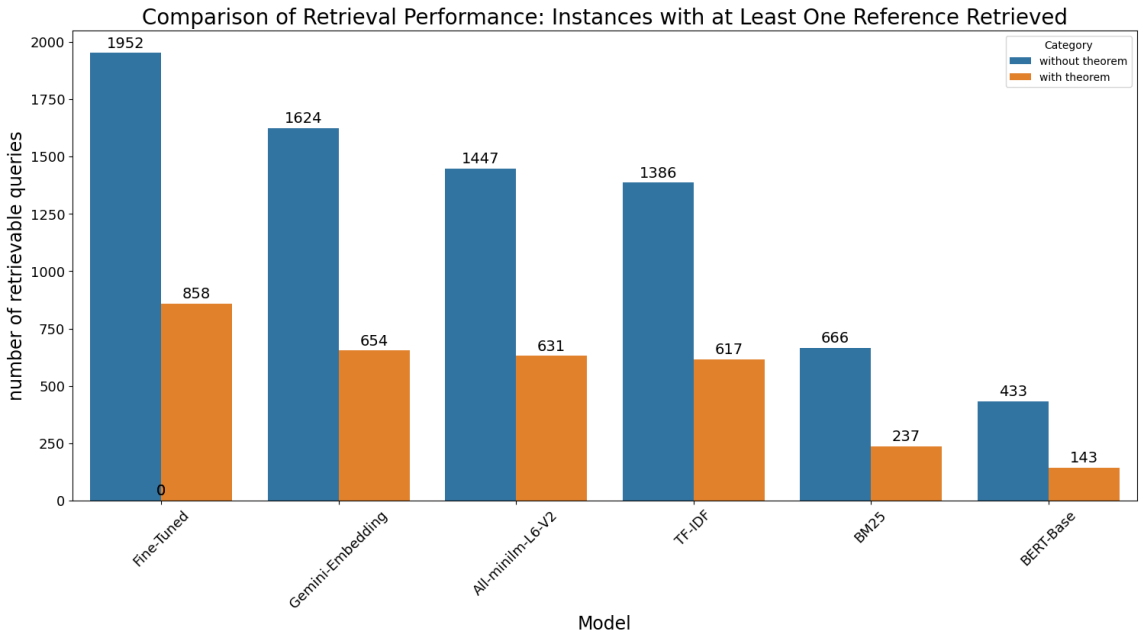


Figure 6: Comparison of Retrievable Queries.

Figure 6 compares the retrieval performance of different models based on the number of queries for which at least one reference was retrieved. The results are shown for two conditions:

With-Theorem Experiment and Without-Theorem Experiment.

- The Fine-Tuned all-MiniLM-L6-v2 retrieves the highest number of references, particularly in the Without-Theorem Database Experiment (1952 queries), significantly outperforming state-of-the-art and other models.
- all-MiniLM-L6-v2 and TF-IDF follow closely, with TF-IDF achieving a retrieval count of 1386 in the Without-Theorem Database Experiment.
- BM25 and BERT-Base show the lowest retrieval counts, with BERT-Base performing the worst overall.
- Across all models, the Without-Theorem Database Experiment consistently results in a higher number of retrievable queries compared to the With-Theorem Database Experiment.

This suggests that the number of references stored in the database influences retrieval effectiveness, with fewer references improving the model's ability to retrieve relevant results for given queries.

With Theorem Database		**After Fine-Tuning**	
		Zero	Non-zero
Before Fine-Tuning	Zero	63	447
	Non-zero	96	2140

Table 3: Effect of Fine-Tuning, reported by Reciprocal Rank (RR) scores in With-Theorem Database experiment

Without Theorem Database		**After Fine-Tuning**	
		Zero	Non-zero
Before Fine-Tuning	Zero	18	194
	Non-zero	35	2499

Table 4: Effect of Fine-Tuning, reported by Reciprocal Rank (RR) scores in Without-Theorem Database experiment

From Table 3 and Table 4 presents the impact of Fine-Tuning in both experimental settings by comparing the Reciprocal Rank (RR) scores before and after Fine-Tuning. We analyze three key cases:

- Zero to Non-zero: The document was not retrieved before Fine-Tuning but was successfully retrieved afterward (447 for With-Theorem Database, 194 for Without-Theorem Database).

- Non-zero to Zero: The document was retrieved before Fine-Tuning but not afterward (96 for With-Theorem Database, 35 for Without-Theorem Database).

- Zero to Zero: The document remained unretrieved even after Fine-Tuning (63 for With-Theorem Database, 18 for Without-Theorem Database).

As shown in Table 3 and Table 4, Fine-Tuning significantly improves retrieval performance in both experiments. It enhances the model's ability to retrieve more relevant documents that were previously missed.

5 Conclusion

This study uses various information retrieval techniques to explore AI-assisted reference retrieval for mathematical theorem proving. By fine-tuning the all-MiniLM-L6-v2 model using a curated dataset from ProofWiki, we achieved significant improvements in retrieval performance. Notably, in the "Without-Theorem Database" setting, our fine-tuned model attained a **mAP of 30.09%** and a **MRR of 0.71**. In the "With-Theorem Database" setting, it achieved a **mAP of 12.46%** and an **MRR of 0.31**. Our approach has slightly outperformed Gemini Embedding, a current state-of-the-art retriever, and other retrieval methods for mAP and MRR in both experiments. In addition, we have noticed the difference between both experiments, showing the promise that lower data source density might have a good effect on the retriever.

6 Discussion

The experiments indicate that excluding theorems from the reference database improves retrieval effectiveness, even though some theorems inherently require theorem-based references. This improvement possibly due to reducing noise and enhancing the relevance of retrieved documents. Additionally, traditional methods such as TF-IDF continue to perform competitively, demonstrating their robustness despite advancements in neural retrieval models.

While the fine-tuned all-MiniLM-L6-v2 significantly improves retrieval accuracy, retrieval effectiveness across all models remains relatively low, especially in the "With-Theorem Database" setting. This suggests potential limitations in the dataset structure, model fine-tuning process, or the inherent complexity of theorem retrieval.

7 Future Works

Our findings emphasize the potential of AI-driven retrieval models in enhancing reference retrieval for theorem proving. The significant performance boost from fine-tuning suggests that domain-specific adaptation is crucial for improving retrieval accuracy in mathematical contexts. Several challenges and areas for future improvement remain:

- **Database Structuring and Quality** The performance gap between the two experimental settings suggests that data quality and structuring play a key role in retrieval success. Further refinement of the reference database, such as curating high-quality mathematical statements and improving data preprocessing, could enhance results.
- **Integration with Large Language Models (LLMs)** While this study focuses on retrieval, integrating these retrieval models into proof generation systems could help in producing more reliable AI-assisted theorem proving. Future work could explore combining retrieval with Retrieval-Augmented Generation (RAG) to enhance reasoning capabilities.
- **Fine-Tuning and Negative Sampling Strategies** Although fine-tuning improves retrieval effectiveness, incorporating more diverse negative samples during training could help refine model precision. Additionally, exploring other similarity-based training objectives may further enhance retrieval efficiency.
- **Beyond Sentence Similarity Retrieval** Future research could explore alternative retrieval architectures, such as contrastive learning-based retrievers for more robust theorem retrieval [12].

Overall, this research highlights the growing role of AI in mathematical theorem proving and sets a foundation for future advancements in automated proof generation and mathematical reasoning.

8 Acknowledgment

We would like to express our sincere gratitude to the Department of Mathematics at Mahidol University for providing the necessary resources and a supportive research environment. Finally, we would like to thank the NaturalProofs team for making their dataset available, which played a crucial role in the development of this research.

References

[1] Janice Ahn, Rishu Verma, Renze Lou, Di Liu, Rui Zhang, and Wenpeng Yin. Large language models for mathematical reasoning: Progresses and challenges. *arXiv preprint arXiv:2402.00157*, 2024.

[2] Zhangir Azerbayev, Bartosz Piotrowski, Hailey Schoelkopf, Edward W Ayers, Dragomir Radev, and Jeremy Avigad. Proofnet: Autoformalizing and formally proving undergraduate-level mathematics. *arXiv preprint arXiv:2302.12433*, 2023.

[3] Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, et al. Training verifiers to solve math word problems. *arXiv preprint arXiv:2110.14168*, 2021.

[4] Bijoyan Das and Sarit Chakraborty. An improved text sentiment classification model using tf-idf and next word negation. *arXiv preprint arXiv:1806.06407*, 2018.

[5] Aniket Deroy and Subhankar Maity. Exploring the mathematical reasoning capabilities of gemini. *Authorea Preprints*, 2024.

[6] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pages 4171–4186, 2019.

[7] Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvasy, Pierre-Emmanuel Mazaré, Maria Lomeli, Lucas Hosseini, and Hervé Jégou. The faiss library. *arXiv preprint arXiv:2401.08281*, 2024.

[8] Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.

[9] Bofei Gao, Feifan Song, Zhe Yang, Zefan Cai, Yibo Miao, Qingxiu Dong, Lei Li, Chenghao Ma, Liang Chen, Runxin Xu, et al. Omni-math: A universal olympiad level mathematic benchmark for large language models. *arXiv preprint arXiv:2410.07985*, 2024.

[10] Elliot Glazer, Ege Erdil, Tamay Besiroglu, Diego Chicharro, Evan Chen, Alex Gunning, Caroline Falkman Olsson, Jean-Stanislas Denain, Anson Ho, Emily de Oliveira Santos, et al. Frontiermath: A benchmark for evaluating advanced mathematical reasoning in ai. *arXiv preprint arXiv:2411.04872*, 2024.

[11] Dan Hendrycks, Collin Burns, Saurav Kadavath, Akul Arora, Steven Basart, Eric Tang, Dawn Song, and Jacob Steinhardt. Measuring mathematical problem solving with the math dataset. *arXiv preprint arXiv:2103.03874*, 2021.

[12] Gautier Izacard, Mathilde Caron, Lucas Hosseini, Sebastian Riedel, Piotr Bojanowski, Armand Joulin, and Edouard Grave. Unsupervised dense information retrieval with contrastive learning. *arXiv preprint arXiv:2112.09118*, 2021.

[13] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, 33:9459–9474, 2020.

[14] Yiheng Liu, Tianle Han, Siyuan Ma, Jiayue Zhang, Yuanyuan Yang, Jiaming Tian, Hao He, Antong Li, Mengshen He, Zhengliang Liu, et al. Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-radiology*, 1(2):100017, 2023.

[15] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[16] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*, 2019.

[17] Stephen Robertson, Hugo Zaragoza, et al. The probabilistic relevance framework: Bm25 and beyond. *Foundations and Trends® in Information Retrieval*, 3(4):333–389, 2009.

[18] Gemini Team, Petko Georgiev, Ving Ian Lei, Ryan Burnell, Libin Bai, Anmol Gulati, Garrett Tanzer, Damien Vincent, Zhufeng Pan, Shibo Wang, et al. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. *arXiv preprint arXiv:2403.05530*, 2024.

[19] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[20] Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. Minilm: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. *Advances in neural information processing systems*, 33:5776–5788, 2020.

[21] Sean Welleck, Jiacheng Liu, Ronan Le Bras, Hannaneh Hajishirzi, Yejin Choi, and Kyunghyun Cho. Naturalproofs: Mathematical theorem proving in natural language. *arXiv preprint arXiv:2104.01112*, 2021.

แอปพลิเคชันเว็บสำหรับการตรวจสอบและประเมินแผลเบาหวาน ด้วยการเรียนรู้เชิงลึก

วรัท เมฆาสุวรรณดำรง[1,*] ธัชธรรม์ กษิดิ์เดชพงศ์[1,*] ปุณยวีร์ อริยกิตติวงศ์[1,*]
สิทธิโชค โสมอ่ำ[2,†] และ เจิมพิภัช เจริญปัญญาศักดิ์[3]

[1]โรงเรียนมหิดลวิทยานุสรณ์ 73170
[2]ภาควิชาคณิตศาสตร์และวิทยาการคำนวณ โรงเรียนมหิดลวิทยานุสรณ์ 73170
[3]สาขาวิชาชีวเวชศาสตร์และวิศวกรรมชีวการแพทย์ คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ 90110

บทคัดย่อ

โรคเบาหวานเป็นปัญหาสุขภาพที่ส่งผลกระทบต่อประชากรจำนวนมาก โดยหนึ่งในภาวะแทรกซ้อนที่สำคัญคือแผลเบาหวาน ซึ่งมีลักษณะหายช้าและอาจนำไปสู่การสูญเสียอวัยวะ หากไม่ได้รับการติดตามดูแลอย่างเหมาะสม การดูแลผู้ป่วยในพื้นที่ห่างไกลโดยอาสาสมัครสาธารณสุขประจำหมู่บ้าน (อสม.) ยังประสบปัญหาด้านภาระงานและข้อจำกัดทางความรู้เฉพาะทาง งานวิจัยนี้จึงมุ่งพัฒนาแอปพลิเคชันเว็บเพื่อช่วยในการ ตรวจสอบและประเมินแผลเบาหวานผ่านภาพถ่ายโดยใช้เทคโนโลยีการเรียนรู้เชิงลึก (Deep Learning) โดยมีวัตถุประสงค์หลักในการจำแนกภาพระหว่างแผลเบาหวานกับภาพที่ไม่ใช่แผลเบาหวาน รวมถึงการวัดขนาดของแผลเพื่อประเมินความรุนแรงอย่างแม่นยำ ชุดข้อมูลที่ใช้ในการพัฒนาโมเดลประกอบด้วยภาพถ่ายเท้าทั้งหมด 1,095 ภาพ โดยในขั้นตอนการจำแนกประเภทของแผล ได้เปรียบเทียบประสิทธิภาพของโมเดล Convolutional Neural Networks (CNN) ได้แก่ MobileNet, VGGNet, EfficientNet และ ResNet ผลการทดลองพบว่า MobileNet มีประสิทธิภาพสูงสุด ด้วยค่า f1-score เท่ากับ 0.99 ในการจำแนกประเภทของแผล

หลังจากนั้น แอปพลิเคชันจะวัดขนาดของแผลด้วยเทคนิค การแบ่งส่วนภาพ โดยประยุกต์ใช้ โมเดล U-Net ร่วมกับ MobileNet โดยผลการทดลองพบว่าได้ค่า Intersection over Union เท่ากับ 0.78 ซึ่งแสดงให้เห็นถึงความแม่นยำในการแยกบริเวณแผลออกจากพื้นหลัง แอปพลิเคชันที่พัฒนาขึ้นนี้จะช่วยให้อสม. สามารถติดตามอาการของผู้ป่วยได้อย่างมีประสิทธิภาพผ่านระบบออนไลน์ ซึ่งจะช่วยส่งเสริมการดูแลรักษาผู้ป่วยเบาหวานในชุมชนห่างไกลให้มีคุณภาพมากยิ่งขึ้น

*ผู้นำเสนอ (keenmeka@gmail.com)
†ผู้ประพันธ์บรรณกิจ (sittichoke.som@mwit.ac.th)

คำสำคัญ: แผลเบาหวาน, การเรียนรู้เชิงลึก, การจำแนกรูปภาพ, การแบ่งส่วนภาพ
2020 MSC: 68T07, 68U10, 68T45, 92C50

1 บทนำ

โรคเบาหวาน(Diabetes) เป็นหนึ่งในกลุ่มโรคไม่ติดต่อ หรือ Non-Communicable Diseases(NCDs) ที่พบได้บ่อยทั่วโลกเกิดจากระดับน้ำตาลในเลือดสูงกว่าปกติ จากการสำรวจพบว่าในประเทศไทยมีผู้เสียชีวิตจากโรคเบาหวาน 200 คนต่อวัน และผู้เสียชีวิตจากภาวะแทรกซ้อนของโรคเบาหวาน 27 คนต่อวัน [1] โดยหนึ่งในภาวะแทรกซ้อนที่พบได้บ่อยคือ แผลเบาหวาน(Diabetic Foot Ulcers: DFUs) เป็นแผลเรื้อรังที่มักจะเกิดบริเวณเท้าของผู้ป่วย เนื่องจากเท้าเป็นบริเวณที่เกิดอาการชาได้ง่ายทำให้มีโอกาสเป็นแผลโดยไม่รู้ตัวสูง โดยแผลเบาหวานมีสาเหตุหลักเกิดจากความผิดปกติของระบบไหลเวียนเลือดหรือเส้นประสาทบริเวณแผล ทำให้แผลเบาหวานจะเป็นแผลที่หายช้า และมีโอกาสที่จะติดเชื้อสูงกว่าแผลปกติ ซึ่งถ้าไม่ได้รับการรักษาอาจเกิดภาวะเนื้อเยื่อตาย แล้วนำไปสู่การถูกตัดขาโดยมีความเสี่ยงสูงถึงร้อยละ 20 ถ้าไม่ได้รับการรักษา และดูแลที่เหมาะสม [2] ในปัจจุบัน หน่วยงานที่ใกล้ชิด และคอยดูแลผู้ป่วยคือ อาสาสมัครสาธารณสุขประจำหมู่บ้าน หรือ อสม. ซึ่งสำคัญมากในกลุ่มผู้ป่วยชุมชนห่างไกล หรือผู้ป่วยติดเตียง โดย อสม. จะเป็นผู้มาเยี่ยมบ้าน ให้คำแนะนำ คัดกรอง ติดตาม และดูแลผู้ป่วยเบื้องต้น โดยทั่วไป อสม. 1 คนจะรับผิดชอบดูแลประชาชนประมาณ 10-20 หลังคาเรือน [3] จากการศึกษาพบว่า อสม. มักมีปัญหาเกี่ยวกับภาระงานที่รับผิดชอบสูง และข้อจำกัดด้านความรู้เฉพาะด้าน [4] ทำให้การประยุกต์ใช้เทคโนโลยีเข้ามาเป็นส่วนช่วยของ อสม. จะสามารถส่งผลดีต่อทั้งตัวผู้ป่วย และ อสม. เป็นอย่างมาก

ในปัจจุบันมีการประยุกต์เทคโนโลยีเข้ามามีส่วนช่วยในด้านการแพทย์เป็นอย่างแพร่หลาย รวมไปถึงการใช้เทคโนโลยีการเรียนรู้เชิงลึก (Deep learning) มาช่วยการวินิจฉัยโรค และช่วยการเก็บข้อมูล เพื่อให้การรักษามีประสิทธิภาพเพิ่มมากขึ้น โดยการเรียนรู้เชิงลึกคือวิธีการเรียนรู้แบบอัตโนมัติของคอมพิวเตอร์ด้วยการเลียนแบบการทำงานของระบบประสาทของมนุษย์ โดยนำระบบโครงข่ายประสาทมาซ้อนกันหลายชั้นเพื่อเรียนรู้ข้อมูลตัวอย่าง [5] โดยมีหลายงานวิจัยที่มีการประยุกต์ใช้เทคโนโลยีเพื่อช่วยแก้ปัญหาแผลเบาหวานต่าง ๆ เช่น งานวิจัยของ Thotad, P. N., Bharamagoudar(2023) ได้มีการเปรียบเทียบประสิทธิภาพการวินิจฉัยแผลของโมเดล CNN EfficientNet ต่าง ๆ ใช้ชุดข้อมูลภาพเท้าทั้งหมด 844 ภาพ แบ่งเป็นภาพเท้าปกติและภาพเท้าที่มีแผลเบาหวาน โดยได้ค่าสูงสุดที่ F1-score 0.99 งานวิจัยนี้ชี้ให้เห็นถึงคุณภาพของเทคโนโลยีการเรียนรู้เชิงลึกในการวินิจฉัยแผลได้ดีมาก [6] และงานวิจัยของ Rania et al. (2022) ได้มีการเปรียบเทียบประสิทธิภาพการแบ่งส่วนภาพของแผลด้วยโมเดล U-Net และโมเดลอื่น ๆ โดยใช้ข้อมูลภาพถ่ายจากโทรศัพท์ และกล้องถ่ายความร้อนภาพทั้งหมด 112 ภาพ โดยแบ่งเป็น 92 ภาพสำหรับการฝึกอบรมและ 22 ภาพสำหรับการทดสอบ ได้ค่า Intersection over Union 98% แสดงถึงความแม่นยำที่สูงแต่มีข้อจำกัดในด้านข้อมูลภาพที่ค่อนข้างน้อย [7]

คณะผู้วิจัยเล็งเห็นถึงปัญหาดังกล่าว และประโยชน์ในการใช้เทคโนโลยีการเรียนรู้เชิงลึกในเชิงการแพทย์ งานวิจัยนี้จึงพัฒนาแอปพลิเคชันเว็บออนไลน์ซึ่งสามารถเข้าถึงได้ง่ายแม้อยู่ในที่ห่างไกล ซึ่งช่วย อสม. ให้สามารถติดตามแผลของผู้ป่วยของตนได้อย่างมีประสิทธิภาพผ่านระบบออนไลน์ และยังช่วยกลุ่มผู้ป่วยในชุมชนห่างไกลสามารถเข้าถึงการดูแลได้อย่างมีคุณภาพมากขึ้น โดยแอปนี้จะใช้โมเดลการเรียนรู้เชิงลึกในการจำแนกรูปภาพ (Image classification) เพื่อที่จะสามารถวินิจฉัยแผลเบาหวานตั้งแต่ระยะแรก ๆ เพื่อลดความเสี่ยงการติดเชื้อของแผลที่จะนำไปสู่การตัดอวัยวะ และโมเดลที่ใช้เทคนิคการแบ่งส่วนภาพ (Image segmentation) เพื่อวัดขนาดแผลเพื่อที่จะช่วยให้ทั้ง อสม. และตัวผู้ป่วยสามารถติดตาม และประเมินแผลเบื้องต้นได้

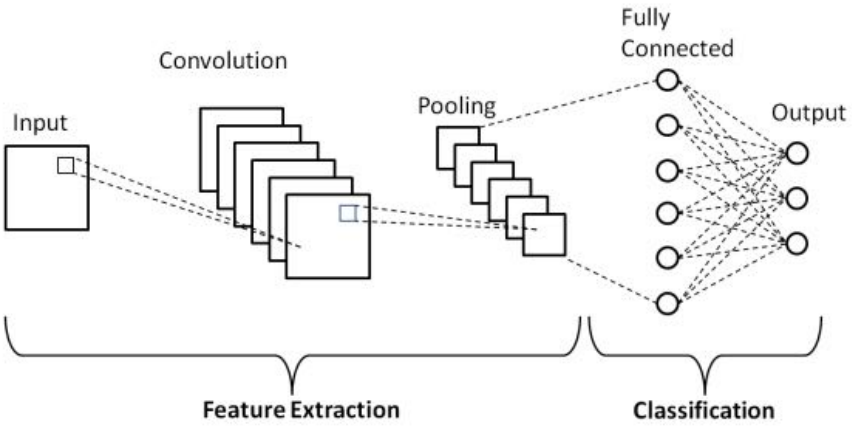
2 ความรู้พื้นฐาน

2.1 การเรียนรู้เชิงลึก

การเรียนรู้เชิงลึก (Deep Learning) เป็นสาขาหนึ่งของปัญญาประดิษฐ์ที่ใช้โครงข่ายประสาทเทียมหลายชั้นในการเรียนรู้และวิเคราะห์ข้อมูลที่ซับซ้อน [8] โดยสามารถนำไปใช้ในงานต่าง ๆ เช่น การรู้จำภาพ การประมวลผลภาษาธรรมชาติ และการวิเคราะห์ข้อมูลเสียง

2.1.1 โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional neural network)

โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Networks: CNNs) เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่ได้รับการออกแบบมาเพื่อประมวลผลข้อมูลที่มีโครงสร้างเป็นตาราง เช่น ภาพถ่ายหรือวิดีโอ โครงข่ายประสาทแบบคอนโวลูชันประกอบไปด้วยส่วนประกอบที่สำคัญ 2 ส่วน คือ Convolutional Layer และ Pooling Layer ซึ่ง Convolutional layer ทำหน้าที่สกัดคุณลักษณะจากข้อมูลนำเข้า (Input) โดยใช้การคำนวณแบบคอนโวลูชัน ซึ่งเป็นการคำนวณแบบ dot product ระหว่างพื้นที่ส่วนย่อยของรูปภาพกับเคอร์นอล (Kernel) [9] เพื่อดึงลักษณะเด่น (Feature Extraction) ออกมาจากรูปภาพ และ Pooling Layer ใช้ในการลดขนาดของข้อมูลที่ได้จาก Convolutional Layer โดยการเลือกค่าสูงสุด (Max Pooling) หรือค่าเฉลี่ย (Average Pooling) ในแต่ละพื้นที่ย่อย เพื่อลดจำนวนพารามิเตอร์และการคำนวณ ทำให้โมเดลมีประสิทธิภาพมากขึ้น



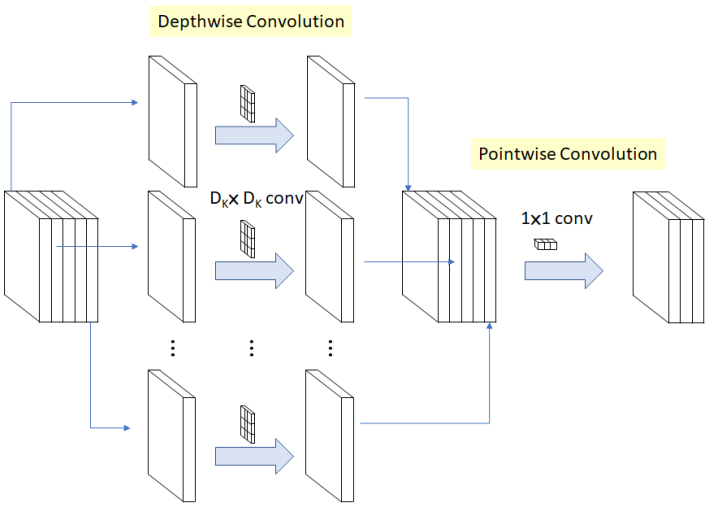
รูปที่ 1 โครงสร้างแบบจำลอง CNN

2.2 แบบจำลองสำหรับการจำแนกประเภท (Models for classification)

มีแบบจำลอง CNN หลายรุ่นที่ถูกพัฒนาขึ้นเพื่อเพิ่มประสิทธิภาพในการจำแนกภาพ โดยแต่ละรุ่นมีโครงสร้างและความสามารถที่แตกต่างกัน

2.2.1 MobileNet

MobileNet เป็นโครงข่าย CNN ที่ออกแบบมาให้มีขนาดเล็กและประหยัดพลังงาน เหมาะสำหรับการใช้งานบนอุปกรณ์พกพาและระบบที่มีทรัพยากรจำกัด จุดเด่นของ MobileNet คือการใช้ Depthwise Separable Convolutions ซึ่งช่วยลดจำนวนพารามิเตอร์และการคำนวณ ทำให้สามารถประมวลผลได้รวดเร็วโดยยังคงประสิทธิภาพในการจำแนกภาพได้ดี [10]



รูปที่ 2 โครงสร้างแบบจำลอง MobileNet

2.2.2 EfficientNet

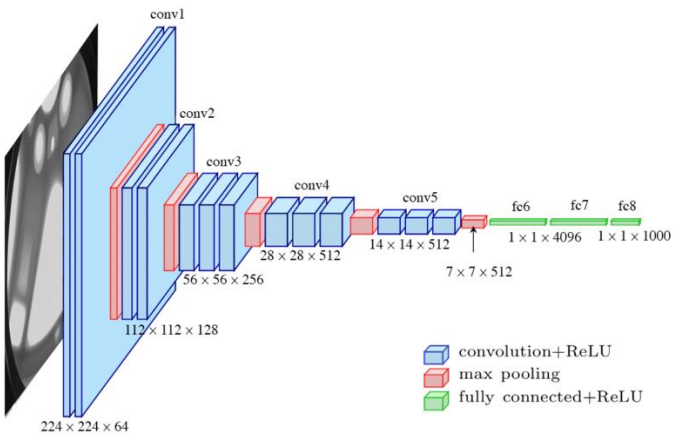
EfficientNet เป็นโครงข่าย CNN ที่พัฒนาขึ้นโดยใช้แนวคิดการปรับขนาดโครงข่ายอย่างมีประสิทธิภาพ โดยสามารถเพิ่มความลึก ความกว้าง และความละเอียดของโครงข่ายได้อย่างเหมาะสม จุดเด่นของ EfficientNet คือการใช้วิธี Compound Scaling ที่ช่วยให้โมเดลมีประสิทธิภาพสูงสุดในการจำแนกภาพ โดยใช้พารามิเตอร์น้อยกว่าโมเดลอื่นที่มีขนาดใกล้เคียงกัน [11]

2.2.3 ResNet

ResNet (Residual Networks) เป็นโครงข่าย CNN ที่มีความลึกมากและใช้การเชื่อมต่อแบบข้ามชั้น (skip connections) เพื่อแก้ปัญหาการลดทอนของการไล่ระดับ (vanishing gradients) ทำให้สามารถฝึกโครงข่ายที่ลึกมาก ๆ ได้อย่างมีประสิทธิภาพ [12]

2.2.4 VGG-Net

VGG-Net เป็นโครงข่าย CNN ที่มีโครงสร้างเรียบง่าย โดยใช้คอนโวลูชันขนาด 3x3 ซ้อนกันหลายชั้น จุดเด่นของ VGG-Net คือความสามารถในการสกัดคุณลักษณะได้อย่างละเอียด และความยืดหยุ่นในการปรับความลึกของโครงข่าย[13]



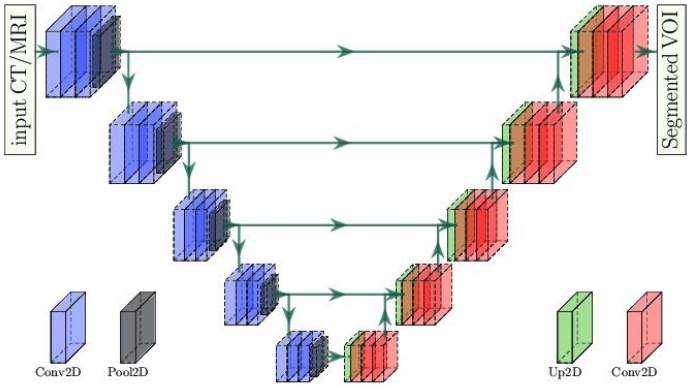
รูปที่ 3 โครงสร้างแบบจำลอง VGG-Net

2.3 แบบจำลองสำหรับการแบ่งส่วนภาพ (Models for image segmentation)

การแบ่งส่วนภาพ (Image Segmentation) เป็นกระบวนการที่แบ่งภาพออกเป็นส่วนย่อย ๆ เพื่อระบุและแยกแยะวัตถุหรือบริเวณที่สนใจ

2.3.1 U-Net

U-Net เป็นโครงข่าย CNN ที่ออกแบบมาสำหรับงานแบ่งส่วนภาพทางการแพทย์ มีโครงสร้างแบบตัว "U" โดยประกอบด้วยส่วนการย่อขนาด (contracting path) และส่วนการขยายขนาด (expanding path) ทำให้สามารถสกัดคุณลักษณะและสร้างแผนที่การแบ่งส่วนที่แม่นยำ [14]



รูปที่ 4 โครงสร้างแบบจำลอง U-Net

2.4 ตัวชี้วัดประสิทธิภาพ (Performance Metrics)

		Prediction	
		Negative (0)	Positive (1)
Actual	Negative (0)	True Negative (TN)	False Positive (FP)
	Positive (1)	False Negative (FN)	True Positive (TP)

รูปที่ 5 เมทริกซ์ความสับสน

2.4.1 เมทริกซ์ความสับสน (Confusion Matrix)

ในงานวิจัยนี้ได้ใช้ตัวชี้วัด 4 ชนิดในการประเมินประสิทธิภาพแบบจำลองการจำแนกประเภทได้แก่ ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วงดุล (F1-Score) ซึ่งสามารถคำนวณได้ตามสมการที่ (2.1) ถึง (2.4) [15]

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (2.1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2.2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (2.3)$$

$$\text{F1 score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision+Recall}} \quad (2.4)$$

2.4.2 Intersection over Union (IoU)

เป็นมาตรวัดที่ใช้ในการประเมินความแม่นยำของการแบ่งส่วนภาพ โดยคำนวณจากสัดส่วนของพื้นที่ที่ทับซ้อนกันระหว่างผลลัพธ์ที่ได้กับค่าจริง เทียบกับพื้นที่รวมทั้งหมด ค่ายิ่งใกล้ 1 แสดงถึงความแม่นยำที่สูงขึ้น [16]

$$\text{IoU score} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (2.5)$$

3 ระเบียบวิธีวิจัย (Methodology)

ในการวิจัยนี้ ผู้จัดทำได้แบ่งขั้นตอนการดำเนินการเป็น 3 ส่วนคือ 1.การจำแนกประเภทรูปภาพ 2.การแบ่งส่วนภาพ 3.การประยุกต์ใช้งาน ดังนี้

3.1 การจำแนกประเภทรูปภาพ (Image Classification)

3.1.1 การเตรียมชุดข้อมูล

ใช้ชุดข้อมูลออนไลน์จากเว็บไซต์ Kaggle [17] ประกอบด้วยภาพถ่ายเท้าปกติและภาพแผลทั่วไป จำนวน 562 ภาพ และภาพถ่ายแผลเบาหวานจำนวน 533 ภาพ จากนั้นนำชุดข้อมูลทั้งสองมาทำการแบ่งเป็น 3 ชุด ได้แก่ ชุดฝึกอบรม (Training set) 80%, ชุดทดสอบ (Test set) 10% และชุดตรวจสอบความถูกต้อง (Validation set) 10% เพื่อใช้ในการฝึกสอนและประเมินผลของโมเดล Machine Learning ต่อไป

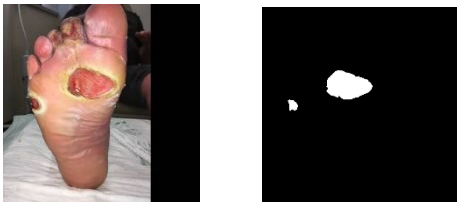
3.1.2 ขั้นตอนการพัฒนาแบบจำลองสำหรับการจำแนกประเภท

ใช้การเรียนรู้แบบจำแนกประเภทภาพ (Binary Image Classification) โดยแบ่งข้อมูลออกเป็น 2 กลุ่ม คือ กลุ่มที่มีแผลเบาหวาน และกลุ่มที่ไม่มีแผลเบาหวาน จากนั้นนำข้อมูลเข้าสู่แบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks - CNN) ที่ประกอบด้วยหลายสถาปัตยกรรม ได้แก่ MobileNetV2, EfficientNetB5, VGGs-Net, ResNet50 และ ResNet101 เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของภาพของแต่ละแบบจำลอง

3.2 การแบ่งส่วนภาพ (Image Segmentation)

3.2.1 การเตรียมชุดข้อมูล

ใช้ชุดข้อมูลออนไลน์ [18] โดยประกอบด้วยภาพแผลเบาหวาน 1010 รูป และภาพขอบเขตของแผลในรูปแบบมาสก์ (Mask) 1010 รูป ดังตัวอย่างในรูปที่ 6 โดยข้อมูลแบ่งเป็น 3 ชุดเช่นเดียวกันกับการจำแนกภาพ คือ Training 80%, Validation 10% และ Test 10%



รูปที่ 6 ตัวอย่างชุดข้อมูลของแผลและมาสก์

3.2.2 ขั้นตอนพัฒนาแบบจำลองสำหรับการแบ่งส่วนภาพ

การแบ่งส่วนภาพใช้เทคนิคการแยกบริเวณแผลออกจากภาพต้นฉบับโดยใช้แบบจำลอง U-Net และ U-Net ที่ประยุกต์ร่วมกับโมเดล MobileNet เป็น encoder เพื่อเพิ่มประสิทธิภาพในการประมวลผลและลดความซับซ้อนของโมเดล ภายหลังจากการ segmentation แล้ว จะมีการวัดพื้นที่แผลในหน่วยเซนติเมตรโดยอ้างอิงจากขนาดเหรียญมาตรฐาน 1 เซนติเมตรที่ปรากฏในภาพ พร้อมทั้งแปลงพื้นที่มาสก์เป็นภาพ HSV เพื่อใช้วิเคราะห์สีของแผล โดยแบ่งออกเป็นสีหลัก ได้แก่ แดง เหลือง ขาว และดำ

3.3 การประยุกต์ใช้งาน (Application)

เพื่อให้สามารถใช้งานระบบได้อย่างสะดวกผ่านเว็บไซต์หรือแอปพลิเคชันบนอุปกรณ์พกพา ได้มีการพัฒนาระบบโดยใช้เทคโนโลยีดังนี้:

- Python: ใช้สำหรับพัฒนาโมเดล AI และการประมวลผลภาพ
- Streamlit: ใช้สร้างหน้าเว็บแอปพลิเคชันที่ใช้งานง่ายและรวดเร็ว
- GitHub: ใช้สำหรับการเก็บและแบ่งปันซอร์สโค้ด รวมถึงการทำงานร่วมกันแบบทีม

4 ผลการศึกษา

4.1 การจำแนกประเภทของแผลเบาหวาน

ในส่วนของการจำแนกประเภทของแผล คณะผู้จัดทำได้เปรียบเทียบประสิทธิภาพของแบบจำลอง Convolutional Neural Networks (CNN) จำนวน 5 แบบ ได้แก่ MobileNet, VGGNet, EfficientNet, ResNet50 และ ResNet100 ผลการประเมินโดยใช้ค่า f1-score บนชุดข้อมูลทดสอบมีรายละเอียดดังนี้:

ตารางที่ 1 ตารางเปรียบเทียบประสิทธิภาพของ pre-train model

Comparison of pre-trained model performance

Model	Accuracy	Precision	Recall	F1-score
VGG-Net16	0.92	0.92	0.92	0.92
MobileNetV2	0.99	0.99	0.99	0.99
EfficientNetB5	0.98	0.98	0.98	0.98
ResNet50	0.93	0.93	0.93	0.93
ResNet100	0.89	0.89	0.89	0.89

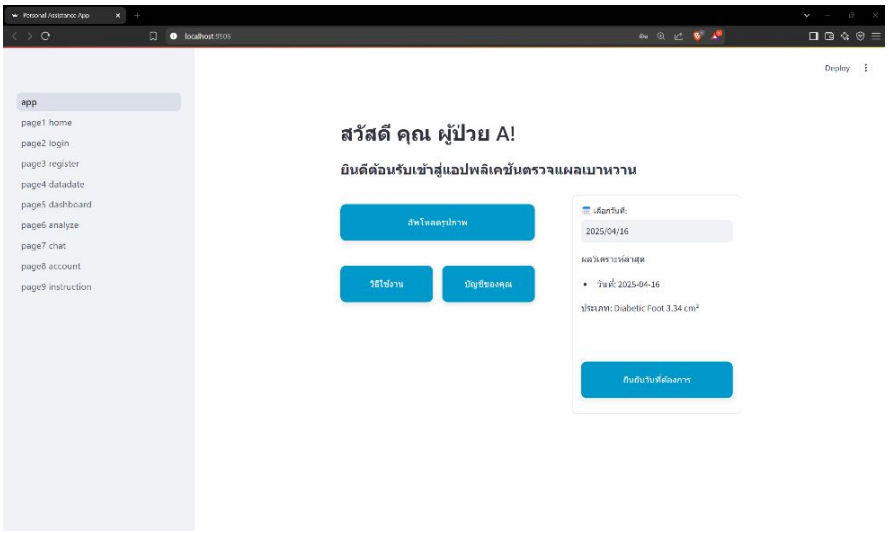
4.2 การแบ่งส่วนภาพแผลเบาหวาน

สำหรับการวัดขนาดแผล คณะผู้จัดทำได้ใช้เทคนิค Image Segmentation โดยประยุกต์ใช้แบบจำลอง U-Net ร่วมกับ MobileNet เป็นตัว Encoder เพื่อลดจำนวนพารามิเตอร์และเพิ่มความเร็วในการประมวลผล

ผลการวัดประสิทธิภาพด้วยค่า Intersection over Union (IoU) บนชุดข้อมูลทดสอบ พบว่าแบบจำลองสามารถระบุขอบเขตแผลได้อย่างแม่นยำ โดยมีค่าเฉลี่ย IoU อยู่ที่ 0.78

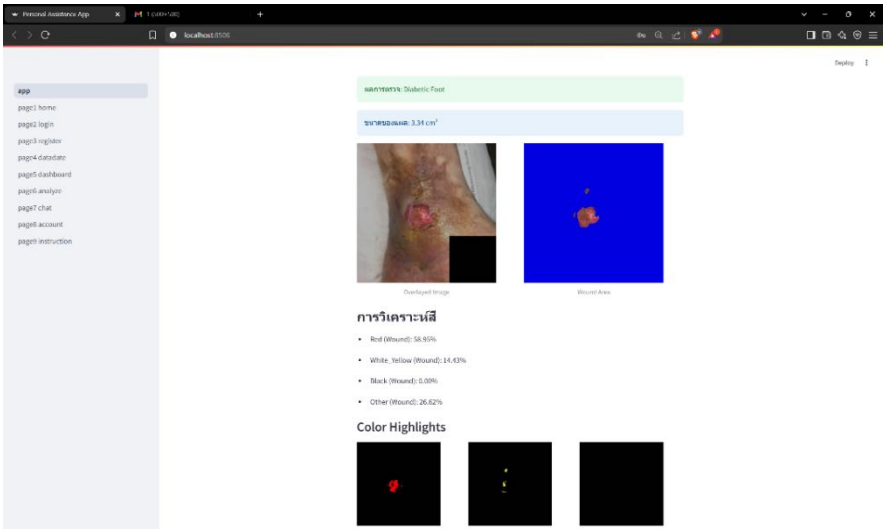
4.3 การทำงานของแอปพลิเคชัน

คณะผู้จัดทำได้พัฒนาแอปพลิเคชันเว็บด้วยภาษา Python ร่วมกับเฟรมเวิร์ก Streamlit เพื่อให้สามารถใช้งานได้ผ่านเบราว์เซอร์ โดยรูปที่ 6 แสดงให้เห็นหน้า dashboard ของแอปพลิเคชัน ซึ่งผู้ใช้งานสามารถอัปโหลดภาพถ่ายของเท้าของตนเอง ศึกษาวิธีการใช้งาน และดูอาการในวันที่ผ่านๆมาได้ โดยเมื่อผู้ใช้ได้อัปโหลดรูปเท้าลงในระบบ แอปพลิเคชันจะประมวลผลโดยใช้แบบจำลองที่ได้จากการทดลองมาจำแนกภาพว่าเป็นแผลเบาหวานหรือไม่ แบ่งส่วนแผลออกจากพื้นหลังภาพและวัดอัตราส่วนของสีในภาพ โดยแสดงผลให้ผู้ใช้ตามในรูปที่ 7 จากนั้นระบบจะ



รูปที่ 7 ภาพบันทึกหน้าจอของ Dashboard ของผู้ป่วย

บันทึกข้อมูลของแผลสำหรับการติดตามผลต่อไป



รูปที่ 8 ภาพบันทึกหน้าจอการแสดงผลของโปรแกรม

5 สรุปและอภิปรายผล

จากการศึกษาและทดลองพัฒนาแอปพลิเคชันสำหรับการตรวจสอบและประเมินแผลเบาหวานด้วยเทคโนโลยีการเรียนรู้เชิงลึก พบว่าแนวทางนี้สามารถใช้งานได้และมีความแม่นยำสูง แบบจำลอง MobileNet สามารถจำแนกประเภทของแผลเบาหวานได้อย่างแม่นยำ โดยมีค่า f1-score เท่ากับ 0.99 ซึ่งสูงมากเมื่อเทียบกับแบบจำลองอื่น ๆ ที่ใช้ในการทดลอง โดยอาจเกิดจากลักษณะของชุดข้อมูลที่ไม่ซับซ้อนและมีขนาดไม่ใหญ่มาก ทำให้แบบจำลองที่มีโครงสร้างเบาอย่าง MobileNet ไม่เกิดการ overfitting และสามารถเรียนรู้ลักษณะพื้นผิวของแผลได้อย่างมีประสิทธิภาพ ขณะที่แบบจำลองที่ซับซ้อนกว่าอาจเรียนรู้มากเกินจำเป็นและประสิทธิภาพลดลงในบริบทนี้ นอกจากนี้แบบจำลอง U-Net ร่วมกับ MobileNet ยังสามารถแยกขอบเขตแผลจากภาพได้แม่นยำด้วยค่า IoU เท่ากับ 0.78

6 ข้อเสนอแนะ

ควรพัฒนาระบบต่าง ๆ ภายในแอปพลิเคชันเพื่อให้ผู้ใช้ใช้แอปพลิเคชันได้ง่ายขึ้น เช่น ระบบบัญชีแยกสำหรับผู้ป่วย และอสม. เพื่อให้ง่ายต่อการแยกใช้งานอย่างชัดเจน ระบบที่บอกให้รู้ถึงข้อมูลของคนที่รับผิดชอบตนในฝั่งของบัญชีผู้ป่วย และในฝั่งของบัญชีอสม. บอกจำนวน และข้อมูลผู้ป่วยที่อยู่ภายใต้การดูแลของตน และระบบการแนะนำวิธีดูแลแผลเบื้องต้นสำหรับผู้ป่วยแต่ละคนเพื่อให้ลดโอกาศที่แผลจะรุกรามและติดเชื้อ

ควรออกแบบหน้าจอการใช้งานแอปพลิเคชันให้มีความเข้าใจง่าย และใช้งานง่ายสำหรับทั้งผู้ป่วยและอสม. จากนั้นทำแบบประเมินการใช้แอปพลิเคชันทั้งในเรื่องความเข้าใจในการใช้งาน ความสะดวก ความถูกต้อง และความพึงพอใจในภาพรวมของแอปพลิเคชัน เพื่อให้ผู้ใช้งานได้มีโอกาศแสดงความคิดเห็นต่อแอปพลิเคชันนี้เพื่อพัฒนาต่อไป

ควรนำแอปพลิเคชันนี้ไปทดลองใช้จริงในกลุ่มอสม. และผู้ป่วย เพื่อให้สามารถรู้ถึงขอบเขตการใช้งานของแอปพลิเคชันรวมถึงแสดงถึงความสามารถต่อสถานการณ์จริงในการใช้งาน

เอกสารอ้างอิง

[1] Wattanapat Hospital Ao Nang. (n.d.). Diabetes mellitus: A disease close to home., from https://aonanghospital.com/diabetes_mellitus

[2] Katherine McDermott, Michael Fang, Andrew J.M. Boulton, Elizabeth Selvin, Caitlin W. Hicks; Etiology, Epidemiology, and Disparities in the Burden of Diabetic Foot Ulcers. Diabetes Care 2 January 2023; 46 (1): 209–221. https://doi.org/10.2337/dci22-0043

[3] Public Relations Group, Department of Health Service Support, Ministry of Public Health. (2023, October 12). 3 หมอ 'คนไทยทุกครอบครัวมีหมอประจำตัว 3 คน'. Public Relations HSS. https://prgroup.hss.moph.go.th/medias/article/1175-3-หมอ-“คนไทยทุกครอบครัวมีหมอประจำตัว-3-คน”

[4] Sittapanich, S. (2022). The effect of diabetes patient caring training program in village health volunteers in Yo Primary Care Unit. Yasothon Medical Journal. https://he04.tci-thaijo.org/index.php/hciyasohos/article/view/312/137

[5] SAS Institute Inc. (n.d.). What is deep learning?. Retrieved May 15, 2025, from https://www.sas.com/en_us/insights/analytics/deep-learning.html

[6] Thotad, P. N., Bharamagoudar, G. R., & Anami, B. S. (2023). Diabetic foot ulcer detection using deep learning approaches. Sensors International, 4, 100210. https://doi.org/10.1016/j.sintl.2022.100210

[7] Niri, R., Lucas, Y., Treuillet, S., & Douzi, H. (2020). Smartphone-based thermal imaging system for diabetic foot ulcer assessment. Journées d’Étude sur la Télésanté, Sorbonne Universités, Paris, France.

[8] ABB. (2020). Deep learning. ABB News. https://new.abb.com/news/detail/58004/deep-learning

[9] Adrian Rosebrock. (2017, March 20). *ImageNet: VGGNet, ResNet, Inception, and Xception with Keras*. PyImageSearch. https://pyimagesearch.com/2017/03/20/imagenet-vggnet-resnet-inception-xception-keras/

[10] Zink, M. (2023, October 18). What is MobileNet? Built In. https://builtin.com/machine-learning/mobilenet

[11] efficient Viso Suite. (2022, September 12). EfficientNet: Scaling CNNs for AI vision. Viso. https://viso.ai/deep-learning/efficientnet/#:~:text=EfficientNet%20is%20a%20Convolutional%20Neural,high%20accuracy%20with%20computational%20efficiency.

[12] GeeksforGeeks. (2022, November 17). Residual Networks (ResNet) – Deep learning. https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/

[13] Bangar, S. (2022, June 28). VGG-Net architecture explained. Medium. https://medium.com/@siddheshb008/vgg-net-architecture-explained-71179310050f

[14] Tomar, N. (2021, January 19). What is UNET? Analytics Vidhya.
https://medium.com/analytics-vidhya/what-is-unet-157314c87634
[15] GeeksforGeeks. (2025, May 16). Understanding the confusion matrix in machine learning.
https://www.geeksforgeeks.org/confusion-matrix-machine-learning/
[16] Rosebrock, A. (2016, November 7). Intersection over Union (IoU) for object detection.
PyImageSearch. https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/
[17] Laith. (2021, August 11). Diabetic foot ulcer (DFU). Kaggle.
https://www.kaggle.com/datasets/laithjj/diabetic-foot-ulcer-dfu
[18] University of Wisconsin-Milwaukee. (2019). Foot ulcer segmentation challenge [Source code]. GitHub. https://github.com/uwm-bigdata/wound-segmentation/tree/master/data/Foot%20Ulcer%20Segmentation%20Challenge

แอปพลิเคชันสำหรับการให้ข้อมูลสมุนไพรและระบุสมุนไพรจากรูปภาพใบสมุนไพรโดยใช้เทคนิคการเรียนรู้ของเครื่อง

ตฤน จันทร์อร่าม[1,*] ณัฐธีร์ ศรุติเจริญวงศ์[1,*] วิชานันท์ ชานันโท[1,†]
สิทธิโชค โสมอ่ำ[1] วรพันธ์ หอมสมบูรณ์[1] และ เจิมพิภัช เจริญปัญญาศักดิ์[2]

[1] โรงเรียนมหิดลวิทยานุสรณ์ ตำบลศาลายา อำเภอพุทธมณฑล จังหวัดนครปฐม 73170

[2] สาขาวิชาเวชศาสตร์และวิศวกรรมชีวการแพทย์หน่วยงาน คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ 90110

บทคัดย่อ

หลังการแพร่ระบาดของไวรัสโควิด-19 ความนิยมในการใช้สมุนไพรไทยร่วมกับยาแผนปัจจุบันเพิ่มสูงขึ้น อย่างไรก็ตาม ปัจจุบันยังขาดเครื่องมือที่ให้ข้อมูลเกี่ยวกับสมุนไพรและอันตรกิริยากับยาแผนปัจจุบันอย่างครบถ้วน งานวิจัยนี้จึงมีเป้าหมายเพื่อพัฒนาแอปพลิเคชันบนมือถือ ซึ่งสามารถให้ข้อมูลเกี่ยวกับสมุนไพรไทย อันตรกิริยากับยาแผนปัจจุบัน และสามารถระบุชนิดสมุนไพรจากภาพถ่ายใบได้ โดยกระบวนการวิจัยเริ่มจากการรวบรวมข้อมูลสมุนไพรกว่า 250 ชนิด ซึ่งในจำนวนนี้ 61 ชนิดมีข้อมูลเกี่ยวกับอันตรกิริยากับยาแผนปัจจุบัน และมีสมุนไพรจำนวน 69 ชนิดที่มีลักษณะคล้ายคลึงกันซึ่งยากต่อการจำแนก จึงถูกนำมาใช้ในการพัฒนาแบบจำลองการเรียนรู้ของเครื่อง ได้แก่ Convolutional Neural Network (CNN), Support Vector Machine (SVM) และ Random Forest ผลการทดลองพบว่า SVM ให้ค่าความแม่นยำสูงสุดที่ 97% ในขณะที่ CNN แม่นยำ 87% แต่มีความเหมาะสมมากกว่าในการใช้งานจริงบนแอปพลิเคชัน เนื่องจากสอดคล้องกับข้อจำกัดด้านประสิทธิภาพของภาษา Kotlin ที่ใช้ในการพัฒนาแอป จึงเลือกใช้โมเดล CNN และถูกออกแบบให้มีฟังก์ชันหลัก ได้แก่ การค้นหาข้อมูลสมุนไพร การตรวจสอบอันตรกิริยา และการระบุชนิดของสมุนไพรจากภาพถ่ายของใบ โดยผู้พัฒนามีแผนจะทดสอบกับกลุ่มผู้ใช้งานจริง เพื่อปรับปรุงและพัฒนาแอปพลิเคชันให้สามารถตอบสนองต่อความต้องการของผู้ใช้ได้ดียิ่งขึ้นในอนาคต

คำสำคัญ: แอปพลิเคชันบนมือถือ, สมุนไพรไทย, อันตรกิริยาสมุนไพรกับยาแผนปัจจุบัน

2020 MSC: 68T05, 68T2

*ผู้นำเสนอ (trin.cha_g33@mwit.ac.th, natee.sar_g33@mwit.ac.th)

†ผู้นำเสนอและผู้ประพันธ์บรรณกิจ (wichanan.cha_g33@mwit.ac.th)

1 บทนำ

ปัจจุบันเทคโนโลยีทางด้านการแพทย์ได้มีการพัฒนาไปมาก โดยเฉพาะทางด้านเภสัชกรรมที่สามารถผลิตยาได้หลากหลายประเภท เช่น ยาที่สามารถยับยั้งหรือฆ่าเชื้อโรคโดยตรง ยาที่ออกฤทธิ์บรรเทาอาการของโรค เป็นต้น แต่อย่างไรก็ดีการแพทย์ทางเลือกอย่างการรักษาโดยอิงความรู้ของตำหรับยาโบราณและสมุนไพรไทยก็ยังมีความสำคัญ ตัวอย่างเช่นในช่วงสถานการณ์การแพร่ระบาดของเชื้อไวรัสโควิด-19 ในประเทศไทยได้มีมีการนำสารสกัดจากสมุนไพรไทยอย่างฟ้าทะลายโจร (Green Chiretta) มาใช้รักษาร่วมกับยาแผนปัจจุบัน เนื่องจากสารสกัดจากฟ้าทะลายโจรมีสรรพคุณช่วยบรรเทาอาการระคายคอ ลดไข้ รวมถึงมีฤทธิ์ยับยั้งการเพิ่มจำนวนของไวรัสโควิด-19 สังเกตว่าการรักษาโดยการผสมผสานระหว่างการใช้ยาแผนปัจจุบันกับศาสตร์การใช้สมุนไพรไทยเป็นการผสมผสานข้อดีของยาแผนปัจจุบันจะออกฤทธิ์รวดเร็วและช่วยบรรเทาอาการป่วยได้ในเวลาอันสั้น เข้ากับข้อดีของยาสมุนไพรที่จะออกฤทธิ์อ่อน ๆ สามารถบรรเทาอาการเบื้องต้นได้ดี นอกจากนี้ยังเป็นการลดปริมาณการใช้การยาแผนปัจจุบัน ซึ่งเป็นสารเคมีสังเคราะห์หากใช้ในปริมาณที่มากเกินพอดีอาจเกิดผลข้างเคียงที่รุนแรง (Bernard Y-H. Thong et al. 2010 [19]) ทั้งยังส่งผลเสียต่ออวัยวะสำคัญ เช่น ตับและไต ซึ่งสามารถนำไปสู่โรคเรื้อรังที่ยากจะรักษา (Thomas YK Chan et al. 2001 [3])

แนวคิดการใช้สมุนไพรไทยร่วมกับการใช้ยาแผนปัจจุบันจึงนับเป็นแนวทางการรักษาที่น่าสนใจ ทั้งมีแนวโน้มที่จะช่วยแก้ไขปัญหาสาธารณสุขหลาย ๆ ประการ อาทิเช่น ปัญหาจำนวนคนไข้ล้นโรงพยาบาล ปัญหาค่ายาแผนปัจจุบันในการรักษาโรคเรื้อรัง เป็นต้น แต่ในสถานการณ์จริงนั้นกลับพบว่ายังไม่สามารถทำได้จริง สาเหตุหลักประการหนึ่ง คือ การขาดความรู้และความเข้าใจที่ถูกต้องในการใช้สมุนไพรไทย โดยการนำสมุนไพรมาใช้ร่วมกับยาแผนปัจจุบันนั้นจำเป็นต้องใช้อย่างเหมาะสม เนื่องจากสารเคมีในยาและสมุนไพรสามารถทำอันตรกิริยาต่อกัน อาจก่อให้เกิดสารใหม่ที่เป็นอันตรายต่อร่างกายโดยตรง เกิดเป็นผลเสียในการรักษาโรค (Thomas YK Chan et al. 2001 [3]) นอกจากนี้ความรู้เกี่ยวกับการสมุนไพรไทยเป็นภูมิปัญญาไทยที่ควรค่าแก่การอนุรักษ์และเผยแพร่และปัจจุบันยังไม่มีแอปพลิเคชันใดที่ให้บริการอย่างครอบคลุมทั้งในด้านการจำแนกชนิดของสมุนไพรและการให้ข้อมูลเกี่ยวกับอันตรกิริยาระหว่างสมุนไพรกับยาอย่างครบถ้วน ในขณะเดียวกัน แอปพลิเคชันที่มีอยู่ส่วนใหญ่ยังขาดการพัฒนาและการอัปเดตมาเป็นเวลานาน ส่งผลให้เกิดปัญหาเกี่ยวกับความเสถียรในการใช้งาน รวมถึงการแสดงข้อมูลที่ล้าสมัยหรือคลาดเคลื่อนซึ่งยังไม่ได้รับการแก้ไขอย่างเหมาะสม ทางผู้จัดทำจะแก้ปัญหาเกี่ยวกับการใช้ยาสมุนไพรร่วมกับยาแผนปัจจุบัน รวมถึงการสร้างฐานข้อมูลที่ทันสมัยและเข้าถึงได้ง่ายเพื่อที่จะให้ความรู้และระบุสมุนไพรไทยพื้นฐาน ดังนั้นทางผู้จัดจึงพัฒนาต้นแบบแอปพลิเคชันที่สามารถระบุสมุนไพรไทยจากรูปภาพใบสมุนไพร พร้อมให้ข้อมูลพื้นฐานเกี่ยวกับสมุนไพร รวมถึงข้อมูลที่มีลักษณะเฉพาะอย่างอันตรกิริยาของสมุนไพรไทยกับยาแผนปัจจุบัน

จากข้อจำกัดด้านการประมวลผลของระบบคอมพิวเตอร์และการรวบรวมข้อมูลสมุนไพรที่มีความน่าเชื่อถือในระยะแรกของการพัฒนาแอปพลิเคชันต้นแบบ ผู้จัดทำจะดำเนินการคัดเลือกสมุนไพรที่มีสรรพคุณทางยาซึ่งเป็นที่นิยมใช้ร่วมกับยาแผนปัจจุบัน จำนวน 69 ชนิด โดยอ้างอิงจากแหล่งข้อมูลทางวิชาการ อาทิเช่น หนังสือ ๑๐๘

สมุนไพรไทยใช้เป็น หายป่วย ตำรับยาอายุวัฒนะ กัน-แก้ สารพัดโรคจากครัวเรือน รวมถึงคำแนะนำจากผู้เชี่ยวชาญ จากนั้นจะสร้างแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning Models) สำหรับการระบุสมุนไพรจากรูปภาพใบสมุนไพร ซึ่งประกอบด้วยแบบจำลองการเรียนรู้ของเครื่องที่ใช้การTransfer Learning ร่วมกับ Convolutional Neural Network (CNN), Support Vector Machine (SVM) และ Random Forest จากนั้นจะเปรียบเทียบประสิทธิภาพเชิงทฤษฎีของแบบจำลองการเรียนรู้ของเครื่องและระบุแบบจำลองที่มีความแม่นยำสูงสุดแล้ว ขั้นตอนถัดไป คือ การปรับแบบจำลองดังกล่าวให้สามารถทำงานร่วมกับแอปพลิเคชันบนโทรศัพท์มือถือได้อย่างเหมาะสมและมีประสิทธิภาพ

ในส่วนถัดไปผู้พัฒนาจะสร้างฐานข้อมูลสมุนไพรประกอบด้วยข้อมูลพื้นฐาน คือ ชื่อสมุนไพร (ชื่อท้องถิ่นและชื่อวิทยาศาสตร์) สรรพคุณและภาพสมุนไพร และข้อมูลอันตรกิริยาของสมุนไพรและยาแผนปัจจุบัน

ส่วนสุดท้าย คือ การพัฒนาต้นแบบแอปพลิเคชันบนโทรศัพท์มือถือ ซึ่งจะมีฟังก์ชันหลัก 2 ฟังก์ชัน คือ ฟังก์ชันที่ประมวลผลรูปถ่ายใบสมุนไพรแล้วระบุสมุนไพร และฟังก์ชันค้นหาข้อมูลสมุนไพร โดยผู้จัดทำมีจุดมุ่งหมายหลักของการศึกษา คือ การหาแบบจำลองการเรียนรู้ที่เหมาะสมที่สามารถระบุสมุนไพรไทยจากใบสมุนไพรได้ถูกต้องและเหมาะสมกับระบบการประมวลโทรศัพท์มือถือ และการสร้างต้นแบบแอปพลิเคชันที่สามารถพัฒนาเป็นแอปพลิเคชันที่ใช้ได้จริง เหมาะสมสำหรับกลุ่มบุคคลทั่วไป

2 ความรู้พื้นฐาน

ในหัวข้อนี้ ผู้จัดทำจะขอเริ่มต้นด้วยการกล่าวถึงแบบจำลองการเรียนรู้ของเครื่องและการวัดประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง (สามารถศึกษาฉบับเต็มได้จาก [12]) จากนั้นจึงจะขอสรุปความรู้พื้นฐานเกี่ยวกับการการพัฒนาแอปพลิเคชั่นบนมือถือด้วยภาษา Kotlin (ผู้สนใจสามารถศึกษาเพิ่มเติมได้จาก [6])

2.1 แนวคิดพื้นฐานของการเรียนรู้ของเครื่อง (Fundamentals of Machine Learning)

การเรียนรู้ของเครื่อง (Machine Learning: ML) เป็นแขนงหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence: AI) ที่มุ่งเน้นการสร้างแบบจำลองทางคณิตศาสตร์เพื่อให้คอมพิวเตอร์สามารถเรียนรู้จากข้อมูล และสามารถทำนายหรือวิเคราะห์ข้อมูลใหม่ได้โดยไม่ต้องโปรแกรมล่วงหน้าทีละขั้นตอน

2.1.1 แนวคิดการเรียนรู้จากข้อมูล (Learning from Data)

หลักขั้นแรกของการพัฒนาแบบจำลองการเรียนรู้ของเครื่อง คือ การฝึกฝนแบบจำลอง (หรือโมเดล) ด้วยข้อมูลตัวอย่าง (training data) ซึ่งประกอบด้วย features (คุณลักษณะ) และ labels (คำตอบหรือคลาส) โดยแบบจำลองจะพยายามหา "ความสัมพันธ์" ระหว่างคุณลักษณะของข้อมูลและผลลัพธ์ เพื่อให้สามารถนำข้อสรุปที่ได้มาใช้ทำนายผลของข้อมูลใหม่ได้ โดยกระบวนการฝึกแบบจำลองการเรียนรู้ของเครื่องสามารถแบ่งได้เป็น 3 ประเภทหลัก คือ

- การเรียนรู้แบบมีผู้สอน (Supervised Learning) : การฝึกแบบจำลองจากชุดข้อมูลที่มี label ตัวอย่างเช่น การจำแนกภาพสมุนไพรว่าเป็นสมุนไพรชนิดใด โดยชุดข้อมูลตัวอย่างจะมีการระบุชนิดที่ถูกต้อง (label) ไว้แล้ว

- การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) : การฝึกแบบจำลองค้นหารูปแบบจากข้อมูลที่ไม่มี label ตัวอย่างเช่น การจัดกลุ่มสมุนไพรตามลักษณะใบ โดยเป้าหมายของการฝึกแบบจำลองแบบนี้ คือ การทำให้แบบจำลองสามารถค้นพบโครงสร้าง แพทเทิร์น หรือความสัมพันธ์ภายในข้อมูลที่ไม่มีการระบุคำตอบไว้

- การเรียนรู้แบบเสริมแรง (Reinforcement Learning) : แบบจำลองจะเรียนรู้ผ่านการลองผิดลองถูก (trial and error) โดยผ่านการรับรางวัลเมื่อทำได้ถูกต้องหรือได้รับบทลงโทษเมื่อผิดพลาด เช่น การฝึกฝนหุ่นยนต์ให้เรียนรู้วิธีเดิน

สำหรับปัญหาการระบุสมุนไพรจากรูปภาพใบสมุนไพรนั้น ผู้จัดทำวางแผนที่จะฝึกแบบจำลองด้วยข้อมูลตัวอย่าง คือ รูปภาพใบสมุนไพรที่มีการระบุชนิดสมุนไพรที่ถูกต้องไว้แล้ว จึงจัดอยู่ในกลุ่มการเรียนรู้แบบมีผู้สอน

2.1.2 การจำแนกประเภท (Classification)

การจำแนกประเภท คือ กระบวนการที่โมเดลทำการคาดการณ์ว่าข้อมูลที่รับมา (input) หนึ่ง ๆ ควรจะถูกจัดอยู่ในกลุ่มใดจากกลุ่มทั้งหมดที่กำหนดไว้ล่วงหน้า ตัวอย่างเช่น ระบบที่จำแนกว่ารูปภาพที่ได้รับมาเป็น “ภาพสุนัข” หรือ “ภาพแมว” เป็นต้น

การจำแนกประเภทจะมีขั้นตอนสำคัญ ดังนี้

- การเตรียมข้อมูล (data preparation) : การทำความสะอาดข้อมูล (data cleaning) และการจัดการข้อมูลให้อยู่ในรูปแบบที่แบบจำลองสามารถนำไปเรียนรู้ได้
- การเลือกโมเดล (model selection) : เลือกอัลกอริทึมให้เหมาะกับลักษณะข้อมูล
- การเลือกคุณสมบัติ (feature selection) : การเลือกคุณลักษณะที่สำคัญต่อการจำแนก ซึ่งแบบจำลองการเรียนรู้ของเครื่องแต่ละชนิดจะมีชุดของคุณสมบัติ (set of features) ที่สามารถปรับค่าได้ต่างกัน
- การฝึกโมเดล (model training) : การใช้ข้อมูลตัวอย่างเพื่อฝึกให้โมเดลเรียนรู้และพยายามหาข้อสรุป
- การทดสอบโมเดล (model testing) : การประเมินประสิทธิภาพของแบบจำลองด้วยข้อมูลที่แบบจำลองไม่เคยนำมาประมวล

นอกจากนี้ ยังมีเทคนิคต่าง ๆ ที่ช่วยปรับปรุงประสิทธิภาพการเรียนรู้ของแบบจำลอง เช่น การทำ Cross-Validation (เพื่อป้องกัน Overfitting) และการทำ Feature Scaling เพื่อปรับค่าของข้อมูลให้อยู่ในช่วงเดียวกัน (สามารถศึกษาเพิ่มเติมได้จาก [6], [8], [22])

2.1.3 เทคนิค Grid search

ในกระบวนการสร้างโมเดล Machine Learning การเลือกค่าพารามิเตอร์ที่เหมาะสม (Hyperparameter Tuning) เป็นสิ่งสำคัญที่ส่งผลต่อประสิทธิภาพของโมเดลอย่างมาก เทคนิคหนึ่งที่นิยมใช้คือ Grid Search ซึ่งเป็นวิธีการค้นหาค่าที่ดีที่สุดโดยการทดลองค่าต่าง ๆ ที่อยู่ในช่วงหรือชุดที่กำหนดไว้ล่วงหน้า

หลักการของ Grid Search คือการสร้าง “ตาราง” หรือ “กริด” ของค่าพารามิเตอร์หลายชุด แล้วนำแต่ละชุดไปฝึกโมเดลและประเมินผลด้วยวิธีการวัดประสิทธิภาพ เช่น Accuracy, F1 Score หรือ Cross Validation Score จากนั้นเลือกค่าที่ให้ผลลัพธ์ดีที่สุด

ตัวอย่างพารามิเตอร์ที่อาจใช้ Grid Search ได้แก่

- ใน SVM: C, gamma, kernel
- ใน Random Forest: n_estimators, max_depth, min_samples_split

Grid Search มักใช้งานร่วมกับ Cross Validation เพื่อให้การประเมินผลมีความแม่นยำและลดความเสี่ยงจากการ overfitting

ข้อดีของ Grid Search คือ สามารถหาค่าที่เหมาะสมได้อย่างเป็นระบบ
ข้อจำกัด คือ ใช้เวลานานและทรัพยากรมากเมื่อพารามิเตอร์มีหลายค่า (combinatorial explosion)

2.1.4 เทคนิคการประเมินโมเดล: Cross Validation

Cross Validation (CV) คือเทคนิคสำคัญที่ใช้ในการประเมินประสิทธิภาพของโมเดล Machine Learning โดยเฉพาะในกรณีที่มีข้อมูลจำกัด เทคนิคนี้ช่วยลดความลำเอียงจากการแบ่งชุดข้อมูลแบบสุ่มเพียงครั้งเดียว และสามารถตรวจสอบความสามารถของโมเดลในการ generalize ไปยังข้อมูลใหม่ได้ดีขึ้น

หลักการของ Cross Validation คือการแบ่งข้อมูลทั้งหมดออกเป็นหลายชุดย่อย (folds) จากนั้นนำแต่ละชุดผลัดกันเป็น validation set ส่วนที่เหลือใช้เป็น training set ตัวอย่างที่นิยมคือ:

- k-Fold Cross Validation: ข้อมูลจะถูกแบ่งเป็น k ส่วนเท่า ๆ กัน ในแต่ละรอบจะเลือก 1 ส่วนเป็น validation set และอีก k-1 ส่วนเป็น training set กระบวนการนี้จะทำซ้ำ k ครั้ง แล้วคำนวณค่าเฉลี่ยของผลลัพธ์ทั้งหมด

การใช้ Cross Validation ควบคู่กับการปรับพารามิเตอร์ เช่นใน Grid Search จะช่วยให้ได้โมเดลที่มีประสิทธิภาพสูงและสามารถใช้งานกับข้อมูลใหม่ได้ดี

2.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) เป็นอัลกอริทึมสำหรับการเรียนรู้แบบมีผู้สอน (Supervised Learning) ที่ได้รับความนิยมสูงในงานจำแนกประเภท (Classification) (อาทิเช่น [7],[18]) หรือในบางกรณีก็สามารถนำมาใช้ในการวิเคราะห์การถดถอย (Regression analysis) ได้ด้วย โดยจุดเด่นของ SVM คือ ความสามารถในการสร้างเส้นแบ่งข้อมูล (decision boundary) ที่เหมาะสมที่สุดในเชิงคณิตศาสตร์ ซึ่งเส้นแบ่งดังกล่าวจะสามารถแยกกลุ่มของข้อมูลออกจากกันได้อย่างมีประสิทธิภาพ โดยเฉพาะกรณีที่ข้อมูลมีมิติสูง (high-dimensional data)

2.2.1 หลักการทำงานของ SVM

หลักการของ SVM มีแนวคิดพื้นฐาน คือ การหาเส้นหรือระนาบ (hyperplane) ที่แบ่งข้อมูลออกเป็นสองกลุ่มโดยให้ margin หรือ ระยะห่างระหว่างขอบเขตกับข้อมูลของแต่ละกลุ่มมีค่ามากที่สุด กล่าวคือ SVM จะพยายามหาขอบเขตที่ "ดีที่สุด" ตามเงื่อนไขนี้
ในกรณีที่ข้อมูลสามารถแยกด้วยเส้นตรงได้ (linearly separable) SVM จะหาเส้นแบ่งที่ห่างจากข้อมูลทั้งสองกลุ่ม (หรือคลาส) ให้มากที่สุด และข้อมูลที่อยู่ใกล้ขอบเขตมากที่สุด (เรียกว่า Support Vectors) จะมีบทบาทในการกำหนดตำแหน่งของขอบเขต

2.2.2 Hyperparameter ของ SVM

การปรับค่า Hyperparameter นั้นมีความสำคัญอย่างยิ่งต่อประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง Support Vector Machine (SVM) โดยเฉพาะเมื่อเลือกใช้ Kernel Function อย่าง Linear หรือ Radial Basis Function (RBF) ซึ่งมีลักษณะการจำแนกข้อมูลที่ต่างกัน ค่าที่สำคัญของ SVM ได้แก่:

- C (Regularization Parameter):
  ค่านี้ควบคุมความสมดุลระหว่างความถูกต้องของการจำแนกข้อมูลในชุดฝึก (training set) และความสามารถในการ generalize กับข้อมูลใหม่ ค่าที่ต่ำจะยอมให้มีการแบ่งผิดบ้างเพื่อให้ได้ margin ที่กว้างขึ้น ในขณะที่ค่าที่สูงจะพยายามแบ่งข้อมูลให้ถูกต้องทั้งหมด ซึ่งอาจทำให้เกิด overfitting

- Kernel Type: ตัวกำหนดวิธีการแปลงข้อมูลเข้าสู่ feature space มิติสูง ในงานวิจัยนี้ได้ทดลองระหว่าง Linear Kernel และ RBF Kernel
    - *Linear Kernel* เหมาะกับข้อมูลที่สามารถแยกได้ด้วยเส้นตรง
    - *RBF Kernel* มีความสามารถสูงในการแยกข้อมูลที่มีลักษณะโค้งหรือซับซ้อนในมิติเดิม
- Gamma (γ): ใช้เฉพาะในกรณีที่ใช้ RBF Kernel (รวมถึง Polynomial และ Sigmoid) ค่านี้ยิ่งสูงจะทำให้แต่ละจุดข้อมูลมีอิทธิพลจำกัดอยู่ใกล้ตัวเอง ซึ่งอาจทำให้โมเดลซับซ้อนมากขึ้นและเสี่ยงต่อ overfitting ขณะที่ค่า gamma ต่ำจะทำให้ขอบเขตของการจำแนกเนียนและกว้างขึ้น

การปรับค่า Hyperparameter เหล่านี้ทำได้โดยใช้เทคนิคเช่น Grid Search หรือ Cross Validation เพื่อหาค่าที่เหมาะสมที่สุดกับข้อมูลในงานวิจัยนั้น ๆ

2.3 Random Forest

Random Forest เป็นอัลกอริทึมในกลุ่มของ Ensemble Learning (เทคนิคใน Machine Learning ที่รวมผลของโมเดลหลายตัวเข้าด้วยกัน ศึกษาเพิ่มเติมได้จาก [2]) ที่ได้รับความนิยมอย่างแพร่หลายในงานจำแนกประเภท (Classification) (อาทิการศึกษาของ [12]) และการวิเคราะห์การถดถอย (Regression analysis) เนื่องจากสามารถเพิ่มความแม่นยำของแบบจำลองและลดปัญหา overfitting ได้อย่างมีประสิทธิภาพ โดยหลักการทำงานของ Random Forest คือ การรวมโมเดลย่อยหลายตัว (โดยเฉพาะ Decision Trees) เข้าด้วยกันเพื่อให้ได้ผลลัพธ์ที่มีความน่าเชื่อถือมากยิ่งขึ้น

2.3.1 แนวคิดของ Decision Tree

Decision Tree เป็นโมเดลที่ใช้แนวคิดแบบลำดับชั้น (hierarchical) โดยเริ่มจากราก (root) และแตกแขนงไปยัง node ต่าง ๆ ตามค่าของ features เพื่อแยกข้อมูลออกเป็นกลุ่มย่อย โดยแต่ละ node จะมีเกณฑ์การตัดสินใจ ตัวอย่างเช่น ค่า threshold ของ feature หนึ่ง ๆ เพื่อแยกข้อมูลให้มีความบริสุทธิ์ (purity) มากที่สุด ซึ่งการตัดสินใจว่าจะใช้ feature ใดเพื่อแบ่งข้อมูลจะขึ้นอยู่กับเกณฑ์ เช่น Gini Index, Entropy/ Information Gain เป็นต้น อย่างไรก็ตาม Decision Tree เพียงต้นเดียวมักมีปัญหา Overfitting โดยเฉพาะเมื่อข้อมูลมี noise หรือข้อมูลตัวอย่างมีจำนวนน้อย

2.3.2 Hyperparameter ของ Random Forest (RF)

Random Forest เป็นแบบจำลองการเรียนรู้ของเครื่องแบบ ensemble ที่ประกอบด้วยต้นไม้ตัดสินใจ (Decision Trees) จำนวนมาก โดยการปรับค่าพารามิเตอร์ที่เหมาะสมจะส่งผลต่อความแม่นยำและความสามารถในการ generalize ของโมเดล โดย Hyperparameter ที่สำคัญของ RF ได้แก่

- n_estimators: จำนวนต้นไม้ที่ใช้ในป่า (forest) ค่ายิ่งมากมักช่วยลดปัญหา overfitting และเพิ่มความเสถียรของโมเดล แต่ก็ใช้เวลาในการฝึกสอนมากขึ้น
- max_depth: ความลึกสูงสุดของแต่ละต้นไม้ หากไม่กำหนด ค่า default จะปล่อยให้ต้นไม้เติบโตจนข้อมูลหมด ซึ่งอาจทำให้ overfitting ได้ การจำกัดความลึกช่วยให้โมเดลเรียนรู้ pattern ที่ไม่ซับซ้อนเกินไป
- min_samples_split: จำนวนข้อมูลขั้นต่ำที่จำเป็นต่อการแยก (split) โหนดหนึ่งออกเป็นสองโหนดย่อย ค่านี้ยิ่งมาก จะทำให้ต้นไม้เติบโตช้าลง และช่วยป้องกัน overfitting
- min_samples_leaf: จำนวนข้อมูลขั้นต่ำที่โหนดปลาย (leaf node) ต้องมี เพื่อให้ต้นไม้ไม่จำแนกข้อมูลยิบย่อยเกินไป ช่วยลดการเรียนรู้ noise จากข้อมูล
- max_features: จำนวน feature สูงสุดที่อนุญาตให้พิจารณาเมื่อตัดสินใจ split ในแต่ละโหนด ค่านี้มีผลต่อความหลากหลายของต้นไม้ใน forest
    - `'sqrt'` (ค่าเริ่มต้นสำหรับการจำแนก): ใช้รากที่สองของจำนวน feature ทั้งหมด
    - `'log2'`: ใช้ลอการิทึมฐานสอง
    - ค่าจำนวนเต็มหรือทศนิยมก็สามารถกำหนดเองได้

การปรับค่า Hyperparameter เหล่านี้สามารถทำได้โดยใช้เทคนิค Grid Search ร่วมกับการประเมินผลแบบ Cross Validation เพื่อค้นหาค่าที่เหมาะสมกับชุดข้อมูลของงานวิจัย

2.4 โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network: CNN)

Convolutional Neural Networks (CNNs) เป็นสถาปัตยกรรมหนึ่งของ Deep Learning ที่ถูกออกแบบมาเพื่อจัดการกับข้อมูลแบบภาพ (image data) โดยเฉพาะ จุดเด่นของ CNN คือความสามารถในการเรียนรู้ลักษณะเฉพาะ (features) จากภาพโดยอัตโนมัติ โดยไม่ต้องอาศัยการออกแบบคุณลักษณะด้วยมือ (manual feature engineering) เหมือนในอัลกอริทึมดั้งเดิม

2.4.1 CNN เหมาะกับงานประมวลผลภาพ (Image processing)

CNN ได้รับการออกแบบให้สามารถรับรู้โครงสร้างเชิงพื้นที่ของภาพได้อย่างมีประสิทธิภาพ เช่น ตำแหน่ง รูปร่าง และรูปแบบที่ปรากฏในพื้นที่ต่าง ๆ ของภาพ การใช้ convolution ช่วยให้โมเดลสามารถเรียนรู้ลักษณะภาพโดยไม่จำเป็นต้องดูทั้งภาพพร้อมกัน ซึ่งทำให้สามารถนำไปใช้กับภาพที่มีขนาดใหญ่ได้

นอกจากนี้ โครงสร้างของ CNN ยังสามารถขยายเป็น deep CNN โดยมีหลาย convolutional layers เพื่อเรียนรู้ลักษณะที่ซับซ้อนขึ้นเรื่อยๆ จาก low-level features (เช่น ขอบ) ไปจนถึง high-level features (เช่น รูปร่างของใบพืช)

2.4.2 การเรียนรู้ features อัตโนมัติ

หนึ่งในข้อได้เปรียบของ CNN คือความสามารถในการเรียนรู้ features โดยตรงจากข้อมูลดิบโดยไม่ต้องออกแบบด้วยมนุษย์ เช่น ในการจำแนกสมุนไพร CNN สามารถเรียนรู้ได้เองว่า "ใบของสมุนไพร A" มีขอบหยัก "สมุนไพร B" มีจุดบนใบ เป็นต้น คุณสมบัตินี้ทำให้ CNN มีความยืดหยุ่นและประสิทธิภาพสูง โดยเฉพาะเมื่อมีข้อมูลจำนวนมากในการฝึก

2.4.3 Hyperparameter ของ CNN

ในการฝึก Convolutional Neural Networks (CNN) เพื่อให้ได้ประสิทธิภาพสูงสุดในการจำแนกรูปภาพหรือข้อมูลเชิงพื้นที่ประเภทอื่น ๆ จำเป็นต้องกำหนดและปรับแต่งชุดของไฮเปอร์พารามิเตอร์ (hyperparameters) อย่างเหมาะสม โดยไฮเปอร์พารามิเตอร์เหล่านี้ส่งผลโดยตรงต่อความสามารถในการเรียนรู้ของโมเดล และมีผลต่อความเร็วในการฝึก รวมถึงความสามารถในการ generalize ไปยังข้อมูลที่ไม่เคยเห็นมาก่อน

ไฮเปอร์พารามิเตอร์ที่สำคัญของ CNN ได้แก่:

- Learning Rate (อัตราการเรียนรู้): กำหนดขนาดของการอัปเดตน้ำหนักในแต่ละรอบของการฝึก
- Batch Size: จำนวนตัวอย่างข้อมูลที่ใช้ในการอัปเดตพารามิเตอร์ต่อครั้ง
- Number of Epochs: จำนวนรอบทั้งหมดที่ข้อมูลจะถูกส่งผ่านโมเดล
- Kernel Size: ขนาดของฟิลเตอร์ที่ใช้ในการคอนโวลูชัน เช่น 3×3, 5×5
- Padding: การเติมค่ารอบภาพเพื่อรักษาขนาด spatial dimension เช่น 'same' หรือ 'valid'
- Dropout Rate: อัตราการปิดการทำงานของนิวรอนบางส่วนในช่วงฝึก เพื่อป้องกัน overfitting
- Activation Function: ฟังก์ชันไม่เชิงเส้น เช่น ReLU, LeakyReLU, หรือ tanh ที่ใช้หลังจาก convolution และ dense layers

การเลือกและปรับแต่งค่าไฮเปอร์พารามิเตอร์เหล่านี้มักทำผ่านกระบวนการ trial-and-error, grid search, หรืออัลกอริธึมขั้นสูงอย่าง Bayesian optimization หรือ Hyperband ทั้งนี้ขึ้นอยู่กับทรัพยากรคำนวณและลักษณะของปัญหาที่กำลังศึกษา

2.5 MobileNetV3-Small

MobileNetV3 เป็นสถาปัตยกรรมโครงข่ายประสาทเทียมแบบ CNN ที่พัฒนาโดย Google โดยมุ่งเน้นให้มีความแม่นยำสูงในขณะที่ยังคงประสิทธิภาพด้านความเร็วและขนาดโมเดลที่เล็ก เหมาะสำหรับการใช้งานบนอุปกรณ์เคลื่อนที่ ในงานวิจัยนี้เลือกใช้ MobileNetV3-Small เป็น feature extractor ซึ่งคือ การใช้ MobileNetV3-small ในการดึงลักษณะเฉพาะ (features) ที่สำคัญจากภาพสมุนไพร โดยการใช้ MobileNetV3-Small ช่วยให้แบบจำลอง

สามารถเรียนรู้รูปแบบเชิงลึกในภาพ เช่น รูปร่างใบ พื้นผิว สีและเส้นลายของใบ เป็นต้น ซึ่งจะทำให้แบบจำลองสามารถเรียนรู้ที่จะแยกแยะสมุนไพรแต่ละชนิดได้อย่างดีขึ้น

2.5.1 การประยุกต์ใช้เป็น Feature Extractor

ในงานวิจัยนี้ MobileNetV3-Small ถูกนำมาใช้ในลักษณะของ pretrained CNN ที่ตัดส่วน fully connected layer ออก แล้วใช้เฉพาะ convolutional blocks เพื่อ ดึงคุณลักษณะ (features) จากภาพสมุนไพร

กระบวนการ

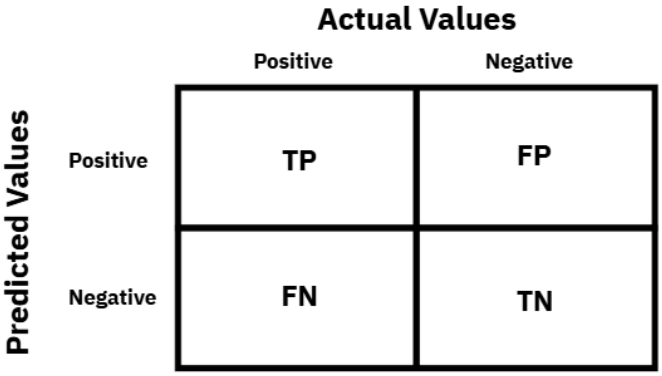
1. อินพุต: ป้อนภาพสมุนไพรขนาดมาตรฐาน เช่น 224×224224 \times 224224×224 พิกเซล เข้าสู่ MobileNetV3-Small
2. การสกัดคุณลักษณะ (Feature Extraction): Feature map สุดท้ายก่อน flattening จะมีค่าที่สะท้อนถึงลักษณะของภาพ เช่น ขอบใบ, ลายเส้น, สี
3. การนำไปใช้งานต่อ: Feature vector ที่ได้สามารถนำไปใช้กับโมเดลจำแนก เช่น SVM, Fully Connected Layer หรือ Random Forest

2.6 การวัดประสิทธิภาพ (Evaluation Metrics)

2.6.1 Confusion Matrix

Confusion Matrix เป็นตารางสรุปผลการจำแนกของโมเดล ประกอบด้วย 4 ค่า

- TP (True Positive): ทำนายถูกว่าเป็นคลาสนั้น
- TN (True Negative): ทำนายถูกว่าไม่ใช่คลาสนั้น
- FP (False Positive): ทำนายผิดว่าเป็นคลาสนั้น
- FN (False Negative): ทำนายผิดว่าไม่ใช่คลาสนั้น



ภาพที่ 2: ภาพ **Confusion Matrix** หรือ **ตารางเปรียบเทียบค่าทำนายกับความจริง** ที่ใช้ในการประเมินประสิทธิภาพของโมเดล Classification จาก (สุพจน์, 2567 [27])

2.6.2 ตัวชี้วัดหลัก (Key Metrics)

การวัดความสามารถในการจำแนกประเภทของโมเดล (สุพจน์, 2567 [27])

1) Accuracy คือ ค่าความถูกต้องโดยรวม

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

เหมาะกับข้อมูลที่มีจำนวนตัวอย่างของแต่ละคลาสใกล้เคียงกัน

2) Precision คือ ค่าความแม่นยำในการทำนายว่าเป็นคลาสเป้าหมาย

$$\text{Precision} = \frac{TP}{TP + FP}$$

บอกว่าในจำนวนที่ทำนายว่า “ใช่” มีจริงๆ กี่อันที่ถูกต้อง

3) Recall (Sensitivity) คือ ความสามารถในการจับคลาสเป้าหมายได้ครบ

$$\text{Recall} = \frac{TP}{TP + FN}$$

ใช้ในกรณีที่ False Negative มีผลกระทบมาก เช่น การวินิจฉัยโรคผิด

4) F1-Score คือ เป็นค่าเฉลี่ยแบบถ่วงน้ำหนักระหว่าง Precision และ Recall

$$\text{F1} - \text{score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

เหมาะกับข้อมูลไม่สมดุล หรือเมื่อต้องการพิจารณาความแม่นยำโดยรวมในมุม balance

2.7 การพัฒนาแอปพลิเคชันใน android studio ด้วยภาษา Kotlin

ในหัวข้อนี้จะกล่าวถึงองค์ประกอบสำคัญในการพัฒนาแอปพลิเคชันบนแพลตฟอร์ม Android ด้วยภาษา Kotlin โดยมุ่งเน้นในสามส่วนหลัก ได้แก่ การออกแบบอินเทอร์เฟซผู้ใช้ (User Interface: UI), การสร้างฟังก์ชันการค้นหาข้อมูลภายในแอป และการประยุกต์ใช้กล้องของอุปกรณ์เพื่อการจำแนกวัตถุ (Identification)เนื้อหาจะอธิบายถึงหลักการทำงานเครื่องมือที่เกี่ยวข้องและแนวทางการประยุกต์ใช้ในเชิงเทคนิคอย่างเป็นระบบ

2.7.1 การพัฒนาอินเทอร์เฟซผู้ใช้ (User Interface Design)

การออกแบบส่วนติดต่อผู้ใช้ (UI) เป็นขั้นตอนสำคัญที่มีผลโดยตรงต่อประสบการณ์ของผู้ใช้งานแอปพลิเคชัน Android ใช้ไฟล์ library ของ jetpack compose 3 เป็นโครงสร้างหลักในการนิยามลักษณะของหน้าจอ และใช้ภาษา Kotlin ในการควบคุมตรรกะเบื้องหลัง Layout ต่าง ๆ การวางโครงสร้างของ Layout เช่น Constraint Layout, Linear Layout และ Relative Layout ช่วยในการจัดตำแหน่งขององค์ประกอบ เช่น ปุ่ม ข้อความ และรูปภาพ ในลักษณะที่เหมาะสมกับขนาดหน้าจอที่หลากหลายของอุปกรณ์ Android

นอกจากนี้ การออกแบบที่ตอบสนองต่อผู้ใช้ (Responsive Design) และการสนับสนุนหลายภาษา (Localization) ก็เป็นส่วนสำคัญในการเพิ่มประสิทธิภาพและความสามารถในการเข้าถึงของแอปพลิเคชัน Responsive Design หมายถึงการออกแบบUI ที่สามารถปรับเปลี่ยนและแสดงผลได้อย่างเหมาะสมตามขนาด

และความละเอียดของหน้าจอในอุปกรณ์ที่แตกต่างกัน เช่น สมาร์ตโฟน แท็บเล็ต หรือจอแนวนอนและแนวตั้ง การกำหนดค่าขนาดและตำแหน่งขององค์ประกอบโดยใช้หน่วยแบบยืดหยุ่น เช่น dp, sp, หรือการตั้งค่า layout_width="match_parent" และ layout_height="wrap_content" รวมถึงการใช้ ConstraintLayout อย่างเหมาะสม จะช่วยให้แอปสามารถปรับตัวตามลักษณะของหน้าจอโดยอัตโนมัติ นอกจากนี้ยังสามารถใช้

โฟลเดอร์เฉพาะ เช่น layout-sw600dp/ หรือ drawable-xxhdpi/ เพื่อจัดแสดงเนื้อหาที่เหมาะสม

กับอุปกรณ์แต่ละประเภท

ในส่วนของ Localization หรือการรองรับหลายภาษานักพัฒนาสามารถแยกข้อความในแอปออกจากโค้ดหลักและจัดเก็บไว้ในไฟล์ทรัพยากร (strings.xml) เพื่อให้สามารถแปลข้อความได้อย่างเป็นระบบ โดยการสร้างไฟล์แยกตามภาษาที่ต้องการ เช่น values-th/strings.xml สำหรับภาษาไทย และ values-en/strings.xml สำหรับภาษาอังกฤษ แล้วใช้คำสั่ง @string/ชื่อข้อความ เรียกใช้ใน UI เพื่อให้ระบบโหลดภาษาที่เหมาะสมกับการตั้งค่าของอุปกรณ์ ผู้ใช้งานโดยอัตโนมัติ การดำเนินการดังกล่าวช่วยเพิ่มความสามารถในการเข้าถึง (Accessibility) และเพิ่มโอกาสในการขยายผู้ใช้งานไปยังกลุ่มเป้าหมายในหลายประเทศ

2.7.2 การพัฒนาฟังก์ชันค้นหา (Search Function Implementation)

ฟังก์ชันค้นหาเป็นองค์ประกอบสำคัญของแอปพลิเคชันที่มีฐานข้อมูลขนาดค่อนข้างใหญ่และผู้ใช้จำเป็นจะต้องค้นหาข้อมูลจำเพาะจากข้อมูลจำนวนมาก เช่น แอปพลิเคชันแสดงรายการวัตถุดิบ พืช หรือข้อมูลทางการแพทย์ โดยทั่วไปจะมีการใช้ SearchView หรือ EditText เพื่อรับข้อมูลจากผู้ใช้ จากนั้นนำข้อมูลไปกรอง (Filter) รายการจากแหล่งข้อมูล เช่น RecyclerView หรือฐานข้อมูล Room แบบเรียลไทม์ กลไกการค้นหานี้สามารถทำงานได้ทั้งในแบบออนไลน์ (ดึงจาก API) และออฟไลน์ (ค้นจากข้อมูลในเครื่อง) เพื่อความรวดเร็วและความยืดหยุ่น การใช้เทคนิคที่เหมาะสม เช่น การ normalize ข้อความ การค้นหาด้วย keyword หรือ fuzzy search ยังสามารถช่วยเพิ่มความแม่นยำในการค้นหาได้อีกด้วย

2.7.3 การใช้งานกล้องเพื่อการจำแนก (Camera Function for Identification)

การนำกล้องมาใช้ในแอปพลิเคชันเพื่อจำแนกวัตถุ เช่น พืช สมุนไพร หรือวัตถุเฉพาะทางอื่น ๆ จำเป็นต้องรวมความสามารถของฮาร์ดแวร์ (Hardware Access) เข้ากับกระบวนการประมวลผลภาพ (Image Processing) ใน Android Studio นักพัฒนาสามารถใช้งาน CameraX หรือ Camera2 API เพื่อเปิดกล้องและดึงข้อมูลภาพในรูปแบบ real-time

หรือจากการถ่ายภาพจากนั้นสามารถนำภาพดังกล่าวไปประมวลผลร่วมกับโมเดลการเรียนรู้ของเครื่อง (เช่น TensorFlow Lite) เพื่อทำการจำแนกวัตถุ ฟังก์ชันนี้มักต้องการการขออนุญาต (Permissions) อย่างเหมาะสม และต้องพิจารณาประสิทธิภาพของอุปกรณ์เพื่อให้การประมวลผลไม่ส่งผลกระทบต่อประสบการณ์ใช้งานของผู้ใช้

2.7.4 การใช้งานแบบการเรียนรู้ของเครื่องในแอปพลิเคชัน (Machine Learning Utilization)

Tensorflow Lite (TFLite) เป็นกรอบการทำงานโมเดลการเรียนรู้ของเครื่องที่เน้นความเร็วประมวลผล ขนาดไฟล์ที่เล็ก และสามารถทำงานร่วมกับ Android Studio ได้ง่าย เหมาะกับการนำไปใช้งานในเครื่องมือ

ที่มีประสิทธิภาพในการประมวลผลจำกัด เช่น โทรศัพท์มือถือ และ อินเตอร์เน็ตแห่งสรรพสิ่ง (Internet Of Things) ทำให้เป็น framework ที่ผู้จัดทำได้นำไปใช้งาน

แต่ TFLite เป็น framework ที่เน้นการทำงานแบบเครือข่ายหน่วยประสาท (Neural-Network) ที่สนับสนุนเพียงโมเดล CNN จากทั้ง 3 โมเดลเท่านั้น อันเนื่องมาจากอีกทั้งสองโมเดล (SVM และ Random Forest) นั้น

1. โครงสร้างการคำนวณที่ไม่ใช่ Tensor: SVM
   ใช้การหาคำตอบเชิงคณิตศาสตร์ผ่านการหาค่าความเหมาะสมกับข้อมูลขอบเขต (support vectors) และ kernel functions ส่วน Random Forest ทำงานผ่านการตัดสินใจแบบลำดับ (tree traversal) และการแยกกิ่ง ซึ่งไม่สามารถแปลงเป็นการคำนวณแบบ tensor ได้อย่างมีประสิทธิภาพ
2. ไม่มีการสนับสนุนโดยตรงจาก TensorFlow: SVM และ Random Forest มักจะถูกสร้างผ่านไลบรารีอย่าง scikit-learn หรือ XGBoost ซึ่งไม่มีเส้นทางมาตรฐานในการแปลงโมเดลเหล่านี้ให้อยู่ในรูปแบบที่ TFLite รองรับ
3. ข้อจำกัดด้านประสิทธิภาพและขนาดโมเดล: โมเดลของ TFLite มักถูกปรับให้มีขนาดเล็กและรันได้รวดเร็ว ซึ่ง CNN สามารถถูกดัดแปลงเพื่อให้เหมาะกับข้อจำกัดของขนาดได้ ขณะที่ SVM และ ะ Random Forest มีลักษณะการคำนวณและหน่วยความจำที่ไม่สม่ำเสมอทำให้การปรับแต่งและรันบนอุปกรณ์ขนาดเล็กทำได้ยาก

กล่าวโดยสรุป การที่ TFLite รองรับ CNN แต่ไม่รองรับ SVM และ Random Forest เป็นผลจากการออกแบบที่เน้นประสิทธิภาพและความเข้ากันได้กับสถาปัตยกรรมของ deep learning ซึ่ง CNN สามารถใช้งานร่วมกับ TFLite ได้อย่างมีประสิทธิภาพ ขณะที่โมเดลการเรียนรู้แบบดั้งเดิมนั้น ไม่เหมาะสมกับแนวทางดังกล่าว

3 งานวิจัยที่เกี่ยวข้อง

ในช่วงหลายปีที่ผ่านมา การผสานองค์ความรู้ดั้งเดิมเกี่ยวกับสมุนไพรเข้ากับเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) และการเรียนรู้ของเครื่อง (Machine Learning) ได้กลายเป็นแนวทางที่สำคัญในการพัฒนาเครื่องมือเพื่อการจำแนกพืชสมุนไพร รวมถึงการใช้สมุนไพรอย่างปลอดภัยเมื่อร่วมกับยาแผนปัจจุบัน

ในด้านความปลอดภัยของการใช้สมุนไพรร่วมกับยาแผนปัจจุบัน สมุนไพรได้รับความนิยมอย่างแพร่หลายทั้งในด้านการส่งเสริมสุขภาพและการรักษาโรค โดยเฉพาะในรูปแบบการใช้ร่วมกับยาแผนปัจจุบัน อย่างไรก็ตาม การใช้สมุนไพรร่วมกับยาทางการแพทย์อาจก่อให้เกิดอันตรกิริยาทางเภสัชจลนศาสตร์หรือเภสัชพลศาสตร์ที่คาดไม่ถึง รวมถึงอาจส่งผลเสียต่อผู้ป่วยอย่างรุนแรง ซึ่งมีงานวิจัยมากมายได้กล่าวถึง เช่น ในปี 2000 Chan และคณะ ([4]) ได้ศึกษาอันตรกิริยาระหว่างสมุนไพร Danshen (Salvia miltiorrhiza) กับยา warfarin และพบว่า Danshen มีฤทธิ์ยับยั้งการแข็งตัวของเลือดหลายกลไก ซึ่งอาจเสริมฤทธิ์ยา warfarin จนทำให้เกิดความเสี่ยงจากการตกเลือด นอกจากนั้นยังมีรายงานโดย Schwaiblmair และคณะในปี 2012 ([18]) เกี่ยวกับภาวะโรคปอดระหว่างเนื้อเยื่อจากยา (Drug-Induced Interstitial Lung Disease: DILD) ซึ่งเกิดจากยาเคมีบำบัด ยาปฏิชีวนะ และยากดภูมิคุ้มกัน โดยเน้นว่าการวินิจฉัยทำได้ยากเนื่องจากไม่มีลักษณะเฉพาะทางคลินิกหรือภาพรังสี รวมถึง (*Pazhayattil & Shirali,*

2014 [19]) ได้ทบทวนความเสี่ยงของยาที่ก่อให้เกิดพิษต่อไต (nephrotoxicity) และชี้ว่าการประเมินผลข้างเคียงจากยาเป็นสิ่งสำคัญในการออกแบบกลยุทธ์การรักษาและการป้องกัน ในส่วนประเด็นของการแพ้ยานั้น Thong และ Tan ([22]) ได้สังเคราะห์วรรณกรรมทางวิชาการเกี่ยวกับปัจจัยเสี่ยงของการแพ้ยาจนพบว่ายังมีความแตกต่างในนิยามและวิธีการรายงานระหว่างงานวิจัย ซึ่งเป็นอุปสรรคต่อการวิเคราะห์และเปรียบเทียบผลได้อย่างเป็นระบบ

สำหรับการจำแนกชนิดของพืชจากลักษณะทางสัณฐานวิทยา อาทิเช่น รูปร่างของใบ ขอบใบ เส้นใบ หรือโครงสร้างของลำต้นและดอก ถือเป็นวิธีการพื้นฐานที่ใช้กันมาอย่างยาวนานในทางพฤกษศาสตร์ แต่การประเมินด้วยสายตามนุษย์มักมีข้อจำกัด ทั้งในแง่ของความแม่นยำ ความต่อเนื่อง และความสามารถในการแยกพืชที่มีลักษณะคล้ายกัน ด้วยเหตุนี้เทคโนโลยีการประมวลผลภาพร่วมกับอัลกอริทึมการเรียนรู้ของเครื่องจึงถูกนำมาใช้ โดยในปี 2023 Bambil และคณะ ([1]) ได้ทดลองใช้อัลกอริทึม AdaBoost, Random Forest, SVM และ Deep Learning เพื่อจำแนกชนิดพืช โดยพบว่า SVM, Random Forest และ Deep Learning ให้ความแม่นยำเกิน 93% และสามารถนำไปประยุกต์ใช้บนอุปกรณ์พกพาได้อย่างมีประสิทธิภาพ แต่หากจะจำเพาะการประยุกต์ใช้งานจริงโดยเน้นเฉพาะพืชสมุนไพร งานของ Muneer และ Fati ในปี 2020 ([14]) ชี้ให้เห็นถึงความเป็นไปได้เป็นอย่างมาก เนื่องจากทั้งสองคนได้ทดลองนำภาพใบสมุนไพรจำนวน 1,000 ภาพมาจำแนก โดยใช้แบบจำลอง SVM และ Deep Learning Neural Network (DLNN) ซึ่งพบว่า DLNN ให้ผลแม่นยำถึง 93% และใช้เวลาประมวลผลน้อยกว่า SVM อย่างมีนัยสำคัญ ในส่วนของแบบจำลองขั้นสูงยิ่งขึ้นอย่าง EfficientNetV2B0 ก็ได้รับการประยุกต์ใช้โดย Putra และคณะ ในปี 2022 ([17]) โดยในการทดลองจะใช้ภาพใบสมุนไพรไทยจำนวน 10 ชนิด รวมภาพสมุนไพรทั้งสิ้น 3,500 ภาพ ผลการศึกษาพบว่าแบบจำลอง EfficeintNetV2B0 สามารถจำแนกชนิดสมุนไพรได้อย่างแม่นยำถึง 99.14% โดยมีค่าความสูญเสียเพียง 1.95% รวมถึงการประยุกต์ใช้ CNN บนเครือข่ายข้อมูลจริงในระดับประเทศปรากฏในงานของ Yalcin และ Razavi ([24]) ซึ่งใช้ข้อมูลจากสถานีเกษตรกว่า 1,200 แห่งในประเทศตุรกี โดยเปรียบเทียบ CNN กับ SVM ที่ใช้ฟีเจอร์ LBP และ GIST มีผลการศึกษาว่า CNN มีประสิทธิภาพเหนือกว่าอย่างชัดเจน

นอกจากนี้นักวิจัยคนไทยยังได้เคยพัฒนาแอปพลิเคชันเพื่อใช้ระบุชนิดสุมนไพร ชื่อว่า Herb ID ด้วยโมเดล MobileNet (*Phophan, 2023* [26]) ซึ่งสามารถจำแนกสมุนไพรไทยจากภาพถ่ายได้อย่างรวดเร็วและแม่นยำ ซึ่งมีผลการทดลองสอดคล้องกับ Kavitha และคณะ ([9]) ที่ได้พัฒนาแอปพลิเคชันแบบเรียลไทม์โดยใช้โมเดล MobileNetV3 ที่มีความแม่นยำถึง 96.47% บนอุปกรณ์พกพา นอกจากนั้นเพื่อส่งเสริมการเรียนรู้และเข้าถึงข้อมูลสมุนไพรอย่างมีระบบ (*Chaikhamwang et al., 2021* [3]) ได้พัฒนาแอปพลิเคชัน Thai Herbs Application ที่ใช้

AutoML Vision API จำแนกภาพพืชสมุนไพร พร้อมระบบ Web Application สำหรับจัดการข้อมูลสมุนไพร อาการและข้อควรระวัง โดยได้รับคะแนนประเมินจากผู้เชี่ยวชาญในระดับสูง

ผู้ทำงานวิจัย	เป้าหมาย	โมเดลที่ใช้	acacuracy
bambil et.al [1]	จำแนกชนิดพืชโดยใช้สี รูปร่าง และเนื้อสัมผัสร่วมกับ machine learning และ ANN	Machine Learning + Artificial	94.70%
Muneer et al. [14]	จำแนกพืชสมุนไพรโดยใช้คุณลักษณะเชิงรูปร่างและเนื้อสัมผัสร่วมกับ deep learning	Deep learning	92.35%
Putra et al. [17]	จำแนกใบสมุนไพรด้วย EfficintNetV2B0	EfficientNetV2B0	95.20%
Kavitha et al. [9]	จำแนกพืชสมุนไพรแบบเรียลไทม์ผ่านกล้องโดยใช้ deep leaning	MobileNetv3	96.47%
		SVM	96.62%
		Random Forest	88.47%
กลุ่มผู้พัฒนา	ทดสอบและหาโมเดลที่มีประสิทธิภาพที่สุดสำหรับการจำแนกสมุนไพรไทย	CNN	87.45%

ตารางเปรียบเทียบประสิทธิภาพของโมเดลจากผู้พัฒนาที่แตกต่างกัน

4 ขั้นตอนการทำวิจัย

การศึกษาบทความเบื้องต้น

บทความเกี่ยวกับอันตรกิริยาระหว่างสมุนไพรกับยาแผนปัจจุบัน

อันตรกริยาของสมุนไพรและผลิตภัณฑ์เสริมอาหารกับยาแผนปัจจุบัน โดย อ.ดร.ภาณุพงศ์ พุทธรักษ์ คณะเภสัชศาสตร์ มหาวิทยาลัยสงขลานครินทร์

สมุนไพรและผลิตภัณฑ์เสริมอาหารได้รับความนิยมเพิ่มขึ้นเนื่องจากผู้บริโคคเชื่อมั่นในความปลอดภัยและประสิทธิภาพจากธรรมชาติ โดยเฉพาะในช่วงการระบาดของโรคโควิด-19 อย่างไรก็ตาม ยังพบปัญหาการใช้ที่ขาดความรู้ ความเข้าใจ หรือไม่มีข้อมูลประกอบการตัดสินใจอย่างถูกต้อง บทความนี้จึงรวบรวมข้อมูลอันตรกิริยาและแนวทางการใช้สมุนไพรเพื่อส่งเสริมการบริโภคที่ปลอดภัยและมีประสิทธิภาพยิ่งขึ้น

การศึกษางานวิจัยการจำแนกพืชด้วยภาพ

การจำแนกภาพใบสมุนไพรด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก Herb Leaf Image Classification with Deep Convolutional Neural Networks โดย มธุรส ผ่านเมือง[1], ชนนิกานต์ รอดมรกฎ[2]
Mathuros Panmuang[1], Chonnikarn Rodmorn[2]
[1]มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
[2]มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ปัจจุบันสมุนไพรได้รับความนิยมในการรักษาอาการเจ็บป่วยมากขึ้น แต่การจำแนกชนิดสมุนไพรจากใบยังเป็นปัญหาเนื่องจากความคล้ายคลึงกันของลักษณะใบและขาดผู้เชี่ยวชาญ งานวิจัยนี้จึงเสนอการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก (CNN) เพื่อจำแนกภาพใบสมุนไพร โดยเปรียบเทียบประสิทธิภาพของโมเดลและวิธีเพิ่มความแม่นยำของข้อมูล ผลการทดลองพบว่าโมเดลสามารถจำแนกใบสมุนไพรได้อย่างมีประสิทธิภาพ โดยมีความแม่นยำสูงถึง 95.40%

การสำรวจฐานข้อมูล

ในการพัฒนาการจำแนกสมุนไพรด้วยภาพถ่ายใบ ผู้พัฒนาได้เลือกชนิดสมุนไพรจากการศึกษาหนังสือ **“108 สมุนไพรไทย ใช้เป็น หายป่วย ตำรับยาอายุวัฒนะ กัน–แก้ สารพัดโรคจากครัวเรือน”** ของ สุธีรา วิไลรัตน์ โดยที่หนังสือเล่มนี้รวบรวมข้อมูลเกี่ยวกับสมุนไพรไทยจำนวน 108 ชนิด พร้อมคำแนะนำในการใช้รักษาโรคทั่วไปในครัวเรือน รวมถึงตำรับยาแผนโบราณและสรรพคุณในการบำรุงร่างกายและชะลอวัย เนื้อหานำเสนออย่างเข้าใจง่าย ใช้ได้จริง เหมาะกับทั้งผู้ที่สนใจสมุนไพรพื้นบ้านและผู้เริ่มต้นดูแลสุขภาพด้วยวิธีธรรมชาติ นอกจากนั้น มีการพัฒนาฐานข้อมูลโดยได้รับความอนุเคราะห์จากกรมการกรมการแพทย์แผนไทยและการแพทย์ทางเลือก ซึ่งมีข้อมูลที่สำคัญหลายประการ เช่น ชื่อ ชื่อวิทยาศาสตร์ สรรพคุณ และข้อควรระวัง เป็นต้น

4.1 การรวบรวมและเตรียมชุดข้อมูล

4.1.1 ผู้พัฒนาได้ขอความอนุเคราะห์ข้อมูลภาพใบสมุนไพรจากกรมการแพทย์แผนไทยและการแพทย์ทางเลือก รวมจำนวน 34,500 ภาพ โดยแบ่งเป็น 27,600 ภาพสำหรับฝึกแบบจำลอง และ 6,900 ภาพสำหรับทดสอบ

4.1.2 คัดเลือกเฉพาะสมุนไพรที่สามารถแยกแยะได้จากใบ และสมุนไพรที่ยังไม่เป็นที่รู้จักแพร่หลายจำนวน 69 ชนิด

4.1.3 ทำการคัดกรองภาพที่มีคุณภาพดี ไม่มีสิ่งรบกวน เช่น ดอกหรือลำต้น และแยกภาพเป็นชุดข้อมูลฝึก (Training Set) และชุดข้อมูลทดสอบ (Test Set) ในอัตราส่วน 80:20

4.2 การฝึกและประเมินโมเดล Machine Learning

การทดลองครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของโมเดลการจำแนกภาพสมุนไพรโดยใช้เทคนิค Machine Learning สามรูปแบบ ได้แก่ Support Vector Machine (SVM), Random Forest (RF) และ Convolutional Neural Network (CNN) โดยมีขั้นตอนสำคัญดังต่อไปนี้

4.2.1 การแบ่งชุดข้อมูล (Train-Test Split)

ข้อมูลทั้งหมดถูกแบ่งออกเป็นสองชุดอย่างชัดเจน:

- Training Set: 80% ของข้อมูลทั้งหมด
- Testing Set: 20% ของข้อมูล (แยกไว้ก่อนขั้นตอนฝึก)

การแบ่งข้อมูลนี้ช่วยให้การประเมินประสิทธิภาพของโมเดลมีความน่าเชื่อถือ และหลีกเลี่ยงปัญหา data leakage

4.2.2 การเตรียมข้อมูล (Preprocessing)

ในทุกโมเดลมีการเตรียมข้อมูลในขั้นตอนต้น ดังนี้:

- Resizing: ปรับขนาดภาพทั้งหมดให้เป็น 224 × 224 พิกเซล
- Normalization:
  - ในโมเดล SVM และ RF: ใช้ preprocess_input จาก MobileNetV3 แปลงค่าพิกเซลจากช่วง [0, 255] เป็น [-1, 1]
  - ในโมเดล CNN: ใช้ rescale=1.0/255 แปลงค่าพิกเซลให้อยู่ในช่วง [0, 1]

4.2.3 การใช้ Data Augmentation

ในโมเดล CNN ได้มีการใช้ Data Augmentation บนชุดฝึก (Training Set) เพื่อเพิ่มความหลากหลายของข้อมูล และลดความเสี่ยงจาก overfitting โดยใช้เทคนิคดังต่อไปนี้:

- การหมุนภาพ (rotation) แบบสุ่มในช่วง ±15 องศา
- การเลื่อนภาพ (shift) ไม่เกิน ±10% ของขนาดภาพ
- การซูม (zoom) ในช่วง ±10% ของภาพ
- การสะท้อนแนวนอน (horizontal flip) แบบสุ่ม

SVM และ Random Forest ไม่ได้ใช้ Data Augmentation เนื่องจากใช้การแปลงภาพเป็นเวกเตอร์ผ่าน feature extractor ล่วงหน้า

4.2.4 Cross-validation

ในขั้นตอนการปรับพารามิเตอร์ของ SVM และ RF ได้ใช้เทคนิค k-Fold Cross Validation (k = 3) ร่วมกับ GridSearchCV เพื่อประเมินประสิทธิภาพของค่าพารามิเตอร์แต่ละชุดอย่างรอบด้าน และลดความเอนเอียงจากการแบ่งข้อมูลเพียงครั้งเดียว

4.2.5 การกำหนดค่าพารามิเตอร์ของโมเดล (Model Hyperparameters)

- SVM:

  C: [0.1, 1, 10] ,gamma: [0.001, 0.01, 0.1] และ kernel: ['linear', 'rbf']

- Random Forest:

  n_estimators: [100, 200, 300] ,max_depth: [None, 10, 20] ,min_samples_split: [2, 5, 10]

  และmin_samples_leaf: [1, 2, 4]

- CNN:

  learning_rate: 0.001 ,dropout_rate: 0.2 ,optimizer: Adam ,loss function: Categorical Crossentropy ,callbacks: EarlyStopping, ReduceLROnPlateau ,epochs: 100 และ

  batch size: 32

4.2.6 การประเมินประสิทธิภาพโมเดล

การประเมินใช้ชุดข้อมูลทดสอบที่แยกไว้ล่วงหน้า โดยใช้ตัวชี้วัดต่อไปนี้:

- Accuracy
- F1 Score (weighted)
- Classification Report
- เวลาในการทดสอบ (Testing Time)

4.3 การสร้างฐานข้อมูลสมุนไพร

4.3.1 รวบรวมข้อมูลสมุนไพรจากกรมการแพทย์แผนไทยฯ และเว็บไซต์ของคณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล

4.3.2 แยกประเภทข้อมูลออกเป็น 2 กลุ่ม คือ:

ประเภทข้อมูล	ฟิลด์หลัก
ข้อมูลทั่วไป สมุนไพร 273 ชนิด	รหัส, ชื่อ, ชื่อวิทยาศาสตร์, ชื่อวงศ์, ชื่อท้องถิ่น, ลักษณะทางพฤกษศาสตร์, นิเวศวิทยา, รส และสรรพคุณ, ส่วนที่ใช้, วิธีใช้, ข้อควรระวัง, เอกสารอ้างอิง
ข้อมูลอันตรกิริยา 65 ชนิด	order, herb_name, herb_subname, usage, drug_name, drug_subname, drug_group, reaction_effect, violence_level, risk, suggestion, link

4.3.3 ตรวจสอบความถูกต้องของการคัดลอกข้อมูล โดยทำการตรวจทานข้อมูลที่รวบรวมและคัดลอกข้อมูลมาจัดเก็บไว้เป็นฐานข้อมูลรูปแบบสกุล CSV

4.3.4 จัดรูปแบบข้อมูลให้อยู่ในไฟล์ CSV และแปลงเป็นไฟล์ SQLite สำหรับใช้ในแอปพลิเคชัน

4.4 การพัฒนาแอปพลิเคชันต้นแบบ (Application prototype)

พัฒนาแอปพลิเคชันบนโทรศัพท์มือถือด้วยภาษา Kotlin ผ่านโปรแกรม Android Studio โดยมีคุณสมบัติหลัก ดังนี้

1. การถ่ายภาพใบสมุนไพรและจำแนกด้วยโมเดล CNN
2. การค้นหาข้อมูลสมุนไพรด้วยระบบ Text Search และ Category Search
3. การแสดงข้อมูลสมุนไพรและอันตรกิริยาในรูปแบบที่เข้าใจง่าย

ขั้นตอนการออกแบบแอปพลิเคชัน มีดังนี้

4.4.1 ติดตั้ง Android Studio และ Tools ที่เกี่ยวข้อง

4.4.2 ศึกษาการเขียน UI เบื้องต้น

4.4.3 ศึกษา Library ที่เกี่ยวข้อง; Room database และ Tensorflow Lite

4.4.4 เริ่มสร้างแอพพลิเคชั่น เริ่มจาก UI และ Permission request ต่างๆ

4.4.5 เขียนโค้ดในส่วนการทำนายสมุนไพร โดยเขียนจากการเก็บภาพ, Pre-processing, ประมวลผล และ UI แสดงผล ตามลำดับ

4.4.6 เขียนโค้ดในส่วนการค้นหาด้วยข้อความ (Text search) โดยเขียนจากการนำเข้า Database, แปลง Database เป็น Class ข้อมูล (Kotlin OOP), ฟังก์ชั่นค้นหา และ UI แสดงผล ตามลำดับ

การทดสอบแอปพลิเคชัน

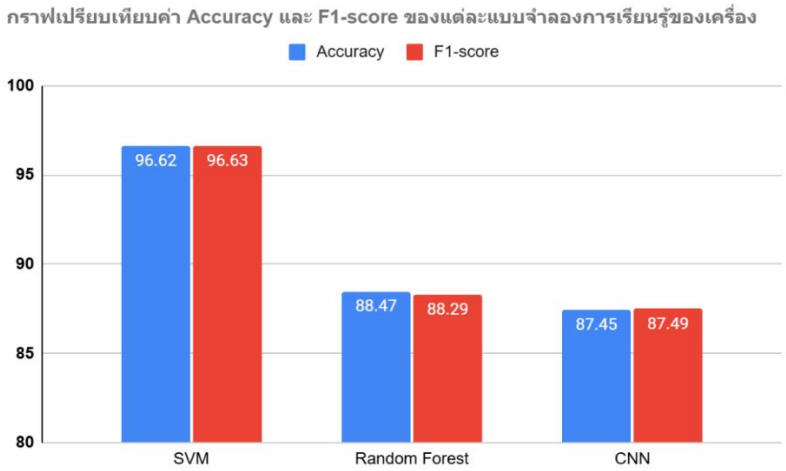
ผู้พัฒนาได้กำหนดกลุ่มผู้ใช้งานทดลองโดย แบ่งเป็นบุคคลทั่วไป 80 คนและผู้เชี่ยวชาญด้านสมุนไพร 20 คน โดยให้ทดลองใช้แอปพลิเคชัน และประเมินความพึงพอใจทั้งความเสถียรของแอปพลิเคชัน การใช้งานโหมดการจำแนกสมุนไพรด้วยรูปถ่ายใบ การค้นหาข้อมูลสมุนไพร และความสวยงามของ UI นอกจากนั้น ผู้ทดลองใช้งานยังสามารถเสนอแนะในเชิงความคิดเห็นปลายเปิดได้ด้วย

5 ผลการศึกษา

ในบทนี้นำเสนอผลการดำเนินงานในแต่ละขั้นตอนที่ได้กล่าวไว้ในบทก่อนหน้าโดยแสดงผลการสร้างแบบจำลองการเรียนรู้ของเครื่อง การพัฒนาฐานข้อมูลสมุนไพร และผลการสร้างแอปพลิเคชันต้นแบบ ดังนี้

5.1 ผลการพัฒนาโมเดลการเรียนรู้ของเครื่องสำหรับการทำนายสมุนไพร

ผู้พัฒนาได้ทำการฝึกแบบจำลอง 3 ประเภท ได้แก่ Support Vector Machine (SVM), Random Forest และ Convolutional Neural Network (CNN) โดยทำการวัดประสิทธิภาพด้วยค่าความแม่นยำ (Accuracy) และค่า F1-score ผลการเปรียบเทียบแสดงดังแผนภูมิแท่งได้ ดังนี้



ภาพที่ 3: แผนภูมิแท่งที่แสดงผลการทดลองการฝึกแบบจำลองการเรียนรู้ของเครื่อง SVM, Random forest และ CNN ซึ่งอธิบายด้วยค่า accuracy, F-1 score ในหน่วย %

แต่หากพิจารณาถึงเวลาในการประมาณผลแล้วจะได้ว่า

1. SVM: เวลาประมวลผล 97.84 วินาที ต่อ 6900 ภาพ
2. Random Forest: เวลาประมวลผล 0.83 วินาที ต่อ 6900 ภาพ
3. CNN: เวลาประมวลผล 582.36 วินาที ต่อ 6900 ภาพ

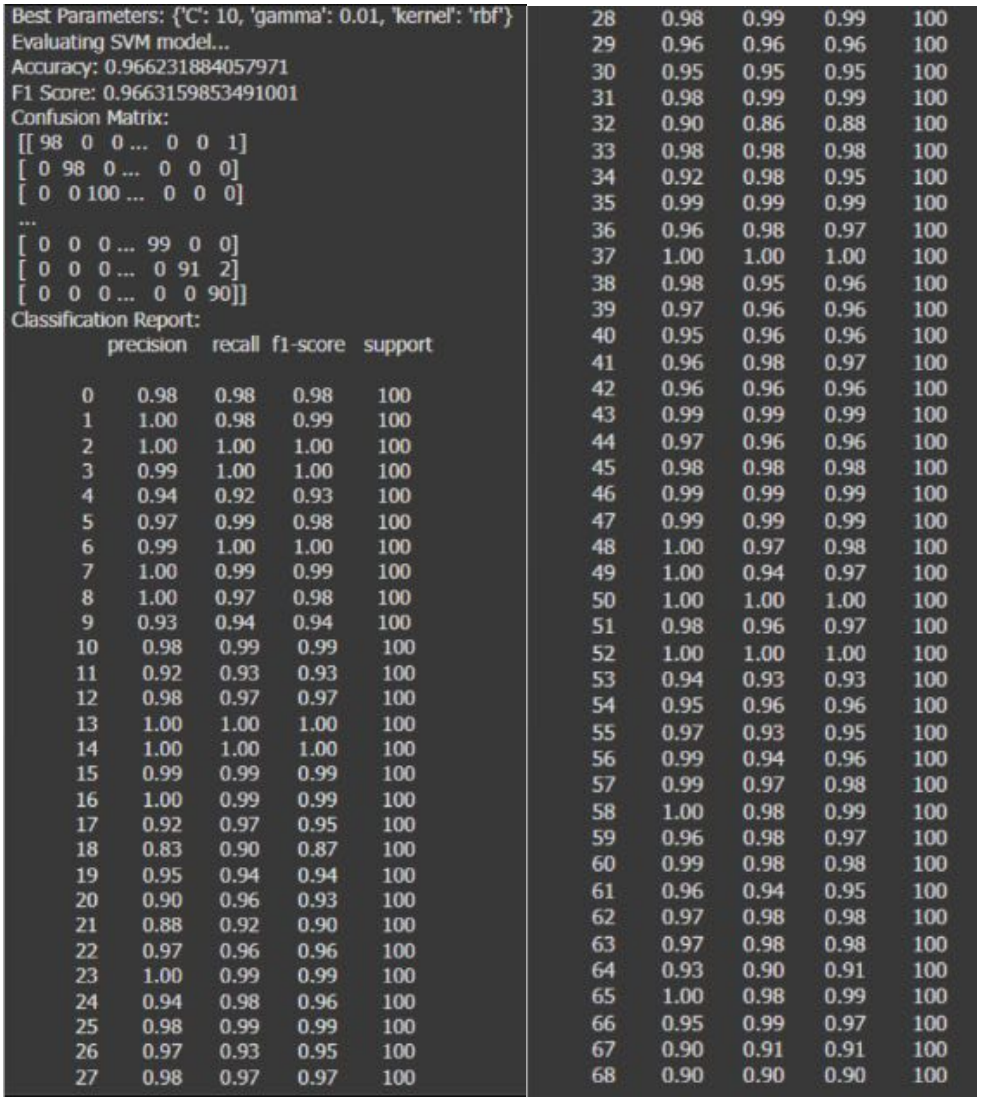
จากผลการทดลองพบว่า

- แบบจำลองการเรียนรู้ของเครื่อง SVM มีประสิทธิภาพมากที่สุด โดยมี Accuracy และ F1 score สูงถึงประมาณ 97 %
- แบบจำลองการเรียนรู้ของเครื่อง Random forest สามารถประมวลผลและระบุสมุนไพรจากภาพใบสมุนไพรได้เร็วที่สุด
- ภาษา Kotlin มี library ที่รองรับแบบจำลองการเรียนรู้ของเครื่อง โดยใช้ library : Scikit-learn ซึ่งจะรองรับแบบจำลองการเรียนรู้ของเครื่องที่เป็น Neuron Netwrok (NN) เท่านั้น จากข้อจำกัดในส่วนนี้

ถึงแม้ว่าแบบจำลอง SVM จะมีความแม่นยำสูงที่สุด แต่ไม่สามารถส่งออกเป็นไฟล์ TFLite ได้ ซึ่งจำเป็นสำหรับการพัฒนาแอปพลิเคชันด้วยภาษา Kotlin ในขณะที่ CNN มีความสามารถในการแปลงเป็นไฟล์ TFLite ได้ ทำให้เหมาะสมสำหรับการนำมาใช้ในแอปพลิเคชันแม้ค่าความแม่นยำจะต่ำกว่าเล็กน้อย ([6], [7])

Classification report ของการฝึกแบบจำลองการเรียนรู้ของเครื่องทั้ง 3 ชนิด

Classification ของแบบจำลองการเรียนรู้ของเครื่อง SVM สำหรับสมุนไพรแต่ละชนิด รวมถึงการแสดงผล accuracy กับ F1 score เฉลี่ย และเวลาในการประมวลผลต่อรูปภาพทั้งหมด รวมถึงพารามิเตอร์ที่ดีที่สุด



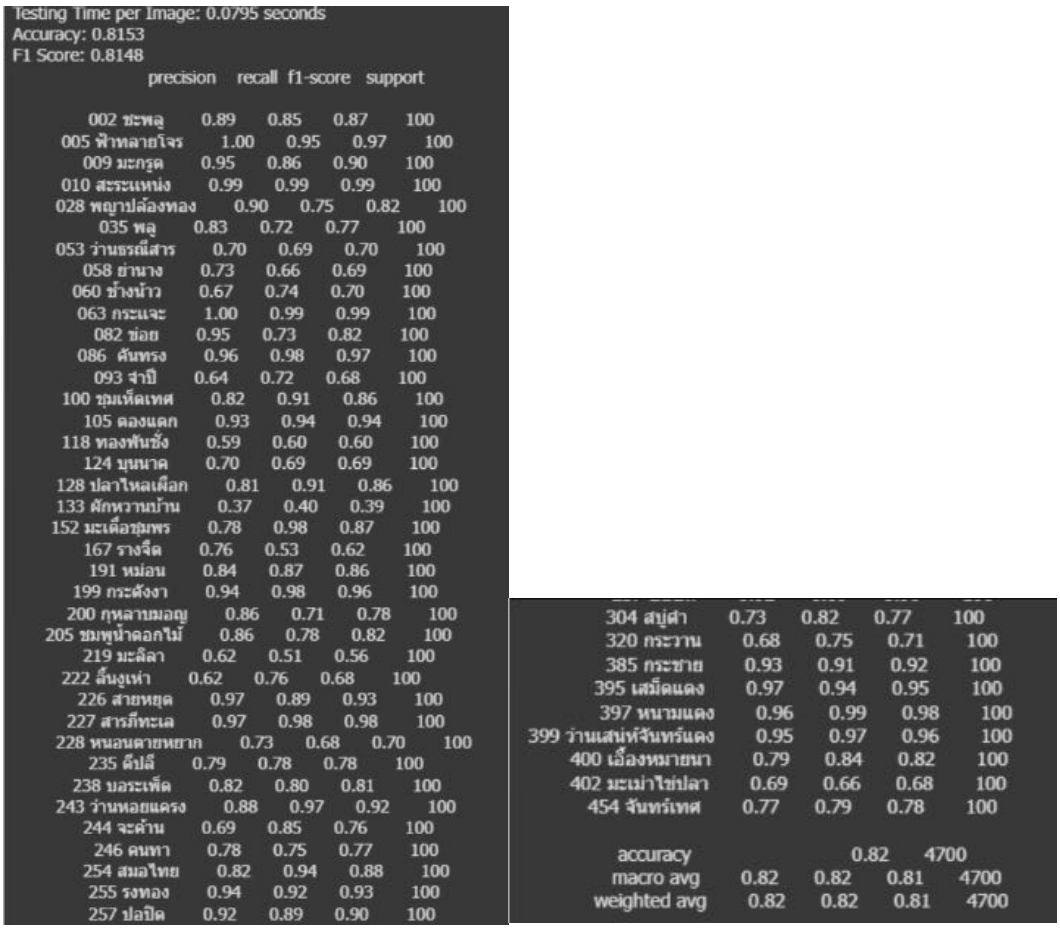
```
Best Parameters: {'C': 10, 'gamma': 0.01, 'kernel': 'rbf'}
Evaluating SVM model...
Accuracy: 0.966231884057971
F1 Score: 0.9663159853491001
Confusion Matrix:
[[ 98   0   0 ...   0   0   1]
 [  0  98   0 ...   0   0   0]
 [  0   0 100 ...   0   0   0]
 ...
 [  0   0   0 ...  99   0   0]
 [  0   0   0 ...   0  91   2]
 [  0   0   0 ...   0   0  90]]
Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.98      0.98       100
           1       1.00      0.98      0.99       100
           2       1.00      1.00      1.00       100
           3       0.99      1.00      1.00       100
           4       0.94      0.92      0.93       100
           5       0.97      0.99      0.98       100
           6       0.99      1.00      1.00       100
           7       1.00      0.99      0.99       100
           8       1.00      0.97      0.98       100
           9       0.93      0.94      0.94       100
          10       0.98      0.99      0.99       100
          11       0.92      0.93      0.93       100
          12       0.98      0.97      0.97       100
          13       1.00      1.00      1.00       100
          14       1.00      1.00      1.00       100
          15       0.99      0.99      0.99       100
          16       1.00      0.99      0.99       100
          17       0.92      0.97      0.95       100
          18       0.83      0.90      0.87       100
          19       0.95      0.94      0.94       100
          20       0.90      0.96      0.93       100
          21       0.88      0.92      0.90       100
          22       0.97      0.96      0.96       100
          23       1.00      0.99      0.99       100
          24       0.94      0.98      0.96       100
          25       0.98      0.99      0.99       100
          26       0.97      0.93      0.95       100
          27       0.98      0.97      0.97       100
          28       0.98      0.99      0.99       100
          29       0.96      0.96      0.96       100
          30       0.95      0.95      0.95       100
          31       0.98      0.99      0.99       100
          32       0.90      0.86      0.88       100
          33       0.98      0.98      0.98       100
          34       0.92      0.98      0.95       100
          35       0.99      0.99      0.99       100
          36       0.96      0.98      0.97       100
          37       1.00      1.00      1.00       100
          38       0.98      0.95      0.96       100
          39       0.97      0.96      0.96       100
          40       0.95      0.96      0.96       100
          41       0.96      0.98      0.97       100
          42       0.96      0.96      0.96       100
          43       0.99      0.99      0.99       100
          44       0.97      0.96      0.96       100
          45       0.98      0.98      0.98       100
          46       0.99      0.99      0.99       100
          47       0.99      0.99      0.99       100
          48       1.00      0.97      0.98       100
          49       1.00      0.94      0.97       100
          50       1.00      1.00      1.00       100
          51       0.98      0.96      0.97       100
          52       1.00      1.00      1.00       100
          53       0.94      0.93      0.93       100
          54       0.95      0.96      0.96       100
          55       0.97      0.93      0.95       100
          56       0.99      0.94      0.96       100
          57       0.99      0.97      0.98       100
          58       1.00      0.98      0.99       100
          59       0.96      0.98      0.97       100
          60       0.99      0.98      0.98       100
          61       0.96      0.94      0.95       100
          62       0.97      0.98      0.98       100
          63       0.97      0.98      0.98       100
          64       0.93      0.90      0.91       100
          65       1.00      0.98      0.99       100
          66       0.95      0.99      0.97       100
          67       0.90      0.91      0.91       100
          68       0.90      0.90      0.90       100

    accuracy                           0.97      6900
   macro avg       0.97      0.97      0.97      6900
weighted avg       0.97      0.97      0.97      6900

Testing Time: 97.84 seconds
```

Classification Report ของแบบจำลองการเรียนรู้ของเครื่อง Random Forest สำหรับสมุนไพรแต่ละชนิด รวมถึงการแสดงผล accuracy กับ F1 score เฉลี่ย และเวลาในการประมวลผลต่อภาพ



Testing Time per Image: 0.0795 seconds
Accuracy: 0.8153
F1 Score: 0.8148

	precision	recall	f1-score	support
002 ชะพลู	0.89	0.85	0.87	100
005 ฟ้าทลายโจร	1.00	0.95	0.97	100
009 มะกรูด	0.95	0.86	0.90	100
010 สะระแหน่ง	0.99	0.99	0.99	100
028 พญาปล้องทอง	0.90	0.75	0.82	100
035 พลู	0.83	0.72	0.77	100
053 ว่านธรณีสาร	0.70	0.69	0.70	100
058 ย่านาง	0.73	0.66	0.69	100
060 ข้างน้าว	0.67	0.74	0.70	100
063 กระแจะ	1.00	0.99	0.99	100
082 ข่อย	0.95	0.73	0.82	100
086 คันทรง	0.96	0.98	0.97	100
093 ฮาบิ	0.64	0.72	0.68	100
100 ชุมเห็ดเทศ	0.82	0.91	0.86	100
105 ตองแตก	0.93	0.94	0.94	100
118 ทองพันชั่ง	0.59	0.60	0.60	100
124 บุนนาค	0.70	0.69	0.69	100
128 ปลาไหลเผือก	0.81	0.91	0.86	100
133 ผักหวานบ้าน	0.37	0.40	0.39	100
152 มะเดื่อชุมพร	0.78	0.98	0.87	100
167 รางจืด	0.76	0.53	0.62	100
191 หม่อน	0.84	0.87	0.86	100
199 กระดังงา	0.94	0.98	0.96	100
200 กุหลาบมอญ	0.86	0.71	0.78	100
205 ชมพูน่ำดอกไม้	0.86	0.78	0.82	100
219 มะลิลา	0.62	0.51	0.56	100
222 ลิ้นงูเห่า	0.62	0.76	0.68	100
226 สายหยุด	0.97	0.89	0.93	100
227 สารภีทะเล	0.97	0.98	0.98	100
228 หนอนตายหยาก	0.73	0.68	0.70	100
235 ดีปลี	0.79	0.78	0.78	100
238 บอระเพ็ด	0.82	0.80	0.81	100
243 ว่านหอยแครง	0.88	0.97	0.92	100
244 จะค้าน	0.69	0.85	0.76	100
246 คนทา	0.78	0.75	0.77	100
254 สมอไทย	0.82	0.94	0.88	100
255 รงทอง	0.94	0.92	0.93	100
257 ปอปิด	0.92	0.89	0.90	100
304 สบู่ดำ	0.73	0.82	0.77	100
320 กระวาน	0.68	0.75	0.71	100
385 กระชาย	0.93	0.91	0.92	100
395 เสม็ดแดง	0.97	0.94	0.95	100
397 หนามแดง	0.96	0.99	0.98	100
399 ว่านเสน่หาจันทร์แดง	0.95	0.97	0.96	100
400 เอื้องหมายนา	0.79	0.84	0.82	100
402 มะเม่าใบปลา	0.69	0.66	0.68	100
454 จันทร์เทศ	0.77	0.79	0.78	100
accuracy			0.82	4700
macro avg	0.82	0.82	0.81	4700
weighted avg	0.82	0.82	0.81	4700

Classification Report ของแบบจำลองการเรียนรู้ของเครื่อง CNN สำหรับสมุนไพรแต่ละชนิด รวมถึงการแสดงผล accuracy กับ F1 score เฉลี่ย และเวลาในการประมวลผลต่อรูปภาพทั้งหมด รวมถึงพารามิเตอร์ที่ดีที่สุด

105 ตองแตก	0.97	0.91	0.94	100
118 ทองพันชั่ง	0.61	0.83	0.71	100
124 บุนนาค	0.89	0.70	0.78	100
128 ปลาไหลเผือก	0.96	0.98	0.97	100
133 ผักหวานบ้าน	0.72	0.52	0.60	100
152 มะเดื่อชุมพร	0.95	0.92	0.93	100
167 รางจืด	0.90	0.73	0.81	100
178 ว่านมหากาฬ	0.96	0.87	0.91	100
191 หม่อน	0.94	0.89	0.91	100
199 กระดังงา	0.93	0.99	0.96	100
200กุหลาบมอญ	0.92	0.83	0.87	100
205 ชมพู่น้ำดอกไม้	0.82	0.89	0.85	100
206 ชองระอา	0.93	0.95	0.94	100
208 ว่านตะขาบหิน	0.95	0.92	0.93	100
212 ผักชีล้อม	0.90	0.81	0.85	100
214 พลูคาว	0.95	0.97	0.96	100
219 มะลิลา	0.93	0.80	0.86	100
222 ลิ้นงูเห่า	0.89	0.92	0.91	100
226 สายหยุด	0.85	0.98	0.91	100
227 สารภีทะเล	0.96	0.96	0.96	100
228 หนอนตายหยาก	0.95	0.91	0.93	100
235 ดีปลี	0.79	0.96	0.86	100
237 หูเสือ	0.95	1.00	0.98	100
238 บอระเพ็ด	0.83	0.85	0.84	100
243 ว่านหอยแครง	0.93	1.00	0.97	100
244 จะค้าน	0.84	0.82	0.83	100
246 คนทา	0.92	0.79	0.85	100
254 สมอไทย	0.82	0.85	0.83	100
255 รงทอง	0.89	0.97	0.93	100
257 ปอปิด	0.83	0.94	0.88	100
304 สบู่ดำ	0.90	0.96	0.93	100
320 กระวาน	0.94	0.88	0.91	100
385 กระชาย	0.93	0.93	0.93	100
387 กระทือ	0.85	0.82	0.84	100
395 เสม็ดแดง	0.98	0.98	0.98	100
397 หนามแดง	0.97	0.90	0.93	100
398 รางพึง	0.72	0.86	0.79	100
399 ว่านเส่นห์จันทร์แดง	0.99	0.98	0.98	100
400 เอื้องหมายนา	0.85	0.94	0.90	100
402 มะเม่าไข่ปลา	0.74	0.82	0.78	100
454 จันทร์เทศ	0.77	0.67	0.72	100
accuracy			0.88	6900
macro avg	0.89	0.88	0.88	6900
weighted avg	0.89	0.88	0.88	6900

สำหรับการอ้างอิงประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องนั้น ทางผู้จัดทำขออ้างอิงข้อมูลการศึกษา Thai Herbal Image Identification: Herb ID จัดทำโดย กรมการแพทย์แผนไทยและแพทย์ทางเลือกกระทรวงสาธารณสุข ([22]) ซึ่งได้พัฒนาแอปพลิเคชันเพื่อระบุสมุนไพรจากภาพถ่าย ชื่อว่า Herb ID โดยแหล่งข้อมูลที่ใช้ในการฝึกแบบการจำลองการเรียนรู้ของเครื่องนั้นเป็นแหล่งเดียวกับที่ทางผู้จัดทำได้นำมาใช้เพื่อพัฒนาแบบจำลองการเรียนรู้ของเครื่อง

พบว่า แบบจำลองการเรียนรู้ของเครื่องของนายเจตสิกมีค่า accuracy 93.6% ซึ่งมีค่าน้อยกว่าของแบบจำลองการเรียนรู้ของเครื่อง SVM ของผู้คณะผู้จัดทำ ซึ่งมีค่า accuracy ที่ 96.62% แต่มากกว่าแบบจำลองการเรียนรู้ของเครื่อง Random Forest และ CNN ที่มีค่า accuracy ซึ่งมีค่า accuracy ที่ 88.47% และ 87.45% ตามลำดับ

4.2 ผลการพัฒนาต้นแบบแอปพลิเคชันต้นแบบ

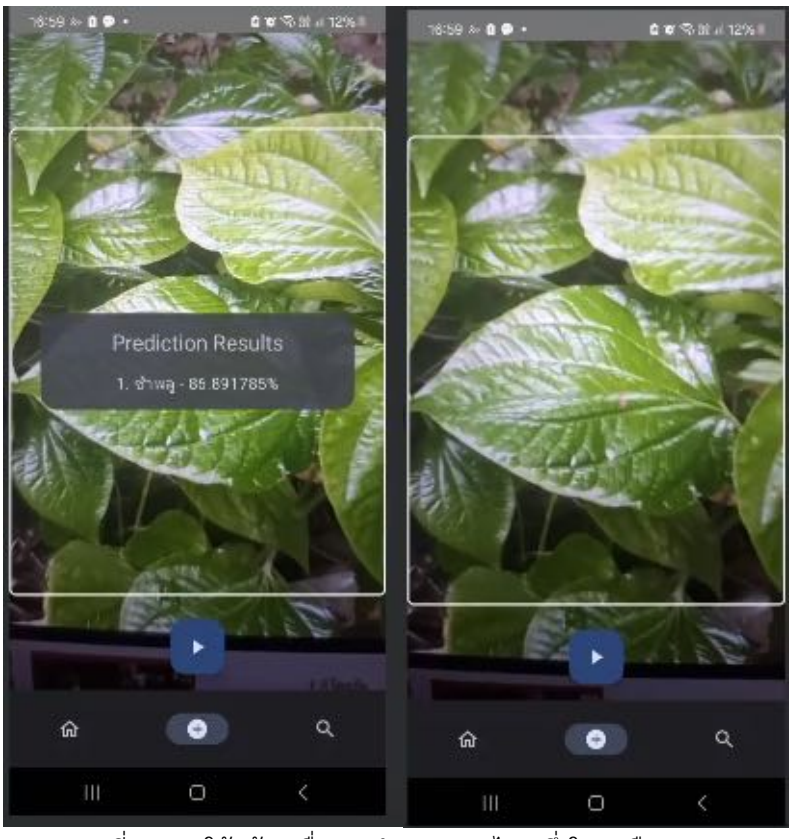
ต้นแบบแอปพลิเคชันแบบถูกพัฒนาขึ้นโดยใช้ภาษา Kotlin ผ่านแพลตฟอร์ม Android Studio โดยมีคุณลักษณะหลักดังนี้

5.3.1 ฟังก์ชัน User Interface (UI)



ภาพที่ 4: home screen

5.3.3 ฟังก์ชันการถ่ายภาพและจำแนกสมุนไพร



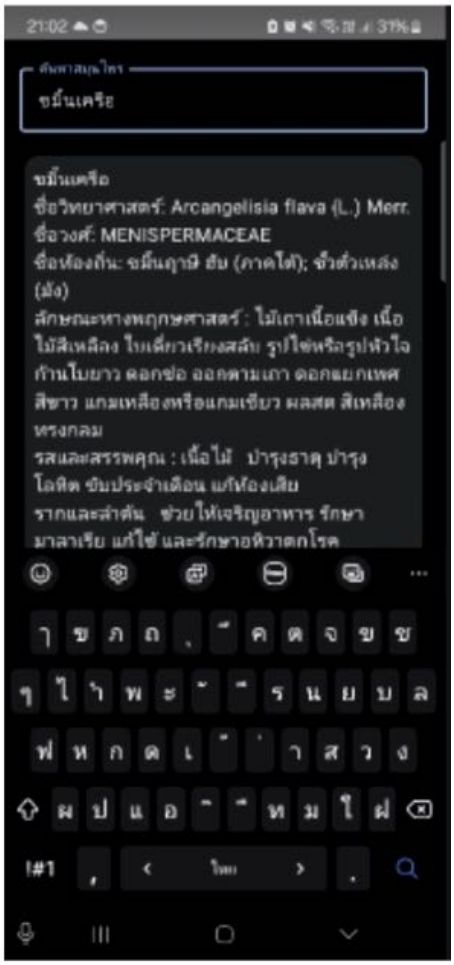
ภาพที่ 5: การใช้กล้องเพื่อการจำแนกสมุนไพร ซึ่งในรูปคือ ชะพลู

ผู้ใช้สามารถถ่ายภาพใบสมุนไพรและระบบจะทำการจำแนกประเภทโดยใช้แบบจำลอง CNN ซึ่งแสดงผลตามภาพที่ 6 ซึ่งผู้ใช้สามารถกดเครื่องหมายบวกในการเลือกใช้งานการถ่ายรูปในบริเวณตรงกลางของขอบล่าง แล้วจึงกดเครื่องหมาย

สามเหลี่ยมเพื่อให้แอปพลิเคชันถ่ายภาพและประมวลผล จากนั้นแอปพลิเคชันจึงแสดงชื่อสมุนไพรที่มีค่า accuracy มากที่สุด พร้อมกับค่า accuracy นั้น ๆ

5.3.4 ฟังก์ชันการค้นหาข้อมูลสมุนไพรและอันตรกิริยาของสมุนไพร

ผู้ใช้งานสามารถค้นหาข้อมูลได้ทั้งจากข้อความ **(Text Search)** และตามหมวดหมู่ **(Category Search)** ตัวอย่างเช่น ต้องการหาข้อมูลของ ขมิ้นเครือก็สามารถพิมพ์ข้อความได้เลย ซึ่งผู้ใช้สามารถกดเครื่องหมายรูปค้นหาในบริเวณส่วนขวาของขอบล่างเพื่อเลือกใช้งานการค้นหาข้อมูลของสมุนไพรที่ต้องการ แล้วจึงพิมพ์ชื่อสมุนไพรที่ต้องการค้นหาลงในแถบ search bar ในบริเวณด้านบน จากนั้นแอปพลิเคชันจะแสดงข้อมูลของสมุนไพร เช่น ชื่อ, ชื่อวิทยาศาสตร์ และ ลักษณะทางพฤกษศาสตร์ เป็นต้น ดังภาพที่ 6



ภาพที่ 6: ตัวอย่างการค้นหาข้อมูลขมิ้นเครือ

6 การอภิปรายผลและการเสนอแนะ

จากผลการศึกษา และวิเคราะห์เปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องในการทำนายชนิดสมุนไพรจากภาพถ่ายใบสมุนไพร ซึ่งเป็นข้อมูลที่ได้รับการอนุเคราะห์จากกรมการแพทย์แผนไทยและการแพทย์ทางเลือก โดยสร้างแบบจำลอง ด้วยวิธีที่แตกต่างกัน 3 วิธี ได้แก่ Support Vector Machine (SVM) Random forest และ โครงข่ายประสาทแบบคอนโวลูชัน (CNN) แบ่งข้อมูลออกเป็นชุดฝึกและชุดทดสอบ ในอัตราส่วน 80:20 สร้างแบบจำลองด้วยชุดข้อมูลฝึก และวัดผลการเปรียบเทียบการทำนายชนิดสมุนไพรโดยใช้ เวลาที่ใช้ทดสอบกับชุดทดสอบ ค่าความแม่นยำ (Accuracy) และ F1-score พบว่า แบบจำลองการเรียนรู้ของเครื่องแบบ SVM ใช้เวลา 97.84 วินาทีต่อ 6900 ภาพ Accuracy 96.62% และ F1-score 96.63% แบบ Random forest ใช้เวลา 0.83 วินาทีต่อ 6900 ภาพ Accuracy 88.47% และ F1-score 88.29% แบบ CNN ใช้เวลา 582.36 วินาทีต่อ 6900 ภาพ Accuracy 87.45% และ F1-score 87.49% สรุปผลการทดลองพบว่า แบบ SVM เป็นวิธีที่มีคาวมแม่นยำในการ

ทำนายสูงที่สุด และแบบ Random forest เป็นวิธีที่ประมวลผลได้เร็วที่สุด ต่อมาผู้วิจัยได้นำแบบจำลองการเรียนรู้ของเครื่องไปใช้กับแอปพลิเคชัน ทำให้พบว่า SVM และ Random forest ไม่สามารถใช้งานร่วมกับแอปพลิเคชันที่พัฒนาด้วย Kotlin ได้ เนื่องจากไม่มีไลบรารี scikit-learn ทำให้ผู้วิจัยใช้แบบจำลองการเรียนรู้ของเครื่องแบบ CNN และพัฒนาระบบค้นหาของแอปพลิเคชันด้วยการนำฐานข้อมูลที่รวบรวมจากสถานบันหรือหน่วยงานที่มีความน่าเชื่อถือ ดังนั้นแอปพลิเคชันต้นแบบมีทั้งหมด 2 ฟังก์ชันหลัก ได้แก่

1.การทำนายสมุนไพรจากภาพถ่ายใบสมุนไพร สามารถทำนายได้ทั้งหมด 69 ชนิด

2.การค้นหาสมุนไพรโดยมีข้อมูลทั่วไปของสมุนไพร 273 ชนิด และ ข้อมูลอันตรกิริยาระหว่างแผนยาปัจจุบันกับสมุนไพร 65 ชนิด

ดังนั้น ข้อเสนอแนะสำหรับการพัฒนาต่อยอดในอนาคต แม้แอปพลิเคชันต้นแบบจะสามารถใช้งานได้จริงและตอบสนองต่อวัตถุประสงค์หลักของงานวิจัย แต่อาจมีจุดที่สามารถพัฒนาเพิ่มเติมเพื่อเพิ่มประสิทธิภาพและความน่าสนใจของระบบ ดังนี้

ด้านการพัฒนาโมเดล

- ควรมีการทดลองใช้โมเดลอื่นเพิ่มเติม เช่น EfficientNet และ Vision Transformer (ViT) เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกภาพสมุนไพรในบริบทที่หลากหลาย นอกจากนี้ควรนำเทคนิค Data Augmentation มาใช้เพิ่มเติม เพื่อเพิ่มความหลากหลายของข้อมูลฝึกให้ครอบคลุมสภาพแวดล้อมที่แตกต่างกัน เช่น การถ่ายภาพในแสงธรรมชาติ แสงน้อย ฉากหลังรก หรือการถ่ายจากมุมเอียง ซึ่งเป็นสถานการณ์ที่อาจพบเจอได้จริงในการใช้งานของผู้ใช้ทั่วไป
- ควรคำนึงถึงกรณีที่ผู้ใช้ถ่ายภาพสมุนไพรในสภาพที่ใบไม้ซ้อนทับกัน หรือมีวัตถุอื่นในภาพ เช่น มือหรือพื้นหลังของสนามหญ้า ซึ่งอาจเป็นอุปสรรคต่อความแม่นยำของระบบ โดยสามารถแก้ปัญหาเหล่านี้ได้ด้วยการนำเทคนิค **Image Segmentation** เช่น การใช้โมเดล U-Net หรือ U²-Net เข้ามาช่วยแยกเฉพาะบริเวณใบพืชออกจากพื้นหลัง ก่อนนำไปประมวลผลโดยโมเดลจำแนกชนิดพืช ทั้งนี้จะช่วยให้โมเดลมีความยืดหยุ่นและแม่นยำมากยิ่งขึ้นเมื่อนำไปใช้งานนอกห้องปฏิบัติการหรือพื้นที่ควบคุม
- แบบจำลองการเรียนรู้ของเครื่อง SVM และ Random Forest ไม่สามารถส่งออกไฟล์ในสกุล TFLite ได้ หากผู้สนใจต้องการจะพัฒนาต่อยอดโดยใช้ภาษา Kotlin ควรจะเลือกแบบจำลองการเรียนรู้ของเครื่องที่มีการพัฒนาต่อยอดโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural network) เช่น MobileNet, EfficientNet, VGGNet เป็นต้น

ด้านข้อมูล

สถานที่สำหรับขยายฐานข้อมูลเกี่ยวกับข้อมูลทั่วไปสมุนไพร เช่น สวนสิริรุกขชาติ มหาวิทยาลัยมหิดล เป็นต้น

ด้านการใช้งานจริง

ควรนำแอปพลิเคชันไปทดสอบกับกลุ่มเป้าหมายที่หลากหลาย เช่น แพทย์แผนไทย นักศึกษา หรือประชาชนทั่วไป พร้อมเก็บข้อมูลความพึงพอใจและข้อเสนอแนะด้านการใช้งาน (UX/UI) เพื่อนำไปปรับปรุงให้เหมาะสมกับผู้ใช้จริงในบริบทที่หลากหลายมากยิ่งขึ้น

กิตติกรรมประกาศ คณะผู้วิจัยขอขอบคุณผู้ทรงคุณวุฒิทุกท่านที่ได้ให้ข้อคิดเห็นและข้อเสนอแนะต่าง ๆ เพื่อปรับปรุงบทความวิจัย และขอขอบคุณ กรมการแพทย์แผนไทยและการแพทย์ทางเลือก ที่ให้คำปรึกษาเกี่ยวกับการสร้างแอปพลิเคชันและอนุญาตให้ใช้ข้อมูลของสมุนไพรในประเทศไทย และคณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล ที่อนุญาตให้ใช้ข้อมูลเกี่ยวกับสมุนไพรไทยและยาแผนปัจจุบันที่มีอันตรกิริยากับสมุนไพรไทย และโรงเรียนมหิดลวิทยานุสรณ์ สำหรับการสนับสนุนการทำวิจัยในครั้งนี้

เอกสารอ้างอิง

[1] Bambil, D., et al. (2020). *Plant species identification using color learning resources, shape, texture, through machine learning and artificial neural networks*. Environmental Systems and Decisions, 40(4), 480–484. https://doi.org/10.1007/s10669-020-09769-w

[2] Breiman, L. (2001). *Random forests*. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

[3] Chaikhamwang, S., Montri, V., & Chailangkan, Y. (2021). *Thai herbs application*. International Journal of Industrial Education and Technology, 3(1), January–December.

[4] Chan, T. Y. (2001). *Interaction between Warfarin and Danshen (Salvia Miltiorrhiza)*. Annals of Pharmacotherapy, 35(4), 501–504.

[5] Cortes, C., & Vapnik, V. (1995). *Support-vector networks*. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018

[6] Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.

[7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[8] Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., ... & Le, Q. V. (2019). *Searching for MobileNetV3*. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 1314–1324). https://doi.org/10.1109/ICCV.2019.00140

[9] Kavitha, S., Kumar, T. S., Naresh, E., Kalmani, V. H., Bamane, K. D., & Pareek, P. K. (2024). *Medicinal plant identification in real-time using deep learning model*. SN Computer Science, 5(1), Article 73. https://doi.org/10.1007/s42979-023-02398-5

[10] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). *Gradient-based learning applied to document recognition*. Proceedings of the IEEE, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

[11] Liaw, A., & Wiener, M. (2002). *Classification and regression by randomForest*. R News, 2(3), 18–22. https://CRAN.R-project.org/doc/Rnews/

[12] Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.

[13] Moghaddam, R. F., & Cheriet, M. (2012). *Application of multi-level classifiers and clustering for automatic script identification*. Pattern Recognition, 45(12), 4158–4168.

[14] Muneer, A., et al. (2020). *Efficient and automated herbs classification approach based on shape and texture features using deep learning*. IEEE Xplore.

[15] Noble, W. S. (2006). *What is a support vector machine?* Nature Biotechnology, 24(12), 1565–1567. https://doi.org/10.1038/nbt1206-1565

[16] Powers, D. M. W. (2011). *Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation*. Journal of Machine Learning Technologies, 2(1), 37–63.

[17] Putra, R. P. S., et al. (2024). *Herbal leaf classification using deep learning model EfficientNetV2B0*. Jurnal Ilmiah Teknologi dan Komputer, 9(2), 301–307. https://doi.org/10.33480/jitk.v9i2.5119

[18] Schwaiblmair, M. (2012). *Drug induced interstitial lung disease*. Open Respiratory Medicine Journal, 6(1), 63–74. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3415629/

[19] Shirali, A., & Pazhayattil, G. S. (2014). *Drug-induced impairment of renal function*. International Journal of Nephrology and Renovascular Disease, 7, 457. https://www.dovepress.com/drug-induced-impairment-of-renal-function-peer-reviewed-fulltext-article-IJNRD

[20] Sokolova, M., & Lapalme, G. (2009). *A systematic analysis of performance measures for classification tasks*. Information Processing & Management, 45(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

[21] Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking model scaling for convolutional neural networks*. In Proceedings of the 36th International Conference on Machine Learning (Vol. 97, pp. 6105–6114). PMLR. https://proceedings.mlr.press/v97/tan19a.html

[22] Thong, B. Y.-H., & Tan, T. C. (2011). *Epidemiology and risk factors for drug allergy*. British Journal of Clinical Pharmacology, 71(5), 684–700. https://bpspubs.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2125.2010.03774.x

[23] Tutorialspoint. (n.d.). *Machine learning - Confusion matrix*. Tutorialspoint. Retrieved April 16, 2025, from https://www.tutorialspoint.com/machine_learning/machine_learning_confusion_matrix.htm

[24] Yalcin, H., & Razavi, S. (2016). *Plant classification using convolutional neural networks*. In 2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics). IEEE.

[25] พิมลพรรณ อนันต์กิจไพศาล. (2564). *108 สมุนไพรไทย ใช้เป็น หายป่วย ตำรับยาอายุวัฒนะ กัน-แก้สารพัดโรค จากครัวเรือน* (พิมพ์ครั้งที่ 4). กรุงเทพฯ: รู้รอบ พับลิชชิ่ง.

[26] โพธิ์พันธุ์ เจตสิก. (2021). *Thai Herbal Image Identification: Herb ID*. กรมการแพทย์แผนไทยและการแพทย์ทางเลือก กระทรวงสาธารณสุข.

[27] สุพจน์ เฮงพระพรหม. (2567). *การเรียนรู้ของเครื่อง (Machine Learning)* (พิมพ์ครั้งที่ 2). นครปฐม: สำนักพิมพ์มหาวิทยาลัยศิลปากร.

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

Properties of the Double Domination Numbers of Some Glued Paths

Punvarot Rattakorn[1,*] and Piyashat Sripratak[2,†]

[1]Graduate Master Degree Program in Teaching Mathematics, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

[2]Department of Mathematics, Faculty of Science
Chiang Mai University, Chiang Mai 50200, Thailand

Abstract

Let G be a graph with vertex set $V(G)$. For $D^2 \subseteq V(G)$, if every vertex $v \in V(G)$ is dominated by at least two vertices in D^2, then D^2 is a *double dominating set* for G. The *double domination number* is the minimum cardinality among those of all double dominating sets of G, denoted by $dd(G)$. In this work, we establish the double domination numbers of some glued graphs whose original graphs are P_n and P_n, where n is a positive integer such that $n \geq 3$, at clones H_1 and H_2 that are connected subgraphs of P_n which are isomorphic to a path of length not less than $n - 6$.

Keywords: double domination number, glued graph, path
2020 MSC: 05C38, 05C69, 05C76

1 Introduction

Glued graphs were introduced by Uiyyasathian [1]. Promsakorn and Uiyyasathian studied the chromatic numbers [2] and edge-chromatic numbers of glued graphs [3]. Uiyyasathian and Saduakdee [4] focused on perfect glued graphs at complete clones. Clique partitions of glued graphs were explored by Jongthawonwuth and Uiyyasathian [5]. Eulerian glued graphs were studied by Boonthong et al. [6].

The definition of domination numbers was established by Berge in 1958 [7]. In 1998, the concept of a dominating set in a graph was described by Haynes et al. [8]. Alikhani et al. [9] worked on the domination numbers of some graphs such as trees. Glebov et al. [10] studied the domination numbers of random graph. The domination numbers of the generalized Petersen graphs were investigated by Behzad et al. [11].

The Domination numbers of glued graphs has been studied by several researchers. Boonmee et al. [12] worked on the domination numbers of glued graphs whose original graphs are C_m and C_n. The domination numbers of glued graphs containing P_n and P_m as the original graphs were investigated by Ruangnai and Panma [13]. Sripratak and Panma [14] presented a lower bound and an upper bound of the domination numbers of glued graphs.

*Speaker (punvarot.r@gmail.com)
†Corresponding author (psripratak@gmail.com)

The concept of double domination in graphs was first defined by Harary and Haynes in 2000 [15]. Blidia et al. [16] worked on the double domination numbers of trees. Henning [17] established an upper bound on the double domination number of a connected graph with minimum degree at least two that is not a C_5. Hajian and Rad [18] presented a lower bound for the double domination numbers of any connected graphs. Double domination number in shadow graphs was studied by Aytaç and Mutlu [20].

To the best of our knowledge, there are no publications on the double domination numbers of glued graphs.

2 Preliminaries

There have been many research topics on operations between two graphs. We consider an operation called *glue operator*, first introduced by Uiyyasathian for solving the maximal-clique partition problems. Let G_1 and G_2 be any two graphs with disjoint vertex sets. Let $H_1 \subseteq G_1$ and $H_2 \subseteq G_2$ be non-empty connected subgraphs such that $H_1 \cong H_2$ with an isomorphism f. *The glued graph of* G_1 *and* G_2 *at* H_1 *and* H_2 *with respect to* f, denoted by $G_1 \lhd_{H_1 \cong_f H_2} \rhd G_2$ is the graph that resulted from identifying H_1 and H_2 under the isomorphism f. Let H be the copy of H_1 and H_2 in the glued graph, we refer to H as the *clones* of the glued graph and G_1 and G_2 as the *original graphs*. The glued graph of G_1 and G_2 at the clone H, written as $G_1 \lhd\rhd G_2$, means that there exists a subgraph H_1 of G_1, a subgraph H_2 of G_2 and an isomorphism f between H_1 and H_2 such that $G_1 \lhd_{H_1 \cong_f H_2} \rhd G_2$ and H is the copy of H_1 and H_2 in the resulting graph. The importance of the gluing operation is that this operation creates a class of graphs that each of them contains the vertices and edges of both original graphs. Results obtained on original graphs can be extended or applied to the larger glued graph. Therefore, this operation allows researchers to leverage existing knowledge and findings to understand the properties and behaviors of more complex graphs.

Among several graph parameters, domination number is one of the most studied parameters in graph theory. It was introduced by Berge in 1958. Let G be a graph. Denote $N[v] = \{u \in V(G) : u = v \text{ or } uv \in E(G)\}$ the *closed neighborhood* of $v \in V(G)$. For vertices $u, v \in V(G)$, vertex v *dominates* vertex u if $u \in N[v]$. A *dominating set* D of graph G is a set of vertices such that for any vertex v in $V(G)$, there is a vertex in D that dominates v. The *domination number*, denoted by $\gamma(G)$, is the minimum cardinality of a dominating set in G. Domination number produces applications in various fields. For example, the concepts and properties of domination is used to solve the queens problem in chess, locate radar stations problem, model social networks, locate facilities that help minimizing cost and providing service, construct a route for school bus, or apply in coding theory [19].

The concept of double domination was introduced by Harary and Haynes in 2000 [15]. Let G be a graph. A set $D^2 \subseteq V(G)$ is a *double dominating set* for G if for any vertex $v \in V(G)$, vertex v is dominated by at least two vertices in D^2. The double domination number, denoted by $dd(G)$, is the minimum cardinality among those of all double dominating sets of G. Double domination number also rich in applications, for example, prisoners and guards problem. The concept of domination is that each prisoner can be seen by some guards. Double domination increases security by requiring that each prisoner is assigned to be monitored by at least two guards.

We consider double domination numbers of glued paths. A *path* is a graph with vertex set $V(P_n) = \{v_1, v_2, \ldots, v_n\}$ and edge set $E(P_n) = \{\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{n-1}, v_n\}\}$ where $n \in \mathbb{N}$.

This path can be represented by $P_n = v_1v_2 \dots v_n$.

In this work, our goal is to offer lower and upper bounds for double domination numbers of some glued path.

3 Main Results

In this section, we offer our main results: the double domination number of some glued paths. For this purpose, we first introduce additional facts which are useful in this research work.

Theorem 3.1. [20] *For $n \geq 2$, $dd(P_n) = \left\lceil \frac{2n+2}{3} \right\rceil$.*

Observation 3.2. [21] *Every double dominating set of a graph contains all vertices of degree one and vertices that are adjacent to them.*

Next, we establish our observation and lemmas that we will use to prove our main theorems. The proofs of lemmas are simple and straightforward, so we omit the proofs.

Observation 3.3. *For any pair of vertices u and v that are adjacent in D^2, u is dominated by at least two vertices which are itself and v. The same goes for v.*

Lemma 3.4. *Let D^2 be a double dominating set of path P_n. For every two adjacent vertices, there must be at least one vertex in D^2.*

Lemma 3.5. *Let D^2 be a double dominating set of path P_n. For $i = 2, 3, \dots, n-1$, let $v_{i-1}, v_i, v_{i+1} \in V(P_n)$, at least two vertices in $\{v_{i-1}, v_i, v_{i+1}\}$ are in D^2.*

Let $P^1 \lhd_{H^1 \cong_f H^2} \rhd P^2$ be a glued graph of $P^1 = P_n$ and $P^2 = P_n$ at clones H^1 and H^2 with respect to isomorphism f, where H^1 and H^2 are connected subgraphs that are isomorphic to P^1 and P^2, respectively. Assume that clone H_1 is a path $v_iv_{i+1} \dots v_j$. We define $L(P^1)$ a path $v_1v_2 \dots v_i$ and $R(P^1)$ a path $v_jv_{j+1} \dots v_n$. We can define $L(P^2)$ and $R(P^2)$ in a similar way. Let $|L(P^1)|$ be a number of edges of $L(P^1)$ and we can define $|L(P^2)|$, $|R(P^1)|$ and $|R(P^2)|$ in a similar way.

We consider when $1 \leq |L(P^1)|, |L(P^2)|, |R(P^1)|, |R(P^2)| \leq 3$. Let (a, b, c, d) be a non-decreasing sequence of $|L(P^1)|, |L(P^2)|, |R(P^1)|$, and $|R(P^2)|$. The graph $G(a, b, c, d)$ is the glued graph $P^1 \lhd_{H^1 \cong_f H^2} \rhd P^2$ with (a, b, c, d) sequence. Let $A = (a_1, a_2, a_3, a_4)$ and $B = (b_1, b_2, b_3, b_4)$. We say that $A \leq B$ if and only if $a_1 < b_1$ or there is an integer $k \in \{1, 2, 3\}$ that $a_1 = a_2 = \dots = a_k = b_1 = b_2 = \dots = b_k$ and $a_{k+1} \leq b_{k+1}$.

In this paper we consider only the cases when $(a, b, c, d) \leq (2, 2, 3, 3)$, and we do not include the cases when $|L(P^1)| = |L(P^2)| = 3$ or when $|R(P^1)| = |R(P^2)| = 3$.

Let $P^1 = P_n = u_1u_2 \dots u_n$ and $P^2 = P_n = v_1v_2 \dots v_n$. We define seven types of glued graphs depending on H^1, H^2, and isomorphism f as follows.

1. $G(1, 1, 1, 1)$
   - $H^1 = u_2u_3 \dots u_{n-1}$ and $H^2 = v_2v_3 \dots v_{n-1}$
   - f is defined by $f(u_i) = v_i$ for all $i = 2, 3, \dots, n-1$

2. $G(1,1,2,2)$ 1st kind (denoted by $G^1(1,1,2,2)$)
   - $H^1 = u_2u_3\dots u_{n-2}$ and $H^2 = v_3v_4\dots v_{n-1}$
   - f is defined by $f(u_i) = v_{i+1}$ for all $i = 2,3,\dots,n-2$
3. $G(1,1,2,2)$ 2nd kind (denoted by $G^2(1,1,2,2)$)
   - $H^1 = u_2u_3\dots u_{n-2}$ and $H^2 = v_2v_3\dots v_{n-2}$
   - f is defined by $f(u_i) = v_i$ for all $i = 2,3,\dots,n-2$
4. $G(1,1,3,3)$
   - $H^1 = u_2u_3\dots u_{n-3}$ and $H^2 = v_4v_5\dots v_{n-1}$
   - f is defined by $f(u_i) = v_{i+2}$ for all $i = 2,3,\dots,n-3$
5. $G(1,2,2,3)$
   - $H^1 = u_2u_3\dots u_{n-3}$ and $H^2 = v_3v_4\dots v_{n-2}$
   - f is defined by $f(u_i) = v_{i+1}$ for all $i = 2,3,\dots,n-3$
6. $G(2,2,2,2)$
   - $H^1 = u_3u_4\dots u_{n-2}$ and $H^2 = v_3v_4\dots v_{n-2}$
   - f is defined by $f(u_i) = v_i$ for all $i = 3,4,\dots,n-2$
7. $G(2,2,3,3)$
   - $H^1 = u_3u_4\dots u_{n-3}$ and $H^2 = v_4v_5\dots v_{n-2}$
   - f is defined by $f(u_i) = v_{i+1}$ for all $i = 3,4,\dots,n-3$

For example, let $P^1 \lhd_{H^1\cong_f H^2}\rhd P^2$ be a glued graph of $P^1 = P_9 = u_1u_2\dots u_9$ and $P^2 = P_9 = v_1v_2\dots v_9$ at clones $H^1 = u_2u_3\dots u_6$ and $H^2 = v_3v_4\dots v_7$ with respect to $f = \{(u_2,v_3),(u_3,v_4),(u_4,v_5),(u_5,v_6),(u_6,v_7)\}$. We get $L(P^1) = u_1u_2$, $R(P^1) = u_6u_7u_8u_9$, $L(P^2) = v_1v_2v_3$, $R(P^2) = v_7v_8v_9$, $|L(P^1)| = 1$, $|R(P^1)| = 3$, $|L(P^2)| = 2$, and $|R(P^2)| = 2$. Hence, the glued graph $P^1 \lhd_{H^1\cong_f H^2}\rhd P^2$ is in the type $G(1,2,2,3)$ (see Figure 1).

Next, we investigate the double domination numbers of these seven types of graphs.

Theorem 3.6. *Let $n \geq 3$ be a positive integer. The double domination number of a graph in type $G(1,1,1,1)$ is $dd(P_n)+2$.*

Proof. Let G be a graph in type $G(1,1,1,1)$. The double domination number of G is denoted by $dd(G(1,1,1,1))$.

For the lower bound, let $D^2_{G(1,1,1,1)}$ be a double dominating set of size $dd(G(1,1,1,1))$ of G. By Observation 3.2, vertices u_1, v_1, u_n, and v_n are in $D^2_{G(1,1,1,1)}$. Suppose for a contradiction that $|D^2_{G(1,1,1,1)}| < dd(P_n)+2$.

Let $D^2_k = D^2_{G(1,1,1,1)}\backslash\{v_1,v_n\}$. It is easy to check that D^2_k is a double dominating set of P^1.

Since $\{v_1,v_n\} \subseteq D^2_{G(1,1,1,1)}$, we get

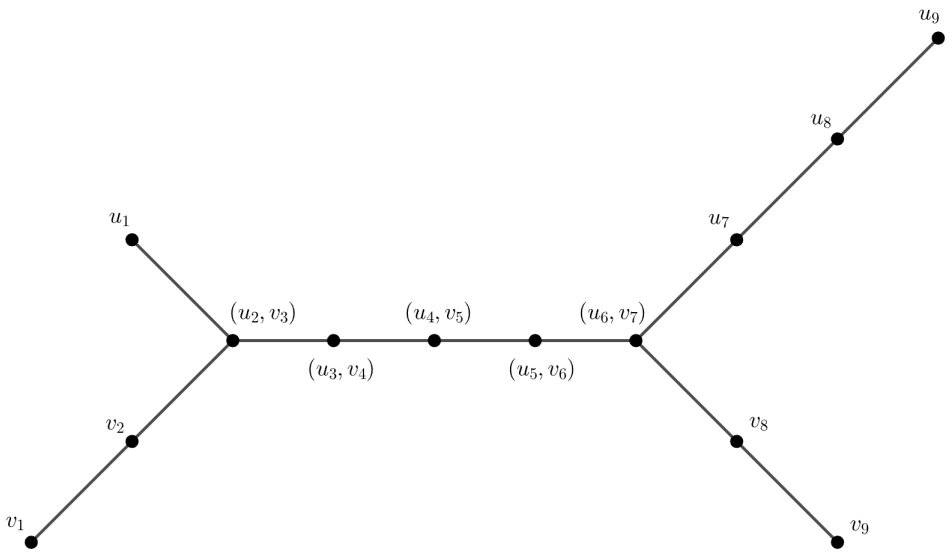


Figure 1: A glued graph in type $G(1, 2, 2, 3)$

$$|D_k^2| = |D^2_{G(1,1,1,1)}| - |\{v_1, v_n\}| < dd(P^1) + 2 - 2 = dd(P^1) = dd(P_n).$$

It contradicts to the fact that $dd(P_n)$ is the minimum cardinality among those of all double dominating sets of P_n. Hence, $dd(G(1, 1, 1, 1)) = |D^2_{G(1,1,1,1)}| \geq dd(P_n) + 2$.

For the upper bound, we construct a double dominating set of a graph in type $G(1, 1, 1, 1)$. Define

$$D^2 = \{u_1, v_1, (u_2, v_2), (u_{n-1}, v_{n-1}), u_n, v_n\} \cup \{(u_i, v_i) | i \not\equiv 0 (\text{mod} 3) \text{ and } i \in \{3, 4, \ldots, n-2\}\}.$$

Case $n \equiv 0 \pmod 3$**:** We get $\left\lceil \frac{2n+2}{3} \right\rceil = \frac{2n+3}{3}$. Therefore,

$$|D^2| = 6 + \left((n-4) - \left(\frac{n-3}{3}\right)\right) = \frac{2n+3}{3} + 2 = \left\lceil \frac{2n+2}{3} \right\rceil + 2.$$

Case $n \equiv 1 \pmod 3$**:** We get $\left\lceil \frac{2n+2}{3} \right\rceil = \frac{2n+4}{3}$. Therefore,

$$|D^2| = 6 + \left((n-4) - \left(\frac{n-4}{3}\right)\right) = \frac{2n+4}{3} + 2 = \left\lceil \frac{2n+2}{3} \right\rceil + 2.$$

Case $n \equiv 2 \pmod 3$**:** We get $\left\lceil \frac{2n+2}{3} \right\rceil = \frac{2n+2}{3}$. Therefore,

$$|D^2| = 6 + \left((n-4) - \left(\frac{n-2}{3}\right)\right) = \frac{2n+2}{3} + 2 = \left\lceil \frac{2n+2}{3} \right\rceil + 2.$$

For $i = 3, 4, \ldots, n-2$,

if $i \equiv 0 \pmod 3$, vertex (u_i, v_i) is dominated by (u_{i-1}, v_{i-1}) and (u_{i+1}, v_{i+1}),
if $i \equiv 1 \pmod 3$, vertex (u_i, v_i) is dominated by (u_i, v_i) and (u_{i+1}, v_{i+1}), and
if $i \equiv 2 \pmod 3$, vertex (u_i, v_i) is dominated by (u_{i-1}, v_{i-1}) and (u_i, v_i).

Since u_1 is adjacent to (u_2, v_2), (u_{n-1}, v_{n-1}) is adjacent to u_n, and all four vertices are in D^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_1 is dominated by v_1 and (u_2, v_2), and v_n is dominated by v_n and (u_{n-1}, v_{n-1}). Then, any vertex in a graph in type $G(1,1,1,1)$ is dominated by at least two vertices in D^2. Therefore, D^2 is a double dominating set of a graph in type $G(1,1,1,1)$. Hence, $dd(G(1,1,1,1)) \leq |D^2| = \left\lceil \frac{2n+2}{3} \right\rceil + 2 = dd(P_n) + 2$. □

Theorem 3.7. *Let $n \geq 4$ be a positive integer. The double domination number of a graph in type $G^1(1,1,2,2)$ is $dd(P_n) + 4$.*

Proof. Let G be a graph in type $G^1(1,1,2,2)$. The double domination number of G is denoted by $dd(G^1(1,1,2,2))$. We can see that $G = P^1 \lhd_{H^1 \cong_f H^2} \rhd P^2$ where $H^1 = u_2u_3 \dots u_{n-2}$, $H^2 = v_3v_4 \dots v_{n-1}$, and $f(u_i) = v_{i+1}$ for all $i = 2, 3, \dots, n-2$ is isomorphic to $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$, where $H_1 = u_2u_3 \dots u_{n-1}$, $H_2 = v_3v_4 \dots v_n$, and $f(u_i) = v_{i+1}$ for all $i = 2, 3, \dots, n-1$.

First, we consider the lower bound. Let $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$ be a double dominating set of size $dd(G^1(1,1,2,2))$ of $G = P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$. By Observation 3.2, vertices $u_1, v_1, v_2, (u_2, v_3), (u_{n-1}, v_n), u_n, v_{n+1}$ and v_{n+2} are in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$. Suppose for a contradiction that $|D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}| < dd(P_n) + 4$.

Let $D^2_k = D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}} \setminus \{v_1, v_2, v_{n+1}, v_{n+2}\}$. Removing v_1 and v_{n+2} does not affect the domination of any vertex in P_n. Removing v_2 and v_{n+1} affects the domination of (u_2, v_3) and (u_{n-1}, v_n), respectively. We can show that D^2_k is a double dominating set of P_n in the similar way as in the proof of Theorem 3.6.

Since $\{v_1, v_2, v_{n+1}, v_{n+2}\} \subseteq D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, we get

$$|D^2_k| = |D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}| - |\{v_1, v_2, v_{n+1}, v_{n+2}\}| < dd(P_n) + 4 - 4 = dd(P_n).$$

It contradicts to the fact that $dd(P_n)$ is the minimum cardinality among those of all double dominating sets of P_n. Hence, $dd(G^1(1,1,2,2) = |D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}| \geq dd(P_n) + 4$.

For the upper bound, we construct a double dominating set of $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$. Define

$$\begin{aligned} D^2 = &\{u_1, v_1, v_2, (u_2, v_3), (u_{n-1}, v_n), u_n, v_{n+1}, v_{n+2}\} \\ &\cup \{(u_i, v_{i+1}) | i \not\equiv 0 (\text{mod} 3) \text{ and } i \in \{3, 4, \dots, n-2\}\}. \end{aligned}$$

We can calculate $|D^2|$ the same way as in the proof of Theorem 3.6, and we get $|D^2| = \left\lceil \frac{2n+2}{3} \right\rceil + 4 = dd(P_n) + 4$.

For $i = 3, 4, \dots, n-2$,
if $i \equiv 0 \pmod 3$, vertex (u_i, v_{i+1}) is dominated by (u_{i-1}, v_i) and (u_{i+1}, v_{i+2}),
if $i \equiv 1 \pmod 3$, vertex (u_i, v_{i+1}) is dominated by (u_i, v_{i+1}) and (u_{i+1}, v_{i+2}), and
if $i \equiv 2 \pmod 3$, vertex (u_i, v_{i+1}) is dominated by (u_{i-1}, v_i) and (u_i, v_{i+1}).

Since u_1 is adjacent to (u_2, v_3), v_1 is adjacent to v_2, (u_{n-1}, v_n) is adjacent to u_n, v_{n+1} is adjacent to v_{n+2}, and all eight vertices are in D^2, by Observation 3.3, each vertex is dominated

by at least two vertices. Then, any vertex in $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$ is dominated by at least two vertices in D^2. Therefore, D^2 is a double dominating set of $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$. Hence, $dd(G^1(1,1,2,2)) \leq |D^2| = \left\lceil \frac{2n+2}{3} \right\rceil + 4 = dd(P_n) + 4$. □

Theorem 3.8. *Let $n \geq 4$ be a positive integer. The double domination number of a graph in type $G^2(1,1,2,2)$ is $dd(P_n) + 3$.*

Proof. Let G be a graph in type $G^2(1,1,2,2)$. The double domination number of G is denoted by $dd(G^2(1,1,2,2))$.

For the lower bound, let $D^2_{G^2(1,1,2,2)}$ be a double dominating set of size $dd(G^2(1,1,2,2))$ of G. By observation 3.2, vertices $u_1, v_1, (u_2, v_2), u_{n-1}, u_n, v_{n-1}$, and v_n are in $D^2_{G^2(1,1,2,2)}$. Suppose for a contradiction that $|D^2_{G^2(1,1,2,2)}| < dd(P_n) + 3$.

Consider vertices (u_{n-2}, v_{n-2}) and (u_{n-3}, v_{n-3}). Assume that $(u_{n-2}, v_{n-2}), (u_{n-3}, v_{n-3}) \notin D^2_{G^2(1,1,2,2)}$, the vertex (u_{n-3}, v_{n-3}) is dominated by at most one vertex in $D^2_{G^2(1,1,2,2)}$, contradicting to (u_{n-3}, v_{n-3}) being dominated by at least two vertices in $D^2_{G^2(1,1,2,2)}$. Therefore, there must be at least one vertex in $\{(u_{n-2}, v_{n-2}), (u_{n-3}, v_{n-3})\}$ in $D^2_{G^2(1,1,2,2)}$.

Let $D^2_k = D^2_{G^2(1,1,2,2)} \backslash \{v_1, v_{n-1}, v_n\}$. Removing v_n does not affect the domination of any vertex in P^1. Removing v_1 affects only the domination of vertex (u_2, v_2). Since (u_2, v_2) is dominated by u_1 and itself, (u_2, v_2) is dominated by at least two vertices in D^2_k. Removing v_{n-1} affects only the domination of vertex (u_{n-2}, v_{n-2}).

It remains to show that vertex (u_{n-2}, v_{n-2}) is dominated by at least two vertices in D^2_k.

Case 1: vertex $(u_{n-2}, v_{n-2}) \in D^2_{G^2(1,1,2,2)}$. Since (u_{n-2}, v_{n-2}) is dominated by u_{n-1} and itself, (u_{n-2}, v_{n-2}) is dominated by at least two vertices in D^2_k.

Case 2: vertex $(u_{n-2}, v_{n-2}) \notin D^2_{G^2(1,1,2,2)}$. Then, $(u_{n-3}, v_{n-3}) \in D^2_{G^2(1,1,2,2)}$. Since (u_{n-2}, v_{n-2}) is dominated by (u_{n-3}, v_{n-3}) and u_{n-1}, we can see that (u_{n-2}, v_{n-2}) is dominated by at least two vertices in D^2_k.

Hence, each vertex in P^1 is dominated by at least two vertices in D^2_k. Therefore, D^2_k is a double dominating set of P^1.

Since $\{v_1, v_{n-1}, v_n\} \subseteq D^2_{G^2(1,1,2,2)}$, we get

$$|D^2_k| = |D^2_{G^2(1,1,2,2)}| - |\{v_1, v_{n-1}, v_n\}| < dd(P^1) + 3 - 3 = dd(P^1) = dd(P_n).$$

It contradicts to the fact that $dd(P_n)$ is the minimum cardinality among those of all double dominating sets of P_n. Hence, $dd(G^2(1,1,2,2)) = |D^2_{G^2(1,1,2,2)}| \geq dd(P_n) + 3$.

For the upper bound, we construct a double dominating set of G. Define

$$D^2 = \{u_1, v_1, (u_2, v_2), u_{n-1}, u_n, v_{n-1}, v_n\} \cup \{(u_i, v_i) | i \not\equiv 0 (\mathrm{mod} 3) \text{ and } i \in \{3, 4, \ldots, n-2\}\}.$$

We can calculate $|D^2|$ in the same way as in the proof of Theorem 3.6, and we get $|D^2| = \left\lceil \frac{2n+2}{3} \right\rceil + 3 = dd(P_n) + 3$.

Similar to what is shown in the proof of Theorem 3.6, for $i = 3, 4, \ldots, n-2$, we can show that (u_i, v_i) is dominated by at least two vertices in D^2. Since u_1 is adjacent to (u_2, v_2), u_{n-1} is adjacent to u_n, v_{n-1} is adjacent to v_n, and all six vertices are in D^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_1 is dominated by v_1 and (u_2, v_2), and (u_{n-2}, v_{n-2}) is dominated by u_{n-1} and v_{n-1}. Then, any vertex in G is dominated by at least

two vertices in D^2. Therefore, D^2 is a double dominating set of G. Hence, $dd(G^2(1,1,2,2)) \leq |D^2| = \left\lceil \frac{2n+2}{3} \right\rceil + 3 = dd(P_n) + 3$. □

Theorem 3.9. *Let $n \geq 10$ be a positive integer. The double domination number of a graph in type $G(1,1,3,3)$ is $\left\lfloor \frac{2(n+4)+2}{3} \right\rfloor + 2$.*

Proof. Let G be a graph in type $G(1,1,3,3)$. The double domination number of G is denoted by $dd(G(1,1,3,3))$. We can see that $G = P^1 \lhd_{H^1 \cong_f H^2} \rhd P^2$ where $H^1 = u_2u_3 \dots u_{n-3}$, $H^2 = v_4v_5 \dots v_{n-1}$, and $f(u_i) = v_{i+2}$ for all $i = 2,3,\dots,n-3$ is isomorphic to $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}$, where $H_1 = u_2u_3 \dots u_{n-1}$, $H_2 = v_4v_5 \dots v_{n+1}$, and $f(u_i) = v_{i+2}$ for all $i = 2,3,\dots,n-1$.

We start by considering the lower bound. Let $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$ be a double dominating set of size $dd(G(1,1,3,3))$ of $G = P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}$. By Observation 3.2, vertices $u_1, v_1, v_2, (u_2, v_4), (u_{n-1}, v_{n+1}), u_n, v_{n+3}$, and v_{n+4} are in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$. Suppose for a contradiction that $|D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}| < \left\lfloor \frac{2(n+4)+2}{3} \right\rfloor + 2$.

Let $D^2_k = D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}} \setminus \{u_1, u_n\}$. Removing u_1 and u_n affects the domination of vertices (u_2, v_4) and (u_{n-1}, v_{n+1}), respectively. There are three cases depending on remainders of dividing $n+4$ by 3.

Case 1: $n + 4 \equiv 0 \pmod 3$. Then, $\left\lfloor \frac{2(n+4)+2}{3} \right\rfloor + 2 = (dd(P_{n+4}) - 1) + 2$.

Case 1.1: $v_3 \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$ and $(u_3, v_5) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$, or $(u_{n-2}, v_n) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$ and $v_{n+2} \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$. By symmetry, we can assume that $v_3 \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$ and $(u_3, v_5) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$.

For $i \in I = \{4, 7, 10, \dots, n-4\}$, let $S_i = \{(u_i, v_{(i+2)}), (u_{i+1}, v_{(i+2)+1}), (u_{i+2}, v_{(i+2)+2})\}$. Since each S_i contains at least two vertices in D^2_k, we have $|D^2_k \cap S_i| \geq 2$. We define D^2_i as follows.

$$D^2_i = \begin{cases} (u_i, v_{(i+2)}), (u_{i+2}, v_{(i+2)+2}) & \text{if } |D^2_k \cap S_i| = 2 \\ S_i & \text{if } |D^2_k \cap S_i| = 3. \end{cases}$$

We construct

$$D^2_N = \{v_1, v_2, (u_2, v_4), (u_3, v_5)\} \cup \bigcup_{i \in I} D^2_i \cup (\{(u_{n-1}, v_{n+1}), v_{n+2}, v_{n+3}, v_{n+4}\} \cap D^2_k).$$

For each vertex in S_i, $(u_i, v_{(i+2)})$ is dominated by $(u_{(i-3)+2}, v_{((i-3)+2)+2})$ and $(u_i, v_{(i+2)})$. Besides, $(u_{i+1}, v_{(i+2)+1})$ is dominated by $(u_i, v_{(i+2)})$ and $(u_{i+2}, v_{(i+2)+2})$, and $(u_{i+2}, v_{(i+2)+2})$ is dominated by $(u_{i+2}, v_{(i+2)+2})$ and $(u_{(i+3)}, v_{(i+3)+2})$. Since v_1 is adjacent to v_2, (u_2, v_4) is adjacent to (u_3, v_5), v_{n+3} is adjacent to v_{n+4}, and all six vertices are in D^2_N, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_3 is dominated by v_2 and (u_2, v_4). We also have (u_{n-1}, v_{n+1}) dominated by (u_{n-2}, v_n) and (u_{n-1}, v_{n+1}), and v_{n+2} dominated by (u_{n-1}, v_{n+1}) and v_{n+3}. Then, any vertex in P_{n+4} is dominated by at least two vertices in D^2_N. Hence, D^2_N is a double dominating set of P_{n+4}.

Since $\{u_1, u_n\} \subseteq D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$, we get

$$|D_N^2| = |D_k^2| = |D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}| - |\{u_1, u_n\}| + |\{(u_3, v_5)\}|$$
$$< (dd(P_{n+4}) - 1) + 2 - 2 + 1 = dd(P_{n+4}).$$

Case 1.2: There is at least one vertex in $\{v_3, (u_3, v_5)\}$ in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$, and there is at least one vertex in $\{(u_{n-2}, v_n), v_{n+2}\}$ in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$. Removing u_1 affects only the domination of vertex (u_2, v_4). Since there is at least one vertex in $\{v_3, (u_3, v_5)\}$ in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$, we get (u_2, v_4) is dominated by itself and one of the vertex in $\{v_3, (u_3, v_5)\}$. Removing u_n affects only the domination of vertex (u_{n-1}, v_{n+1}). Since there is at least one vertex in $\{(u_{n-2}, v_n), v_{n+2}\}$ in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$, we get (u_{n-1}, v_{n+1}) is dominated by itself and one of the vertex in $\{(u_{n-2}, v_n), v_{n+2}\}$. Then, any vertex in P_{n+4} is dominated by at least two vertices in D_k^2. Hence, D_k^2 is a double dominating set of P_{n+4}.

Since $\{u_1, u_n\} \subseteq D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$, we get

$$|D_k^2| = |D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}| - |\{u_1, u_n\}|$$
$$< (dd(P_{n+4}) - 1) + 2 - 2 = dd(P_{n+4}) - 1 < dd(P_{n+4}).$$

Both cases contradict to the fact that $dd(P_{n+4})$ is the minimum cardinality among those of all double dominating sets of (P_{n+4}).

Case 2: $n + 4 \equiv 1 \pmod 3$. Then, $\left\lfloor \frac{2(n+4)+2}{3} \right\rfloor + 2 = (dd(P_{n+4}) - 1) + 2$.

Case 2.1: $v_3 \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$ and $(u_3, v_5) \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$, or $(u_{n-2}, v_n) \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$ and $v_{n+2} \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$. By symmetry, we can assume that $v_3 \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$ and $(u_3, v_5) \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$. Consider vertices (u_3, v_5), (u_4, v_6), and (u_5, v_7). Since $(u_3, v_5) \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$, by Lemma 3.5, both (u_4, v_6) and (u_5, v_7) are in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$.

For $i \in I = \{5, 8, 11, \ldots, n-4\}$, we define S_i and D_i^2 in the same way as in Case 1.1. We construct

$$D_N^2 = \{v_1, v_2, (u_2, v_4), (u_3, v_5), (u_4, v_6)\} \cup \bigcup_{i \in I} D_i^2 \cup (\{(u_{n-1}, v_{n+1}), v_{n+2}, v_{n+3}, v_{n+4}\} \cap D_k^2).$$

Similar to Case 1.1, for $i = 5, 8, 11, \ldots, n-4$, we can show that each vertex in S_i is dominated by at least two vertices in D_N^2.

Since v_1 is adjacent to v_2, (u_2, v_4) is adjacent to (u_3, v_5), v_{n+3} is adjacent to v_{n+4}, and all six vertices are in D_N^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_3 is dominated by v_2 and (u_2, v_4). Besides, (u_4, v_6) is dominated by (u_3, v_5) and (u_5, v_7). We also have (u_{n-1}, v_{n+1}) dominated by (u_{n-2}, v_n) and (u_{n-1}, v_{n+1}), and v_{n+2} dominated by (u_{n-1}, v_{n+1}) and v_{n+3}. Then, any vertex in P_{n+4} is dominated by at least two vertices in D_N^2. Hence, D_N^2 is a double dominating set of P_{n+4}.

We get the same contradiction as in Case 1.1.

Case 2.2: There is at least one vertex in $\{v_3, (u_2, v_5)\}$ in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$, and there is at least one vertex in $\{(u_{n-2}, v_n), v_{n+2}\}$ in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+4}}$. We can show in the similar way as in Case 1.2 that D_k^2 is a double dominating set of P_{n+4}, and get the same contradiction.

Case 3: $n+4 \equiv 2 \pmod 3$. Then, $\left\lfloor \frac{2(n+4)+2}{3} \right\rfloor + 2 = dd(P_{n+4}) + 2$.

Case 3.1: $v_3 \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$ and $(u_3, v_5) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$, or $(u_{n-2}, v_n) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$ and $v_{n+2} \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$. Consider vertices (u_3, v_5), (u_4, v_6), and (u_5, v_7). Since $(u_3, v_5) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$, by Lemma 3.5, both (u_4, v_6) and (u_5, v_7) are in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$.

For $i \in I = \{6, 9, 12, \ldots, n-4\}$, we define S_i and D^2_i in the same way as in Case 1.1. We construct

$$D^2_N = \{v_1, v_2, (u_2, v_4), (u_3, v_5), (u_5, v_7)\} \cup \bigcup_{i \in I} D^2_i \cup (\{(u_{n-1}, v_{n+1}), v_{n+2}, v_{n+3}, v_{n+4}\} \cap D^2_k).$$

Similar to Case 1.1, for $i = 6, 9, 12, \ldots, n-4$, we can show that each vertex in S_i is dominated by at least two vertices in D^2_N.

Since v_1 is adjacent to v_2, (u_2, v_4) is adjacent to (u_3, v_5), v_{n+3} is adjacent to v_{n+4}, and all six vertices are in D^2_N, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_3 is dominated by v_2 and (u_2, v_4). Besides, (u_4, v_6) is dominated by (u_3, v_5) and (u_5, v_7), and (u_5, v_7) is dominated by (u_5, v_7) and (u_6, v_8). We also have (u_{n-1}, v_{n+1}) dominated by (u_{n-2}, v_n) and (u_{n-1}, v_{n+1}), and v_{n+2} dominated by (u_{n-1}, v_{n+1}) and v_{n+3}. Then, any vertex in P_{n+4} is dominated by at least two vertices in D^2_N. Hence, D^2_N is a double dominating set of P_{n+4}.

Since $\{u_1, u_n\} \subseteq D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$, we get

$$|D^2_N| = |D^2_k| = |D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}| - |\{u_1, u_n\}| < dd(P_{n+4}) + 2 - 2 = dd(P_{n+4}).$$

It contradicts to the fact that $dd(P_{n+4})$ is the minimum cardinality among those of all double dominating sets of P_{n+4}.

Case 3.2: There is at least one vertex in $\{v_3, (u_3, v_5)\}$ in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$, and there is at least one vertex in $\{(u_{n-2}, v_n), v_{n+2}\}$ in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$. We can show in the similar way as in Case 1.2 that D^2_k is a double dominating set of P_{n+4}.

Since $\{u_1, u_n\} \subseteq D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$, we get

$$|D^2_k| = |D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}| - |\{u_1, u_n\}| < dd(P_{n+4}) + 2 - 2 = dd(P_{n+4}).$$

It contradicts to the fact that $dd(P_{n+4})$ is the minimum cardinality among those of all double dominating sets of P_{n+4}.

From all cases, $dd(G(1,1,3,3)) = |D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}| \geq \left\lfloor \frac{2(n+4)+2}{3} \right\rfloor + 2$.

For the upper bound, we construct a double dominating set of $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}$. Define

$$\begin{aligned} D^2 = &\{u_1, v_1, v_2, (u_2, v_4), (u_{n-1}, v_{n+1}), u_n, v_{n+3}, v_{n+4}\} \\ &\cup \{(u_i, v_{i+2}) | i \not\equiv 0 (\text{mod} 3) \text{ and } i \in \{3, 4, \ldots, n-2\}\}. \end{aligned}$$

We can calculate $|D^2|$ in the similar way as in the proof of Theorem 3.6, and we get $|D^2| = \left\lfloor \frac{2(n+4)+2}{3} \right\rfloor + 2$.

For $i = 3, 4, \ldots, n-2$,

if $i \equiv 0 \pmod 3$, vertex (u_i, v_{i+2}) is dominated by $(u_{i-1}, v_{(i+2)-1})$ and $(u_{i+1}, v_{(i+2)+1})$,

if $i \equiv 1 \pmod 3$, vertex (u_i, v_{i+2}) is dominated by (u_i, v_{i+2}) and $(u_{i+1}, v_{(i+2)+1})$, and
if $i \equiv 2 \pmod 3$, vertex (u_i, v_{i+2}) is dominated by $(u_{i-1}, v_{(i+2)-1})$ and (u_i, v_{i+2}).

Since u_1 is adjacent to (u_2, v_4), v_1 is adjacent to v_2, v_{n+3} is adjacent to v_{n+4}, and all six vertices are in D^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, (u_{n-1}, v_{n+1}) is dominated by itself and u_n, and v_{n+2} is dominated by (u_{n-1}, v_{n+1}) and v_{n+3}.

Then, for any vertex in $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}$, the vertex is dominated by at least two vertices in D^2. Therefore, D^2 is a double dominating set of $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}$.

Hence, $dd(G(1,1,3,3)) \leq |D^2| = \left\lfloor \dfrac{2(n+4)+2}{3} \right\rfloor + 2$. □

Theorem 3.10. *Let $n \geq 12$ be a positive integer. The double domination number of a graph in type $G(1,2,2,3)$ is*

$$dd(G(1,2,2,3)) \geq \begin{cases} dd(P_n) + 3 & , n \equiv 1 \pmod 3 \\ dd(P_n) + 4 & , \textit{otherwise.} \end{cases}$$

Proof. Let G be a graph in type $G(1,2,2,3)$. The double domination number of G is denoted by $dd(G(1,2,2,3))$.

For the lower bound, let $D^2_{G(1,2,2,3)}$ be a double dominating set of size $dd(G(1,2,2,3))$ of G. By Observation 3.2, vertices $u_1, v_1, v_2, (u_2, v_3), u_{n-1}, u_n, v_{n-1}$, and v_n are in $D^2_{G(1,2,2,3)}$. There are three cases depending on remainders of dividing n by 3.

Let $D^2_k = D^2_{G(1,2,2,3)} \backslash \{v_1, v_2, v_{n-1}, v_n\}$. Removing v_1 and v_n does not affect the domination of any vertex in P^1. Removing v_2 affects the domination of (u_2, v_3). Since (u_2, v_3) is dominated by u_1 and itself, (u_2, v_3) is dominated by at least two vertices in D^2_k. Removing v_{n-1} affects the domination of (u_{n-3}, v_{n-2}). Note that we need to verify that (u_{n-3}, v_{n-2}) is dominated by at least two vertices in D^2_k.

Case 1: $n \equiv 0 \pmod 3$. Suppose for a contradiction that $|D^2_{G(1,2,2,3)}| < dd(P_n) + 4$.

Case 1.1: At most one vertex in $\{(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-2}\}$ is in $D^2_{G(1,2,2,3)}$. If $(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}) \notin D^2_{G(1,2,2,3)}$, the vertex (u_{n-3}, v_{n-2}) is dominated by at most one vertex in $D^2_{G(1,2,2,3)}$, which is vertex v_{n-1}. This contradicts to (u_{n-3}, v_{n-2}) being dominated by at least two vertices in $D^2_{G(1,2,2,3)}$. Therefore, there must be at least one vertex in $\{(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-2}\}$ in $D^2_{G(1,2,2,3)}$.

Assume that $(u_{n-3}, v_{n-2}), u_{n-2} \notin D^2_{G(1,2,2,3)}$, the vertex u_{n-2} is dominated by at most one vertex in $D^2_{G(1,2,2,3)}$, which is vertex u_{n-1}. This contradicts u_{n-2} being dominated by at least two vertices in $D^2_{G(1,2,2,3)}$. Assume that $(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}) \notin D^2_{G(1,2,2,3)}$, the vertex (u_{n-4}, v_{n-3}) is dominated by at most one vertex in $D^2_{G(1,2,2,3)}$, which is vertex (u_{n-5}, v_{n-4}). This contradicts to (u_{n-4}, v_{n-3}) being dominated by at least two vertices in $D^2_{G(1,2,2,3)}$. Hence, we consider only the case where $(u_{n-3}, v_{n-2}) \in D^2_{G(1,2,2,3)}$, $(u_{n-4}, v_{n-3}) \notin D^2_{G(1,2,2,3)}$, and $u_{n-2} \notin D^2_{G(1,2,2,3)}$.

Consider vertices $(u_{n-6}, v_{n-5}), (u_{n-5}, v_{n-4})$, and (u_{n-4}, v_{n-3}). Since $(u_{n-4}, v_{n-3}) \notin D^2_{G(1,2,2,3)}$, by Lemma 3.5, (u_{n-6}, v_{n-5}) and (u_{n-5}, v_{n-4}) are in $D^2_{G(1,2,2,3)}$.

For $i \in I = \{3, 6, 9, \ldots, n-9\}$, let $S_i = \{(u_i, v_{(i+1)}), (u_{i+1}, v_{(i+1)+1}), (u_{i+2}, v_{(i+1)+2})\}$. Since each S_i contains at least two vertices in D^2_k, we have $|D^2_k \cap S_i| \geq 2$. We define D^2_i as follows.

$$D_i^2 = \begin{cases} (u_i, v_{(i+1)}), (u_{i+2}, v_{(i+1)+2}) & \text{if } |D_k^2 \cap S_i| = 2 \\ S_i & \text{if } |D_k^2 \cap S_i| = 3. \end{cases}$$

We construct

$$D_N^2 = \{u_1, (u_2, v_3)\} \cup \bigcup_{i \in I} D_i^2 \cup \{(u_{n-6}, v_{n-5}), (u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-1}, u_n\}.$$

For each vertex in S_i, we get $(u_i, v_{(i+1)})$ dominated by $(u_{(i-3)+2}, v_{((i-3)+1)+2})$ and $(u_i, v_{(i+1)})$. Besides, $(u_{i+1}, v_{(i+1)+1})$ is dominated by $(u_i, v_{(i+1)})$ and $(u_{i+2}, v_{(i+1)+2})$, and $(u_{i+2}, v_{(i+1)+2})$ is dominated by $(u_{i+2}, v_{(i+1)+2})$ and $(u_{(i+3)}, v_{(i+3)+1})$.

Since u_1 is adjacent to (u_2, v_3), u_{n-1} is adjacent to u_n, (u_{n-4}, v_{n-3}) is adjacent to (u_{n-3}, v_{n-2}) and all six vertices are in D_N^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, (u_{n-6}, v_{n-5}) is dominated by (u_{n-7}, v_{n-6}) and (u_{n-6}, v_{n-5}). We also have (u_{n-5}, v_{n-4}) dominated by (u_{n-6}, v_{n-5}) and (u_{n-4}, v_{n-3}), and u_{n-2} dominated by (u_{n-3}, v_{n-2}) and u_{n-1}. Any vertex in P^1 is dominated by at least two vertices in D_N^2. Hence, D_N^2 is a double dominating set of P^1.

Since $\{v_1, v_2, v_{n-1}, v_n\} \subseteq D^2_{G(1,2,2,3)}$, we get

$$|D_N^2| = |D_k^2| = |D^2_{G(1,2,2,3)}| - |\{v_1, v_2, v_{n-1}, v_n\}| < dd(P^1) + 4 - 4 = dd(P^1) = dd(P_n).$$

It contradicts to the fact that $dd(P_n)$ is the minimum cardinality among those of all double dominating sets of P_n.

Case 1.2: At least two vertices in $\{(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-2}\}$ are in $D^2_{G(1,2,2,3)}$. Removing v_{n-1} affects the domination of (u_{n-3}, v_{n-2}).

If $(u_{n-3}, v_{n-2}) \in D^2_{G(1,2,2,3)}$ and $u_{n-2} \in D^2_{G(1,2,2,3)}$, we have (u_{n-3}, v_{n-2}) dominated by (u_{n-3}, v_{n-2}) and u_{n-2}. If $(u_{n-4}, v_{n-3}) \in D^2_{G(1,2,2,3)}$ and $u_{n-2} \in D^2_{G(1,2,2,3)}$, we have (u_{n-3}, v_{n-2}) is dominated by (u_{n-4}, v_{n-3}) and u_{n-2}. If $(u_{n-3}, v_{n-2}) \in D^2_{G(1,2,2,3)}$ and $(u_{n-4}, v_{n-3}) \in D^2_{G(1,2,2,3)}$, we have (u_{n-3}, v_{n-2}) is dominated by (u_{n-4}, v_{n-3}) and (u_{n-3}, v_{n-2}). From all these cases, (u_{n-3}, v_{n-2}) is dominated by at least two vertices in D_k^2.

Hence, any vertex in P^1 is dominated by at least two vertices in D_k^2. Therefore, D_k^2 is a double dominating set of P^1. We get the same contradiction as in Case 1.1.

From both cases, $dd(G(1,2,2,3)) = |D^2_{G(1,2,2,3)}| \geq dd(P_n) + 4$.

Case 2: $n \equiv 1 \pmod 3$. Suppose for a contradiction that $|D^2_{G(1,2,2,3)}| < dd(P_n) + 3$.

Case 2.1: At most one vertex in $\{(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-2}\}$ is in $D^2_{G(1,2,2,3)}$. Similar to Case 1.1, we consider only the case where $(u_{n-3}, v_{n-2}) \in D^2_{G(1,2,2,3)}$, $(u_{n-4}, v_{n-3}) \notin D^2_{G(1,2,2,3)}$ and $u_{n-2} \notin D^2_{G(1,2,2,3)}$. We construct $D_N^2 = D_k^2 \cup \{u_{n-2}\}$. Therefore, (u_{n-3}, v_{n-2}) is dominated by two vertices in D_N^2 that are (u_{n-3}, v_{n-2}) and u_{n-2}. Hence, any vertex in P^1 is dominated by at least two vertices in D_N^2. Therefore, D_N^2 is a double dominating set of P^1.

Since $\{v_1, v_2, v_{n-1}, v_n\} \subseteq D^2_{G(1,2,2,3)}$, we get

$$\begin{aligned} |D_N^2| &= |D_k^2 \cup \{u_{n-2}\}| = |D^2_{G(1,2,2,3)}| - |\{v_1, v_2, v_{n-1}, v_n\}| + |\{u_{n-2}\}| \\ &< (dd(P^1) + 3) - 4 + 1 = dd(P^1) = dd(P_n). \end{aligned}$$

Case 2.2: At least two vertices in $\{(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-2}\}$ are in $D^2_{G(1,2,2,3)}$. Similar to Case 1.2 we can show that (u_{n-3}, v_{n-2}) is dominated by at least two vertices in $D^2_{G(1,2,2,3)}$.

Hence, any vertex in P^1 is dominated by at least two vertices in D_k^2. Therefore, D_k^2 is a double dominating set of P^1.

Since $\{v_1, v_2, v_{n-1}, v_n\} \subseteq D^2_{G(1,2,2,3)}$, we get

$$\begin{aligned}|D_k^2| &= |D^2_{G(1,2,2,3)}| - |\{v_1, v_2, v_{n-1}, v_n\}| \\ &< (dd(P^1) + 3) - 4 = dd(P^1) - 1 = dd(P_n) - 1 < dd(P_n).\end{aligned}$$

From both cases, it contradicts to the fact that $dd(P_n)$ is the minimum cardinality among those of all double dominating sets of P_n. Hence, $dd(G(1,2,2,3)) = |D^2_{G(1,2,2,3)}| \geq dd(P_n) + 3$.

Case 3: $n \equiv 2 \pmod 3$. Suppose for a contradiction that $|D^2_{G(1,2,2,3)}| < dd(P_n) + 4$.

Case 3.1: At most one vertex in $\{(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-2}\}$ is in $D^2_{G(1,2,2,3)}$. Similar to Case 1.1, we consider only the case where $(u_{n-3}, v_{n-2}) \in D^2_{G(1,2,2,3)}$, $(u_{n-4}, v_{n-3}) \notin D^2_{G(1,2,2,3)}$ and $u_{n-2} \notin D^2_{G(1,2,2,3)}$.

Consider vertices (u_3, v_4) and (u_4, v_5). Assume that $(u_3, v_4), (u_4, v_5) \notin D^2_{G(1,2,2,3)}$, we get (u_3, v_4) is dominated by at most one vertex in $D^2_{G(1,2,2,3)}$, which is (u_2, v_3). This contradicts to (u_3, v_4) being dominated by at least two vertices in $D^2_{G(1,2,2,3)}$. Therefore, there must be at least one vertex in $\{(u_3, v_4), (u_4, v_5)\}$ in $D^2_{G(1,2,2,3)}$. Consider vertices $(u_{n-6}, v_{n-5}), (u_{n-5}, v_{n-4})$, and (u_{n-4}, v_{n-3}). Since $(u_{n-4}, v_{n-3}) \notin D^2_{G(1,2,2,3)}$, by Lemma 3.5, (u_{n-6}, v_{n-5}) and (u_{n-5}, v_{n-4}) are in $D^2_{G(1,2,2,3)}$.

For $i \in I = \{5, 8, 11, \ldots, n-9\}$, let $S_i = \{(u_i, v_{(i+1)}), (u_{i+1}, v_{(i+1)+1}), (u_{i+2}, v_{(i+1)+2})\}$. Since each S_i contains at least two vertices in D_k^2, we have $|D_k^2 \cap S_i| \geq 2$. We define D_i^2 and D_L^2 as follows.

$$D_i^2 = \begin{cases} \{(u_i, v_{(i+1)}), (u_{i+2}, v_{(i+1)+2})\} & \text{if } |D_k^2 \cap S_i| = 2 \\ S_i & \text{if } |D_k^2 \cap S_i| = 3, \text{ and} \end{cases}$$

$$D_L^2 = \begin{cases} \{(u_4, v_5)\} & \text{if } (u_3, v_4) \notin D_k^2 \text{ and } (u_4, v_5) \in D_k^2 \\ \{(u_4, v_5)\} & \text{if } (u_3, v_4) \in D_k^2 \text{ and } (u_4, v_5) \notin D_k^2 \\ \{(u_3, v_4), (u_4, v_5)\} & \text{if } (u_3, v_4) \in D_k^2 \text{ and } (u_4, v_5) \in D_k^2. \end{cases}$$

We construct

$$D_N^2 = \{u_1, (u_2, v_3)\} \cup D_L^2 \cup \bigcup_{i \in I} D_i^2 \cup \{(u_{n-6}, v_{n-5}), (u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-1}, u_n\}.$$

Similar to Case 1.1, for $i = 5, 8, 11, \ldots, n-9$, we can show that each vertex in S_i is dominated by at least two vertices in D_N^2. Since u_1 is adjacent to (u_2, v_3), u_{n-1} is adjacent to u_n, (u_{n-4}, v_{n-3}) is adjacent to (u_{n-3}, v_{n-2}), and all six vertices are in D_N^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, (u_2, v_3) is dominated by (u_2, v_3) and (u_4, v_5), and (u_4, v_5) is dominated by (u_4, v_5) and (u_5, v_6). Besides, (u_{n-6}, v_{n-5}) is dominated by (u_{n-7}, v_{n-6}) and (u_{n-6}, v_{n-5}). We also have (u_{n-5}, v_{n-4}) dominated by (u_{n-6}, v_{n-5}) and (u_{n-4}, v_{n-3}), and u_{n-2} dominated by (u_{n-3}, v_{n-2}) and u_{n-1}. Then, any vertex in P^1 is dominated by at least two vertices in D_N^2. Hence, D_N^2 is a double dominating set of P^1.

Since $\{v_1, v_2, v_{n-1}, v_n\} \subseteq D^2_{G(1,2,2,3)}$, we get

$$|D_N^2| = |D_k^2| = |D^2_{G(1,2,2,3)}| - |\{v_1, v_2, v_{n-1}, v_n\}| < dd(P^1) + 4 - 4 = dd(P^1) = dd(P_n).$$

It contradicts to the fact that $dd(P_n)$ is the minimum cardinality among those of all double dominating sets of P_n.

Case 3.2: At least two vertices in $\{(u_{n-4}, v_{n-3}), (u_{n-3}, v_{n-2}), u_{n-2}\}$ are in $D^2_{G(1,2,2,3)}$. Similar to Case 1.2 we can show that (u_{n-3}, v_{n-2}) is dominated by at least two vertices in $D^2_{G(1,2,2,3)}$. Hence, any vertex in P^1 is dominated by at least two vertices in D^2_k. Therefore, D^2_k is a double dominating set of P^1. We get the same contradiction as in Case 3.1.

From both cases, $dd(G(1,2,2,3)) = |D^2_{G(1,2,2,3)}| \geq dd(P_n) + 4$.

Now we consider the upper bound. There are three cases depending on remainders of dividing n by 3.

Case $n \equiv 0 \pmod 3$: Define

$$D^2 = \{u_1, v_1, v_2, (u_2, v_3), u_{n-2}, u_{n-1}, u_n, v_{n-1}, v_n\} \cup \{(u_i, v_{i+1}) | i \not\equiv 0 (\mathrm{mod} 3) \text{ and } i \in \{3, 4, \ldots, n-3\}\}.$$

We can calculate $|D^2|$ in the similar way as in the proof of Theorem 3.6, and we get $|D^2| = \left\lceil \frac{2n+2}{3} \right\rceil + 4 = dd(P_n) + 4$. Similar to what is shown in the proof of Theorem 3.7, for $i = 3, 4, \ldots, n-3$, we can show that (u_i, v_{i+1}) is dominated by at least two vertices in D^2.

Since u_1 is adjacent to (u_2, v_3), v_1 is adjacent to v_2, u_{n-1} is adjacent to u_n, v_{n-1} is adjacent to v_n, and all eight vertices are in D^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, (u_{n-3}, v_{n-2}) is dominated by (u_{n-4}, v_{n-3}) and u_{n-2}, and u_{n-2} is dominated by u_{n-2} and u_{n-1}. Then, any vertex in G is dominated by at least two vertices in D^2. Therefore, D^2 is a double dominating set of G. Hence, $dd(G(1,2,2,3)) \leq |D^2| = dd(P_n) + 4$.

Case $n \equiv 1 \pmod 3$: Define

$$D^2 = \{u_1, v_1, v_2, (u_2, v_3), u_{n-1}, u_n, v_{n-1}, v_n\} \cup \{(u_i, v_{i+1}) | i \not\equiv 0 (\mathrm{mod} 3) \text{ and } i \in \{3, 4, \ldots, n-3\}.$$

We can calculate $|D^2|$ in the similar way as in the proof of Theorem 3.6, and we get $|D^2| = \left\lfloor \frac{2n+2}{3} \right\rfloor + 3 = dd(P_n) + 3$. Similar to what is shown in the proof of Theorem 3.7, for $i = 3, 4, \ldots, n-3$, we can show that (u_i, v_{i+1}) is dominated by at least two vertices in D^2.

Since u_1 is adjacent to (u_2, v_3), v_1 is adjacent to v_2, u_{n-1} is adjacent to u_n, v_{n-1} is adjacent to v_n, and all eight vertices are in D^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, (u_{n-3}, v_{n-2}) is dominated by (u_{n-3}, v_{n-2}) and v_{n-1}, and u_{n-2} is dominated by (u_{n-3}, v_{n-2}) and u_{n-1}. Then, any vertex in G is dominated by at least two vertices in D^2. Therefore, D^2 is a double dominating set of G. Hence, $dd(G(1,2,2,3)) \leq |D^2| = dd(P_n) + 3$.

Case $n \equiv 2 \pmod 3$: Define D^2 in the same way as in the case when $n \equiv 1 \pmod 3$. We can calculate $|D^2|$ in the similar way as in the proof of Theorem 3.6, and we get $|D^2| = \left\lceil \frac{2n+2}{3} \right\rceil + 4 = dd(P_n) + 4$. Similar to what is shown in the proof of Theorem 3.7, for $i = 3, 4, \ldots, n-3$, we can show that (u_i, v_{i+1}) is dominated by at least two vertices in D^2.

Since u_1 is adjacent to (u_2, v_3), v_1 is adjacent to v_2, u_{n-1} is adjacent to u_n, v_{n-1} is adjacent to v_n, and all eight vertices are in D^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, (u_{n-3}, v_{n-2}) is dominated by (u_{n-3}, v_{n-2}) and v_{n-1}, and u_{n-2} is dominated by (u_{n-3}, v_{n-2}) and u_{n-1}. Then, any vertex in G is dominated by at least two vertices in D^2. Therefore, D^2 is a double dominating set of G. Hence, $dd(G) \leq |D^2| = \left\lceil \frac{2n+2}{3} \right\rceil + 4$. □

The proof of the following theorem can be done in a similar way as the previous theorems. So, we omit the proof.

Theorem 3.11. *Let $n \geq 5$ be a positive integer. The double domination number of a graph in type $G(2,2,2,2)$ is $dd(P_n)+4$.*

Theorem 3.12. *Let $n \geq 11$ be a positive integer. The double domination number of a graph in type $G(2,2,3,3)$ is $\left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4$.*

Proof. Let G be a graph in type $G(2,2,3,3)$. The double domination number of G is denoted by $dd(G(2,2,3,3))$. We can see that $G = P^1 \lhd_{H^1 \cong_f H^2} \rhd P^2$ where $H^1 = u_3u_4 \dots u_{n-3}$, $H^2 = v_4v_5 \dots v_{n-2}$, and $f(u_i) = v_{i+1}$ for all $i = 3,4,\dots,n-3$ is isomorphic to $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$, where $H_1 = u_3u_4 \dots u_{n-2}$, $H_2 = v_4v_5 \dots v_{n-1}$, and $f(u_i) = v_{i+1}$ for all $i = 3,4,\dots,n-2$.

For the lower bound, let $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$ be a double dominating set of size $dd(G(2,2,3,3))$ of G. By Observation 3.2, $u_1, u_2, v_1, v_2, u_{n-1}, u_n, v_{n+1}$, and v_{n+2} are in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+4}}$. Suppose for a contradiction that $|D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}| < \left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4$.

Let $D^2_k = D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}} \setminus \{u_1, u_2, u_{n-1}, u_n\}$. Removing u_1 and u_n does not affect the domination of any vertex in P_{n+2}. Removing u_2 and u_{n-1} affects the domination of (u_3, v_4) and (u_{n-2}, v_{n-1}), respectively. There are three cases depending on remainders of dividing $n+2$ by 3.

Case 1: $n+2 \equiv 0 \pmod 3$. Then, $\left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4 = (dd(P_{n+2}) - 1) + 4$.

Case 1.1: At most one vertex in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ is in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, or at most one vertex in $\{(u_{n-3}, v_{n-2}), (u_{n-2}, v_{n-1}), v_n\}$ is in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$. By symmetry, we can assume that at most one vertex in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ is in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$.

Assume that $v_3, (u_3, v_4), (u_4, v_5) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, the vertex (u_3, v_4) is dominated by at most one vertex in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, which is u_2. This contradicts to (u_3, v_4) being dominated by at least two vertices in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$. Therefore, there must be at least one vertex in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$.

Assume that $v_3, (u_3, v_4) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, the vertex v_3 is dominated by at most one vertex in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, which is v_2. This contradicts to v_3 being dominated by at least two vertices in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$. Assume that $(u_3, v_4), (u_4, v_5) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, the vertex (u_4, v_5) is dominated by at most one vertex in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, which is (u_5, v_6). This contradicts to v_3 being dominated by at least two vertices in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$. Hence, we consider only the case where $(u_3, v_4) \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, $v_3 \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, and $(u_4, v_5) \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$.

Consider vertices (u_{n-2}, v_{n-1}) and v_n. Assume that $(u_{n-2}, v_{n-1}), v_n \notin D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, we get v_n is dominated by at most one vertex in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, which is v_{n+1}. It contradicts to v_n being dominated by at least two vertices in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$. Therefore, there must be at least one vertex in $\{(u_{n-2}, v_{n-1}), v_n\}$ in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$.

For $i \in I = \{5, 8, 11, \ldots, n-5\}$, let $S_i = \{(u_i, v_{(i+1)}), (u_{i+1}, v_{(i+1)+1}), (u_{i+2}, v_{(i+1)+2})\}$. Since each S_i contains at least two vertices in D_k^2, we have $|D_k^2 \cap S_i| \geq 2$. We define D_i^2 and D_R^2 as follows.

$$D_i^2 = \begin{cases} (u_i, v_{(i+1)}), (u_{i+2}, v_{(i+1)+2}) & \text{if } |D_k^2 \cap S_i| = 2 \\ S_i & \text{if } |D_k^2 \cap S_i| = 3, \text{ and} \end{cases}$$

$$D_R^2 = \begin{cases} \{(u_{n-2}, v_{n-1})\} & \text{if } (u_{n-2}, v_{n-1}) \notin D_k^2 \text{ and } v_n \in D_k^2 \\ \{(u_{n-2}, v_{n-1})\} & \text{if } (u_{n-2}, v_{n-1}) \in D_k^2 \text{ and } v_n \notin D_k^2 \\ \{(u_{n-2}, v_{n-1}), v_n\} & \text{if } (u_{n-2}, v_{n-1}) \in D_k^2 \text{ and } v_n \in D_k^2. \end{cases}$$

We construct

$$D_N^2 = \{v_1, v_2, (u_3, v_4), (u_4, v_5)\} \cup \bigcup_{i \in I} D_i^2 \cup D_R^2 \cup \{v_{n+1}, v_{n+2}\}.$$

Similar to Case 1.1 in the proof of Theorem 3.10, for $i = 5, 8, 11, \ldots, n-5$, we can show that each vertex in S_i is dominated by at least two vertices in D_N^2.

Since v_1 is adjacent to v_2, (u_3, v_4) is adjacent to (u_4, v_5), v_{n+1} is adjacent to v_{n+2}, and all six vertices are in D_N^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_3 is dominated by v_2 and (u_3, v_4), We also have (u_{n-2}, v_{n-1}) dominated by (u_{n-3}, v_{n-2}) and (u_{n-2}, v_{n-1}), and v_n dominated by (u_{n-2}, v_{n-1}) and v_{n+1}. Then, any vertex in $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$ is dominated by at least two vertices in D_N^2. Hence, D_N^2 is a double dominating set of P_{n+2}.

Since $\{u_1, u_2, u_{n-1}, u_n\} \subseteq D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, we get

$$\begin{aligned} |D_N^2| = |D_k^2| &= |D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}| - |\{u_1, u_2, u_{n-1}, u_n\}| + |\{(u_4, v_5)\}| \\ &< (dd(P_{n+2}) - 1) + 4 - 4 + 1 = dd(P_{n+2}). \end{aligned}$$

It contradicts to the fact that $dd(P_{n+2})$ is the minimum cardinality among those of all double dominating sets of P_{n+2}.

Case 1.2: At least two vertices in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ are in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, and at least two vertices in $\{(u_{n-3}, v_{n-2}), (u_{n-2}, v_{n-1}), v_n\}$ are in $D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$.

Removing u_2 affects the domination of (u_3, v_4). If $v_3 \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$ and $(u_4, v_5) \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, since (u_3, v_4) is dominated by v_3 and (u_4, v_5), we get (u_3, v_4) is dominated by at least two vertices in D_k^2. If $v_3 \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$ and $(u_3, v_4) \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, since (u_3, v_4) is dominated by v_3 and (u_3, v_4), we get (u_3, v_4) is dominated by at least two vertices in D_k^2. If $(u_3, v_4) \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$ and $(u_4, v_5) \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, since (u_3, v_4) is dominated by (u_3, v_4) and (u_4, v_5), we get (u_{n-3}, v_{n-2}) is dominated by at least two vertices in D_k^2.

Removing u_{n-1} affects the domination of (u_{n-2}, v_{n-1}). If $(u_{n-3}, v_{n-2}) \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$ and $v_n \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, since (u_{n-2}, v_{n-1}) is dominated by (u_{n-3}, v_{n-2}) and v_n, we get (u_{n-2}, v_{n-1}) is dominated by at least two vertices in D_k^2. If $(u_{n-3}, v_{n-2}) \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$ and $(u_{n-2}, v_{n-1}) \in D^2_{P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}}$, since (u_{n-2}, v_{n-1}) is dominated by (u_{n-3}, v_{n-2}) and (u_{n-2}, v_{n-1}), we get (u_{n-2}, v_{n-1}) is dominated by at least two vertices in D_k^2. If (u_{n-2}, v_{n-1})

$\in D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$ and $v_n \in D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, since (u_{n-2}, v_{n-1}) is dominated by v_n and (u_{n-2}, v_{n-1}), we get v_n is dominated by at least two vertices in D^2_k.

Then, any vertex in P_{n+2} is dominated by at least two vertices in D^2_k. Hence, D^2_k is a double dominating set of P_{n+2}.

Since $\{u_1, u_2, u_{n-1}, u_n\} \subseteq D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, we get

$$\begin{aligned} |D^2_k| &= |D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}| - |\{u_1, u_2, u_{n-1}, u_n\}| \\ &< (dd(P_{n+2}) - 1) + 4 - 4 = dd(P_{n+2}) - 1 < dd(P_{n+2}). \end{aligned}$$

It contradicts to the fact that $dd(P_{n+2})$ is the minimum cardinality among those of all double dominating sets of P_{n+2}.

Case 2: $n + 2 \equiv 1 \pmod 3$. Then, $\left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4 = (dd(P_{n+2}) - 1) + 4$.

Case 2.1: At most one vertex in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ is in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, or at most one vertex in $\{(u_{n-3}, v_{n-2}), (u_{n-2}, v_{n-1}), v_n\}$ is in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$. By symmetry, we can assume that at most one vertex in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ is in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$.

Similar to Case 1.1, we consider only the case where $(u_3, v_4) \in D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, $v_3 \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, and $(u_4, v_5) \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, and there must be at least one vertex in $\{(u_{n-2}, v_{n-1}), v_n\}$ in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$. Consider vertices $(u_4, v_5), (u_5, v_6)$, and (u_6, v_7). Since $(u_4, v_5) \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, by Lemma 3.5, both (u_5, v_6) and (u_6, v_7) are in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$.

For $i \in I = \{6, 9, 12, \ldots, n-5\}$, we define S_i, D^2_i, and D^2_R in the same way as in Case 1.1. We construct

$$D^2_N = \{v_1, v_2, (u_3, v_4), (u_4, v_5), (u_5, v_6)\} \cup \bigcup_{i \in I} D^2_i \cup D^2_R \cup \{v_{n+1}, v_{n+2}\}.$$

Similar to Case 1.1 in the proof of Theorem 3.10, for $i = 6, 9, 12, \ldots, n-5$, we can show that each vertex in S_i is dominated by at least two vertices in D^2_N.

Since v_1 is adjacent to v_2, (u_3, v_4) is adjacent to (u_4, v_5), v_{n+1} is adjacent to v_{n+2}, and all six vertices are in D^2_N, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_3 is dominated by v_2 and (u_3, v_4). We also have (u_{n-2}, v_{n-1}) dominated by (u_{n-3}, v_{n-2}) and (u_{n-2}, v_{n-1}), and v_n dominated by (u_{n-2}, v_{n-1}) and v_{n+1}. Then, any vertex in $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$ is dominated by at least two vertices in D^2_N. Hence, D^2_N is a double dominating set of P_{n+2}.

We get the same contradiction as in Case 1.1.

Case 2.2: At least two vertices in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ are in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, and at least two vertices in $\{(u_{n-3}, v_{n-2}), (u_{n-2}, v_{n-1}), v_n\}$ are in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$. Removing u_2 and u_{n-1} affects the domination of (u_3, v_4) and (u_{n-2}, v_{n-1}), respectively. We can show that (u_3, v_4) and (u_{n-2}, v_{n-2}) are dominated by at least two vertices in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$ in the similar way as in Case 1.2. Then, any vertex in P_{n+2} is dominated by at least two vertices in D^2_k. Hence, D^2_k is a double dominating set of P_{n+2}. We get the same contradiction as in Case 1.2.

Case 3: $n + 2 \equiv 2 \pmod 3$. Then, $\left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4 = dd(P_{n+2}) + 4$.

Case 3.1: At most one vertex in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ is in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, or at most one vertex in $\{(u_{n-3}, v_{n-2}), (u_{n-2}, v_{n-1}), v_n\}$ is in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$. By symmetry, we can assume that at most one vertex in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ is in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$.

Similar to Case 1.1, we consider only the case where $(u_3, v_4) \in D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, $v_3 \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, and $(u_4, v_5) \notin D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, and there must be at least one vertex in $\{(u_{n-2}, v_{n-1}), v_n\}$ in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$. Considering in the same way as Case 2.1, both (u_5, v_6) and (u_6, v_7) are in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$.

For $i \in I = \{7, 10, 13, \ldots, n-5\}$, we define S_i, D^2_i, and D^2_R in the same way as in Case 1.1. We construct

$$D^2_N = \{v_1, v_2, (u_3, v_4), (u_4, v_5), (u_6, v_7)\} \cup \bigcup_{i \in I} D^2_i \cup D^2_R \cup \{v_{n+1}, v_{n+2}\}.$$

Similar to Case 1.1 in the proof of Theorem 3.10, for $i = 7, 10, 13, \ldots, n-5$, we can show that each vertex in S_i is dominated by at least two vertices in D^2_N.

Since v_1 is adjacent to v_2, (u_3, v_4) is adjacent to (u_4, v_5), v_{n+1} is adjacent to v_{n+2}, and all six vertices are in D^2_N, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_3 is dominated by v_2 and (u_3, v_4), (u_5, v_6) is dominated by (u_4, v_5) and (u_6, v_7). Besides, (u_6, v_7) is dominated by (u_6, v_7) and (u_7, v_8). We also have (u_{n-2}, v_{n-1}) dominated by (u_{n-3}, v_{n-2}) and (u_{n-2}, v_{n-1}), and v_n dominated by (u_{n-2}, v_{n-1}) and v_{n+1}. Then, any vertex in $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$ is dominated by at least two vertices in D^2_N. Hence, D^2_N is a double dominating set of P_{n+2}.

Since $\{u_1, u_2, u_{n-1}, u_n\} \subseteq D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, we get

$$\begin{aligned} |D^2_N| = |D^2_k| &= |D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}| - |\{u_1, u_2, u_{n-1}, u_n\}| \\ &< dd(P_{n+2}) + 4 - 4 = dd(P_{n+2}). \end{aligned}$$

It contradicts to the fact that $dd(P_{n+2})$ is the minimum cardinality among those of all double dominating sets of P_{n+2}.

Case 3.2: At least two vertices in $\{v_3, (u_3, v_4), (u_4, v_5)\}$ are in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$, and at least two vertices in $\{(u_{n-3}, v_{n-2}), (u_{n-2}, v_{n-1}), v_n\}$ are in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$. Removing u_2 and u_{n-1} affects the domination of (u_3, v_4) and (u_{n-2}, v_{n-1}), respectively. We can show that (u_3, v_4) and (u_{n-2}, v_{n-2}) are dominated by at least two vertices in $D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}$ in the similar way as in Case 1.2. Then, any vertex in P_{n+2} is dominated by at least two vertices in D^2_k. Hence, D^2_k is a double dominating set of P_{n+2}. We get the same contradiction as in Case 3.1.

From all cases, $dd(G(2,2,3,3)) = |D^2_{P_n \lhd H_1 \cong_f H_2 \rhd P_{n+2}}| \geq \left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4$.

For the upper bound, there are two cases depending on remainders of dividing n by 3.

Case $n + 2 \equiv 0 \pmod 3$ and $n + 2 \equiv 1 \pmod 3$: we construct a double dominating set of $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$. Define

$$\begin{aligned} D^2 = &\{u_1, u_2, v_1, v_2, u_{n-1}, u_n, v_{n+1}, v_{n+2}\} \\ &\cup \{(u_i, v_{i+1}) | i \not\equiv 1 (\text{mod} 3) \text{ and } i \in \{3, 4, \ldots, n-2\}\}. \end{aligned}$$

We can calculate $|D^2|$ in the similar way as in the proof of Theorem 3.6, and we get $|D^2| = \left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4$.

For $i = 4, 5, \ldots, n-3$,

if $i \equiv 0 \pmod 3$, vertex (u_i, v_{i+1}) is dominated by $(u_{i-1}, v_{(i+1)-1})$ and (u_i, v_{i+1}),

if $i \equiv 1 \pmod 3$, vertex (u_i, v_{i+1}) is dominated by $(u_{i-1}, v_{(i+1)-1})$ and $(u_{i+1}, v_{(i+1)+1})$, and

if $i \equiv 2 \pmod 3$, vertex (u_i, v_{i+1}) is dominated by (u_i, v_{i+1}) and $(u_{i+1}, v_{(i+1)+1})$.

Since u_1 is adjacent to u_2, v_1 is adjacent to v_2, u_{n-1} is adjacent to u_n, v_{n+1} is adjacent to v_{n+2}, and all eight vertices are in D^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_3 is dominated by v_2 and (u_3, v_4), and (u_3, v_4) is dominated by (u_3, v_4) and u_2. We also have (u_{n-2}, v_{n-1}) dominated by (u_{n-2}, v_{n-1}) and u_{n-1}, and v_n dominated by (u_{n-2}, v_{n-1}) and v_{n+1}. Then, any vertex in $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$ is dominated by at least two vertices in D^2. Therefore, D^2 is a double dominating set of $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$.

Case $n+2 \equiv 2 \pmod 3$: we construct a double dominating set of $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$. Define

$$D^2 = \{u_1, u_2, v_1, v_2, u_{n-1}, u_n, , v_n, v_{n+1}, v_{n+2}\} \cup \{(u_i, v_{i+1}) | i \not\equiv 1 (\text{mod} 3) \text{ and } i \in \{3, 4, \ldots, n-2\}\}.$$

We can calculate $|D^2|$ in the similar way as in the proof of Theorem 3.6, and we get $|D^2| = \left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4$.

For $i = 4, 5, \ldots, n-3$, we can show that (u_i, v_{i_1}) is dominated by at least two vertices in D^2 in the same way as in the previous case.

Since u_1 is adjacent to u_2, v_1 is adjacent to v_2, u_{n-1} is adjacent u_n, v_{n+1} is adjacent to v_{n+2}, and all eight vertices are in D^2, by Observation 3.3, each vertex is dominated by at least two vertices. Moreover, v_3 is dominated by v_2 and (u_3, v_4), and (u_3, v_4) is dominated by (u_3, v_4) and u_2. We also have (u_{n-2}, v_{n-1}) dominated by (u_{n-3}, v_{n-2}) and u_{n-1}, and v_n dominated by v_n and v_{n+1}. Then, any vertex in $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$ is dominated by at least two vertices in D^2. Therefore, D^2 is a double dominating set of $P_n \lhd_{H_1 \cong_f H_2} \rhd P_{n+2}$.

From both cases, $dd(G(2,2,3,3)) \leq |D^2| = \left\lfloor \frac{2(n+2)+2}{3} \right\rfloor + 4$. □

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments.

References

[1] C. Uiyyasathian, Maximal-Clique Partition, (Doctoral dissertation), University of Colorado at Denver, USA, 2003 https://digital.auraria.edu/work/sc/67772794-aea1-40f7-bf6a-6e8d651275fe.

[2] C. Promsakon, C. Uiyyasathian, Chromatic numbers of glued graphs, Thai Journal of Mathematics 4 (3) (2006) 75–81.

[3] C. Promsakon, C. Uiyyasathian, Edge-chromatic numbers of glued graphs, Thai Journal of Mathematics 4 (2) (2006) 395–401.

[4] C. Uiyyasathian, S. Saduakdee, Perfect glued graphs at complete clones, Journal of Mathematics Research 1 (1) (2009) 25–30. (2006) 395–401.

[5] U. Jongthawonwuth, C. Uiyyasathian, Clique partitions of glued graphs, Journal of Mathematics Research 2 (2) (2010) 104–111.

[6] V. Boonthong, P. Putthapiban, P. Chaisuriya, P. Pacheenburawana, Eulerian glued graphs, Thai Journal of Mathematics 8 (1) (2010) 103–109.

[7] C. Berge, Theory of Graphs and its Applications, Wiley, New York, 1964 (Original work published 1958).

[8] T.W. Haynes, S. Hedetniemi, P. Slater, Fundamentals of Domination in Graphs (1st ed.), CRC Press, Boca Raton, 1998.

[9] S. Alikhania, Y. H. Penga, K. A. M. Atana, On the domination number of some graphs, International Mathematical Forum 3 (38) (2008) 1879–1884.

[10] R. Glebov, A. Liebenau, T. Szabó, On the concentration of the domination number of the random graph, SIAM Journal on Discrete Mathematics 29 (3) (2015).

[11] A. Behzada, M. Behzadb, C. E. Praegerc, On the domination number of the generalized Petersen graphs, Discrete Mathematics 308 (2008) 603–610.

[12] P. Boonmee, J. Ma-In, S. Panma, Domination numbers of amalgamations of cycles at connected subgraphs, Journal of Mathematics (2022) 1–11.

[13] M. Ruangnai, S. Panma, The domination number of a graph $Pk((k1, k2), (k3, k4))$, Communications in Mathematics and Applications 10 (2019) 745–762.

[14] P. Sripratak, S. Panma, On the bounds of the domination numbers of glued graphs, Thai Journal of Mathematics 19 (4) (2021) 1719–1728.

[15] F. Harary, T. W. Haynes, Double domination in graphs, Ars. Combin. (55) (2000) 201–213.

[16] M. Blidia, M. Chellali, T. W. Haynes, Characterizations of trees with equal paired and double domination numbers, Discrete Mathematics 306 (2006) 1840–1845.

[17] M. A. Henning, Graphs with large double domination numbers, Discussiones Mathematicae Graph Theory 25 (2005) 13—28.

[18] M. Hajian, N. J. Rad, A new lower bound on the double domination number of a graph, Discrete Applied Mathematics 254 (2019) 280—282.

[19] J. Rani, P. Mor, Domination in graph and some of its applications in various fields, Journal of Emerging Technologies and Innovative Research 6 (4) (2019) 322–326.

[20] A. Aytaç, A. Mutlu, Double domination in shadow graphs, Bull. Int. Math. Virtual Inst., 12 (2) (2022), 345–352.

[21] M. Chellali, T.W. Haynes, A note on the double domination number in trees, AKCE J. Graphs. Combin., 3 (2) (2006) 147–150.

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

Equitable Colorings of Cartesian Products of Fans and Tripartite Graphs

Annob Jobpan[1,*] and Kittikorn Nakprasit[1,†]

[1]Department of Mathematics, Faculty of Science,
Khon Kaen University, Khon Kaen 40002, Thailand

Abstract

A *proper k-coloring* of G is a mapping $c : V(G) \to \{1, 2, \ldots, k\}$ such that $c(x) \neq c(y)$ for $xy \in E(G)$. A graph G is *k-colorable* if there is a proper k-coloring. The *color class* V_i where $i \in \{1, 2, \ldots, k\}$ is the set of vertices with color i.

An *equitable k-coloring* is a proper k-coloring with color classes $V_1, V_2, \ldots, V_k$ for which the numbers of vertices in any two color classes differ by at most 1. A graph G is *equitably k-colorable* if there is an equitable k-coloring. The smallest positive integer k such that G is equitably k-colorable is the *equitable chromatic number* of G, denoted by $\chi_=(G)$. The *equitable chromatic threshold* of G, denoted by $\chi^*_=(G)$, is the smallest positive integer n such that G is equitably n'-colorable for all $n' \geq n$.

In this work, we study the bounds or exact values of the equitable chromatic number and the equitable chromatic threshold of the Cartesian products of Fans and complete tripartite graphs.

Keywords: Cartesian product, equitable coloring, equitable chromatic number, equitable chromatic threshold

2020 MSC: 05C15, 05C76

1 Introduction

Throughout this paper, the graphs are finite, simple, and undirected graphs. If $G = (V(G), E(G))$ or $G = (V, E)$ is a graph, $V(G)$ and $E(G)$ denote its vertex set and edge set, respectively. For a real number x, let $\lceil x \rceil$ and $\lfloor x \rfloor$ denote the smallest integer greater than or equal to x and the largest integer less than or equal to x, respectively.

A *proper k-coloring* of G is a mapping $c : V(G) \to \{1, 2, \ldots, k\}$ such that $c(x) \neq c(y)$ for $xy \in E(G)$. A graph G is *k-colorable* if there is a proper k-coloring. The *color class* V_i where $i \in \{1, 2, \ldots, k\}$ is the set of vertices with color i. The *chromatic number* of G, denoted by $\chi(G)$, is the smallest positive integer k such that G is k-colorable. An *equitable k-coloring* is a proper k-coloring with color classes $V_1, V_2, \ldots, V_k$ for which the numbers of vertices in any two color classes differ by at most 1. A graph G is *equitably k-colorable* if there is an equitable k-coloring. If a graph G has n vertices that are equitably k-colorable, then each color class $V_1, V_2, \ldots, V_k$ is of size

*Speaker (annobjo@kkumail.com)
†Corresponding author (kitnak@kku.ac.th)

$\lfloor \frac{n}{k} \rfloor$ or $\lceil \frac{n}{k} \rceil$; more precisely, the color class V_i has size $\lfloor \frac{n+i-1}{k} \rfloor (= \lceil \frac{n-k+i}{k} \rceil)$ for $i \in \{1, 2, \ldots, k\}$ when $|V_1| \leq |V_2| \leq \cdots \leq |V_k|$. The smallest positive integer k such that G is equitably k-colorable is the *equitable chromatic number* of G, denoted by $\chi_=(G)$. The *equitable chromatic threshold* of G, denoted by $\chi^*_=(G)$, is the smallest positive integer n such that G is equitably n'-colorable for all $n' \geq n$. It is evident from the definition that $\chi(G) \leq \chi_=(G) \leq \chi^*_=(G)$ for any graph G.

In 1998, Chen, Lih, and Yan [2] studied the equitable chromatic numbers of Cartesian products of paths, cycles, and complete graphs. In 2006, Furmańczyk [3] studied the equitable coloring in Cartesian products of hypercubes, cycles, paths, bipartite graphs, tripartite graphs, and r-partite graphs. In 2012, Lin and Chang [4] obtained exact values or upper bounds of $\chi_=(G \square H)$ and $\chi^*_=(G \square H)$ when G and H are stars, paths, cycles, or complete bipartite graphs. In 2014, Yan, Lin, and Wang [5] proved a conjecture of Lin and Chang [4] which states that the Cartesian product of any bipartite graphs G and H is equitably k-colorable for $k \geq 4$. In 2014, Zuo, Wu, and Zhang [6] studied the equitable coloring in Cartesian products of wheels with complete bipartite graphs.

In 2015, Zuo, Wu, and Zhang [7] studied the equitable coloring in Cartesian products of Fans with bipartite graphs. In the present paper, we study the bounds or exact values of the equitable chromatic number and the equitable chromatic threshold of the Cartesian products of Fans and complete tripartite graphs. The results of Fans with 4 or 5 vertices generalize the results in [7] with the same number of vertices.

2 Preliminaries

This section presents definitions that are essential in this work. For definitions of graph-related terminology not mentioned in this paper, we refer the reader to [1].

Definition 2.1. ([1]) For an integer $n \geq 1$, a *path* P_n is a graph of order n and size $n-1$ whose vertices can be labeled by $v_1, v_2, \ldots, v_n$ with edges $v_i v_{i+1}$ for $i = 1, 2, \ldots, n-1$.

Definition 2.2. ([1]) A graph G is a *complete graph* if every two distinct vertices are adjacent. The complete graph of order n is denoted by K_n.

Definition 2.3. ([1]) A graph $G(V_1, V_2, \ldots, V_k)$ is a *k-partite graph* if $V(G)$ can be partitioned into k independent sets (no two vertices in the same set are adjacent) $V_1, V_2, \ldots, V_k$ (called *partite sets*). The graph $G(V_1, V_2)$ is called a *bipartite graph* and $G(V_1, V_2, V_3)$ is called a *tripartite graph*.

Definition 2.4. ([1]) A *complete k-partite graph* G is a k-partite graph where two vertices are adjacent if and only if they belong to different partite sets. If $|V_i| = n_i$ for $1 \leq i \leq k$, then G is denoted by $K_{n_1, n_2, \ldots, n_k}$ where the order of $n_1, n_2, \ldots, n_k$ does not matter. The graph K_{n_1, n_2} is called a *complete bipartite graph* and K_{n_1, n_2, n_3} is called a *complete tripartite graph*.

Definition 2.5. A *Fan graph* $F_{n'+1}$ is defined as the join graph $K_1 \vee P_{n'}$.

Definition 2.6. ([7]) The *Cartesian product* of graphs $G_1(V_1, E_1)$ and $G_2(V_2, E_2)$ denoted by $G_1 \square G_2$, has the vertex set $\{(u, x) \,|\, u \in V_1, x \in V_2\}$ and $\{(u, x), (v, y)\}$ is an edge if and only if $u = v$ and $xy \in E_2$ or $x = y$ and $uv \in E_1$.

The following results by Zuo, Wu, and Zhang [7] are the inspiration of our study. It presented bounds or exact values of the equitable chromatic number and the equitable chromatic threshold of the Cartesian products $F_{n'+1}\Box K_{m,n}$.

Let s, l, m, n, and n' be positive integers where $m \geq n$.

Theorem 2.7. *([7]) Suppose that $m \geq n \geq 1, l \geq 2$, and $k \geq 4$. If $n' = 2l$, then $F_{n'+1}\Box K_{m,n}$ is equitably k-colorable.*

Theorem 2.8. *([7]) Suppose that $m \geq n \geq 1$.*
(1) If $n' = 4$ and $2n - m \in \{0, \pm 1, \pm 2\}$, or
(2) if $n' = 2l$ where $l \geq 3$ and $m = 2n$,
then $F_{n'+1}\Box K_{m,n}$ is equitably 3-colorable, and $\chi^*_=(F_{n'+1}\Box K_{m,n}) = 3$.

Theorem 2.9. *([7]) Suppose that $m \geq n \geq 1$, $l \geq 1$, and $k \geq 4$. If $n' = 2l+1$, then $F_{n'+1}\Box K_{m,n}$ is equitably k-colorable.*

Theorem 2.10. *([7]) Suppose that $m \geq n \geq 1$.*
(1) If $n' = 3$ and $2n - m \in \{0, \pm 1, \pm 2\}$, or
(2) if $n' = 2l + 1$ where $l \geq 2$ and $m = 2n$,
then $F_{n'+1}\Box K_{m,n}$ is equitably 3-colorable, and $\chi^*_=(F_{n'+1}\Box K_{m,n}) = 3$.

Inspired by these results, in Section 3, we find the bounds or exact values of the equitable chromatic number and the equitable chromatic threshold of the Cartesian products of Fans and complete tripartite graphs.

3 Main Results

Let l, r, s, t, and n' be positive integers. A graph $F_{n'+1}$ represents the Fan with vertex set $V(F_{n'+1}) = \{v, v_1, v_2, ..., v_{n'}\}$ and graph $K_{r,s,t}$ represents a complete tripartite graph with three parts $X = \{x_1, x_2, ..., x_r\}$, $Y = \{y_1, y_2, ..., y_s\}$, and $Z = \{z_1, z_2, ..., z_t\}$ where $r \leq s \leq t$. We find the bounds or exact values of $\chi_=(F_{n'+1}\Box K_{r,s,t})$ and $\chi^*_=(F_{n'+1}\Box K_{r,r,r})$. Clearly, $\chi^*_=(F_{n'+1}\Box K_{r,s,t}) \geq \chi_=(F_{n'+1}\Box K_{r,s,t}) \geq 3$.

We use the following the notation and definition in the proofs.

Definition 3.1. Let $P_1, P_2, \ldots, P_\alpha, Q_1, Q_2, \ldots, Q_\beta$ be sets of vertices of a graph.

We use the notation $P_i \to Q_{j_1}, Q_{j_2}, \ldots, Q_{j_k}$ when vertices in P_i are adjacent to only vertices in $Q_{j_1}, Q_{j_2}, \ldots, Q_{j_k}$. Furthermore, the *gap-size* between P_i and Q_j, denoted by $S(P_i, Q_j)$, is $|P_{i+1}| + \cdots + |P_\alpha| + |Q_1| + \cdots + |Q_{j-1}|$.

Theorem 3.2. *Suppose that $1 \leq r \leq s \leq t$.*
(1) If $n' = 2$, or
(2) if $n' = 4$ and $t - r \leq 1$, or
(3) if $n' = 2l$ where $l \geq 3$ and $r = s = t$,
then $F_{n'+1}\Box K_{r,s,t}$ is equitably 3-colorable.

Proof. The Cartesian product graph $F_{n'+1}\Box K_{r,s,t}$ has its structure illustrated in Figure 1.

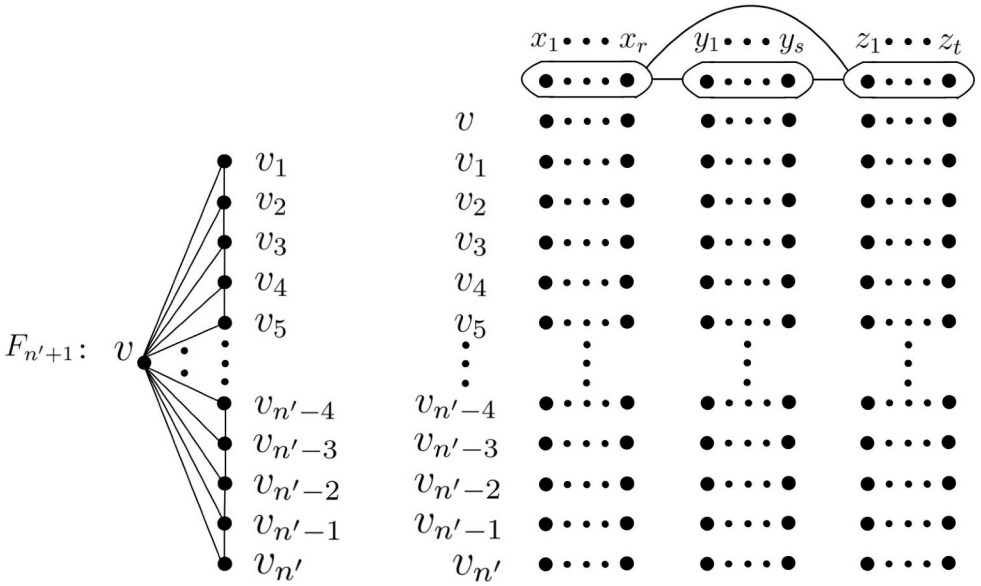


Figure 1: The Cartesian product graph $F_{n'+1}\Box K_{r,s,t}$

(1) Assume that $n' = 2$. We can partition the vertex set $V(F_{n'+1}\Box K_{r,s,t})$ into the following three independent sets:

$$\begin{aligned}
V_1 = &\ \{(v, x_j) : 1 \le j \le r\} \cup \{(v_1, y_j) : 1 \le j \le s\} \cup \{(v_2, z_j) : 1 \le j \le t\},\\
V_2 = &\ \{(v, y_j) : 1 \le j \le s\} \cup \{(v_1, z_j) : 1 \le j \le t\} \cup \{(v_2, x_j) : 1 \le j \le r\}, \text{ and}\\
V_3 = &\ \{(v, z_j) : 1 \le j \le t\} \cup \{(v_1, x_j) : 1 \le j \le r\} \cup \{(v_2, y_j) : 1 \le j \le s\}.
\end{aligned}$$

Then $|V_1| = |V_2| = |V_3| = r + s + t$.

Therefore, $F_{n'+1}\Box K_{r,s,t}$ is equitably 3-colorable.

(2) Assume that $n' = 4$ and $t - r \le 1$. We can partition the vertex set $V(F_{n'+1}\Box K_{r,s,t})$ into the following three independent sets:

$$\begin{aligned}
V_1 = &\ \{(v, x_j) : 1 \le j \le r\} \cup \{(v_i, y_j) : i = 1, 3 \text{ and } 1 \le j \le s\} \cup\\
&\ \{(v_i, z_j) : i = 2, 4 \text{ and } 1 \le j \le t\},\\
V_2 = &\ \{(v, y_j) : 1 \le j \le s\} \cup \{(v_i, z_j) : i = 1, 3 \text{ and } 1 \le j \le t\} \cup\\
&\ \{(v_i, x_j) : i = 2, 4 \text{ and } 1 \le j \le r\}, \text{ and}\\
V_3 = &\ \{(v, z_j) : 1 \le j \le t\} \cup \{(v_i, x_j) : i = 1, 3 \text{ and } 1 \le j \le r\} \cup\\
&\ \{(v_i, y_j) : i = 2, 4 \text{ and } 1 \le j \le s\}.
\end{aligned}$$

Then $|V_1| = r + 2s + 2t, |V_2| = s + 2t + 2r$, and $|V_3| = t + 2r + 2s$.

Since $t - r \le 1$, this partition is an equitable 3-coloring.

Therefore, $F_{n'+1}\Box K_{r,s,t}$ is equitably 3-colorable.

(3) Assume that $n' = 2l$ where $l \ge 3$ and $r = s = t$. We can partition the vertex set

$V(F_{n'+1}\square K_{r,s,t})$ into the following three independent sets:

$$\begin{aligned}
V_1 = &\ \{(v,x_j) : 1 \le j \le r\} \cup \{(v_i,y_j) : i = 1,3,\dots,2l-1 \text{ and } 1 \le j \le s\} \cup \\
&\ \{(v_i,z_j) : i = 2,4,\dots,2l \text{ and } 1 \le j \le t\}, \\
V_2 = &\ \{(v,y_j) : 1 \le j \le s\} \cup \{(v_i,z_j) : i = 1,3,\dots,2l-1 \text{ and } 1 \le j \le t\} \cup \\
&\ \{(v_i,x_j) : i = 2,4,\dots,2l \text{ and } 1 \le j \le r\}, \text{ and} \\
V_3 = &\ \{(v,z_j) : 1 \le j \le t\} \cup \{(v_i,x_j) : i = 1,3,\dots,2l-1 \text{ and } 1 \le j \le r\} \cup \\
&\ \{(v_i,y_j) : i = 2,4,\dots,2l \text{ and } 1 \le j \le s\}.
\end{aligned}$$

Then $|V_1| = r + ls + lt, |V_2| = s + lt + lr$, and $|V_3| = t + lr + ls$.
Since $r = s = t$, it follows that $F_{n'+1}\square K_{r,s,t}$ is equitably 3-colorable. □

Theorem 3.3. *Suppose that* $1 \le r \le s \le t$.
(1) If $n' = 1$ *or* $n' = 3$ *and* $t - r \le 1$, *or*
(2) if $n' = 2l + 1$ *where* $l \ge 2$ *and* $r = s = t$,
then $F_{n'+1}\square K_{r,s,t}$ *is equitably* 3*-colorable.*

Proof. (1) Assume that $n' = 1$ and $t - r \le 1$. We can partition the vertex set $V(F_{n'+1}\square K_{r,s,t})$ into the following three independent sets:

$$\begin{aligned}
V_1 = &\ \{(v,x_j) : 1 \le j \le r\} \cup \{(v_1,y_j) : 1 \le j \le s\}, \\
V_2 = &\ \{(v,y_j) : 1 \le j \le s\} \cup \{(v_1,z_j) : 1 \le j \le t\}, \text{ and} \\
V_3 = &\ \{(v,z_j) : 1 \le j \le t\} \cup \{(v_1,x_j) : 1 \le j \le r\}.
\end{aligned}$$

Then $|V_1| = r + s, |V_2| = s + t$, and $|V_3| = t + r$.
Since $t - r \le 1$, this partition is an equitable 3-coloring.
Therefore, $F_{n'+1}\square K_{r,s,t}$ is equitably 3-colorable.
Assume that $n' = 3$ and $t - r \le 1$. We can partition the vertex set $V(F_{n'+1}\square K_{r,s,t})$ into the following three independent sets:

$$\begin{aligned}
V_1 = &\ \{(v,x_j) : 1 \le j \le r\} \cup \{(v_i,y_j) : i = 1,3 \text{ and } 1 \le j \le s\} \cup \{(v_2,z_j) : 1 \le j \le t\}, \\
V_2 = &\ \{(v,y_j) : 1 \le j \le s\} \cup \{(v_i,z_j) : i = 1,3 \text{ and } 1 \le j \le t\} \cup \{(v_2,x_j) : 1 \le j \le r\}, \text{ and} \\
V_3 = &\ \{(v,z_j) : 1 \le j \le t\} \cup \{(v_i,x_j) : i = 1,3 \text{ and } 1 \le j \le r\} \cup \{(v_2,y_j) : 1 \le j \le s\}
\end{aligned}$$

Then $|V_1| = r + 2s + t, |V_2| = s + 2t + r$, and $|V_3| = t + 2r + s$.
Since $t - r \le 1$, this partition is an equitable 3-coloring.
Therefore, $F_{n'+1}\square K_{r,s,t}$ is equitably 3-colorable.
(2) Assume that $n' = 2l + 1$ where $l \ge 2$ and $r = s = t$. We can partition the vertex set $V(F_{n'+1}\square K_{r,s,t})$ into the following three independent sets:

$$\begin{aligned}
V_1 = &\ \{(v,x_j) : 1 \le j \le r\} \cup \{(v_i,y_j) : i = 1,3,\dots,2l+1 \text{ and } 1 \le j \le s\} \cup \\
&\ \{(v_i,z_j) : i = 2,4,\dots,2l \text{ and } 1 \le j \le t\}, \\
V_2 = &\ \{(v,y_j) : 1 \le j \le s\} \cup \{(v_i,z_j) : i = 1,3,\dots,2l+1 \text{ and } 1 \le j \le t\} \cup \\
&\ \{(v_i,x_j) : i = 2,4,\dots,2l \text{ and } 1 \le j \le r\}, \text{ and} \\
V_3 = &\ \{(v,z_j) : 1 \le j \le t\} \cup \{(v_i,x_j) : i = 1,3,\dots,2l+1 \text{ and } 1 \le j \le r\} \cup \\
&\ \{(v_i,y_j) : i = 2,4,\dots,2l \text{ and } 1 \le j \le s\}.
\end{aligned}$$

Then $|V_1| = r + (l+1)s + lt, |V_2| = s + (l+1)t + lr$, and $|V_3| = t + (l+1)r + ls$.
Since $r = s = t$, it follows that $F_{n'+1}\square K_{r,s,t}$ is equitably 3-colorable. □

Theorem 3.4. *(1) If $n' = 2$ and $k \geq 6$, then $F_{n'+1}\Box K_{r,r,r}$ is equitably k-colorable.*
(2) If $n' = 4$ and $k \geq 3$, then $F_{n'+1}\Box K_{r,r,r}$ is equitably k-colorable.

Proof. (1) Assume that $n' = 2$.

Setting. We sort the vertices of $F_{n'+1}\Box K_{r,r,r}$ as follows:

$$(v, x_1), (v, x_2), \ldots, (v, x_r), (v_1, y_1), (v_1, y_2), \ldots, (v_1, y_r), (v_2, z_1), (v_2, z_2), \ldots, (v_2, z_r),$$
$$(v, y_1), (v, y_2), \ldots, (v, y_r), (v_1, z_1), (v_1, z_2), \ldots, (v_1, z_r), (v_2, x_1), (v_2, x_2), \ldots, (v_2, x_r),$$
$$(v, z_1), (v, z_2), \ldots, (v, z_r), (v_1, x_1), (v_1, x_2), \ldots, (v_1, x_r), (v_2, y_1), (v_2, y_2), \ldots, (v_2, y_r).$$

We can partition the vertex set $V(F_{n'+1}\Box K_{r,r,r})$ into V_1, V_2, and V_3 such that

$$V_1 = A_1 \cup A_2 \cup A_3$$

where $A_1 = \{(v, x_j) : 1 \leq j \leq r\}$, $A_2 = \{(v_1, y_j) : 1 \leq j \leq r\}$, $A_3 = \{(v_2, z_j) : 1 \leq j \leq r\}$,

$$V_2 = B_1 \cup B_2 \cup B_3$$

where $B_1 = \{(v, y_j) : 1 \leq j \leq r\}$, $B_2 = \{(v_1, z_j) : 1 \leq j \leq r\}$, $B_3 = \{(v_2, x_j) : 1 \leq j \leq r\}$, and

$$V_3 = C_1 \cup C_2 \cup C_3$$

where $C_1 = \{(v, z_j) : 1 \leq j \leq r\}$, $C_2 = \{(v_1, x_j) : 1 \leq j \leq r\}$, $C_3 = \{(v_2, y_j) : 1 \leq j \leq r\}$. Then $|V_1| = |V_2| = |V_3| = 3r$. Note that V_1, V_2, and V_3 are independent sets.

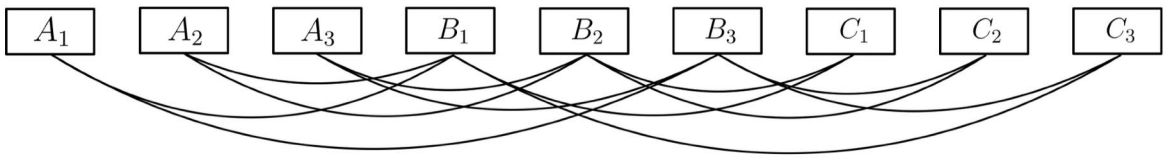


Figure 2: Theorem 3.4 when $n' = 2$

(See Figure 2.) Consider $A_1, A_2, A_3, B_1, B_2, B_3$, we have that

$$A_1 \to B_1, B_3; \quad A_2 \to B_1, B_2, \text{ and } A_3 \to B_2, B_3.$$

It is straightforward to see that $S(A_1, B_1) = 2r, S(A_1, B_3) = 4r, S(A_2, B_1) = r, S(A_2, B_2) = 2r, S(A_3, B_2) = r$, and $S(A_3, B_3) = 2r$.

Consequently, the minimum value of $S(A_i, B_j)$ above is r.

(See Figure 2.) Consider $B_1, B_2, B_3, C_1, C_2, C_3$, we have that

$$B_1 \to C_1, C_3; \quad B_2 \to C_1, C_2, \text{ and } B_3 \to C_2, C_3.$$

It is straightforward to see that $S(B_1, C_1) = 2r, S(B_1, C_3) = 4r, S(B_2, C_1) = r, S(B_2, C_2) = 2r, S(B_3, C_2) = r$, and $S(B_3, C_3) = 2r$.

Consequently, the minimum value of $S(B_i, C_j)$ above is r. It follows that any set of consecutive vertices of size at most r is independent. Since $n' = 2$, we have that $|V(F_{n'+1}\Box K_{r,r,r})| = 9r$.

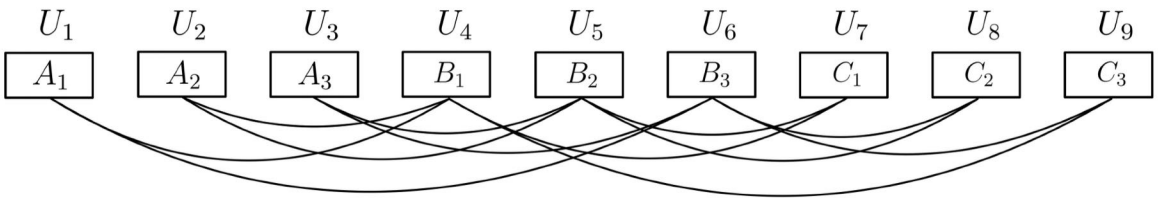


Figure 3: Theorem 3.4 when $n' = 2$ and $k = 9$

Let

$$\sigma_t = \left\lfloor \frac{9r + t - 1}{k} \right\rfloor \text{ where } 1 \leq t \leq k.$$

Case 1: $k \geq 9$. We have that

$$\sigma_t = \left\lfloor \frac{9r + t - 1}{k} \right\rfloor \leq \sigma_k = \left\lfloor \frac{9r + k - 1}{k} \right\rfloor = \left\lceil \frac{9r}{k} \right\rceil \leq \left\lceil \frac{9r}{9} \right\rceil = r.$$

Thus, $\sigma_t \leq \sigma_k \leq r$ for $1 \leq t \leq k$. We partition the vertex set $V(F_{n'+1} \square K_{r,r,r})$ into $U_1, U_2, \ldots, U_k$ according to the sorting in the setting where

$$|U_i| = \left\lfloor \frac{9r + i - 1}{k} \right\rfloor \text{ where } 1 \leq i \leq k.$$

Since $|U_i| = \sigma_i \leq r$ for each $i \in \{1, 2, \ldots, k\}$, it follows that each U_i is an independent set from the observation above. (See Figure 3.)

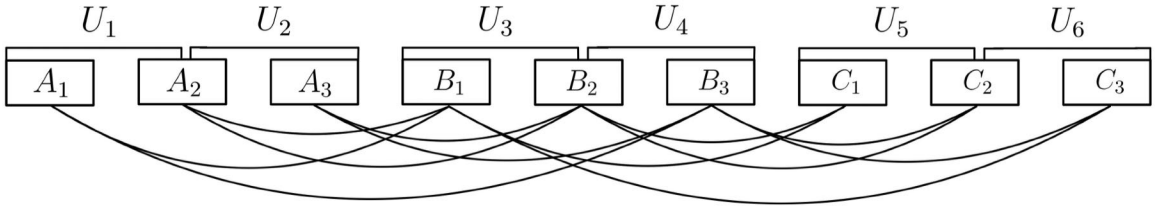


Figure 4: Theorem 3.4 when $n' = 2$ and $k = 6$

Case 2: $k = 6$. We partition the vertex set $V(F_{n'+1} \square K_{r,r,r})$ into U_1, U_2, U_3, U_4, U_5, and U_6. (See Figure 4.)

Let

$$U_1 = A_1 \cup A_2'$$

where $A_2' = \{(v_1, y_j) : 1 \leq j \leq \lfloor \frac{r}{2} \rfloor\}$,

$$U_2 = A_2'' \cup A_3$$

where $A_2'' = \{(v_1, y_j) : \lfloor \frac{r}{2} \rfloor + 1 \leq j \leq r\}$,

$$U_3 = B_1 \cup B_2'$$

where $B_2' = \{(v_1, z_j) : 1 \leq j \leq \lfloor \frac{r}{2} \rfloor\}$,

$$U_4 = B_2'' \cup B_3$$

where $B_2'' = \{(v_1, z_j) : \lfloor \frac{r}{2} \rfloor + 1 \leq j \leq r\}$,

$$U_5 = C_1 \cup C_2'$$

where $C_2' = \{(v_1, x_j) : 1 \leq j \leq \lfloor \frac{r}{2} \rfloor\}$, and

$$U_6 = C_2'' \cup C_3$$

where $C_2'' = \{(v_1, x_j) : \lfloor \frac{r}{2} \rfloor + 1 \leq j \leq r\}$. It follows that $|U_1| = |U_3| = |U_5| = r + \lfloor \frac{r}{2} \rfloor$ and $|U_2| = |U_4| = |U_6| = r + (r - \lfloor \frac{r}{2} \rfloor) = 2r - \lfloor \frac{r}{2} \rfloor$. Thus, the number of vertices in any two sets differ by at most 1. Since $U_1, U_2 \subseteq V_1, U_3, U_4 \subseteq V_2$, and $U_5, U_6 \subseteq V_3$, it follows that U_1, U_2, U_3, U_4, U_5, and U_6 are independent sets.

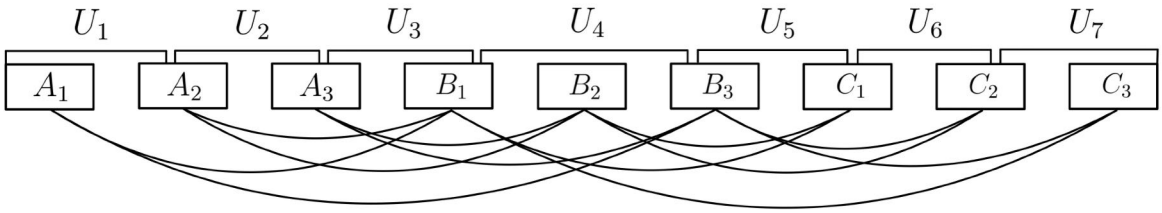


Figure 5: Theorem 3.4 when $n' = 2$ and $k = 7$

Case 3: $k = 7$. (See Figure 5.) We partition the vertex set $V(F_{n'+1} \square K_{r,r,r})$ into $U_1, U_2, U_3, U_4, U_5, U_6$, and U_7 according to the sorting in the setting and

$$|U_i| = \left\lfloor \frac{9r + i - 1}{7} \right\rfloor \text{ for each } i \in \{1, 2, 3, 4, 5, 6, 7\}.$$

It follows that $U_1 \subseteq A_1 \cup A_2, U_2 \subseteq A_2 \cup A_3, U_3 \subseteq A_3 \cup B_1, U_4 \subseteq B_1 \cup B_2 \cup B_3, U_5 \subseteq B_3 \cup C_1, U_6 \subseteq C_1 \cup C_2$, and $U_7 \subseteq C_2 \cup C_3$. Since $U_1, U_2 \subseteq V_1, U_4 \subseteq V_2$, and $U_6, U_7 \subseteq V_3$, it follows that U_1, U_2, U_4, U_6, and U_7 are independent sets. Since vertices in A_3 and B_1 are not adjacent, it follows that U_3 is an independent set. Since vertices in B_3 and C_1 are not adjacent, it follows that U_5 is an independent set.

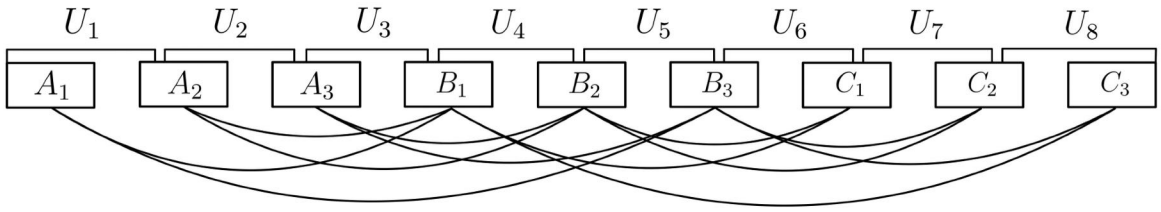


Figure 6: Theorem 3.4 when $n' = 2$ and $k = 8$

Case 4: $k = 8$. (See Figure 6.) We partition the vertex set $V(F_{n'+1}\square K_{r,r,r})$ into $U_1, U_2, U_3, U_4, U_5, U_6, U_7$, and U_8 according to the sorting in the setting and

$$|U_i| = \left\lfloor \frac{9r + i - 1}{8} \right\rfloor \text{ for each } i \in \{1, 2, 3, 4, 5, 6, 7, 8\}.$$

It follows that $U_1 \subseteq A_1 \cup A_2, U_2 \subseteq A_2 \cup A_3, U_3 \subseteq A_3 \cup B_1, U_4 \subseteq B_1 \cup B_2, U_5 \subseteq B_2 \cup B_3, U_6 \subseteq B_3 \cup C_1, U_7 \subseteq C_1 \cup C_2$, and $U_8 \subseteq C_2 \cup C_3$. Since $U_1, U_2 \subseteq V_1, U_4, U_5 \subseteq V_2$, and $U_7, U_8 \subseteq V_3$, it follows that U_1, U_2, U_4, U_5, U_7, and U_8 are independent sets. Since vertices in A_3 and B_1 are not adjacent, it follows that U_3 is an independent set. Since vertices in B_3 and C_1 are not adjacent, it follows that U_6 is an independent set.

Therefore, according to the vertex sorting above, the vertex set of $F_{n'+1}\square K_{r,r,r}$ can be partitioned into k independent sets in a way that the numbers of vertices in any two sets differ by at most 1. Hence $F_{n'+1}\square K_{r,r,r}$ is equitably k-colorable when $k \geq 6$.

(2) Assume that $n' = 4$.

Setting. We sort the vertices of $F_{n'+1}\square K_{r,r,r}$ as follows:

$(v, x_1), (v, x_2), \ldots, (v, x_r), (v_2, z_1), (v_2, z_2), \ldots, (v_2, z_r), (v_1, y_1), (v_1, y_2), \ldots, (v_1, y_r),$
$(v_3, y_1), (v_3, y_2), \ldots, (v_3, y_r), (v_4, z_1), (v_4, z_2), \ldots, (v_4, z_r), (v_2, x_1), (v_2, x_2), \ldots, (v_2, x_r),$
$(v_1, z_1), (v_1, z_2), \ldots, (v_1, z_r), (v, y_1), (v, y_2), \ldots, (v, y_r), (v_4, x_1), (v_4, x_2), \ldots, (v_4, x_r),$
$(v_3, z_1), (v_3, z_2), \ldots, (v_3, z_r), (v_1, x_1), (v_1, x_2), \ldots, (v_1, x_r), (v_2, y_1), (v_2, y_2), \ldots, (v_2, y_r),$
$(v_4, y_1), (v_4, y_2), \ldots, (v_4, y_r), (v_3, x_1), (v_3, x_2), \ldots, (v_3, x_r), (v, z_1), (v, z_2), \ldots, (v, z_r).$

We can partition the vertex set $V(F_{n'+1}\square K_{r,r,r})$ into V_1, V_2, and V_3 such that

$$V_1 = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5$$

where $A_1 = \{(v, x_j) : 1 \leq j \leq r\}$, $A_2 = \{(v_2, z_j) : 1 \leq j \leq r\}$, $A_3 = \{(v_1, y_j) : 1 \leq j \leq r\}$, $A_4 = \{(v_3, y_j) : 1 \leq j \leq r\}$, $A_5 = \{(v_4, z_j) : 1 \leq j \leq r\}$,

$$V_2 = B_1 \cup B_2 \cup B_3 \cup B_4 \cup B_5$$

where $B_1 = \{(v_2, x_j) : 1 \leq j \leq r\}$, $B_2 = \{(v_1, z_j) : 1 \leq j \leq r\}$, $B_3 = \{(v, y_j) : 1 \leq j \leq r\}$, $B_4 = \{(v_4, x_j) : 1 \leq j \leq r\}$, $B_5 = \{(v_3, z_j) : 1 \leq j \leq r\}$, and

$$V_3 = C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_5$$

where $C_1 = \{(v_1, x_j) : 1 \leq j \leq r\}$, $C_2 = \{(v_2, y_j) : 1 \leq j \leq r\}$, $C_3 = \{(v_4, y_j) : 1 \leq j \leq r\}$, $C_4 = \{(v_3, x_j) : 1 \leq j \leq r\}$, $C_5 = \{(v, z_j) : 1 \leq j \leq r\}$. Then $|V_1| = |V_2| = |V_3| = 5r$. Note that V_1, V_2, and V_3 are independent sets.

(See Figure 7.) Consider $A_1, A_2, \ldots, A_5, B_1, B_2, \ldots, B_5$, we have that

$$A_1 \to B_1, B_3, B_4; \quad A_2 \to B_1, B_2, B_5; \quad A_3 \to B_2, B_3; \quad A_4 \to B_3, B_5, \text{ and } A_5 \to B_4, B_5.$$

It is straightforward to see that $S(A_1, B_1) = 4r, S(A_1, B_3) = 6r, S(A_1, B_4) = 7r, S(A_2, B_1) = 3r, S(A_2, B_2) = 4r, S(A_2, B_5) = 7r, S(A_3, B_2) = 3r, S(A_3, B_3) = 4r, S(A_4, B_3) = 3r, S(A_4, B_5) = 5r, S(A_5, B_4) = 3r$, and $S(A_5, B_5) = 4r$.

Consequently, the minimum value of $S(A_i, B_j)$ above is $3r$.

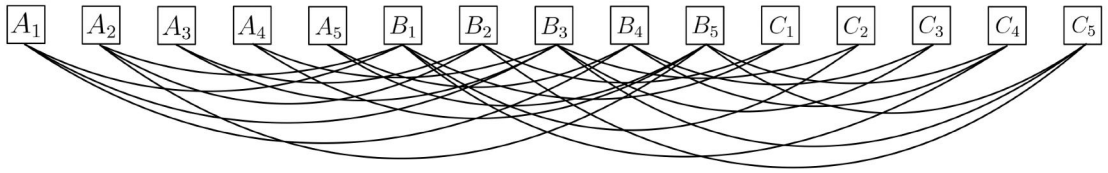


Figure 7: Theorem 3.4 when $n' = 4$

(See Figure 7.) Consider $B_1, B_2, \ldots, B_5, C_1, C_2, \ldots, C_5$, we have that

$$B_1 \to C_1, C_2, C_4; \quad B_2 \to C_1, C_5; \quad B_3 \to C_2, C_3, C_5; \quad B_4 \to C_3, C_4, \text{ and } B_5 \to C_4, C_5.$$

It is straightforward to see that $S(B_1, C_1) = 4r, S(B_1, C_2) = 5r, S(B_1, C_4) = 7r, S(B_2, C_1) = 3r, S(B_2, C_5) = 7r, S(B_3, C_2) = 3r, S(B_3, C_3) = 4r, S(B_3, C_5) = 6r, S(B_4, C_3) = 3r, S(B_4, C_4) = 4r, S(B_5, C_4) = 3r$, and $S(B_5, C_5) = 4r$.

Consequently, the minimum value of $S(B_i, C_j)$ above is $3r$. It follows that any set of consecutive vertices of size at most $3r$ is independent. Since $n' = 4$, we have that $|V(F_{n'+1} \square K_{r,r,r})| = 15r$.

Let

$$\sigma_t = \left\lfloor \frac{15r + t - 1}{k} \right\rfloor \text{ where } 1 \leq t \leq k.$$

Applying Theorem 3.2 (2), we have that $F_{n'+1} \square K_{r,r,r}$ is equitably 3-colorable.

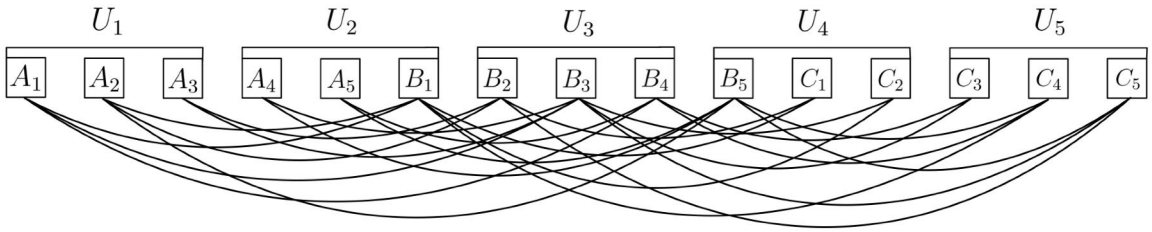


Figure 8: Theorem 3.4 when $n' = 4$ and $k = 5$

Case 1: $k \geq 5$. We have that

$$\sigma_t = \left\lfloor \frac{15r + t - 1}{k} \right\rfloor \leq \sigma_k = \left\lfloor \frac{15r + k - 1}{k} \right\rfloor = \left\lceil \frac{15r}{k} \right\rceil \leq \left\lceil \frac{15r}{5} \right\rceil = 3r.$$

Thus, $\sigma_t \leq \sigma_k \leq 3r$ for $1 \leq t \leq k$. We partition the vertex set $V(F_{n'+1} \square K_{r,r,r})$ into $U_1, U_2, \ldots, U_k$ according to the sorting in the setting where

$$|U_i| = \left\lfloor \frac{15r + i - 1}{k} \right\rfloor \text{ where } 1 \leq i \leq k.$$

Since $|U_i| = \sigma_i \leq 3r$ for each $i \in \{1, 2, \ldots, k\}$, it follows that each U_i is an independent set from the observation above. (See Figure 8.)

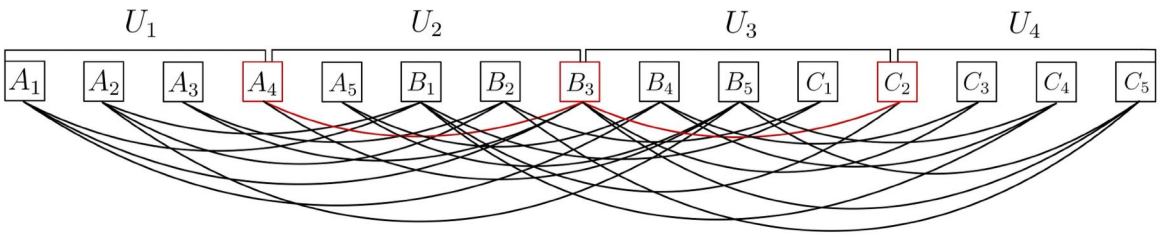


Figure 9: In Theorem 3.4 when $n' = 4$ and $k = 4$, we need to analyze the adjacency of vertices in A_4 and B_3, and the adjacency of vertices in B_3 and C_2

Case 2: $k = 4$. We partition the vertex set $V(F_{n'+1}\square K_{r,r,r})$ into U_1, U_2, U_3, and U_4. (See Figure 9.)

Let

$$U_1 = A_1 \cup A_2 \cup A_3 \cup A_4'$$

where $A_4' = \{(v_3, y_j) : 1 \le j \le \lfloor \frac{3r}{4} \rfloor\}$,

$$U_2 = A_4'' \cup A_5 \cup B_1 \cup B_2 \cup B_3'$$

where $A_4'' = \{(v_3, y_j) : \lfloor \frac{3r}{4} \rfloor + 1 \le j \le r\}$, $B_3' = \{(v, y_j) : 1 \le j \le \lfloor \frac{r}{2} \rfloor\}$,

$$U_3 = B_3'' \cup B_4 \cup B_5 \cup C_1 \cup C_2'$$

where $B_3'' = \{(v, y_j) : \lfloor \frac{r}{2} \rfloor + 1 \le j \le r\}$, $C_2' = \{(v_2, y_j) : 1 \le j \le \lfloor \frac{r}{4} \rfloor\}$, and

$$U_4 = C_2'' \cup C_3 \cup C_4 \cup C_5$$

where $C_2'' = \{(v_2, y_j) : \lfloor \frac{r}{4} \rfloor + 1 \le j \le r\}$. It follows that $|U_1| = 3r + \lfloor \frac{3r}{4} \rfloor, |U_2| = 3r + (r - \lfloor \frac{3r}{4} \rfloor) + \lfloor \frac{r}{2} \rfloor, |U_3| = 3r + (r - \lfloor \frac{r}{2} \rfloor) + \lfloor \frac{r}{4} \rfloor$, and $|U_4| = 3r + (r - \lfloor \frac{r}{4} \rfloor)$. Thus, the number of vertices in any two sets differ by at most 1. Since $U_1 \subseteq V_1$ and $U_4 \subseteq V_3$, it follows that U_1 and U_4 are independent sets.

Consider an edge from $U_2 = A_4'' \cup A_5 \cup B_1 \cup B_2 \cup B_3'$. From the reasons above, an edge may occur only if one endpoint is in A_4'' and the other is in B_3'. That is we have j such that $(v_3, y_j) \in A_4''$ when $j \ge \lfloor \frac{3r}{4} \rfloor + 1$ and $(v, y_j) \in B_3'$ when $j \le \lfloor \frac{r}{2} \rfloor$, a contradiction. We conclude that such an edge does not exist. Hence U_2 is an independent set.

Consider an edge from $U_3 = B_3'' \cup B_4 \cup B_5 \cup C_1 \cup C_2'$. From the reasons above, an edge may occur only if one endpoint is in B_3'' and the other is in C_2'. That is we have j such that $(v, y_j) \in B_3''$ when $j \ge \lfloor \frac{r}{2} \rfloor + 1$ and $(v_2, y_j) \in C_2'$ when $j \le \lfloor \frac{r}{4} \rfloor$, a contradiction. We conclude that such an edge does not exist. Hence U_3 is an independent set.

Therefore, according to the vertex sorting above, the vertex set of $F_{n'+1}\square K_{r,r,r}$ can be partitioned into k independent sets in a way that the numbers of vertices in any two sets differ by at most 1. Hence $F_{n'+1}\square K_{r,r,r}$ is equitably k-colorable when $k \ge 3$. □

Theorem 3.5. *If $n' = 1$ or $n' = 3$, and $k \ge 3$, then $F_{n'+1}\square K_{r,r,r}$ is equitably k-colorable.*

Proof. Assume that $n' = 1$.

Setting. We sort the vertices of $F_{n'+1}\square K_{r,r,r}$ as follows:

$(v_1, y_1), (v_1, y_2), \ldots, (v_1, y_r), (v, x_1), (v, x_2), \ldots, (v, x_r),$
$(v_1, z_1), (v_1, z_2), \ldots, (v_1, z_r), (v, y_1), (v, y_2), \ldots, (v, y_r),$
$(v_1, x_1), (v_1, x_2), \ldots, (v_1, x_r), (v, z_1), (v, z_2), \ldots, (v, z_r).$

We can partition the vertex set $V(F_{n'+1}\square K_{r,r,r})$ into V_1, V_2, and V_3 such that

$$V_1 = A_1 \cup A_2$$

where $A_1 = \{(v_1, y_j) : 1 \leq j \leq r\}$, $A_2 = \{(v, x_j) : 1 \leq j \leq r\}$,

$$V_2 = B_1 \cup B_2$$

where $B_1 = \{(v_1, z_j) : 1 \leq j \leq r\}$, $B_2 = \{(v, y_j) : 1 \leq j \leq r\}$, and

$$V_3 = C_1 \cup C_2$$

where $C_1 = \{(v_1, x_j) : 1 \leq j \leq r\}$, $C_2 = \{(v, z_j) : 1 \leq j \leq r\}$. Then $|V_1| = |V_2| = |V_3| = 2r$. Note that V_1, V_2, and V_3 are independent sets.

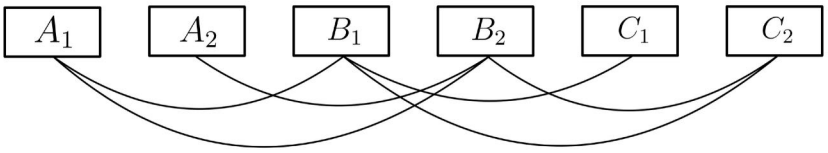


Figure 10: Theorem 3.5 when $n' = 1$

(See Figure 10.) Consider A_1, A_2, B_1, B_2, we have that

$$A_1 \to B_1, B_2 \text{ and } A_2 \to B_2.$$

It is straightforward to see that $S(A_1, B_1) = r, S(A_1, B_2) = 2r$, and $S(A_2, B_2) = r$. Consequently, the minimum value of $S(A_i, B_j)$ above is r.
(See Figure 10.) Consider B_1, B_2, C_1, C_2, we have that

$$B_1 \to C_1, C_2 \text{ and } B_2 \to C_2.$$

It is straightforward to see that $S(B_1, C_1) = r, S(B_1, C_2) = 2r$, and $S(B_2, C_2) = r$.
Consequently, the minimum value of $S(B_i, C_j)$ above is r. It follows that any set of consecutive vertices of size at most r is independent. Since $n' = 1$, we have that $|V(F_{n'+1}\square K_{r,r,r})| = 6r$.
Let

$$\sigma_t = \left\lfloor \frac{6r + t - 1}{k} \right\rfloor \text{ where } 1 \leq t \leq k.$$

Applying Theorem 3.3 (1), we have that $F_{n'+1}\square K_{r,r,r}$ is equitably 3-colorable.
Case 1: $k \geq 6$. We have that

$$\sigma_t = \left\lfloor \frac{6r + t - 1}{k} \right\rfloor \leq \sigma_k = \left\lfloor \frac{6r + k - 1}{k} \right\rfloor = \left\lceil \frac{6r}{k} \right\rceil \leq \left\lceil \frac{6r}{6} \right\rceil = r.$$

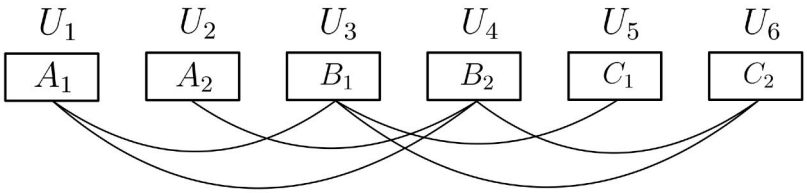


Figure 11: Theorem 3.5 when $n' = 1$ and $k = 6$

Thus, $\sigma_t \leq \sigma_k \leq r$ for $1 \leq t \leq k$. We partition the vertex set $V(F_{n'+1} \square K_{r,r,r})$ into $U_1, U_2, \ldots, U_k$ according to the sorting in the setting where

$$|U_i| = \left\lfloor \frac{6r + i - 1}{k} \right\rfloor \text{ where } 1 \leq i \leq k.$$

Since $|U_i| = \sigma_i \leq r$ for each $i \in \{1, 2, \ldots, k\}$, it follows that each U_i is an independent set from the observation above. (See Figure 11.)

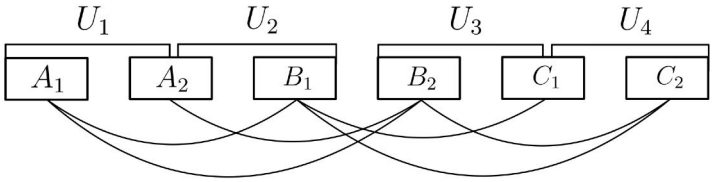


Figure 12: Theorem 3.5 when $n' = 1$ and $k = 4$

Case 2: $k = 4$. We partition the vertex set $V(F_{n'+1} \square K_{r,r,r})$ into U_1, U_2, U_3, and U_4. (See Figure 12.)

Let

$$U_1 = A_1 \cup A_2'$$

where $A_2' = \{(v, x_j) : 1 \leq j \leq \lfloor \frac{r}{2} \rfloor\}$,

$$U_2 = A_2'' \cup B_1$$

where $A_2'' = \{(v, x_j) : \lfloor \frac{r}{2} \rfloor + 1 \leq j \leq r\}$,

$$U_3 = B_2 \cup C_1'$$

where $C_1' = \{(v_1, x_j) : 1 \leq j \leq \lfloor \frac{r}{2} \rfloor\}$, and

$$U_4 = C_1'' \cup C_2$$

where $C_1'' = \{(v_1, x_j) : \lfloor \frac{r}{2} \rfloor + 1 \leq j \leq r\}$. It follows that $|U_1| = |U_3| = r + \lfloor \frac{r}{2} \rfloor$ and $|U_2| = |U_4| = r + (r - \lfloor \frac{r}{2} \rfloor) = 2r - \lfloor \frac{r}{2} \rfloor$. Thus, the number of vertices in any two sets differ by at most 1. Since $U_1 \subseteq V_1$ and $U_4 \subseteq V_3$, it follows that U_1 and U_4 are independent sets. Since vertices in A_2 and B_1 are not adjacent, it follows that U_2 is an independent set. Since vertices in B_2 and C_1 are not adjacent, it follows that U_3 is an independent set.

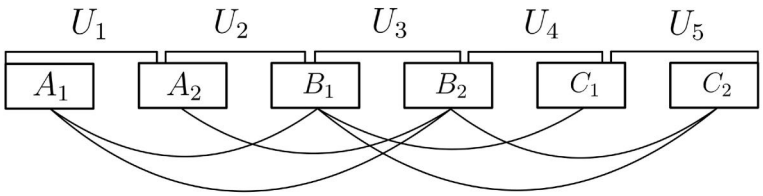


Figure 13: Theorem 3.5 when $n' = 1$ and $k = 5$

Case 3: $k = 5$. (See Figure 13.) We partition the vertex set $V(F_{n'+1}\square K_{r,r,r})$ into U_1, U_2, U_3, U_4, and U_5 according to the sorting in the setting and

$$|U_i| = \left\lfloor \frac{6r + i - 1}{5} \right\rfloor \text{ for each } i \in \{1, 2, 3, 4, 5\}.$$

It follows that $U_1 \subseteq A_1 \cup A_2, U_2 \subseteq A_2 \cup B_1, U_3 \subseteq B_1 \cup B_2, U_4 \subseteq B_2 \cup C_1$, and $U_5 \subseteq C_1 \cup C_2$. Since $U_1 \subseteq V_1, U_3 \subseteq V_2$, and $U_5 \subseteq V_3$, it follows that U_1, U_3, and U_5 are independent sets. Since vertices in A_2 and B_1 are not adjacent, it follows that U_2 is an independent set. Since vertices in B_2 and C_1 are not adjacent, it follows that U_4 is an independent set.

Assume that $n' = 3$.

Setting. We sort the vertices of $F_{n'+1}\square K_{r,r,r}$ as follows:

$(v, x_1), (v, x_2), \ldots, (v, x_r), (v_2, z_1), (v_2, z_2), \ldots, (v_2, z_r), (v_1, y_1), (v_1, y_2), \ldots, (v_1, y_r),$
$(v_3, y_1), (v_3, y_2), \ldots, (v_3, y_r), (v_2, x_1), (v_2, x_2), \ldots, (v_2, x_r), (v_1, z_1), (v_1, z_2), \ldots, (v_1, z_r),$
$(v, y_1), (v, y_2), \ldots, (v, y_r), (v_3, z_1), (v_3, z_2), \ldots, (v_3, z_r), (v_1, x_1), (v_1, x_2), \ldots, (v_1, x_r),$
$(v_2, y_1), (v_2, y_2), \ldots, (v_2, y_r), (v_3, x_1), (v_3, x_2), \ldots, (v_3, x_r), (v, z_1), (v, z_2), \ldots, (v, z_r).$

We can partition the vertex set $V(F_{n'+1}\square K_{r,r,r})$ into V_1, V_2, and V_3 such that

$$V_1 = A_1 \cup A_2 \cup A_3 \cup A_4$$

where $A_1 = \{(v, x_j) : 1 \le j \le r\}$, $A_2 = \{(v_2, z_j) : 1 \le j \le r\}$, $A_3 = \{(v_1, y_j) : 1 \le j \le r\}$, $A_4 = \{(v_3, y_j) : 1 \le j \le r\}$,

$$V_2 = B_1 \cup B_2 \cup B_3 \cup B_4$$

where $B_1 = \{(v_2, x_j) : 1 \le j \le r\}$, $B_2 = \{(v_1, z_j) : 1 \le j \le r\}$, $B_3 = \{(v, y_j) : 1 \le j \le r\}$, $B_4 = \{(v_3, z_j) : 1 \le j \le r\}$, and

$$V_3 = C_1 \cup C_2 \cup C_3 \cup C_4$$

where $C_1 = \{(v_1, x_j) : 1 \le j \le r\}$, $C_2 = \{(v_2, y_j) : 1 \le j \le r\}$, $C_3 = \{(v_3, x_j) : 1 \le j \le r\}$, $C_4 = \{(v, z_j) : 1 \le j \le r\}$. Then $|V_1| = |V_2| = |V_3| = 4r$. Note that V_1, V_2, and V_3 are independent sets.

(See Figure 14.) Consider $A_1, A_2, A_3, A_4, B_1, B_2, B_3, B_4$, we have that

$$A_1 \to B_1, B_3; \quad A_2 \to B_1, B_2, B_4; \quad A_3 \to B_2, B_3, \text{ and } A_4 \to B_3, B_4.$$

It is straightforward to see that $S(A_1, B_1) = 3r, S(A_1, B_3) = 5r, S(A_2, B_1) = 2r, S(A_2, B_2) = 3r, S(A_2, B_4) = 5r, S(A_3, B_2) = 2r, S(A_3, B_3) = 3r, S(A_4, B_3) = 2r$, and $S(A_4, B_4) = 3r$.

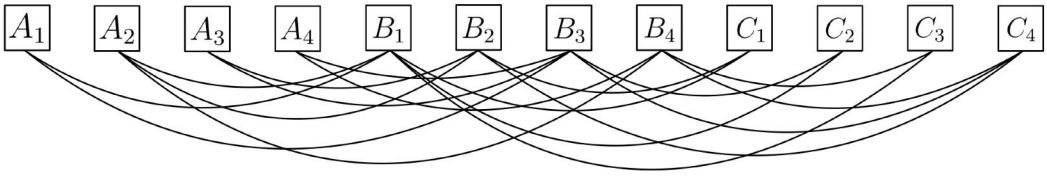


Figure 14: Theorem 3.5 when $n' = 3$

Consequently, the minimum value of $S(A_i, B_j)$ above is $2r$.
(See Figure 14.) Consider $B_1, B_2, B_3, B_4, C_1, C_2, C_3, C_4$, we have that

$$B_1 \to C_1, C_2, C_3; \quad B_2 \to C_1, C_4; \quad B_3 \to C_2, C_4, \text{ and } B_4 \to C_3, C_4.$$

It is straightforward to see that $S(B_1, C_1) = 3r, S(B_1, C_2) = 4r, S(B_1, C_3) = 5r, S(B_2, C_1) = 2r, S(B_2, C_4) = 5r, S(B_3, C_2) = 2r, S(B_3, C_4) = 4r, S(B_4, C_3) = 2r$, and $S(B_4, C_4) = 3r$.

Consequently, the minimum value of $S(B_i, C_j)$ above is $2r$. It follows that any set of consecutive vertices of size at most $2r$ is independent. Since $n' = 3$, we have that $|V(F_{n'+1} \square K_{r,r,r})| = 12r$.

Let

$$\sigma_t = \left\lfloor \frac{12r + t - 1}{k} \right\rfloor \text{ where } 1 \leq t \leq k.$$

Applying Theorem 3.3 (1), we have that $F_{n'+1} \square K_{r,r,r}$ is equitably 3-colorable.

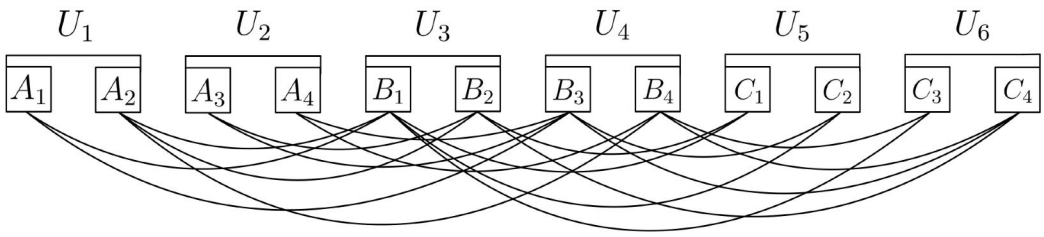


Figure 15: Theorem 3.5 when $n' = 3$ and $k = 6$

Case 1: $k \geq 6$. We have that

$$\sigma_t = \left\lfloor \frac{12r + t - 1}{k} \right\rfloor \leq \sigma_k = \left\lfloor \frac{12r + k - 1}{k} \right\rfloor = \left\lceil \frac{12r}{k} \right\rceil \leq \left\lceil \frac{12r}{6} \right\rceil = 2r.$$

Thus, $\sigma_t \leq \sigma_k \leq 2r$ for $1 \leq t \leq k$. We partition the vertex set $V(F_{n'+1} \square K_{r,r,r})$ into $U_1, U_2, \ldots, U_k$ according to the sorting in the setting where

$$|U_i| = \left\lfloor \frac{12r + i - 1}{k} \right\rfloor \text{ where } 1 \leq i \leq k.$$

Since $|U_i| = \sigma_i \leq 2r$ for each $i \in \{1, 2, \ldots, k\}$, it follows that each U_i is an independent set from the observation above. (See Figure 15.)

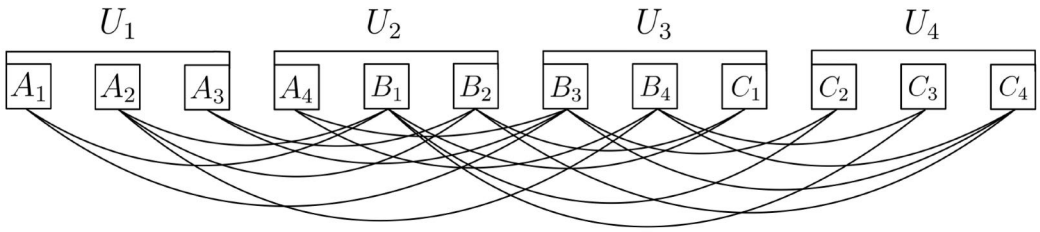


Figure 16: Theorem 3.5 when $n' = 3$ and $k = 4$

Case 2: $k = 4$. It follows that $\dfrac{12r}{4} = 3r$. (See Figure 16.)

We partition the vertex set $V(F_{n'+1}\square K_{r,r,r})$ into U_1, U_2, U_3, and U_4 as follows: $U_1 = A_1 \cup A_2 \cup A_3, U_2 = A_4 \cup B_1 \cup B_2, U_3 = B_3 \cup B_4 \cup C_1$, and $U_4 = C_2 \cup C_3 \cup C_4$. If follows that $|U_1| = |U_2| = |U_3| = |U_4| = 3r$. Since $U_1 \subseteq V_1$ and $U_4 \subseteq V_3$, it follows that U_1 and U_4 are independent sets. Since vertices in A_4, B_1, and B_2 are not adjacent, it follows that U_2 is an independent set. Since vertices in B_3, B_4, and C_1 are not adjacent, it follows that U_3 is an independent set.

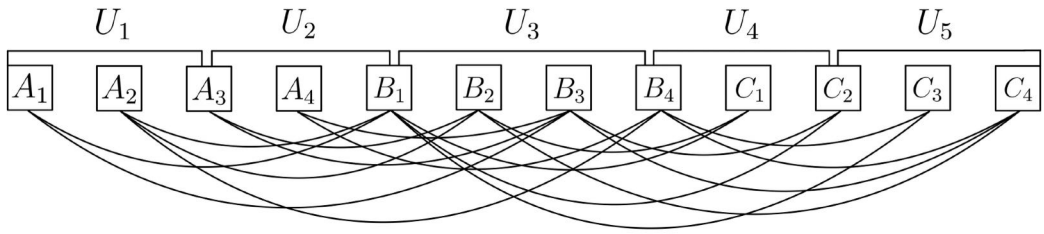


Figure 17: Theorem 3.5 when $n' = 3$ and $k = 5$

Case 3: $k = 5$. (See Figure 17.) We partition the vertex set $V(F_{n'+1}\square K_{r,r,r})$ into U_1, U_2, U_3, U_4, and U_5 according to the sorting in the setting and

$$|U_i| = \left\lfloor \frac{12r + i - 1}{5} \right\rfloor \text{ for each } i \in \{1, 2, 3, 4, 5\}.$$

If follows that $U_1 \subseteq A_1 \cup A_2 \cup A_3, U_2 \subseteq A_3 \cup A_4 \cup B_1, U_3 \subseteq B_1 \cup B_2 \cup B_3 \cup B_4, U_4 \subseteq B_4 \cup C_1 \cup C_2$, and $U_5 \subseteq C_2 \cup C_3 \cup C_4$. Since $U_1 \subseteq V_1, U_3 \subseteq V_2$, and $U_5 \subseteq V_3$, it follows that U_1, U_3, and U_5 are independent sets. Since vertices in A_3, A_4, and B_1 are not adjacent, it follows that U_2 is an independent set. Since vertices in B_4, C_1, and C_2 are not adjacent, it follows that U_4 is an independent set.

Therefore, according to the vertex sorting above, the vertex set of $F_{n'+1}\square K_{r,r,r}$ can be partitioned into k independent sets in a way that the numbers of vertices in any two sets differ by at most 1. Hence $F_{n'+1}\square K_{r,r,r}$ is equitably k-colorable when $k \geq 3$. □

4 Discussion

In summary, we obtain the bounds or exact values of the equitable chromatic numbers and the equitable chromatic thresholds of the Cartesian products of Fans and complete tripartite graphs as in Table 1:

Conditions	$\chi_{=}(F_{n'+1}\Box K_{r,s,t})$	$\chi^{*}_{=}(F_{n'+1}\Box K_{r,r,r})$
$n' = 1, 3, 4$ and $r \leq s \leq t \leq r+1$	3	3
$n' = 2$ and $r \leq s \leq t$	3	≤ 6

Table 1: The bounds or exact values of $\chi_{=}(F_{n'+1}\Box K_{r,s,t})$ and $\chi^{*}_{=}(F_{n'+1}\Box K_{r,r,r})$

Using the same idea as in the proof of Theorem 3.4 with case analysis, we can show that $F_3\Box K_{r,r,r}$ is equitably 5-colorable. The result and the proof will appear in the master thesis of the first author. It is still open whether the graph is equitably 4-colorable or not.

One may generalize the concept of sorting in the proof to show that $F_{n'+1}\Box K_{r,r,r}$ is equitably k-colorable when $k = 3$ or $k \geq 6$. This result and the proof will also appear in the thesis. It is still open whether the graph is equitably k-colorable when $k = 4$ or 5.

For future research, it would be interesting to study the equitable colorings in $F_{n+1}\Box K_{r,s,t}$ where $1 \leq r \leq s \leq t$ and $t - r \geq 2$.

Note that $K_{m,n}$ when $m = 2n$ is a spanning subgraph of $K_{n,n,n}$. Thus, the results in Theorem 2.8 (1) and Theorem 2.10 (1) when $m = 2n$ are generalized by our results.

Acknowledgment. The first author would like to thank the Development and Promotion of Science and Technology Talents Project (DPST) for support.

References

[1] G. Chartrand, L. Lesniak, P. Zhang, Graphs and Digraphs, 6th ed., Taylor & Francis Group LLC, New York, 2015.

[2] B.-L. Chen, K.-W. Lih, J.-H. Yan, Equitable coloring of interval graphs and products of graphs, `arXiv:0903.1396v1`.

[3] H. Furmańczyk, Equitable coloring of graph products, Opuscula Math. 26 (1) (2006) 31–44.

[4] W.-H. Lin, G. J. Chang, Equitable colorings of Cartesian products of graphs, Discrete Appl. Math. 160 (2012) 239–247.

[5] Z. Yan, W.-H. Lin, W. Wang, The equitable chromatic threshold of the Cartesian product of bipartite graphs is at most 4, Taiwanese J. Math. 18 (3) (2014) 773–780.

[6] L. Zuo, F. Wu, S. Zhang, Equitable colorings of Cartesian products graphs of wheels with complete bipartite graphs, WSEAS Trans. Math. 13 (2014) 236–245.

[7] L. Zuo, F. Wu, S. Zhang, Equitable colorings of Cartesian products of fans with bipartite graphs, Ital. J. Pure Appl. Math. 35 (2015) 91–100.

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

Simple Necessary Conditions for Achieving Byzantine Agreement over (2,3)-Uniform Hypergraphs

Supphachai Sonwai[1,*] and Wutichai Chongchitmate[1,†]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

Byzantine agreement problem is one of the most fundamental problems in the field of distributed computing. In this problem, a group of parties attempts to reach an agreement after exchanging messages in the presence of faults. More specifically, the Byzantine agreement problem deals with a synchronous network of n parties, in which at most t parties are faults or corrupt parties. Those who are not corrupt parties are called honest parties. At the beginning, each party starts with an initial binary value. In each round, each party sends messages to others by an authenticated communication channel, while corrupt parties can send arbitrary messages. The primary goal of this problem is to achieve Byzantine agreement, which means that after exchanging messages, all honest parties must decide on a single value with certain conditions. The communication among every party can be modeled as a graph, where each party is represented by a vertex, while a communication channel is represented by an edge. In our communication model, there may exist a broadcast channel for a group of three parties. The resulting communication can be modeled as a (2,3)-uniform hypergraph, i.e., each hyperedge has size two or three. In this work, we focus on simplifying the conditions for achieving Byzantine agreement over (2,3)-uniform hypergraphs from Ravikant et al. (2004) and Khan and Vaidya (2021), which are difficult to verify whether a (2,3)-uniform hypergraph satisfies the conditions. We provide the simple necessary conditions for achieving Byzantine agreement over (2,3)-uniform hypergraphs.

Keywords: Byzantine agreement, Byzantine fault tolerance, consensus, (2,3)-uniform hypergraph

2020 MSC: 68M15, 05C65, 68M14

1 Introduction

In a distributed system, a basic problem arises when multiple parties communicate with each other. In some situations, it is important that all parties agree on the same thing, or decide to do the same thing, for example, voting. In practice, the parties mentioned above are not necessarily a real human. In many situations, parties can be referred to processors in a computer system.

*Speaker (sonwai.ss@outlook.co.th)
†Corresponding author (wutichai.ch@chula.ac.th)

One example is the flight control system in aircrafts as well as spacecrafts, where data are needed to be correct in order to operate flight safely. Another example is the transaction in a blockchain of decentralized cryptocurrency.

The byzantine agreement problem is one of the most fundamental problems in the field of distributed computing. Byzantine agreement is also referred to by some authors as Byzantine consensus and Byzantine fault tolerance. The problem simply attempts to reach an agreement among a group of parties with the presence of faults, sometimes called corrupt parties, who would like to disrupt the information for each party. By reaching an agreement, we mean that every non-corrupt party must decide on the same value. The problem has a wide range of applications including cryptocurrency. Byzantine fault tolerance (BFT) is extremely important for the blockchain technology. It plays a vital role in enabling a network to perform computations, even though some parties may behave maliciously. BFT ensures that every party in the network arrive at the same decision.

The Byzantine generals problem [1] was first introduced and formalized by Lamport et al. in 1980s. The term "Byzantine" in the problem basically comes from the historical context of the Byzantine Empire. Let us imagine that we are in a battle in the medieval warfare. Many divisions of the Byzantine army are camped outside the wall of its enemy city. Each division is commanded by its own general. It is usual to assume that all soldiers in a division always obey the orders from its general. All the generals can communicate with one another only by sending messages via a runner or a messenger. Communication proceeds in rounds. In each round, each general can send messages, while also receiving messages from other generals. After exchanging messages, the generals must decide on a plan of action. However, some of them may be traitors, so they are trying to prevent the loyal generals from reaching an agreement. The goal of this problem is that all loyal generals achieve an agreement on a single decision.

The Byzantine agreement problem is quite similar to the Byzantine generals problem. More specifically, the Byzantine agreement deals with a synchronous network of n parties, where at most t (constant) parties are faults or corrupt parties. The terms "faults" and "corrupt parties" may be used interchangably. Those who are not corrupt are said to be honest. We model adversary as a fictitious outside party who chooses parties to be corrupted or controlled. Alternatively, it can be assumed that all corrupt parties cooperate with each other. At the beginning, each party holds its own binary input. In each round, each party sends messages to others via an authenticated communication channel, while corrupt parties can exhibit arbitrary behaviours, i.e., send arbitrary messages. Every sent message is received correctly and each receiver knows the message sender. We say that Byzantine agreement (tolerating at most t faults/corrupt parties) is achievable provided that after a number of rounds of communication, all honest parties must decide on a single value at the end of the protocol that satisfies the following conditions.

1. Agreement: All honest parties must decide on the same value.
2. Validity: If all honest parties start with the same value $b \in \{0, 1\}$, then all honest parties must decide on b.
3. Termination: All honest parties eventually decide.

The communication can be modeled as a graph, where parties are represented by vertices, while each edge represents a communication channel between two parties. There have been

several studies on Byzantine agreement over various models of communication. Pease et al. [2] studied the condition for achieving Byzantine agreement over complete communication model, i.e., complete graph, and proved that Byzantine agreement is achievable if and only if the number of corrupt parties is less than one-third of the total number of all parties.

Then Dolev [3] expanded the scope of the problem to a case of incomplete graph and proved that Byzantine agreement is achievable over any communication graph if and only if the number of corrupt parties is less than the total number of all parties and also less than one-half of the (vertex) connectivity.

Fitzi and Maurer [4] generalized the problem by allowing a broadcast channel among every three parties, while all channels for every pair of parties also exist. The result improved the upper bound of the number of corrupt parties. It was shown that if there exists a local broadcast channel among every 3 parties, then Byzantine agreement is achievable over complete communication graph if and only if the number of corrupt parties is less than one-half of the total number of all parties. The use of 3-party broadcast channels makes the communication graph become a hypergraph.

Ravikant et al. [5] further generalized the problem by trying to remove some broadcast channels, which implies that some groups of 3 parties may not have a broadcast channel. In this model, some 3-hyperedges can be missing. Khan and Vaidya [6] verified and modified the result in [5]. However, it is not easy to verify whether a hypergraph satisfies the conditions in [5,6]. For this reason, we aim to simplify those conditions for achieving Byzantine agreement over certain classes of hypergraphs, called $(2,3)$-uniform hypergraphs.

In Section 2, we shall give definitions relating to graphs and hypergraphs as well as some important results used in this work. The communication model are also provided in the next section. More details about the conditions for achieving Byzantine agreement for different models of communication are included in Section 3, and Section 4 contains the main results of this work.

2 Preliminaries

In this section, we give definitions required for understanding the results in section 4.

2.1 Notations

We use the following notations throughout this work.

- Let P denote the set of all parties in the network.
- Let E denote the set of communication channels for parties in P.

2.2 Graphs and Hypergraphs

As we model the communication among all parties by using graph, we first recall basic concepts of graph theory. [7]

Definition 2.1. A (simple, undirected) *graph* G is an unordered pair of disjoint sets (P, E), where P is a non-empty finite set and E is a subset of the set $P^{(2)}$ of unordered pairs of P. The set $P = P(G)$ is called the *vertex set*, and the set $E = E(G)$ is called the *edge set*. Elements of P are called *vertices*, and elements of E are called *edges*. The number of edges of G is denoted by $e(G)$.

Remark 2.2. In most graph theory textbooks, the vertex set is usually denoted by V. However, in this context, vertices are used to represent parties in the network, so the set of vertices will be denoted by P to avoid confusion.

Definition 2.3. Let $G(P, E)$ be a graph. An edge $\{x, y\} \in E$ (may write briefly as xy) is said to *join* the vertices x and y. In this case, x and y are said to be *adjacent*, and we also say that x and y are *neighbours* of each other and the *endvertices* of the edge $\{x, y\}$. We simply write $G(P, E)$ for the graph G with the vertex set P and the edge set E.

Definition 2.4. A *complete graph* K_n is a graph of n vertices in which every pair of distinct vertices is connected by a unique edge, namely, K_n has the vertex set P with $|P| = n$, and has the edge set $P^{(2)}$.

Definition 2.5. We say that $F(P', E')$ is a *subgraph* of a graph $G(P, E)$ if $P' \subset P$ and $E' \subset E$.

Definition 2.6. Let x be a vertex in a graph $G(P, E)$. The *neighbourhood* of x, denoted by $N(x)$, is defined to be the set of all neighbours of x, i.e., $N(x) := \{y \in P : xy \in E\}$. The *degree* of x, denoted by $d(x)$, is defined to be the number of its neighbours, namely, $d(x) := |N(x)|$. If $d(x) = 0$, x is said to be an *isolated* vertex. Furthermore, the *maximum degree* of G, denoted by $\Delta(G)$, is defined by $\Delta(G) := \max\{d(x) : x \in P\}$. Similarly, the *minimum degree* of G, denoted by $\delta(G)$, is defined by $\delta(G) := \min\{d(x) : x \in P\}$.

The following proposition is a basic properties of graph that provides a relationship between the degrees of vertices and the number of edges in a graph.

Proposition 2.7 (Handshaking Lemma)**.** *Let $G(P, E)$ be a graph. Then*

$$\sum_{x \in P} d(x) = 2e(G). \tag{2.1}$$

Proof. We prove by double counting. We count the pair (x, e), where x is a vertex that belongs to an edge $e \in E$. There are 2 endvertices for each edge $e \in E$, so we have $2e(G)$. For each vertex $x \in G$, there are $d(x)$ edges containing x. All add up to $\sum_{x \in P} d(x)$. □

Definition 2.8. Let x, y be vertices in a graph $G(P, E)$. An *x-y path* in G is a sequence $x = x_1, x_2, ..., x_k, x_{k+1} = y$ of distinct vertices of G, where $x_i x_{i+1} \in E$ for all $1 \leq i \leq k$. Then we say that G is *connected* provided that there is an x-y path in G for all $x, y \in P$.

Definition 2.9. Let $G(P, E)$ be a graph. Suppose that $Q \subset P$. The *subgraph induced (or spanned) by Q*, denoted by $G[Q]$, is a graph whose the vertex set is Q and the edge set is $E \cap Q^{(2)}$. The graph $G - Q$ is defined by $G[P \setminus Q]$.

Definition 2.10. Let $G(P, E)$ be a connected graph with at least two vertices. We define the *(vertex) connectivity* of G, denoted by $\kappa(G)$, by the smallest $|S|$ such that $S \subset P$ and $G - S$ is disconnected (or a single vertex for the case of complete graph). We say that G is k-connected if $\kappa(G) \geq k$.

Next, we give elementary concepts of hypergraphs [5, 8] required for understanding the communication model in this work.

Definition 2.11. Let P be a finite set. An (undirected) *hypergraph* H is an unordered pair of disjoint sets (P, E), where E is a subset of the power set 2^P of all possible (non-empty) subsets of P. The set $P = P(H)$ is called the *vertex set*, and the set $E = E(H)$ is called the *hyperedge set*. Elements of P are called *vertices*, and elements of E are called *hyperedges*. We say that a hyperedge $e \in E$ is a k-hyperedge if $|e| = k$. The number of hyperedges of H is denoted by $e(H)$. If $\{x, y\} \subset e$ for some $e \in E$, x and y are said to be *neighbours* of each other. We simply write $H(P, E)$ for the graph H with the vertex set P and the edge set E.

Definition 2.12. Let x be a vertex in a hypergraph $H(P, E)$. The *degree* of x, denoted by $d(x)$, is defined to be the number of hyperedges which contains x, i.e., $d(x) := |\{e \in E : x \in e\}|$. If $d(x) = 0$, x is said to be an *isolated* vertex.

Definition 2.13. A hypergraph $H(P, E)$ is said to be a *k-uniform* hypergraph if all hyperedges of H are k-hyperedges. A hypergraph $H(P, E)$ is said to be a *$(k_1, k_2, ..., k_l)$-uniform* hypergraph if each hyperedge of H is k_i-hyperedge for some $i \in \{1, 2, ..., l\}$.

Definition 2.14. For a hypergraph $H(P, E)$, the *2-section* of H, denoted by $[H]_2$, is defined by a graph $[H]_2(P, E')$, where $E' := \{\{x, y\} : x \neq y \wedge \exists e \in E, \{x, y\} \subseteq e\}$.

2.3 Communication Model

We recall that the Byzantine agreement problem is a problem concerning communication in a network. Analogously to the original version of Byzantine agreement problem [1,2] mentioned in the previous section, we assume from now on that the network consists of n parties. There can be at most t faults/corrupt parties among these parties, where t is constant. The communication model of the network are formally described as follows. Communication proceeds in round. In each round of communication, a party sends and receives messages by a communication channel between two parties. Apart from these channels, there may be a local broadcast channel for a group of three parties. The local broadcast channel allows a sender to transmit messages to all receivers within a group of three parties. Throughout this work, we model the communication by a $(2, 3)$-uniform hypergraph $H(P, E)$, where each vertex in the vertex set P represents each party in the network, an edge (2-hyperedge) xy represents a communication channel between x and y, and a 3-hyperedge $\{x, y, z\}$ (or sinply written as xyz) represents a local broadcast channel for x, y and z.

3 Byzantine Agreement

In this section, we review the previous results of Byzantine agreement over some models of communication. The study of Byzantine agreement began in the late 1970s and early 1980s. We start with the foundational work on Byzantine agreement problem.

Theorem 3.1. *[1, 2] Byzantine agreement is achievable over complete graph if and only if $n > 3t$.*

This result gives the upper bound of the number of corrupt parties. Dolev [3] generalized the problem to a case of incomplete graph.

Theorem 3.2. *[3] Byzantine agreement is achievable over arbitrary graph $G(P, E)$ if and only if $n > 3t$ and $\kappa(G) > 2t$.*

Fitzi and Maurer [4] attempted to improve the upper bound of the number of corrupt parties by allowing a broadcast channel among every three parties, while the communication graph is complete. The concept of 3-party broadcast channels leads to the use of hypergraphs to model the communication. The communication are now modeled by the complete $(2,3)$-uniform hypergraph.

Theorem 3.3. *[4] Suppose that there exists a local broadcast channel among every 3 parties. Then Byzantine agreement is achievable over complete graph if and only if $n > 2t$.*

Now, the number of corrupt parties can be raised up to half of the total number of all parties. Ravikant et al. [5] further generalized the problem to a case of incomplete $(2,3)$-uniform hypergraph. Khan and Vaidya [6] verified and modified the result in [5].

Theorem 3.4. *[5, 6] Byzantine agreement is achievable over a hypergraph $(2,3)$-uniform hypergraph $H(P,E)$ if and only if all the following hold:*

1. *$n > 2t$*
2. *$[H]_2$ is either complete or $(2t+1)$-connected.*
3. *For every $P_1, P_2, P_3 \subset P$ such that $P_1 \cup P_2 \cup P_3 = P$ and $|P_1| = |P_2| = |P_3| = t$, there exists a 3-hyperedge across $P_1 \setminus (P_2 \cup P_3)$, $P_2 \setminus (P_1 \cup P_3)$ and $P_3 \setminus (P_1 \cup P_2)$.*

However, it is difficult to verify whether a $(2,3)$-uniform hypergraph $H(P,E)$ satisfies the last condition in Theorem 3.4. We call it Condition A. In the next section, we mainly focus on simplifying this condition.

Remark 3.5. We observe that if $t < \frac{n}{3}$, then for any $P_1, P_2, P_3 \subset P$ with size t each,

$$|P_1 \cup P_2 \cup P_3| \leq |P_1| + |P_2| + |P_3| = 3t < n,$$

so in the case when $n > 3t$, such three subsets P_1, P_2, P_3 whose union is P cannot exist.

4 Main Results

In this section, we provide simple necessary conditions for achieving Byzantine agreement over $(2,3)$-uniform hypergraphs. Before we proceed to the results, we give a few more notations. For any $(2,3)$-uniform hypergraph $H(P,E)$, we write $E = E_2 \cup E_3$, where E_2 denotes the set of all 2-hyperedges of H, and E_3 denotes the set of all 3-hyperedges of H. Denote $e_2(H) := |E_2|$ and denote $e_3(H) := |E_3|$. We also define H_3 to be the hypergraph with the vertex set P and the hyperedge set E_3.

We start this section with the generalization of the handshaking lemma to a 3-uniform hypergraph.

Proposition 4.1. *Let $H(P,E)$ be a 3-uniform hypergraph. Then*

$$\sum_{x \in P} d(x) = 3e_3(H). \tag{4.1}$$

Proof. Similar to Proposition 2.7. □

By defining the 3-uniform hypergraph H_3 above, in order to verify whether a $(2,3)$-uniform hypergraph $H(P,E)$ satisfies Condition A, this allows us to consider a hypergraph containing only 3-hyperedges instead of $H(P,E)$.

Remark 4.2. Let $H(P,E)$ be a $(2,3)$-uniform hypergraph. $H(P,E)$ satisfies Condition A if and only if $H_3(P,E_3)$ satisfies Condition A.

The following result is the key property to obtain all subsequent results as from now including the main results in this section.

Lemma 4.3. *Let P be a set with $|P| = n$, where $2t < n \leq 3t$. Suppose that $P_1, P_2, P_3 \subset P$, where $P_1 \cup P_2 \cup P_3 = P$ and $|P_1| = |P_2| = |P_3| = t$. Then P_1, P_2 and P_3 have at most $3t - n$ common vertices.*

Proof. We observe that all common vertices of P_1, P_2 and P_3 are in $(P_1 \cap P_2) \setminus P_3$, $(P_1 \cap P_3) \setminus P_2$, $(P_2 \cap P_3) \setminus P_1$ and $P_1 \cap P_2 \cap P_3$, which are disjoint sets. We obtain that

$$\begin{aligned}
&|(P_1 \cap P_2) \setminus P_3| + |(P_1 \cap P_3) \setminus P_2| + |(P_2 \cap P_3) \setminus P_1| + |P_1 \cap P_2 \cap P_3| \\
&\quad = |P_1 \cap P_2| - |P_1 \cap P_2 \cap P_3| + |P_1 \cap P_3| - |P_1 \cap P_2 \cap P_3| + |P_2 \cap P_3| \\
&\qquad - |P_1 \cap P_2 \cap P_3| + |P_1 \cap P_2 \cap P_3| \\
&\quad = |P_1 \cap P_2| + |P_1 \cap P_3| + |P_2 \cap P_3| - 2|P_1 \cap P_2 \cap P_3| \\
&\quad = |P_1| + |P_2| + |P_3| - |P_1 \cup P_2 \cup P_3| + |P_1 \cap P_2 \cap P_3| - 2|P_1 \cap P_2 \cap P_3| \\
&\quad = 3t - n - |P_1 \cap P_2 \cap P_3| \\
&\quad \leq 3t - n.
\end{aligned}$$

Hence, P_1, P_2 and P_3 have at most $3t - n$ common vertices. □

The above lemma leads to the following observation, which guides the proof of the main results afterward. To approach the desired simple necessary condition, we do not prove implication directly, but we attempt to construct a $(2,3)$-uniform hypergraph that does not satisfy Condition A.

Lemma 4.4. *Let $H(P,E)$ be a $(2,3)$-uniform hypergraph, where $2t < n \leq 3t$. If $H_3(P,E_3)$ has at least $n - 2t$ isolated vertices, then $H(P,E)$ does not satisfy Condition A.*

Proof. Suppose that $H_3(P,E_3)$ has at least $n - 2t$ isolated vertices. Then by Lemma 4.3, for any three subsets $P_1, P_2, P_3 \subset P$ such that $P_1 \cup P_2 \cup P_3 = P$ and $|P_1| = |P_2| = |P_3| = t$, there are at most $3t - n$ common vertices. Form P_1, P_2 and P_3 as follows. Take exactly $3t - n$ vertices and place them in both P_1 and P_2. Since $3t - n < t$, there are other $n - 2t$ places available for vertices in P_1 and P_2. Put all $n - 2t$ isolated vertices of $H_3(P,E_3)$ in P_1. Each of $n - (3t - n) - (n - 2t) = n - t$ remaining vertices are put in either P_2 or P_3. Notice that all $3t - n$ common vertices in both P_1 and P_2. From the above construction, we obtain that $P_1 \cup P_2 \cup P_3 = P$ and $|P_1| = |P_2| = |P_3| = t$. Let $x \in P_1$. Then x is either a common vertex or an isolated vertex in $H_3(P,E_3)$. If x is a common vertex, then x cannot belong to any of $P_1 \setminus (P_2 \cup P_3)$, $P_2 \setminus (P_1 \cup P_3)$ and $P_3 \setminus (P_1 \cup P_2)$. Note that x belongs to no 3-hyperedges if it is an isolated vertex. Hence, $H(P,E)$ does not satisfy Condition A. □

Corollary 4.5. *Let $H(P,E)$ be a $(2,3)$-uniform hypergraph, where $2t < n \leq 3t$. Suppose that $H(P,E)$ satisfies Condition A. Then $H_3(P,E_3)$ has less than $n - 2t$ isolated vertices.*

Proof. This corollary follows from the contraposition of Lemma 4.4. □

Lemma 4.6. *Let $H(P, E)$ be a $(2, 3)$-uniform hypergraph with $2t + 1$ vertices. If there are vertices $x, y \in P$ such that $\{x, y\}$ is contained in less than t 3-hyperedges, then $H(P, E)$ does not satisfy Condition A.*

Proof. Let $x, y \in P$ be vertices such that $xyz_1, xyz_2, ..., xyz_s$ are the only 3-hyperedges containing both x and y, where $s < t$. Form P_1, P_2, P_3 as follows. We put x in P_1 and put y in P_2. Place $z_1, z_2, ..., z_s$ in both P_1 and P_2. Take exactly $t - s - 1$ vertices from $P \setminus \{x, y, z_1, z_2, ..., z_s\}$ and set them in both P_1 and P_2. The remaining vertices $P \setminus (P_1 \cup P_2)$ are put in P_3. We can see that $P_1 \cup P_2 \cup P_3 = P$ and $|P_1| = |P_2| = |P_3| = t$. It follows that the only vertex in $P_1 \setminus (P_2 \cup P_3)$ is x. Similarly, $P_2 \setminus (P_1 \cup P_3) = \{y\}$. However, we know that all 3-hyperedges containing both x and y do not contain any vertices in P_3 nor in $P_3 \setminus (P_1 \cup P_2)$. Thus, there cannot be a 3-hyperedge that passes through all $P_1 \setminus (P_2 \cup P_3)$, $P_2 \setminus (P_1 \cup P_3)$ and $P_3 \setminus (P_1 \cup P_2)$, i.e., $H(P, E)$ does not satisfy Condition A. □

Corollary 4.7. *Suppose that a $(2, 3)$-uniform hypergraph $H(P, E)$ consisted of $2t+1$ vertices satisfies Condition A. Then any pair of distinct vertices of P must be part of at least t 3-hyperedges.*

Proof. This corollary follows from the contraposition of Lemma 4.6. □

Remark 4.8. The converse of Corollary 4.7 does not hold. Consider the following example. Let $H(P, E)$ be a $(2, 3)$-uniform hypergraph of 7 vertices, where $P = \{1, 2, 3, 4, 5, 6, 7\}$ and the 3-hyperedge set E_3 contains only the following 3-hyperedges.

124	125	126	127	134	135	136	137
146	147	156	157	234	235	236	237
245	267	345	367	456	457	467	567

Now, let $t = 3$. Then $n = 2t + 1$. It is not hard to see that each of all possible $\binom{7}{2}$ pairs of elements of P is contained in at least t 3-hyperedges. We construct P_1, P_2 and P_3 as follows.

$$\begin{aligned} P_1 &= \{1, 2, 3\}, \\ P_2 &= \{1, 4, 5\}, \\ P_3 &= \{1, 6, 7\}. \end{aligned}$$

Clearly from the construction above, $P_1 \cup P_2 \cup P_3 = P$ and $|P_1| = |P_2| = |P_3| = t$. Also,

$$\begin{aligned} P_1 \setminus (P_2 \cup P_3) &= \{2, 3\}, \\ P_2 \setminus (P_1 \cup P_3) &= \{4, 5\}, \\ P_3 \setminus (P_1 \cup P_2) &= \{6, 7\}. \end{aligned}$$

However, none of above 3-hyperedges passes through all $P_1 \setminus (P_2 \cup P_3)$, $P_2 \setminus (P_1 \cup P_3)$ and $P_3 \setminus (P_1 \cup P_2)$. Thus, the hypergraph $H(P, E)$ defined above cannot satisfy Condition A.

Corollary 4.7 also gives a lower bound of the number of 3-hyperedges when $n = 2t + 1$, i.e., if a $(2, 3)$-uniform hypergraph $H(P, E)$ satisfies Condition A, then by Proposition 4.1, we have that

$$e_3(H) \geq \frac{1}{3}\binom{2t+1}{2}t = \frac{(2t+1)(2t)t}{6} = \frac{2t^3 + t^2}{3}. \tag{4.2}$$

The inequality on the number of 3-hyperedges in the hypergraph similar to Equation 4.2 can be a good candidate for a simpler necessary condition to verify than Condition A in [5, 6].

Question 4.9. In a similar way to Lemma 4.6 and Corollary 4.7, can we find such conditions when $2t+1 < n \leq 3t$?

The following theorem will answer the above question and lay a path directly to our main theorem.

Lemma 4.10. *Let $H(P, E)$ be a $(2, 3)$-uniform hypergraph, where $2t < n \leq 3t$. For any $S \subset P$, define*

$$F_S := \{e \in E_3(H) : \exists x, y \in S, x \neq y \wedge \{x, y\} \subset e\}.$$

If there exists $S \subset P$ with $|S| = 2(n-2t)$ such that $|F_S| < 3t - n + 1$, then $H(P, E)$ does not satisfy Condition A.

Proof. Suppose that there exists $S \subset P$ such that $|S| = 2(n-2t)$ and $|F_S| < n - 3t + 1$. Define

$$U_S := \bigcup_{e \in F_S} e \setminus S = \bigcup_{e \in F_S} (e \setminus S) = \bigcup_{\substack{\{x,y,z\} \in F_S, \\ x,y \in S}} \{z\}$$

We can see that $|U_S| < 3t - n + 1$. Indeed,

$$|U_S| \leq |F_S| < 3t - n + 1.$$

Assume that $S = \{x_1, x_2, ..., x_{n-2t}, y_1, y_2, ..., y_{n-2t}\}$ and that $U_S = \{z_1, z_2, ..., z_k\}$, where $k < 3t - n + 1$. Form P_1, P_2, P_3 as follows. Place all $z_1, z_2, ..., z_k$ in both P_1 and P_2. Next, put $x_1, x_2, ..., x_{n-2t}$ in only P_1, and put $y_1, y_2, ..., y_{n-2t}$ in only P_2, which is possible because $|P_1| = |P_2| = n-2t+k < n-2t+3t-n+1 < t+1$. If $|P_1|$ and $|P_2|$ are not equal to t, further add exactly $t-(n-2t)-k = 3t-n-k$ vertices from $P \setminus (S \cup U_S)$ to both P_1 and P_2 so that $|P_1| = |P_2| = t$. Lastly, take $P_3 = P \setminus (P_1 \cup P_2)$. Since there are all $2(n-2t)+k+3t-n-k = n-t$ distinct vertices together in P_1 and P_2, we obtain that $|P_3| = |P| - |P_1 \cup P_2| = t$. Now, $|P_1| = |P_2| = |P_3| = t$. It is not difficult to see that $P_1 \cup P_2 \cup P_3 = P$. We also notice that

$$\begin{aligned} P_1 \setminus (P_2 \cup P_3) &= \{x_1, x_2, ..., x_{n-2t}\}, \\ P_2 \setminus (P_1 \cup P_3) &= \{y_1, y_2, ..., y_{n-2t}\}. \end{aligned}$$

However, all 3-hyperedges containing any pair from $\{x_1, x_2, ..., x_{n-2t}, y_1, y_2, ..., y_{n-2t}\}$ have all vertices in only P_1 and P_2. This completes the proof. □

Corollary 4.11. *Let $H(P, E)$ be a $(2, 3)$-uniform hypergraph, where $2t < n \leq 3t$. Suppose that $H(P, E)$ satisfies Condition A. For any $S \subset P$, define*

$$F_S := \{e \in E_3(H) : \exists x, y \in S, x \neq y \wedge \{x, y\} \subset e\}.$$

Then for any $S \subset P$ with $|S| = 2(n-2t)$, we have that

$$|F_S| \geq 3t - n + 1,$$

i.e., the total number of 3-hyperedges that contains at least one pair of elements of S must be at least $3t - n + 1$.

Proof. This corollary follows from the contraposition of Lemma 4.10. □

Corollary 4.12. *Let $H(P,E)$ be a $(2,3)$-uniform hypergraph, where $2t < n \leq 3t$. Suppose that $H(P,E)$ satisfies Condition A. Then*

$$e_3(H) \geq \frac{\binom{n}{2(n-2t)}(3t-n+1)}{\binom{n}{2(n-2t)} - \binom{n-3}{2(n-2t)} - 3\binom{n-3}{2(n-2t)-1}}.$$

.

Proof. By Corollary 4.11, for any $S \subset P$ with $|S| = 2(n-2t)$, we have $|F_S| \geq 3t-n+1$, which implies that

$$\sum_{\substack{S\subset P,\\ |S|=2(n-2t)}} |F_S| \geq \binom{n}{2(n-2t)}(3t-n+1).$$

We count the pair (S,e), where $S \subset P$, $|S| = 2(n-2t)$ and e is a 3-hyperedge that contains a pair of elements of S. First, for any $S \subset P$ with $|S| = 2(n-2t)$, there exist $|F_S|$ 3-hyperedges, so all add up to

$$\sum_{\substack{S\subset P,\\ |S|=2(n-2t)}} |F_S|.$$

Now, we let $e := \{x,y,z\} \in E_3(H)$. Then the set of all possible pairs of elements in e is $e^{(2)} = \{\{x,y\},\{y,z\},\{x,z\}\}$. Instead of counting the number of S with $|S| = 2(n-2t)$ that has pairs in e directly, we consider its complement. We observe that e is not counted for S if neither $\{x,y\}$, nor $\{y,z\}$, nor $\{x,z\}$ is contained in S, that is, S contains at most 1 element from e. The number of $(2(n-2t))$-subsets that contains no elements from e is equal to $\binom{n-3}{2(n-2t)}$. The number of $(2(n-2t))$-subsets that contains exactly one element from e is equal to $3\binom{n-3}{2(n-2t)-1}$. It follows that the total number of $(2(n-2t))$-subsets that contain at least 1 element from e is equal to $\binom{n}{2(n-2t)} - \binom{n-3}{2(n-2t)} - 3\binom{n-3}{2(n-2t)-1}$. Summing over $e \in E_3$ yields

$$\sum_{e\in E_3}\left\{\binom{n}{2(n-2t)} - \binom{n-3}{2(n-2t)} - 3\binom{n-3}{2(n-2t)-1}\right\} = \sum_{\substack{S\subset P,\\ |S|=2(n-2t)}} |F_S|.$$

Since there are all e_3 3-hyperedges, we obtain that

$$e_3\left(\binom{n}{2(n-2t)} - \binom{n-3}{2(n-2t)} - 3\binom{n-3}{2(n-2t)-1}\right) = \sum_{\substack{S\subset P,\\ |S|=2(n-2t)}} |F_S|.$$

Therefore,

$$e_3 \geq \frac{\binom{n}{2(n-2t)}(3t-n+1)}{\binom{n}{2(n-2t)} - \binom{n-3}{2(n-2t)} - 3\binom{n-3}{2(n-2t)-1}}$$

as desired. □

Combining all results together gives the following theorem, which provides a simple necessary condition for achieving Byzantine agreement over $(2,3)$-uniform hypergraphs.

Theorem 4.13. *Suppose that Byzantine agreement is achievable over a $(2,3)$-uniform hypergraph $H(P,E)$, where $H_3(P,E_3)$ has s isolated vertices. Then all the following hold:*

1. $n > 2t$
2. $[H]_2$ *is either complete or* $(2t+1)$*-connected.*
3. $s < n - 2t$ *when* $2t < n \leq 3t$.
4. $e_3(H) \geq \frac{\binom{n}{2(n-2t)}(3t-n+1)}{\binom{n}{2(n-2t)}-\binom{n-3}{2(n-2t)}-3\binom{n-3}{2(n-2t)-1}}$.

Proof. This follows directly from Theorem 3.4, Corollary 4.5 and Corollary 4.12. □

5 Discussion

In this work, we simplified conditions for achieving Byzantine agreement over $(2,3)$-uniform hypergraphs from [5, 6], and provided a simple necessary condition for achieving Byzantine agreement over $(2,3)$-uniform hypergraphs in Theorem 4.13, which gives the lower bound of the number of 3-hyperedges. Nevertheless, the scope of this work was limited to only $(2,3)$-uniform hypergraphs. In [6], Khan and Vaidya further proved that the conditions in Theorem 3.4 are both necessary and sufficient for arbitrary hypergraphs. Intuitively, one interesting open problem is simplifying those conditions in [5,6] to achieve Byzantine agreement over not only a specific class of hypergraphs but any hypergraphs.

Acknowledgment. The authors would like to thank the Department of Mathematics and Computer Science, Chulalongkorn University, and the Development and Promotion of Science and Technology Talents Project (DPST) scholarship for the financial support.

References

[1] L. Lamport, R. Shostak, M. Pease, The Byzantine Generals Problem, ACM Trans. Program. Lang. Syst. 4(3) (1982) 382–401.

[2] M. Pease, R. Shostak, L. Lamport, Reaching Agreement in the Presence of Faults, J. ACM 27(2) (1980) 228–234.

[3] D. Dolev, The Byzantine Generals Strike Again, Algorithms. 3(1) (1982) 14–30.

[4] M. Fitzi, U. Maurer, From partial consistency to global broadcast, in: F.F. Yao and E.M. Luks (Eds.), Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, STOC '00, Portland, OR, ACM, New York, 2000, pp. 494–503.

[5] D.V.S. Ravikant, V. Muthuramakrishnan, V. Srikanth, K. Srinathan, C.P. Rangan, On Byzantine Agreement over (2,3)-Uniform Hypergraphs, in: R. Guerraoui (Ed.), Proceedings of the Eighteenth International Symposium on Distributed Computing, DISC 2004, Amsterdam, The Netherlands, Springer, Berlin, Heidelberg, Germany, 2004, pp. 450–464.

[6] M.S. Khan, N.H. Vaidya, Byzantine Consensus in Directed Hypergraphs. <`https://arxiv.org/pdf/2109.01205`>, 2021, (accessed 02.09.21).

[7] B. Bollobás, Modern Graph Theory, Springer, New York, 1998.

[8] A. Bretto, Hypergraph Theory, Springer, Cham, 2013.

A Distinct Length Path Decomposition of Graphs

Sittichoke Songsa-ard[1] and Atthakorn Sakda[1,*]

[1]Mathematics Programs, Faculty of Science and Technology,
Suratthani Rajabhat University, Surat Thani 84100, Thailand

Abstract

A *distinct length path decomposition* of a graph G is a partition of $E(G)$ into paths such that every pair of paths has a distinct length. In this paper, we establish that the complete graph K_n and its subgraphs, including $K_n - C_{m+1}$, $K_n - P_{m+1}$, and $K_n - S_m$, admit distinct length path decompositions. Our approach is constructive, as we explicitly construct a set of paths satisfying the decomposition criteria. This method not only demonstrates the existence of such decompositions but also provides an algorithmic framework for their construction.

Keywords: path decomposition, distinct length, decomposition
2020 MSC: 05C38, 05C70

1 Introduction

An $\mathcal{H}$-decomposition of a graph G is a partition of the edge set of G, denoted as $E(G) = E_1 \cup E_2 \cup \cdots \cup E_k$, such that the subgraph of G induced by E_i is isomorphic to a member of $\mathcal{H}$ for every $i \in \{1, 2, \ldots, k\}$. If the members of $\mathcal{H}$ are paths or cycles, the decomposition is called a *path decomposition* or *cycle decomposition*, respectively.

There exists a strong similarity between the techniques for decomposing a graph into paths and those for decomposing it into cycles. In general, research on graph decompositions into paths or cycles can be categorized into two main types. The first class consists of graph decompositions into paths or cycles of equal length, which can be further categorized into two groups. The first group includes decompositions into paths or cycles of uniform length, as studied in [1], [2], [3], [4] and [5]. The second group focuses on minimizing the number of paths or cycles required to cover the graph, which relates to Gallai's Conjecture, as discussed in [6], [7] and [8]. The second class consists of decompositions into paths or cycles of varying lengths, which may or may not be distinct. Studies such as [9] and [10] explore this category, considering different approaches to handling cycles and paths of non-uniform length.

In 2008, Bryant and Horsley [9] demonstrated the existence of dense packings of the complete graph, where the edges are partitioned into pairwise disjoint cycles of arbitrarily specified length. In 2010, they [10] established the necessary and sufficient conditions for decomposing the complete graph K_n into paths of lengths $m_1, m_2, \ldots, m_t$ when $m_1 + m_2 + \cdots + m_t \leq \frac{n(n-1)}{2}$.

We define a *distinct length path decomposition* of a graph G as a partition of $E(G)$ into paths such that every pair of paths has a distinct length. In this paper, we focus on distinct length path decompositions of the complete graph and some of its subgraphs.

*Speaker and corresponding author (atthakorn.sak@sru.ac.th)

2 Preliminaries

If H is a subgraph of G, we denote by $G-H$ the subgraph of G obtained by removing all edges of H while retaining its vertices. We denote by K_n the complete graph on n vertices, and by $S_m = [v_0; v_1, v_2, \ldots, v_m]$ the star graph, consisting of a central vertex v_0 adjacent to m leaves $v_1, v_2, \ldots, v_m$. Similarly, we denote by $P_n = [v_0, v_1, \ldots, v_n]$ the path with n edges and $n+1$ vertices. The cycle C_n is denoted as $C_n = (v_0, v_1, \ldots, v_{n-1})$, indicating a cycle with n edges where v_0 is repeated to signify closure. From the perspective of viewing P and C as subgraphs of G, P is called a *Hamiltonian path* if it contains all the vertices of G, while C is called a *Hamiltonian cycle* if it contains all the vertices of G. If P (or C) is a path (or cycle), we denote by $l(P)$ (or $l(C)$) the number of edges in the path P (or the cycle C).

Throughout this paper, the notation $[x]_n$ denotes the equivalence class of x modulo n, that is, $[x]_n = \{y \in \mathbb{Z} \mid y \equiv x \pmod{n}\}$. In particular, we may write $y = [x]_n$ to mean that y is the least non-negative residue of x modulo n, i.e., the remainder when x is divided by n.

In this paper, we assume that $\{0, 1, 2, \ldots, 2n-1\}$ is the vertex set of the complete graph K_{2n}, and the vertex set of the complete graph K_{2n+1} is $\{0, 1, 2, \ldots, 2n-1, \infty\}$. To construct a *distinct length path decomposition*, we begin with classical results regarding the decomposition of complete graphs. Specifically, the complete graph K_{2n} can be decomposed into n paths of length $2n-1$, as stated in Theorem 2.1, and the complete graph K_{2n+1} can be decomposed into n cycles of length $2n+1$, as stated in Theorem 2.2. These results appear in Bermond [11] and serve as a foundation for developing our distinct length path decomposition, for which we provide new constructive proofs to facilitate their application in our framework.

Theorem 2.1. *The complete graph K_{2n} has a decomposition into n Hamiltonian paths. Moreover, for any $i = 0, 1, \ldots, n-1$, the ith path $P^i = [p_0^i, p_1^i, \ldots, p_{2n-1}^i]$ where*

$$p_k^i = \left[i + (-1)^k \lfloor \frac{k+1}{2} \rfloor\right]_{2n}.$$

Proof. First, we compute the difference between consecutive terms: $p_k^i - p_{k-1}^i$. From the property

$$\sum_{j=0}^{k} (-1)^j j = (-1)^k \lfloor \frac{k+1}{2} \rfloor,$$

we obtain

$$p_k^i - p_{k-1}^i = \begin{cases} -k, & \text{if } k \text{ is odd,} \\ k, & \text{if } k \text{ is even.} \end{cases}$$

Thus, all vertices in P^i are distinct for any given i. Moreover, for any i, j with $i \neq j$, we have $p_k^i \neq p_k^j$ for all k, ensuring that no vertex appears in more than one path at the same position.

We need to show that for every pair of distinct vertices a, b from $\{0, 1, \ldots, 2n-1\}$, there exists some $i \in \{0, 1, \ldots, n-1\}$ and some $k \in \{1, 2, \ldots, 2n-1\}$ such that $\{a, b\} = \{p_k^i, p_{k-1}^i\}$. We define i as follows:

$$i = \begin{cases} \left[\frac{b+a+1}{2}\right]_n & \text{if } b+a \text{ is odd,} \\ \left[\frac{b+a}{2}\right]_n & \text{if } b+a \text{ is even.} \end{cases}$$

Since the sequence p_k^i is defined by

$$p_k^i = \left[i + (-1)^k \lfloor \frac{k+1}{2} \rfloor\right]_{2n},$$

for any $i \in \{0, 1, \ldots, n-1\}$ and a vertex a, we can can determine the position k_a of a in a path P^i as follows:

$$k_a = \begin{cases} 0, & \text{if } a - i = 0, \\ 2(i-a) - 1, & \text{if } a - i < 0, \\ 2(a-i), & \text{if } 0 < a - i < n, \\ 2(2n + i - a) - 1, & \text{if } n \geq a - i. \end{cases}$$

This confirms that every edge $\{a, b\}$ appears in some path P^i.

Finally, we prove the uniqueness of each edge occurrence. Since every edge in K_{2n} appears in some path P^i, the total number of edges in K_{2n}, which is $\binom{2n}{2} = n(2n-1)$, must be less than or equal to the total number of edges in all P^i's, for $i = 0, 1, \ldots, n-1$.

Each path P^i contains exactly $2n-1$ edges, and there are n such paths, giving a total of $n(2n-1)$ edges. Hence, the number of edges covered by the union of all P^i's is exactly equal to the number of edges in K_{2n}.

Therefore, each edge in K_{2n} must appear in exactly one path P^i, and at exactly one position, completing the proof.

□

Theorem 2.2 (Lucus' Theorem)**.** *The complete graph K_{2n+1} has a decomposition into n Hamiltonian cycles.*

Proof. Let $V(K_{2n+1}) = \{0, 1, \ldots, 2n-1, \infty\}$. Considering the complete subgraph G on the vertex set $\{0, 1, \ldots, 2n-1\}$, by Theorem 2.1, for all $i = 0, 1, \ldots, n-1$, P^i is a Hamiltonian path of G, where

$$P^i = [p_0^i, p_1^i, \ldots, p_{2n-1}^i].$$

The remaining edges are $\{\{j, \infty\} : 0 \leq j \leq 2n-1\}$, connecting every vertex in G to ∞. Since the endpoints of each path P^i are i and $n+i$, we define the cycle

$$C^i = (p_0^i, p_1^i, \ldots, p_{2n-1}^i, \infty),$$

which forms a Hamiltonian cycle for each $i \in \{0, 1, \ldots, n-1\}$.

□

3 Main Results

In this section, we establish a series of results on the *distinct length path decomposition* of complete graphs and their subgraphs. A *distinct length path decomposition* refers to a partition of the edge set of a graph into paths such that no two paths have the same length. This type of decomposition provides a structured way to analyze complete graphs and their subgraphs, offering insights into how edges can be efficiently distributed among paths of unique lengths.

We first prove that any complete graph, whether of even-order or odd-order, admits a *distinct length path decomposition*. These results form the foundation for extending our study to subgraphs of complete graphs. In particular, we consider the removal of a cycle from the complete graph and show that the remaining graph still admits a *distinct length path decomposition*. We then extend this further by analyzing the case where a path, instead of a cycle, is removed, and demonstrate that the resulting graph continues to have a *distinct length path decomposition*.

Each of these results is established through a constructive approach that leverages the structural properties of *pairwise partitions* from previously defined sets $\mathcal{P}_{2n+1}$ and $\mathcal{P}^0_{2n-1}$. By carefully selecting and organizing the paths, we ensure that all path lengths remain distinct throughout the decomposition.

To establish the distinct length path decomposition of complete graphs and their subgraphs, we first introduce a key structural property used throughout our proofs. This involves a systematic way to partition the vertex set into pairs, ensuring that each pair contributes to a well-defined path length distribution. The following definition formalizes this concept.

Definition 3.1. Let $n, k \in \mathbb{N}$. We define

$$N_k = \{1, 2, \ldots, k\}$$

and let

$$\mathcal{P}_{2n+1} = \{\{i, 2n+1-i\} : i \in N_n\}.$$

Then, the following properties hold:

1. $\bigcup \mathcal{P}_{2n+1} = \bigcup_{i=1}^{n} \{i, 2n+1-i\} = N_{2n}$.
2. For any distinct $P, Q \in \mathcal{P}_{2n+1}$, we have $P \cap Q = \emptyset$.
3. For each $P \in \mathcal{P}_{2n+1}$, we have $\sum P = 2n+1$.

Similarly, let

$$\mathcal{P}^0_{2n-1} = \{\{i, 2n-1-i\} : i \in N_{n-1} \cup \{0\}\}.$$

Then, the following properties hold:

1. $\bigcup \mathcal{P}^0_{2n-1} = \bigcup_{i=0}^{n} \{i, 2n-1-i\} = N_{2n-1} \cup \{0\}$.
2. For any distinct $P, Q \in \mathcal{P}^0_{2n-1}$, we have $P \cap Q = \emptyset$.
3. For each $P \in \mathcal{P}^0_{2n-1}$, we have $\sum P = 2n-1$.

With the structure of $\mathcal{P}$ established, we now proceed to prove the existence of a distinct length path decomposition in complete graphs. We first consider the case of an even-order complete graph and demonstrate that it admits a distinct length path decomposition.

Theorem 3.2. *The complete graph K_{2n} has a distinct length path decomposition.*

Proof. By Theorem 2.2, K_{2n} has a decomposition with a collection of Hamiltonian paths $\{P^i : i \in N_{n-1} \cup \{0\}\}$. For any $i \in N_{n-1}$, define

$$P^i_- = [p^i_0, \ldots, p^i_i] \quad \text{and} \quad P^i_+ = [p^i_i, \ldots, p^i_{2n-1}].$$

Thus the lengths of paths P^i_- and P^i_+ are $l(P^i_-) = i$ and $l(P^i_+) = 2n-1-i$, respectively. Therefore $\{\{0, l(P^0)\}\} \cup \{\{l(P^i_-), l(P^i_+)\} : i \in N_{n-1}\} = \mathcal{P}^0_{2n-1}$, so K_{2n} has a distinct length path decomposition by $\{P^0, P^1_-, P^1_+, \ldots, P^{n-1}_-, P^{n-1}_+\}$. □

Theorem 3.3. *The complete graph K_{2n+1} has a distinct length path decomposition.*

Proof. By Theorem 2.1, K_{2n+1} has a decomposition with a collection of Hamiltonian cycles $\{C^i : i \in N_{n-1} \cup \{0\}\}$. For any $i \in N_{n-1} \cup \{0\}$, define

$$P_-^i = [\infty, p_0^i, \ldots, p_i^i] \quad \text{and} \quad P_+^i = [p_i^i, \ldots, p_{2n-1}^i, \infty].$$

Thus the lengths of paths P_-^i and P_+^i are $l(P_-^i) = i+1$ and $l(P_+^i) = 2n-i$, respectively. Therefore $\{\{l(P_-^i), l(P_+^i)\} : i \in N_{n-1} \cup \{0\}\} = \mathcal{P}_{2n+1}$, so K_{2n+1} has a distinct length path decomposition by $\{P_-^0, P_+^0, P_-^1, P_+^1, \ldots, P_-^{n-1}, P_+^{n-1}\}$. □

Having established that the complete graph admits a distinct length path decomposition, we now extend our analysis to the case where a cycle is removed from the graph. Specifically, we show that even after deleting certain cycles, the remaining structure still allows for a distinct length path decomposition.

Theorem 3.4. *Let $n > 2$. A graph $G = K_{2n} - C_{m+1}$ for any $m = 2, 4, \ldots, 2n-1$ has a distinct length path decomposition.*

Proof. By Theorem 2.1, K_{2n} has a distinct length path decomposition by a collection $\{P^i : i = 0, 1, \ldots, n-1\}$ where each path P^i is given by $P^i = [p_0^i, p_1^i, \ldots, p_{2n-1}^i]$ with $p_k^i = \left[i + (-1)^k \lfloor \frac{k+1}{2} \rfloor\right]_{2n}$.

Without loss of generality, consider the cycle C_{m+1} defined as $C_{m+1} = (p_0^0, p_1^0, \ldots, p_m^0)$. Since C_{m+1} is a cycle, it follows that $p_m^0 \neq p_0^0$ and $p_m^0 \neq p_1^0$. Then C_{m+1} can be separated into two paths: $[p_0^0, p_1^0, \ldots, p_m^0]$ and $[p_0^0, p_m^0]$.

Thus, P^0 still remains as $[p_m^0, p_{m+1}^0, \ldots, p_{2n-1}^0]$ with length $l(P^0) = 2n-1-m$. Next, we modify the collection $\{P^i : i \in N_{n-1}\}$ by adjusting the deleted edges associated with C_{m+1}.

Case $n = 2t+1$ (n is odd): First, we determine i such that the edge $\{p_0^0 = 0, p_m^0\}$ belongs to P^i using the following formula:

$$i = \begin{cases} \left[\frac{p_m^0+1}{2}\right]_n & \text{if } p_m^0 \text{ is odd,} \\ \left[\frac{p_m^0}{2}\right]_n & \text{if } p_m^0 \text{ is even.} \end{cases}$$

For each $i \in N_{n-1}$, since p_k^i is given by $p_k^i = \left[i + (-1)^k \lfloor \frac{k+1}{2} \rfloor\right]_{2n}$, it follows that $p_{2i-2}^i = 2i-1, p_{2i-1}^i = 0$ and $p_{2i}^i = 2i$. Next, we cut the path P^i at $p_{2i-1}^i = 0$

Let $P_-^i = [p_0^i, p_1^i, \ldots, p_{2i-2}^i = 2i-1, p_{2i-1}^i = 0]$ and $P_+^i = [p_{2i-1}^i = 0, p_{2i}^i = 2i, \ldots, p_{2n-1}^i]$. Then $l(P_-^i) = 2i-1$ and $l(P_+^i) = 2n-2i$.

If $1 \leq i \leq t$, then $2 \leq 2i \leq 2t \leq n-1$. Consider,

$$P^{2i} = [p_0^{2i} = 2i, p_1^{2i} = 2i-1, p_2^{2i}, \ldots, p_{2n-1}^{2i}].$$

Thus we can remove the edge $\{2i, 2i-1\}$ in P^{2i} and $l(P^{2i} - \{2i\}) = 2n-2$.

Let P_*^i and P_{**}^i be defined as follows:

$$P_*^i = \begin{cases} P_-^i & \text{if } p_m^0 \text{ is odd,} \\ P_+^i & \text{if } p_m^0 \text{ is even.} \end{cases} \qquad P_{**}^i = \begin{cases} P_+^i & \text{if } p_m^0 \text{ is odd,} \\ P_-^i & \text{if } p_m^0 \text{ is even.} \end{cases}$$

We remove the edge connecting p_m^0 to 0 in P_*^i and attach $\{2i-1, 2i\}$ to obtain the new paths:

$$P_{*,2i}^i = P_*^i \cup \{2i-1, 2i\}.$$

Then $\{l(P^i_{*,2i}), l(P^i_{**})\} = \{2i-1, 2n-2i\}$, so the resulting path lengths remain $2i-1$ and $2n-2i$.

We now obtain the following collection of new paths: $P^0 = [p^0_m, \ldots, p^0_{2n-1}]$(of length $2n-1-m$), $P^{2i} - \{2i\}$(of length $2n-2$) and $P^i_{*,2i}, P^i_{**}$ with lengths $\{l(P^i_{*,2i}), l(P^i_{**})\} = \{2i-1, 2n-2i\}$. Additionally, for $j = 1, 2, \ldots, n-1$ with $j \neq i, 2i$, we retain P^j with length $2n-1$.

Since the number of remaining length pairs in $\mathcal{P}^0_{2n-1}$ after removing $\{m, l(P^0)\}$, $\{l(P^i_{*,2i}), l(P^i_{**})\}$, $\{1, l(P^{2i} - \{2i\})\}$ is $n-3$ and $|\{P^j : j \in N_{n-1} - \{i, 2i\}| = n-3$, we can cut each path P^j such that its lengths correspond to a pair of path lengths $\{a, b\}$ from the remaining elements in $\mathcal{P}^0_{2n-1} - \{\{m, l(P^0)\}, \{l(P^i_{*,2i}), l(P^i_{**})\}, \{1, l(P^{2i} - \{2i\})\}\}$. Thus, in this case, $K_{2n} - C_{m+1}$ has a distinct length path decomposition.

If $t < i \leq n-1$, then $0 \leq 2i - n \leq n-2$. Since n is odd, $2i - n \neq 0$, so $0 < 2i - n \leq n-2$ Consider,

$$P^{2i-n} = [p^{2i-n}_0, \ldots, p^{2i-n}_{2n-2} = 2i-1, p^{2i-n}_{2n-1} = 2i].$$

Thus we can cut the edge $\{2i-1, 2i\}$ in P^{2i-n} and $l(P^{2i-n} - \{2i\}) = 2n-2$.

Let P^i_* and P^i_{**} be defined as follows:

$$P^i_* = \begin{cases} P^i_- & \text{if } p^0_m \text{ is odd,} \\ P^i_+ & \text{if } p^0_m \text{ is even.} \end{cases} \quad P^i_{**} = \begin{cases} P^i_+ & \text{if } p^0_m \text{ is odd,} \\ P^i_- & \text{if } p^0_m \text{ is even.} \end{cases}$$

We remove the edge connecting p^0_m to 0 in P^i_* and attach $\{2i-1, 2i\}$ to obtain the new paths:

$$P^i_{*,2i} = P^i_* \cup \{2i-1, 2i\}.$$

Then $\{l(P^i_{*,2i}), l(P^i_{**})\} = \{2i-1, 2n-2i\}$, so the resulting path lengths remain $2i-1$ and $2n-2i$.

We now obtain a collection of new paths: $P^0 = [p^0_m, \ldots, p^0_{2n-1}]$(of length $2n-1-m$), $P^{2i-n} - \{2i\}$(of length $2n-2$) and $P^i_{*,2i}, P^i_{**}$ with lengths $\{l(P^i_{*,2i}), l(P^i_{**})\} = \{2i-1, 2n-2i\}$, along with P^j(of length $2n-1$) when $j = 1, 2, \ldots, n-1$ and $j \neq i, 2i$. Applying the same argument as in the previous case, $K_{2n} - C_{m+1}$ has a distinct length path decomposition.

Case $n = 2t$ (n is even): Similarly, we can compute i for the edge $\{0, p^0_m\}$ is in P^i. Let $P^i_- = [p^i_0, p^i_1, \ldots, p^i_{2i-2} = 2i-1, p^i_{2i-1} = 0]$ and $P^i_+ = [p^i_{2i-1} = 0, p^i_{2i} = 2i, p^i_{2i+1}, \ldots, p^i_{2n-1}]$. Then the length of P^i_- and P^i_+ are $2i-1$ and $2n-2i$, respectively.

If $1 \leq i < t$, then $2 \leq 2i < n-1$. Consider the path:

$$P^{2i} = [p^{2i}_0 = 2i, p^{2i}_1 = 2i-1, p^{2i}_2, \ldots, p^{2i}_{2n-1}].$$

By removing the edge $\{2i, 2i-1\}$ from P^{2i}, we obtain a path $P^{2i} - \{2i\}$ of length $2n-2$.

Following the approach in the case $n = 2t+1$, we construct $P^i_{*,2i}$ and P^i_{**} such that $\{l(P^i_{*,2i}), l(P^i_{**})\} = \{2i-1, 2n-2i\}$. Then $K_{2n} - C_{m+1}$ has a distinct length path decomposition.

Now, if $t < i < n-1$, then $0 < 2i - n < n-2$. Consider,

$$P^{2i-n} = [p^{2i-n}_0, \ldots, p^{2i-n}_{2n-2} = 2i-1, p^{2i-n}_{2n-1} = 2i].$$

Removing the edge $\{2i-1, 2i\}$ from P^{2i-n} and the length of $P^{2i-n} - \{2i\}$ is $2n-2$.

Using a similar construction as in the case $n - 2t + 1$, we define $P^i_{*,2i}$ and P^i_{**} such that $\{l(P^i_{*,2i}), l(P^i_{**})\} = \{2i-1, 2n-2i\}$. Therefore, $K_{2n} - C_{m+1}$ has a distinct length path decomposition.

In the final case, when $i = t$. we have:

$$t = \begin{cases} \left[\frac{p_m^0+1}{2}\right]_n & \text{if } p_m^0 \text{ is odd,} \\ \left[\frac{p_m^0}{2}\right]_n & \text{if } p_m^0 \text{ is even.} \end{cases}$$

This implies that either $p_m^0 = n = p_{2n-1}^0$ or $p_m^0 = n-1 = p_{2n-2}^0$, and we obtain $P^{2t-n} = P^0$.

If $p_m^0 = n-1$, which implies $p_{2n-2}^t = n-1$, then P^0 simplifies to the path $[n-1, n]$. Let $P_-^t - \{0\} = [p_0^t, \ldots, p_{2t-2}^t = n-1]$ and $P_+^t = [p_{2t-1}^t = 0, p_{2t}^t = n, \ldots, p_{2n-1}^t]$. Since n is included in P_+^t, we can extend $P_-^t - \{0\}$ by appending $[n-1, n]$, forming the new path:

$$P_{-,n}^t = [p_0^t, \ldots, p_{2t-2}^t = n-1, n].$$

As a result, the new collection of new paths consists of: $P_{-,n}^t$ and P_+^t with lengths n and $n-1$, respectively, along with P^j(of length $2n-1$) for $j = 1, 2, \ldots, n-1$ and $j \neq t$. We then separate P^j to two paths with the lengths $\{j, 2n-1-j\}$, ensuring that $K_{2n} - C_{m+1}$ has a distinct length path decomposition.

If $p_m^0 = n = p_{2n-1}^0$, then P^0 is removed since p_{2n-1}^0 is an endpoint of P^0. Consider $P_-^t = [p_0^t, \ldots, p_{2t-2}^t = n-1, p_{2t-1}^t = 0]$ and $P_+^t - \{0\} = [p_{2t}^t = n, \ldots, p_{2n-1}^t]$.

Since $n > 2$, $t > 1$, we have $t < n-1$, ensuring that $P^t \neq P^{n-1}$. Next, we identify the edge $\{2n-3, 0\}$ in the path $P^{n-1} = [p_0^{n-1}, \ldots, p_{2n-4}^{n-1} = 2n-3, p_{2n-3}^{n-1} = 0, p_{2n-2}^{n-1} = 2n-2, p_{2n-1}^{n-1}]$.

Let $P_-^{n-1} = [p_0^{n-1}, \ldots, p_{2n-4}^{n-1} = 2n-3]$ and $P_+^{n-1} = [p_{2n-3}^{n-1} = 0, p_{2n-2}^{n-1} = 2n-2, p_{2n-1}^{n-1}]$ with lengths $2n-4$ and 2, respectively. Since $n-1 < 2n-3$ and $2n-3 \in P_+^t$, we attach the edge $\{2n-3, 0\}$ to P_-^t, forming the extended path:

$$P_{-,2n-3}^t = [p_0^t, \ldots, p_{2t-2}^t = n-1, p_{2t-1}^t = 0, 2n-3].$$

The updated collection of new paths consists of P_-^{n-1}, P_+^{n-1} with lengths $2n-4$ and 2 and $P_{-,2n-3}^t, P_+^t$ with lengths n and $n-1$, along with P^j(of length $2n-1$) when $j = 1, 2, \ldots, n-1$ and $j \neq t, n-1$. Therefore $K_{2n} - C_{m+1}$ has a distinct length path decomposition. □

Next, we consider the case where a cycle is removed from an odd-order complete graph. While the general strategy remains similar to the even-order case, the odd-order case presents additional structural challenges. In the even case, constructing a distinct length path decomposition requires carefully cutting some paths and reattaching them to ensure that all paths have unique lengths. In contrast, the odd case benefits from the presence of an extra vertex, allowing a more direct decomposition without requiring such modifications. Despite these differences, we show that the remaining structure after removing a cycle from K_{2n+1} still admits a distinct length path decomposition.

Theorem 3.5. *A graph $G = K_{2n+1} - C_{m+1}$ for any $m = 2, 4, \ldots, 2n$ has a distinct length path decomposition.*

Proof. By Theorem 2.2, K_{2n+1} has a distinct length path decomposition by a collection $\{C^i : i \in N_{n-1}\}$ where $C^i = (p_0^i, p_1^i, \ldots, p_{2n-1}^i, \infty)$ and $p_k^i = \left[i + (-1)^k \lfloor \frac{k+1}{2} \rfloor\right]_{2n}$.

If $m = 2n$, without loss of generality, let $C_{m+1} = (p_0^0, p_1^0, \ldots, p_{2n-1}^0, \infty)$. Then C^0 is deleted, for each $i \in N_{n-1}$, define

$$P_-^i = [\infty, p_0^i, \ldots, p_i^i] \quad \text{and} \quad P_+^i = [p_i^i, \ldots, p_{2n-1}^i, \infty].$$

Thus $l(P_-^i) = i+1$ and $l(P_+^i) = 2n-i$. Therefore $\{\{l(P_-^i), l(P_+^i)\} : i \in N_{n-1}\} = \mathcal{P}_{2n+1} - \{\{1, 2n-2\}\}$, so $K_{2n+1} - C_{2n+1}$ has a distinct path decomposition.

If $m < 2n$, without loss of generality, let $C_{m+1} = (p_0^0, p_1^0, \ldots, p_m^0)$. Thus $p_m^0 \neq p_0^0$ and $p_m^0 \neq p_1^0$ because C_{m+1} is a cycle. Then C_{m+1} can be separated into two paths: $[p_0^0, p_1^0, \ldots, p_m^0]$ and $[p_0^0, p_m^0]$, so C^0 still remains as $P^0 := [p_m^0, \ldots, p_{2n-1}^0, \infty, p_0^0]$, where $l(P^0) = 2n+1-m$. Next, we find j such that the edge $\{p_0^0 = 0, p_m^0\}$ is in C^j by following formula:

$$j = \begin{cases} \left[\frac{p_m^0+1}{2}\right]_n & \text{if } p_m^0 \text{ is odd,} \\ \left[\frac{p_m^0}{2}\right]_n & \text{if } p_m^0 \text{ is even.} \end{cases}$$

Let

$$P^j = \begin{cases} [p_{2j-1}^j, \ldots, p_{2n-1}^j, \infty, p_0^j, \ldots, p_{2j-2}^j], & \text{if } \ p_m^0 \text{ is odd,} \\ [p_{2j}^j, \ldots, p_{2n-1}^j, \infty, p_0^j, \ldots, p_{2j-1}^j], & \text{if } \ p_m^0 \text{ is even.} \end{cases}$$

Hence $l(P^j) = 2n$. Since $\left|\mathcal{P}_{2n+1} - \{\{1, 2n\}, \{m-1, 2n+1-m\}\}\right| = n-2$ and $|\{P^i : i \in N_{n-1} - \{j\}\}| = n-2$, $K_{2n+1} - C_{m+1}$ has a distinct length path decomposition. □

So far, we have demonstrated that removing a cycle from a complete graph still allows for a distinct length path decomposition. We now extend our study to the case where a path is removed instead of a cycle. Unlike cycles, which inherently provide a closed structure, paths introduce new endpoints that must be carefully managed in the decomposition process.

However, it is not difficult to see that the remaining structure in K_{2n+1} continues to admit a distinct length path decomposition. This can be achieved simply by relabeling the removed path so that it matches one of the paths already present in Theorem 3.3. In the following theorem, we provide an alternative relabeling method that illustrates the construction more explicitly.

Theorem 3.6. *A graph $G = K_{2n+1} - P_{m+1}$ where $m \in N_{n-1}$ has a distinct length path decomposition.*

Proof. By Theorem 2.1, there is a collection of cycles $\{C^i : i \in N_{n-1} \cup \{0\}\}$ where

$$C^i = (p_0^i, p_1^i, \ldots, p_{2n-1}^i, \infty).$$

Without loss of generality, assume that $P_{m+1} = [p_0^0, p_1^0, \ldots, p_m^0]$ $(l(P_{m+1}) = m)$. Then C^0 remains as $[p_m^0, p_{m+1}^0, \ldots, p_{2n-1}^0, \infty, p_0^0]$ with the length $2n+1-m$.

Then $|\mathcal{P}_{2n+1} - \{\{m, 2n+1-m\}\}| = n-1$, so for any $i \in N_{n-1} - \{m-1, 2n+2-m\}$, define

$$P_-^i = [\infty, p_0^i, \ldots, p_i^i] \quad \text{and} \quad P_+^i = [p_i^i, p_{i+1}^i, \ldots, p_{2n-1}^i, \infty].$$

For $i_0 = \min\{m, 2n+1-m\}$, define $P_-^{i_0} = [\infty, p_0^{i_0}]$ and $P_+^{i_0} = [p_0^{i_0}, p_1^{i_0}, \ldots, p_{2n-1}^{i_0}, \infty]$. Then $l(P_-^i) = i+1$, $l(P_+^i) = 2n-i$, $l(P_-^{i_0}) = 1$ and $l(P_+^{i_0}) = 2n$. Then $K_{2n+1} - P_{m+1}$ has distinct length path decomposition. □

We next consider the case where a path is removed from an even-order complete graph. As in previous results, the even-order case presents additional challenges compared to the odd case. In particular, the absence of an extra vertex in K_{2n} requires a more intricate reorganization of paths to ensure a distinct length path decomposition. Nevertheless, the result can again be obtained by relabeling the removed path to match the appropriate structure in Theorem 3.2. In the next theorem, we describe an alternative relabeling approach to demonstrate the construction.

Theorem 3.7. *A graph $G = K_{2n} - P_{m+1}$ where $m \in N_{n-1}$ has a distinct length path decomposition.*

Proof. By Theorem 2.2, there is a collection of paths $\{P^i : i \in N_{n-1} \cup \{0\}\}$ where

$$P^i = [p_0^i, p_1^i, \ldots, p_{2n-1}^i].$$

Without loss of generality, assume that $P_{m+1} = [p_0^0, p_1^0, \ldots, p_m^0]$ $(l(P_{m+1}) = m)$. Then P^0 remains as $[p_m^0, p_{m+1}^0, \ldots, p_{2n-1}^0]$ with the length $2n - 1 - m$.

Then $|\mathcal{P}_{2n-1}^0 - \{\{m, 2n-1-m\}\}| = n - 1$, so for any $i \in N_{n-1} - \{m, 2n-1-m\}$, define

$$P_-^i = [p_0^i, \ldots, p_i^i] \quad \text{and} \quad P_+^i = [p_i^i, p_{i+1}^i, \ldots, p_{2n-1}^i].$$

Then $l(P_-^i) = i$, $l(P_+^i) = 2n - 1 - i$ and hence for $i_0 = \min\{m, 2n-1-m\}$,

$$\{\{l(P_-^i), l(P_+^i)\} : i \in N_{n-1} - \{i_0\}\} \cup \{\{0, l(P^{i_0}) = 2n-1\}\} = \mathcal{P}_{2n-1}^0 - \{\{m, 2n-1-m\}\}$$

Then $K_{2n} - P_{m+1}$ has distinct length path decomposition. □

Finally, we consider the case of a complete graph with a star graph removed, which we denote as S_m. If $m = 1, 2$, it is evident that S_m is a path, allowing us to apply the previous theorems to obtain a distinct length path decomposition. Therefore, we focus on the case where $m \geq 3$, dividing our analysis into two cases: the even-order complete graph and the odd-order complete graph. To establish that these graphs admit a distinct length path decomposition, we employ a technique that pairs the lengths of two subpaths. The following theorems confirm the validity of this approach:

Theorem 3.8. *A graph $G = K_{2n} - S_m$ where $m = 3, \ldots, 2n-1$ has a distinct length path decomposition.*

Proof. By Theorem 2.2, there exists a collection of paths $\{P^i : i \in N_{n-1} \cup \{0\}\}$, where each path is given by:

$$P^i = [p_0^i, p_1^i, \ldots, p_{2n-1}^i].$$

We aim to show that after removing the edges of the star graph S_m, the remaining graph still admits a distinct length path decomposition.

Case m is even: Without loss of generality, assume that $S_m = [0; 1, 2, \ldots, m]$ and that the edges are given by

$$E(K_{1,m}) = \{\{0, j\} : j \in N_m\}.$$

The removal of these edges affects the paths containing vertex 0 since it was originally part of several paths in the decomposition. To ensure a valid distinct length path decomposition, we modify these paths accordingly.

For each $i \in N_{\frac{m}{2}}$, we define the subpaths obtained after removing vertex 0:

$$P_-^i = [p_0^i, p_1^i, \ldots, p_{2i-2}^i], \quad P_+^i = [p_{2i}^i, \ldots, p_{2n-1}^i].$$

Their respective lengths satisfy

$$l(P_-^i) = 2i - 2, \quad l(P_+^i) = 2n - 2i - 1.$$

Considering P^{i-1} and P^i for $i = 2, \ldots, \frac{m}{2}$, we obtain:

$$P_-^{i-1} = [p_0^{i-1}, p_1^{i-1}, \ldots, p_{2i-4}^{i-1}], \quad P_+^{i-1} = [p_{2i-2}^{i-1}, \ldots, p_{2n-1}^{i-1}].$$

with path lengths

$$l(P_-^{i-1}) = 2i - 4, \quad l(P_+^{i-1}) = 2n - 2i + 1.$$

Noting that

$$l(P_+^{i-1}) + l(P_-^i) = (2n - 2i + 1) + (2i - 2) = 2n - 1,$$

it follows that the remaining paths preserve the distinct length property since they still conform to the allowable set $\mathcal{P}_{2n-1}^0$.

Since the remaining available pairs of lengths satisfy

$$\left|\mathcal{P}_{2n-1}^0 - \left(\{\{l(P_+^{i-1}), l(P_-^i)\} : i = 2, \ldots, \frac{m}{2}\} \cup \{\{m, 2n - m - 1\}\}\right)\right| = n - \frac{m}{2}$$

where $l(P_-^1) = 0$ and $l(P_+^{\frac{m}{2}}) = 2n - m - 1$, the number of paths needed for a valid decomposition remains unchanged. Hence, $K_{2n} - S_m$ still has a distinct length path decomposition.

Case m is odd : Without loss of generality, assume that $S_m = [0; 1, 2, \ldots, m - 1, 2n - 1]$ and that the edges are given by

$$E(S_m) = \{\{0, j\} : j \in N_{m-1}\} \cup \{\{0, 2n - 1\}\}.$$

The removal of these edges affects the paths containing vertex 0 since it was originally part of several paths in the decomposition. To ensure a valid distinct length path decomposition, we modify these paths accordingly.

To construct the subpaths, for each $i \in N_{\frac{m-1}{2}}$, we define

$$P_-^i = [p_0^i, p_1^i, \ldots, p_{2i-2}^i], \quad P_+^i = [p_{2i}^i, \ldots, p_{2n-1}^i].$$

Their corresponding lengths are given by

$$l(P_-^i) = 2i - 2, \quad l(P_+^i) = 2n - 2i - 1.$$

Examining P^{i-1} and P^i for $i = 2, \ldots, \frac{m}{2}$, we obtain

$$l(P_-^{i-1}) = 2i - 4, \quad l(P_+^{i-1}) = 2n - 2i + 1.$$

Observing that

$$l(P_+^{i-1}) + l(P_-^i) = (2n - 2i + 1) + (2i - 2) = 2n - 1,$$

we conclude that the modified paths preserve the distinct length condition.

Additionally, after removing the edge $\{0, 2n - 1\}$ from P^0, the remaining path is

$$P^0 = [p_1^0 = 2n - 1, p_2^0, \ldots, p_{2n-1}^0]$$

with length $2n - 2$.

Since the remaining available pairs of lengths satisfy

$$\left|\mathcal{P}_{2n-1}^0 - \left(\{\{l(P_+^{i-1}), l(P_-^i)\}\}_{i=2}^{\frac{m-1}{2}} \bigcup \{\{1, 2n - 2\}, \{m - 1, 2n - m\}\}\right)\right| = n - \frac{m-1}{2} - 1,$$

where $l(P_-^1) = 0$ and $l(P_+^{\frac{m-1}{2}}) = 2n - m$, the number of paths required for a valid decomposition remains unchanged. Hence, $K_{2n} - S_m$ still has a distinct length path decomposition. □

Theorem 3.9. *A graph $G = K_{2n+1} - S_m$ where $m = 3, \ldots, 2n$ has a distinct length path decomposition.*

Proof. By Theorem 2.1, there is a collection of cycles $\{C^i : i \in N_{n-1} \cup \{0\}\}$, where each cycle is given by

$$C^i = (p_0^i, p_1^i, \ldots, p_{2n-1}^i, \infty).$$

Without loss of generality, assume that $S_m = [0; 1, 2, \ldots, m]$ and that the edges are given by

$$E(S_m) = \{\{0, j\} : j \in N_m\}.$$

The removal of these edges affects the cycles containing vertex 0. To ensure a valid distinct length path decomposition, we modify these paths accordingly.

Case m is even: For every $i = 1, \ldots, \frac{m}{2}$, we define a path from C^i by

$$PC^i = [p_{2i}^i, p_{2i+1}^i, \ldots, p_{2n-1}^i, \infty, p_0^i, \ldots, p_{2i-2}^i],$$

where $l(PC^i) = 2n - 1$.

If $m = 2n$, then $K_{2n+1} - K_{1,2n}$ has a distinct length path decomposition since we separate P^i by $\mathcal{P}_{2n-1}^0$.

If $m \neq 2n$, then we can separate PC^i by PC_-^i and PC_+^i with $l(PC_-^i) = 2i$ and $l(PC_+^i) = 2(n-i) - 1$ ($l(PC_-^i) \neq l(PC_+^i)$). Thus $l(PC_+^{i-1}) + l(PC_-^i) = (2(n-i)+1) + 2i = 2n+1$ for any $i = 2, 3, \ldots, \frac{m}{2}$. Then Since the remaining available pairs of lengths satisfy

$$\left|\mathcal{P}_{2n+1} - \left(\{\{l(PC_+^{i-1}), l(CP_-^i)\}\}_{i=2}^{\frac{m}{2}} \bigcup \{\{m+2, 2n-m-1\}\}\right)\right| = n - \frac{m}{2},$$

where $l(CP_-^1) = 0$ and $l(CP_+^{\frac{m}{2}}) = 2n - m - 1$. Hence, $K_{2n+1} - S_m$ still has a distinct length path decomposition.

Case m is odd: For every $i = 1, \ldots, \frac{m-1}{2}$, we define a path from C^i by

$$PC^i = [p_{2i}^i, p_{2i+1}^i, \ldots, p_{2n-1}^i, \infty, p_0^i, \ldots, p_{2i-2}^i],$$

where $l(PC^i) = 2n - 1$ and

$$PC^{\frac{m+1}{2}} = [p_m^{\frac{m+1}{2}} = 0, p_{m+1}^{\frac{m+1}{2}} = m+1, p_{2i+1}^{\frac{m+1}{2}}, \ldots, p_{2n-1}^{\frac{m+1}{2}}, \infty, p_0^{\frac{m+1}{2}}, \ldots, p_{2i-2}^{\frac{m+1}{2}}],$$

where $l(PC^{\frac{m+1}{2}}) = 2n$.

Then we can separate PC^i by PC_-^i and PC_+^i with $l(PC_-^i) = 2i$ and $l(PC_+^i) = 2(n-i) - 1$ ($l(PC_-^i) \neq l(PC_+^i)$). Thus $l(PC_+^{i-1}) + l(PC_-^i) = (2(n-i)+1) + 2i = 2n+1$ for any $i = 2, 3, \ldots, \frac{m-1}{2}$. Then Since the remaining available pairs of lengths satisfy

$$\left|\mathcal{P}_{2n+1} - \left(\{\{l(PC_+^{i-1}), l(CP_-^i)\}\}_{i=2}^{\frac{m-1}{2}} \bigcup \{\{1, 2n\}, \{m+1, 2n-m\}\}\right)\right| = n - \frac{m-1}{2} - 1,$$

where $l(CP_-^1) = 0$ and $l(CP_+^{\frac{m-1}{2}}) = 2n - m$. Hence, $K_{2n+1} - K_{1,m}$ still has a distinct length path decomposition. □

4 Discussion

The study of whether other graphs can admit a distinct length path decomposition remains an interesting question. The first aspect to consider is the longest path length within the graph, which corresponds to its *diameter*. If the diameter of a graph is k, but the number of edges exceeds $\frac{k(k-1)}{2}$, then the graph cannot have a distinct length path decomposition.

Furthermore, if we apply the same path construction method as in this paper, we must not only consider the number of edges but also analyze how its complement fits within the paths or cycles used. Specifically, we must examine the role of the complement in relation to the paths in Theorem 2.1 and the cycles in Theorem 2.2.

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments.

References

[1] B. Alspach and H. Gavlas, Cycle decompositons of K_n and $K_n - I$, J. Combin. Theory Ser. B 81 (2001), 77–99.

[2] F. Botler, G.O. Mota, M.T.I. Oshiro, Y. Wakabayashi, Decomposing regular graphs with prescribed girth into paths of given length, European J. Combin. 66(2017) 28–36.

[3] F. Botler, G.O. Mota, M.T.I. Oshiro, Y. Wakabayashi, Decomposing highly connected graphs into paths of length five, Discrete Appl. Math. 245(2018) 128–138.

[4] F. Botler and A. Talon, Decomposition of 8-regular graphs into paths of length 4, Electronic Notes in Discrete Math. 61(2017) 147–153.

[5] T. W. Shyu, Decomposition of complete graphs into paths and stars, Discrete Math. 310(2010) 2164—2169.

[6] F. Botler and A. Jiménez, On path decompositions of $2k$-regular graphs, Discrete Math. 340(2017) 1405–1411.

[7] Y. Chu, G. Fan, and C. Zhou, Path decompositions of triangle-free graphs. Discrete Math. 345(7), (2022).112866.

[8] C. Lu,and N. Yi, The problem of path decomposition for graphs with treewidth at most 4. Discrete Math. 347(6), (2024).113957.

[9] D. Bryant and D. Horsley, Packing cycles in complete graphs, J. Combin. Theory Ser. B 98(2008) 1014–1037.

[10] D. Bryant, Packing paths in complete graphs, J. Combin. Theory Ser. B 100(2010) 206–215.

[11] J.-C. Bermond, Hamiltonian Decompositions of Graphs, Directed Graphs and Hypergraphs. Ann. of Discrete Math., 1978, 3, pp.21-28.

Integrating Active Learning with The Learning Pit Model to Foster Mathematical Resilience

Gritchanat Sangsupakorn[1,*] and Ratchanikorn Chonchaiya[2]

[1]Science Education, Faculty of Science,
King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand
[2]Department of Mathematics, Faculty of Science,
King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand

Abstract

This study explored the integration of Active Learning with the Learning Pit Model as a framework for designing lesson plans that foster students' mathematical resilience. Five lesson plans on algebraic vectors were developed aimed at high school students, each following a purposefully structured instructional sequence. The lesson plans were evaluated through the congruence evaluation form and were continuously refined through consultations with experts, including an assessment specialist and an experienced high school mathematics teacher. Their feedback ensured the lessons were engaging, pedagogically sound, and effective in fostering resilience. The study employed a qualitative approach, focusing on expert evaluations. The findings from experts review highlight how structured lesson plans encourage students to embrace mathematical challenges in class. The discussion provides insights into lesson plans designs and some limitations. This study offers a practical framework for mathematics educators seeking to design engaging lessons that nurture students' mathematical resilience.

Keywords: Active Learning, The Learning Pit Model, Mathematical resilience
2020 MSC: 97C70, 97D40

1 Introduction

1.1 Significance of The Study

The researcher had the opportunity to observe learning management, including teaching at an extra large secondary school. When there are self-organized tests, many students engage in casual conversation. Upon receiving the test paper, many students expressed anxious through their words. Instead of attempting, a significant number of students randomly selected a single answer choice for all questions without reading them. Despite their lack of effort, they did not display any visible signs of concern. Previously, the tests were allowed to leave early, some submitted nearly blank paper within the first five minutes. When the researcher adjusted the rule to require students to stay for the full duration, the same group resorted to sleeping throughout

*Speaker and corresponding author (gritchanat.sangsu@kmutt.ac.th)

the test duration. Whether their test scores were consistently near zero, they exhibited no disappointment yet easily accepted the results. While these behaviors may appear unremarkable at first glance, they could indicate a sense of "learned helplessness" [1], where students disengage from academic challenges due to a persistent belief that their efforts will not lead to success.

Learned helplessness is a psychological condition in which individuals perceive that their efforts have no impact on outcomes, resulting in disengagement and lack of motivation. [1]. In the context of learning mathematics, learned helplessness is closely related to mathematical anxiety, as students who repeatedly experience failure may cause the belief that success is beyond their control. Over time, the students learn that their effort does not achieve the outcome of the correct answer. Thus, they conclude that their effort is not enough to achieve success in mathematics and next time the learner expends less effort [2]. In contrast, **mathematical resilience** provides a constructive approach to overcoming such emotional barriers [3]. Students with mathematical resilience persist despite difficulties [1, 3, 4], collaborate with peers, and effectively articulate their mathematical thinking [3]. They actively seek support, develop new skills as needed, and recognize that sustained effort increases their likelihood of success [1, 4]. By fostering mathematical resilience, students can mitigate the negative effects of learned helplessness and mathematical anxiety, ultimately enhancing their engagement and confidence in mathematics [3].

Active Learning encourages students to engage with course contents through various interactive methods such as problem-solving exercises, small group discussions, case studies, and other activities that require them to apply what they are learning [5]. Unlike passive learning, where students primarily listen to an instructor, active learning promotes engagement through participation and discussion [6, 7], As students face challenges in this process, proper support helps them build resilience in their learning. **The Learning Pit Model** further aids in this process by making students more comfortable with challenges, helping them develop problem-solving strategies, and encouraging them to articulate their learning [8]. The main role of the model is to design lessons that are challenging to the students, while the instructor provides support in the form of guidance or encouragement, thus resulting in more effort, persistence, and comprehension of the lessons [9,10]. To do so, the model can be done neatly by preparing some questions related yet crucial to the lesson when they are in the pit. Taking a look back on the different pathways to climb up the pit as a talk on what they have learned from the beginning of the class is a practical way to construct and integrate both their acquired and shared knowledge [11].

Based on the aforementioned concepts, this study aims to establish a guideline for developing tools that integrate Active Learning with the Learning Pit Model to enhance students' mathematical resilience. By designing learning activities that encourage students to confront challenges and overcome emotional barriers, this approach seeks to foster persistence and confidence in mathematical learning. Developing such tools could provide a meaningful framework for cultivating resilience, ensuring that students not only engage with mathematical concepts but also develop the mathematical resilience necessary to navigate further academic challenges in mathematics effectively.

1.2 The Purpose of The Study

To develop lesson plans integrating Active Learning with the Learning Pit Model which aimed at fostering students' mathematical resilience in specific mathematical contents.

2 Preliminaries

2.1 Active Learning

Active Learning is learning that allows students to talk, listen, read, write, and reflect as they approach the course contents through exercises, small group discussions, case studies, and other activities [5, 12–14], all of which require students to apply what they are learning [5, 12]. Active Learning engages students in the process of learning through activities, opposed to passively listening to an expert [6, 7, 12, 14]. It often involves group work [6, 7, 14] and emphasizes higher-order thinking [12, 15]. Active Learning consists of four stages [15–17]:

Stage 1 *Trigger*: the main goal of this stage is to introduce a topic in an unexpected way yet should facilitate and promote meaningful learning. It can be achieved by raising a wide range of questions, dilemmas, or perceptions. A trigger should be complicated but relevant to the learner according to the concept. The trigger can be completed individually, in pairs, or small groups.

Stage 2 *Activity*: After the students work on the trigger stage that has been assigned to them. This stage could be done briefly or extensively but takes up a significant portion of the lesson. The duration of time dedicated to this stage is determined by the type of trigger used and the educational goals of the trigger.

Stage 3 *Discussion*: After ideas as products appeared during the activity stage are presented, students refine their understanding of concepts and ideas through problems or statements raised related to the concepts. The instructor does not judge students' positions and opinions. At the same time, classmates are encouraged to react to and express their opinions, as well as constructive criticism, in response to the various ideas or materials presented. The instructor highlights the main concepts presented by students and principles derived from these ideas.

Stage 4 *Summary*: this stage of the model contextualizes the topic and highlights the concepts that were discussed. It is handled differently than the previous three stages. First, it is significantly shorter. Second, whereas the students are the main actors in the first three stages, the instructor takes center stage in this stage. The instructor sums up and highlights key concepts, teaching ideas, conceptual frameworks, and other related topics raised and discussed in the previous three stages.

2.2 The Learning Pit Model

The Learning Pit Model is the model [10] introduced in "The Learning Challenge" and represent another interpretation on Vygotsky's Zone of Proximal Development [11]. The model helps students to become more confident with challenges, develop problem-solving strategies, and communicate their learning [8]. These are key messages for the model [9]:

1. *"Learning can be frustrating"*: progress appears elusive when acquiring new knowledge, but challenges are a natural part of the learning journey. These feelings, instead of becoming

disheartened, understand that The Learning Pit Model can help you persist, and explore alternative approaches, investing more effort and determination.

2. *"Learning should be hard"*: The Learning Pit Model's role is to design lessons that will challenge students while instructors will continue to provide encouragement, guidance, and instruction, also urges them to step beyond their comfort zone, and engage in deeper learning with the support of instructors. Overcoming initial setbacks strengthens understanding and adaptability.

3. *"Let's talk about our learning"*: The Learning Pit fosters discussions about both the content and process of learning. Reflecting on learning strategies and making necessary adjustments enhances comprehension and growth.

The Learning Pit model has four stages [11]:

stage 1 *Concept*: the concepts may be acquired from various sources such as media, conversation, observation, or curriculum. As students can grasp the concept or concepts teachers want them to investigate the upcoming challenge whether it is question, observation, or else when teachers want them to investigate, The Learning Pit Model can work.

stage 2 *Conflict*: the core principle of the Learning Pit Model is to get students "into the pit" by raising some cognitive conflict such as some observation or misunderstanding. These stimulate inquiry and reasoning yet reasonableness, indicating The Learning Pit Model is an exemplary model for educational challenges.

stage 3 *Construct*: after a while, as students navigate their way through the pit, students will begin to connect the ideas to the challenging questions or observations related yet beneficial to what they will overcome in the challenge that might involve considering alternatives, connecting ideas, and explaining cause and effect. Ultimately culminating in a eureka moment where newfound clarity is achieved. This sense of discovery is one of the reasons why passing down the pit is worth the effort.

stage 4 *Consider*: students articulate their progression of learning from simplistic ideas to the recognition of more complex and conflicting ideas, and finally to a deeper understanding of how these ideas interrelate. At this stage, they can consider the best ways to relate and apply their new understanding to various contexts.

2.3 Mathematical Resilience

The term "Mathematical Resilience" is defined as the quality by which students approach mathematics with confidence in a successful outcome to their effortful work, persistence in the face of difficulty and a willingness to discuss, reflect and research [2, 3, 19]. The construct of mathematical resilience allows learners to manage and protect themselves from unhelpful emotions that may arise when mathematics becomes difficult to learn [2].

Students with mathematical resilience will persist in the face of difficulties [3, 4, 18], will collaborate with their peers, and will have the language skills required to express their understandings [3]. Mathematically resilient students will also develop new skills as needed, as well as actively seek the assistance and support of others [4, 18]. They will understand that the more they work on mathematics, the better their chances of success [2]. When considering how to help students develop mathematical resilience, four aspects are considered [2, 4, 20]:

1. *Value*: the perception that mathematics is a valuable and useful study subject. The more students believe mathematics is valuable, the more enthusiastic they will be to learn and overcome the challenges that arise with it.

2. *Growth mindset*: the belief that one can learn mathematics and that no one is born with or without the ability to learn mathematics. This factor refers to the belief that mathematical knowledge is not fixed and will continue to improve. Students will seek out a challenge and implement strategies to overcome challenges.

3. *An understanding of how to work at mathematics*: that progress in mathematics requires struggle even for people with mastering mathematical skills, curiosity, and determination, as well as learning to manage the emotions that come with learning something new.

4. *Knowing how to recruit support*: an appreciation of the value of collaboration, the use of collaborative agency to aid in the struggle to increase mathematical knowledge, skills, and understanding.

3 Methodology

3.1 Target Group

The target group was 54 students in the specific extra large school in Bangkok in the first semester of the academic year 2024.

3.2 Instruments

The instruments used in the study consist of 2 instruments. First, the experimental instruments were the lesson plan integrated Active Learning with the Learning Pit model, which promotes the mathematical resilience of the students. Second, the instrument used to evaluate the effectiveness of the teaching plans was a teaching plan evaluation form.

3.3 Instruments Creation and Development

3.3.1 The Integration of Active Learning with The Learning Pit Model

Active Learning with the Learning Pit Model combines the principles of Active Learning and the main framework of The Learning Pit Model, By examining the actions carried out in the classroom, this model can be divided into four main stages.

Stage 1 *Trigger*: Students are presented with the essential concepts that will be stated throughout the class and which are presented in challenging yet often triggered curiosities.

Stage 2 *Activity*: This stage is consisted of 2 phases. *Concept* is the phase where students learn the main contents which are important to be used as tools that lead to the answer for the trigger. Next, *Conflict* is the phase where students are given a question related to the content presented in the concept. The question should be designed to create confusion or be related to some misconception that might happen.

Stage 3 *Discussion*: This stage consists of 2 phases. *Construct* is the phase where students form into groups and work out collaboratively on the problem that they have learned. The problems can be designed to be challenging yet solvable under the lesson the student has learned.

In the meantime, students are free to ask their teacher for guidance or clarification without fear of being judged. The teacher provides guidance and constructive advice to their work. Then, *Consider* is the phase where students answer some questions to summarize what they have learned from the previous activity. The question led to the understanding of the problem raised and what they have exchanged their opinion on.

Stage 4 *Summary*: In this stage, the teacher takes the main role of summarizing and emphasising the importance of all the contents and problems that were learned in the classroom.

3.3.2 Teaching Plans on Algebraic Vectors Integrating Active Learning with The Learning Pit Model

To develop the tools to study the intervention on students' mathematical resilience, five lesson plans implemented with Active Learning and the Learning Model were designed, the designated time for each stage is provided by table 1.

Table 1: Time limit for each lesson plan

Main stage	Time in minutes
Trigger	10
Activity: Concept and Conflict	40
Discussion: Construct and Consider	35
Summary	15

After planning the time limit, five lesson plans which integrate Active Learning with the Learning Pit Model on vectors in high school were constructed. The contents were aligned to Thai Basic Education Core Curriculum in Mathematics Learning B.E.2551 (Revised edition B.E.2560) curriculum. The lessons for each class was illustrated by table 2.

Table 2: Lessons for each lesson plan

Number of lesson plans	Lessons
1	Vectors on a plan and its basic algebraic operations
2	Magnitude of vectors and unit vectors
3	Dot product and angle between vectors
4	Cross product and area of a parallelogram
5	Some properties of products of vectors

3.3.3 Expert Evaluation

The tool used to evaluate the efficiency of the lesson plans was developed based on a review of related literature. A congruence evaluation form was constructed to assess five lesson plans on algebraic vectors, which were designed to integrate Active Learning with the Learning Pit Model to foster students' mathematical resilience. The evaluation instrument employed a five-point Likert scale and consisted of 15 items across five aspects, namely the alignment with

learning objective, contents arrangement, learning activities, utilizing the appropriate media, and assessment and evaluation method.

3.4 Data Collection and Analysis

To assess the effectiveness of the developed lesson plans, experts feedback were obtained and a pilot administration of the scale were conducted. The lesson plans, designed with Active Learning and the Learning Pit Model, were continuously reviewed and refined based on constructive comments from experts. The expert group consisted of two individuals, which are a retired government official who specialized in educational assessment and evaluation and an experienced secondary school teacher with over ten years of teaching expertise. Their feedback was qualitatively analyzed to ensure that the lesson plans were suitable to foster mathematical resilience. In addition, the lesson plans receive score ranging from 4.50 to 5.00 out of 5.00 on the congruence evaluation form, indicating a very good level of quality.

Since lesson plans integrating Active Learning with the Learning Pit Model have not been previously explored, this study presents an innovative approach aimed at improving students' mathematical resilience. The lesson plans were continuously criticized, verified, and evaluated and refined until the lesson plans were suitable for pilot testing.

4 Findings from Experts Review

By breaking down mathematical resilience aspects, the lesson plans integrated with Active Learning with the Learning Pit Model can be constructed by each main stage as follows.

Stage 1 *Trigger*: various kinds of implementation can be used to this stage such as visualization, hands-on materials, demonstrations, guess-and-check strategies, and experiments with theorems. The trigger is to emphasize the things that need to be focused on. Mostly, these approaches lead to some crucial questions raised in the class afterwards that set the theme for the lessons in that class. Since these are triggering students to see the importance of the upcoming contents, bringing to the expected result that students would improve the value aspect of mathematical resilience.

Stage 2 *Activity*: the *Concept* phase introduces relevant lessons in a suitable portion yet sufficient to both connect and understand the Trigger clearly and be ready to deal with "the pit" in the next stage. Then, the *Conflict* performs mostly on the demonstrations which initiate doubts. Responding to the doubt proposed, a single question will be raised immediately. The question would be done neatly whether it is multiple choices or short-answer question. It is crucial that the question must be created correspond to the concepts and provide the sufficient information that the concept might solve that question. Student will answer the question once and talk to the whole class about the solution, reason, or reasonableness by not hesitating about the answer must be correct. This stage highlights the importance of growth and value in developing mathematical resilience.

Stage 3 *Discussion*: in the *Construct* phase, students engage in small group discussion, working on tasks that are mostly arranged into the set of questions and some observations that are plenty in amount and the use of concepts. In the meantime, the teacher would generously support the student in their discussion and not judge the student's opinion. After a period of time, the *Consider* phase shifts the focus to a whole-class discussion, allowing students to reflect on any observations that may lead to interesting theories, building upon what they have learned in previous stages and phases. An understanding of how to work at mathematics would be firstly the expected aspect of mathematical resilience as their improvement following with growth and knowing how to recruit support.

Stage 4 *Summary*: teacher revisits the question posed in the conflict phases, then students answer that question again to compare their understandings before and after the discussion, this brings to their belief that they can progressively learn mathematics. After that, the teacher plays a crucial role in summarizing the knowledge gained throughout the lessons, presenting it concisely yet effectively to ensure a clear understanding. The expected aspect of mathematical resilience is value.

Table 3 indicates an example of the lesson plan which integrates Active Learning with the Learning Pit Model on algebraic vectors in dot product and angle between vectors.

5 Discussion

To develop the lesson plans implemented with Active Learning with the Learning Pit Model that affects students' mathematical resilience, there are several crucial parts that should be discussed.

5.1 Designing The Lesson Plan that Aligned to The Integration of Active Learning with the Learning Pit Model and Its Impact to Fostering Mathematical Resilience

To design the lesson plans integrating Active Learning with the Learning Pit Model to foster students' mathematical resilience, there are concerning viewpoints along with the thematic points in the literature that support the viewpoints. The following viewpoints were provided for each stage.

Stage 1 *Trigger*: Students are at the pit's entry when they are confused at the start of class. The initial confusion is effectively necessary for deeper understanding, which corresponds to Chi and Wylie's ICAP framework, which emphasizes that engagement that is constructive and interactive leads to better outcomes [22]. Since it sets the stage for students to perceive mathematics as an object that is worth exploring, it aligns with the value aspect of mathematical resilience. Furthermore, students' curiosity brings a desire to understand the problem and yet resolve the confusion initially raised, which is aligned with Heiko and Tanya [21].

Stage 2 *Activity*: In the *Concept* phase, as students are encouraged to explore solutions to the conflict, gathering knowledge in suitable portions and clearly explaining the concepts as the necessary tools to climb out of the pit supports the growth mindset aspect of mathematical resilience, as found by Joseph and Emily [23]. Importantly, the phase should address the confusion

Table 3: An example of the lesson plan integrating with Active Learning with the Learning Pit Model on algebraic vectors in dot product and angle between vectors

stage	phase	description
Trigger	-	Students observe a demonstration where two vectors as column matrices are multiplied together (for example, the multiplication of two vectors in two dimensional vector space $\begin{bmatrix}7\\8\end{bmatrix}\begin{bmatrix}5\\-2\end{bmatrix}$). Initially, students notice that the multiplication is undefined due to the inner dimensions. This leads to the introduction of new multiplication operators, focusing on the dot product.
Activity	Concept	The definition of the dot product is introduced to clarify how vectors can be multiplied together with working on examples. Then, the theorem related to dot product, the angle between two vectors is presented together with some examples, providing the foundation for deeper exploration.
	Conflict	The teacher assigns two vectors and allows students to use a web application to find the angle between these vectors. After a few seconds, the teacher pretends to "guess" the correct angle type (acute, perpendicular, or obtuse) without visible calculations. This creates cognitive conflict, as students wonder how the teacher quickly arrived at the answer. To resolve this, students observe different vector pairs and first record their observations about the dot product values.
Discussion	Construct	Students are grouped into teams of six and assigned three pairs of vectors. These pairs are designed so that the dot product values are positive, zero, and negative. Each group calculates the angle between vectors.
	Consider	Students write the results on the board, including dot product value, calculated angle, and angle type. A class-wide discussion takes place where students observe some relation between dot product value and type of angle. At this time, teacher helps students clear the air by guiding the related theorem ,$\vec{u} \cdot \vec{v} = \left\|\vec{u}\right\| \left\|\vec{v}\right\| \cos(\theta)$, leads to the conjecture about it relation, teacher now guiding and verify the correctness and show students how it relation works
Summary	-	The teacher summarizes key concepts and observations in a structured manner. Student understanding is assessed through both multiple-choice question and written question. This ensures that students can apply their knowledge effectively.

raised in the Trigger stage, allowing students to realize the applicability of the mathematical concept, to do so enhances the value aspect of mathematical resilience as students perceive mathematics as worth engaging, which is aligned with Joseph and Emily, and Hazzan and Zazkis [23,24]. Next, the *Conflict* phase also introduces more challenging cognitive conflicts, which students might be confused about how the conflicts were resolved and cannot be immediately solved. This also reflects an understanding of how to work with mathematics aspect of mathematical resilience, seemingly the value is less focused than expected. Moreover, to confront challenges head-on and persist through confusion can improve students' growth mindset aspect, which is consistent with Hazzan and Zazkis [24].

Stage 3 *Discussion*: In the *Construct* phase, which referred to students being in the pit. Working together helps students share strategies, clarify some misunderstandings, and also build confidence in their ability to solve problems. This also builds students growth mindset aspect. Moreover, since students struggle in the pit, knowing how to recruit support is emphasized. Collaboration plays a role in overcoming challenges, which is emphasized by Abdu and Schwarz's findings. [25]. In the *Consider* phase, it reinforces the idea that progress comes through effort and collaboration and then shares it with their class, allowing students to recognize their progress. Students build an understanding. This stage also aligns with the growth mindset aspect, as students realize that their initial confusion does not define their abilities, but instead, persistence through challenges leads to improvement, as highlighted in studies by [26]. In addition, as students were engaged in challenging tasks, these promoted meaningful learning and the development of a growth mindset, as the alignment of National Council of Teachers of Mathematics (NCTM) [27].

Stage 4 *Summary*: This stage solidifies the value aspects of mathematical resilience through their belief that what they have learned is what they would use, which corresponds to [28]. Also, by reflecting on the process students have made, students can overcome challenges through their effort, which reinforces the belief that learning mathematics involves struggle. This also leads to the improvement of growth mindset, which also corresponds to [28], which is not proposed to the construction of the lesson plans.

5.2 Some Challenges of the Integration of Active Learning with The Learning Pit Models

Since the integration of Active Learning with the Learning Pit Model was intended to foster mathematical resilience, some challenges must be addressed for effective implementation.

First, the time limit according to the curriculum may limit the implementation potential since Active Learning with the Learning Pit Model requires students to experience cognitive struggle while students are in the pit, which involves extensive periods of time to take care of students' discussion and support. Students' deep understanding through the Learning Pit Model often exceeds the duration provided in the traditional teaching plans or curriculum, which is aligned with Nottingham and Larsson [29].

Second, the implementation depends on the characteristics and competencies of the teacher. According to Active Learning, teachers must be capable of planning on choosing a trigger, de-

signing learning activities, guiding, and supporting students. These tasks require pedagogical expertise and subject knowledge, which corresponds to Andrews, Auerbach, and Grant [30].

Third, demographic factors of students, particularly class size and student diversity, may influence the impact of Active Learning with the Learning Pit Model. It can be difficult to provide support as students navigate their way in the pit, especially in the large classroom or classrooms with varying learner needs. Also, a student's background plays a role in determining how effectively resilience is fostered, noted by Russo and Sullivan et al. [31]

6 Suggestion

The experts' feedback focused on optimizing each stage of the lesson plan, from effectively triggering curiosity to structuring conflicts and facilitating productive discussions. Experts suggest that an effective trigger should be directly connected to what students will learn within the class, ensuring there is no long gap that mixes unrelated content. The concepts introduced should serve as effective tools for solving the upcoming construct phase. In designing the conflict, it should create frustration but remain solvable, allowing students enough time to analyze before receiving kind guidance. A well-designed conflict should act as a "bomb" that confuses students, such as by presenting statements related to common misconceptions or key content that requires careful application. Encouraging discussions about reasonableness while students work through the conflict helps them anticipate and avoid future mistakes, fostering a safety net for their understanding. During the construction phase, the metaphor of exploring an unknown path applies, before venturing into unfamiliar territory, students must first prepare their "tools" or knowledge. Teachers must carefully assess how deep the "pit" is, ensuring that the questions are appropriately challenging yet manageable, guiding students toward discovering the best way to climb out of the learning pit.

Acknowledgement. The author is grateful to the experts who spend their time providing constructive comments and suggestions to enhance the lesson plan, as well as to the advisor who offers continuous feedback.

References

[1] J. Goodall, S. Johnston-Wilder, Overcoming mathematical helplessness and developing mathematical resilience in parents: An illustrative case study, in: Proceedings, pp. 526–529, 2015.

[2] C. Lee, S. Johnston-Wilder, Understanding emotions in mathematical thinking and learning, in: Proceedings, pp. 269–291, 2017.

[3] S. Johnston-Wilder, C. Lee, Developing mathematical resilience, in: BERA Annual Conference 2010, 1–4 Sep 2010, University of Warwick.

[4] The Centre for Research in Education and Educational Technology (CREET), Developing mathematical resilience, The Open University.

[5] M. Takele, Practices and challenges of active learning methods in mathematics classes of upper primary schools, Int. J. Educ. Learn. 11 (13) (2020) 26–29.

[6] E. Lugosi, G. Uribe, Active learning strategies with positive effects on students' achievements in undergraduate mathematics education, Int. J. Math. Educ. Sci. Technol. 51 (2020).

[7] D.C. Owens, et al., Student motivation from and resistance to active learning with the learning pit model rooted in essential science practices.

[8] J. Russo, et al., Primary teacher attitudes towards productive struggle in mathematics in remote learning versus classroom-based settings, Educ. Sci. 2021.

[9] A short guide to the Learning Pit [Online]. Available: `https://www.learningpit.org/guide/`.

[10] The Learning Pit [Online]. Available: `https://www.learningpit.org/guide/`.

[11] J. Nottingham, The learning challenge: A quick guide, in: An Introduction to the Learning Pit by Its Creator, James Nottingham.

[12] I. Vale, A. Barbosa, Active learning strategies for an effective mathematics teaching and learning, Eur. J. Sci. Math. Educ. 11 (3) (2023) 574.

[13] A. Alkilany, The impact of the use of active learning strategies in the development of mathematical thinking among students and the trend towards mathematics, J. Educ. Pract. 8 (36) (2017) 13.

[14] Z. Ješková, Active learning in STEM education with regard to the development of inquiry skills, Educ. Sci. 12 (2022) 686.

[15] O. Hazzan, et al., Active-learning-based teaching model, in: Guide to Teaching Computer Science, Springer, pp. 17–18.

[16] P. Muangkheaw, Development of economics problem-solving abilities of Matthayomsuka 6 students taught by using active learning process with KWL plus technique, p. 32, 2018.

[17] A. Pongsak, Peer coaching to promote the learning management competency of Thai language teachers based on active learning theory to develop reading comprehension skills of primary students, p. 64, 2019.

[18] S. Johnston-Wilder, J. Brindley, P. Dent, A survey of mathematics anxiety and mathematical resilience among existing apprentices, London: Gatsby Charitable Foundation, p. 5, 2014.

[19] F. Duah, Mathematics resilience: What is known in the pre-tertiary mathematics education research and what we have found researching non-mathematics specialists, Proceedings of the British Society for Research into Learning Mathematics 37 (2017) 1–2.

[20] S. Johnston-Wilder, Measuring mathematical resilience: An application of the construct of resilience to the study of mathematics, in: AERA 2013, San Francisco, California, 2013.

[21] H. Dietrich, T. Evans, Traditional lectures versus active learning – A false dichotomy?, STEM Educ. 2 (4) (2022) 275–292.

[22] M.T.H. Chi, R. Wylie, The ICAP framework: Linking cognitive engagement to active learning outcomes, Educ. Psychol. 49 (4) (2014) 219–243.

[23] J. DiNapoli, E.K. Miller, Recognizing, supporting, and improving student perseverance in mathematical problem-solving: The role of conceptual thinking scaffolds, J. Math. Behav. 66 (2022) 100965.

[24] O. Hazzan, R. Zazkis, Reducing abstraction: The case of school mathematics, Educ. Stud. Math. 58 (1) (2005) 101–119.

[25] R. Abdu, B. Schwarz, Split up, but stay together: Collaboration and cooperation in mathematical problem solving, Instr. Sci. 48 (2020) 313–336.

[26] J. Dixon, T. Mahoney, Productive struggle in mathematics: Utilising challenging tasks to establish richer mathematical learning experiences, 29 (2024) 10–15.

[27] National Council of Teachers of Mathematics, Principles to actions: Ensuring mathematical success for all, Reston, VA: Author, 2014.

[28] A.J.B. Hutauruk, et al., Achievement of students mathematical resilience through problem based learning model with metacognitive approach, J. Phys.: Conf. Ser. 1315 (2019) 012051.

[29] J. Nottingham, M. Larsson, The Learning Challenge: How to Guide Your Students Through the Learning Pit to Achieve Deeper Understanding, Corwin Press, 2019.

[30] T.C. Andrews, A.J.J. Auerbach, E.F. Grant, Exploring the relationship between teacher knowledge and active-learning implementation in large college biology courses, CBE—Life Sci. Educ. 18 (4) (2019) ar52. https://doi.org/10.1187/cbe.19-01-0010.

[31] R. Russo, P. Sullivan, et al., Teachers' views on productive struggle in mathematics: A cross-cultural analysis, Teach. Teach. Educ. 114 (2022) 103689.

Developing an Understanding of "a half" in 3- to 4-Year-Old Children Lacking Cardinality Knowledge Through Play-Based Learning Activities

Sinchai Jansem[1,*] and Amika Wachirawin[1,†]

[1]Elementary Program, Faculty of Education, Suan Dusit University,
Bangkok 10300, Thailand

Abstract

Children who have not yet grasped the concept of cardinality often encounter difficulties in determining the precise number of objects within a set. Nevertheless, research indicates that they may still be capable of recognizing the equivalence between two sets. Unlike numerical counting, the notion of "a half" is intrinsically linked to the concept of equivalent sets. Based on this premise, it was hypothesized that children who are unable to ascertain the number of objects in a set might still comprehend the concept of 'a half.' To investigate this hypothesis, a teaching experiment was conducted utilizing a play-based learning approach to examine how young children lacking an understanding of cardinality develop the concept of 'a half.' The study involved four participants: two three-year-olds and two four-year-olds. The learning activities, grounded in a play-based learning approach, were designed to enhance children's understanding of "a half" through two primary tasks: 'Sharing a half' and 'Taking a half.' Initially, none of the children were able to successfully perform either task. However, following engagement in the activities, they were able to accurately execute both 'Sharing a half' and 'Taking a half.' These findings suggest that even in the absence of cardinality comprehension, young children can develop an understanding of 'a half.'

Keywords: cardinality, fraction value concept, numeracy development
2020 MSC: 97F20

*Speaker (sinchai_jan@dusit.ac.th)
†Corresponding author (amika_wac@dusit.ac.th)

1 Introduction

The concept of "a half" represents a fraction, a mathematical notion that often poses challenges for young children, and even adults, to comprehend. Despite this complexity, terms such as "half" or "quarter" are frequently introduced to young children. When children encounter the term "half," it is uncertain whether they grasp that it signifies "one out of two equal parts." The understanding of fractions is not inherently intuitive for children, as this concept does not naturally emerge from the process of counting. According to Gabriel et al. [1], the difficulties children face with fractions arise from the necessity to integrate conceptual and procedural knowledge, which are typically taught separately from the development of counting abilities.

Counting plays a crucial role in the development of mathematical concepts among preschool-aged children, serving as a fundamental strategy for enhancing numeracy skills and establishing a foundation for more advanced mathematical comprehension [2]. In the early stages of counting, children often recite number words without fully grasping their quantitative significance, indicating an undeveloped concept of cardinality. However, this does not suggest an absence of numeracy concepts. Research indicates that some children begin to develop numerical understanding even before language acquisition. For instance, infants can discern differences in the cardinal values of objects within a set. Studies demonstrate that children aged 2 to 3 years exhibit some understanding of numerical concepts before they can count accurately. Silver et al. [3] propose an emerging ability to identify relative quantities, even if precise counting is not yet possible. Furthermore, several studies have shown that six-month-old infants can distinguish between arrays with a 1:2 ratio but encounter difficulty with a 2:3 ratio [4-6]. Additionally, research has shown that five-month-old infants can perform simple addition and subtraction operations on small sets of objects [7].

For preschool levels, teachers aim to develop children's numeracy skills in two aspects: cardinality and ordinal relations. Both cardinality and ordinal relations contribute to children's understanding of numerical magnitude, but in complementary ways. Cardinality helps children understand the absolute quantity represented by a number, while ordinal relations allow them to compare and order numbers. Together, these concepts form a crucial foundation for children's numerical development and later mathematical skills [8-10]. In addition to these foundational skills, young children also begin to develop an understanding of numerical equivalence—the concept that two sets contain the same number of objects—even before they become proficient counters [11,12]. Some children, despite not yet being able to count, can recognize when two sets are equal in quantity. For instance, if candies are arranged in an array as depicted in the diagram below:



Most preschoolers recognize that the sets have the same number of candies, based on the one-to-one alignment of the individual pieces [13].

Mathematics educators widely acknowledge that the development of early numeracy skills and concepts at the preschool level significantly influences subsequent mathematics achievement. Empirical research has demonstrated a strong correlation between early numeracy skills and long-term success in mathematics [14]. The critical importance of early numeracy is further underscored by studies indicating that challenges in early numeracy can impede the acquisition of mathematical skills in later childhood. This understanding has prompted the creation of early mathematics interventions and assessment tools aimed at supporting the numeracy development of young children [15]. Fractions are widely recognized as one of the most challenging topics in primary school mathematics. Multiple studies highlight the difficulties students face in understanding and working with fractions, particularly those with mathematics difficulties [16,17].

To enhance preschoolers' comprehension of fraction values, educators should integrate activities that reinforce both fraction and whole number concepts. Contemporary pedagogical methods, such as play-based learning, game-based strategies, and the integration of technology, have proven effective in introducing early numeracy skills. These methodologies can be adapted to introduce fundamental fraction concepts in a manner appropriate for young children, thereby building on their foundational numeracy skills. For example, Cohrssen and Niklas [18] examine the implementation of play-based mathematics games in preschool settings. The NT Preschool Math Games have demonstrated efficacy in improving children's mathematical abilities, including number recognition, counting, and problem-solving. Although these games do not specifically target fractions, similar game-based approaches could be modified to introduce basic fraction concepts. Furthermore, hands-on activities that emphasize the magnitude of fractions can significantly enhance children's conceptual understanding [19].

Given that the notion of “a half” is rooted in the concept of two equivalent sets, which young children begin to develop at an early age, the study was constructed based on the following conditional statement: “If children do not understand cardinality but grasp the concept of equivalence, then they will be able to understand the concept of a half.” From this foundational idea, learning activities were designed to instruct children who lack an understanding of cardinality and to examine how they develop the concept of “a half.”

The learning activities developed were based on play-based learning methodologies, which underscore the significance of play in the cognitive and developmental processes of children. This pedagogical approach acknowledges the intrinsic link between play and learning, particularly in early childhood [20]. Within the framework of play-based learning, children participate in activities that are simultaneously enjoyable and educational, facilitating exploration, experimentation, and the discovery of new concepts in an engaging and natural manner. This approach can integrate various strategies, including serious games, outdoor activities, and creative arts, to instruct disciplines such as science, technology, engineering, and mathematics (STEAM) [21, 22]. To effectively implement play-based learning, educators may employ several key strategies: 1) cultivating a nurturing and responsive educator-child relationship, 2) promoting social-emotional learning through daily interactions and practices, 3) leveraging the physical environment to stimulate learning, and 4) collaborating with caregivers [23].

This study posited that young children might develop an understanding of the concept of "a half" prior to fully comprehending the notion of cardinality. The research specifically aimed to investigate how children, who have not yet achieved mastery of cardinality, could acquire the concept of "a half" through engagement in learning activities grounded in a play-based learning approach. Figure 1 below illustrates the research conceptual framework.

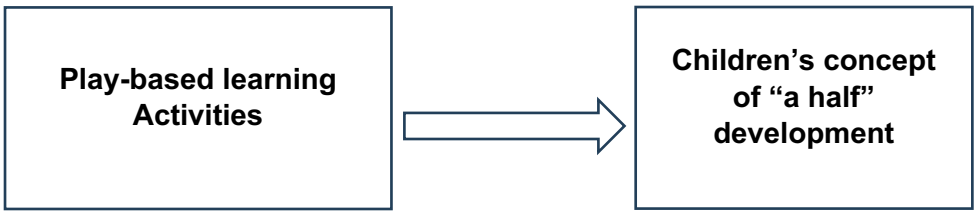


Figure 1: Research Framework

2 Methods

The objective of this study was to examine how children lacking an understanding of cardinality comprehend the concept of "a half." A case study was conducted by gathering data from the children of the researcher's colleagues, employing qualitative methods for data analysis.

2.1 Participants

The study involved four participants, two of whom were three years old and the other two were four years old. All participants were children of either the authors' colleagues, who

are faculty members at the Faculty of Education, Suan Dusit University, or other individuals affiliated with the university. The parents of the children also collaborated with the authors, helping during the children's participation in the activities.

The four participants were selected from an initial group of seven children. Initially, all seven children were interviewed to assess their understanding of cardinality and their counting skills. While all participants were capable of counting, only four demonstrated a lack of understanding of the concept of cardinality. Consequently, these four children were chosen to participate in the study.

2.2 Procedure

To examine how young children who did not understand cardinality develop concept of "a half", we developed an assessment protocol for examining these children's concept of a half. The protocol comprised of 3 parts: 1) Equivalence, 2) Sharing a half and 3) Taking a half. Figure 2 (on the next page) depicts the protocol.

The items in part 1, Equivalence, were designed for determining children's concept equivalence of 2 equal sets. The part 2, Sharing a half, and part 3, Taking a half, were for determining children conceptual of "a half." Each part is composed of 8 items. The assessment protocol was used in both pre- and post-assessments.

Participants of the same age were paired together. Each pair participated in eight sessions. The first session (15 minutes) was dedicated to a pre-assessment, while sessions two through seven focused on teaching, with each session lasting approximately 25 minutes. The final session (15 minutes) was used for a post-assessment. One of the authors, a kindergarten teacher, led the activities and worked directly with the children. The other author and the parents observed and took notes. Occasionally, parents assisted by clarifying questions for their children.

To facilitate children's understanding of the concept of 'a half,' learning activities were designed utilizing a play-based learning approach. In these activities, both teachers and parents guided the children through various games and discussions, using questions specifically developed by the authors. The primary objectives of these learning activities were twofold: 1) To enhance the children's ability to apply one-to-one correspondence, thereby enabling them to divide objects, such as candies, into two equal groups, and 2) To cultivate an understanding of the concept of "a half," encompassing the notions of "sharing a half" and "taking a half." Figure 3 illustrates the objectives and strategies employed to support the children's comprehension of this concept.

ASSESSMENT PROTOCOL

Part 1. EQUAVALECE

1. The teacher places 4 candies by arranging into 2 columns.



Then asks, “Are the two groups have same amount of candies?”

2. The teacher places 5 candies by arranging in 2 columns.



Then asks, “Are the two groups have same amount of candies?”

3. The teacher places 6 candies by arranging in 2 columns.



Then asks, “Are the two groups have same amount of candies?”

4. The teacher places 6 candies by arranging in 2 columns.



Then asks, “Are the two groups have same amount of candies?”

5-8. The teacher places numbers of candies on the table (4, 5, 6 and 7) and asks; “Can you divide these candies into 2 equal piles?”

Part 2. SHARING A HALF.

1. The teacher places 4 candies by arranging into 2 columns.



Then asks the both children to share a half of candies.

2. The teacher places 5 candies by arranging in 2 columns.



Then asks the both children to share a half of candies.

3. The teacher places 6 candies by arranging in 2 columns.



Then asks the both children to share a half of candies.

4. The teacher places 6 candies by arranging in 2 columns.



Then asks the each kid to share a half of candies.

5-8. Repeat 1 – 4 but place candies randomly.

Part 3. TAKING A HALF

1. The teacher places 4 candies by arranging in 2 columns.



Then asks one of them to take a half of candies.

2. The teacher places 5 candies by arranging in 2 columns.



Then asks one of them to take a half of candies.

3. The teacher places 6 candies by arranging in 2 columns.



Then ask one of them to take a half of candies.

4. The teacher places 6 candies by arranging in 2 columns.



Then asks one of them to take a half of candies.

5-8. Repeat 1 – 4 but place candies randomly.

Figure 2: Assessment Protocol

Teaching Objective	Strategies
1. Children are able to divide the candies into two groups	1. One-to-one dealing
2. Children are able to share a half of candies.	1. Divide the candies into two groups. [If this process does not succeed (odd number), sharing a half cannot be done.] 2. Take all candies from one group for yourself and give another group to the other.
3. Children are able to take a half of candies	1. Divide the candies into two groups. [If this process does not succeed (odd number), sharing a half cannot be done.] 2. Take all candies from one group for yourself and leave the other group stayed.

Figure 3: Teaching Strategy

2.3 Data analysis

Children's responses to the tasks in the pre- and post-assessment were documented through field notes and video recordings. To analyze the data, the authors examined both verbal and nonverbal communication expressed by the children. Each child's response to each item was evaluated using a binary scoring system, where a score of "1" was assigned for correct responses and "0" for incorrect responses. Initially, both authors conducted independent analyses of the data before comparing their findings. In cases where discrepancies arose, discussions were held until a consensus was reached, ensuring the reliability and consistency of the analysis.

3 Results

Tables 1 and 2 present the assigned scores for each child. Specifically, Table 1 displays the scores for each item from pre- and post-assessments, while Table 2 provides the total scores for each child across both assessments. The children labeled as C1 and C2 in the tables were three years old, whereas C3 and C4 were four years old.

Table 1: The score of each item for pre- and post-assessments

Children	Assessment	EQUIVALENCE								SHARING A HALF								TAKING A HALF							
		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
C1	pre	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	post	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
C2	pre	1	1	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	post	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1	0	1	1	1	1
C3	pre	1	1	1	1	1	0	0	0	1	1	1	1	0	0	0	0	1	0	1	0	0	0	0	0
	post	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C4	pre	1	1	1	1	1	0	1	0	1	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0
	post	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 2: The total score for pre- and post-assessments

Children	Equivalence		Sharing a half		Taking a half	
	Pre (8)	Post (8)	Pre (8)	Post (8)	Pre (8)	Post (8)
C1	3	7	0	4	0	4
C2	5	8	1	4	0	6
C3	5	8	4	8	2	8
C4	6	8	2	8	3	8

Equivalence

The results from the pre-assessment revealed that all the children were able to recognize two equal sets, but only when the objects were arranged in two columns. They demonstrated an inability to divide the candies into two equal groups independently. In the post-assessment, all the children successfully recognized two equal sets when the objects were arranged in two columns. Regarding the task of dividing the candies into two equal groups, all four-year-old children performed correctly. However, one of the three-year-old children (C1) made an error on item 8, where the total number of candies was seven.

Sharing a half

The pre-assessment scores of the three-year-old children indicated that they were initially unable to perform the task of sharing a half in any given situation. However, after participating in the activities, they demonstrated an understanding of one-to-one correspondence and grasped the concept of a half. In the post-assessment, both children successfully performed the task of sharing a half, yet they received scores of 4/8 each. The errors primarily occurred in situations where the teachers presented an odd number of candies, such as seven candies (item 4). In this case, child C1 distributed three candies to herself and

three to the other child. When asked by the teacher, “Did you share half of the candies?” the child responded, “Yes, because we both get the same.”

For the four-year-old children, the pre-assessment scores were 2 and 4. These children were able to correctly share a half only when the candies were pre-arranged into two equal columns by the teacher but struggled with items 6–8. However, in the post-assessment, both children correctly performed the task of sharing a half. Child C2 initially struggled with items 4 and 8 but was able to provide the correct responses after the teacher explained the concept of dividing an odd number into two equal groups. The post-assessment scores of 8/8 demonstrated that these children had fully comprehended the process of “sharing a half” of the given candies.

Taking a half

The pre-assessment results indicated that the three-year-old children did not initially understand how to take half of the candies. However, in the post-assessment, both children successfully performed the task when the candies were presented in even numbers (items 1, 3, 5, and 7). For items 2 and 4, which involved odd-numbered candies arranged in two columns, both children selected candies from the column with the larger number. Notably, child C2 correctly performed the task for items 6 and 8.

For the four-year-old children, the pre-assessment scores were 2 and 3, demonstrating that they were only able to correctly share half of the candies when the teacher pre-arranged them into two equal columns. However, the post-assessment scores of 8/8 indicated that they had developed a comprehensive understanding of how to take half of the candies in all presented situations.

4 Conclusions and Discussion

Given that some children develop an understanding of equivalence prior to grasping cardinality and considering that the concept of "a half" is grounded in the notion of equivalence, it was hypothesized that children lacking an understanding of cardinality might still comprehend the concept of "a half." Consequently, this study aimed to investigate how young children who do not understand cardinality develop the concept of "a half."

The findings indicated that children were able to comprehend the concept of "a half" after engaging in the activities of "Sharing a half" and "Taking a half." The pre-assessment scores demonstrated that initially, none of the children understood the meaning of "a half." However, the children developed an understanding of the concept of “a half” after participating in activities with the teacher. Notably, even though these children did not grasp the concept of cardinality, and the teacher did not attempt to explain or facilitate their understanding of cardinality, the children demonstrated an understanding of "a half" without possessing an understanding of cardinality. This evidence is consistent with other research

suggesting that the development of fraction concepts is not primarily based on counting ability. Instead, understanding fractions may be facilitated by a ratio-processing system (RPS) that is attuned to the magnitudes of nonsymbolic ratios [24-26]. Findings from developmental psychology and neuroscience further corroborate the existence of an RPS capable of processing nonsymbolic ratio magnitudes, potentially providing a foundational basis for fraction comprehension [27].

In the post-assessment of the "Sharing a Half" tasks with three-year-old children, it was noted that the children encountered difficulties when tasked with dividing a set of candies in half, especially when the total number of candies was odd. In these instances, they consistently set aside one candy and proceeded to evenly distribute the remaining candies. The teacher sought clarification from the children, and one child articulated, "If we did not leave one candy on the table, we cannot have the candies equally." This behavior indicates that the children were attempting to achieve an equitable distribution of the candies. Therefore, it can be inferred that the three-year-old children exhibited an understanding of the concept of "Sharing a Half."

In the development of children's numeracy concepts, there is often an emphasis on whole numbers, including counting numbers and zero. However, recent research by Karjcsi et al. [28] suggests that preschoolers are capable of understanding more complex mathematical concepts, such as empty sets. This finding highlights the importance of introducing mathematical concepts, such as fractions, at the preschool level. Although fractions are traditionally perceived as challenging for children, evidence indicates that a well-structured curriculum can facilitate the learning of fractions [29]. By engaging children in activities that introduce fraction concepts early, they can develop an understanding of values such as "a half" or "a quarter." Establishing these foundational concepts in preschool can enhance children's comprehension of fractions, thereby aiding their understanding in primary school. The primary challenge for educators is to develop a curriculum that effectively enables young children to learn fractions.

References

[1] G. Gabriel, F. Coché, D. Szucs, V. Carette, B. Rey, A. Content, A componential view of children's difficulties in learning fractions, Front. Psychol. 4 (2013) 715. https://doi.org/10.3389/fpsyg.2013.00715.

[2] P. Aunio, P. Räsänen, Core numerical skills for learning mathematics in children aged five to eight years – a working model for educators, Eur. Early Child. Educ. Res. J. 24 (5) (2015) 684–704. https://doi.org/10.1080/1350293x.2014.996424.

[3] A.M. Silver, E. Votruba-Drzal, N. Cabrera, E.J. Braham, C.S. Tamis-Lemonda, H.J. Bachman, L. Elliott, M.E. Libertus, Measuring Emerging Number Knowledge in Toddlers, Front. Psychol. 12 (2021). https://doi.org/10.3389/fpsyg.2021.703598.

[4] F. Xu, S. Goddard, E.S. Spelke, Number sense in human infants, Dev. Sci. 8 (1) (2004) 88–101. https://doi.org/10.1111/j.1467-7687.2005.00395.x.

[5] K. Vanmarle, K. Wynn, Six-month-old infants use analog magnitudes to represent duration, Dev. Sci. 9 (5) (2006) F41–F49. https://doi.org/10.1111/j.1467-7687.2006.00508.x.

[6] J.S. Lipton, E.S. Spelke, Origins of number sense. Large-number discrimination in human infants, Psychol. Sci. 14 (5) (2003) 396–401. https://doi.org/10.1111/1467-9280.01453.

[7] G. Decarli, L. Veggiotti, M.D. de Hevia, The link between number and action in human infants, Sci. Rep. 12 (2022) 3371. https://doi.org/10.1038/s41598-022-07389-9.

[8] N.R. Scalise, G.B. Ramani, Symbolic Magnitude Understanding Predicts Preschoolers' Later Addition Skills, J. Cogn. Dev. 22 (2) (2021) 185–202. https://doi.org/10.1080/15248372.2021.1888732.

[9] K. Vanmarle, Y. Mou, D.C. Geary, F.W. Chu, J. Rouder, J.H. Seok, Attaching meaning to the number words: contributions of the object tracking and approximate number systems, Dev. Sci. 21 (1) (2016). https://doi.org/10.1111/desc.12495.

[10] C. Xu, E.A. Maloney, H.P. Osana, V. Simms, J. Wylie, S.-L. Skwarchuk, S. Di Lonardo Burr, J.-A. Lefevre, A direct comparison of two measures of ordinal knowledge among 8-year-olds, J. Numer. Cogn. 9 (2) (2023) 253–267. https://doi.org/10.5964/jnc.10201.

[11] K. Mix, Similarity and numerical equivalence appearances count, Cogn. Dev. 14 (2) (1999) 269–297.

[12] K. Mix, J. Huttenlocher, S.C. Levine, Do preschool children recognize auditory-visual numerical correspondences?, Child Dev. 67 (4) (1996) 1592–1608.

[13] J. Piaget, The child's conception of number, Norton, New York, 1965 (Original work published 1941).

[14] K. Missall, R.L. Hojnoski, G.I.L. Caskie, P. Repasky, Home Numeracy Environments of Preschoolers: Examining Relations Among Mathematical Activities, Parent Mathematical Beliefs, and Early Mathematical Skills, Early Educ. Dev. 26 (3) (2014) 356–376. https://doi.org/10.1080/10409289.2015.968243.

[15] D. Purpura, A. Napoli, Early numeracy and literacy: Untangling the relation between specific components, Math. Think. Learn. 17 (2-3) (2015) 197–218.

[16] D.J. Hacker, J.R. Levin, S.A. Kiuhara, A metacognitive intervention for teaching fractions to students with or at-risk for learning disabilities in mathematics, ZDM 51 (4) (2019) 601–612. https://doi.org/10.1007/s11858-019-01040-0.

[17] J. Tian, R.S. Siegler, Fractions Learning in Children With Mathematics Difficulties, J. Learn. Disabil. 50 (6) (2016) 614–620. https://doi.org/10.1177/0022219416662032.

[18] C. Cohrssen, F. Niklas, Using mathematics games in preschool settings to support the development of children's numeracy skills, Int. J. Early Years Educ. 27 (3) (2019) 322–339. https://doi.org/10.1080/09669760.2019.1629882.

[19] F. Gabriel, B. Szucs, C. Content, R. C. Palmers, A. D. Butterworth, Developing Children's Understanding of Fractions: An Intervention Study, *Mind Brain Educ.* 6 (2) (2012) 137–146.

[20] M. Theobald, D. Jones, J. Bourne, J. Einarsdóttir, S. Ross, H. Knaggs, S. Danby, C. Carter-Jones, Children's Perspectives of Play and Learning for Educational Practice, Educ. Sci. 5 (4) (2015) 345–362. https://doi.org/10.3390/educsci5040345.

[21] R. Lamrani, E. Abdelwahed, Game-based learning and gamification to improve skills in early years education, Comput. Sci. Inf. Syst. 17 (1) (2019) 339–356. https://doi.org/10.2298/csis190511043l.

[22] E. Skaraki, Creating a watermill through steam activities for preschool children in the school yard, Adv. Mob. Learn. Educ. Res. 3 (2) (2023) 748–753. https://doi.org/10.25082/amler.2023.02.001.

[23] C. Blewitt, A. O'Connor, A. Mousa, R. Green, A. Ifanti, H. Morris, H. Skouteris, K. Jackson, A. Nolan, "It's Embedded in What We Do for Every Child": A Qualitative Exploration of Early Childhood Educators' Perspectives on Supporting Children's Social and Emotional Learning, Int. J. Environ. Res. Public Health 18 (4) (2021) 1530. https://doi.org/10.3390/ijerph18041530.

[24] A. Viegut, I. Resnick, D. Miller-Cotto, N.S. Newcombe, N.C. Jordan, Tracking informal fraction knowledge and its correlates across first grade, Dev. Psychol. 59 (10) (2023) 1739–1756. https://doi.org/10.1037/dev0001581

[25] D.H. Bailey, R.S. Siegler, D.C. Geary, Early predictors of middle school fraction knowledge, Dev. Sci. 17 (5) (2014) 775–785. https://doi.org/10.1111/desc.12155.

[26] S.A. Hecht, K.J. Vagi, Sources of group and individual differences in emerging fraction skills, J. Educ. Psychol. 102 (4) (2010) 843–859. https://doi.org/10.1037/a0019824.

[27] E.M. Hubbard, P.G. Matthews, Ratio-based perceptual foundations for rational numbers, and perhaps whole numbers, too?, Behav. Brain Sci. 44 (2021) e192. https://doi.org/10.1017/S0140525X2100114X.

[28] A. Krajcsi, P. Kojouharova, G. Lengyel, Development of Preschoolers' Understanding of Zero, Front. Psychol. 12 (2021) 583734. https://doi.org/10.3389/fpsyg.2021.583734.

[29] L.S. Fuchs, et al., Improving at-risk learners' understanding of fractions, J. Educ. Psychol. 105 (3) (2013) 683–700. https://doi.org/10.1037/a0032446.

การศึกษาสภาพปัญหาทักษะการสอนด้านเทคโนโลยีดิจิทัลของครู สำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของวิชาชีพครูในโรงเรียนจังหวัดสุรินทร์

ประฎิชญา สร้อยจิตร[1,†] สุพัตรา วะยะลุน[2] อรุณธิดา สีเชียงหา[1,*] และ วิภาวดี มูลไชยสุข[1]

[1]สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสุรินทร์ 32000

[2]สาขาวิชาเทคโนโลยีคอมพิวเตอร์ คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏสุรินทร์ 32000

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาสภาพปัญหาทักษะการสอนด้านเทคโนโลยีดิจิทัลของครู สำหรับศตวรรษที่ 21 กลุ่มตัวอย่างในการวิจัยครั้งนี้ คือ ครูผู้สอนคณิตศาสตร์ จำนวน 40 คน 27 โรงเรียน เครื่องมือ ได้แก่ แบบสอบถามสภาพปัญหาทักษะการสอนด้านเทคโนโลยีดิจิทัลของครู แบบสัมภาษณ์แบบกึ่งโครงสร้าง ผลการวิจัยพบว่า 1) ด้านการเข้าถึง ต้องการให้โรงเรียนมีการสนับสนุนและส่งเสริมให้ท่านเข้ารับการอบรมพัฒนาเกี่ยวกับการใช้เครื่องคอมพิวเตอร์และโปรแกรมคอมพิวเตอร์สำหรับการสอนคณิตศาสตร์ ค่าเฉลี่ย เท่ากับ 3.95 และค่าส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.88 2) ด้านความเข้าใจ สามารถเลือกโปรแกรมเพื่อผลิตสื่อการเรียนการสอนได้อย่างมีประสิทธิภาพ ค่าเฉลี่ยเท่ากับ 3.60 และค่าส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.81 3) ด้านการใช้ การใช้เทคโนโลยีช่วยในการจัดเรียนการสอน ค่าเฉลี่ย เท่ากับ 3.68 และค่าส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.89 4) ด้านการสร้าง สร้างสื่อการเรียนการสอนโดยใช้เทคโนโลยีดิจิทัลได้อย่างคล่องแคล่ว ค่าเฉลี่ย เท่ากับ 3.58 และค่าส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.84

คำสำคัญ: ทักษะการสอน, เทคโนโลยีดิจิทัล, ศตวรรษที่ 21

2020 MSC: 97D40, 97U50, 97B50

*ผู้นำเสนอ (aruntida.si@srru.ac.th)

†ผู้ประพันธ์บรรณกิจ (Pradidchaya@gmail.com)

1 บทนำ

การจัดการศึกษาต้องยึดหลักว่าผู้เรียนทุกคนมีความสามารถเรียนรู้และพัฒนาตนเองได้และถือว่าผู้เรียนมีความสำคัญที่สุด กระบวนการจัดการศึกษาต้องส่งเสริมให้ผู้เรียนสามารถพัฒนาตามธรรมชาติและเต็มศักยภาพ เป้าหมายหลักของการพัฒนาการศึกษาของประเทศไทย คือ การพัฒนาผู้เรียนให้เป็นบุคคลที่มีคุณภาพด้วยกระบวนการเรียนรู้เพื่อความเจริญงอกงามของบุคคลและสังคมโดยถ่ายทอดความรู้ การฝึกอบรม การสืบสานทางวัฒนธรรม การสร้างสรรค์ จรรโลง ความก้าวหน้าทางวิชาการ การสร้างองค์ความรู้อันเกิดจากการจัดสภาพแวดล้อมสังคม การเรียนรู้และปัจจัย เกื้อหนุนให้บุคคลเรียนรู้อย่างต่อเนื่องตลอดชีวิต (พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2553) ทักษะความพร้อมในการดำรงชีวิตในศตวรรษที่ 21 มีทักษะครอบคลุม ครบถ้วนในสาระวิชารวมถึงทักษะที่จำเป็น 3 ด้าน ได้แก่ 1) ทักษะการเรียนรู้ด้านนวัตกรรม 2) ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี และ 3) ทักษะชีวิตและอาชีพ (การดำรงชีพ) ซึ่งทักษะทั้ง 3 ที่กล่าวมานี้ นับเป็นทักษะที่จำเป็นที่ทุกคนต้องเรียนรู้เพื่อปรับตัวให้เข้ากับศตวรรษที่ 21 โดยเฉพาะอย่างยิ่งทักษะการเรียนรู้ด้านคอมพิวเตอร์ เทคโนโลยี และสื่อสารสนเทศ นับว่าเป็นหนึ่งในทักษะที่สำคัญในการพัฒนาการเรียนรู้ตามรูปแบบ 7Cs ได้แก่ Communications, Information & Media Literacy (ทักษะด้านการสื่อสารสารสนเทศ และรู้เท่าทันสื่อ) Computing & ICT Literacy (ทักษะด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศและการสื่อสาร) (วิจารณ์ พานิช, 2555)

ครูในศตวรรษที่ 21 นี้ เน้นแบ่งปันองค์ความรู้เสริมสร้างประสบการณ์ พัฒนาศักยภาพด้านความรู้ ทักษะความสามารถ รังสรรค์องค์ความรู้สร้างคุณค่าให้ตนเองส่งต่อให้ผู้เรียน ผ่านรูปแบบจัดการเรียนการสอนแนวใหม่ (New Normal Class room) โดย ใช้เทคโนโลยีเป็นสื่อ ส่งเสริมทักษะใหม่ให้กับผู้เรียน ทักษะด้านเทคโนโลยีสารสนเทศ เป็นทักษะที่จำเป็นสำหรับการรับมือต่อการเปลี่ยนแปลงของโลกและสังคมในอนาคต ครูเองก็ต้องเข้าใจในบทบาทตนเองและต้องเร่งศึกษาค้นคว้าการใช้เทคโนโลยีสมัยใหม่ให้ชำนาญ ประกอบด้วย การใช้อินเทอร์เน็ต (Internet of Things: IOT) การประมวลผลแบบกลุ่มเมฆ (Cloud Computing) และเข้าใจการใช้ข้อมูลขนาดใหญ่ (Big Data) เทคโนโลยีเหล่านี้จะช่วยอำนวยความสะดวก ลดการใช้ทรัพยากร รวมทั้งตอบสนองความต้องการด้านต่าง ๆ ของผู้เรียนได้มากมาย ถึงเวลาที่ครูต้องปรับเปลี่ยนวิธีการสอน และใช้เทคโนโลยีเข้ามาช่วยออกแบบการสอนแทน รวมถึงการแก้ปัญหาการเรียนรู้ ซึ่งเป็นส่วนหนึ่งของการจัดสภาพแวดล้อมการเรียนรู้ที่ทำให้เกิดการเรียนรู้ เกิดความร่วมมือ แรงจูงใจ ตรงตามความต้องการที่เปลี่ยนแปลงไปของผู้เรียน เปิดพื้นที่ไว้สำหรับการแลกเปลี่ยนเรียนรู้ เพิ่มประสิทธิภาพของเนื้อหาความรู้และกระบวนการเรียนรู้ที่มีการจัดสอนของครูในห้องเรียนอย่างเหมาะสมในศตวรรษนี้ (ฐิติวัสส์ สุขป้อม และคณะ, 2563) ข้อจำกัดของครูไทยที่ไม่สามารถมี “ทักษะสมรรถนะ” ได้เต็มศักยภาพ เนื่องจาก

ครูไทยไม่ได้ถูกเตรียมตัวมาให้มี “ทักษะสมรรถนะ” นอกจากนั้น ต้องสร้างบรรยากาศให้โรงเรียนเป็นองค์กรแห่งการเรียนรู้ ครูสามารถเรียนรู้จากเพื่อนร่วมงาน ชุมชน ผู้ปกครอง หรือผู้เรียนเพื่อพัฒนาตนเองอย่างต่อเนื่อง ความก้าวหน้าในวิชาชีพครูต้องขึ้นกับสมรรถนะทางวิชาชีพที่ส่งผลให้ผู้เรียนมีผลสัมฤทธิ์การเรียนรู้ที่ดี มีคุณภาพ และสร้างเครือข่ายความร่วมมือและการเรียนรู้ของครูระหว่างสถานศึกษาโดยมีสำนักงานเขตพื้นที่การศึกษาและศึกษาธิการจังหวัดเป็นเจ้าภาพ ทั้งนี้มีครูอีกกลุ่มหนึ่งขาดทักษะการใช้ “ดิจิตอลแพลตฟอร์ม” ทุกคนสามารถถ่ายทอดได้ แต่ขาดการทำให้บทเรียนน่าสนใจ เป็นการเรียนแบบไม่มีชีวิต ขณะเดียวกันการเลื่อนชั้นอัตโนมัติของไทยก็เป็นอีกหนึ่งปัญหาเนื่องจากเป็นการเลื่อนชั้นแบบไม่มีระบบ เด็กทุกคนเลื่อนชั้นแต่ครูกลับสอนอยู่ที่เดิมไม่ได้ตามเด็กไปทั้งที่ครูเหล่านี้รู้จักเด็กของตนเองทำให้ขาดการเรียนรู้อย่างต่อเนื่อง (กรุงเทพธุรกิจ, 2564) ภารกิจสำคัญมหาวิทยาลัยราชภัฏสุรินทร์ตามมาตรา 7 แห่งพระราชบัญญัติมหาวิทยาลัยราชภัฏ พ.ศ. 2547 คือ ให้มหาวิทยาลัยเป็นสถาบันอุดมศึกษาเพื่อการพัฒนาท้องถิ่นที่เสริมสร้างพลังปัญญาของแผ่นดิน ฟื้นฟูพลังการเรียนรู้เชิดชูภูมิปัญญาท้องถิ่น สร้างสรรค์ศิลปวิทยา เพื่อความเจริญก้าวหน้าอย่างมั่นคงและยั่งยืนของปวงชน มีส่วนร่วมในการจัดการ การบำรุงรักษา การใช้ประโยชน์จากทรัพยากรธรรมชาติและสิ่งแวดล้อมอย่างสมดุลและยั่งยืน โดยมีวัตถุประสงค์ให้การศึกษา ส่งเสริมวิชาการและวิชาชีพชั้นสูง ทำการสอน วิจัย ให้บริการทางวิชาการแก่สังคม ปรับปรุง ถ่ายทอดและพัฒนาเทคโนโลยี ทำนุบำรุงศิลปะและวัฒนธรรม ผลิตครูและส่งเสริม วิทยฐานะครู (แผนยุทธศาสตร์ มหาวิทยาลัยราชภัฏสุรินทร์ ระยะ 5 ปี พ.ศ. 2566-2570, 2565)

จากปัญหาดังกล่าวผู้วิจัยและคณะเล็งเห็นถึงการพัฒนาอย่างเร่งด่วน คณะผู้วิจัยได้มีประสบการณ์การในการจัดโครงการพัฒนาสมรรถนะครูคณิตศาสตร์ในจังหวัดสุรินทร์ จากการสัมภาษณ์และการพัฒนาครูพบว่า ครูมีความสามารถในการถ่ายทอดความรู้คณิตศาสตร์เป็นอย่างดีแต่ครูขาดทักษะการใช้เทคโนโลยีดิจิทัลที่เกิดประสิทธิภาพในการสอน ตลอดเพื่อให้ครูมีทักษะเพื่ออนาคตทันต่อการเปลี่ยนแปลงของโลกแต่การสอนยังคงมีประสิทธิภาพและการวิจัยในครั้งนี้สอดคล้องกับภารกิจยุทธศาสตร์ของมหาวิทยาลัยราชภัฏสุรินทร์ที่เห็นถึงความสำคัญของการพัฒนาครูในท้องถิ่นและชุมชน จึงสนใจศึกษาปัญหาทักษะการสอนด้านเทคโนโลยีดิจิทัลของครูสำหรับศตวรรษที่ 21 ครูในโรงเรียนจังหวัดสุรินทร์ เพื่อเป็นประโยชน์ต่อการศึกษาของชาติและพัฒนาครูในจังหวัดสุรินทร์ต่อไป

2 วิธีการดำเนินงาน

ในการวิจัยครั้งนี้ผู้วิจัยมุ่งที่จะศึกษาสภาพปัญหาทักษะการสอนโดยใช้เทคโนโลยีดิจิทัลของครูและการเรียนการสอนและการจัดการศึกษาสำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของครูในโรงเรียนจังหวัดสุรินทร์ กลุ่ม

ตัวอย่าง ได้แก่ ศึกษานิเทศก์ ผู้บริหารสถานศึกษา ครูผู้สอนคณิตศาสตร์ ครูพี่เลี้ยงสาขาวิชาคณิตศาสตร์ให้กับนักศึกษาฝึกประสบการณ์ของมหาวิทยาลัยราชภัฏสุรินทร์ จำนวน 40 คน จาก 27 โรงเรียนในจังหวัดสุรินทร์ โดยใช้วิธีการคัดเลือกแบบเฉพาะเจาะจงจากครูผู้สอนวิชาคณิตศาสตร์ผู้สนใจที่ทำการสมัครเข้าร่วมโครงการวิจัย ซึ่งผู้วิจัยได้ดำเนินการดังนี้

1. การศึกษาสภาพปัญหาและรวมถึงสภาพปัญหาทั่วไปในการวิจัย ศึกษาเอกสาร งานวิจัยที่เกี่ยวข้องโดยมีการศึกษาจากแหล่งข้อมูล

2. การศึกษาภาคสนาม เนื่องจากการวิจัยครั้งนี้เป็นวิจัยด้านทักษะการสอนด้านเทคโนโลยีดิจิทัลของครูในโรงเรียนจังหวัดสุรินทร์ จึงเน้นการเก็บข้อมูลเพื่อนำมาพัฒนา โดยผู้วิจัยได้แบ่งการรวบรวมข้อมูลไว้ดังนี้

2.1 การสัมภาษณ์เชิงลึก ผู้วิจัยทำการเก็บรวบรวมข้อมูลโดยการสัมภาษณ์แบบกึ่งโครงสร้าง สัมภาษณ์กลุ่มศึกษานิเทศก์ ผู้บริหารสถานศึกษา ครูผู้สอนคณิตศาสตร์ และครูพี่เลี้ยงสาขาวิชาคณิตศาสตร์ โรงเรียนในจังหวัดสุรินทร์ ซึ่งสังกัดเขตพื้นที่การศึกษามัธยมศึกษา รวมทั้งสิ้นจำนวน 85 โรงเรียน โดยจังหวัดสุรินทร์มีการแบ่งเขตพื้นที่เป็น 3 เขต

3. เครื่องมือที่ใช้ในการรวบรวมข้อมูล

3.1 แบบสอบถาม (Questionnaires) ที่ผู้วิจัยทำการสร้างขึ้นโดยพิจารณาจากเครื่องมือที่มีผู้วิจัยมาก่อนภายใต้กรอบแนวคิดเชิงทฤษฎีที่ได้จากการศึกษาวิเคราะห์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง รวมทั้งให้สอดคล้องกับคำจำกัดความในการวิจัยที่ กำหนดประกอบด้วยชุดคำถาม ดังนี้

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถามเป็นคำถามเกี่ยวกับ อายุ ระดับการศึกษา ประสบการณ์ในการทำงาน และตำแหน่งหน้าที่ในสถานศึกษา มีลักษณะเป็นแบบ ตรวจสอบ

ตอนที่ 2 สภาพปัญหาและสภาพทักษะการสอนด้านเทคโนโลยีดิจิทัลของครูในโรงเรียนจังหวัดสุรินทร์ ข้อคำถาม

ตอนที่ 3 ข้อเสนอแนะอื่น ๆ

3.2 แบบสัมภาษณ์แบบกึ่งโครงสร้าง (Semi-Structured Interview) เกี่ยวกับปัญหาด้านทักษะการสอนด้านเทคโนโลยีดิจิทัลของครูในโรงเรียนจังหวัดสุรินทร์นำแบบสัมภาษณ์แบบกึ่งโครงสร้างที่สมบูรณ์แล้วไปใช้เก็บรวบรวมข้อมูลจากผู้บริหารสถานศึกษา ครูผู้สอนคณิตศาสตร์ ในสังกัดสำนักงานเขตพื้นที่การศึกษาในจังหวัดสุรินทร์

3.3 แบบสอบถามสภาพปัญหาและสภาพทักษะการสอนด้านเทคโนโลยีดิจิทัลของครูในโรงเรียนจังหวัดสุรินทร์ เก็บแบบสอบถาม กลุ่มศึกษานิเทศก์ ผู้บริหารสถานศึกษา ครูผู้สอนคณิตศาสตร์ และครูพี่เลี้ยงสาขาวิชาคณิตศาสตร์ โรงเรียนในจังหวัดสุรินทร์ ซึ่งสังกัดเขตพื้นที่การศึกษาสุรินทร์ รวมทั้งสิ้นจำนวน 85 โรงเรียน โดยจังหวัดสุรินทร์มีการแบ่งเขตพื้นที่เป็น 3 เขต โดยใช้การสุ่มเพื่อเก็บแบบสอบถาม จำนวน 40 ชุด โดยใช้แบบสอบถามที่ผ่านการตรวจสอบของผู้เชี่ยวชาญ

3 ผลการศึกษา

ผลวิเคราะห์การศึกษาสภาพปัญหาทักษะการสอนโดยใช้เทคโนโลยีดิจิทัลของครูและการเรียนการสอนและการจัดการศึกษาสำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของครูในโรงเรียนจังหวัดสุรินทร์ โดยใช้ 1) แบบสอบถาม 2) แบบสัมภาษณ์แบบกึ่งโครงสร้าง ที่ผู้วิจัยสร้างขึ้น กลุ่มตัวอย่างที่ใช้ในการวิจัย ได้แก่ ครูผู้สอน สังกัดในจังหวัดสุรินทร์ จำนวน 40 คน แสดงผลการวิเคราะห์ตอนที่ 1 จำนวนและร้อยละข้อมูลผู้ตอบแบบสอบถามครูผู้สอนสังกัดในจังหวัดสุรินทร์ ผลวิเคราะห์การศึกษาสภาพปัญหาทักษะการสอนโดยใช้เทคโนโลยีดิจิทัลของครูและการเรียนการสอนและการจัดการศึกษาสำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของครูในโรงเรียนจังหวัดสุรินทร์

พบว่าค่าเฉลี่ยสูงสุดรายข้อ เป็นเพศหญิง ร้อยละ 80.00 วุฒิการศึกษาปริญญาโท ร้อยละ 52.50 ตำแหน่ง ครูชำนาญการพิเศษ ร้อยละ 35.00 และสอนในระดับประถมศึกษามากที่สุด ร้อยละ 65.00

ตารางที่ 1 ผลวิเคราะห์การศึกษาสภาพปัญหาทักษะการสอนโดยใช้เทคโนโลยีดิจิทัลของครูและการเรียนการสอนและการจัดการศึกษาสำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของครูในโรงเรียนจังหวัดสุรินทร์

รายการ	**ระดับความสามารถ**		
	$\overline{X}$	SD	**การแปลผล**
1. ด้านการเข้าถึง			
1.1 ท่านสามารถใช้งานอุปกรณ์และเครื่องมือประกอบการจัดการเรียนการสอนได้อย่างคล่องแคล่ว เช่น แท็บเล็ต คอมพิวเตอร์ สมาร์ทโฟน โปรเจคเตอร์	4.20	0.72	มาก
1.2 โรงเรียนที่ท่านสอนมีอินเทอร์เน็ตที่มีประสิทธิภาพเข้าถึงข้อมูลแบบออนไลน์ได้อย่างรวดเร็ว	4.15	0.70	มาก
1.3 ท่านสามารถสืบค้นข้อมูลในอินเทอร์เน็ตได้อย่างคล่องแคล่ว	4.55	0.68	มากที่สุด
1.4 ท่านสามารถรู้จักโปรแกรมทางคอมพิวเตอร์ที่ช่วยใช้ในการสอนได้	4.18	0.98	มาก

รายการ	ระดับความสามารถ		
	$\overline{X}$	SD	การแปลผล
1.5 โรงเรียนมีการสนับสนุนและส่งเสริมให้ท่าน เข้ารับการอบรมพัฒนาเกี่ยวกับการใช้เครื่อง คอมพิวเตอร์และโปรแกรมคอมพิวเตอร์สำหรับการสอน	3.95	0.88	ปานกลาง
2. ด้านความเข้าใจ			
2.1 สามารถเลือกใช้สื่อการสอนด้านเทคโนโลยีดิจิทัลได้อย่างมีประสิทธิภาพ	3.63	0.63	ปานกลาง
2.2 ท่านเข้าใจความต้องการของผู้เรียนในการใช้สื่อเทคโนโลยีดิจิทัล	3.73	0.64	ปานกลาง
2.3 สามารถเลือกโปรแกรมเพื่อผลิตสื่อการเรียนการสอนได้อย่างมีประสิทธิภาพ	3.60	0.81	ปานกลาง
2.4 ท่านสามารถเปลี่ยนแปลงและปรับตัวให้เข้ากับเทคโนโลยี	3.85	0.80	ปานกลาง
2.5 ท่านเข้าใจความแตกต่างระหว่างบุคคลของผู้เรียนในการใช้สื่อเทคโนโลยีดิจิทัลเพื่อช่วยในการสอน	4.30	0.46	มาก
3. ด้านการใช้			
3.1 ท่านใช้เทคโนโลยีช่วยในการจัดเรียนการสอน เช่น โปรแกรม GSP, Latex, GeoGebra, Kahoot, Wordwall, Quizizz	3.68	0.89	ปานกลาง
3.2 ท่านสามารถสำรอง จัดเก็บ ส่งข้อมูลในระบบออนไลน์ได้ เช่น Google Drive	3.80	0.76	ปานกลาง
3.3 ท่านใช้เทคโนโลยีดิจิทัลในการสร้างเครือข่ายการเรียนรู้ทางวิชาชีพได้ เช่น Google Meet, Zoom	4.18	0.75	มาก
3.4 ท่านมีทักษะในการใช้เทคโนโลยีเพื่อช่วยวัดและประเมินผลในการเรียนได้	3.70	0.69	ปานกลาง
3.5 ท่านใช้เทคโนโลยีในการจัดการชั้นเรียนในการจัดการชั้นเรียน	3.73	0.72	ปานกลาง
4. ด้านการสร้าง			

รายการ	ระดับความสามารถ		
	$\overline{X}$	SD	การแปลผล
4.1 ท่านสร้างสื่อการเรียนการสอนโดยใช้เทคโนโลยีดิจิทัลได้อย่างคล่องแคล่ว	3.58	0.84	ปานกลาง
4.2 ท่านใช้เครื่องมือสื่อเทคโนโลยีดิจัลในจัดการเรียนการสอนได้ทุกแฟรตฟอร์ม	3.75	0.74	ปานกลาง
4.3 ท่านสามารถผลิตเนื้อหาในการสอนผ่านเครื่องมือดิจิทัลได้อย่างหลากหลายเครื่องมือ	3.88	0.91	ปานกลาง
4.4 ท่านสามารถใช้โปรแกรมสร้างเว็บหรือบทเรียนบนเครือข่ายออนไลน์ได้	3.90	0.87	ปานกลาง
4.5 ท่านมีการเผยแพร่ความรู้ด้านทักษะการสอนเทคโนโลยีดิจิทัลทางสื่อสังคมออนไลน์ให้กับผู้อื่น เช่น Facebook , YouTube	4.63	0.63	มากที่สุด
เฉลี่ยรวม	**3.95**	**0.75**	**ปานกลาง**

จากตารางที่ 1 พบว่า ผลวิเคราะห์การศึกษาสภาพปัญหาทักษะการสอนโดยใช้เทคโนโลยีดิจิทัลของครูและการเรียนการสอนและการจัดการศึกษาสำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของครูในโรงเรียนจังหวัดสุรินทร์ ทำการศึกษา 4 ด้าน ดังนี้

1) ด้านการเข้าถึง ครูสามารถสืบค้นข้อมูลในอินเทอร์เน็ตได้อย่างคล่องแคล่ว อยู่ในระดับมากที่สุด ($\overline{X}$ =4.55, SD=0.68) ค่าเฉลี่ยรองลงมาคือ ครูสามารถใช้งานอุปกรณ์และเครื่องมือประกอบการจัดการเรียนการสอนได้อย่างคล่องแคล่ว เช่น แท็บเล็ต คอมพิวเตอร์ สมาร์ทโฟน โปรเจคเตอร์ ($\overline{X}$ =4.20, SD=0.72) และ โรงเรียนมีการสนับสนุนและส่งเสริมให้ครูเข้ารับการอบรม พัฒนาเกี่ยวกับการใช้เครื่องคอมพิวเตอร์และโปรแกรมคอมพิวเตอร์สำหรับการสอน มีค่าเฉลี่ยต่ำที่สุด ($\overline{X}$ =3.95, SD=0.88)

2) ด้านความเข้าใจ ครูเข้าใจความแตกต่างระหว่างบุคคลของผู้เรียนในการใช้สื่อเทคโนโลยีดิจิทัลเพื่อช่วยในการสอน มีค่าเฉลี่ยสูงสุด ($\overline{X}$ =4.30, SD=0.46) ค่าเฉลี่ยรองลงมาคือ ครูสามารถเปลี่ยนแปลงและปรับตัวให้เข้ากับ

เทคโนโลยี ($\overline{X}$ =3.85, SD=0.80) และสามารถเลือกโปรแกรมเพื่อผลิตสื่อการเรียนการสอนได้อย่างมีประสิทธิภาพ มีค่าเฉลี่ยต่ำที่สุด ($\overline{X}$ =3.60, SD=0.81)

3) ด้านการใช้ ครูสามารถใช้เทคโนโลยีดิจิทัลในการสร้างเครือข่ายการเรียนรู้ทางวิชาชาชีพได้ เช่น Google Meet, Zoom อยู่ในระดับมากที่สุด ($\overline{X}$ =4.18, SD=0.75) ค่าเฉลี่ยรองลงมาคือ ครูสามารถสำรอง จัดเก็บ ส่งข้อมูลในระบบออนไลน์ได้ เช่น Google Drive ($\overline{X}$ =3.80, SD=0.76) และครูใช้เทคโนโลยีช่วยในการจัดเรียนการสอน เช่น โปรแกรม GSP, Latex, GeoGebra, Kahoot, Wordwall, Quizizz มีค่าเฉลี่ยต่ำที่สุด ($\overline{X}$ =3.68, SD=0.89)

4) ด้านการสร้าง ครูมีการเผยแพร่ความรู้ด้านทักษะการสอนเทคโนโลยีดิจิทัล ทางสื่อสังคมออนไลน์ให้กับผู้อื่น เช่น Facebook, YouTube อยู่ในระดับมากที่สุด ($\overline{X}$ =4.63, SD=0.63) ค่าเฉลี่ยรองลงมา คือ ครูสามารถใช้โปรแกรมสร้างเว็บหรือบทเรียนบนเครือข่ายออนไลน์ได้ ($\overline{X}$ =3.90, SD=0.87) และครูสร้างสื่อการเรียนการสอนโดยใช้เทคโนโลยีดิจิทัลได้อย่างคล่องแคล่ว มีค่าเฉลี่ยต่ำที่สุด ($\overline{X}$ =3.58, SD=0.84)

ผลการวิเคราะห์การสัมภาษณ์แบบกึ่งโครงสร้าง พบว่า ปัญหาด้านทักษะการสอนโดยใช้เทคโนโลยีดิจิทัลของครูที่พบมากที่สุด 3 ประการ คือ 1) ครูขาดการสนับสนุนในด้านอุปกรณ์จากโรงเรียน 2) ครูไม่สามารถนำเทคโนโลยีมาใช้ในการสอนให้เกิดประสิทธิภาพในชั้นเรียนได้ 3) ครูขาดการสนับสนุนในการพัฒนาและอบรมในด้านทักษะเทคโนโลยีดิจิทัล

4 สรุปผลและอภิปรายผล

จากผลการศึกษาวิจัย เรื่อง การศึกษาสภาพปัญหาทักษะการสอนด้านเทคโนโลยีดิจิทัลของครู สำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของวิชาชีพครูในโรงเรียนจังหวัดสุรินทร์ ผู้วิจัยได้นำประเด็นที่ค้นพบมาอภิปรายดังนี้ ผลวิเคราะห์สภาพปัญหาทักษะการสอนด้านเทคโนโลยีดิจิทัลของครู สำหรับศตวรรษที่ 21 บนฐานชุมชนการเรียนรู้ของวิชาชีพครูในโรงเรียนจังหวัดสุรินทร์ ผู้วิจัยและคณะได้ศึกษาพบว่า โดยรวมและรายด้าน อยู่ในระดับมาก เรียงลำดับค่าเฉลี่ยจากมากไปน้อย ได้แก่ ด้านการใช้ ด้านการเข้าถึง และ ด้านการสร้าง สามารถสรุปเป็นรายด้านดังนี้ ด้านการเข้าถึง ครูสามารถสืบค้นข้อมูลในอินเทอร์เน็ตได้อย่างคล่องแคล่ว อยู่ในระดับมากที่สุด ค่าเฉลี่ยรองลงมาคือ ครูสามารถใช้งานอุปกรณ์และเครื่องมือประกอบการจัดการเรียนการสอนได้อย่างคล่องแคล่ว เช่น แท็บเล็ต คอมพิวเตอร์ สมาร์ทโฟน โปรเจคเตอร์ และโรงเรียนมีการสนับสนุนและส่งเสริมให้ครูเข้ารับการอบรมพัฒนาเกี่ยวกับการใช้เครื่องคอมพิวเตอร์และโปรแกรมคอมพิวเตอร์สำหรับการสอน มีค่าเฉลี่ยต่ำที่สุด ด้านความเข้าใจ ครูเข้าใจความแตกต่างระหว่างบุคคลของผู้เรียนในการใช้สื่อเทคโนโลยีดิจิทัลเพื่อช่วยในการสอน มีค่าเฉลี่ยสูงสุด ค่าเฉลี่ยรองลงมาคือ ครูสามารถเปลี่ยนแปลงและปรับตัวให้เข้ากับเทคโนโลยี และสร้างสื่อการเรียนการสอน

โดยใช้เทคโนโลยีดิจิทัลได้อย่างคล่องแคล่วมีค่าเฉลี่ยต่ำที่สุด ด้านการใช้ ครูสามารถใช้เทคโนโลยีดิจิทัลในการสร้างเครือข่ายการเรียนรู้ทางวิชาชีพได้ เช่น Google Meet, Zoom อยู่ในระดับมากที่สุด ค่าเฉลี่ยรองลงมาคือ ครูสามารถสำรอง จัดเก็บ ส่งข้อมูล ในระบบออนไลน์ได้ เช่น Google Drive และครูใช้เทคโนโลยีช่วยในการจัดเรียนการสอน เช่น โปรแกรม GSP, Latex, GeoGebra, Kahoot, Wordwall, Quizizz มีค่าเฉลี่ยต่ำที่สุด ด้านการสร้าง ครูมีการเผยแพร่ความรู้ด้านทักษะการสอนเทคโนโลยีดิจิทัลทางสื่อสังคมออนไลน์ให้กับผู้อื่น เช่น Facebook, YouTube อยู่ในระดับมากที่สุด ค่าเฉลี่ยรองลงมา คือ ครูสามารถใช้โปรแกรมสร้างเว็บหรือบทเรียนบนเครือข่ายออนไลน์ได้ และครูสร้างสื่อการเรียนการสอนโดยใช้เทคโนโลยีดิจิทัลได้อย่างคล่องแคล่ว มีค่าเฉลี่ยต่ำที่สุด โดยสรุป 1) ด้านการเข้าถึง จากผลการสรุปจะเห็นว่า ครูมีความสามารถในการค้นหาข้อมูลทางอินเทอร์เน็ตอยู่ในระดับดี แต่ความสามารถในการใช้เครื่องมือการสอนในห้องเรียนมีความไม่สอดคล้องกัน แสดงให้เห็นว่าแม้ครูจะสามารถเข้าถึงข้อมูลได้ดี แต่ยังขาดทักษะในการนำข้อมูลที่ค้นหาไปใช้ในกระบวนการสอนจริง ๆ ซึ่งอาจส่งผลต่อคุณภาพการเรียนการสอน 2) ด้านความเข้าใจ ครูมีความเข้าใจในความแตกต่างของผู้เรียน แต่การสร้างสื่อการเรียนการสอนยังต้องการพัฒนาการเข้าใจผู้เรียนเป็นปัจจัยสำคัญในกระบวนการเรียนการสอน แต่การขาดทักษะในการสร้างสื่อการเรียนการสอนดิจิทัลอาจเป็นอุปสรรคในการปรับใช้ความรู้ที่มีอยู่ 3) ด้านการใช้ ครูมีทักษะในการสร้างเครือข่ายการเรียนรู้ผ่านเทคโนโลยี แต่การใช้โปรแกรมเฉพาะสำหรับการสอนยังต้องการการพัฒนาซึ่งการใช้เทคโนโลยีในการสื่อสารและแลกเปลี่ยนความรู้ช่วยเพิ่มประสิทธิภาพในการสอน แต่หากไม่มีการพัฒนาทักษะในโปรแกรมที่เฉพาะเจาะจง การสอนอาจไม่สามารถตอบสนองต่อความต้องการของนักเรียนได้ 4) ด้านการสร้าง ครูมีความสามารถในการเผยแพร่ความรู้ผ่านโซเชียลมีเดียอยู่ในระดับดี แต่การสร้างสื่อการเรียนการสอนโดยใช้เทคโนโลยีดิจิทัลได้อย่างคล่องแคล่วคะแนนยังต่ำ ซึ่งสอดคล้องกับ (ศิริวรรณ ฉัตรมณีรุ่งเจริญ และวรางคนา ทองนพคุณ, 2557) 5) ผลการวิเคราะห์การสัมภาษณ์แบบกึ่งโครงสร้าง พบว่า ปัญหาด้านทักษะการสอนโดยใช้เทคโนโลยีดิจิทัลของครู ที่พบมากที่สุด 3 ประการคือ 1) ครูขาดการสนับสนุนในด้านอุปกรณ์จากโรงเรียน เช่น ครูมีอุปกรณ์แท็บเล็ต คอมพิวเตอร์ แต่โรงเรียนไม่มีอินเทอร์เน็ตที่มีประสิทธิภาพ หรือไม่มีอุปกรณ์ที่จะรองรับการใช้งาน 2) ครูไม่สามารถนำเทคโนโลยีมาใช้ในการสอนให้เกิดประสิทธิภาพในชั้นเรียนได้ เช่น ครูไม่รู้วิธีที่หลากหายการวัดและประเมินผลโดยใช้เทคโนโลยี และ 3) ครูขาดการสนับสนุนในการพัฒนาและอบรมในด้านทักษะเทคโนโลยีดิจิทัล ครูต้องการการสนับสนุนในด้านงบประมาณในการฝึกอบรมที่ครูสนใจ หรือหน่วยงานไม่มีการพัฒนาสนับสนุนฝึกอบรมด้านทักษะเทคโนโลยีดิจิทัล สอดคล้องกับงานวิจัยของ บันเย็น เพ็งกระจ่าง (2561 : 75 - 76) ได้ทำการศึกษา เรื่อง การพัฒนาครูด้านการจัดการเรียนการสอนในศตวรรษที่ 21 ของโรงเรียนสารสาสน์วิเทศ คลองหลวง สังกัดสำนักงาน คณะกรรมการส่งเสริมการศึกษาเอกชน ผลการศึกษาพบว่า ปัญหาและอุปสรรคการพัฒนาครูด้านการจัดการเรียนการสอนในศตวรรษที่ 21 มี 4 ประการ คือ 1) ครูขาดการพัฒนาและฝึกอบรม เช่น หลักสูตร นวัตกรรมใหม่ และเทคโนโลยีใหม่ 2) ขาดแคลนงบประมาณในการสร้างสื่อ นวัตกรรมใหม่ ๆ 3) ครูสอนเนื้อหาแค่ในตารางเรียน ครูที่มีประสบการณ์สูงไม่ยอมช่วยเหลือครูที่เข้ามาใหม่และไม่เสนอแนะช่วยเหลือ และ 4) การประสานงานที่ไม่ค่อยตรงกัน

ดังนั้น การเผยแพร่ความรู้เป็นสิ่งที่มีคุณค่า แต่การสร้างสื่อการเรียนการสอนที่มีคุณภาพจะช่วยให้การเรียนการสอนมีประสิทธิภาพยิ่งขึ้น ครูมีความสามารถในระดับปานกลางส่งผลให้ครูต้องได้รับการพัฒนาทักษะการสอนด้านเทคโนโลยีดิจิทัลเพื่อให้ตอบสนองความต้องการของผู้เรียนและสอนอย่างมีประสิทธิภาพ

กิตติกรรมประกาศ ผู้แต่งขอขอบคุณมหาวิทยาลัยราชภัฏสุรินทร์ ที่ให้ทุนในการวิจัยครั้งนี้ และขอขอบคุณผู้ทรงคุณวุฒิทุกท่านที่ได้ให้ข้อคิดเห็นและข้อเสนอแนะต่าง ๆ เพื่อปรับปรุงบทความวิจัยนี้

เอกสารอ้างอิง

[1] คุณัชญา ทิพยานุรักษ์สกุล, การศึกษาสภาพปัญหา และแนวทางแก้ปัญหาการเรียนการสอนออนไลน์สำหรับนักเรียนระดับประถมศึกษาในโรงเรียนเอกชน, วารสารวิชาการสถาบันการอาชีวศึกษาภาคเหนือ 3. 7(14) (2566) 347--364.

[2] ฐิติวัสส์ สุขป้อม และคณะ, ครูกับเทคโนโลยีการสอนในศตวรรษที่ 21, วารสารบัณฑิตวิทยาลัย มหาวิทยาลัยราชภัฏจันทรเกษม. 15(2) (2563) 1--17.

[3] เทคโนโลยีการสอนในศตวรรษที่ 21, วารสารบัณฑิตวิทยาลัย มหาวิทยาลัยราชภัฏจันทรเกษม. 15(2) 1--17.

[4] นันทวัน จันทร์กลิ่น, การศึกษาปัญหาและแนวทางการบริหารจัดการคุณภาพในการพัฒนาทักษะผู้เรียนในศตวรรษที่ 21: โรงเรียนบ้านเนินมะปราง สำนักงานเขตพื้นที่การศึกษาประถมศึกษาพิษณุโลก เขต 2, (วิทยานิพนธ์ปริญญามหาบัณฑิต), มหาวิทยาลัยราชภัฏอุบลราชธานี, 2557.

[5] วิจารณ์ พานิช, วิถีการสร้างการเรียนรู้เพื่อศิษย์ในศตวรรษที่ 21, กรุงเทพฯ: มูลนิธิสด ศรีสฤษดิ์วงศ์, 2555.

[6] วรพจน์ วงศ์กิจรุ่งเรือง และ อธิป จิตฤกษ์, ทักษะแห่งอนาคตใหม่: การศึกษาเพื่อศตวรรษที่ 21, กรุงเทพฯ: โอเพ่นเวิลด์ส, 2554.

[7] วรลักษณ์ คำหว่าง, แนวทางพัฒนาทักษะครูในศตวรรษที่ 21 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาในจังหวัดพิษณุโลก, วารสารมหาวิทยาลัยราชภัฏลำปาง. 6(1) (2559) 129--138.

[8] สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 33, แบบรายงานผลการดำเนินกิจกรรมพัฒนาวิชาชีพแบบชุมชนแห่งการเรียนรู้ทางวิชาชีพ (Professional Learning Community: PLC), 2562.

[9] สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 33, แบบรายงานผลการดำเนินกิจกรรมพัฒนาวิชาชีพแบบชุมชนแห่งการเรียนรู้ทางวิชาชีพ (Professional Learning Community: PLC) ประจำปีการศึกษา 2563, 2565.

[10] ศิริวรรณ ฉัตรมณีรุ่งเจริญ และ วรางคณา ทองนพคุณ, ทักษะแห่งศตวรรษที่ 21 ความท้าทายในอนาคต, คณะครุศาสตร์ มหาวิทยาลัยภูเก็ต, 2557.

[11] ปรียา สงค์ประเสริฐ, ผลการพัฒนาระบบประกันคุณภาพภายในสถานศึกษาด้วยกระบวนการนิเทศแบบ 3PPS, ในรายงานผลการวิจัย, สำนักงานเขตพื้นที่การศึกษาประถมศึกษานครศรีธรรมราช เขต 1, 2565.

[12] มหาวิทยาลัยราชภัฏสุรินทร์, แผนยุทธศาสตร์มหาวิทยาลัยราชภัฏสุรินทร์ ระยะ 5 ปี พ.ศ. 2566–2570, สุรินทร์: กองนโยบายและแผน, 2566.

[13] B. S. Bloom, Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain, New York, NY: David McKay Company, Inc., 1956.

[14] T. J. Sergiovanni, Building Community in Schools, San Francisco, CA: Jossey-Bass, 1994.

การพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์โดย การจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4

เวธน์วศิล พิรุณกาญจน์[1,*] อัคเรศ เตี๋ยวสกุล[1,*] กาญจนา เจริญช่วย[2] กรวิกา ก้องกุล[3,†] อลงกรณ์ แซ่ตั้ง[3]
เกษราภรณ์ แซ่ตั้ง[3] และ ปรีดาภรณ์ กาญจนสำราญวงศ์[3]

[1]สาขาวิชาคณิตศาสตร์ คณะศึกษาศาสตร์ มหาวิทยาลัยทักษิณ 90000
[2]กลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนหาดใหญ่วิทยาลัย 90110
[3]สาขาวิชาคณิตศาสตร์และการจัดการข้อมูล คณะวิทยาศาสตร์และนวัตกรรมดิจิทัล มหาวิทยาลัยทักษิณ 93210

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์ 1) เพื่อพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ให้มีคะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์เฉลี่ยไม่น้อยกว่าร้อยละ 70 ของคะแนนเต็ม และมีจำนวนนักเรียนที่ผ่านเกณฑ์อย่างน้อยร้อยละ 70 ของจำนวนนักเรียนทั้งหมด 2) เพื่อพัฒนาผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ให้มีคะแนนผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์เฉลี่ยไม่น้อยกว่าร้อยละ 70 ของคะแนนเต็ม และมีค่าดัชนีประสิทธิผลไม่น้อยกว่า 0.5 โดยมีกลุ่มเป้าหมาย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 4/11 โรงเรียนหาดใหญ่วิทยาลัย สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาสงขลา สตูล ภาคเรียนที่ 2 ปีการศึกษา 2567 จำนวน 39 คน รูปแบบการวิจัยเป็นวิจัยเชิงปฏิบัติการ จำนวน 3 วงจร เครื่องมือที่ใช้วิจัยประกอบด้วย แผนจัดการเรียนรู้ เรื่อง ความน่าจะเป็น โดยใช้การจัดการเรียนรู้ด้วยวิธีการแบบเปิด แบบทดสอบวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ เรื่อง ความน่าจะเป็น และแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง ความน่าจะเป็น สถิติในการวิเคราะห์ข้อมูลโดยใช้ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ร้อยละ ผลการวิจัยพบว่า 1) นักเรียนมีคะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์เฉลี่ยเท่ากับ 19.46 คิดเป็นร้อยละ 81.08 ของคะแนนเต็ม และมีจำนวนนักเรียนที่ผ่านเกณฑ์ 31 คน คิดเป็นร้อยละ 79.49 ของจำนวนนักเรียนทั้งหมด และ 2) นักเรียนมีคะแนนผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์เฉลี่ยเท่ากับ 10.62 คิดเป็นร้อยละ 70.80 ของคะแนนเต็ม และมีค่าดัชนีประสิทธิผลเท่ากับ 0.63 ซึ่งสูงกว่าเกณฑ์ที่กำหนด แสดงให้เห็นพัฒนาการในการเรียนรู้ของนักเรียนที่มีความรู้เพิ่มขึ้นคิดเป็นร้อยละ 63

*ผู้นำเสนอ (akkaraidtiawsakul@gmail.com)
†ผู้ประพันธ์บรรณกิจ (konvika@tsu.ac.th)

คำสำคัญ: การจัดการเรียนรู้ด้วยวิธีการแบบเปิด, ความสามารถในการแก้ปัญหาทางคณิตศาสตร์, ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์

2020 MSC: 97D40, 97D60

1 บทนำ

คณิตศาสตร์มีบทบาทที่สำคัญต่อความสำเร็จในการเรียนรู้ในศตวรรษที่ 21 ช่วยให้มนุษย์มีความคิดริเริ่มสร้างสรรค์ คิดอย่างมีเหตุผล เป็นระบบ มีแบบแผน สามารถวิเคราะห์ปัญหาหรือสถานการณ์ได้อย่างรอบคอบและถี่ถ้วน ช่วยให้คาดการณ์ วางแผน ตัดสินใจ แก้ปัญหา ได้อย่างถูกต้องเหมาะสม และสามารถนำไปใช้ในชีวิตจริงได้อย่างมีประสิทธิภาพ รวมทั้ง คณิตศาสตร์ยังเป็นเครื่องมือในการศึกษาด้านวิทยาศาสตร์ เทคโนโลยี และศาสตร์อื่น ๆ อันเป็นรากฐานในการพัฒนาทรัพยากรบุคคลของชาติให้มีคุณภาพและพัฒนาเศรษฐกิจของประเทศให้ทัดเทียมกับนานาชาติ หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560) มุ่งส่งเสริมให้นักเรียนมีทักษะที่จำเป็นสำหรับการเรียนรู้ในศตวรรษที่ 21 เป็นสำคัญ นั่นคือ การเตรียมนักเรียนให้มีทักษะด้านการแก้ปัญหา ทั้งนี้ การจัดการเรียนรู้คณิตศาสตร์ที่ประสบความสำเร็จนั้น จะต้องเตรียมนักเรียนให้มีความพร้อมที่จะเรียนรู้สิ่งต่าง ๆ ดังนั้น สถานศึกษาควรจัดการเรียนรู้ให้เหมาะสมตามศักยภาพของนักเรียน หลักสูตรแกนกลางการศึกษาขั้นพื้นฐานมุ่งให้นักเรียนเกิดสมรรถนะสำคัญด้านความสามารถในการแก้ปัญหา คือ สามารถแก้ปัญหาและอุปสรรคต่าง ๆ ที่เผชิญได้อย่างถูกต้องเหมาะสมบนพื้นฐานของหลักเหตุผล คุณธรรมและข้อมูลสารสนเทศ เข้าใจความสัมพันธ์และการเปลี่ยนแปลงของเหตุการณ์ต่าง ๆ ในสังคม นำความรู้ที่ได้จากการศึกษามาใช้ในการป้องกันและแก้ไขปัญหาที่ประสบ ซึ่งการตัดสินใจที่มีประสิทธิภาพต้องคำนึงถึงผลกระทบที่เกิดขึ้นต่อตนเอง สังคม และสิ่งแวดล้อม [1] การแก้ปัญหาทางคณิตศาสตร์เป็นกระบวนการที่นักเรียนควรเรียนรู้ ฝึกฝน และพัฒนาให้เกิดขึ้น ทั้งนี้ การแก้โจทย์ปัญหาทางคณิตศาสตร์จะช่วยให้นักเรียนมีแนวทางการคิดที่หลากหลาย มีนิสัยกระตือรือร้น และมีความมั่นใจในการแก้ปัญหาที่เผชิญอยู่ทั้งภายในและภายนอกห้องเรียน เป็นทักษะพื้นฐานที่นักเรียนสามารถนำไปใช้ในชีวิตประจำวันได้ [2] นอกจากนี้ การแก้ปัญหาทางคณิตศาสตร์ยังเป็นกระบวนการทางความคิดที่จะนำความรู้ ความสามารถทางคณิตศาสตร์มาผสมผสานกับประสบการณ์ในการแก้โจทย์ปัญหาที่ผู้แก้โจทย์ปัญหามีอยู่ ช่วยให้การแก้โจทย์มีระบบระเบียบ เป็นขั้นตอน และทำให้การแก้โจทย์ปัญหามีประสิทธิภาพยิ่งขึ้น [3]

จากการสังเกตพฤติกรรมการเรียนรู้ การตอบคำถาม และการแสดงวิธีทำเกี่ยวกับการแก้โจทย์ปัญหาของนักเรียนในการจัดการเรียนรู้รายวิชา ค31102 คณิตศาสตร์พื้นฐาน เรื่อง หลักการนับเบื้องต้น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนหาดใหญ่วิทยาลัย ภาคเรียนที่ 2 ปีการศึกษา 2567 พบว่า นักเรียนส่วนใหญ่สามารถทำความเข้าใจและแก้โจทย์ที่มีลักษณะตรงไปตรงมาได้เป็นอย่างดี โดยเฉพาะโจทย์ที่มีรูปแบบใกล้เคียงกับตัวอย่างที่เรียน อย่างไรก็ตาม เมื่อนำโจทย์ปัญหาที่มีความซับซ้อนมากขึ้นมาใช้ในการประเมินผล นักเรียนจะมีปัญหาเกี่ยวกับ

การวิเคราะห์โจทย์ปัญหา กระบวนการแก้ปัญหาไม่เป็นลำดับขั้นตอน ตลอดจนผลลัพธ์ที่ได้ไม่สมเหตุสมผล นักเรียนส่วนใหญ่ตอบคำถามในชั้นเรียนโดยเลือกวิธีแก้ปัญหาตามนักเรียนคนอื่นในชั้นเรียน โดยขาดการพิจารณาปัญหาและวิธีแก้ปัญหาอย่างรอบคอบ ไม่มีการเชื่อมโยงความรู้เดิมกับสถานการณ์ใหม่ และขาดการวางแผนการแก้ปัญหาอย่างเป็นระบบ

ผู้วิจัยจึงเห็นความสำคัญของการจัดการเรียนรู้ที่มุ่งเน้นกระบวนการคิดในการแก้ปัญหาอย่างเป็นระบบ เพื่อพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์ จากการศึกษาวิธีการจัดการเรียนรู้ที่ส่งเสริมกระบวนการคิด การจัดการเรียนรู้จากอดีตจนถึงปัจจุบันมีรูปแบบที่หลากหลาย รูปแบบการจัดการเรียนรู้มีความสำคัญต่อการส่งเสริมให้นักเรียนสามารถพัฒนาตนเองเต็มตามศักยภาพ การเลือกรูปแบบการจัดการเรียนรู้ที่มีความชัดเจนเหมาะสมกับนักเรียนจะช่วยให้นักเรียนสามารถเรียนรู้และพัฒนาตนเองได้ ในการจัดการเรียนรู้ควรเน้นให้นักเรียนได้ฝึกทักษะกระบวนการคิด ทักษะการแสวงหาความรู้ด้วยตนเองจากแหล่งเรียนรู้ที่หลากหลาย การฝึกปฏิบัติและการประยุกต์ความรู้สู่การปฏิบัติจริง [4] การจัดการเรียนรู้ด้วยวิธีการแบบเปิด (Open Approach) มีผลต่อการพัฒนาการสอนคณิตศาสตร์ของญี่ปุ่นและพัฒนาผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียน โดยเน้นกระบวนการแก้ปัญหาที่มีลักษณะเป็นปัญหาปลายเปิด ซึ่งเป็นสื่อที่ให้นักเรียนพัฒนาความรู้ความเข้าใจ ทักษะการคิด และทักษะการแก้ปัญหา เป้าหมายของการจัดการเรียนรู้ด้วยวิธีการแบบเปิด คือ มุ่งเน้นให้นักเรียนทุกคนเรียนคณิตศาสตร์ด้วยความสามารถของแต่ละคนด้วยตนเอง นักเรียนสามารถพัฒนาหรือสร้างผลงานทางคณิตศาสตร์ และกระบวนการเรียนรู้ของตนเองอย่างมีคุณภาพและประสิทธิภาพ ในการจัดการเรียนรู้ครูจำเป็นต้องทำความเข้าใจแนวคิดของนักเรียนให้มากที่สุดเพื่อเป็นแนวทางสนับสนุน กระตุ้น และจัดประสบการณ์การเรียนรู้ให้นักเรียนได้พัฒนาการเรียนรู้ได้อย่างเต็มความสามารถ [5] นอกจากนี้ยังพบว่า การจัดการเรียนรู้ด้วยวิธีการแบบเปิดสามารถพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ตลอดจนพัฒนาทักษะอื่น ๆ ในวิชาคณิตศาสตร์ เป็นการจัดการเรียนรู้ที่นักเรียนจะได้คิดวางแผนและแก้ปัญหาด้วยตนเองอย่างหลากหลาย หรืออาจกล่าวได้ว่าการจัดการเรียนรู้ด้วยวิธีการแบบเปิดเป็นแนวทางในการจัดการเรียนรู้หนึ่งที่จะช่วยสนับสนุนความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ตลอดจนการพัฒนาทักษะอื่น ๆ ในวิชาคณิตศาสตร์และทำให้ผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์สูงขึ้น [6]

จากที่กล่าวมาข้างต้น ทำให้ผู้วิจัยสนใจที่จะนำการจัดการเรียนรู้ด้วยวิธีการแบบเปิดซึ่งประกอบด้วย 4 ขั้นตอนหลัก ได้แก่ 1) การนำเสนอปัญหา ซึ่งเป็นการกระตุ้นให้นักเรียนตั้งคำถามและพยายามหาวิธีแก้ไข 2) การสำรวจและแก้ปัญหา ที่เปิดโอกาสให้นักเรียนใช้แนวทางของตนเองในการแก้ปัญหา 3) การอภิปรายและเปรียบเทียบแนวคิด ซึ่งช่วยให้นักเรียนได้เรียนรู้แนวทางที่หลากหลายจากเพื่อน และ 4) การสรุปและเชื่อมโยงความรู้ ที่ช่วยให้นักเรียนสร้างความเข้าใจที่ลึกซึ้งยิ่งขึ้น มาใช้ในการพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ในการจัดการเรียนรู้รายวิชา ค31102 คณิตศาสตร์พื้นฐาน เรื่อง ความน่าจะเป็น นักเรียนชั้นมัธยมศึกษาปีที่ 4/11 โรงเรียนหาดใหญ่วิทยาลัย

2 วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ให้มีคะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์เฉลี่ยไม่น้อยกว่าร้อยละ 70 ของคะแนนเต็ม และมีจำนวนนักเรียนที่ผ่านเกณฑ์อย่างน้อยร้อยละ 70 ของจำนวนนักเรียนทั้งหมด

2. เพื่อพัฒนาผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ให้มีคะแนนผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์เฉลี่ยไม่น้อยกว่าร้อยละ 70 ของคะแนนเต็ม และมีค่าดัชนีประสิทธิผลไม่น้อยกว่า 0.5

3 วิธีดำเนินการวิจัย

การวิจัยเชิงปฏิบัติการในชั้นเรียน (Classroom Action Research) มีรายละเอียดดังต่อไปนี้

3.1 รูปแบบการวิจัย

การวิจัยในครั้งนี้เป็นการวิจัยเชิงปฏิบัติการในชั้นเรียน โดยใช้หลักการและขั้นตอนการวิจัยเชิงปฏิบัติการตามแนวคิดของ Kemmis & McTaggard [7] จำนวน 3 วงจร ซึ่งแต่ละวงจรปฏิบัติการประกอบด้วย 4 ขั้นตอน ได้แก่ 1) ขั้นวางแผน (Plan) 2) ขั้นปฏิบัติการ (Act) 3) ขั้นสังเกต (Observe) และ 4) ขั้นสะท้อนผลการปฏิบัติ (Reflect)

3.2 กลุ่มเป้าหมาย

กลุ่มเป้าหมาย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 4/11 ซึ่งเป็นนักเรียนในโครงการพลังสิบ ที่มุ่งพัฒนาสมรรถนะของนักเรียน ทางด้านวิทยาศาสตร์ คณิตศาสตร์ และเทคโนโลยีผ่านกระบวนการของหลักสูตร โรงเรียนหาดใหญ่วิทยาลัย สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาสงขลา สตูล ภาคเรียนที่ 2 ปีการศึกษา 2567 จำนวน 39 คน เป็นนักเรียนชาย 10 คน และนักเรียนหญิง 29 คน ประกอบด้วย นักเรียนที่มีผลการเรียนในระดับเก่ง 19 คน ปานกลาง 17 คน และอ่อน 3 คน โดยพิจารณาจากผลการเรียนรายวิชาคณิตศาสตร์พื้นฐาน ในภาคเรียนที่ 1 ปีการศึกษา 2567 ข้อมูลดังแสดงในตารางที่ 1 ต่อไปนี้

ตารางที่ 1: ระดับความสามารถของนักเรียนชั้นมัธยมศึกษาปีที่ 4/11 โรงเรียนหาดใหญ่วิทยาลัย จำแนกโดยผลการเรียนรายวิชาคณิตศาสตร์พื้นฐาน ในภาคเรียนที่ 1 ปีการศึกษา 2567

ระดับ	เก่ง	ปานกลาง	อ่อน	รวม
ผลการเรียน	2.00 - 2.50	2.51-3.49	3.50-4.00	
นักเรียนหญิง (คน)	16	10	3	29
นักเรียนชาย (คน)	3	7	-	10
รวม	19	17	3	39

3.3 ตัวแปรในการวิจัย

ตัวแปรต้น คือ การจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ตามแนวคิดของไมตรี อินทร์ประสิทธิ์

ตัวแปรตาม ได้แก่ ความสามารถในการแก้ปัญหาทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์

3.4 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยประกอบด้วย

3.4.1 แผนจัดการเรียนรู้ เรื่อง ความน่าจะเป็น โดยใช้การจัดการเรียนรู้ด้วยวิธีการแบบเปิด จำนวน 3 แผน โดยใช้เวลาในการจัดการเรียนรู้ทั้งหมด 8 คาบ คาบละ 50 นาที

3.4.2 แบบทดสอบวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ เรื่อง ความน่าจะเป็น เป็นแบบทดสอบอัตนัย จำนวน 2 ข้อ วัดความสามารถใน 4 องค์ประกอบ ได้แก่ ความเข้าใจปัญหา การเลือกวิธีการแก้ปัญหา การใช้วิธีการแก้ปัญหา และการสรุปคำตอบ กำหนดเกณฑ์การให้คะแนนโดยใช้รูบริค โดยกำหนดการให้คะแนนในแต่ละองค์ประกอบ เป็น 0 1 2 และ 3 ข้อละ 12 คะแนน รวม 24 คะแนน

3.4.3 แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง ความน่าจะเป็น เป็นแบบทดสอบปรนัยชนิดเติมคำตอบ จำนวน 10 ข้อ แต่ละข้อ ตอบถูกได้ 0.5 คะแนน ตอบผิดได้ 0 คะแนน รวม 5 คะแนน และแบบทดสอบอัตนัย จำนวน 5 ข้อ กำหนดเกณฑ์การให้คะแนนรายข้อ ข้อละ 2 คะแนน รวม 10 คะแนน คะแนนรวม 15 คะแนน

3.5 การวิเคราะห์ข้อมูล

3.5.1 วิเคราะห์ข้อมูลจากแบบทดสอบความสามารถในการแก้ปัญหาทางคณิตศาสตร์ เรื่อง ความน่าจะเป็น หลังการจัดการเรียนรู้ด้วยวิธีการแบบเปิด โดยการหาค่าเฉลี่ยและร้อยละ จากคะแนนของแบบทดสอบ แล้วนามาเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ คือ มีคะแนนเฉลี่ยไม่น้อยกว่าร้อยละ 70 ของคะแนนเต็ม และมีจำนวนนักเรียนที่ผ่านเกณฑ์ร้อยละ 70 ขึ้นไป ของจำนวนนักเรียนทั้งหมด

3.5.2 วิเคราะห์ข้อมูลจากแบบทดสอบผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง ความน่าจะเป็น ก่อนและหลังการจัดการเรียนรู้โดยใช้การจัดการเรียนรู้ด้วยวิธีการแบบเปิด โดยการหาค่าเฉลี่ยและร้อยละ ของคะแนนแบบทดสอบ แล้วนามาเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ คือ มีคะแนนเฉลี่ยไม่น้อยกว่าร้อยละ 70 ของคะแนนเต็ม รวมทั้งหาค่าดัชนีประสิทธิผลเทียบเกณฑ์ไม่น้อยกว่า 0.5 เพื่อพิจารณาพัฒนาการในการเรียนรู้ที่เพิ่มขึ้นของนักเรียน เมื่อ ดัชนีประสิทธิผล = (ผลรวมของคะแนนหลังเรียนทุกคน - ผลรวมของคะแนนก่อนเรียนทุกคน)/(ผลรวมของคะแนนเต็มทุกคน - ผลรวมของคะแนนก่อนเรียนทุกคน)

4 ผลการวิจัย

หลังจากการดำเนินการวิจัยตามวงจรปฏิบัติการเสร็จสิ้นตามแผนการจัดการเรียนรู้ที่ 1-3 ที่ได้กำหนดไว้ ผู้วิจัยได้นำเสนอผลการวิเคราะห์ข้อมูลตามลำดับ ดังนี้

1. ผลการพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4

ตารางที่ 2: ผลการทดสอบความสามารถในการแก้ปัญหาทางคณิตศาสตร์

จำนวนนักเรียน	คะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์				ส่วนเบี่ยงเบนมาตรฐาน	ร้อยละของคะแนนเฉลี่ย	จำนวนนักเรียนที่ผ่านเกณฑ์	
	เต็ม	สูงสุด	ต่ำสุด	เฉลี่ย			จำนวน (คน)	ร้อยละ
39	24	24	13	19.46	4.17	81.08	31	79.49

จากตารางที่ 2 ผลการทดสอบความสามารถในการแก้ปัญหาทางคณิตศาสตร์ หลังการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 พบว่า นักเรียนมีคะแนนเฉลี่ยเท่ากับ 19.46 คิดเป็นร้อยละ 81.08 ของคะแนนเต็ม และมีนักเรียนที่ผ่านเกณฑ์จำนวน 31 คน คิดเป็นร้อยละ 79.49 ของนักเรียนทั้งหมด ซึ่งสูงกว่าเกณฑ์ที่กำหนด

ตารางที่ 3: ผลการทดสอบความสามารถในการแก้ปัญหาทางคณิตศาสตร์รายองค์ประกอบ

ความสามารถในการแก้ปัญหาทางคณิตศาสตร์	จำนวนนักเรียน	คะแนน		ส่วนเบี่ยงเบนมาตรฐาน	ร้อยละของคะแนนเฉลี่ย	จำนวนนักเรียนที่ผ่านเกณฑ์	
		เต็ม	เฉลี่ย			จำนวน (คน)	ร้อยละ
ความเข้าใจปัญหา	39	6	6.00	0.00	100.00	39	100.00
การเลือกวิธีการแก้ปัญหา	39	6	4.74	1.43	79.00	28	71.79
การใช้วิธีการแก้ปัญหา	39	6	4.71	1.38	78.50	25	64.10
การสรุปคำตอบ	39	6	4.00	1.78	66.67	16	41.03

จากตารางที่ 3 ผลการทดสอบความสามารถในการแก้ปัญหาทางคณิตศาสตร์รายองค์ประกอบ หลังการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 พบว่า องค์ประกอบที่นักเรียนมีร้อยละของคะแนนเฉลี่ยสูงสุด คือ ความเข้าใจปัญหา มีคะแนนเฉลี่ยเท่ากับ 6 คิดเป็นร้อยละ 100 รองลงมา คือ การเลือกวิธีการแก้ปัญหา มีคะแนนเฉลี่ยเท่ากับ 4.74 คิดเป็นร้อยละ 79 การใช้วิธีการแก้ปัญหา มีคะแนนเฉลี่ยเท่ากับ 4.71 คิดเป็นร้อยละ 78.50 และองค์ประกอบที่นักเรียนมีร้อยละของคะแนนเฉลี่ยต่ำที่สุด คือ การสรุปคำตอบ มีคะแนนเฉลี่ยเท่ากับ 4.00 คิดเป็นร้อยละ 66.67 ตามลำดับ ทั้งนี้มีองค์ประกอบที่มีจำนวนนักเรียนที่ผ่านเกณฑ์มากที่สุด คือ ความเข้าใจปัญหา นักเรียนผ่านเกณฑ์ทุกคนจำนวน 39 คน รองลงมา คือ การเลือกวิธีการแก้ปัญหา มีนักเรียนที่ผ่านเกณฑ์จำนวน 28 คน คิดเป็นร้อยละ 71.79 การใช้วิธีการแก้ปัญหา มีนักเรียนที่ผ่านเกณฑ์จำนวน 25 คน คิดเป็นร้อยละ 64.10 และการสรุปคำตอบ มีนักเรียนที่ผ่านเกณฑ์จำนวน 16 คน คิดเป็นร้อยละ 41.03 ตามลำดับ ทั้งนี้ นักเรียนมีร้อยละของคะแนนเฉลี่ยผ่านเกณฑ์มากกว่าร้อยละ 70 ของคะแนนเต็ม ใน 3 องค์ประกอบ ได้แก่ ความเข้าใจปัญหา การเลือกวิธีการแก้ปัญหา และการใช้วิธีการแก้ปัญหา และมีจำนวนนักเรียนที่ผ่านเกณฑ์มากกว่าร้อยละ 70 ของจำนวนนักเรียนทั้งหมด ใน 2 องค์ประกอบ ได้แก่ ความเข้าใจปัญหาและการเลือกวิธีการแก้ปัญหา

2. ผลการพัฒนาผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4

ตารางที่ 4: ผลการทดสอบผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์

ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์	จำนวนนักเรียน	คะแนนผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์		ส่วนเบี่ยงเบนมาตรฐาน	ร้อยละของคะแนนเฉลี่ย	ดัชนีประสิทธิผล
		เต็ม	เฉลี่ย			
ก่อนเรียน	39	15	3.23	1.71	21.53	0.63
หลังเรียน	39	15	10.62	2.87	70.80	

จากตารางที่ 4 ผลการทดสอบผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ก่อนและหลังการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 พบว่า นักเรียนมีคะแนนเฉลี่ยก่อนเรียนและหลังเรียนเท่ากับ 3.23 และ 10.62 คิดเป็นร้อยละ 21.53 และ 70.80 ของคะแนนเต็ม ตามลำดับ โดยนักเรียนมีคะแนนผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์เฉลี่ยหลังเรียนสูงกว่าเกณฑ์ที่กำหนด นอกจากนี้ ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนมีค่าดัชนีประสิทธิผลเท่ากับ 0.63 ซึ่งสูงกว่าเกณฑ์ที่กำหนด แสดงให้เห็นพัฒนาการในการเรียนรู้ของนักเรียนที่มีความรู้เพิ่มขึ้น 0.63 หรือคิดเป็นร้อยละ 63

4 สรุปผลและอภิปรายผล

จากผลการพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 สรุปผลการวิจัยได้ดังต่อไปนี้

1. นักเรียนมีคะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์เฉลี่ยเท่ากับ 19.46 คิดเป็นร้อยละ 81.08 ของคะแนนเต็ม และมีจำนวนนักเรียนที่ผ่านเกณฑ์ 31 คน คิดเป็นร้อยละ 79.49 ของจำนวนนักเรียนทั้งหมด ซึ่งสูงกว่าเกณฑ์ที่กำหนด เมื่อพิจารณาคะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์รายองค์ประกอบ พบว่า นักเรียนมีร้อยละของคะแนนเฉลี่ยผ่านเกณฑ์มากกว่าร้อยละ 70 ของคะแนนเต็ม ใน 3 องค์ประกอบ ได้แก่ ความเข้าใจปัญหา การเลือกวิธีการแก้ปัญหา และการใช้วิธีการแก้ปัญหา โดยองค์ประกอบที่นักเรียนมีร้อยละของคะแนนเฉลี่ยสูงสุด คือ ความเข้าใจปัญหา คิดเป็นร้อยละ 100 รองลงมา คือ การเลือกวิธีการแก้ปัญหา คิดเป็นร้อยละ 79 และการใช้วิธีการแก้ปัญหา คิดเป็นร้อยละ 78.50 ตามลำดับ สำหรับองค์ประกอบที่นักเรียนมีร้อยละของคะแนนเฉลี่ยต่ำที่สุด คือ การสรุปคำตอบ คิดเป็นร้อยละ 66.67 ทั้งนี้ มีจำนวนนักเรียนที่ผ่านเกณฑ์มากกว่าร้อยละ 70 ของจำนวนนักเรียนทั้งหมด ใน 2 องค์ประกอบ ได้แก่ ความเข้าใจปัญหาและการเลือกวิธีการแก้ปัญหา โดยองค์ประกอบที่มีจำนวนนักเรียนที่ผ่านเกณฑ์มากที่สุด คือ ความเข้าใจปัญหา นักเรียนผ่านเกณฑ์ทุกคน คิดเป็นร้อยละ 100 รองลงมา คือ การเลือกวิธีการแก้ปัญหา มีจำนวนนักเรียนที่ผ่านเกณฑ์ คิดเป็นร้อยละ 71.79 นอกจากนี้ยังมีอีก 2 องค์ประกอบ ได้แก่ การใช้วิธีการแก้ปัญหา มีจำนวนนักเรียนที่ผ่านเกณฑ์ คิดเป็นร้อยละ 64.10 และการสรุปคำตอบ มีจำนวนนักเรียนที่ผ่านเกณฑ์ คิดเป็นร้อยละ 41.03 ซึ่งมีจำนวนนักเรียนที่ผ่านเกณฑ์น้อยกว่าร้อยละ 70 ของจำนวนนักเรียนทั้งหมด

2. นักเรียนมีคะแนนผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์เฉลี่ยเท่ากับ 10.62 คิดเป็นร้อยละ 70.80 ของคะแนนเต็ม และมีค่าดัชนีประสิทธิผลเท่ากับ 0.63 ซึ่งสูงกว่าเกณฑ์ที่กำหนด แสดงให้เห็นพัฒนาการในการเรียนรู้ของนักเรียนที่มีความรู้เพิ่มขึ้น 0.63 หรือคิดเป็นร้อยละ 63

ในการวิจัยครั้งนี้เป็นการพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์โดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด เรื่อง ความน่าจะเป็น ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 จากกระบวนการวิจัยเชิงปฏิบัติการและผลการวิจัย ผู้วิจัยนำเสนอการอภิปรายผลการการวิจัย ดังต่อไปนี้

1. นักเรียนมีคะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์เฉลี่ยเท่ากับ 19.46 คิดเป็นร้อยละ 81.08 ของคะแนนเต็ม และมีจำนวนนักเรียนที่ผ่านเกณฑ์ 31 คน คิดเป็นร้อยละ 79.49 ของจำนวนนักเรียนทั้งหมด ซึ่งสูงกว่าเกณฑ์ที่กำหนด ทั้งนี้เนื่องมาจากนักเรียนในกลุ่มเป้าหมายได้รับการฝึกทักษะในการแก้ปัญหาทางคณิตศาสตร์และฝึกปฏิบัติตามการจัดการเรียนรู้ด้วยวิธีการแบบเปิด อย่างต่อเนื่องตามแผนการจัดการเรียนรู้ จำนวน 3 แผนการจัดการเรียนรู้ รวม 8 ชั่วโมง อีกทั้งยังพบว่า นักเรียนมีคะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์รายองค์ประกอบ โดยองค์ประกอบที่นักเรียนมีร้อยละของคะแนนเฉลี่ยสูงสุด คือ ความเข้าใจปัญหา คิดเป็นร้อยละ 100 รองลงมา คือ การเลือกวิธีการแก้ปัญหา คิดเป็นร้อยละ 79 การใช้วิธีการแก้ปัญหา คิดเป็นร้อยละ 78.50 และ

องค์ประกอบที่นักเรียนมีร้อยละของคะแนนเฉลี่ยต่ำที่สุด คือ การสรุปคำตอบ คิดเป็นร้อยละ 66.67 ตามลำดับ ทั้งนี้ การจัดการเรียนรู้ด้วยวิธีการแบบเปิดจะช่วยให้นักเรียนได้มีส่วนร่วมและพัฒนาตนเองในหลาย ๆ ด้าน โดยเฉพาะอย่างยิ่งในด้านการคิด ซึ่งเป็นการคิดที่หลากหลาย อีกทั้งในด้านการอภิปรายแลกเปลี่ยนเรียนรู้ และมีความคิดสร้างสรรค์จากการแก้ปัญหา เนื่องจากการจัดการเรียนรู้ด้วยวิธีการแบบเปิดเน้นการสร้างสถานการณ์ปัญหาให้เป็นปลายเปิด ซึ่งสอดคล้องกับแนวคิดของ Becker & Shimada [8] ที่ว่าวิธีการแบบเปิดว่าเป็นวิธีการที่ผู้เรียนมีส่วนร่วมในปัญหาที่กำหนดให้มีคำตอบที่ถูกต้องอย่างหลากหลาย เรียกว่า ปัญหาปลายเปิด (Open-Ended Problem) เพื่อเตรียมประสบการณ์การจัดการเรียนรู้ในการค้นหาความรู้หรือสิ่งใหม่ ๆ โดยผ่านการรวบรวมความรู้ ทักษะ หรือวิธีคิดที่ได้รับจากการเรียนรู้ก่อนหน้านี้ของผู้เรียนเอง กระบวนการแก้ปัญหาทางคณิตศาสตร์จึงมีความสำคัญอย่างยิ่ง ในการที่จะพัฒนาผู้เรียนให้มีแนวทางการคิดที่เป็นขั้นตอน มีทักษะที่สามารถนำไปใช้ในการแก้ปัญหาในชีวิตประจำวันได้ ซึ่งเป็นกระบวนการที่มีลำดับขั้นตอนชัดเจนสามารถช่วยให้นักเรียนได้พัฒนาความสามารถในการแก้ปัญหา คิดวางแผนการแก้ปัญหา การตรวจสอบผลที่ได้ ฝึกการทำงานที่มีขั้นตอนและเป็นระบบ จากที่พบว่า องค์ประกอบที่นักเรียนมีร้อยละของคะแนนเฉลี่ยสูงสุด คือ ความเข้าใจปัญหา ซึ่งในองค์ประกอบนี้สามารถบอกความเข้าใจของนักเรียนจากสถานการณ์ปัญหาที่กำหนด โดยผู้วิจัยได้สร้างสถานการณ์ปัญหาแบบปลายเปิดให้นักเรียนมีแนวคิดในการหาคำตอบที่หลากหลาย อีกทั้งผู้วิจัยมีการใช้คำถามเพื่อเป็นการกระตุ้นให้นักเรียนได้เกิดการคิด การวิเคราะห์ปัญหา จึงทำให้นักเรียนสามารถเข้าใจสถานการณ์ปัญหาต่าง ๆ ได้ดี คำถามที่ใช้เริ่มจากสิ่งที่เป็นพื้นฐานเพื่อสร้างความสนใจ กระตุ้นให้เกิดความสงสัยเชื่อมโยงสู่การตรวจสอบทำความเข้าใจปัญหา คำถามเกี่ยวกับประสบการณ์ คำถามที่ให้จำแนกแยกแยะ จัดกลุ่ม หรือเปรียบเทียบ ส่งเสริมให้นักเรียนวิเคราะห์ ขยายความ โดยการอธิบายแนวคิดของข้อมูลต่าง ๆ รวมทั้งยังเป็นการตั้งคำถามเพื่อตรวจสอบ เพื่อให้นักเรียนได้คิดทบทวน ตรวจสอบคำตอบ และตอบคำถามโดยให้ข้อมูลเพิ่มเติม ซึ่งสอดรับกับแนวคิดของจุฑาทิพย์ หาญกุดตุ้ม และคณะ [9] ที่ว่าการใช้คำถามเป็นฐานในการเรียนรู้เป็นการกระตุ้นความความสนใจของผู้เรียนได้ทั้งทางตรงและทางอ้อมทำให้ผู้เรียนได้คิด วิเคราะห์สถานการณ์ปัญหา และยังพบว่าองค์ประกอบที่นักเรียนมีร้อยละของคะแนนเฉลี่ยต่ำที่สุด คือ การสรุปคำตอบ เนื่องจากเป็นองค์ประกอบสุดท้ายของการแก้ปัญหาทางคณิตศาสตร์ และนักเรียนบางคนยังขาดความเข้าใจในการตรวจสอบคำตอบหรือผลสรุปได้อย่างถูกต้อง ทั้งนี้ ความถูกต้องของการสรุปคำตอบ ยังเป็นผลมาจากองค์ประกอบก่อนหน้าทั้งการเลือกวิธีการแก้ปัญหา และการใช้วิธีการแก้ปัญหา ที่จะต้องผ่านการคิดและดำเนินการอย่างเป็นลำดับขั้นตอนอย่างถูกต้อง ซึ่งความเข้าใจที่ถูกต้องในการตรวจสอบความสมเหตุสมผลและผลสรุปที่ถูกต้องย่อมส่งผลต่อการมองย้อนกลับเพื่อตรวจสอบองค์ประกอบก่อนหน้าที่ส่งผลต่อความถูกต้องของคำตอบ

2. นักเรียนมีคะแนนผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์เฉลี่ยเท่ากับ 10.62 คิดเป็นร้อยละ 70.80 ของคะแนนเต็ม และมีค่าดัชนีประสิทธิผลเท่ากับ 0.63 ซึ่งสูงกว่าเกณฑ์ที่กำหนด แสดงให้เห็นพัฒนาการในการเรียนรู้ของนักเรียนที่มีความรู้เพิ่มขึ้น 0.63 หรือคิดเป็นร้อยละ 63 ทั้งนี้เนื่องมาจากนักเรียนในกลุ่มเป้าหมายได้รับการฝึกทักษะในการแก้ปัญหาทางคณิตศาสตร์และฝึกปฏิบัติโดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิดอย่างต่อเนื่องตามแผนการจัดการเรียนรู้ จำนวน 3 แผนการจัดการเรียนรู้ รวม 8 ชั่วโมง การจัดการเรียนรู้ด้วยวิธีการแบบเปิด ประกอบด้วย 4

ขั้นตอน [5] ได้แก่ 1) ขั้นนำเสนอปัญหาปลายเปิด ในขั้นนี้ครูจะนำเสนอสถานการณ์ปัญหาให้กับนักเรียน เพื่อให้นักเรียนเกิดความสนใจในการเรียนและตระหนักว่าสถานการณ์ปัญหาดังกล่าวเป็นปัญหาของตนเองที่จะต้องแก้ปัญหาให้สำเร็จ 2) ขั้นเรียนรู้ด้วยตนเองของนักเรียน ในขั้นนี้นักเรียนลงมือแก้ปัญหาด้วยตนเองโดยการแก้ปัญหาสามารถกระทำเป็นรายบุคคลหรือเป็นกลุ่มก็ได้ ทำให้นักเรียนเกิดการเรียนรู้ด้วยตนเอง นักเรียนสามารถใช้วิธีการต่าง ๆ ที่หลากหลายในการคิดเพื่อหาคำตอบด้วยตนเอง ครูมีหน้าที่คอยอำนวยความสะดวกให้กับนักเรียนและคอยสังเกตพฤติกรรมของนักเรียนขณะดำเนินการแก้ปัญหา 3) ขั้นอภิปรายทั้งชั้นเรียนและเปรียบเทียบแนวคิดของนักเรียน ในขั้นนี้นักเรียนจะนำเสนอวิธีการคิดแก้ปัญหาของตนเองหน้าชั้นเรียนให้เพื่อนและครูได้ร่วมกันอภิปรายและตรวจสอบความถูกต้อง โดยครูมีหน้าที่เลือกลำดับการนำเสนอของนักเรียนโดยพิจารณาลำดับการนำเสนอจากการแก้ปัญหาขั้นพื้นฐานไปยังการแก้ปัญหาขั้นสูง การอภิปรายทั้งชั้นเรียนทำให้นักเรียนอธิบายแนวคิดของตนเองอีกครั้งหนึ่ง ทำให้นักเรียนมีการตรวจสอบและทบทวนแนวคิดของตนเองที่เกิดขึ้น และ 4) ขั้นสรุปโดยเชื่อมโยงแนวคิดของนักเรียน ในขั้นนี้ครูเป็นผู้สรุปความรู้ เพื่อเชื่อมโยง แนวคิดของนักเรียนและทำให้นักเรียนเกิดแนวคิดความเข้าใจในเรื่องที่สอนคาบนั้น ซึ่งสอดคล้องกับเป้าหมายของบทเรียนที่ต้องการให้เกิดขึ้นกับนักเรียน รวมถึงส่งผลให้นักเรียนได้ตระหนักถึงวิธีการต่าง ๆ ในการแก้ปัญหาที่นำมาสู่แนวคิดที่สำคัญ จะเห็นได้ว่าการจัดการเรียนรู้ด้วยวิธีการแบบเปิดบทบาทสำคัญจะอยู่ที่ผู้เรียน โดยครูเป็นเพียงผู้ที่คอยอำนวยความสะดวกในชั้นเรียน นักเรียนทุกคนในชั้นเรียนมีส่วนร่วมในกิจกรรม มีการแลกเปลี่ยนความรู้ซึ่งกันและกันภายในกลุ่ม บรรยากาศในการเรียนการสอนไม่ตึงเครียด การจัดการเรียนรู้ด้วยวิธีการแบบเปิดจะใช้สถานการณ์ปัญหาปลายเปิดเป็นจุดเริ่มต้นของการจัดการเรียนรู้ ซึ่งลักษณะของความเปิดจะมี 3 ลักษณะ ได้แก่ 1) กระบวนการเปิด ปัญหาประเภทนี้มีวิธีการที่หลากหลายในการแก้ปัญหา 2) ผลลัพธ์เปิด ปัญหาประเภทนี้มีคำตอบที่หลากหลาย และ 3) แนวทางการพัฒนาปัญหาเปิด หลังจากที่ผู้เรียนแก้ปัญหาแล้วสามารถพัฒนาปัญหาใหม่ ๆ โดยการเปลี่ยนเงื่อนไขหรือใช้เหตุผลจากปัญหาเดิม ซึ่งเป็นไปตามแนวคิดของ ไมตรี อินทร์ประสิทธิ์ [10] ที่ว่า ปัญหาปลายเปิดมีลักษณะเป็นสถานการณ์ปัญหา มีทั้งคำตอบที่หลากหลาย มีกระบวนการแก้ปัญหาที่หลากหลาย และสามารถพัฒนาไปเป็นปัญหาอื่นได้ ซึ่งลักษณะเด่นดังกล่าวนี้ทำให้นักเรียนที่มีความสามารถแตกต่างกันในชั้นเรียนสามารถแก้ปัญหาได้ตามความถนัดและความสนใจของตนเอง หรือกล่าวได้ว่าปัญหาปลายเปิด เปิดโอกาสให้นักเรียนจำนวนมากสามารถแก้ปัญหาได้อย่างเต็มศักยภาพ เพราะนักเรียนสามารถสร้างปัญหาของตนได้จากสถานการณ์ปัญหาปลายเปิดที่กำหนดให้ และเปิดโอกาสให้นักเรียนสามารถมีประสบการณ์ในการแก้ปัญหาแต่ละครั้ง ซึ่งเป็นการเพิ่มความเชื่อมั่นของนักเรียนจำนวนมากว่า ทุกคนสามารถเรียนคณิตศาสตร์ได้ นอกจากนี้ยังสอดคล้องกับการศึกษาของมัญชุพร กัลยาสนธิ์ และคณะ [11] ที่พบว่าผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนโดยการจัดการเรียนรู้ด้วยวิธีการแบบเปิด หลังเรียนสูงกว่าเกณฑ์ร้อยละ 70 ซึ่งเกิดจากนักเรียนได้มีประสบการณ์ที่หลากหลายในการแก้ปัญหาปลายเปิด มีแนวคิดที่หลากหลายที่เกิดจากการคิดร่วมกันในกลุ่ม ส่งผลให้เกิดการเรียนรู้อย่างมีประสิทธิภาพ ทั้งนี้ ปวีณ ไวยโภคา และคณะ [12] ยังได้ศึกษาผลการจัดกิจกรรมการเรียนรู้โดยใช้วิธีการแบบเปิดร่วมกับเทคนิคเพื่อนคู่คิดที่มีต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์และความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 พบว่า ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์และ

ความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 หลังได้รับการจัดกิจกรรมการเรียนรู้โดยใช้วิธีการแบบเปิดร่วมกับเทคนิคเพื่อนคู่คิดสูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เนื่องมาจากการจัดการเรียนรู้ด้วยวิธีการแบบเปิดในการวิจัยครั้งนี้เป็นการจัดกิจกรรมการเรียนรู้โดยการกำหนดปัญหาปลายเปิดขึ้นมา เพื่อกระตุ้นความคิดของนักเรียน ให้คิดหาคำตอบได้อย่างกว้างขวาง ทำให้นักเรียนได้เปิดกว้างทางความคิด สามารถค้นพบคำตอบได้ด้วยวิธีการของตนเอง ทำให้นักเรียนรู้สึกท้าทายในการคิดหาคำตอบและไม่เครียดจนเกินไปเพราะนักเรียนมีอิสระทางความคิด อีกทั้งการนำเทคนิคเพื่อนคู่คิดมาช่วยส่งเสริมการจัดกิจกรรมการเรียนรู้ด้วยวิธีการแบบเปิดทำให้นักเรียนได้มีส่วนร่วมในการปฏิบัติกิจกรรม กล้าแสดงความคิดเห็น แลกเปลี่ยนความรู้ความคิดและคำตอบของตนเองกับคู่ของตน นำไปสู่การปรับปรุงและพัฒนาวิธีการเรียนรู้ให้ดียิ่งขึ้น ดังนั้น จากการศึกษาข้างต้นจึงกล่าวได้ว่า การจัดการเรียนรู้ด้วยวิธีการแบบเปิดทำให้ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนมีค่าเพิ่มขึ้น การจัดการเรียนรู้บรรลุตามวัตถุประสงค์ อีกทั้ง การจัดการเรียนรู้ด้วยวิธีการแบบเปิดส่งผลให้นักเรียนมีคะแนนความสามารถในการแก้ปัญหาทางคณิตศาสตร์อยู่ในระดับสูงกว่าเกณฑ์ที่กำหนดไว้ จึงทำให้นักเรียนมีผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์สูงกว่าเกณฑ์ที่กำหนดด้วย

กิตติกรรมประกาศ ผู้แต่งขอขอบคุณผู้ทรงคุณวุฒิทุกท่านที่ได้ให้ข้อคิดเห็นและข้อเสนอแนะต่าง ๆ เพื่อปรับปรุงบทความวิจัยนี้

เอกสารอ้างอิง

[1] กระทรวงศึกษาธิการ, ตัวชี้วัดและสาระการเรียนรู้แกนกลางกลุ่มสาระการเรียนรู้คณิตศาสตร์ (ฉบับปรับปรุง 2560) ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551, โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย, กรุงเทพมหานคร, 2560.

[2] สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, ทักษะ/กระบวนการทางคณิตศาสตร์, โรงพิมพ์คุรุสภาลาดพร้าว, กรุงเทพมหานคร, 2560.

[3] สุประวีณ์ สังข์ทอง, มนตรี วงษ์สะพาน, การพัฒนาความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 โดยใช้การจัดการเรียนรู้แบบสืบเสาะหาความรู้ ร่วมกับชุดฝึกทักษะการคิดแก้ปัญหาแบบร่วมมือ เรื่อง ความน่าจะเป็น, วารสารมหาจุฬานาครทรรศน์. 7 (9) (2563) 222--235.

[4] ขวัญชัย ขัวนา, ธารทิพย์ ขัวนา, การจัดการเรียนการสอนเพื่อส่งเสริมทักษะการเรียนรู้ในศตวรรษที่ 21, วารสารบัณฑิตศึกษา. 16 (73) (2562) 13--22.

[5] ไมตรี อินทร์ประสิทธิ์, กระบวนการแก้ปัญหาในคณิตศาสตร์ระดับโรงเรียน Processes of Problem Solving in School Mathematics (พิมพ์ครั้งที่ 2), ศูนย์วิจัยคณิตศาสตร์ศึกษา มหาวิทยาลัยขอนแก่น, ขอนแก่น, 2565.

[6] ปรีชา เนาว์เย็นผล, กิจกรรมการเรียนการสอนคณิตศาสตร์ โดยใช้สถานการณ์ปัญหาปลายเปิด สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1, (ปริญญานิพนธ์ดุษฎีบัณฑิต), มหาวิทยาลัยศรีนครินทรวิโรฒ, 2554.

[7] S. Kemmis, R. McTaggart, The Action Research Planner (3rd ed), Midwest, California, 1999.

[8] J. P. Becker, S. Shimada, The Open-Ended Approach: A New Proposal for Teaching Mathematics, National Council of Teachers of Mathematics, Virginia, 1997.

[9] จุฑาทิพย์ หาญกุดตุ้ม, สมภารธัชธรณ์ ศิโลศรีไช, สัญญา เคณาภูมิ, การจัดการเรียนรู้โดยใช้คำถามเป็นฐาน, วารสารราชภัฏสุรินทร์วิชาการ. 1 (5) (2566) 67--82.

[10] ไมตรี อินทร์ประสิทธิ์, การสอนโดยใช้วิธีการแบบเปิดในชั้นเรียนญี่ปุ่น, วารสารศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น. 1 (1) (2547) 1--1.

[11] มัญชุพร กัลยาสนธิ์, ธัญญลักษณ์ เขจรภักดิ์, ณัฏฐชัย จันทชุม, การพัฒนาผลสัมฤทธิ์ทางการเรียนรู้คณิตศาสตร์โดยใช้วิธีการสอนแบบเปิด (Open Approach) ของนักเรียนชั้นประถมศึกษาปีที่ 3, วารสารวิชาการศึกษาศาสตร์ ศรีนครินทรวิโรฒ. 23 (1) (2565) 19--36.

[12] ปวีณ ไวยโภคา, เวชฤทธิ์ อังกนะภัทรขจร, ผลาดร สุวรรณโพธิ์, ผลการจัดกิจกรรมการเรียนรู้โดยใช้วิธีการแบบเปิด (Open Approach) ร่วมกับเทคนิคเพื่อนคู่คิด (Think Pair Share) ทีมีต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์และความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่4, วารสารวิจัยทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ. 17 (2) (2565) 105--117.

การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ ที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียน และเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 วิทยาลัยเทคนิคมีนบุรี

อภิสิทธิ์ กินรีวงศ์[1,*] และ ภาสุดา ภาคาผล[1,†]

[1]ภาควิชาหลักสูตรและการสอน คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง 10240

บทคัดย่อ

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อ (1) เปรียบเทียบผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติ ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษากับเกณฑ์ร้อยละ 70 และ (2) ศึกษาเจตคติของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่มีต่อการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติ หลังได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา โดยใช้กลุ่มตัวอย่างเป็นนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 แผนกเมคคาทรอนิกส์และหุ่นยนต์ ที่กำลังศึกษาอยู่ในภาคเรียนที่ 2 ปีการศึกษา 2567 วิทยาลัยเทคนิคมีนบุรี กรุงเทพมหานคร จำนวน 31 คน ที่ได้มาโดยการสุ่มแบบกลุ่ม เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ประกอบด้วย (1) แผนการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา จำนวน 3 แผน (2) แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติ และ (3) แบบสอบถามเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ วิเคราะห์ข้อมูลด้วยค่าเฉลี่ยเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน และสถิติ T – Test Dependent

ผลการวิจัยพบว่า (1) นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ มีผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ผ่านเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ (2) นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษามีเจตคติต่อการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติ ในภาพรวมอยู่ในระดับดี ($\bar{x}$ = 4.44 และ S.D. = 0.66)

คำสำคัญ: การจัดการเรียนรู้, สะเต็มศึกษา, ผลสัมฤทธิ์ทางการเรียน, เจตคติ

2020 MSC: 97D40, 97D50

*ผู้นำเสนอ (apisitkin1@gmail.com)

†ผู้ประพันธ์บรรณกิจ(pasuda.p@rumail.ru.ac.th)

1 บทนำ

ปัจจุบันประเทศไทยให้ความสำคัญกับการพัฒนาทรัพยากรมนุษย์ จึงได้มีการกำหนดเกี่ยวกับคุณลักษณะคุณภาพที่พึงประสงค์เพื่อให้สถานศึกษาทุกแห่งยึดเป็นกรอบสำหรับสร้างคนไทย 4.0 ในมาตรฐานการศึกษาของชาติที่เน้นการ *"ธำรงความเป็นไทยและแข่งขันได้ในเวทีโลก"* โดยคุณลักษณะหนึ่ง คือ การส่งเสริมให้พลเมืองมีความสามารถเป็นผู้ร่วมสร้างนวัตกรรม (Innovative co-creator) เพื่อเพิ่มโอกาสและมูลค่าให้กับตนเองและสังคม [1] สอดคล้องกับแผนการศึกษาแห่งชาติ พุทธศักราช 2560 – 2579 ที่กำหนดวิสัยทัศน์ว่า *"คนไทยทุกคนได้รับการศึกษาและเรียนรู้ตลอดชีวิตอย่างมีคุณภาพ ดำรงชีวิตอย่างเป็นสุข สอดคล้องกับหลักปรัชญาของเศรษฐกิจพอเพียง และการเปลี่ยนแปลงของโลกศตวรรษที่ 21"* เนื่องด้วยการเปลี่ยนแปลงในด้านต่าง ๆ ทั้งในระดับประเทศและนานาชาติที่เต็มไปด้วยปัญหาที่มีความซับซ้อน ความรู้ทางด้านวิทยาศาสตร์ เทคโนโลยี วิศวกรรมศาสตร์ และคณิตศาสตร์ ถือเป็นศาสตร์ที่มีความสำคัญและเป็นเครื่องมือหนึ่งในการพัฒนาพลเมืองให้มีคุณภาพและมีทักษะการเรียนรู้ที่สำคัญที่จำเป็นต่อการดำรงชีวิตในศตวรรษที่ 21 จึงเกิดแนวทางในการจัดการเรียนรู้ที่เรียกว่า *"สะเต็ม"* ขึ้นมา [2]

สะเต็มศึกษา หมายถึง แนวทางในการจัดการเรียนรู้ที่มีการบูรณาการความรู้ทางด้านวิทยาศาสตร์ เทคโนโลยี วิศวกรรมศาสตร์ และคณิตศาสตร์ โดยเน้นการนำความรู้ไปใช้แก้ปัญหาในชีวิตจริง รวมทั้งการพัฒนากระบวนการและผลผลิตใหม่ที่เป็นประโยชน์ต่อการดำเนินชีวิต ยิ่งไปกว่านี้ ผู้เรียนยังได้เรียนรู้ทักษะสำคัญบางประการที่ไม่ใช่ทักษะเฉพาะในวิทยาศาสตร์ เทคโนโลยี วิศวกรรมศาสตร์ และคณิตศาสตร์ แต่เป็นทักษะสากลที่สามารถนำไปใช้ได้ในการแก้ปัญหาที่มีลักษณะเป็นพหุวิชาอีกด้วย [3] ซึ่งสอดคล้องกับ Stohlmann, et al. [4] ที่กล่าวถึงการจัดการเรียนรู้แบบสะเต็มศึกษาว่าเป็นการเรียนรู้แบบบูรณาการที่ผสานความรู้และทักษะหลากหลายด้านผ่านการทำกิจกรรมที่เน้นการลงมือปฏิบัติ (activity-based learning) หรือการทำโครงงาน (project-based learning) ซึ่งช่วยให้นักเรียนได้พัฒนาทักษะการคิด ทักษะการใช้เทคโนโลยีสารสนเทศ ทักษะการแก้ปัญหา และทักษะการสื่อสาร นอกจากนี้ผู้เรียนยังเกิดความเข้าใจในความรู้แบบองค์รวมที่สามารถเชื่อมโยงหรือประยุกต์ใช้ในชีวิตประจำวันได้ โดยสะเต็มศึกษามีความสำคัญต่อความสำเร็จในอนาคตของนักเรียน

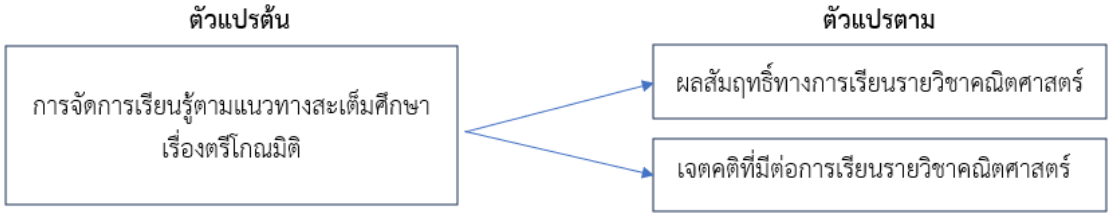
คณิตศาสตร์มีบทบาทสำคัญอย่างยิ่งต่อความสำเร็จในการเรียนรู้ในศตวรรษที่ 21 เนื่องจากคณิตศาสตร์ช่วยพัฒนาความคิดอย่างมีเหตุผล เป็นระบบ มีแบบแผน และเสริมสร้างความคิดริเริ่มสร้างสรรค์ อีกทั้ง ยังส่งเสริมทักษะการวิเคราะห์ปัญหาอย่างรอบคอบ ช่วยในการคาดการณ์ วางแผน ตัดสินใจ และแก้ปัญหาได้เหมาะสม ซึ่งสามารถนำไปประยุกต์ใช้ในชีวิตจริงได้อย่างมีประสิทธิภาพ [5] อีกทั้ง ความรู้ทางคณิตศาสตร์เป็นความรู้ที่อยู่คู่กับชีวิตมนุษย์ตั้งแต่ตื่นนอนจนถึงก่อนเข้านอน มนุษย์ได้ใช้คณิตศาสตร์ทั้งที่ตั้งใจและไม่ตั้งใจ ทั้งที่รู้ตัวและไม่รู้ตัว [6] ประกอบกับรายงานผลการสอบของสำนักงานวิชาการและมาตรฐานการศึกษา ที่ระบุถึงปัญหาผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ของนักเรียนว่าเป็นปัญหาที่ต้องได้รับการแก้ไขอย่างเร่งด่วนอีกปัญหาหนึ่ง จากผลการประเมินคุณภาพภายนอกสถานศึกษาระดับการศึกษาขั้นพื้นฐานของสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา พบผลการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 วิชาคณิตศาสตร์

ตั้งแต่ปี 2561-2566 มีคะแนนเฉลี่ยเท่ากับ 30.04, 26.73, 25.46, 24.47, 24.39 และ 25.38 ตามลำดับ ซึ่งจะเห็นได้ว่าคะแนนเฉลี่ยวิชาคณิตศาสตร์ของนักเรียน มีคะแนนเฉลี่ยอยู่ในระดับที่ต่ำ

อนึ่ง เจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ก็มีความสำคัญสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนของนักเรียน เพราะเจตคติเป็นสิ่งที่คอยกำหนดการแสดงออกด้านพฤติกรรมของนักเรียนว่าจะไปในทิศทางบวกหรือทิศทางลบขึ้นอยู่กับประสบการณ์ ความเข้าใจ หรือความคาดหวังที่นักเรียนมีต่อรายวิชาคณิตศาสตร์ ซึ่งจะช่วยเสริมสร้างทักษะที่จำเป็นในชีวิตและการทำงานในอนาคต และยังช่วยให้นักเรียนสามารถเรียนรู้และพัฒนาความสามารถด้านคณิตศาสตร์ได้อย่างมีประสิทธิภาพ ซึ่งสอดคล้องกับการศึกษาของเปรียบฟ้า ด้วงนุ่ม [7] ได้สร้างกิจกรรมการเรียนรู้ตามแนวทางสะเต็มศึกษาโดยใช้แนวคิด เรื่องพาราโบลา สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 ศึกษาความสามารถในการเรียนรู้ตามแนวทางสะเต็มศึกษาโดยใช้แนวคิด เรื่องพาราโบลา ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 และศึกษาเจตคติที่มีต่อกิจกรรมการเรียนรู้ตามแนวสะเต็มศึกษาโดยใช้แนวคิด เรื่องพาราโบลาของนักเรียนชั้นมัธยมศึกษาปีที่ 5 ผลการวิจัยพบว่า นักเรียนชั้นมัธยมศึกษาปีที่ 5 ที่เรียนด้วยกิจกรรมการเรียนรู้ตามแนวสะเต็มศึกษาโดยใช้แนวคิด เรื่องพาราโบลา มีความสามารถในการเรียนรู้ตามแนวสะเต็มศึกษาโดยใช้แนวคิดเรื่องพาราโบลา ไม่น้อยกว่าร้อยละ 65 ของคะแนนรวม มีจำนวนมากกว่าร้อยละ 65 ของจำนวนนักเรียนทั้งหมดที่ระดับนัยสำคัญ .01 และมีเจตคติต่อกิจกรรมการเรียนรู้ตามสะเต็มศึกษาโดยใช้แนวคิด เรื่องพาราโบลา อยู่ในเกณฑ์ดี

จากความสำคัญของปัญหาที่กล่าวมาข้างต้น ผู้วิจัยจึงมีความสนใจนำการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา มาใช้ในการพัฒนาผลสัมฤทธิ์ทางการเรียนและเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 วิทยาลัยเทคนิคมีนบุรี ที่มีจุดเด่นในด้านการจัดการเรียนรู้เพื่อช่วยพัฒนาทักษะกระบวนการคิด ทักษะการใช้เทคโนโลยีสารสนเทศ ทักษะการแก้ปัญหา และทักษะการสื่อสาร ซึ่งทักษะดังกล่าวเป็นทักษะการเรียนรู้ในศตวรรษที่ 21 ยิ่งไปกว่านั้น การใช้การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษาเรื่องตรีโกณมิติ จะเน้นให้ผู้เรียนมีบทบาทสำคัญต่อการดำเนินกิจกรรมการเรียนรู้ของตนเอง สามารถช่วยกระตุ้นความสนใจของนักเรียน ทำให้นักเรียนมีความตื่นเต้นและเกิดความสนุกสนานกับการเรียนรู้ เข้าใจเนื้อหา เห็นคุณค่าและประโยชน์ของเนื้อหาผ่านการทดลอง การเขียน การอภิปรายรายการแก้ปัญหา อันนำไปสู่การประยุกต์ใช้ความรู้ไปใช้ในสถานการณ์หรือปัญหาที่กำหนดให้เกิดประสิทธิภาพสูงสุดต่อไปได้

การศึกษาการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ ที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนและเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 วิทยาลัยเทคนิคมีนบุรี ผู้วิจัยได้ศึกษาทฤษฎี แนวคิด และงานวิจัยที่เกี่ยวข้องเพื่อกำหนดกรอบแนวคิดของการวิจัยในครั้งนี้ โดยมีกรอบแนวคิด ดังนี้



รูปที่ 1: กรอบแนวคิดของการวิจัย

1.1 วัตถุประสงค์ของการวิจัย

1) เพื่อเปรียบเทียบผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษากับเกณฑ์ร้อยละ 70

2) เพื่อศึกษาเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 หลังได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา

1.2 สมมติฐานของการวิจัย

1) นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษามีผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติ ผ่านเกณฑ์มากกว่าร้อยละ 70

2) นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 หลังได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษามีเจตคติต่อการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติ อยู่ในระดับดี

2 การทบทวนวรรณกรรม

การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เป็นการจัดการเรียนการสอนที่มีการบูรณาการกระบวนการออกแบบเชิงวิศวกรรมเข้ากับการเรียนรู้ใน 3 สาขาวิชา ได้แก่ วิชาวิทยาศาสตร์ คณิตศาสตร์ และเทคโนโลยีของนักเรียน เพื่อพัฒนาความรู้ความเข้าใจและฝึกกระบวนการทางวิทยาศาสตร์ คณิตศาสตร์ และเทคโนโลยี โดยมีจุดมุ่งหมายคือการพัฒนาทักษะที่จำเป็นสำหรับศตวรรษที่ 21 เช่น การคิดวิเคราะห์ การแก้ปัญหา ความคิดสร้างสรรค์ และการทำงานร่วมกัน ส่งเสริมให้ผู้เรียนนำความรู้ไปประยุกต์ใช้ในชีวิตจริงผ่านกิจกรรมเชิงปฏิบัติและการทำโครงการ ยิ่งไปกว่านั้น สะเต็มศึกษายังมีความสำคัญในการเตรียมพลเมืองให้พร้อมสำหรับอาชีพในสาขาที่เกี่ยวข้องกับวิทยาศาสตร์และเทคโนโลยี ซึ่งเป็นที่ต้องการในตลาดงานปัจจุบัน [8, 9] ประกอบด้วยขั้นตอน 6 ขั้นตอน ได้แก่ (1) ขั้นระบุปัญหา (Problem Identification) นักเรียนร่วมกันศึกษาสถานการณ์ วิเคราะห์และทำความเข้าใจปัญหา ภายใต้เงื่อนไขหรือข้อจำกัดที่กำหนด (2) ขั้นรวบรวมข้อมูลและแนวคิดที่เกี่ยวข้องกับปัญหา (Related Information Search) นักเรียนสืบค้นข้อมูลที่เกี่ยวข้องกับปัญหาจากแหล่งข้อมูลต่าง ๆ (3) ขั้นออกแบบวิธีการแก้ปัญหา (Solution Design) นักเรียนต้องเขียนร่วมกันร่างต้นแบบ กำหนดขั้นตอนการทำงาน เลือกใช้วัสดุอุปกรณ์ในการออกแบบชิ้นงาน สื่อหรือแหล่งการเรียนรู้ เพื่อช่วยในการแก้ปัญหา (4) ขั้นวางแผนและดำเนินการแก้ปัญหา (Planning and Development) นักเรียนดำเนินการงานตามลำดับขั้นตอนของการสร้างชิ้นงานหรือวิธีการใช้ในการแก้ปัญหาที่ออกแบบไว้ในขั้นที่ 3 (5) ขั้นทดสอบ ประเมินผล และปรับปรุงแก้ไขชิ้นงานหรือวิธีการแก้ปัญหา (Testing, Evaluation and Design Improvement) นักเรียนร่วมกันทดสอบ ปรับปรุง และแก้ไขชิ้นงานหรือวิธีการเพื่อให้ได้ผลที่แม่นยำมากขึ้น และ (6) ขั้นนำเสนอวิธีการแก้ปัญหา ผลการแก้ปัญหา หรือชิ้นงาน (Presentation) นักเรียนนำเสนอแนวคิดและขั้นตอนการแก้ปัญหาของการสร้างชิ้นงานหรือวิธีการแก้ปัญหา ให้ผู้อื่นเข้าใจและได้ข้อเสนอแนะ [10, 11]

ผลสัมฤทธิ์ทางการเรียน เป็นความรู้ความสามารถ ที่สามารถวัดได้ในหลายด้านของนักเรียน โดยความรู้ความสามารถดังกล่าวอาจเกิดจากนักเรียนได้ศึกษาเรียนรู้ ประสบการณ์ และการเปลี่ยนแปลงพฤติกรรม ซึ่งสามารถจำแนกผลสัมฤทธิ์ทางการเรียนจากการจัดกิจกรรมการเรียนรู้ กระบวนการจัดการเรียนรู้ตามความเหมาะสม โดยครูผู้สอนต้องศึกษาแนวทางในการสร้างเครื่องมือวัดให้มีคุณภาพ สอดคล้องตามวัตถุประสงค์ และด้านที่ต้องการวัดในทำนองเดียวกัน ผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ จะเป็นความรู้ คุณลักษณะเมื่อนักเรียนเกิดการเรียนรู้ ซึ่งประกอบไปด้วยความรู้ ความเข้าใจ ความสามารถในการคิดคำนวณและความสามารถในการแก้ปัญหา การใช้สูตรและหลักการทางคณิตศาสตร์ หรือการเชื่อมโยงทฤษฎีต่าง ๆ ในชีวิตประจำวัน โดยผลสัมฤทธิ์นี้สามารถวัดได้จากการทดสอบ การสอบ การทำการบ้าน หรือกิจกรรม ที่ช่วยพัฒนาความรู้และทักษะที่จำเป็นของนักเรียน ดังนั้น การจัดการเรียนรู้ วิธีการจัดการเรียนรู้ สื่อ กระบวนการทางคณิตศาสตร์ ทำให้เกิดแรงจูงใจ รวมถึงการจัดกิจกรรมการเรียนรู้โดยใช้เทคนิควิธีการที่หลากหลายและมีประสิทธิภาพของครูผู้สอนนั้น สามารถทำให้นักเรียนมีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ที่สูงขึ้น [12]

เจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ เป็นพฤติกรรมที่แสดงออกเพื่อจะตอบสนองต่อรายวิชาคณิตศาสตร์ อันมีผลมาจากความรู้สึกของนักเรียนที่มีต่อการเรียนรายวิชาคณิตศาสตร์ ซึ่งสามารถแสดงออกมาให้เห็นได้ทั้งพฤติกรรมทางบวกและพฤติกรรมทางลบ เจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์อาจเปลี่ยนแปลงได้เมื่อนักเรียนได้รับประสบการณ์ใหม่ที่แตกต่างจากเดิม การจัดการเรียนรู้รายวิชาคณิตศาสตร์ในปัจจุบัน ครูผู้สอนต้องมีการพิจารณาด้านเจตคติร่วมไปกับด้านวิชาการหรือการให้ความรู้ในรายวิชา สถาบันส่งเสริมการจัดการเรียนรู้วิทยาศาสตร์และเทคโนโลยี [13] กล่าวว่า นักเรียนแสดงพฤติกรรมที่จะตอบสนองต่อวิชาคณิตศาสตร์ เป็นผลมาจากเจตคติต่อการเรียนวิชาคณิตศาสตร์ ซึ่งส่งผลให้ความรู้สึกของนักเรียนที่มีต่อวิชาคณิตศาสตร์ รวมทั้ง ความพร้อมหรือไม่พร้อมที่จะเรียนวิชาคณิตศาสตร์ การจัดกิจกรรมการเรียนการสอนที่แปลกใหม่ ส่งเสริมให้นักเรียนมีปฏิสัมพันธ์ในชั้นเรียน มีส่วนร่วมในการคิดการทำ [8]

3 วิธีดำเนินการวิจัย

3.1 ประชากรและกลุ่มตัวอย่าง

3.1.1 ประชากรที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่กำลังศึกษาอยู่ในภาคเรียนที่ 2 ปีการศึกษา 2567 วิทยาลัยเทคนิคมีนบุรี กรุงเทพมหานคร จำนวนนักเรียน 485 คน

3.1.2 กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 แผนกเมคคาทรอนิกส์ที่กำลังศึกษาอยู่ในภาคเรียนที่ 2 ปีการศึกษา 2567 วิทยาลัยเทคนิคมีนบุรี กรุงเทพมหานคร จำนวนนักเรียน 31 คน โดยใช้วิธีการสุ่มแบบแบ่งกลุ่ม (Cluster Random Sampling) ซึ่งการจัดกลุ่มเรียนแต่ละกลุ่มเป็นแบบคละความสามารถของนักเรียน

3.2 เครื่องมือที่ใช้ในการวิจัย

3.2.1 แผนการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา จำนวน 3 แผน ประกอบด้วย

1) แผนการจัดการเรียนรู้ที่ 1 เรื่อง มุมและการวัด จำนวน 4 ชั่วโมง

2) แผนการจัดการเรียนรู้ที่ 2 เรื่อง อัตราส่วนตรีโกณมิติ จำนวน 2 ชั่วโมง

3) แผนการจัดการเรียนรู้ที่ 3 เรื่อง ฟังก์ชันตรีโกณมิติของวงกลมหนึ่งหน่วย จำนวน 4 ชั่วโมง

ในแต่ละแผนการจัดการเรียนรู้จะประกอบด้วยขั้นตอน 6 ขั้นตอน ได้แก่ ขั้นระบุปัญหา ขั้นรวบรวมข้อมูลและแนวคิดที่เกี่ยวข้องกับปัญหา ขั้นออกแบบวิธีการแก้ปัญหา ขั้นวางแผนและดำเนินการแก้ปัญหา ขั้นทดสอบ ประเมินผลและปรับปรุงแก้ไขชิ้นงานหรือวิธีการแก้ปัญหา และขั้นนำเสนอวิธีการแก้ปัญหา ผลการแก้ปัญหา หรือชิ้นงาน โดยผู้วิจัยได้นำแผนการจัดการเรียนรู้ทั้งหมดให้ผู้เชี่ยวชาญด้านการจัดการเรียนรู้คณิตศาสตร์และการวัดผล จำนวน 3 ท่าน พิจารณาเกี่ยวกับความชัดเจนและความเหมาะสมของแต่ละแผนการจัดการเรียนรู้มาปรับปรุงแก้ไข จากนั้นทำการทดลองใช้กับนักเรียนที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 15 คน และตรวจสอบความเหมาะสม ปรับปรุงแก้ไขอีกครั้ง เพื่อนำไปใช้กับกลุ่มตัวอย่างในการวิจัยต่อไป

3.2.2 แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ เรื่อง ตรีโกณมิติ ซึ่งเป็นแบบทดสอบแบบปรนัย ชนิด 4 ตัวเลือก รวมทั้งหมด 30 ข้อ ที่สร้างขึ้นตามตารางวิเคราะห์จุดประสงค์การเรียนรู้เพื่อวัดพฤติกรรมด้านพุทธิพิสัย 3 ระดับ คือ ความรู้ความจำ (Computation) ความเข้าใจ (Comprehension) และการประยุกต์ (Application) ที่มีค่าดัชนีของความสอดคล้อง (Index of item - Objective Congruence: IOC) ตั้งแต่ 0.66 – 1.00 ค่าความยากง่าย (p) อยู่ระหว่าง 0.25 – 0.60 ค่าจำแนก (r) อยู่ระหว่าง 0.20 – 0.35 และมีค่าความเชื่อมั่น (Reliability) เท่ากับ 0.94

3.2.3 แบบสอบถามเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ ประกอบด้วย 3 ด้าน ได้แก่ ด้านความรู้สึกต่อการเรียนรายวิชาคณิตศาสตร์ ด้านความรู้สึกต่อการจัดการเรียนการสอน และด้านความรู้สึกต่อผู้สอน ด้านละ 10 ข้อ รวมทั้งหมด 30 ข้อ และใช้แบบวัดเจตคติแบบแบ่งช่วงสเกลของลิเคิร์ท (Likert) ซึ่งถูกแบ่งออกเป็น 5 ระดับ ได้แก่ ดีมาก ดี ปานกลาง น้อย และน้อยมาก โดยผู้เชี่ยวชาญด้านการจัดการเรียนรู้หรือการวัดผลทางการศึกษาหรือจิตวิทยาการศึกษา จำนวน 3 ท่าน พิจารณาเกี่ยวกับความเหมาะสมด้านภาษา

3.3 การเก็บรวบรวมข้อมูล

ผู้วิจัยใช้เวลาในการดำเนินการทดลอง เป็นเวลาทั้งหมด 10 ชั่วโมง และใช้เวลาในทดสอบหลังเรียน เป็นเวลา 1 ชั่วโมง ซึ่งรายละเอียดการดำเนินการทดลองมีดังนี้

1) ผู้วิจัยชี้แจงให้ผู้เรียนทราบถึงการเรียนรู้โดยการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เพื่อที่ผู้เรียนจะปฏิบัติตนได้ถูกต้อง

2) ผู้วิจัยดำเนินการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา ตามแผนการจัดการเรียนรู้ที่ผู้วิจัยสร้างขึ้น จำนวน 3 แผน รวมเวลา 10 ชั่วโมง

3) เมื่อสิ้นสุดการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา ให้นักเรียนทำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนและแบบสอบถามเจตคติต่อการเรียนรายวิชาคณิตศาสตร์ เป็นเวลา 1 ชั่วโมง บันทึกผลเป็นคะแนนหลังเรียน (Posttest) เพื่อวัดผลสัมฤทธิ์ทางการเรียนและเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์

4) นำคะแนนมาวิเคราะห์ข้อมูลทางสถิติและทำการทดสอบสมมติฐาน

3.4 การวิเคราะห์ข้อมูล

การวิจัยครั้งนี้ ผู้วิจัยได้ทำการวิเคราะห์ข้อมูลทางสถิติ ดังนี้

1) นำคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ เรื่อง ตรีโกณมิติ มาวิเคราะห์หาค่าเฉลี่ยเลขคณิต ($\bar{x}$) ค่าร้อยละ และส่วนเบี่ยงเบนมาตรฐาน (S.D.) และเปรียบเทียบกับเกณฑ์ร้อยละ 70 โดยใช้สถิติ T – Test Dependent

2) นำคะแนนจากแบบสอบถามเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ มาวิเคราะห์หาค่าเฉลี่ยเลขคณิต ($\bar{x}$) และส่วนเบี่ยงเบนมาตรฐาน (S.D.)

3.5 ตัวอย่างแผนการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา

ในงานวิจัยนี้ ผู้วิจัยจะขอนำเสนอแผนการจัดการเรียนรู้ที่ 1 เรื่อง มุมและการวัด ซึ่งมีขั้นตอนในการจัดการเรียนรู้ตามแนวสะเต็มศึกษา 6 ขั้นตอน ดังนี้

ขั้นที่ 1 ขั้นระบุปัญหา ครูกระตุ้นให้นักเรียนเกิดความสนใจด้วยคำถามพร้อมกับร่วมกันยกตัวอย่างประกอบและครูเปิดคลิปวิดีทัศน์ให้นักเรียนได้สังเกตและอภิปรายร่วมกัน จากนั้นครูนำเข้าสู่กิจกรรม แบ่งกลุ่มนักเรียนออกเป็นกลุ่ม กลุ่มละ 3 – 4 คน และแจกใบกิจกรรม เรื่องการวัดตำแหน่งของกลุ่มดาวและวัตถุบนฟ้าให้กับนักเรียนทุกคน เพื่อให้นักเรียนแต่ละคนศึกษาหาแนวทางคำตอบของตนเองตามเงื่อนไขที่ครูกำหนด ดังนี้

> ***สถานการณ์ปัญหา:*** ในวันที่ 7 ธันวาคม 2567 “ดาวพฤหัสบดี” ใกล้โลกที่สุดในรอบปี ซึ่งจะโคจรมาอยู่ในตำแหน่งตรงข้ามดวงอาทิตย์ (Jupiter Opposition) หมายถึง ตำแหน่งที่ดวงอาทิตย์ โลก และดาวพฤหัสบดี เรียงอยู่ในแนวเส้นตรงเดียวกัน มีโลกอยู่ตรงกลาง ส่งผลให้ดาวพฤหัสบดีอยู่ในตำแหน่งใกล้โลกที่สุดในรอบปี ห่างจากโลกประมาณ 611.76 ล้านกิโลเมตร ในวันดังกล่าว เมื่อดวงอาทิตย์ลับขอบฟ้า ดาวพฤหัสบดีจะปรากฏสว่างทางทิศตะวันออก สังเกตได้ด้วยตาเปล่า นานตลอดคืนจนถึงรุ่งเช้า
>
> นางสาวดาวเรืองออกไปดูดาวที่ทุ่งหญ้าห่างไกลจากตัวเมือง ซึ่งในคืนนั้นท้องฟ้าใสและไม่มีเมฆปกคลุม และพบว่าดาวที่ส่องสว่างที่สุดในคืนนั้นคือดาวพฤหัสบดี หากดาวเรืองมองท้องฟ้าต่อไปในทิศทางเดียวกันนี้ และต้องการดูดาวในเวลา 19.00 – 24.00 น. จะสามารถมองเห็นกลุ่มดาวต่าง ๆ ได้อย่างไรบ้าง

ขั้นที่ 2 ขั้นรวบรวมข้อมูลและแนวคิดที่เกี่ยวข้องกับปัญหา ครูชวนนักเรียนสนทนาเกี่ยวกับปัญหากระตุ้นนักเรียนตอบคำถาม จากนั้นครูให้นักเรียนศึกษาข้อมูลด้วยตนเองในหนังสือเรียน และสามารถใช้เทคโนโลยีแหล่งสืบค้น สื่อการสอน และอินเทอร์เน็ตในการค้นหาข้อมูลอื่น ๆ ตามโจทย์ปัญหาและเงื่อนไขที่ครูกำหนดไว้

ขั้นที่ 3 ขั้นออกแบบวิธีการแก้ปัญหา ครูให้แต่ละกลุ่มร่วมกันสรุปข้อมูลที่สืบค้นมา และร่วมกันออกแบบแผนที่ดาวลงในใบกิจกรรม

ขั้นที่ 4 ขั้นวางแผนและดำเนินการแก้ปัญหา ครูให้นักเรียนแต่ละกลุ่มวางแผนการออกแบบแผนที่ดาวตามแบบร่างที่ร่างไว้ในใบกิจกรรม และนักเรียนเริ่มทำการสร้างแผนที่ดาวตามแผนการที่วางไว้ ครูจะทำหน้าที่สังเกตพฤติกรรมการทำงานร่วมกันของนักเรียน พร้อมกับให้คำแนะนำอย่างใกล้ชิด

ขั้นที่ 5 ขั้นทดสอบ ประเมินผล และปรับปรุงแก้ไขชิ้นงานหรือวิธีการแก้ปัญหา นักเรียนแต่ละกลุ่มตรวจสอบปรับปรุงแก้ไขแผนที่ดาวของตนเองตามแบบที่ได้ออกแบบไว้แล้ว และนำมาทดสอบตามเงื่อนไขวันและเวลาตามที่ครูได้กำหนด เพื่อทำการสรุป บันทึกผลลัพธ์ที่ได้ ตอบคำถาม และแนวทางแก้ไขลงในใบกิจกรรม

ขั้นที่ 6 ขั้นนำเสนอวิธีการแก้ปัญหา ผลการแก้ปัญหา หรือชิ้นงาน นักเรียนแต่ละกลุ่มนำเสนอผลลัพธ์ที่ได้จากการทำกิจกรรมผ่านโปรแกรม Canva และครูให้นักเรียนสาธิตการแก้ปัญหาตามเงื่อนไขที่ครูกำหนดโดยใช้แผนที่ดาวของกลุ่มตนเอง พร้อมกับร่วมกันสรุปแนวคิดและสิ่งที่ได้จากกิจกรรม

4 ผลการวิเคราะห์ข้อมูล

จากการวิเคราะห์ข้อมูล ผู้วิจัยได้ทำการนำเสนอผลการวิเคราะห์ข้อมูลตามลำดับ ดังนี้

4.1 ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์

4.1.1 ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ กับเกณฑ์ร้อยละ 70 โดยใช้การวิเคราะห์ด้วยสถิติ T – Test Dependent

ตารางที่ 1 ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา

การทดสอบ	*N*	$\bar{x}$	*S.D.*	μ	*t*	*df*	Sig.
คะแนนผลสัมฤทธิ์ทางการเรียน	31	24.68	2.66	21	7.688	30	<0.05*

*มีนัยสำคัญทางสถิติที่ระดับ 0.05

จากตารางที่ 1 พบว่า คะแนนผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ ผ่านเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยนักเรียนที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ มีคะแนนผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์เฉลี่ยเท่ากับ 24.68 ซึ่งผ่านเกณฑ์ที่กำหนดร้อยละ 70 หรือ 21 คะแนน จากคะแนนเต็ม 30 คะแนน ($\bar{x}$ = 24.68, SD = 2.66)

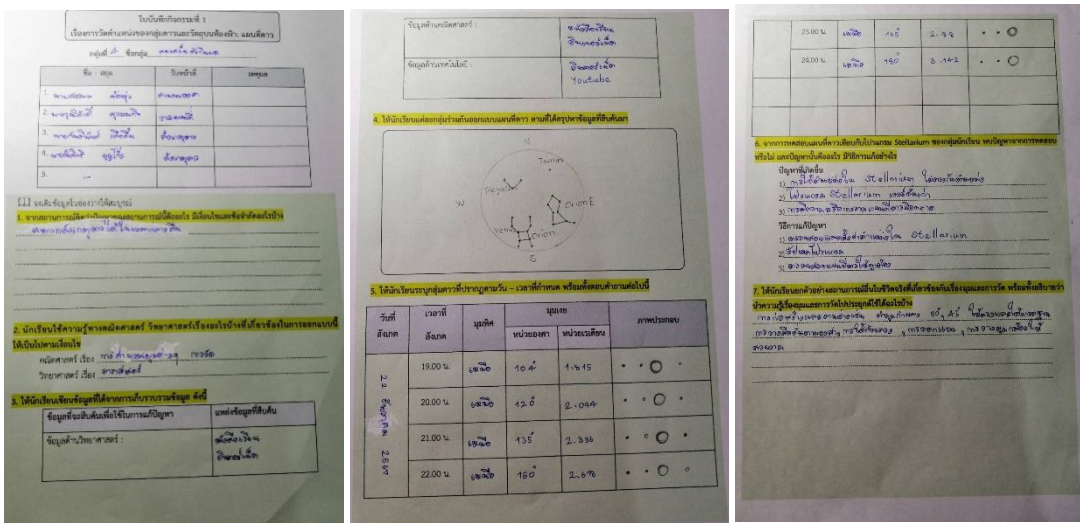
4.1.2 ผลการวิเคราะห์คะแนนใบบันทึกกิจกรรมของนักเรียนที่ได้รับจากการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษารายบุคคล ไปหาคะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

ตารางที่ 2 ผลการวิเคราะห์คะแนนใบบันทึกกิจกรรมของนักเรียนที่ได้รับจากการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษารายบุคคล

แหล่งที่มาของคะแนน	คะแนนเต็ม	คะแนนเฉลี่ย ($\bar{x}$)	คะแนนเฉลี่ย (ร้อยละ)	ส่วนเบี่ยงเบนมาตรฐาน (S.D.)
ใบบันทึกกิจกรรมที่ 1	10	9.55	95.5	0.72
ใบบันทึกกิจกรรมที่ 2	10	9.55	95.5	0.81
ใบบันทึกกิจกรรมที่ 3	10	9.61	96.1	0.62
ภาพรวม	**10**	**9.57**	**95.7**	**0.71**

จากตารางที่ 2 พบว่า คะแนนจากใบบันทึกกิจกรรมของนักเรียนที่ได้รับจากการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา มีคะแนนเฉลี่ยเท่ากับ 28.71 คิดเป็นร้อยละ 95.7 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.51 ผ่านเกณฑ์มากกว่าร้อยละ 70 สรุปได้ว่า นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ มีผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ ผ่านเกณฑ์มากกว่าที่กำหนดร้อยละ 70 หรือ 21 คะแนน จากคะแนนเต็ม 30 คะแนน

ภาพที่ 1 ตัวอย่างใบบันทึกกิจกรรมของนักเรียนที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องมุมและการวัด



จากภาพที่ 1 ตัวอย่างใบบันทึกกิจกรรมที่ 1 เรื่องการวัดตำแหน่งของกลุ่มดาวและวัตถุบนท้องฟ้า: แผนที่ดาว ของนักเรียนที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา แผนการจัดการเรียนรู้ที่ 1 เรื่องมุมและการวัด โดยรวมมีคะแนนเฉลี่ยเท่ากับ 9.55 จากคะแนนเต็ม 10 คะแนน ($\bar{x}$ = 9.55, SD = 0.72)

4.2 ผลการศึกษาเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์

ผลการศึกษาเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ โดยใช้การวิเคราะห์ค่าสถิติบรรยาย ได้แก่ คะแนนเฉลี่ย ($\bar{x}$) ส่วนเบี่ยงเบนมาตรฐาน (S.D.) และการแปลความหมายตามเกณฑ์ที่กำหนดไว้

ตารางที่ 3: ผลการศึกษาเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ของนักเรียนที่ได้รับจากการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ

รายด้าน	$\bar{x}$	S.D.	แปลความหมาย
ด้านความรู้สึกต่อการเรียนรายวิชาคณิตศาสตร์	4.55	0.60	ดีมาก
ด้านความรู้สึกต่อการจัดการเรียนการสอน	4.21	0.70	ดี
ด้านความรู้สึกต่อผู้สอน	4.57	0.62	ดีมาก
ภาพรวม	**4.44**	**0.66**	**ดี**

จากตารางที่ 3 พบว่า คะแนนเฉลี่ยเจตคติที่มีต่อคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ โดยรวมอยู่ในระดับดี ($\bar{x}$ = 4.44, SD = 0.66) และเมื่อพิจารณาแต่ละด้านเรียงจากมากไปน้อย พบว่า นักเรียนมีคะแนนเฉลี่ยเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ด้านความรู้สึกต่อผู้สอนมากที่สุด ($\bar{x}$ = 4.57, SD = 0.62) รองลงมา คือ ด้านความรู้สึกต่อการเรียนรายวิชาคณิตศาสตร์ ($\bar{x}$ = 4.55, SD = 0.60) และด้านความรู้สึกต่อการจัดการเรียนการสอนน้อยที่สุด ($\bar{x}$ = 4.21, SD = 0.70)

5 สรุปผลและอภิปรายผล

จากการวิจัยเรื่อง การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ ที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนและเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 วิทยาลัยเทคนิคมีนบุรี สรุปผลการวิจัยได้ดังนี้

1) นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ มีผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ผ่านเกณฑ์มากกว่าร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2) นักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา เรื่องตรีโกณมิติ มีเจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์อยู่ในระดับดี

โดยผลของการวิจัยสามารถอภิปรายผลมีรายละเอียดดังต่อไปนี้

1) ผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์ เรื่องตรีโกณมิติ ของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 ที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษาผ่านเกณฑ์มากกว่าร้อยละ 70 อย่างมีนัยสำคัญทางสถิติ

ที่ระดับ .05 อาจเนื่องมาจากการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษาเป็นการจัดการเรียนรู้ที่เห็นความสำคัญในการพัฒนาความรู้ความเข้าใจและฝึกทักษะด้านวิทยาศาสตร์ คณิตศาสตร์ และเทคโนโลยี โดยให้นักเรียนได้นำความรู้มาใช้ในการออกแบบวิธีการหรือกระบวนการ เพื่อแก้ปัญหา เพื่อให้ได้เทคโนโลยี ซึ่งเป็นผลผลิตจากกระบวนการออกแบบเชิงวิศวกรรม เป็นลำดับขั้นตอน โดยครูผู้สอนคอยกระตุ้น ชี้แนะแนวทางให้นักเรียนพยายามคิด วิเคราะห์หาทางแก้ปัญหาด้วยตนเอง ให้นักเรียนเป็นผู้ลงมือกระทำมากกว่าเป็นผู้รับสาร สรุปเป็นองค์ความรู้ของตนเองได้ สามารถเชื่อมโยงความรู้และทักษะที่จำเป็นต่อการดำเนินชีวิตในศตวรรษที่ 21

สอดคล้องกับงานวิจัยของวรรณธนะ ปัดชา และสืบสกุล อยู่ยืนยง [14] พบว่า กลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา มีผลสัมฤทธิ์ทางการเรียน เรื่องอัตราส่วนตรีโกณมิติ สูงกว่ากลุ่มที่ได้รับการจัดการเรียนการสอนแบบ สสวท. อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และกลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา หลังเรียนมีทักษะทางด้านสะเต็มศึกษาสูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เนื่องจากการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษาดังกล่าว เป็นกิจกรรมที่เน้นให้นักเรียนเป็นผู้คิด สังเกต ลงมือปฏิบัติจริง และสร้างองค์ความรู้ด้วยตนเองจากโจทย์ปัญหาที่ได้รับ มีกิจกรรมที่หลากหลายทำให้นักเรียนมีความสนใจ กระตือรือร้นในการทำกิจกรรม ในทำนองเดียวกัน กมลฉัตร กล่อมอิ่ม [15] ได้เปรียบเทียบผลสัมฤทธิ์ทางการเรียนของนักเรียนที่เรียนด้วยการจัดการเรียนรู้แบบบูรณาการสะเต็มศึกษากับการจัดการเรียนรู้แบบปกติ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4 พบว่า นักเรียนที่ได้รับการจัดการเรียนรู้แบบบูรณาการสะเต็มศึกษามีผลสัมฤทธิ์ทางการเรียนหลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01 และนักเรียนที่ได้รับการจัดการเรียนรู้แบบบูรณาการสะเต็มศึกษามีผลสัมฤทธิ์ทางการเรียนสูงกว่านักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 โดยนักเรียนได้ค้นหาประสบการณ์ความรู้ ออกแบบวิธีการแก้ปัญหาตามเงื่อนไขที่กำหนดได้ด้วยตนเอง มีครูเป็นผู้ชี้แนะแนวทาง จนสามารถสร้างองค์ความรู้และเชื่อมโยงเนื้อหาหลายวิชาเข้ากับชีวิตจริงได้ นอกจากนี้ยังสอดคล้องกับงานวิจัยของ Lina Fouad Jawad และ Ban Hassan Majeed [16] พบว่า กลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา มีผลสัมฤทธิ์ทางการเรียนรายวิชาคณิตศาสตร์และทักษะด้านความคิดสร้างสรรค์ สูงกว่ากลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้แบบปกติ อย่างมีนัยสำคัญที่ระดับ .05 เนื่องจากการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษาช่วยสร้างบรรยากาศ ความกระตือรือร้น ดึงดูดความสนใจของนักเรียน และช่วยส่งเสริมแรงจูงใจในการเรียน ความคิดสร้างสรรค์ และการสร้างชิ้นงานของนักเรียนได้ จะเห็นว่า การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษามีรูปแบบการเรียนการสอนแบบลงมือปฏิบัติผ่านการเรียนรู้จากสถานการณ์จริง การสร้างสรรค์ชิ้นงานสร้างนวัตกรรมจริง

2) เจตคติที่มีต่อการเรียนรายวิชาคณิตศาสตร์ หลังได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษาเรื่องตรีโกณมิติ อยู่ในระดับดี อาจเนื่องมาจากการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษาส่งเสริมการเรียนรู้ร่วมกัน การจัดกิจกรรม การเรียนการสอนที่แปลกใหม่ ส่งเสริมให้นักเรียนมีปฏิสัมพันธ์ในชั้นเรียน มีส่วนร่วมในการคิด การทำ การแก้ปัญหาตามสถานการณ์ที่กำหนด ส่งผลทำให้นักเรียนเกิดความสนุกกับการเรียนรู้หลังเรียน นอกจากนี้

การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษายังทำให้นักเรียนเข้าใจปัญหาที่เกี่ยวกับคณิตศาสตร์เป็นรูปธรรมมากขึ้น และสามารถเชื่อมโยงเนื้อหาในชีวิตประจำวันได้ มีความสุขเมื่อได้เรียนวิชาคณิตศาสตร์ทุกครั้ง

สอดคล้องกับงานวิจัยของดุสิต ทองสุขนอก และณัฐกานต์ ศาสตร์สูงเนิน [17] พบว่า นักเรียนที่ได้รับการจัดการเรียนรู้ตามแนวคิดสะเต็มศึกษา เรื่อง โยเกิร์ตข้าว มีคะแนนเจตคติต่อวิทยาศาสตร์เฉลี่ยหลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญที่ระดับ .05 เนื่องจากการจัดการเรียนรู้ตามแนวคิดสะเต็มศึกษาเป็นการจัดการเรียนรู้ที่มีความหมายต่อผู้เรียน โดยเฉพาะขั้นตอนการระบุปัญหา ทำให้นักเรียนมองเห็นปัญหา เผชิญปัญหา และแก้ปัญหาโดยคำนึงถึงบริบทแวดล้อมที่สัมพันธ์กับความเป็นจริง ในทำนองเดียวกัน งานวิจัยของวรรณธนะ ปัดชา และสืบสกุล อยู่ยืน ย ง [14] ที่พบว่า กลุ่มนักเรียนที่ได้รับการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษามีความพึงพอใจโดยภาพรวมอยู่ในระดับมาก เนื่องจากการจัดการเรียนรู้ตามแนวทางสะเต็มได้เน้นวิธีการทำงานแบบกลุ่ม และก่อให้เกิดทักษะการแก้ปัญหา และนักเรียนสามารถเชื่อมโยงสาระวิชาต่าง ๆ ในการแก้ปัญหาได้อีกด้วย

6 ข้อเสนอแนะ

ข้อเสนอแนะทั่วไป

1) การจัดการเรียนรู้ตามแนวทางสะเต็มศึกษา ครูผู้สอนจะทำหน้าที่เป็นผู้คอยช่วยเหลือ เป็นแหล่งข้อมูลที่ดี เมื่อนักเรียนมีคำถามจะสามารถให้คำแนะนำได้ ประเมินการทำงานของนักเรียนอย่างใกล้ชิด ให้นักเรียนได้ลงมือปฏิบัติจริง และควรนำปัญหาที่ท้าทายและซับซ้อนมาให้ผู้เรียนได้แก้ไข เช่น ปัญหาการออกแบบที่ต้องใช้การคำนวณทางคณิตศาสตร์ เพื่อให้ผู้เรียนได้ฝึกฝนและพัฒนาทักษะการแก้ปัญหาที่มีความหมายและท้าทาย

2) การจัดการเรียนรู้ควรมีการผสมผสานทั้งเนื้อหาทางคณิตศาสตร์กับการประยุกต์ใช้ในสถานการณ์จริง และกระตุ้นให้นักเรียนมีความสนใจในการเรียนรู้แบบเชื่อมโยงและใช้งานได้จริงในชีวิตประจำวันและในอาชีพ

3) การจัดการเรียนรู้ควรมีอุปกรณ์ต่าง ๆ เพื่อให้มีความหลากหลายทางในแต่ละกลุ่ม ส่งเสริมอิสระทางความคิด และฝึกให้นักเรียนใช้เทคโนโลยีในการสืบค้นข้อมูล

ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

1) ควรศึกษาผลการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษากับเนื้อหาทางคณิตศาสตร์เรื่องอื่น ๆ ที่มีความเกี่ยวข้องกันเพื่อพัฒนาผลสัมฤทธิ์ทางการเรียนและเจตคติต่อการเรียนของนักเรียน เช่น ความน่าจะเป็น แคลคูลัส และสถิติ เป็นต้น

2) ควรศึกษาผลการจัดการเรียนรู้ตามแนวทางสะเต็มศึกษากับตัวแปรที่เกี่ยวข้องอื่น ๆ ในทักษะและกระบวนการทางคณิตศาสตร์ เช่น ความสามารถในการแก้ปัญหา ความคิด สร้างสรรค์ การให้เหตุผล การสื่อสารและการสื่อความหมายทางคณิตศาสตร์ เป็นต้น

กิตติกรรมประกาศ งานวิจัยฉบับนี้ สำเร็จลุล่วงได้อย่างสมบูรณ์เป็นเพราะผู้วิจัยได้รับความกรุณาอย่างยิ่งจากอาจารย์ ดร.ภาสุดา ภาคาผล อาจารย์ที่ปรึกษางานวิจัย ที่กรุณาให้คำแนะนำอย่างดียิ่งและเป็นประโยชน์ต่อการดำเนินการวิจัย ตลอดจนเสียสละเวลาในการตรวจทานอย่างละเอียดจนงานวิจัยฉบับนี้สำเร็จลุล่วงด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้ ขอขอบพระคุณนางปิยลักษณ์ รัตนวิทย์ นางสาวชนิดาภา พิมพ์ลา และนายธนะรัชต์ จำปาวัลย์ ที่กรุณาเป็นผู้เชี่ยวชาญในการตรวจสอบคุณภาพของเครื่องมือที่ใช้ในการวิจัยและให้ข้อเสนอแนะที่เป็นประโยชน์ต่อการดำเนินการวิจัย ส่งผลให้งานวิจัยฉบับนี้สมบูรณ์ยิ่งขึ้น ขอขอบคุณกลุ่มตัวอย่าง ได้แก่ นักเรียนในแผนกช่างอิเล็กทรอนิกส์ และแผนกช่างเมคคาทรอนิกส์ ระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 วิทยาลัยเทคนิคมีนบุรีทุกคน ที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูลในงานวิจัยครั้งนี้เป็นอย่างดี สุดท้ายนี้ คุณค่าและประโยชน์จากงานวิจัยฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาพระคุณของบิดามารดา ที่อบรมบ่มนิสัยให้ผู้วิจัยมีความเข้มแข็งอดทนและไม่ย่อท้อต่ออุปสรรคทั้งปวง รวมทั้งครู อาจารย์ทุกท่านที่ได้อบรมสั่งสอนและประสิทธิ์ประสาทวิชาให้แก่ผู้วิจัย

เอกสารอ้างอิง

[1] สำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ. (2561). มาตรฐานการศึกษาของชาติ พ.ศ. 2561.

[2] สำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ. (2560). แผนการศึกษาแห่งชาติ พ.ศ. 2560 – 2579.

[3] ปัฐมาภรณ์ พิมพ์ทอง. (2566). การจัดการเรียนรู้สะเต็มศึกษา. พิมพ์ครั้งที่ 2.—กรุงเทพฯ: สุนทรฟิล์ม.

[4] Stohlmann, M., Moore, T. J., & Roehrig, G. H. (2012). Considerations for teaching integrated STEM education. Journal of Pre-College Engineering Education Research (J-PEER), 2(1), 4.

[5] กิรณา จิรโชติเดโช, นพพร แหยมแสง, และวรนุช แหยมแสง. (2562). การศึกษาผลสัมฤทธิ์ทางการเรียนและความพึงใจที่มีต่อการจัดการเรียนรู้วิชาคณิตศาสตร์ตามแนวคิดทฤษฎีการสร้างองค์ความรู้ด้วยตนเองเรื่องความน่าจะเป็นของนักเรียนชั้นมัธยมศึกษาปีที่ 6. วารสารครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 47(3), 64 – 81.

[6] อัมพร ม้าคนอง. (2557). คณิตศาสตร์สำหรับครูมัธยม. กรุงเทพฯ: ศูนย์ส่งเสริมการพัฒนานวัตกรรม ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

[7] เปรียบฟ้า ด้วงนุ่ม. (2560). กิจกรรมการเรียนรู้ตามแนวสะเต็มศึกษาโดยใช้แนวคิดเรื่อง พาราโบลา สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5. (วิทยานิพนธ์ กศ.ม. (คณิตศาสตร์ศึกษา)). มหาวิทยาลัยศรีนครินทรวิโรฒ, กรุงเทพฯ.

[8] จินตนา เบ้าสุภี และพรรณวิไล ดอกไม้. (2566). การจัดการเรียนรู้แบบสะเต็มศึกษาเพื่อส่งเสริมความคิดสร้างสรรค์ และเจตคติต่อวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1. วารสารสหวิทยาการวิจัยและวิชาการ, 3(5), 569 – 588.

[9] ชวิศ เชื้อธวัช. (2563). การศึกษาความสามารถในการเชื่อมโยงทางคณิตศาสตร์ เรื่องภาคตัดกรวยของนักเรียนชั้นมัธยมศึกษาปีที่ 5 ผ่านกิจกรรมการเรียนรู้ตามแนวสะเต็มศึกษา. (ปริญญานิพนธ์มหาบัณฑิต). มหาวิทยาลัยศรีนครินทรวิโรฒ, กรุงเทพฯ.

[10] National Research Council. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas: National Academies Press.

[11] สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2558). คู่มืออบรมครูสะเต็มศึกษา. กรุงเทพฯ: สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, กระทรวงศึกษาธิการ.

[12] ปวริศา โคติวงศ์. (2564). การศึกษาผลสัมฤทธิ์ทางการเรียนและเจตคติต่อการเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่ได้รับการจัดการเรียนรู้แบบร่วมมือเทคนิคจิ๊กซอว์ ร่วมกับการใช้ชุดการเรียนรู้แบบอุปนัย. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยศรีนครินทรวิโรฒ, กรุงเทพฯ.

[13] สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2555). เอกสารส่งเสริมความรู้วิชาคณิตศาสตร์ เรื่องพีชคณิต (พิมพ์ครั้งที่4. ed.): กรุงเทพฯ: ไฮเอ็ดพับลิชชิ่ง.

[14] วรรณธนะ ปัดชา และสืบสกุล อยู่ยืนยง. (2559). ผลสัมฤทธิ์ทางการเรียนจากการจัดการเรียนรู้แบบสะเต็มศึกษา เรื่อง อัตราส่วนตรีโกณมิติ. ฉบับภาษาไทย สาขามนุษยศาสตร์ สังคมศาสตร์ และศิลปะ ปีที่ 9(3). 830 – 839.

[15] กมลฉัตร กล่อมอิ่ม. (2559). การจัดการเรียนรู้แบบบูรณาการสะเต็มศึกษาสำหรับนักศึกษาวิชาชีพครู. วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร, 18(4), 334-348.

[16] Lina Fouad Jawad, and Ban Hassan Majeed. (2021). The impact of teaching by using STEM approach in the development of creative thinking and mathematical achievement among the students of the fourth scientific class. International Journal of Interactive Mobile Technologies, 15(13), 172 – 188.

[17] ดุสิต ทองสุขนอก และณัฐกานต์ ศาสตร์สูงเนิน (2565). การพัฒนาผลสัมฤทธิ์ทางการเรียน ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ และเจตคติต่อวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาชั้นปีที่ 3 ด้วยการจัดการเรียนรู้ตามแนวคิดสะเต็มศึกษา เรื่อง โยเกิร์ตข้าว. วารสารบัณฑิตศึกษา มหาวิทยาลัยราชภัฏสกลนคร, 19(84), 128 – 138.

ผลการจัดการเรียนรู้แบบ Concrete – Pictorial – Abstract ร่วมกับเทคนิคเพื่อนคู่คิด ที่ส่งผลต่อความสามารถในการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน

อนุชา เสาวคนธ์[1,*] ธารีรัตน์ ธนัตถ์พาณิชย์[2] และ วราภรณ์ จาตนิล[2]

[1]สาขาวิชาคณิตศาสตร์ศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี 10140

[2]ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี 10140

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษา 1) ความสามารถในการให้เหตุผลทางคณิตศาสตร์ 2) ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ และ 3) พัฒนาการสัมพันธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ การจัดกิจกรรมการเรียนรู้แบบ CPA (Concrete-Pictorial-Abstract) ร่วมกับเทคนิคเพื่อนคู่คิดเป็นการบูรณาการกระบวนการเรียนรู้อย่างเป็นลำดับขั้นจากรูปธรรมสู่นามธรรมควบคู่กับการแลกเปลี่ยนความคิดเห็นระหว่างผู้เรียน ซึ่งเป็นแนวทางหนึ่งในการพัฒนากระบวนการคิดวิเคราะห์ การให้เหตุผล และเสริมสร้างการเรียนรู้แบบมีส่วนร่วม กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนนาหลวง กรุงเทพมหานคร จำนวน 54 คน ได้มาโดยวิธีการสุ่มแบบกลุ่ม เครื่องมือที่ใช้ในการวิจัยประกอบด้วย แผนการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด จำนวน 6 แผน และแบบทดสอบวัดความสามารถในการให้เหตุผลทางคณิตศาสตร์ และแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ก่อนและหลังเรียน เรื่อง การเรียงสับเปลี่ยน ผลการวิจัยพบว่า 1) ความสามารถในการให้เหตุผลทางคณิตศาสตร์ มีคะแนนเฉลี่ยคิดเป็นร้อยละ 76.09 ซึ่งสูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 2) ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์มีคะแนนเฉลี่ยคิดเป็นร้อยละ 78 ซึ่งสูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ 3) นักเรียนที่ได้คะแนนพัฒนาการสัมพันธ์ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์อยู่ในระดับพัฒนาการสูงถึงสูงมาก มีจำนวน 49 คน คิดเป็นร้อยละ 90.7 ของจำนวนนักเรียนทั้งหมด ซึ่งเป็นไปตามเกณฑ์ที่กำหนด

คำสำคัญ: การจัดการเรียนรู้แบบ CPA, เทคนิคเพื่อนคู่คิด, การเรียงสับเปลี่ยน, การให้เหตุผลทางคณิตศาสตร์, ผลสัมฤทธิ์ทางการเรียน

2020 MSC: 97C70, 97D40

*ผู้นำเสนอและผู้ประพันธ์บรรณกิจ (anuchasaowakon@gmail.com)

1 บทนำ

การให้เหตุผลทางคณิตศาสตร์เป็นสิ่งที่ส่งเสริมให้นักเรียนได้ใช้การเรียบเรียงข้อมูลต่าง ๆ ที่เกี่ยวข้องกับวิชาคณิตศาสตร์ที่มีอยู่ก่อนแล้ว ให้เกิดความสัมพันธ์กันเพื่อส่งผลให้ข้อความที่อธิบายเหล่านั้นมีความต่อเนื่องกันจนทำให้เกิดข้อสรุปทางคณิตศาสตร์ขึ้นมาอย่างสมเหตุสมผลและเหมาะสมกับบริบทนั้น ๆ [1] การส่งเสริมให้นักเรียนได้ฝึกฝนการให้เหตุผลทางคณิตศาสตร์อย่างสม่ำเสมอมีความสำคัญอย่างยิ่งต่อการพัฒนาทักษะและกระบวนการทางคณิตศาสตร์อื่น ๆ ให้มีประสิทธิภาพมากยิ่งขึ้น ซึ่งจะนำไปสู่ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ที่สูงขึ้น ครูคณิตศาสตร์จึงควรตระหนักถึงความสำคัญของการให้เหตุผลทางคณิตศาสตร์ซึ่งมีลักษณะเป็นนามธรรมและมีความซับซ้อนในการทำความเข้าใจ ทั้งนี้หากเกิดกระบวนการให้เหตุผลทางคณิตศาสตร์ในลักษณะที่ไม่ถูกต้องอาจส่งผลเชิงลบต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนได้เช่นกัน [2] นอกจากนี้ การให้เหตุผลทางคณิตศาสตร์ยังเป็นทักษะที่มีความสำคัญในการศึกษาหน่วยการเรียนรู้เรื่อง การเรียงสับเปลี่ยน ซึ่งจำเป็นต้องอาศัยการให้เหตุผลทางคณิตศาสตร์ประกอบกับหลักการทางคณิตศาสตร์ เพื่อวิเคราะห์และอธิบายถึงความสมเหตุสมผลในสถานการณ์ต่าง ๆ

จากการศึกษาสภาพปัญหาที่เกิดขึ้นในการจัดการเรียนการสอนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนโรงเรียนนาหลวง ในภาคการศึกษาที่ 2 ปีการศึกษา 2566 พบประเด็นจากการวิเคราะห์การทำแบบทดสอบของนักเรียน คือนักเรียนส่วนใหญ่ไม่สามารถให้เหตุผลทางคณิตศาสตร์ได้อย่างถูกต้อง อีกทั้งยังขาดการแสดงภาพแนวคิดประกอบการหาคำตอบที่ส่งผลต่อความถูกต้องของผลลัพธ์ที่ได้ ทั้งนี้ มีนักเรียนเพียงจำนวน 5 คน คิดเป็นร้อยละ 9.25 เท่านั้น ที่สามารถให้เหตุผลทางคณิตศาสตร์ได้อย่างถูกต้องพร้อมทั้งแสดงภาพแนวคิดในการหาคำตอบจนได้ผลลัพธ์ที่ถูกต้อง จากผลการวิจัยของ Stylianou และ Sliver [3] ซึ่งได้ทำการศึกษาเปรียบเทียบการใช้การแสดงภาพในการแก้ปัญหาคณิตศาสตร์ระดับสูงระหว่างนักคณิตศาสตร์ กับนักเรียนหรือนักศึกษา พบว่านักคณิตศาสตร์ใช้การแสดงภาพทางคณิตศาสตร์เป็นเครื่องมือสำคัญในการคิดและวางแผนในกระบวนการแก้ปัญหา ส่วนนักเรียนหรือนักศึกษานั้นไม่ได้ใช้การแสดงภาพในการแก้ปัญหาเนื่องจากข้อมูลมีความซับซ้อนเป็นอย่างมาก ดังนั้น การรวบรวมข้อมูลและแสดงภาพแนวคิดควบคู่กับการให้เหตุผลทางคณิตศาสตร์ที่สอดคล้องกันจึงมีความสำคัญ และส่งผลโดยตรงต่อความถูกต้องของคำตอบ ในการพัฒนาการเรียนการสอนจึงควรมุ่งเน้นที่การเชื่อมโยงระหว่างการให้เหตุผลทางคณิตศาสตร์และการแสดงภาพแนวคิดร่วมกับการเรียนรู้ เรื่อง การเรียงสับเปลี่ยน

การจัดการเรียนรู้แบบ Concrete Pictorial Abstract (CPA) เป็นกระบวนการสอนที่ช่วยให้นักเรียนเข้าใจความคิดรวบยอดหรือแนวคิดทางคณิตศาสตร์อย่างค่อยเป็นค่อยไป โดยเริ่มจากสิ่งที่เป็นรูปธรรมที่สามารถจับต้องได้และนำสู่สิ่งที่เป็นนามธรรมมากขึ้น ประกอบด้วยขั้นที่ 1 ขั้นรูปธรรม (Concrete) ครูจะใช้วัสดุหรือสิ่งของจริงที่นักเรียนสามารถสัมผัสและมองเห็นได้ ขั้นที่ 2 ขั้นภาพประกอบ (Pictorial) ครูจะใช้ภาพประกอบ หรือสัญลักษณ์ต่าง ๆ เพื่อเชื่อมโยงความเข้าใจจากของจริงไปสู่สิ่งที่เป็นนามธรรมมากขึ้น และขั้นที่ 3 ขั้นนามธรรม (Abstract) ครูจะนำเสนอแนวคิดในรูปแบบของสัญลักษณ์ สมการ หรือนามธรรมอื่น ๆ [4] การจัดการเรียนรู้แบบ CPA จึงเป็นกระบวนการเรียนรู้ที่เหมาะอย่างยิ่งสำหรับเรื่อง การเรียงสับเปลี่ยน เพราะการเรียนเรื่องนี้มีความซับซ้อนเกี่ยวกับการนับและลำดับของวัตถุ การเริ่มต้นด้วยวัตถุที่จับต้องได้จะช่วยให้นักเรียนเห็นภาพจริงของการจัดลำดับ ช่วยให้นักเรียน

เข้าใจเกี่ยวกับการสลับตำแหน่งของวัตถุทำให้เกิดการจัดเรียงที่แตกต่างกัน ขั้น Pictorial จะเป็นการฝึกฝนให้นักเรียนใช้ภาพแสดงแนวคิดเพื่อช่วยสร้างความเข้าใจ ภาพเหล่านี้ช่วยให้นักเรียนเห็นรูปแบบของการจัดเรียงที่เกิดขึ้นโดยไม่ต้องใช้วัตถุจริงและเชื่อมโยงไปสู่การใช้สูตรการเรียงสับเปลี่ยนได้อย่างมีเหตุผล

จากการสังเกตพฤติกรรมการเรียนรู้ของนักเรียนโรงเรียนนาหลวงพบว่านักเรียนจะมีความกระตือรือร้นให้ความร่วมมือและมีส่วนร่วมทุกครั้งที่มีการใช้สื่อการสอนหรือมีการลงมือปฏิบัติเกี่ยวกับการเรียนรู้ต่าง ๆ ไปพร้อม ๆ กับครู โดยเฉพาะในการทำกิจกรรมต่าง ๆ ที่มีการจับคู่หรือทำงานเป็นกลุ่ม ทำให้ทุกคนได้มีปฏิสัมพันธ์กับเพื่อน ได้ฝึกการทำงานเป็นทีม และการรับฟังความคิดเห็นของผู้อื่น เทคนิคเพื่อนคู่คิด เป็นเทคนิคหนึ่งที่ส่งเสริมให้นักเรียนได้มีส่วนร่วมในการเรียนรู้อย่างแท้จริง ทำให้นักเรียนมีปฏิสัมพันธ์ซึ่งกันและกัน เกิดความมั่นใจและกล้าที่จะแสดงออก [5] โดยผ่านกระบวนการคิด การแลกเปลี่ยนความคิดเห็นกับเพื่อน และการนำเสนอความคิดหน้าชั้นเรียน โดยเทคนิคเพื่อนคู่คิดมีขั้นตอน คือ ขั้นที่ 1 คิด (Think) ครูจะตั้งคำถามหรือประเด็นที่น่าสนใจให้นักเรียนคิดอย่างลึกซึ้ง นักเรียนแต่ละคนจะได้รับโอกาสในการคิดวิเคราะห์และอธิบายความคิดของตนเองลงในกระดาษหรือสมุด ขั้นที่ 2 คู่ (Pair) ครูจะแบ่งนักเรียนเป็นคู่ ในขั้นตอนนี้ นักเรียนจะได้แลกเปลี่ยนความคิดเห็นและอภิปรายร่วมกับเพื่อนในกลุ่มของตน เพื่อหาข้อสรุปร่วมกัน โดยแต่ละคนจะได้นำเสนอความคิดที่ตนได้คิดไว้ในขั้นแรก ขั้นที่ 3 นำเสนอ (Share) เมื่อนักเรียนได้อภิปรายและหาข้อสรุปร่วมกันในคู่ของตนเองแล้ว ผู้แทนกลุ่มจะออกมานำเสนอความคิดของกลุ่มหน้าชั้นเรียน ในขณะที่กลุ่มอื่น ๆ จะได้ฟังและแสดงความคิดเห็นเพิ่มเติม เกิดการแลกเปลี่ยนความคิดภายในชั้นเรียน ซึ่งสอดคล้องกับพฤติกรรมการเรียนรู้ของนักเรียนโรงเรียนนาหลวง

จากที่กล่าวมาข้างต้น ทำให้ผู้วิจัยสนใจที่จะศึกษาความสามารถในการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 โดยใช้วิธีการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด ซึ่งเป็นการบูรณาการกระบวนการเรียนรู้อย่างเป็นลำดับขั้นจากรูปธรรมสู่นามธรรมควบคู่กับการแลกเปลี่ยนความคิดเห็นระหว่างนักเรียน อันเป็นแนวทางการพัฒนากระบวนการคิดวิเคราะห์ การให้เหตุผล และเสริมสร้างการเรียนรู้แบบมีส่วนร่วมให้แก่นักเรียน

วัตถุประสงค์ของการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยมีวัตถุประสงค์ของการวิจัย ดังนี้

1) เพื่อศึกษาความสามารถในการให้เหตุผลทางคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด

2) เพื่อศึกษาผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด

3) เพื่อศึกษาคะแนนพัฒนาการสัมพัทธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 เรื่อง การเรียงสับเปลี่ยน หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด

สมมติฐานการวิจัย

1) ความสามารถในการให้เหตุผลทางคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด สูงกว่าเกณฑ์ร้อยละ 70

2) ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด สูงกว่าเกณฑ์ร้อยละ 70

3) นักเรียนชั้นมัธยมศึกษาปีที่ 5 ที่ได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิดมีคะแนนพัฒนาการสัมพัทธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน อยู่ในระดับพัฒนาการระดับสูงขึ้นไปคิดเป็นจำนวนไม่น้อยกว่าร้อยละ 70 ของจำนวนนักเรียนทั้งหมด

ขอบเขตของการวิจัย

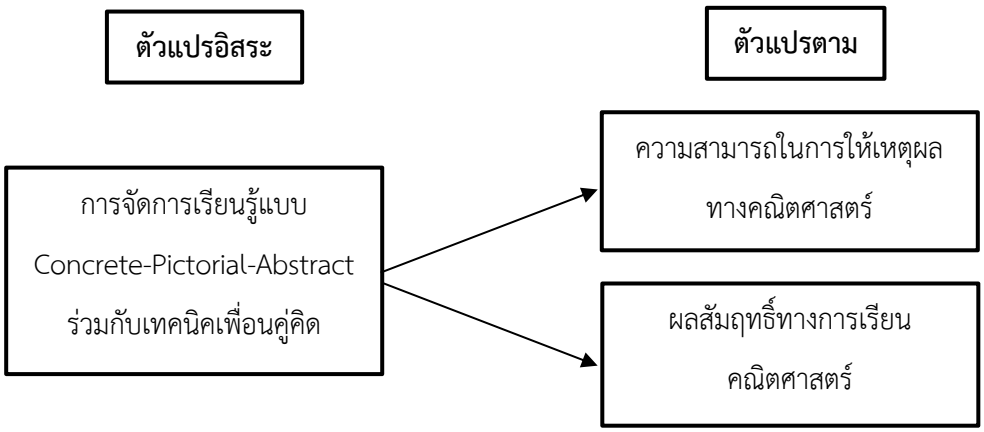
- ประชากรที่ใช้ในการวิจัย

ประชากรในการวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนนาหลวง สำนักงานเขตทุ่งครุ จังหวัดกรุงเทพมหานคร ภาคเรียนที่ 2 ปีการศึกษา 2567 ซึ่งจัดห้องเรียนตามความถนัดของนักเรียน จำนวน 6 ห้องเรียน รวมนักเรียนทั้งหมด 286 คน

- กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างในการวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนนาหลวง สำนักงานเขตทุ่งครุ จังหวัดกรุงเทพมหานคร ภาคเรียนที่ 2 ปีการศึกษา 2567 จำนวน 1 ห้องเรียน ทั้งหมด 54 คน ซึ่งได้มาจากการสุ่มตัวอย่างแบบกลุ่ม (Cluster random sampling)

กรอบแนวคิดการวิจัย



เครื่องมือสำหรับการปฏิบัติการวิจัย

1. แผนการจัดการเรียนรู้แบบ Concrete – Pictorial – Abstract (CPA) ร่วมกับเทคนิคเพื่อนคู่คิด เรื่อง การเรียงสับเปลี่ยน สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 ประกอบด้วยแผนการจัดการเรียนรู้ จำนวน 6 แผน ที่ผ่านการตรวจคุณภาพจากผู้เชี่ยวชาญจำนวน 5 ท่าน เพื่อประเมินและตรวจสอบความเหมาะสมของแผนการจัดการเรียนรู้ ผลการประเมินปรากฏว่ามีค่าดัชนีความสอดคล้อง (IOC) เท่ากับ 1.00

ในแต่ละแผนของการจัดการเรียนรู้แบบ Concrete – Pictorial – Abstract ร่วมกับเทคนิคเพื่อนคู่คิด เรื่อง การเรียงสับเปลี่ยน ประกอบด้วย 3 ขั้นตอนหลัก ดังนี้

1) ขั้นการสอนเชิงรูปธรรม **(Concrete)** เป็นขั้นที่ให้นักเรียนลงมือปฏิบัติด้วยตนเอง โดยใช้สื่อการสอนและอุปกรณ์เป็นเครื่องมือในการสร้างความเข้าใจพื้นฐาน ในขั้นนี้เริ่มจากการกระตุ้นให้นักเรียนเกิดกระบวน **การคิด (Think)** ผ่านสถานการณ์ปัญหาที่กำหนด ทำให้นักเรียนได้สำรวจ สังเกต และค้นหาแนวทางการหาคำตอบด้วยตนเอง
2) ขั้นการสอนเชิงรูปภาพ **(Pictorial)** เป็นขั้นที่ให้นักเรียนเขียนรูปภาพหรือแผนภาพแทนความคิดในการหาคำตอบของปัญหาในแต่ละข้อด้วยตนเอง จากนั้นให้นักเรียน **จับคู่ (Pair)** เพื่อแลกเปลี่ยนและอภิปรายแนวคิดร่วมกันเพื่อหาข้อสรุปของคู่ตนเอง ในตอนท้ายของขั้นนี้ นักเรียนจำนวน 1 คู่ ที่ได้จากการสุ่มจะมา **นำเสนอ (Share)** แนวคิดและเหตุผลของคู่ตนเองต่อเพื่อนคู่อื่น ๆ ในชั้นเรียน เพื่อแลกเปลี่ยนเรียนรู้พร้อมกับเชื่อมโยงไปสู่การสังเกตผลลัพธ์และประเด็นสำคัญ
3) ขั้นการสอนเชิงนามธรรม **(Abstract)** เป็นขั้นที่ให้นักเรียนสร้างองค์ความรู้และถอดความคิดจากการดำเนินการขั้นก่อนหน้า นำมาแสดงโดยการเขียนสัญลักษณ์ทางคณิตศาสตร์ และประยุกต์สำหรับการหาคำตอบของปัญหาในสถานการณ์ต่าง ๆ ในแต่ละข้อด้วยตนเอง

เครื่องมือสำหรับใช้วัดและประเมินผล

1) แบบทดสอบวัดความสามารถในการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์หลังเรียน เรื่อง การเรียงสับเปลี่ยน

แบบทดสอบวัดความสามารถในการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์หลังเรียน เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 เป็นแบบทดสอบอัตนัยจำนวน 15 ข้อ ที่ใช้ทดสอบหลังการจัดการเรียนรู้ ที่มีลักษณะคู่ขนานกับแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ก่อนเรียนและมีข้อคำถามที่ใช้วัดความสามารถในการให้เหตุผลทางคณิตศาสตร์ โดยส่งให้ผู้เชี่ยวชาญจำนวน 5 ท่าน เพื่อตรวจสอบความสอดคล้องระหว่างข้อสอบกับตัวชี้วัด หลังจากการปรับปรุงแก้ไขตามคำแนะนำของผู้เชี่ยวชาญแล้วจึงนำไปทดลองใช้กับนักเรียนชั้นมัธยมศึกษาปีที่ 6 จำนวน 54 คน ที่ไม่ใช่กลุ่มตัวอย่างและได้เคยเรียนรู้ เรื่อง การเรียงสับเปลี่ยนมาแล้ว จากนั้นนำคะแนนที่ได้มาวิเคราะห์ข้อมูลและคัดเลือกข้อสอบที่ผ่านเกณฑ์ตามกำหนด จำนวน 7 ข้อ (เป็น

แบบทดสอบที่มีการวัดผลทั้งการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนในข้อคำถามเดียวกันทั้ง 7 ข้อ) โดยแบบทดสอบวัดความสามารถในการให้เหตุผลทางคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยนซึ่งได้ค่าดัชนีความสอดคล้อง (IOC) เท่ากับ 1 มีค่าความยากง่าย (p) ตั้งแต่ 0.50 ถึง 0.62 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.43 ถึง 0.67 และค่าความเชื่อมั่นเท่ากับ 0.96 และแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยนได้ค่าดัชนีความสอดคล้อง (IOC) เท่ากับ 0.6 มีค่าความยากง่าย (p) ตั้งแต่ 0.52 ถึง 0.73 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.70 ถึง 0.83 และค่าความเชื่อมั่นเท่ากับ 0.95 ทั้งนี้ ได้กำหนดเกณฑ์การประเมินการให้เหตุผลทางคณิตศาสตร์ของแบบทดสอบ ไว้ดังนี้

ระดับ	คำอธิบาย
3	นักเรียนแสดงแนวคิดหรือเหตุผลในการสร้างข้อสรุปทางคณิตศาสตร์ได้อย่างสมเหตุสมผลและถูกต้องครบทุกประเด็น
2	นักเรียนแสดงแนวคิดหรือเหตุผลในการสร้างข้อสรุปทางคณิตศาสตร์ได้ถูกต้องบางประเด็น ซึ่งยังไม่เพียงพอต่อการสร้างข้อสรุปดังกล่าว
1	นักเรียนแสดงร่องรอยแนวคิดหรือให้เหตุผลในการสร้างข้อสรุปทางคณิตศาสตร์ แต่ยังไม่ถูกต้อง และไม่สมเหตุสมผล
0	นักเรียนไม่แสดงแนวคิดหรือเหตุผลในการสร้างข้อสรุปทางคณิตศาสตร์

2) แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ก่อนเรียน เรื่อง การเรียงสับเปลี่ยน

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ก่อนเรียน เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 เป็นแบบทดสอบอัตนัย จำนวน 15 ข้อที่ใช้ทดสอบก่อนเรียน โดยส่งให้ผู้เชี่ยวชาญทั้ง 5 ท่าน พิจารณาเพื่อตรวจสอบความสอดคล้องระหว่างข้อสอบกับตัวชี้วัด ซึ่งได้ค่าดัชนีความสอดคล้อง (IOC) เท่ากับ 0.6 หลังจากการปรับปรุงแก้ไขตามคำแนะนำของผู้เชี่ยวชาญแล้วจึงนำไปทดลองใช้กับนักเรียนชั้นมัธยมศึกษาปีที่ 6 จำนวน 54 คน ที่ไม่ใช่กลุ่มตัวอย่างและได้เรียนรู้เรื่อง การเรียงสับเปลี่ยน มาแล้ว จากนั้นได้วิเคราะห์หาคุณภาพและและคัดเลือกข้อสอบอยู่ในเกณฑ์ตามกำหนด จำนวน 7 ข้อ โดยมีค่าความยากง่าย (p) ตั้งแต่ 0.46 ถึง 0.72 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.70 ถึง 0.76 และค่าความเชื่อมั่นเท่ากับ 0.92

วิธีการดำเนินงานวิจัยและการเก็บรวบรวมข้อมูล

หลังจากสร้างเครื่องมือวิจัยเสร็จสมบูรณ์และผ่านการประเมินจริยธรรมในมนุษย์ ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลจากนักเรียนกลุ่มตัวอย่าง ดังนี้

1) ผู้วิจัยให้นักเรียนทำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ก่อนเรียน เรื่อง การเรียงสับเปลี่ยน จำนวน 7 ข้อ ใช้เวลา 50 นาที

2) ผู้วิจัยดำเนินการจัดการเรียนรู้ตามแผนที่ได้กำหนดไว้ จำนวน 6 แผน แผนละ 100 นาที รวม 600 นาที มีรายละเอียดดังนี้

การเรียงสับเปลี่ยนเชิงเส้นของสิ่งของที่แตกต่างกัน	จำนวน	2	แผน
การเรียงสับเปลี่ยนเชิงเส้นของสิ่งของที่ไม่แตกต่างกัน	จำนวน	2	แผน
การเรียงสับเปลี่ยนเชิงวงกลม	จำนวน	2	แผน

3) ผู้วิจัยให้นักเรียนทำแบบทดสอบวัดความสามารถในการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์หลังเรียน เรื่อง การเรียงสับเปลี่ยน จำนวน 7 ข้อ ใช้เวลา 50 นาที

การวิเคราะห์ข้อมูล

1) สำหรับการศึกษาความสามารถในการให้เหตุผลทางคณิตศาสตร์ และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนหลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด ผู้วิจัยได้นำผลคะแนนจากแบบทดสอบหลังเรียนมาวิเคราะห์ข้อมูล หาค่าเฉลี่ยเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน แล้วนำมาเปรียบเทียบกับเกณฑ์ร้อยละ 70 เพื่อทำการทดสอบสมมติฐานที่ตั้งไว้ว่า ความสามารถในการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด คิดเป็นร้อยละ 76.09 และคิดเป็นร้อยละ 78 ตามลำดับ ซึ่งสูงกว่าเกณฑ์ร้อยละ 70 โดยใช้การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างเดียว (t-test for one sample) ที่ระดับนัยสำคัญ .05

2) สำหรับการศึกษาพัฒนาการสัมพันธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนหลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด ผู้วิจัยได้นำผลคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนก่อนและหลังเรียนมาวิเคราะห์ข้อมูล หาคะแนนพัฒนาการสัมพันธ์ตามหลักการคำนวณของศิริชัย กาญจนวาสี [6] และแปลความหมายจากคะแนนที่เกิดขึ้นตามเกณฑ์การประเมิน เพื่อทำการทดสอบสมมติฐานที่ตั้งไว้ว่า คะแนนพัฒนาการสัมพันธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด อยู่ในระดับพัฒนาการระดับสูงขึ้นไป คิดเป็นจำนวนไม่น้อยกว่าร้อยละ 70 ของนักเรียนทั้งหมด

ผลการวิจัย

1. ความสามารถในการให้เหตุผลทางคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน

ผลการวิเคราะห์ความสามารถในการให้เหตุผลทางคณิตศาสตร์ คะแนนที่ได้จากการทำแบบทดสอบวัดความสามารถในการให้เหตุผลทางคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด มีค่าเฉลี่ยเลยคณิตเท่ากับ 15.98 คิดเป็นร้อยละ 76.09 ของคะแนนเต็ม และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 3.858 ซึ่งสูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผลปรากฏดังตารางที่ 3

ตารางที่ 3 แสดงผลการเปรียบเทียบความสามารถในการให้เหตุผลทางคณิตศาสตร์ของนักเรียนหลังได้รับการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด เรื่อง การเรียงสับเปลี่ยน กับเกณฑ์ร้อยละ 70

การทดสอบ	จำนวน (คน)	คะแนนเต็ม	ค่าเฉลี่ยเลขคณิต	ร้อยละ	ส่วนเบี่ยงเบนมาตรฐาน	ค่า t-test	df	Sig.
หลังการเรียนรู้	54	21	15.98	76.09	3.858	1.869*	53	0.0335

*มีนัยสำคัญทางสถิติที่ระดับ .05

2. ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน

ผลการวิเคราะห์ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ คะแนนที่ได้จากการทำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด มีค่าเฉลี่ยเลยคณิตเท่ากับ 5.46 คิดเป็นร้อยละ 78 ของคะแนนเต็ม และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.193 ซึ่งสูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผลปรากฏดังตารางที่ 4

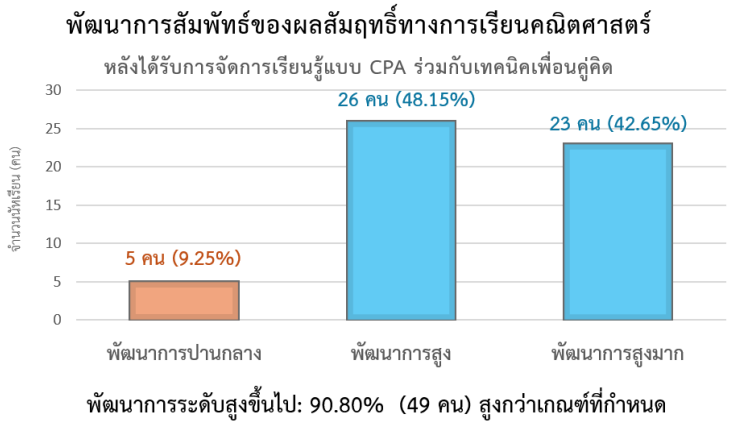
ตารางที่ 4 แสดงผลการเปรียบเทียบความสามารถในการให้เหตุผลทางคณิตศาสตร์ของนักเรียนหลังได้รับการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด เรื่อง การเรียงสับเปลี่ยน กับเกณฑ์ร้อยละ 70

การทดสอบ	จำนวน (คน)	คะแนนเต็ม	ค่าเฉลี่ยเลขคณิต	ร้อยละ	ส่วนเบี่ยงเบนมาตรฐาน	ค่า t-test	df	Sig.
หลังการเรียนรู้	54	7	5.46	78	1.193	2.852*	53	0.003

*มีนัยสำคัญทางสถิติที่ระดับ .05

3. คะแนนพัฒนาการสัมพัทธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน

ผลการวิเคราะห์พัฒนาสัมพัทธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ นักเรียนชั้นมัธยมศึกษาปีที่ 5 ที่ได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิดมีคะแนนพัฒนาการสัมพัทธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน เมื่อเทียบกับเกณฑ์ตามแนวคิดของ ศิริชัย กาญจนวาสี อยู่ในระดับพัฒนาการระดับสูงขึ้นไปคิดเป็นจำนวนไม่น้อยกว่าร้อยละ 70 ของนักเรียนทั้งหมด รายละเอียดแสดงดังรูปที่ 1



รูปที่ 1 : กราฟแสดงระดับพัฒนาการสัมพันธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์

จากรูปที่ 1 พบว่า ระดับพัฒนาการสัมพันธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด สามารถจำแนกได้ดังนี้ ระดับพัฒนาการปานกลาง จำนวน 5 คน คิดเป็นร้อยละ 9.25 ของจำนวนนักเรียนทั้งหมด ระดับพัฒนาการสูงจำนวน 26 คน คิดเป็นร้อยละ 48.15 ของจำนวนนักเรียนทั้งหมด และระดับพัฒนาการสูงมาก จำนวน 23 คน คิดเป็นร้อยละ 42.65 ของจำนวนนักเรียนทั้งหมด โดยพบว่านักเรียนที่มีระดับพัฒนาการตั้งแต่ระดับสูงขึ้นไป มีจำนวนรวมทั้งสิ้น 49 คน คิดเป็นร้อยละ 90.80 ของจำนวนนักเรียนทั้งหมด แสดงว่า การจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด ส่งผลให้นักเรียนมีพัฒนาการสัมพันธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์อยู่ในระดับสูงขึ้นไป มีจำนวนไม่น้อยกว่าร้อยละ 70 ของจำนวนนักเรียนทั้งหมด

สรุปผลและอภิปรายผล

ผลการจัดการเรียนรู้แบบ Concrete – Pictorial – Abstract ร่วมกับเทคนิคเพื่อนคู่คิด เรื่อง การเรียงสับเปลี่ยน ที่ส่งผลต่อความสามารถในการให้เหตุผลทางคณิตศาสตร์และผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 สามารถสรุปผลและอภิปรายผล ดังนี้

1. สรุปผล

1) ความสามารถในการให้เหตุผลทางคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด สูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2) ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด สูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3) นักเรียนชั้นมัธยมศึกษาปีที่ 5 ที่ได้รับการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิดมีคะแนนพัฒนาการสัมพัทธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน อยู่ในระดับพัฒนาการระดับสูงขึ้นไปคิดเป็นจำนวนร้อยละ 90.80 ของจำนวนนักเรียนทั้งหมด

2 อภิปรายผล

1) ความสามารถในการให้เหตุผลทางคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA (Concrete-Pictorial-Abstract) ร่วมกับเทคนิคเพื่อนคู่คิด (Think-Pair-Share) สูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งสอดคล้องกับสมมติฐานที่วางไว้ เนื่องจากการจัดการเรียนรู้แบบ Concrete – Pictorial – Abstract ร่วมกับเทคนิคเพื่อนคู่คิด ช่วยให้นักเรียนได้พัฒนาความเข้าใจคณิตศาสตร์อย่างเป็นลำดับขั้น ผ่านการทำกิจกรรมที่เริ่มต้นจากการเรียนรู้จากสิ่งที่เป็นรูปธรรม ที่ช่วยให้นักเรียนได้ประมวลผลความคิดและสร้างความเข้าใจพื้นฐานจากสิ่งที่จับต้องได้ด้วยตนเอง นำไปสู่การเชื่อมโยงประสบการณ์เชิงรูปธรรมกับการใช้ภาพ เพื่อสร้างตัวแทนสำหรับการสื่อสารแนวคิดและเหตุผลทางคณิตศาสตร์ ในขั้นนี้ เทคนิคเพื่อนคู่คิดมีบทบาทสำคัญในการส่งเสริมการพัฒนาความสามารถในการให้เหตุผล โดยเฉพาะในขั้นตอนการแลกเปลี่ยนความคิดที่นักเรียนได้มีโอกาสอธิบายเหตุผล อภิปราย และตรวจสอบความถูกต้องของการให้เหตุผลร่วมกับเพื่อน กระบวนการนี้ช่วยให้นักเรียนได้เห็นมุมมองที่หลากหลาย ได้ตั้งคำถาม และได้อธิบายแนวคิดของตนเอง ทำให้เกิดการพัฒนาไปสู่การสร้างแนวคิดเชิงนามธรรมที่นักเรียนสามารถใช้สัญลักษณ์และการอธิบายเชิงคณิตศาสตร์ในการให้เหตุผลได้อย่างเป็นระบบ นอกจากนี้ จากการตรวจแบบทดสอบหลังเรียน พบว่านักเรียนมักมีการวาดภาพหรือแผนภาพต้นไม้ประกอบการหาคำตอบและการให้เหตุผล เพื่อพิจารณาความสอดคล้องและความสมเหตุสมผลของการแก้ปัญหา ตรงตามที่ นลิตา สอนวารี และคณะ [7] กล่าวว่า การจัดกิจกรรมการเรียนรู้ด้วยเทคนิค CPA ทำให้นักเรียนส่วนใหญ่มีความเข้าใจมโนมติทางคณิตศาสตร์ที่ถูกต้องสมบูรณ์ขึ้น เพราะนักเรียนได้ปฏิบัติกิจกรรมจากสิ่งที่เป็นรูปธรรม จึงทำให้เข้าใจเนื้อหาได้ถูกต้องไม่มีความคลาดเคลื่อน อีกทั้งนักเรียนยังสามารถเขียนสรุป อธิบายหลักการด้วยตนเองได้ ทำให้เกิดแรงจูงใจที่ดีในกิจกรรมที่จะเรียนรู้และนักเรียนสามารถเชื่อมโยงเข้าสู่สิ่งที่เป็นนามธรรมได้ รวมถึง อังคณา ทิพยรุ่งโรจน์ [8] กล่าวว่า การจัดการเรียนรู้ตามแนวคิด CPA ซึ่งให้นักเรียนได้เรียนรู้โดยเริ่มจากสิ่งที่เป็นรูปธรรมจับต้องได้ จากนั้นเชื่อมโยงสิ่งที่จับต้องได้ให้อยู่ในรูปแบบของภาพ และเชื่อมโยงให้อยู่ในรูปสัญลักษณ์ทางคณิตศาสตร์ในที่สุด ทำให้นักเรียนเข้าใจคณิตศาสตร์จากสิ่งที่เป็นรูปธรรมไปยังสิ่งที่เป็นนามธรรม ส่งผลให้นักเรียนเกิดการเรียนรู้อย่างมีความหมาย และสอดคล้องกับ ธัญญา แนวดง และคณะ [9] ได้ศึกษาการจัดการเรียนรู้ด้วยเทคนิคเพื่อนคู่คิด ที่มีต่อความสามารถในการให้เหตุผลทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 พบว่าความสามารถในการให้เหตุผลทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่ได้รับการจัดการเรียนรู้ด้วยเทคนิคเพื่อนคู่คิด สูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2) ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง การเรียงสับเปลี่ยน ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 หลังได้รับการจัดการเรียนรู้แบบ CPA (Concrete-Pictorial-Abstract) ร่วมกับเทคนิคเพื่อนคู่คิด (Think-Pair-Share) มีคะแนนเฉลี่ยคิดเป็นร้อยละ 78 ของคะแนนเต็ม สูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 นอกจากนี้

นักเรียนร้อยละ 90.80 ของจำนวนนักเรียนทั้งหมด มีคะแนนพัฒนาการสัมพัทธ์ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์อยู่ในระดับพัฒนาการสูงถึงสูงมาก ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ผลการวิจัยดังกล่าวสามารถอธิบายได้ว่า การจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด ช่วยลดความซับซ้อนของเนื้อหาส่งผลให้นักเรียนเข้าใจเนื้อหา เรื่องการเรียงสับเปลี่ยนได้ดียิ่งขึ้น โดยกระบวนการเรียนรู้ที่เป็นขั้นตอน เริ่มจากขั้น Concrete ที่นักเรียนได้สัมผัสและเรียนรู้ผ่านวัตถุที่จับต้องได้ มีการใช้อุปกรณ์หรือสัญลักษณ์ในการจัดเรียงรูปแบบต่าง ๆ ทำให้นักเรียนเข้าใจแนวคิดผ่านประสบการณ์ตรง ต่อมาในขั้น Pictorial นักเรียนได้เรียนรู้ผ่านรูปภาพที่เป็นสื่อในการเชื่อมโยงจากรูปธรรมสู่นามธรรม และในขั้นสุดท้าย Abstract นักเรียนสามารถเข้าใจสัญลักษณ์และสูตรทางคณิตศาสตร์เกี่ยวกับการเรียงสับเปลี่ยนได้อย่างมีความหมาย ในขณะเดียวกันการใช้เทคนิคเพื่อนคู่คิด ช่วยส่งเสริมการเรียนรู้แบบมีส่วนร่วมและการสื่อสารทางคณิตศาสตร์ โดยในขั้น Think นักเรียนได้คิดวิเคราะห์โจทย์ปัญหาการเรียงสับเปลี่ยนด้วยตนเอง ซึ่งช่วยพัฒนาทักษะการคิดอย่างอิสระ ส่วนขั้น Pair นักเรียนได้แลกเปลี่ยนแนวคิดและวิธีการแก้ปัญหากับเพื่อน ตรวจสอบความเข้าใจซึ่งกันและกัน และในขั้น Share นักเรียนได้นำเสนอและสรุปความรู้ร่วมกันในชั้นเรียน กระบวนการนี้ช่วยสร้างบรรยากาศการเรียนรู้ที่เอื้อต่อการแลกเปลี่ยนความคิดเห็น ลดความวิตกกังวลในการเรียนวิชาคณิตศาสตร์ซึ่งเป็นอุปสรรคสำคัญต่อการเรียนรู้ของนักเรียนหลายคน อีกทั้งยังส่งเสริมการสื่อสารทางคณิตศาสตร์ผ่านการอธิบายแนวคิด ทำให้ได้รับข้อมูลย้อนกลับจากเพื่อนและครู จึงสามารถปรับปรุงและพัฒนาความเข้าใจของตนเองได้อย่างต่อเนื่อง ส่งผลให้นักเรียนมีผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ที่สูงขึ้น สอดคล้องกับงานวิจัยของ ชาญชัย ชนไพโรจน์ และคณะ[10] ได้ศึกษาการพัฒนาบทเรียน เรื่อง ปริซึมและทรงกระบอก โดยใช้กระบวนการจัดการเรียนรู้ CPA เพื่อเพิ่มผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 2 พบว่า ผลสัมฤทธิ์ทางการเรียน เรื่อง ปริซึมและทรงกระบอกของนักเรียนชั้นมัธยมศึกษาปีที่ 2 หลังเรียนโดยใช้กระบวนการการจัดการเรียนรู้ CPA สูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 รวมถึง รัตนาภรณ์ ไทยพู และคณะ [11] ได้ศึกษาการจัดการเรียนรู้ตามแนวคิด CPA เพื่อพัฒนาผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ สำหรับนักเรียนมัธยมศึกษาปีที่ 2 พบว่า ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ได้รับการจัดการเรียนรู้ตามแนวคิด CPA สูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และสอดคล้องกับ อัจจิมา บุญพัตร์ [12] ได้ศึกษาผลการจัดการเรียนรู้โดยใช้เทคนิคเพื่อนคู่คิดที่มีต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 พบว่านักเรียนชั้นมัธยมศึกษาปีที่ 1 มีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์หลังได้รับการจัดการเรียนรู้โดยใช้เทคนิคเพื่อนคู่คิดสูงกว่าเกณฑ์ร้อยละ 70 อย่างมีนัยสำคัญทางสถิติที่ .05

ข้อเสนอแนะ

ข้อเสนอแนะเพื่อการนำผลการวิจัยไปใช้

1) การเตรียมสื่อและอุปกรณ์การเรียนรู้ถือเป็นปัจจัยสำคัญในการจัดการเรียนรู้แบบ CPA ผู้สอนควรเลือกและออกแบบสื่อที่สามารถช่วยให้นักเรียนเข้าใจแนวคิดทางคณิตศาสตร์ได้อย่างเป็นรูปธรรมโดยคำนึงถึงความสอดคล้องกับขั้นตอนการเรียนรู้ รวมทั้งคำนึงถึงความเหมาะสมกับผู้เรียน

2) ครูควรชี้แจงรายละเอียดของรูปแบบการเรียนรู้ อธิบายขั้นตอนและบทบาทของนักเรียนอย่างชัดเจน สร้างความเข้าใจถึงเป้าหมายของการเรียนรู้ และกระตุ้นให้นักเรียนเกิดความกระตือรือร้นในการเรียนรู้ ตลอดจนให้การสนับสนุนและส่งเสริมกำลังใจอย่างต่อเนื่อง

3) ครูควรออกแบบห้องเรียนให้เป็นพื้นที่ปลอดภัย เอื้อต่อการแลกเปลี่ยนเรียนรู้ ส่งเสริมให้นักเรียนกล้าคิด กล้าแสดงออก และมีส่วนร่วมในกิจกรรมการเรียนรู้ โดยทำหน้าที่เป็นผู้อำนวยความสะดวกและสร้างบรรยากาศที่เป็นมิตร สนับสนุนการเรียนรู้ร่วมกันของนักเรียน

ข้อเสนอแนะสำหรับการวิจัยในครั้งต่อไป

1) จากผลการวิจัยพบว่าผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์มีระดับพัฒนาการสูงถึงสูงมาก จึงควรพิจารณานำรูปแบบการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิดไปประยุกต์ใช้กับเนื้อหาคณิตศาสตร์อื่น ๆ โดยเฉพาะเนื้อหาที่มีความเป็นนามธรรมสูง เช่น เรื่องความน่าจะเป็น เพื่อส่งเสริมความเข้าใจและพัฒนาความสามารถทางคณิตศาสตร์ของนักเรียนให้มีประสิทธิภาพยิ่งขึ้น

2) การวิจัยในอนาคตควรดำเนินการติดตามผลการเรียนรู้ในระยะยาว เพื่อประเมินความคงทนของความรู้และความสามารถในการให้เหตุผลทางคณิตศาสตร์ภายหลังการจัดการเรียนรู้แบบ CPA ร่วมกับเทคนิคเพื่อนคู่คิด

3) การวิจัยในอนาคตควรศึกษาตัวแปรอื่น ๆ ที่อาจมีความสัมพันธ์หรือส่งผลต่อความสามารถในการให้เหตุผลทางคณิตศาสตร์ อาทิ ความคิดสร้างสรรค์ ทักษะการแก้ปัญหา หรือแรงจูงใจทางการเรียน ซึ่งจะช่วยให้เข้าใจปัจจัยที่ส่งผลต่อการพัฒนาความสามารถทางคณิตศาสตร์ได้อย่างรอบด้าน

4) การวิจัยในอนาคตควรนำเทคโนโลยีหรือสื่อดิจิทัลบูรณาการร่วมกับรูปแบบการจัดการเรียนรู้แบบ CPA เพื่อเพิ่มความน่าสนใจและประสิทธิภาพในการเรียนรู้

กิตติกรรมประกาศ งานวิจัยฉบับนี้สำเร็จลุล่วงไปด้วยดี ผู้วิจัยขอขอบคุณสำนักการศึกษา กรุงเทพมหานคร ที่มอบทุนการศึกษาในประเทศ ระดับปริญญาโท สำหรับข้าราชการครูกรุงเทพมหานครและบุคลากรทางการศึกษา กรุงเทพมหานคร ประจำปีงบประมาณ พ.ศ. 2566 เป็นอย่างยิ่ง

เอกสารอ้างอิง

[1] เมขลิน อมรรัตน์, ประโยชน์ของคณิตศาสตร์ในชีวิตประจำวัน ตอนที่ 2 การให้เหตุผล. <https://www.scimath.org/article-mathematics/item/11465-2-11465>, 2563 (สืบค้น 9 ธ.ค. 67).

[2] สุวัฒนา เอี่ยมอรพรรณ, วิธีและเทคนิคการสอนคณิตศาสตร์เพื่อพัฒนาการคิด, โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร, 2549.

[3] D. Stylianou, E. Silver, The role of visual representations in mathematical problem solving: An examination of expert-novice similarities and differences Mathematical, 6(4) (2004) 353–387.

[4] สุธีรา จันทร์เกตุ, การพัฒนากิจกรรมการเรียนรู้ตามแนวคิด Concrete-Pictorial-Abstract (CPA) เพื่อเสริมสร้างมโนทัศน์ ความสามารถในการสื่อสาร สื่อความหมายทางคณิตศาสตร์และการนำเสนอ เรื่อง ทศนิยม สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4, (การค้นคว้าอิสระมหาบัณฑิต), มหาวิทยาลัยนเรศวร, 2561.

[5] ทิพย์วรรณ จันทร์เขียว, การจัดการเรียนรู้โดยใช้กระบวนการแก้ปัญหาเชิงสร้างสรรค์ร่วมกับเทคนิคเพื่อนคู่คิด ที่มีต่อความสามารถในการแก้ปัญหาเชิงสร้างสรรค์ทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1, (ปริญญานิพนธ์มหาบัณฑิต), มหาวิทยาลัยศรีนครินทรวิโรฒ, 2564.

[6] ศิริชัย กาญจนวาสี, การคำนวณคะแนนพัฒนาการ (Gain Scores), วารสารสมาคมวิจัยสังคมศาสตร์แห่งประเทศไทย. 1(1) (2557) 1-20.

[7] นลิตา สอนวารี, ปกรชัย เมืองโคตร, รุ่งโรจน์ ศรีจันทร์แก้ว, ผลของการใช้เทคนิคการเรียนรู้แบบหลักการ CPA ที่มีต่อทักษะการแก้โจทย์ปัญหาทางคณิตศาสตร์ เรื่อง โจทย์ปัญหาทฤษฎีบทพีทาโกรัสของนักเรียนชั้นมัธยมศึกษาปีที่ 2, วารสารวิจัยทางการศึกษามหาวิทยาลัยศรีนครินทรวิโรฒ 19(2) (2567) 42-55.

[8] อังคณา ทิพยรุ่งโรจน์, การพัฒนาชุดกิจกรรมการเรียนรู้คณิตศาสตร์โดยใช้การจัดการเรียนรู้ตามแนวคิด Concrete-Pictorial-Abstract และแนวคิด Model-Eliciting Activities เรื่อง ทศนิยม เพื่อยกระดับผลสัมฤทธิ์ทางการเรียนและทักษะการแก้ปัญหาทางคณิตศาสตร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 5, (วิทยานิพนธ์มหาบัณฑิต), มหาวิทยาลัยราชภัฏยะลา, 2566.

[9] ธัญญา แนวดง, คงรัฐ นวลแปง, คมสัน ตรีไพบูลย์, ผลการจัดการเรียนรู้แบบอุปนัยร่วมกับเทคนิค Think-Pair - Share ที่มีต่อมโนทัศน์และความสามารถในการให้เหตุผลทางคณิตศาสตร์, วารสารมนุษย์และสังคมศาสตร์มหาวิทยาลัยราชภัฏสุรินทร์. 20(1) (2561) 29-41.

[10] ชาญชัย ชนไพโรจน์, ปวีณา ขันธ์ศิลา, ประภาพร หนองหารพิทักษ์, การพัฒนาบทเรียนเรื่อง ปริซึมและทรงกระบอก โดยใช้กระบวนการจัดการเรียนรู้ CPA ร่วมกับโปรแกรม GeoGebra เพื่อเพิ่มผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นมัธยมศึกษาปีที่ 2, วารสารครุศาสตร์อุตสาหกรรมสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง. 23(2) (2567) 15-26.

[11] รัตนาภรณ์ ไทยพู, รสริน เจิมไธสง, พรภิรมย์ หลงทรัพย์, การจัดการเรียนรู้ตามแนวคิด Concrete-Pictorial-Abstract ร่วมกับบาร์โมเดลเพื่อพัฒนาผลสัมฤทธิ์ทางการเรียน รายวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2, วารสารวิชาการมหาวิทยาลัยราชภัฏเพชรบุรี. 13(2) (2566) 45-53.

[12] อัจจิมา บุญพัตร์, ผลการจัดการเรียนรู้โดยใช้ปัญหาเป็นฐานร่วมกับเทคนิคเพื่อนคู่คิดที่มีต่อผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์และความยึดมั่นผูกพันในการเรียนของนักเรียนชั้นมัธยมศึกษาปีที่ 1, (วิทยานิพนธ์มหาบัณฑิต), มหาวิทยาลัยธุรกิจบัณฑิตย์, 2566.

ผลการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ที่มีต่อความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 4

เริงฤทธิ์ รัตนชวางกูร[1,*] และ รัชนิกร ชลไชยะ[2]

[1]สาขาวิทยาศาสตร์ศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี 10140

[2]ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี 10140

บทคัดย่อ

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาผลการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ที่มีต่อความสามารถในการเชื่อมโยงคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 4 โดยใช้ระเบียบวิธีวิจัยเชิงทดลองเบื้องต้นแบบผสมผสาน เครื่องมือวิจัยประกอบด้วย แบบวัดความสามารถในการเชื่อมโยงทางคณิตศาสตร์ (ค่าความเชื่อมั่นเท่ากับ 0.748) คำถามสัมภาษณ์ บันทึกหลังการสอน และแผนการจัดการเรียนรู้ ประชากรเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 จำนวน 299 คน จาก 7 ห้อง ภาคการเรียนที่ 1 ปีการศึกษา 2567 ของโรงเรียนชายล้วน โดยผู้วิจัยใช้วิธีการสุ่มตัวอย่างแบบแบ่งกลุ่มเพื่อเลือกกลุ่มตัวอย่างในห้องเรียนจำนวน 36 คน โดยแต่ละห้องเรียนมีนักเรียนแบบคละความสามารถ ซึ่งข้อมูลถูกวิเคราะห์ด้วยสถิติพรรณนา ได้แก่ ค่าเฉลี่ย ร้อยละ ค่าความเบี่ยงเบนมาตรฐานและคะแนนเพิ่มสัมพัทธ์ และการทดสอบทางสถิติ ได้แก่ Wilcoxon signed-rank test ผลการวิจัยพบว่า นักเรียนที่ผ่านการจัดการเรียนรู้แบบห้องเรียนกลับด้านโดยใช้เทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed มีการพัฒนาความสามารถในการเชื่อมโยงทางคณิตศาสตร์เพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 อย่างไรก็ตาม นักเรียนบางส่วนมีความต้องการเรียนกับครูมากกว่าการเรียนรู้ด้วยตนเองผ่านวิดีโอการสอน เนื่องจากความเคยชินจากการเรียนรู้แบบบรรยาย ดังนั้นการจัดการเรียนรู้แบบห้องเรียนกลับด้านจึงควรได้รับความร่วมมือทั้งนักเรียนและครูเพื่อเพิ่มประสิทธิภาพในการเรียนรู้อย่างมีประสิทธิผล

คำสำคัญ: ความสามารถในการเชื่อมโยงทางคณิตศาสตร์, เทคนิคตั๋วออก, เว็บไซต์ TED-Ed, ห้องเรียนกลับด้าน

2020 MSC: 97D40

*ผู้นำเสนอและผู้ประพันธ์บรรณกิจ (roengrit.ratt@kmutt.ac.th)

1 บทนำ

1.1 ที่มาและความสำคัญ

ความสามารถในการเชื่อมโยงทางคณิตศาสตร์เป็นทักษะการคิดทางคณิตศาสตร์ขั้นสูง [1] ซึ่งเป็นสิ่งที่จำเป็นต่อการพัฒนาการเรียนรู้ของนักเรียน ช่วยให้นักเรียนเข้าใจแนวคิดทางคณิตศาสตร์และพัฒนาความเข้าใจในศาสตร์ต่าง ๆ ผ่านการเชื่อมโยงแนวคิดทางคณิตศาสตร์กับแนวคิดทางคณิตศาสตร์ในด้านอื่น ๆ หรือวิชาต่าง ๆ หรือเชื่อมโยงแนวคิดทางคณิตศาสตร์กับปัญหาโลกจริง [2-4] จากการศึกษาของ NCTM (National Council of Teachers of Mathematics) [39] ได้ระบุว่า การเชื่อมโยงทางคณิตศาสตร์เป็นหนึ่งในมาตรฐานกระบวนการทางคณิตศาสตร์ที่สำคัญ ซึ่งช่วยให้นักเรียนเห็นคณิตศาสตร์เป็นองค์ความรู้ที่เชื่อมโยงกันเป็นหนึ่งเดียว ไม่ใช่เพียงเนื้อหาที่แยกส่วน เมื่อนักเรียนมีความสามารถในการเชื่อมโยงทางคณิตศาสตร์ที่สูงขึ้น ส่งผลให้นักเรียนมีความรู้และความเข้าใจในบทเรียนทางคณิตศาสตร์ที่ดีขึ้น อีกทั้งนักเรียนยังสามารถนำแนวคิดทางคณิตศาสตร์ไปประยุกต์ใช้ในการแก้ปัญหาในโลกจริงได้ [1, 5, 7] ซึ่งความสามารถในการเชื่อมโยงทางคณิตศาสตร์เป็นกุญแจสำคัญในการสร้างความเข้าใจทางคณิตศาสตร์ ตลอดจนสร้างความเชื่อมโยงระหว่างแนวคิดและกระบวนการทางคณิตศาสตร์และศาสตร์อื่น ๆ [3, 5, 6] ดังนั้นความสามารถในการเชื่อมโยงทางคณิตศาสตร์จึงเป็นทักษะสำคัญสำหรับนักเรียน เพื่อเสริมสร้างความเข้าใจในแนวคิดทางคณิตศาสตร์ อีกทั้งยังช่วยพัฒนาความเข้าใจในศาสตร์อื่น ๆ และสามารถนำไปประยุกต์ใช้ในการแก้ปัญหาในโลกจริงอีกด้วย [5]

ในปีการศึกษา 2564 ผู้วิจัยได้เข้าฝึกสอนที่โรงเรียนแห่งหนึ่งในกรุงเทพมหานคร ซึ่งโรงเรียนนี้เป็นโรงเรียนชายล้วนขนาดใหญ่ที่มีสภาพแวดล้อมที่เอื้อต่อการเรียนรู้และการทำกิจกรรมต่าง ๆ ในโรงเรียน เช่น การเล่นกีฬา เป็นต้น อีกทั้งนักเรียนยังมีการแข่งขันด้านการเรียนค่อนข้างสูง ส่งผลให้นักเรียนมุ่งเน้นการเรียนที่ใช้สำหรับการสอบเพียงอย่างเดียว อีกทั้งนักเรียนส่วนใหญ่ไม่ให้ความสำคัญกับการมีส่วนร่วมในชั้นเรียน เช่น การแสดงความคิดเห็น การตอบคำถาม เป็นต้น ทำให้นักเรียนขาดการมีปฏิสัมพันธ์ระหว่างครูและนักเรียน ซึ่งสอดคล้องกับผู้วิจัยที่ได้ทำการสอนนักเรียนมัธยมศึกษาปีที่ 4 ที่โรงเรียนนี้ในปีการศึกษา 2564 พบว่า นักเรียนมีพื้นฐานทางด้านคณิตศาสตร์ระดับปานกลาง ซึ่งนักเรียนสามารถทำแบบฝึกหัดหรือแก้โจทย์ปัญหาทางคณิตศาสตร์ได้ อย่างไรก็ตาม เมื่อผู้วิจัยต้องการให้นักเรียนอธิบายถึงแนวคิดทางคณิตศาสตร์หรือความรู้ที่นำมาใช้ในการแก้โจทย์ปัญหานั้นนักเรียนกลับไม่สามารถอธิบายได้ เนื่องจากนักเรียนมุ่งเน้นเพียงกระบวนการแก้โจทย์ปัญหาแต่ไม่ได้เข้าใจถึงแนวคิดทางคณิตศาสตร์ นอกจากนี้นักเรียนได้เผชิญกับโจทย์ปัญหาคณิตศาสตร์จากโลกจริง นักเรียนมักจะรอครูแก้โจทย์ปัญหานี้ตลอดโดยที่ไม่แสดงถึงความพยายามในการแก้โจทย์ปัญหา เพราะนักเรียนไม่สามารถประยุกต์ความรู้ทางคณิตศาสตร์ในการแก้โจทย์ปัญหาได้ เมื่อนักเรียนเริ่มแก้โจทย์ปัญหาในโลกจริงด้วยตนเอง นักเรียนมักพูดว่า “ครูครับ พวกผมต้องแก้โจทย์ปัญหานี้อย่างไรบ้าง” “ครูครับ ช่วยบอกแนวทางในการแก้โจทย์ปัญหานี้หน่อยครับ” และ “ครูครับ เฉลยวิธีแก้โจทย์ปัญหานี้เลยได้ไหมครับ ผมไม่รู้ว่าจะทำอย่างไรดี” นอกจากนี้เมื่อนักเรียนต้องทำโจทย์ปัญหาทางคณิตศาสตร์ด้วย

ตนเอง นักเรียนมักต้องการให้ครูช่วยตรวจสอบความถูกต้องของคำตอบอยู่เสมอ เพราะนักเรียนขาดความมั่นใจในการโจทย์ปัญหาทางคณิตศาสตร์ ทำให้นักเรียนต้องหมั่นตรวจสอบความถูกต้องของคำตอบตลอดเวลา

นอกจากนี้ผู้วิจัยได้สัมภาษณ์ครูที่สอนในโรงเรียนนี้เป็นระยะเวลาอย่างน้อย 5 ปี ทั้งหมด 2 คน เกี่ยวกับรูปแบบการเรียนการสอนและพฤติกรรมการเรียนของนักเรียนในโรงเรียนนี้ โดยวิธีการสอนในรายวิชาคณิตศาสตร์ของโรงเรียนนี้เป็นการสอนแบบบรรยายคือ ครูจะสอนเพียงเนื้อหาและทฤษฎีต่าง ๆ ในบทเรียนแค่ละคาบ หลังจากนั้นครูจะมอบหมายให้นักเรียนทำแบบฝึกหัดจากเอกสารประกอบการเรียนเพียงอย่างเดียว ซึ่งครูไม่ได้มีกิจกรรมการเรียนรู้หรือสื่อการเรียนอื่น ๆ นำมาใช้กับการเรียนการสอบของวิชาคณิตศาสตร์ เนื่องจากจำนวนชั่วโมงในการเรียนไม่สอดคล้องกับปริมาณเนื้อหาของบทเรียน เพราะโรงเรียนนี้มีวันหยุดค่อนข้างเยอะ อีกทั้งกิจกรรมภายในโรงเรียนและกีฬาประเพณีที่ใช้เวลาค่อนข้างนาน จึงเป็นสาเหตุที่ให้เวลาในการเรียนของนักเรียนนั้นน้อยลงเป็นอย่างมาก ดังนั้นครูในโรงเรียนนี้จึงมีความจำเป็นอย่างยิ่งที่ต้องสอนแบบบรรยายเพื่อที่ครูจะสามารถสอนให้ครบตามเนื้อหาได้ ซึ่งการสอนแบบบรรยายนั้นทำให้นักเรียนไม่มีปฏิสัมพันธ์กับครูและไม่มีส่วนร่วมในชั้นเรียน ส่งผลให้นักเรียนไม่สามารถเข้าใจคณิตศาสตร์ได้อย่างลึกซึ้ง [8] อีกทั้งการเรียนรู้แบบบรรยายไม่มีประสิทธิภาพต่อการเรียนรู้ของนักเรียน รวมถึงไม่ส่งเสริมความคิดสร้างสรรค์ของนักเรียนอีกด้วย [9-11] ในส่วนของพฤติกรรมการเรียนรู้ของนักเรียน พบว่านักเรียนไม่ค่อยให้ความร่วมมือในชั้นเรียน นอกจากนี้นักเรียนยังมีพฤติกรรมไม่เหมาะสมในชั้นเรียน เช่น พูดเสียงดัง เป็นต้น ส่งผลให้ครูต้องตั้งกฎในห้องเรียนเพื่อปรับเปลี่ยนพฤติกรรมของนักเรียนให้ดีขึ้น บางครั้งการใช้กฎในห้องเรียนไม่อาจปรับเปลี่ยนพฤติกรรมเหล่านี้ได้ ครูจึงต้องตักเตือนนักเรียนเพื่อให้หยุดพฤติกรรมไม่เหมาะสมในชั้นเรียน อย่างไรก็ตามการที่ครูใช้วิธีการแบบนี้ไม่ส่งเสริมต่อการเรียนรู้และแรงจูงใจของนักเรียน [12] นอกจากนี้ครูยังกล่าวอีกว่าคณิตศาสตร์เป็นวิชาที่นักเรียนไม่ชอบอย่างยิ่งเพราะนักเรียนมีมุมมองต่อวิชาคณิตศาสตร์ว่าเป็นวิชาที่มีสูตรในการคำนวณจำนวนมากและไม่สามารถทราบได้ว่าจะนำความรู้ทางคณิตศาสตร์ไปประยุกต์ได้อย่างไรนอกจากการสอบเพียงอย่างเดียว จึงเป็นเหตุให้นักเรียนไม่มีความต้องการเรียนวิชาคณิตศาสตร์ จากข้อมูลทั้งหมดที่กล่าวมาข้างต้นสามารถสรุปได้ว่านักเรียนมีความสามารถในการเชื่อมโยงทางคณิตศาสตร์ในระดับต่ำ จึงจำเป็นต้องพัฒนาความสามารถนี้เพื่อให้นักเรียนเกิดกระบวนการคิดทางคณิตศาสตร์และสามารถนำแนวคิดทางคณิตศาสตร์ไปประยุกต์ใช้กับปัญหาโลกจริงได้

จากการทบทวนวรรณกรรมพบว่า รูปแบบการจัดการเรียนรู้ที่สามารถนำมาใช้เพื่อพัฒนาความสามารถในการเชื่อมโยงทางคณิตศาสตร์ เช่น การจัดการเรียนรู้แบบปัญหาเป็นฐาน การจัดการเรียนรู้แบบเกมเป็นฐาน การจัดกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้าน เป็นต้น ในการศึกษาครั้งนี้ ผู้วิจัยได้เลือกการจัดการเรียนรู้แบบห้องเรียนกลับด้าน เนื่องจากการเรียนรูปแบบนี้มีแนวโน้มช่วยลดเวลาเรียนในชั้นเรียน และเพิ่มกิจกรรมที่เน้นการมีส่วนร่วมในชั้นเรียนของนักเรียน ห้องเรียนกลับด้าน (Flipped classroom) เป็นรูปแบบการเรียนรู้ที่นักเรียนสามารถจัดการเวลาเรียนของตนเองได้ โดยรูปแบบการเรียนนี้สามารถแบ่งออกเป็น 2 ส่วน ได้แก่ การเรียนรู้นอกชั้นเรียน เป็นการเรียนรู้ผ่านการศึกษาวิดีโอการสอนหรือสื่อออนไลน์ที่ครูได้จัดเตรียมให้กับนักเรียนก่อนเข้าในชั้นเรียน และการเรียนรู้ในชั้นเรียน เป็นการเรียนรู้ผ่านกิจกรรมในชั้นเรียนที่ส่งเสริมการมีส่วนร่วมในชั้นเรียนของนักเรียน เช่น การตอบ

คำถามในชั้นเรียน การทำกิจกรรมกลุ่ม การทำโจทย์ปัญหา เป็นต้น [13, 14] ห้องเรียนกลับด้านเปิดโอกาสให้นักเรียนได้จัดสรรการเรียนรู้ด้วยตนเองและสร้างองค์ความรู้ก่อนเข้าห้องเรียน [14, 18] ช่วยพัฒนาประสิทธิภาพในการเรียนรู้และความมั่นใจของนักเรียนหลังจากได้เรียนผ่านการเรียนรู้แบบห้องเรียนกลับด้าน [16] อย่างไรก็ตามการจัดกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านจะสำเร็จได้ขึ้นอยู่กับการให้ความร่วมมือทั้งครูและนักเรียน โดยครูจะต้องออกกิจกรรมในชั้นเรียนที่ส่งเสริมให้นักเรียนต้องเตรียมล่วงหน้าและมีส่วนร่วมในชั้นเรียนอย่างมีประสิทธิภาพ ส่วนนักเรียนจะต้องรับผิดชอบในการดูวิดีโอการสอนก่อนเข้าชั้นเรียนด้วย [19]

สำหรับสื่อออนไลน์ที่ใช้ในการเรียนรู้นอกชั้นเรียนของการจัดกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านควรเป็นสื่อการเรียนรู้ที่มีเนื้อหาครอบคลุม มีระดับความยากง่ายที่นักเรียนสามารถศึกษาและสามารถทำความเข้าใจได้ด้วยตนเอง อีกทั้งสื่อออนไลน์ต้องมีเนื้อหาที่น่าสนใจและนักเรียนสามารถประเมินตนเองได้หลังจากที่ศึกษาด้วยตัวเอง [22, 23] เว็บไซต์ TED-Ed เป็นแพลตฟอร์มที่มุ่งเน้นการสร้างวิดีโอด้านการศึกษา [17, 24, 25] ซึ่งครูสามารถใช้วิดีโอเพื่อสร้างบทเรียนที่น่าสนใจ อีกทั้งยังสามารถเพิ่มคำถามเพื่อให้นักเรียนตอบและแสดงให้เห็นว่าพวกเขาได้เรียนรู้แล้วโดยการรับชมวิดีโอ TED-Ed เนื่องจากบทเรียนใน TED-Ed ประกอบด้วย 4 ส่วน ได้แก่ Watch, Think, Dig, และ Discuss ทำให้นักเรียนมีปฏิสัมพันธ์กับการเรียนรู้นอกชั้นเรียน [24, 25]

นอกจากนี้การประเมินนักเรียนอย่างต่อเนื่องช่วยเพิ่มทั้งความเข้าใจในบทเรียนและประสิทธิภาพในการเรียนรู้ อีกทั้งยังช่วยให้นักเรียนมีส่วนร่วมในชั้นเรียนอย่างสม่ำเสมอ [26] ตั๋วออก (Exit Ticket) เป็นแบบประเมินความก้าวหน้าที่ส่งเสริมการเรียนรู้ของนักเรียน [20, 21] ครูมอบหมายให้นักเรียนทำทันทีหลังจบบทเรียนหรือคาบเรียนและสนับสนุนให้นักเรียนได้เรียนรู้ด้วยตนเอง ซึ่งเป็นวิธีการประเมินแบบหนึ่งที่ง่ายในการประเมินการเรียนรู้ของนักเรียนและแสดงหลักฐานว่านักเรียนเข้าใจหรือไม่เข้าใจบทเรียนใด นอกจากนี้ยังช่วยให้นักเรียนได้ทบทวนความรู้เกี่ยวกับเนื้อหาที่ยังไม่เข้าใจด้วย [20, 21, 27]

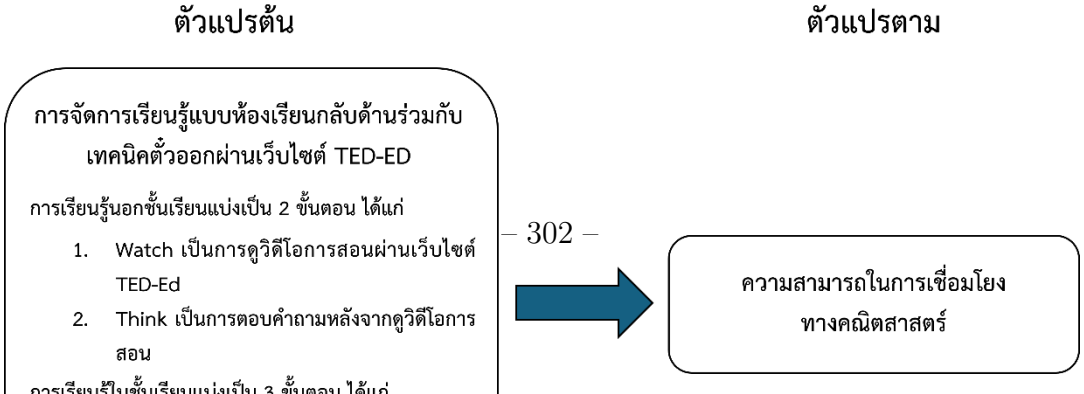
จากปัญหาและแนวคิดที่กล่าวมาข้างต้น ผู้วิจัยเลือกใช้การจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ที่มีผลต่อความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 4

1.2 วัตถุประสงค์

เพื่อศึกษาผลการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ที่มีต่อความสามารถในการเชื่อมโยงคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 4

1.3 สมมติฐานของงานวิจัย

ภายหลังจากการจัดการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่าเว็บไซต์ TED-Ed นักเรียนชั้นมัธยมศึกษาปีที่ 4 มีความสามารถในการเชื่อมโยงทางคณิตศาสตร์ตั้งแต่ร้อยละ 50 ขึ้นไป



รูปที่ 1: กรอบแนวคิดของงานวิจัย

1.4 ผลลัพธ์ที่คาดหวัง

1) นักเรียนที่เข้าร่วมการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed สามารถประยุกต์ใช้ความรู้และสามารถเชื่อมโยงความรู้ทางคณิตศาสตร์ในการแก้ปัญหาในชีวิตจริง
2) นักเรียนที่เข้าร่วมการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed มีความมั่นใจเพิ่มขึ้นในการแก้ปัญหาในชีวิตจริง

2 ความรู้พื้นฐาน

2.1 ห้องเรียนกลับด้าน

ห้องเรียนกลับด้าน (Flipped Classroom) เป็นการเรียนรู้ที่มีนักเรียนเป็นศูนย์กลาง โดยจะแบ่งรูปแบบการเรียนออกเป็น 2 ส่วน ได้แก่ การเรียนรู้นอกชั้นเรียนและการเรียนรู้ในชั้นเรียน [13, 15] การเรียนรู้นอกชั้นเรียนเป็นการเรียนรู้ก่อนเข้าชั้นเรียน โดยที่ครูจะให้นักเรียนไปทบทวนบทเรียนจากสื่อออนไลน์ เช่น YouTube, Khan Academy, TED-Ed เป็นต้น วิดีโอการสอนหรือเอกสารการเรียนที่ครูได้จัดสรรให้ก่อนที่จะเข้าร่วมกิจกรรมในชั้นเรียน [13-17] ส่วนการเรียนรู้ในชั้นเรียนเป็นการเรียนรู้ที่นำความรู้พื้นฐานที่ได้จากการเรียนรู้นอกชั้นเรียนมาประยุกต์ใช้กับกิจกรรมในชั้นเรียน โดยครูสามารถดึงดูดนักเรียนให้เข้าร่วมกิจกรรมการเรียนรู้ผ่านการทำแบบฝึกหัด โครงงาน การอภิปรายและการแก้ปัญหาในชั้นเรียน [13, 14, 16]

2.2 ตั๋วออก

ตั๋วออก (Exit Ticket) เป็นแบบประเมินความก้าวหน้าที่ส่งเสริมการเรียนรู้ของนักเรียน โดยครูจะมอบหมายให้นักเรียนทำทันทีหลังจบบทเรียนหรือคาบเรียน และสนับสนุนให้นักเรียนเรียนรู้ด้วยตนเอง [20, 21] หลังจากที่

นักเรียนทำตั๋วออกเสร็จครูจะทำการประเมินนักเรียนว่านักเรียนเข้าใจหรือไม่เข้าใจบทเรียนไหนบ้าง เพื่อให้นักเรียนได้ทบทวนบทเรียนที่ตนเองยังไม่เข้าใจทำให้นักเรียนเกิดความเข้าใจมากยิ่งขึ้น [20, 21, 27, 28]

2.3 เว็บไซต์ TED-Ed

เว็บไซต์ TED-Ed เป็นแพลตฟอร์มที่มุ่งเน้นการสร้างวิดีโอด้านการศึกษา [17, 24, 25] ซึ่งครูสามารถใช้วิดีโอเพื่อสร้างบทเรียนที่น่าสนใจ อีกทั้งยังสามารถเพิ่มคำถามเพื่อให้นักเรียนตอบและแสดงให้เห็นว่าพวกเขาได้เรียนรู้แล้ว โดยการรับชมวิดีโอ TED-Ed เนื่องจากบทเรียนใน TED-Ed ประกอบด้วย 4 ส่วน ได้แก่ Watch, Think, Dig, และ Discuss ทำให้นักเรียนมีปฏิสัมพันธ์กับการเรียนรู้นอกชั้นเรียน [24, 25]

2.4 ความสามารถในการเชื่อมโยงทางคณิตศาสตร์

ความสามารถในการเชื่อมโยงทางคณิตศาสตร์ (Mathematical Connection Ability) เป็นความสามารถที่เชื่อมโยงระหว่างแนวคิดทางคณิตศาสตร์กับศาสตร์อื่น หรือคณิตศาสตร์ด้านอื่น ๆ หรือในชีวิตจริง [2, 3, 29] ซึ่งประกอบไปด้วยแนวคิดที่สำคัญในการทำความเข้าใจและพัฒนาระหว่างแนวคิดและกระบวนการทางคณิตศาสตร์ [3] เมื่อนักเรียนมีความสามารถในการเชื่อมโยงทางคณิตศาสตร์ที่สูงขึ้น ส่งผลให้นักเรียนมีความรู้และความเข้าใจเนื้อหาในคณิตศาสตร์ดีขึ้น อีกทั้งยังสามารถนำแนวคิดทางคณิตศาสตร์ไปประยุกต์ใช้ในการแก้ปัญหาในโลกจริงได้ [1, 5, 7, 29]

3 วิธีดำเนินวิจัย

งานวิจัยนี้เป็นการวิจัยเชิงทดลองเบื้องต้นแบบผสมผสาน โดยการเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูลของงานวิจัยนี้ประกอบไปด้วยข้อมูลเชิงปริมาณและข้อมูลเชิงคุณภาพ

3.1 ประชากรและกลุ่มตัวอย่าง

3.1.1 ประชากร

ประชากรที่ใช้ในงานวิจัยนี้คือ นักเรียนชั้นมัธยมศึกษาปีที่ 4 ภาคการศึกษาปีที่ 1 ปีการศึกษา 2567 โรงเรียนชายล้วน ทั้งหมด 7 ห้องเรียน จำนวนนักเรียนทั้งหมด 299 คน

3.1.2 กลุ่มตัวอย่าง

กลุ่มตัวอย่างจะใช้วิธีการสุ่มตัวอย่างแบบกลุ่มโดยเลือกกลุ่มตัวอย่าง 1 ห้อง จำนวนนักเรียนทั้งหมด 36 คน ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ภาคการศึกษาปีที่ 1 ปีการศึกษา 2567 โรงเรียนชายล้วน กรุงเทพมหานคร ประเทศไทย

3.2 เครื่องมือวิจัย

เครื่องมือวิจัยที่ใช้สำหรับงานวิจัยนี้ประกอบไปด้วย แผนการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ทั้งหมด 10 แผน แบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์ก่อนเรียนและหลังเรียน

สำหรับแผนการจัดการกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ได้ผ่านการตรวจสอบความสอดคล้องของแผนการสอน (IOC) โดยใช้เกณฑ์ 5 ระดับ จากผู้เชี่ยวชาญทั้งหมด 3 คน ซึ่งคะแนนอยู่ในช่วง 4.00-5.00 (ระดับดีมาก) ส่วนแบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์ได้ผ่านการตรวจสอบความสอดคล้องของเนื้อหา (IOC) จากผู้เชี่ยวชาญทั้งหมด 3 คน โดยมีคะแนนอยู่ในช่วง 0.67-1.00 และได้ทดสอบเครื่องมือวิจัยกับนักเรียนมัธยมศึกษาปีที่ 5 ผู้ซึ่งไม่ใช่กลุ่มตัวอย่าง โดยค่าความยากง่ายของแบบทดสอบอยู่ในช่วง 0.23-0.63 ค่าอำนาจจำแนกของแบบทดสอบอยู่ในช่วง 0.30-0.78 และค่าความเชื่อมั่น (Alpha Cronbach coefficient) ของแบบทดสอบเท่ากับ 0.748

3.3 นิยามศัพท์เฉพาะ

1. การจัดกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านเป็นการเรียนรู้ที่แบ่งออกเป็น 2 ส่วน ได้แก่ การเรียนรู้นอกชั้นเรียนและการเรียนรู้ในชั้นเรียน การเรียนรู้นอกชั้นเรียนเป็นการเรียนรู้จากการศึกษาวิดีโอการสอนผ่านเว็บไซต์ TED-Ed ส่วนการเรียนรู้ในชั้นเรียนเป็นการเรียนรู้โดยใช้การแก้ปัญหาโลกจริง ซึ่งรูปแบบการเรียนรู้ทั้ง 2 รูปแบบ มีขั้นตอนดังต่อไปนี้

การเรียนรู้นอกชั้นเรียน: การเรียนรู้ผ่านเว็บไซต์ TED-Ed ประกอบด้วย 2 ขั้นตอน

1) Watch เป็นการศึกษาวิดีโอการสอนที่ครูได้มอบหมายไว้เพื่อศึกษาก่อนเข้าชั้นเรียน
2) Think เป็นการตอบคำถามหลังจากศึกษาวิดีโอการสอนเสร็จแล้ว

การเรียนรู้ในชั้นเรียนมีขั้นตอนทั้งหมด 3 ขั้นตอน

1) ขั้นทบทวนความรู้เป็นการทบทวนความรู้หลังจากที่นักเรียนศึกษาวิดีโอการสอนจากเว็บไซต์ TED-Ed
2) ขั้นแก้ปัญหาโลกจริง
3) ขั้นสรุปเป็นการสรุปความรู้ที่ได้จากการแก้ปัญหาโลกจริง อีกทั้งนักเรียนต้องใบตั๋วออกก่อนจบคาบเรียน

2. ใบตั๋วออกเป็นแบบประเมินความก้าวหน้าที่ส่งเสริมการเรียนรู้ของนักเรียนก่อนจบคาบเรียน โดยคำถามในใบตั๋วออกแบ่งออกเป็น 2 ส่วน ได้แก่ ความมั่นใจในเนื้อหาทางคณิตศาสตร์เป็นการวัดความเข้าใจในบทเรียน ซึ่งคำถามจะมีความสอดคล้องกับบทเรียนในแต่ละคาบเพื่อตรวจสอบความเข้าใจและข้อผิดพลาดของนักเรียน และความมั่นใจในการเรียนรู้ทางคณิตศาสตร์เป็นการวัดความมั่นใจในบทเรียน ซึ่งเป็นคำถามเพื่อตรวจสอบความมั่นใจในบทเรียนในแต่ละคาบ

3. เว็บไซต์ TED-Ed เป็นแพลตฟอร์มที่มุ่งเน้นการสร้างวิดีโอด้านการศึกษา ซึ่งครูสามารถใช้วิดีโอเพื่อสร้างบทเรียนที่น่าสนใจ ซึ่งประกอบด้วย 2 ส่วน ได้แก่ Watch และ Think ในส่วน Watch เป็นการศึกษาวิดีโอการ

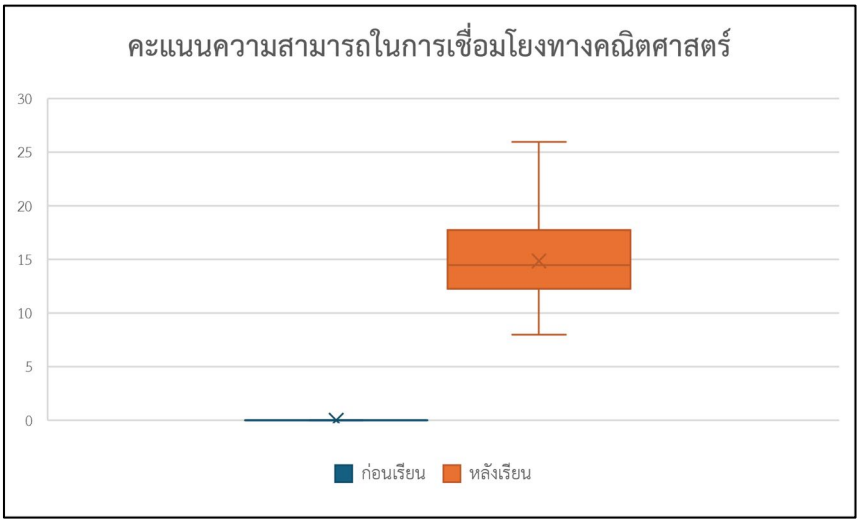
สอนที่ครูได้มอบหมายไว้เพื่อศึกษาก่อนเข้าชั้นเรียน ซึ่งเนื้อหาในวิดีโอการสอนจะสอดคล้องกับหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 โดยมีทั้งหมด 17 วิดีโอ ส่วน Think เป็นการตอบคำถามหลังจากศึกษาวิดีโอการสอนเสร็จแล้วเพื่อตรวจสอบความเข้าใจของนักเรียน

4. ความสามารถในการเชื่อมโยงทางคณิตศาสตร์เป็นความสามารถที่เชื่อมโยงระหว่างแนวคิดทางคณิตศาสตร์กับศาสตร์อื่น หรือคณิตศาสตร์ด้านอื่น ๆ หรือในชีวิตจริง ซึ่งประเมินได้จากข้อเขียนของแบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์ทั้งหมด 4 ข้อ

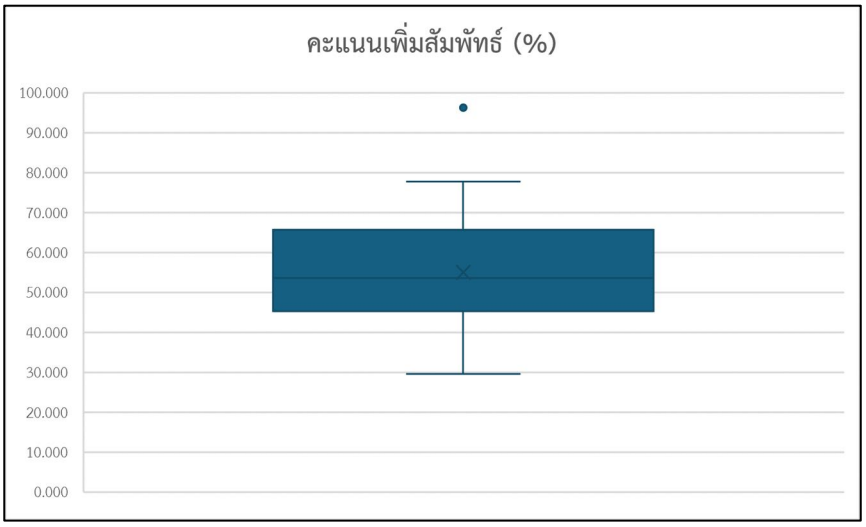
4 ผลการวิจัย

4.1 ผลการทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียน

ในส่วนนี้จะแสดงคะแนนความสามารถในการเชื่อมโยงทางคณิตศาสตร์เฉลี่ยก่อนเรียนและหลังเรียนของนักเรียนทั้งหมด โดยแบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์เป็นข้อสอบแบบแสดงวิธีทำมีทั้งหมด 4 ข้อ คะแนนเต็ม 27 คะแนน การพัฒนาของนักเรียนคำนวณโดยใช้คะแนนเพิ่มสัมพัทธ์ (relative gain score) หลังจากนักเรียนเข้าร่วมการจัดกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed แสดงดังรูปที่ 2 และ 3



รูปที่ 2: แผนภาพกล่องแสดงข้อมูลของคะแนนความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียนก่อนเรียนและหลังเรียน



รูปที่ 3: แผนภาพกล่องแสดงข้อมูลของคะแนนเพิ่มสัมพัทธ์จากคะแนนความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียน

จากรูปที่ 2 และ 3 จะเห็นได้ว่า ค่าเฉลี่ยคะแนนเพิ่มสัมพัทธ์ของนักเรียนเท่ากับ 55.04% ค่าเฉลี่ยคะแนนหลังเรียนเท่ากับ 14.86 คะแนน และคะแนนก่อนเรียนของนักเรียนทุกคนเท่ากับ 0 คะแนน เนื่องจากแบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์เป็นข้อสอบแบบอัตนัยที่ต้องแสดงวิธีทำอย่างละเอียดและประยุกต์ความรู้ทางคณิตศาสตร์กับชีวิตจริง ส่งผลให้นักเรียนไม่สามารถคาดเดาคำตอบได้ ดังนั้นนักเรียนทุกคนจึงได้คะแนนเท่ากับ 0 คะแนน

เนื่องจากคะแนนความสามารถในการเชื่อมโยงทางคณิตศาสตร์ก่อนเรียนของนักเรียนทุกคนเท่ากับ 0 คะแนน จึงไม่สามารถแสดงให้เห็นว่าคะแนนก่อนเรียนเป็นการแจกแจงแบบปกติ ดังนั้นผู้วิจัยจึงเลือกการทดสอบทางสถิติ Wilcoxon signed-rank test ซึ่งเป็นสถิตแบบไม่อิงพารามิเตอร์เพื่อวิเคราะห์ความแตกต่างอย่างมีนัยสำคัญของข้อมูลทั้งสอง

ตารางที่ 1: ผลการวิเคราะห์ข้อมูลโดยใช้การทดสอบทางสถิติโดยวิธี Wilcoxon signed-rank test เพื่อวิเคราะห์ความแตกต่างอย่างมีนัยสำคัญของคะแนนความสามารถในการเชื่อมโยงทางคณิตศาสตร์ก่อนเรียนและหลังเรียน

	ผลต่างระหว่างก่อนเรียนและหลังเรียน
Z	−5.237
Sig. (2 − tailed)	0.000

การทดสอบสมมติฐานแสดงดังต่อไปนี้

$$H_0: M_1 = M_2$$

$$H_1: M_1 < M_2$$

โดยที่ M_1 เป็นค่ามัธยฐานของคะแนนก่อนเรียน

M_2 เป็นค่ามัธยฐานของคะแนนหลังเรียน

จากตารางที่ 1 ค่า Z เท่ากับ -5.237 และค่า Sig. $(2-\text{tailed})$ เท่ากับ 0.000 เนื่องจากค่า Sig. $(1-\text{tailed})$ เท่ากับ $\frac{\text{Sig.}(2-\text{tailed})}{2} = 0.000$ ซึ่งค่านัยสำคัญมีค่าน้อยกว่า .05 ดังนั้นสมมติฐานว่าง (H_0) ถูกปฏิเสธและสมมติฐานทางเลือก (H_1) ถูกยอมรับ จึงสามารถสรุปได้ว่าความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียนเพิ่มขึ้นอย่างมีนัยสำคัญ

4.2 การวิเคราะห์ข้อมูลเชิงคุณภาพ

ในส่วนนี้จะแสดงการเก็บรวบรวมข้อมูลและผลการวิเคราะห์ข้อมูลเชิงคุณภาพของผู้วิจัย โดยแบ่งออกเป็น 2 ส่วน ดังนี้

4.2.1 ผลการวิเคราะห์ข้อมูลเชิงคุณภาพจากการสัมภาษณ์

4.2.2 ผลการวิเคราะห์ข้อมูลเชิงคุณภาพจากบันทึกหลังการสอน

4.2.1 ผลการวิเคราะห์ข้อมูลเชิงคุณภาพจากการสัมภาษณ์

ในส่วนนี้จะแสดงการเก็บข้อมูลและผลการวิเคราะห์ข้อมูลเชิงคุณภาพของผู้วิจัยจากการสัมภาษณ์ ซึ่งคำถามสัมภาษณ์มีทั้งหมด 6 ข้อ และได้เลือกนักเรียน 9 คน จาก 36 คน ที่ให้ความร่วมมือในการสัมภาษณ์ โดยข้อมูลการสัมภาษณ์ของคำถามแต่ละข้อแสดงดังต่อไปนี้

คำถามข้อที่ 1 “นักเรียนคิดว่า ระหว่างการเรียนจากคลิปวิดีโอการสอนผ่านทาง TED-Ed กับการเรียนจากคลิปวิดีโอการสอนแบบปกติมีความแตกต่างกันอย่างไร”

จากการสัมภาษณ์คำถามข้อที่ 1 พบว่า นักเรียน 6 คน ใน 9 คน คิดว่าการเรียนวิดีโอการสอนผ่านเว็บไซต์ TED-Ed นั้นมีความแตกต่างจากการดูวิดีโอการสอนแบบทั่วไป เนื่องจากเนื้อหาที่มีความกระชับ ชัดเจนและตรงกับที่เรียน มีการตรวจสอบความเข้าใจของตนเองผ่านช่อง Think ของ TED-Ed ทำให้รู้ได้ว่ามีตรงไหนบ้างที่นักเรียนยังไม่เข้าใจ เมื่อเปรียบเทียบกับการดูวิดีโอการสอนแบบทั่วไปเนื้อหาจะไม่ค่อยกระชับและไม่ตรงกับที่เรียนด้วย ส่วนนักเรียน 3 คน ใน 9 คน คิดว่าไม่ได้มีความแตกต่างกัน เพราะการเรียนรูปแบบนี้เป็นการเรียนรู้ผ่านคลิปวิดีโอเหมือนกัน

คำถามข้อที่ 2 “นักเรียนคิดว่า อะไรบ้างที่นักเรียนชอบและไม่ชอบกับการเรียนรู้แบบ Flipped Classroom”

จากการสัมภาษณ์คำถามข้อที่ 2 พบว่า สิ่งที่นักเรียนชอบคือ การที่สามารถเรียนเนื้อหาได้ล่วงหน้าผ่านการดูวิดีโอ ทำให้สามารถเตรียมตัวและเข้าใจบทเรียนได้ดีขึ้น การได้มีส่วนร่วมในการทำกิจกรรมในห้องและมีของรางวัล

แจกด้วย ส่วนสิ่งที่ไม่ชอบคือ ไม่ค่อยสะดวกที่จะดูวิดีโอและอยากเรียนในห้องเรียนมากกว่า เพราะเมื่อเกิดข้อสงสัยสามารถถามครูได้ตลอด ถ้าดูวิดีโอแล้วเกิดข้อสงสัยนั้นไม่สามารถถามใครได้เลย

คำถามข้อที่ 3 “นักเรียนคิดว่า การเรียนวิชาคณิตศาสตร์ด้วยการทำโจทย์ปัญหาจากโลกจริงช่วยเสริมสร้างความเข้าใจในวิชาคณิตศาสตร์หรือไม่ เพราะเหตุใด”

จากการสัมภาษณ์คำถามข้อที่ 3 พบว่า การเรียนวิชาคณิตศาสตร์ด้วยการทำโจทย์ปัญหาจากโลกจริงช่วยให้นักเรียนมีความเข้าใจมากยิ่งขึ้น ทำให้นักเรียนสามารถประยุกต์ใช้ความรู้ทางคณิตศาสตร์มาช่วยแก้ไขปัญหาในชีวิตจริง นอกจากนี้ยังช่วยให้เห็นแนวคิดทางคณิตศาสตร์ ซึ่งช่วยให้สามารถวิเคราะห์โจทย์ปัญหาได้มากขึ้น โดยนักเรียนกล่าวว่า “เหมือนเอาเรื่องที่ต้องเจอทุกวันมาโยงเกี่ยวกันทำให้เห็นภาพมากขึ้น ถ้าเจอเหตุการณ์แบบที่เรื่องในห้องก็สามารถนำการแก้ปัญหาในห้องมาใช้ได้เลย ไม่ต้องเสียเวลาในการคิดแก้ปัญหา”

คำถามข้อที่ 4 “นักเรียนคิดว่า การทำ Exit Tickets หลังเรียนจบคาบช่วยให้นักเรียนมีความมั่นใจความสามารถของตัวเองเพิ่มขึ้นหรือไม่ เพราะเหตุใด”

จากการสัมภาษณ์คำถามข้อที่ 4 พบว่า การทำ Exit Tickets หลังเรียนจบคาบช่วยให้นักเรียนส่วนใหญ่ความมั่นใจความสามารถของตัวเองเพิ่มขึ้น เนื่องจาก Exit Tickets สามารถประเมินนักเรียนได้ว่า นักเรียนมีความเข้าใจในบทเรียนแต่ละคาบมากแค่ไหน เพื่อเป็นการกระตุ้นให้นักเรียนได้เกิดการพัฒนาตนเองอย่างต่อเนื่องและยังช่วยให้นักเรียนสามารถจดจำบทเรียนในแต่ละคาบได้ แต่นักเรียน 2 คน ใน 9 คน คิดว่าการทำ Exit Tickets ไม่ค่อยส่งผลต่อความมั่นใจ เนื่องด้วยเหมือนเป็นเพียงตัวช่วยทบทวนบทเรียนและถ้านักเรียนทำไม่ได้จะทำให้รู้สึกไม่มั่นใจตัวเอง อย่างไรก็ตาม ความมั่นใจนั้นจะเพิ่มขึ้นหรือลดลงนั้นขึ้นอยู่กับความสามารถของแต่ละคนว่าตัวเองรู้สึกมั่นใจในแต่ละบทเรียนอย่างไร

คำถามข้อที่ 5 “นักเรียนคิดว่า สาเหตุใดที่ทำให้เพื่อนหรือตัวนักเรียนเองไม่ได้ดูคลิปวิดีโอก่อนเข้าคาบเรียน และส่งผลอย่างไรต่อการทำกิจกรรมในชั้นเรียนบ้าง”

จากการสัมภาษณ์คำถามข้อที่ 5 พบว่า ว่า สาเหตุที่ทำให้นักเรียนไม่ได้ดูวิดีโอก่อนเรียนเกิดจากการที่นักเรียนไม่มีเวลาว่างและนักเรียนรู้สึกเบื่อหน่ายจากการที่ต้องเรียนจากวิดีโอและในชั้นเรียนจึงไม่อยากที่จะเรียนในวิดีโอ อีกทั้งการบ้านแต่ละวิชาเยอะและอยากเรียนกับครูในห้องเรียนมากกว่า นอกจากนี้ไม่มีใครที่คอยกระตุ้นนักเรียนศึกษาวิดีโอทำให้นักเรียนไม่ได้สนใจที่จะศึกษาวิดีโอ ซึ่งการที่ไม่ได้ดูวิดีโอก่อนเรียนค่อนข้างส่งผลในชั้นเรียนเพราะนักเรียนไม่มีความรู้พื้นฐานสำหรับทำกิจกรรมในชั้นเรียน ซึ่งความรู้พื้นฐานนั้นได้จากการดูวิดีโอ ทำให้นักเรียนไม่สามารถตอบคำถามหรือไม่เข้าใจกิจกรรมที่กำลังทำในห้องเรียนนั้นเอง อย่างไรก็ตามผู้วิจัยได้แก้ไขปัญหานี้โดยติดตามและกระตุ้นนักเรียนสำหรับการศึกษาวิดีโอการสอนก่อนเข้าเรียนอย่างสม่ำเสมอ นอกจากนี้ผู้วิจัยได้มีการทบทวนความรู้จากการศึกษาวิดีโอการสอนก่อนเข้าเรียนในขั้นทบทวนความรู้ทุกคาบ เพื่อให้นักเรียนมีความรู้สำหรับการทำกิจกรรมการเรียนรู้ในชั้นเรียนและไม่ให้นักเรียนเสียประโยชน์ต่อการเรียน

คำถามข้อที่ 6 “นักเรียนมีความรู้สึกอย่างไรกับการเรียนคณิตศาสตร์ด้วยรูปแบบการเรียนแบบนี้แล้วอะไรที่ทำให้นักเรียนรู้สึกแบบนั้น และนักเรียนอยากเรียนคณิตศาสตร์ด้วยรูปแบบการเรียนแบบนี้หรือแบบปกติมากกว่ากัน เพราะอะไร”

จากการสัมภาษณ์คำถามข้อที่ 6 พบว่า นักเรียน 5 คน ใน 9 คน รู้สึกดีกับการเรียนแบบนี้ เพราะได้เตรียมตัวสำหรับการเรียนในแต่ละคาบว่าแต่ละคาบจะเรียนเรื่องอะไรบ้างและได้จัดการเวลาเรียนของตัวเองด้วย แต่นักเรียน 4 คน ใน 9 คน รู้สึกไม่คุ้นชินกับการเรียนแบบนี้ เพราะตลอดระยะเวลาที่ผ่านมาไม่เคยเรียนผ่านวิดีโอการสอนมาก่อน ทำให้ไม่ชอบการเรียนในรูปแบบนี้มากและไม่ได้มีเวลาในการปรับตัวด้วย โดยนักเรียนกล่าวว่า “เพราะยังไม่ค่อยชินและปรับตัวไม่ทันเลยไม่ชอบ” นักเรียนส่วนใหญ่ชอบการเรียนแบบนี้เพราะได้ทำความเข้าใจบทเรียนก่อนที่จะเรียนในห้องและสามารถดูย้อนหลังได้ อย่างไรก็ตามนักเรียนบางคนชอบการเรียนแบบปกติ เพราะถ้าไม่เข้าใจตรงไหนสามารถถามครูได้ตลอด แต่การดูวิดีโอเมื่อเกิดความไม่เข้าใจบทเรียนก็ไม่รู้ว่าจะต้องถามใครได้เลย โดยนักเรียนกล่าวว่า “ถึงแม้การดูวิดีโอการสอนจะสามารถถามครูได้ผ่านทางไลน์ แต่การที่ได้พูดคุยต่อหน้ากันจะทำให้เข้าใจมากกว่า”

4.2.2 ผลการวิเคราะห์ข้อมูลเชิงคุณภาพจากบันทึกหลังการสอน

ในส่วนนี้จะแสดงการเก็บข้อมูลและผลการวิเคราะห์ข้อมูลเชิงคุณภาพของผู้วิจัยจากบันทึกหลังการสอน โดยจะสังเกตพฤติกรรมการเรียนของนักเรียนในแต่ละคาบ ซึ่งในบทความนี้จะแสดงบันทึกหลังการสอนคาบที่ 1 และ 5 เพื่อดูพฤติกรรมและการพัฒนาความสามารถในการเชื่อมโยงของนักเรียนดังตารางที่ 2 และ 3

ตารางที่ 2: บันทึกหลังการสอนคาบที่ 1 เรื่อง ความหมาย สัญลักษณ์และชนิดของเมทริกซ์

ขั้นตอนกิจกรรม	**พฤติกรรมของนักเรียน**
ขั้นทบทวนความรู้	นักเรียนทั้งหมดร่วมกันตอบคำถามในชั้นเรียนเกี่ยวกับ ความหมาย สัญลักษณ์และชนิดของเมทริกซ์ และร่วมตอบคำถามในเว็บแอปพลิเคชัน Quizizz ซึ่งนักเรียน 20 คน จาก 36 คน สามารถตอบคำถามทั้งหมดได้ เช่น เมทริกซ์ขนาด $m \times n$ มีจำนวนสมาชิกเท่ากับเท่าใด แต่นักเรียนจำนวน 16 คน จาก 36 คน ไม่สามารถตอบคำถามเกี่ยวกับขนาดและการจำแนกชนิดเมทริกซ์ เช่น เมทริกซ์ $A = \begin{bmatrix} 3 & 1 & 5 \\ 1 & 0 & 0 \\ -1 & 2 & 4 \end{bmatrix}$ มีขนาดเท่าไร เป็นต้น ครูได้ทำการอธิบายเพิ่มเติมเกี่ยวกับขนาดและชนิดของเมทริกซ์จากวิดีโอเมทริกซ์ Ep. 1 และ 2 ในเว็บไซต์ TED-Ed เพื่อให้นักเรียนเข้าใจมากยิ่งขึ้น หลังจากนั้นครูได้นำเสนอสถานการณ์ตัวอย่างเกี่ยวกับการจัดการข้อมูลในตลาดเชื่อมโยงกับความรู้ทางคณิตศาสตร์เรื่องเมทริกซ์
ขั้นแก้ปัญหาโลกจริง	นักเรียนจำนวน 10 คน จาก 36 คน ได้ร่วมแสดงความคิดเห็นและตอบคำถามเกี่ยวกับการแก้ปัญหาจากสถานการณ์ตัวอย่างว่า “ในการจัดการข้อมูลควรนำข้อมูลจัดให้อยู่ในรูปตารางก่อน” ต่อมาครูได้นำความรู้ที่ได้จากวิดีโอ TED-Ed มาเชื่อมโยงกับสถานการณ์ตัวอย่างเพื่อใช้

	แก้ปัญหา หลังจากนักเรียนได้แนวคิดจากครูแล้วนักเรียนเริ่มทำใบงานที่ 1 ด้วยตนเอง พบว่านักเรียนทั้งหมดสามารถเขียนข้อมูลในตารางให้อยู่ในรูปเมทริกซ์ได้ แต่นักเรียนส่วนใหญ่ยังไม่เข้าใจความสำคัญในการกำหนดชื่อแต่ละเมทริกซ์ เนื่องจากการกำหนดชื่อของเมทริกซ์ช่วยให้เข้าใจว่าแต่ละเมทริกซ์ต้องการสื่อถึงอะไรบ้าง หลังจากนั้นให้นักเรียนออกมาแสดงวิธีทำหน้าชั้นเรียนมีเพียงนักเรียนจำนวน 5 คน จาก 36 คน ที่ออกมาแสดงวิธีทำหน้าชั้นเรียน เนื่องจากนักเรียนกลัวที่จะแสดงวิธีทำผิด
ขั้นสรุป	นักเรียนทั้งหมดร่วมกันสรุปความรู้ที่ได้จากการทำกิจกรรมและแต่ละคนได้ทำใบงาน Exit Ticket เพื่อตรวจสอบความเข้าใจของตนเองหลังจากที่ได้ทำกิจกรรม พบว่านักเรียนจำนวน 23 คน ที่ตอบคำถามถูกต้อง หลังจากนักเรียนทำใบงาน Exit Ticket เสร็จแล้วนักเรียนแต่ละคนได้ร่วมกันแบ่งปันความรู้และหาข้อผิดพลาดของตนเอง

ตารางที่ 3: บันทึกหลังการสอนคาบที่ 5 เรื่อง การหาดีเทอร์มิแนนต์ขนาด 2×2

ขั้นตอนกิจกรรม	**พฤติกรรมของนักเรียน**
ขั้นทบทวนความรู้	นักเรียนทุกคนยังไม่เข้าใจเกี่ยวกับการหาดีเทอร์มิแนนต์ขนาด 2×2 หลังจากที่ได้ดูวิดีโอการสอนจาก TED-Ed เรื่องการหาดีเทอร์มิแนนต์ขนาด 2×2 ดังนั้นครูจึงต้องอธิบายและสรุปรวบยอดแนวคิดอีกครั้งและให้นักเรียนทุกคนได้ร่วมกันตอบคำถามในชั้นเรียน ตัวอย่างเช่น กำหนดให้เมทริกซ์ $A = \begin{bmatrix} 1 & 2 \\ 3 & -2 \end{bmatrix}$ นักเรียนคิดว่า ดีเทอร์มิแนนต์ของเมทริกซ์ A เท่ากับเท่าไร ซึ่งมีนักเรียนจำนวน 10 คน จาก 36 คน สามารถตอบคำถามถูกต้อง หลังจากนั้นครูได้นำเสนอสถานการณ์ตัวอย่างเกี่ยวกับการหาพื้นที่สี่เหลี่ยมด้านขนานโดยเชื่อมโยงกับความรู้ทางคณิตศาสตร์เรื่องเมทริกซ์
ขั้นแก้ปัญหาโลกจริง	นักเรียนจำนวน 20 คน จาก 36 คน ได้ร่วมแสดงความคิดเห็นและตอบคำถามเกี่ยวกับการแก้ปัญหาจากสถานการณ์ตัวอย่างว่า “การหาพื้นที่สี่เหลี่ยมด้านขนานใช้สูตร ความยาวของฐานคูณกับความสูง” ต่อมาครูได้นำความรู้ที่ได้จากวิดีโอ TED-Ed มาเชื่อมโยงกับสถานการณ์ตัวอย่างและแสดงที่มาของสูตรเพื่อให้นักเรียนได้เข้าใจแนวคิดและนำไปประยุกต์ใช้กับโจทย์ปัญหาใบงานที่ 5 หลังจากนักเรียนได้แนวคิดจากครูแล้วนักเรียนเริ่มทำใบงานที่ 5 ด้วยตนเอง พบว่านักเรียนจำนวน 20 คน จาก 36 คน สามารถนำแนวคิดที่ได้จากครูไปประยุกต์ใช้ในใบงานที่ 5 แต่ยังมีนักเรียนบางส่วนที่สับสนกับวิธีการใช้ดีเทอร์มิแนนต์ขนาด 2×2 ในการหาพื้นที่สี่เหลี่ยมด้านขนาน หลังจากนั้นให้นักเรียนออกมาแสดงวิธีทำหน้าชั้นเรียนมีนักเรียน

	จำนวน 12 คน จาก 36 คน ที่ออกมาแสดงวิธีทำหน้าชั้นเรียน เนื่องจากมีขนมเป็นของรางวัลสำหรับนักเรียนที่มีส่วนร่วมในการตอบคำถาม
ขั้นสรุป	นักเรียนทั้งหมดร่วมกันสรุปความรู้ที่ได้จากการทำกิจกรรมและแต่ละคนได้ทำใบงาน Exit Ticket เพื่อตรวจสอบความเข้าใจของตนเองหลังจากที่ได้ทำกิจกรรม พบว่านักเรียนจำนวน 21 คน ที่ตอบคำถามถูกต้อง หลังจากนักเรียนทำใบงาน Exit Ticket เสร็จแล้วนักเรียนแต่ละคนได้ร่วมกันแบ่งปันความรู้และหาข้อผิดพลาดของตนเอง

จากตารางที่ 2 จะเห็นได้ว่า นักเรียนส่วนใหญ่ยังไม่ได้ให้ความสำคัญในการตั้งชื่อเมทริกซ์ แสดงให้เห็นว่านักเรียนยังไม่สามารถตีความข้อมูลได้ อีกทั้งนักเรียนไม่ค่อยให้ความร่วมมือกับการตอบคำถามในชั้นเรียน มีนักเรียนเพียงบางคนเท่านั้นที่มีส่วนร่วมในการตอบคำถาม ดังนั้นผู้วิจัยจึงได้ปรับเปลี่ยนกิจกรรมในชั้นเรียนเพื่อให้นักเรียนมีส่วนร่วมในการตอบคำถามมากยิ่งขึ้น เช่น การติดตามและกระตุ้นนักเรียนสำหรับการศึกษาวิดีโอการสอนก่อนเข้าเรียนอย่างสม่ำเสมอ เป็นต้น จากตารางที่ 3 จะเห็นได้ว่าหลังจากผู้วิจัยได้ปรับเปลี่ยนกิจกรรมในชั้นเรียนทำให้นักเรียนให้ความร่วมมือกับการตอบคำถามในชั้นเรียนมากขึ้น

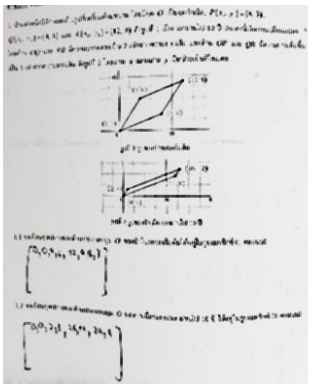
5 สรุปผลและอภิปรายผล

5.1 อภิปรายผล

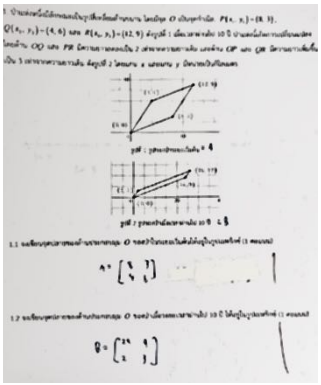
จากการเก็บรวบรวมข้อมูลความสามารถในการเชื่อมโยงทางคณิตศาสตร์และการทดสอบสมมติฐานทางสถิติโดยวิธี Wilcoxon signed-rank test พบว่านักเรียนที่เข้าร่วมการจัดการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed มีความสามารถในการเชื่อมโยงทางคณิตศาสตร์เพิ่มขึ้นอย่างมีนัยสำคัญ จากรูปที่ 2 พบว่าคะแนนความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียนทุกคนเท่ากับ 0 คะแนน เนื่องจากแบบทดสอบนี้เป็นข้อสอบแบบอัตนัยที่ต้องมีการแสดงวิธีทำอย่างละเอียดและนักเรียนยังความรู้ที่ไม่มากพอที่จะประยุกต์ใช้กับโจทย์ปัญหาโลกจริงทำให้นักเรียนหลายคนเลือกที่จะไม่แสดงวิธีทำ อย่างไรก็ตามมีนักเรียนบางคนที่ได้แสดงถึงความพยายามในการแสดงวิธีทำดังรูปที่ 4 (ก) และ 5 (ก) แต่พบว่านักเรียนไม่สามารถแสดงวิธีทำได้อย่างถูกต้อง เกิดจากที่นักเรียนยังไม่สามารถตีความโจทย์ปัญหาโลกจริงได้และยังไม่สามารถประยุกต์ใช้ความรู้ทางคณิตศาสตร์ในการแก้โจทย์ปัญหาโลกจริง ส่งผลให้นักเรียนไม่สามารถตอบคำถามได้อย่างถูกต้องจึงได้ 0 คะแนน หลังจากนักเรียนได้เข้าร่วมการจัดการเรียนรู้แบบนี้แล้ว นักเรียนสามารถเชื่อมโยงแนวคิดทางคณิตศาสตร์เรื่องเมทริกซ์ในการแก้ปัญหาในโลกจริงดังรูปที่ 4 (ข) และ 5 (ข) ทำให้นักเรียนสามารถประยุกต์ความรู้ทางคณิตศาสตร์ในการแก้ปัญหาในโลกจริงและเข้าใจแนวคิดทางคณิตศาสตร์ดีขึ้น ดังนั้นการจัดการเรียนรู้แบบห้องเรียนกลับด้านช่วยให้

นักเรียนแก้ปัญหาทางคณิตศาสตร์ได้ดีขึ้นและช่วยพัฒนาความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียนดีขึ้น [31, 32]

การเรียนรู้แบบห้องเรียนกลับด้านเป็นการเรียนรู้ที่แบ่งเป็น 2 ส่วน ได้แก่ การเรียนรู้นอกชั้นเรียนและการเรียนรู้ในชั้นเรียน สำหรับการเรียนรู้นอกชั้นเรียนเป็นการให้นักเรียนศึกษาความรู้ด้วยตนเองด้วยการดูวิดีโอการสอนผ่านเว็บไซต์ TED-Ed ซึ่งการดูวิดีโอการสอนผ่านเว็บไซต์ TED-Ed ประกอบไปด้วย 2 ส่วน ได้แก่ Watch และ Think ส่วนที่ 1 Watch จะให้นักเรียนดูวิดีโอการสอนผ่านเว็บไซต์ TED-Ed และส่วนที่ 2 Think จะให้นักเรียนตอบคำถามหลังจากดูวิดีโอการสอนผ่านเว็บไซต์ TED-Ed เพื่อตรวจสอบความเข้าใจของนักเรียน จากบทสัมภาษณ์ข้อที่ 1 พบว่าเมื่อนักเรียนศึกษาวิดีโอการสอนผ่านเว็บไซต์ TED-Ed นักเรียนสามารถตรวจสอบความเข้าใจได้จาก Think ซึ่งเป็นส่วนที่ให้นักเรียนตอบคำถามหลังจากดูวิดีโอเพื่อตรวจสอบความเข้าใจของนักเรียนในแต่ละบทเรียนว่าเข้าใจมากน้อยเพียงใดหลังจากดูวิดีโอการสอนแล้ว ซึ่งช่วยให้นักเรียนสามารถประเมินตนเองว่ามีข้อผิดพลาดหรือเนื้อหาใดที่ยังไม่เข้าใจหรือไม่ ซึ่งสอดคล้องกับ Douglas [33] ที่กล่าวว่าการเรียนรู้วิดีโอการสอนผ่าน TED-Ed ช่วยให้นักเรียนเรียนรู้ในเชิงลึกและเข้าใจเนื้อหาได้ลึกซึ้งยิ่งขึ้นผ่านกิจกรรมวิดีโอออนไลน์ ทำให้นักเรียนสามารถนำความรู้ที่ได้รับจากวิดีโอไปใช้ในการทำกิจกรรมในชั้นเรียนได้ จากบทสัมภาษณ์ข้อที่ 2 และ 6 พบว่านักเรียนที่ดูวิดีโอการสอนก่อนเข้าชั้นเรียนช่วยให้เตรียมความพร้อมสำหรับกิจกรรมในชั้นเรียนและเข้าใจบทเรียนได้ดีขึ้น นอกจากนี้นักเรียนยังสามารถจัดการเวลาเรียนเพื่อดูวิดีโอการสอนได้อย่างเหมาะสมและสามารถดูซ้ำได้เรื่อย ๆ ซึ่งสอดคล้องกับ Munir และคณะ [34] ที่กล่าวว่านักเรียนที่ดูวิดีโอการสอนก่อนเข้าชั้นเรียนช่วยพัฒนาทักษะที่สำคัญ เช่น การคิดวิเคราะห์และการสื่อสาร เป็นต้น อีกทั้งช่วยเพิ่มประสิทธิภาพในการเรียนรู้ของนักเรียนด้วย

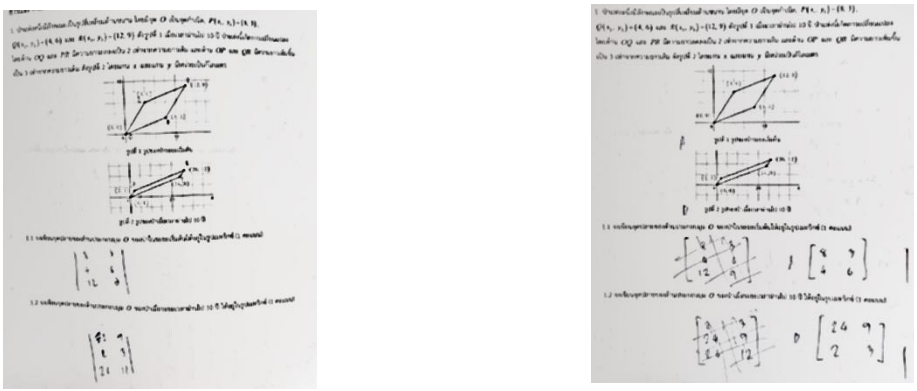


(ก)



(ข)

รูปที่ 4: การพัฒนาความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียน ก ที่เข้าร่วมการจัดการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed (ก) แบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์ก่อนเรียน (ข) แบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์หลังเรียน

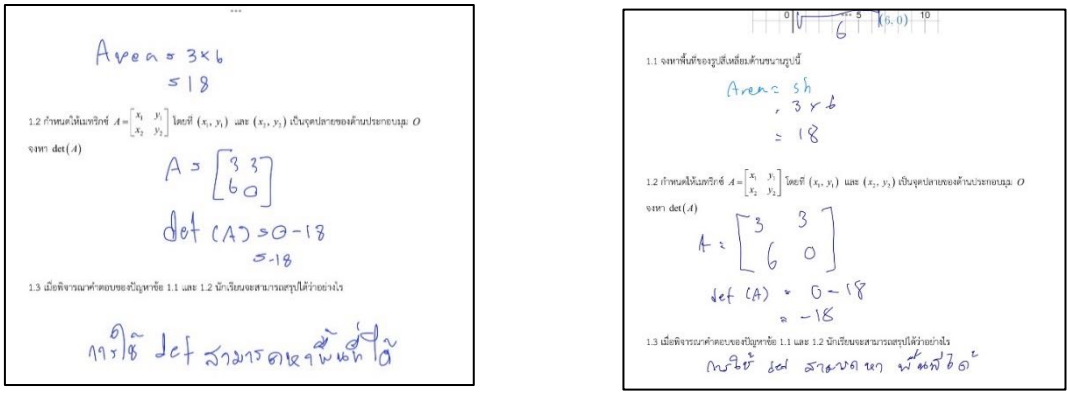


(ก) (ข)

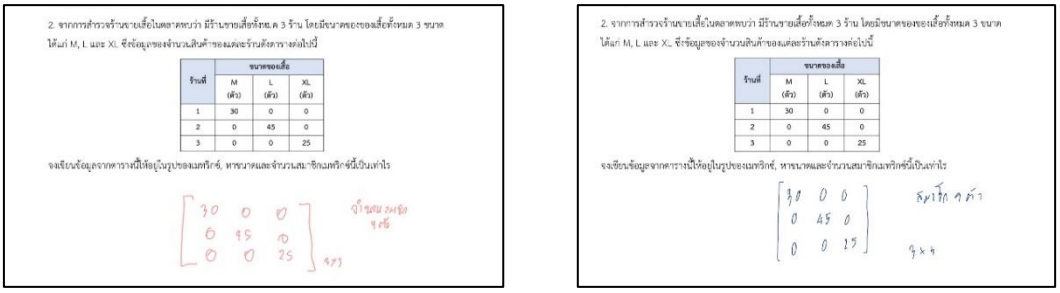
รูปที่ 5: การพัฒนาความสามารถในการเชื่อมโยงทางคณิตศาสตร์ของนักเรียน ข ที่เข้าร่วมการจัดการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed (ก) แบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์ก่อนเรียน (ข) แบบทดสอบความสามารถในการเชื่อมโยงทางคณิตศาสตร์หลังเรียน

สำหรับการเรียนรู้ในชั้นเรียนแบ่งออกเป็น 3 ขั้นตอน ได้แก่ ขั้นทบทวนความรู้ ขั้นแก้ปัญหาโลกจริง และขั้นสรุป ในขั้นแก้ปัญหาโลกจริงจะให้นักเรียนแลกเปลี่ยนความคิดเห็น อภิปราย และนำเสนอแนวทางการแก้ปัญหา ครูจะให้คำแนะนำในการแก้ปัญหาโดยใช้ความรู้ที่ได้รับจากการดูวิดีโอการสอนออนไลน์จากเว็บไซต์ TED-Ed เพื่อเชื่อมโยงกับการแก้ปัญหาในชีวิตจริง โดยนักเรียนนำเสนอแนวทางการแก้ปัญหาตามความรู้ของตน จากนั้นครูจะนำเสนอแนวคิดในการแก้ปัญหาที่เชื่อมโยงกับเมทริกซ์จากวิดีโอการสอน ต่อมาให้นักเรียนพยายามแก้ปัญหาโดยใช้แนวคิดที่ครูนำเสนอและเปรียบเทียบกับแนวคิดของนักเรียน ดังนั้นนักเรียนจึงสามารถเชื่อมโยงความรู้ที่ได้รับจากวิดีโอการสอนกับการแก้ปัญหาในชีวิตจริงได้ จากตารางที่ 2 และรูปที่ 6 พบว่านักเรียนไม่ได้ให้ความสำคัญในการกำหนดชื่อเมทริกซ์แต่ละเมทริกซ์ ซึ่งเป็นสิ่งที่สำคัญในการตีความข้อมูลจากโจทย์ปัญหาในโลกจริง แสดงให้เห็นว่านักเรียนยังไม่สามารถเชื่อมโยงความรู้ทางคณิตศาสตร์กับชีวิตจริงได้ โดยมุ่งเน้นเพียงกระบวนการในการทำโจทย์ปัญหาเท่านั้น แต่ขาดความเข้าใจในแนวคิดทางคณิตศาสตร์ที่ใช้ในการทำโจทย์ปัญหาได้เลย หลังจากนั้นผู้วิจัยได้ปรับรูปแบบกิจกรรมเพื่อปรับเปลี่ยนพฤติกรรมของนักเรียน จากตารางที่ 3 และรูปที่ 7 พบว่านักเรียนเริ่มให้ความสำคัญในการกำหนดชื่อเมทริกซ์แต่ละเมทริกซ์และได้ระบุแนวคิดทางคณิตศาสตร์ที่ใช้ในการทำโจทย์ปัญหา ทำให้เห็นได้ว่านักเรียนเริ่มสามารถเชื่อมโยงความรู้ทางคณิตศาสตร์ได้ จากบทสัมภาษณ์ข้อที่ 3 พบว่าการเรียนรู้คณิตศาสตร์ด้วยการแก้ปัญหาโลกจริงช่วยเสริมสร้างความเข้าใจทางคณิตศาสตร์ของนักเรียน ช่วยให้นักเรียนเข้าใจแนวคิดทางคณิตศาสตร์มากขึ้นและเสริมสร้างการคิดอย่างเป็นระบบ ส่งผลให้นักเรียนสามารถนำความรู้ทางคณิตศาสตร์ไปประยุกต์ใช้ในการวิเคราะห์และแก้ปัญหาในโลกแห่งความเป็นจริงที่สอดคล้องกับกิจกรรมการเรียนรู้แต่ละกิจกรรมได้ การเรียนรู้แบบแก้ปัญหาโลกจริงช่วยให้นักเรียนสร้างความเข้าใจที่ลึกซึ้งยิ่งขึ้นเกี่ยวกับคณิตศาสตร์ที่ใช้ในชีวิตประจำวัน นอกจากนี้ยังช่วยพัฒนาความสามารถเชื่อมโยงทางคณิตศาสตร์ของนักเรียนไปสู่อีกระดับที่สูงขึ้นอีก

ด้วย [35, 36] จากบทสัมภาษณ์ข้อที่ 4 พบว่า เมื่อนักเรียนทำใบตั๋วออกในแต่ละคาบ ทำให้นักเรียนสามารถตรวจสอบความเข้าใจของตนเองเกี่ยวกับข้อผิดพลาดต่างๆ เพื่อพัฒนาตนเองให้เข้าใจมากขึ้น ใบตั๋วออกเป็นเครื่องมือประเมินผลที่มีประสิทธิภาพในการเรียนรู้ อีกทั้งยังช่วยให้นักเรียนปรับปรุงและพัฒนาความเข้าใจของตนเอง [21, 37, 38]



รูปที่ 6: ตัวอย่างใบงานของนักเรียนในคาบเรียนที่ 1



รูปที่ 7: ตัวอย่างใบงานของนักเรียนในคาบเรียนที่ 5

อย่างไรก็ตามนักเรียนส่วนใหญ่ยังคงไม่ให้ความร่วมมือในชั้นเรียน เนื่องจากนักเรียนยังไม่มีความรู้ขั้นพื้นฐานในการทำกิจกรรมในชั้นเรียนที่ได้จากการศึกษาวิดีโอการสอนจากเว็บไซต์ TED-Ed โดยสาเหตุที่นักเรียนไม่ศึกษาวิดีโอการสอนเกิดจากนักเรียนไม่มีเวลาว่าง รู้สึกเบื่อหน่ายจากการเรียนทั้งในรูปแบบวิดีโอและในชั้นเรียน อีกทั้งนักเรียนมีความต้องการที่จะเรียนกับครูในชั้นเรียนมากกว่าการศึกษาวิดีโอการสอนนอกห้องเรียน นอกจากนี้การศึกษาวิดีโอการสอนนอกห้องเรียนไม่มีใครที่กระตุ้นนักเรียนในการศึกษาวิดีโอการสอนนอกห้องเรียน ซึ่งสอดคล้องกับบทสัมภาษณ์ข้อที่ 5 ว่าสาเหตุที่นักเรียนไม่ศึกษาวิดีโอการสอนคือ นักเรียนไม่มีเวลาว่าง รู้สึกเบื่อหน่ายจากการเรียนทั้งวิดีโอและในชั้นเรียน ภาระงานจากวิชาต่าง ๆ ค่อนข้างเยอะ นอกจากนี้นักเรียนยังกล่าวว่า “อยากใช้เวลาส่วนตัวที่เหลือในการพักผ่อนจากการเรียนมาทั้งวัน ไม่ใช่ต้องเอาเวลานั้นมาเรียนเพิ่มเติม” และจากบทสัมภาษณ์ข้อที่ 6 นักเรียนกล่าวว่า “เพราะยังไม่ค่อยชินและปรับตัวไม่ทัน จึงอยากเรียนกับครูมากกว่าดูวิดีโอการสอน” นั้นจึงเป็นเหตุผลที่นักเรียนส่วน

ใหญ่ไม่ศึกษาวิดีโอการสอนก่อนเข้าชั้นเรียนและเลือกที่จะรอครูสอนในชั้นเรียน แต่กิจกรรมในชั้นเรียนนั้นจะต้องใช้ความรู้พื้นฐานที่ได้จากการเรียนรู้นอกชั้นเรียนทำให้นักเรียนไม่สามารถตอบคำถามในชั้นเรียนได้เลย อีกทั้งนักเรียนไม่สามารถทำกิจกรรมในชั้นเรียนด้วย ส่งผลต่อประสิทธิภาพในการเรียนรู้ของนักเรียนทำให้ความสามารถในการเชื่อมโยงทางคณิตศาสตร์พัฒนาได้เพียงระดับหนึ่งเท่านั้น การจัดกิจกรรมการเรียนรู้แบบห้องเรียนกลับด้านจะประสบความสำเร็จนั้นขึ้นอยู่กับความร่วมมือของทั้งครูและนักเรียน ครูต้องออกแบบกิจกรรมในห้องเรียนที่ส่งเสริมให้นักเรียนเตรียมตัวล่วงหน้าและมีส่วนร่วมในชั้นเรียนอย่างมีประสิทธิภาพ นักเรียนต้องรับผิดชอบในการศึกษาวิดีโอการสอนก่อนเข้าชั้นเรียน จึงจะช่วยพัฒนาความสามารถในการเชื่อมโยงทางคณิตศาสตร์ให้มีประสิทธิภาพมากยิ่งขึ้น [16, 19]

5.2 สรุปผล

จากผลการวิจัยสามารถสรุปได้ว่า นักเรียนที่เข้าร่วมการจัดการเรียนรู้แบบห้องเรียนกลับด้านร่วมกับเทคนิคตั๋วออกผ่านเว็บไซต์ TED-Ed ความสามารถในการเชื่อมโยงทางคณิตศาสตร์เพิ่มขึ้น อย่างไรก็ตามทั้งครูและนักเรียนควรมีส่วนร่วมในการจัดการชั้นเรียนซึ่งกันและกันไม่สามารถขึ้นอยู่กับฝ่ายใดฝ่ายหนึ่งเท่านั้น โดยที่ครูจะต้องออกแบบกิจกรรมในห้องเรียนที่ส่งเสริมให้นักเรียนเตรียมตัวล่วงหน้าและมีส่วนร่วมในชั้นเรียนอย่างมีประสิทธิภาพ อีกทั้งครูต้องกระตุ้นนักเรียนให้เกิดการเรียนรู้นอกชั้นเรียนมากขึ้น ส่วนนักเรียนต้องรับผิดชอบในการศึกษาวิดีโอการสอนก่อนเข้าชั้นเรียนและต้องมีวินัยในตนเองเพื่อที่จะสามารถพัฒนาการเรียนรู้ทางคณิตศาสตร์และความสามารถในการเชื่อมโยงทางคณิตศาสตร์ให้มีประสิทธิภาพมากยิ่งขึ้น

5.3 ข้อเสนอแนะ

1. บริบทของการศึกษามีความแคบเกินไปเนื่องจากชั้นเรียนที่ผู้วัยได้ทำการวิจัยนั้นเป็นโรงเรียนชายล้วนควรคำนึงถึงความหลากหลาย เช่น เพศ อายุ และชาติพันธุ์ให้มากขึ้น

2. การเรียนรู้แบบห้องเรียนกลับด้านควรมีใครสักคนที่คอยกระตุ้นนักเรียนให้ศึกษาวิดีโอก่อนเข้าเรียนอย่างต่อเนื่อง เช่น พ่อแม่นักเรียน ครู เป็นต้น หรือใช้แพลตฟอร์มสำหรับการศึกษาวิดีโอที่สามารถคอยกระตุ้นนักเรียนได้ตลอดเวลาเมื่อนักเรียนไม่ได้ดูวิดีโอก่อนเรียน อีกทั้งควรให้นักเรียนได้ปรับความคุ้นชินกับการเรียนในลักษณะนี้ก่อนเพื่อให้ผลลัพธ์ในการวิจัยมีประสิทธิภาพมากขึ้น

3. ควรลดจำนวนวิดีโอในการศึกษาแต่ละคาบ เพื่อลดภาระที่นักเรียนจะได้เรียนรู้ด้วยตัวเอง

กิตติกรรมประกาศ ขอขอบคุณผู้เชี่ยวชาญทั้ง 3 ท่าน ได้แก่ รศ. ดร.ภวธน เขมะวิชานุรัตน์ ว่าที่ร้อยตรีธงชัย เนตรสว่าง และว่าที่ร้อยตรีอนุเทพ เทพปัน ที่สละเวลาในการตรวจสอบเครื่องมือวิจัยและให้คำแนะนำในการปรับปรุงเครื่องมือให้มีความถูกต้องมากยิ่งขึ้นและขอขอบคุณอย่างสุดซึ้งที่โรงเรียนเทพศิรินทร์และนักเรียนได้ให้ความอนุเคราะห์ในการทำวิจัยในครั้งนี้ให้สำเร็จลุล่วงไปด้วยดี

เอกสารอ้างอิง

[1] Sarkam, I. Sujadi, S. Subnti, Mathematical connections ability in solving trigonometry problems based on logical-mathematical intelligence level, Journal of Physics: Conference Series, (1188) (2019).

[2] M. Oktaviani, A. Dudung, Mathematical Connection Ability: An Analysis Based on Test Forms, International Journal of Advanced Science and Technology, 6 (29) (2020) 4694-4701.

[3] A. K. Kenedi, Y. Helsa, Y. Ariani, M. Zainil, S. Hendri, Mathematical Connection of Elementary School Students to Solve Mathematical Problems, Journal on Mathematics Education, 1 (10) (2019) 69-80.

[4] H. Hendriana, U. R. Slamet, U. Sumarmo, Mathematical connection ability and self-confidence (An experiment on junior high school students through contextual teaching and learning with mathematical manipulative), International Journal of Education, 1 (8) (2014) 1-11.

[5] S. Rohmah, T. A. Kusmayadi, L. Fitriana, Mathematical connection ability of junior high school students viewed from mathematical resilience, Journal of Physics: Conference Series, (1538) (2020).

[6] U. H. Ningrum, Mulyono, Isnaarto, Mathematical Connection Ability Based on Self-Efficacy in IDEAL Problem Solving Model Assisted by ICT, Unnes Journal of Mathematics Education Research, 2 (9) (2020) 139-146.

[7] A. G. E. Putri, D. U. Wutsqa, Students' Mathematical Connection Ability in Solving Real-world Problem, IOP Conference Series: Journal of Physics: Conference Series, (1320) (2019).

[8] H. G. Schmindt, S. L. Wagener, G. A.C.M. Smeets, L. M. Keemink, H. T. Molen, On the Use and Misuse of Lectures in Higher Education, Health Professions Education, 1 (1) (2015) 12-18.

[9] C. S. Gosavi, S. Arora, Active Learning Strategies for Engaging Students in Higher Education, Journal of Engineering Education Transformations, Special issue (36) (2022).

[10] R. Kay, T. MacDonald, M. DiGiuseppe, A comparison of lecture-based, active, and flipped classroom teaching approaches in higher education, Journal of Computing in Higher Education, (31) (2019) 449-471.

[11] A. J. Kugler, H. P. Gogineni, L. S. Garavalia, Learning Outcome and Student Preferences with Flipped vs Lecture/Case Teaching Model in a Block Curriculum", American Journal of Pharmaceutical Education, 8 (83) (2019).

[12] M. S. Hussain, S. A. Khan, Self-efficacy of teachers: A review of the literature", Jamshedpur Research Review, 50 (1) (2022) 110-116.

[13] F. Ozdamil, G. Asiksoy, Flipped classroom approach, World Journal on Educational Technology: Current Issues, 2 (8) (2016) 98-105.

[14] C.-L. Lai, G.-J. Hwang, A Self-regulated flipped classroom approach to improving students' learning performance in a mathematics course, International Journal Computers and Education, (100) (2016) 126-140.

[15] A. A. Tawfik, C. Lilly, Using a Flipped Classroom Approach to Support Problem-Based Learning", Technology, Knowledge and Learning, (20) (2015) 299-315.

[16] X. Wei, I.-L. Cheng, N.-S. Chen, X. Yang, Y. Lin, Y. Dong, X. Zhai, Kinshuk, Effect of the flipped classroom on the mathematics performance of middle school students, Education Technology Research and Development, (68) (2020) 1461-1484.

[17] R. K. Kadambaevna, S. Bekchanqizi, R. Muniboeva, D. Z. NavFaqizi, A. S. Amatchonova, F. G. Erkinovna, Implementing Blended Learning Technology "Flipped Class" on the Basis of the Platform "Ted-Ed", Elementary Education Online, 5 (20) (2021) 443-443.

[18] K. Umam, T. Nusantara., I. N. Parta, E. Hidayanto, An Application of Classroom in mathematics Teacher Education Programme, International Journal of Interactive Mobile Technologies, 3 (13) (2019).

[19] T. Låg, R. G. Sæle, Does the Flipped Classroom Improve Student Learning and Satisfaction? A Systematic Review and Meta-Analysis, American Educational Research Association Open, 3 (5) (2019) 1-17.

[20] J. P. Prieto, A. Escobar, Exit Tickets' Effect on Engagement in College Classrooms, Proceedings of EDULEARN16 Conference, (2016) 5915-5918.

[21] A. Danley, A. McCoy, R. Weed, Exit Ticket Open the Door to University Learning, A Journal of Scholarly Teaching, (11) (2016) 48-58.

[22] Z. Zainuddin, C. J. Perera, Supporting students' self-directed learning in the flipped classroom through the LMS TES BlendSpace, On the Horizon The International Journal of Learning Futures, 4 (26) (2018) 281-290.

[23] S. Han, Flipped classroom: Challenges and benefits of using social media in English language teaching and learning, Frontiers in Psychology, (13) (2022).

[24] L. Edwards, What Is TED-Ed and How Does It Work For Education? <https://www.techlearning.com/how-to/what-is-ted-ed-and-how-does-it-work-for-education?fbclid=IwAR10joqpIvsqZmUoRSGNTA6eQH4BjQvmanOuA1rP0hpwyOmt2fsSP1ijB4>, 2021 (accessed 18.10.2021).

[25] R. Lehner, Creating Mathematical Lessons Using TED-Ed, The Mathematics Teacher, 6 (110) (2017) 466-468.

[26] N. Holmes, Engaging with assessment: Increasing student engagement through continuous assessment, Active Learning in Higher Education, 1 (19) (2018) 23-34.

[27] T. Saleem, F. Kausar, M. Mahmood, Title: One-Minute Paper, Muddiest Pont, and Exit Ticket Formative Evaluation Methods Used by Teachers in Daily Classroom Teaching", Journal of Critical Reviews, 2 (8) (2021) 792-803.

[28] J. Paz-Albo, J. M. R. Ruiz, R. Bernárdez-Vilaboa, P. Huerta-Zavala, A. Hervás-Escobar, The Impact of Socrative Exit Tickets on Initial Teacher Training", College Teaching, 4 (70) (2022) 413-421.

[29] Baiduri, O. R. U. Putri, I. Alfani, Mathematical Connection Process of Students with High Mathematics Ability in Solving PISA Problems", European Journal of Education Research, 4 (9) (2020) 1527-1537.

[30] D. Rahmawati, Budiyono, D. R. S. Saputro, Analysis of student's mathematical connection ability in linear equation system with two variables", Journal of Physis: Conference Series, 1 (1211) (2019).

[31] M. Jamaludidin, M. Mutaji, B. S. Bachri, A. P. Sutarto, The Role of Gender and Flex Blended Learning and Mathematics Ability, International Journal of Research in Education and Science, 5 (13) (2023) 873-881.

[32] S. Sengsamak, R. Chonchaiya, C. Toopsuwan, A Study on Flipped Classroom based on Donald and Maki's Approach for Developing Mathematical Connection in Exponential and Logarithmic Functions to Real-Life Situations for Grade 10 Students", International Conference The Future of Education, (2021).

[33] R. Douglas, TED-Ed web-based blended learning to support teaching: An action research project, International Conference on Higher Education Advances, (2) (2016).

[34] M.T. Munir, S. Baroutian, B. Young, S. Carter, Flipped classroom with cooperative learning as a cornerstone, Education for Chemical Engineers, (23) (2018) 25-33.

[35] J. Garcia-Garcia, C. Dolores, Exploring pre-university students' mathematical connections when solving Calculus application problems, International Journal of Mathematical Education in Science and Technology, (6) (52) (2021)

[36] M. Hasbi, A. Lukito, R. Sulaiman, M. Muzaini, Improving the Mathematical Connection Ability of Middle-School Students through Realistic Mathematics Approach, Journal of Mathematical Pedagogy, 1 (1) (2019) 37-46.

[37] V. Paliwal, USING EXIT TICKET TO LEARN FROM MISTAKES AND DEVELOPING GROWTH MINDSET IN AN ALGEBRA CLASS, Proceeding of the forty-fifth annual meeting of North American Chapter of the International Group for the Psychology of Mathematics Education, (2) (2023).

[38] Yogyantoro, Developing Entrance ticket and Exit Card as Formative Assessment for Promoting Learner Autonomy of Struggling Students, Budapest International Research and Critics Institute-Journal, 4 (6) (2023) 2132-2147.

[39] National Council of Teachers of Mathematics, Principles and Standards for School Mathematics, National Council of Teachers of Mathematics, (2000).

Numerical Simulation of Contaminant in Transient Groundwater System

Visarut Huayshelake[1,*] and Khamron Mekchay[1]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

This work aims to study the transportation of contaminants in a groundwater system, which is one of important problems in environmental science. The study relies on numerical simulations of a system of partial differential equations; the transport equation for contaminant flow and transient groundwater flow. The numerical methods are handled by the discontinuous Galerkin (DG) method for spatial coordinates, and together with the second order two-stage Alexander scheme for the time evolution system. The implementation for the simulations is performed based on DUNE library. The study of contaminant is considered with and without pump and treat (P&T) techniques, where the simulation results are provided via examples together with accuracy validation of the numerical method.

Keywords: Discontinuous Galerkin, Contaminant Transportation, Transient Groundwater
2020 MSC: 76M99, 76S05, 65M99

1 Introduction

Groundwater contamination is one of the major environmental issues that treatens to global water resources and human health. According to a survey of the United States Geological Survey, approximately one of five tested groundwater wells have at least one chemical contaminant [1] which are hazadous to daily consumption. Therefore, to address this issue, factors such as the type of groundwater flow, the properties of the aquifer, and the techniques used to treat the contaminants are considered when developing a mathematical model to predict the future scenario.

Mathematical models are essential tools to simulate a complex phenomenon, such as the transport of contaminants in groundwater. One commonly used model is the advection-diffusion equation, which describes the mechanism of contaminant transport in groundwater; however, this equation cannot explain the influence of external factors like a changing in groundwater velocity. To address this limitation, the transient groundwater equation is applied, as it accounts for time-varying flow. The connection between these two equations is described by a groundwater hydraulic head gradient, which represents a spatial change in the hydraulic head and relates to the advection term in the contaminant transport equation.

*Speaker and corresponding author (visarut.hu@gmail.com)

Since the governing equations are a system of partial differential equations (PDEs), numerical methods, such as the Finite Element Method (FEM) and Discontinuous Galerkin (DG) method are commonly employed to solve them. The FEM is a widely used numerical technique for simulating groundwater contaminant in domain with irregular geometries. For instance, one of the first models that applied this method to solve the coupled system of both two equations was developed by Yeh [2] in 1981. Building upon this, Khebchareon [3] extended this approach by developing a two-dimensional finite element to solve the coupled system of the steady-state groundwater flow equation and the advection-dispersion equation. Similarly, Smaoiu [4] impremented an FEM code to solve transient groundwater in heterogeneous media.

Since groundwater contamination models must effectively capture advection-dominated transport to provide reliable predictions, to address this challenge, the Discontinuous Galerkin (DG) method is introduced for solving this problems. The Discontinuous Galerkin (DG) method, introduced in the 1970s, employs discontinuous piecewise polynomial approximations to achive high orders of accuracy and has proven effective for advection-dominated problems [5]. In 2003, Wheeler et al. [6] applied this method to simulate a far-field nuclear waste management problem. Subsequently, Dawson, Sun, and Wheeler [7] investigated a system of equations in which the transport algorithm is compatible in 2004. Later, Povich [8] extended its application to coupled equations with variable density in 2012.

Regarding groundwater contaminanation remediation, the most commonly used method is the Pump and Treat (P&T) technique [9]. This technique is the groundwater well established approach involves extracting contaminanted water, transporting it to a designated treatment area, and then discharging the treated water back into the groundwater system. For example, in 2008, Sharief et al [10] applied this concept to optimize the pumping rate for cleaning up a confined aquifer with FEM. Additionally, Ahmadi et al [11] investigated the effect of pumping rate on treating contaminant groundwater with Finite Volume Method (FVM).

From these advantages in both numerical methods and remediation techniques, this study investigates the effectiveness of the P&T method for remediating contaminant transport within an anisotropic aquifer under transient ground waterflow condition. The numerical simulations are handled by the discontinuous Galerkin (DG) method for spatial coordinates, and by the second order two-stage Alexander scheme for the time evolution system.

2 Contaminant Flow in Groundwater Model

In this simulation, for more real-world realistic mathematical model, the governing equations are the advection-diffusion equation and the transient groundwater flow equation. The advection-diffusion equation describes the transport of contaminant caused by the advection and diffusion processes in the groundwater system, and this equation is derived from the conservation of mass law. For spatial space $(x, y) \in \Omega \subset \mathbb{R}^2$, and temporal space $t \in [0, T)$, let a function $c(x, y, t)$ be the contaminant concentration in the groundwater. The advection-diffusion equation can be expressed in mathematical as:

$$\frac{\partial c}{\partial t} + \nabla \cdot (c\mathbf{u}) - \nabla \cdot (D\nabla c) = f_c, \tag{2.1}$$

where $\mathbf{u}(x, y, t)$ is the groundwater velocity, $D(x, y)$ is the diffusion coefficient, and f_c is contaminant source or sink function.

The transient groundwater flow equation explains the groundwater movement. This equation is derived from Darcy's law and continuity equation. For a saturated porous media, let $h(x, y, t)$ be a hydraulic head function. The transient groundwater flow equation can be written as follows [12]:

$$S_s \frac{\partial h}{\partial t} - \nabla \cdot (K \nabla h) = f_h, \tag{2.2}$$

where S_s is the specific storage coefficient, $K(x, y)$ is the hydraulic conductivity coefficient, and f_h is the groundwater source or sink function.

The groundwater velocity $\mathbf{u}$ in equation (2.1) can be expressed in the gradient respected to a spatial domain by using the Darcy's law:

$$\mathbf{u} = -\frac{K \nabla h}{\phi}, \tag{2.3}$$

where a constant ϕ is the aquifer porosity.

To build the general system of the mathematical model, let $\Omega \subset \mathbb{R}^2$ be a bounded domain, $\partial\Omega = \Gamma_D \cup \Gamma_N$ be the boundary of Ω which is subjected to Dirichlet boundary (Γ_D) and Neumann boundary (Γ_N), and n be the outward unit normal vector. The governing equations can be written as:

$$\begin{aligned}
\frac{\partial c}{\partial t} + \nabla \cdot (c\mathbf{u}) - \nabla \cdot (D \nabla c) &= f_c && \text{for } \Omega \times (0, T), && (2.4)\\
S_s \frac{\partial h}{\partial t} - \nabla \cdot (K \nabla h) &= f_h && \text{for } \Omega \times (0, T), && (2.5)\\
c(x, y, t) &= g_c && \text{on } \Gamma_D \times (0, T), && (2.6)\\
h(x, y, t) &= g_h && \text{on } \Gamma_D \times (0, T), && (2.7)\\
\frac{\partial c}{\partial n} &= 0 && \text{on } \Gamma_N \times (0, T), && (2.8)\\
\frac{\partial h}{\partial n} &= 0 && \text{on } \Gamma_N \times (0, T), && (2.9)\\
c(x, y, 0) &= c^0 && \text{on } \Omega, && (2.10)\\
h(x, y, 0) &= h^0 && \text{on } \Omega, && (2.11)
\end{aligned}$$

where g_c and g_h is contaminant concentration and hydraulic head corresponded to Dirichlet boundary condition, and c^0 and h^0 is the initial for concentration and hydraulic head, respectively.

3 Numerical Methods

For numerically solving the system of governing equations, the governing equation is changed to a weak form. Let $H^1(\Omega)$ be the Sobolev space defined by

$$H^1(\Omega) := \left\{ f \in L^2(\Omega) : \partial f \in L^2(\Omega) \right\}, \tag{3.1}$$

where $L^2(\Omega)$ is the space of square integrable functions and ∂f is the first order weak derivative of the function f. Let Ψ and Ξ denote the Sobolev spaces for contaminant concentation function

$c(x, y, t)$ and hydraulic head function $h(x, y, t)$, respectively. For test functions $\psi \in \Psi$ and $\xi \in \Xi$, weak forms of equations (2.4) to (2.11) are:

$$\int_\Omega \frac{\partial c}{\partial t}\psi dx + \int_\Omega \nabla \cdot (c\mathbf{u})\psi dx + \int_\Omega D\nabla c \cdot \nabla \psi dx + \int_{\Gamma_D} g_c \psi ds = \int_\Omega f_c \psi dx, \tag{3.2}$$

$$\int_\Omega S_s \frac{\partial h}{\partial t}\xi dx + \int_\Omega K\nabla h \cdot \nabla \xi dx + \int_{\Gamma_D} g_h \xi ds = \int_\Omega f_h \xi dx. \tag{3.3}$$

In this study, the Discontinuous Galerkin method (DG) is used to solve the governing equation in spatial domain. Let $\mathcal{T}_h$ be the finite element mesh consisting of elements T, and let $\mathcal{E}_h$ be the set of all element interfaces. The notations $(\cdot,\cdot)_{\tau_h}$ and $(\cdot,\cdot)_{\mathcal{E}_h}$ represent the discrete L^2 inner products with respect to $\mathcal{T}_h$ and $\mathcal{E}_h$, respectively. Then, the semi-discrete scheme is written as

$$\left(\frac{\partial c_h}{\partial t}, \psi\right)_{\tau_h} = (c_h\mathbf{u}, \nabla\psi)_{\tau_h} - \left((c_h^{up}\mathbf{u}) \cdot n, [\psi]\right)_{\mathcal{E}_h} - (D\nabla c_h, \nabla\psi)_{\tau_h} + (\{D\nabla c_h \cdot n\}, [\psi])_{\mathcal{E}_h} + (f_c, \psi)_{\tau_h}, \tag{3.4}$$

$$\left(S_s \frac{\partial h_h}{\partial t}, \xi\right)_{\tau_h} = -(K\nabla h_h, \nabla\xi)_{\tau_h} + (\{K\nabla h_h \cdot n\}, [\xi])_{\mathcal{E}_h} + (f_h, \xi)_{\tau_h}, \tag{3.5}$$

where $c_h^{up}\mathbf{u}$ is an upwind flux, $\{v\}$ is an average of function v, and $[v]$ is a jump of v.

Let $\{t_n\}_{n=0}^N$ be the discrete time points partitioning the interval $[0, T]$, with a time step defined as $\Delta t = t_{n+1} - t_n$. The contaminant concentration and hydraulic head at time t_n are denoted by c^n and h^n, respectively. For time discretization, the second-order and two stage Alexander scheme which known as a diagonally implicit Runge-Kutta (RK) method [13] is applied to discretize the time domain. This scheme provides a stability, making well-suited for numerical solution of the governing equations. The general implicit equation, for solving the contaminant concentration c of the next time step is given by

$$c_1^n = c^n + \alpha\Delta t f(t_n + \alpha\Delta t, c_1^n), \tag{3.6}$$

$$c_2^n = c^n + \Delta t(1-\alpha) f(t_n + \alpha\Delta t, c_1^n) + \alpha\Delta t f(t_n + \Delta t, c_2^n), \tag{3.7}$$

$$c^{n+1} = c^n + \Delta t(1-\alpha) f(t_n + \alpha\Delta t, c_1^n) + \alpha\Delta t f(t_n + \Delta t, c_2^n) \tag{3.8}$$

where α is a weight which equals to $1 - \frac{\sqrt{2}}{2}$.

4 Numerical Experiment

In this study, the Distributed and Unified Numerical Environment (DUNE) is used to implement the governing equations, by using the DG scheme with the second order and two stage Alexander scheme, to the programming language. DUNE is a free and open toolbox based on C++ language for the grid-based numerical imprementation such as fininte element (FEM), finite volume (FVM), or discontinuous galerkin (DG) methods [14]. The computational setup includes an Intel i7-12700 processor with 16 GB of RAM, and equilateral rectangle element for the mesh.

The model parameters for all simulations in this study are shown in Table 1.

Table 1: Input data for all simulations

Parameters	Value
Diffusion coefficient ($D(x, y)$)	5.46×10^{-8} m^2/s
Specific storage coefficient (S_s)	1
Hydraulic conductivity ($K(x, y)$)	2.69×10^{-6} to 1.3×10^{-3} m/s
Aquifer porosity (ϕ)	1
Time step (Δt)	1 day (86400 s)
Final time (T)	2000 days
Domain (Ω)	$[0, 100]^2$ m^2
Element dimension ($\Delta x, \Delta y$)	1 m $\times$1 m

To represent the heterogeneity of an aquifer, the hydraulic conductivity ($K(x, y)$) is generated by randomization from a log-normal distribution, defined by the equation

$$K(x, y) = \exp\left(Y(x, y)\right), \tag{4.1}$$

where $Y(x, y)$ is a Gaussian random field which is the normal distribution with mean -9.5 and varriance 1. To simulate the anisotropic aquifer, the correlation lengths are set to 20 m. in the x-direction and 5 m. in the y-direction. Basesd on these setting, the aquifer is represented to be a mixture of clay and fine sand, with the range of hydraulic conductivity is approximately 2.69×10^{-6} m/s to 1.3×10^{-3} m/s [15,16]. Figure 1 shows the result of the randomized hydraulic conductivity used in model simulation.

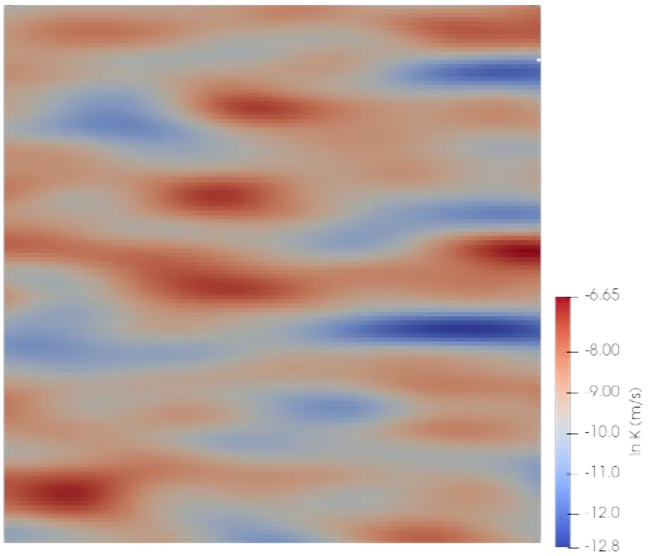


Figure 1: Aquifer texture generated by randomized hydraulic conductivity

4.1 Accuracy Validation

In this section, the model validation is discussed to study the accuracy of numerical solutions for the contaminant concentration function $c(x, y, t)$. Since this problem does not have an analytical solution, a convergence of numerical solution is performed. The numerical solution should converge to the numerical solution from the finest mesh. The finest mesh, $h = 0.00125$ meter with sufficiently small time step $\Delta t = 1$ day is used as a benchmark for comparing the numerical solutions obtained from the coarser mesh sizes. The compared mesh size that used in this study are $h = 0.02, 0.01, 0.005$, and $h = 0.0025$ meter.

For the accuracy validation, let $c_h(x, y, t)$ be the contaminant concentration for the mesh size h at grid point (x, y), $E_{max}(c_h)$ be a discrete maximum norm, $E_{L^2}(c_h)$ be a discrete L^2 norm, and $\text{EOC}(h_i)$ be an order of convergence which define by the following definitions:

$$E_{max}(c_h, t) = \max_{(x,y)\in\Omega_h} |c_{exact}(x, y, t) - c_h(x, y, t)|, \tag{4.2}$$

$$E_{L^2}(c_h, t) = \sqrt{\frac{\sum_{(x,y)\in\Omega_h} (c_{exact}(x, y, t) - c_h(x, y, t))^2}{n_h}}, \tag{4.3}$$

where n_h is the number of grid point in Ω_h and

$$\text{EOC}(h_i) = \frac{\log(E_{h_{i-1}}/E_{h_i})}{\log(h_{i-1}/h_i)}, \tag{4.4}$$

where E_{h_i} denotes the maximum error over time $t \in [0, T]$ for E_{max} and E_{L^2} with mesh size h_i.

The validated model is configured to simulate contaminant transport in transient flow, as described in Example 4.1.

Example 4.1 (**Contaminant in transient flow**).

In this example, the boundary for the hydraulic head of groundwater h is set to 1 m. on Γ_1 and 0.8 m. on Γ_3. The other boundaries are set to be Neumann boundary conditions which the net flux equals to zero. For the contaminant concentation, the boundary condition is set to outflow boundary. The initial plume is defined on $(28, 36) \times (42, 50)$ with a concentration 1 mg/L (see Figure 2a).

Table 2: Errors and rates of order of convergence for example 4.1

h	$E_{max}(c)$	$\text{EOC}(h_i)$	$E_{L^2}(c)$	$\text{EOC}(h_i)$	Time (sec)
0.02	0.8435	-	0.0429	-	206
0.01	0.7619	0.1468	0.0395	0.1191	1065
0.005	0.5249	0.5376	0.0290	0.4458	6386
0.0025	0.1556	1.7542	0.0103	1.4934	45557

From the Table 2, both errors measured by E_{max} and E_{L^2} decrease as the mesh size h decreases. This suggests that the simulations converge and stable when finer grids are used, which validate the accuracy of the method (see also Figure 3a and 3b); however, these smaller

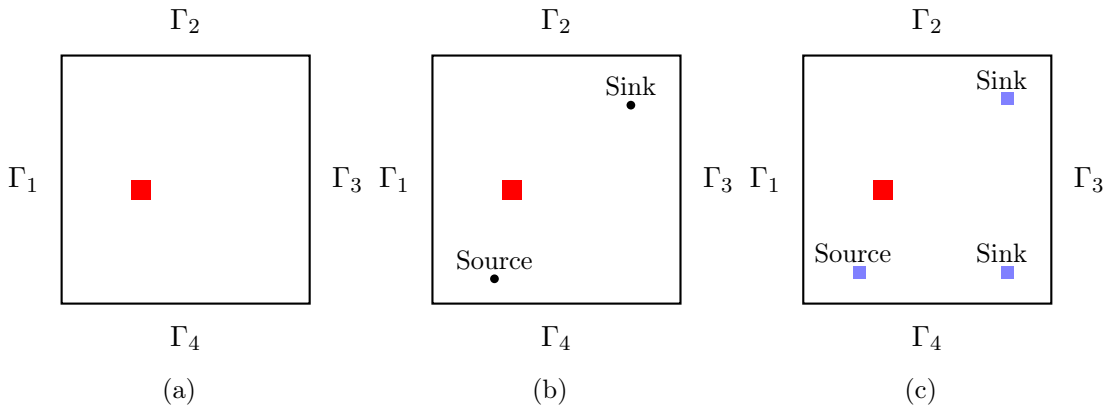


Figure 2: (a) Problem domain for groundwater pollutant with no treatment, (b) a single-point P&T technique, and (c) a regional zone P&T technique

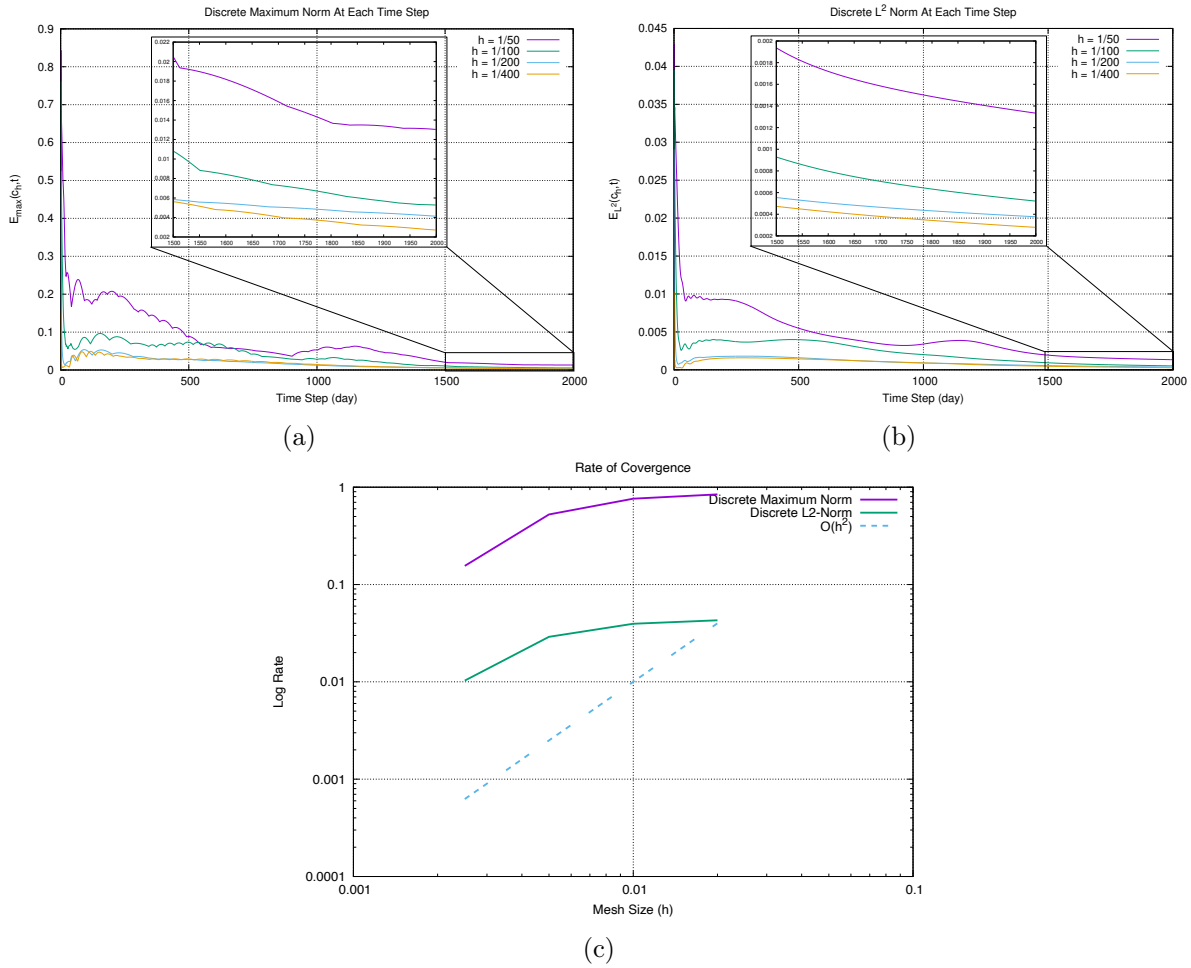


Figure 3: (a) Discrete maximum norm, (b) L^2 norm, and (c) rate of convergence of Example 4.1

grid sizes require more processing time for simulation. Moreover, as shown in Figure 3c, The EOC is approximately predicted to follow $O(h^2)$ as the grid size h decreases.

Figure 4 shows the simulations results for contaminant concentration at different times. The contaminant appears on the source of the contaminant at the initial time (Figure 4a), and then transports away from the source by the groundwater flow (Figure 4b and 4d), where the spreading of the plume (a high concentration of contaminant region) is influenced by the hydraulic of conductivity. Figure 5 illustrates that simulated velocity field at various times; at the initial time, the flow is relatively uniform (Figure 5a). As time progresses the flow paths are distinct (Figure 5b), and the velocity field of groundwater reaches a steady state around 120 days where the velocity field seems to be identical (see Figure 5c and 5d).

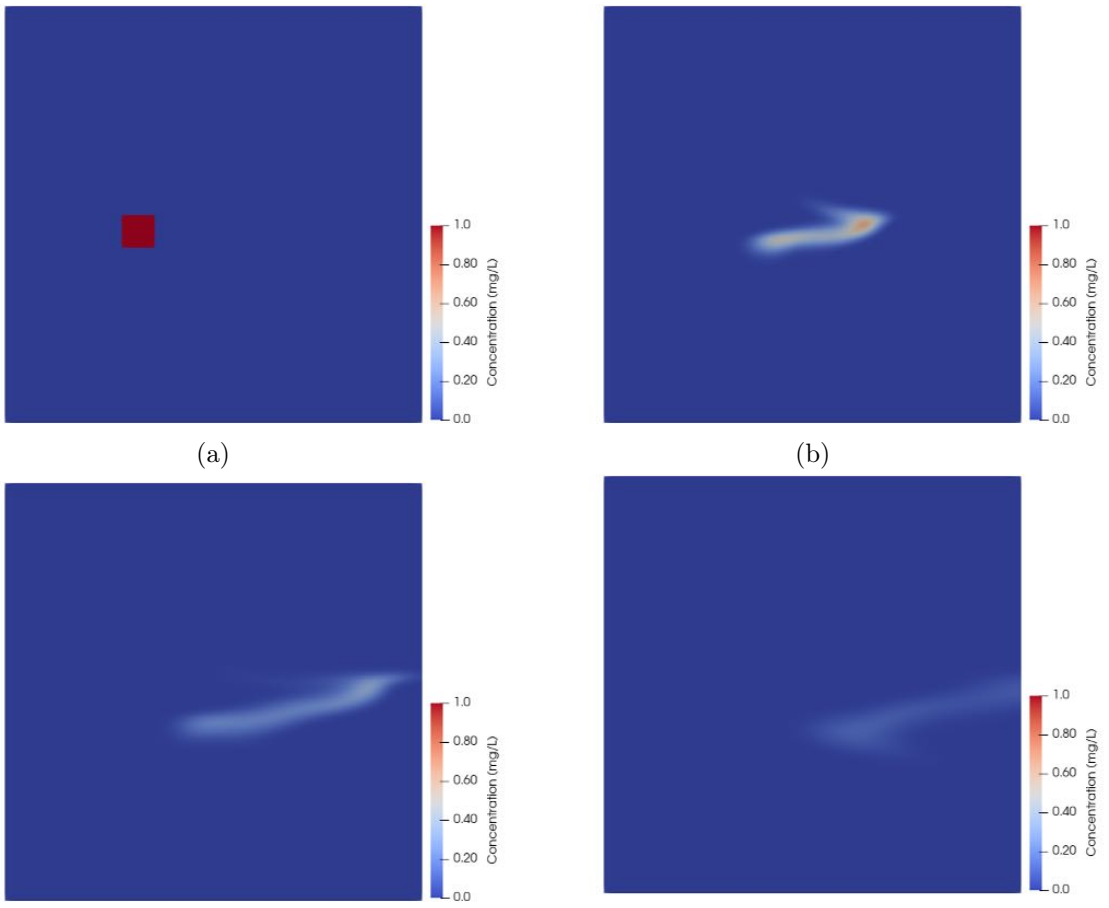


Figure 4: (a) Pollutant flow at time $t = 0$ day, (b) 500 days, (c) 1000 days, and (d) 2000 days

4.2 Simulations of Groundwater Contamination Remediation with P&T technique

This study aims to investigate groundwater contaminant treatment, the Pump and Treat methods (P&T). Two simulation scenarios; single point and region zone, are presented for analyzing a behavior of contaminant extraction in groundwater system.

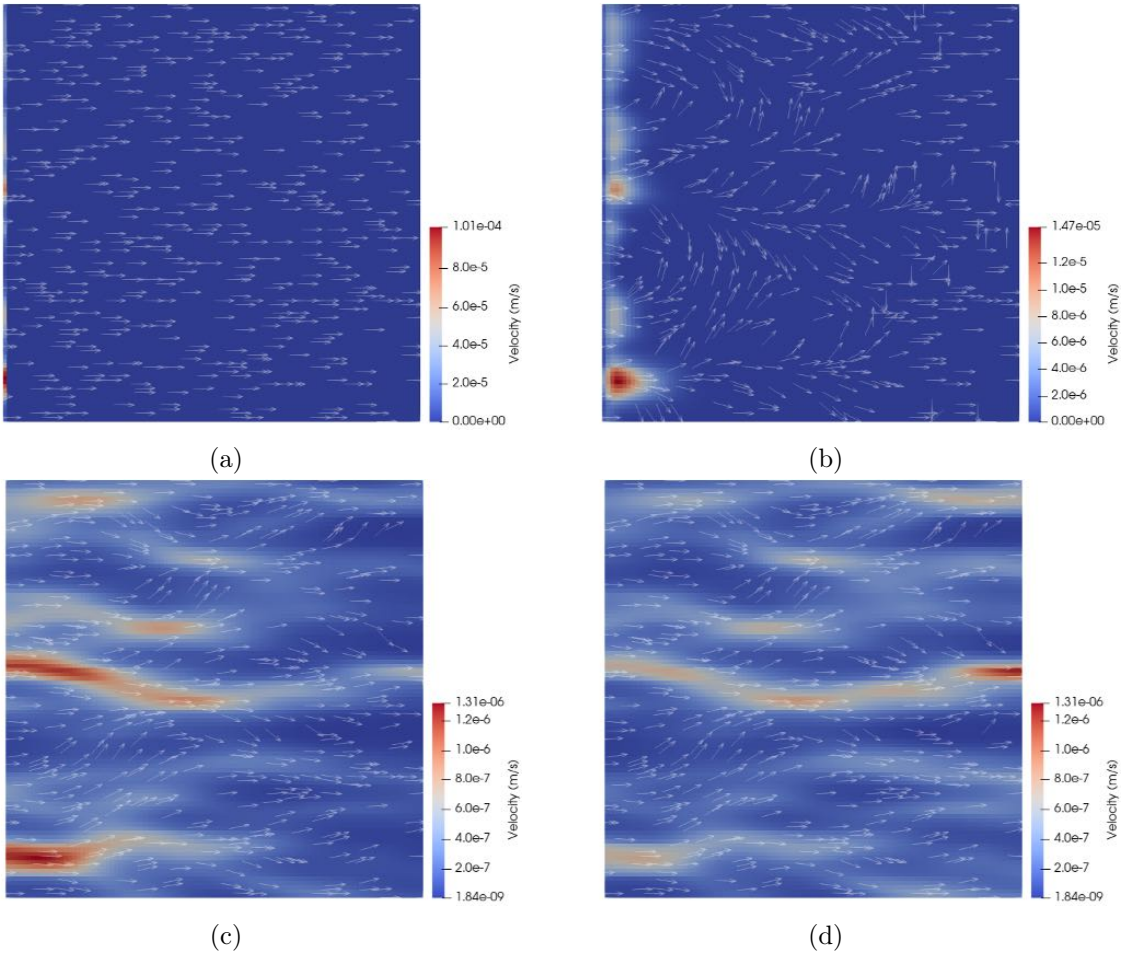


Figure 5: (a) Velocity field at time $t = 0$ day, (b) 1 day, (c) 120 days, and (d) 2000 days

Example 4.2 (**P&T with points sink and source**).

In this example, the effect of a single-point P&T technique is studied. One groundwater source point and one groundwater sink point are added to the studied area, the domain Ω, and operated throughout the simulation continuously. As shown in Figure 2b, the source point is set up at $(25, 10)$ with the rate of injection 9.25×10^{-5} m^3/s, and the sink point is set up at $(80, 80)$ with the rate of extraction 1.01×10^{-5} m^3/s.

Figure 6 demonstrates that the groundwater sink effectively removes contaminants from the system. As illustrate in Figure 6c, the concentration drops significantly to 0.003 mg/L at 1000 days and further decreses to 4×10^{-6} mg/L at the day 2000 (Figure 6d). This reduction is due to the high velocity of groundwater source, which pushes the contaminant towards the sink point. Figure 7 shows how the source point influences the flow pattern of groundwater, which reaches to a steady state after 120 days (see Figure 7c and 7d).

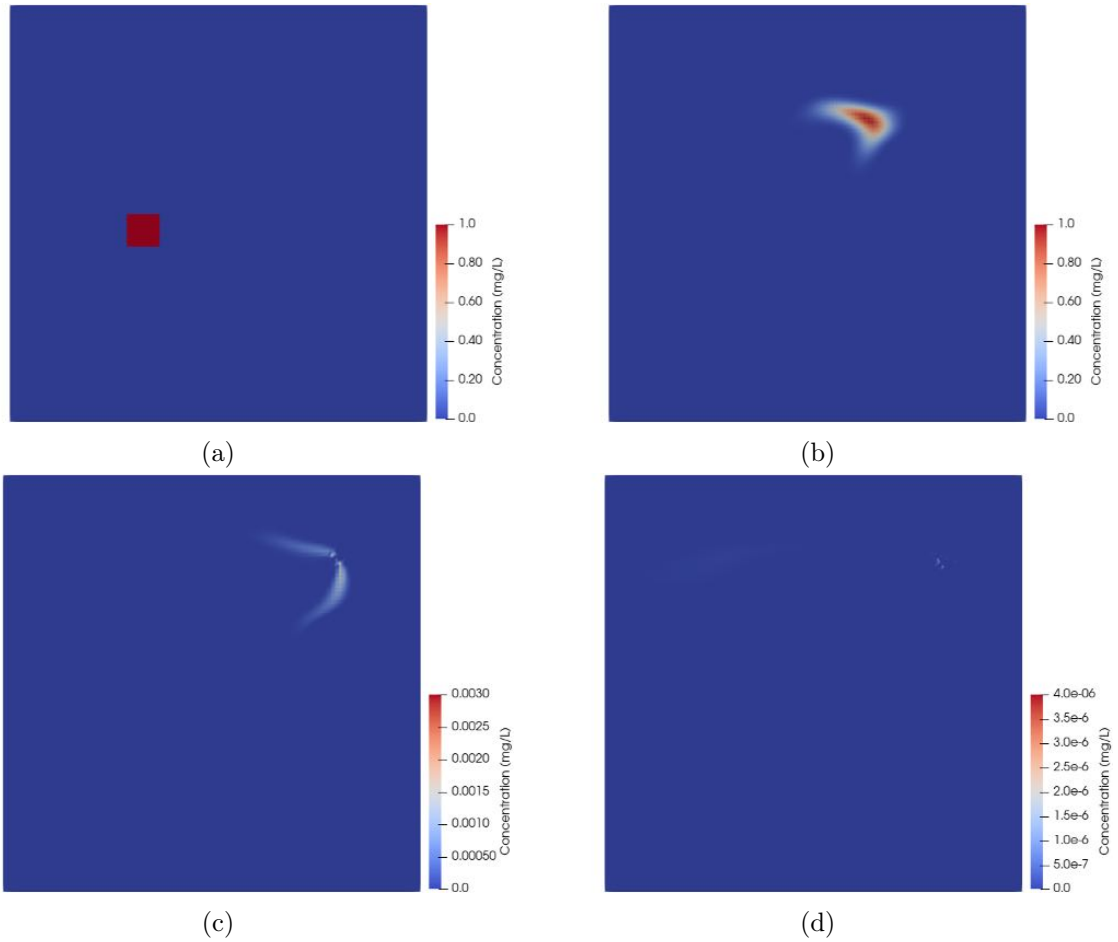


Figure 6: Contaminant concentration at time $t = 0$ day (a), 500 days (b), 1000 days (c), and 2000 days (d)

Example 4.3 (**P&T with regions sink and source**).

In this example, the sink and source point have replaced by region zones with dimension 5×5 m^2. and another sink region zone has added in the domain Ω, as shown in Figure 2c. Parameters for each zone are given in Table 3 .

Table 3: Parameters for simulation in example 4.3

Parameters	Source	Sink1	Sink2
Position	$(20, 25) \times (10, 15)$	$(80, 85)^2$	$(80, 85) \times (10, 15)$
Rate (m^3/s)	9.25×10^{-5}	1.01×10^{-5}	6.01×10^{-5}
Hydraulic conductivity (m/s)	1×10^{-3}	1×10^{-3}	1×10^{-5}

Figure 8 show that the sink region can remove the contaminant from the system. As shown in Figure 8b, the plume is advected from the contaminant source towards the right size of the

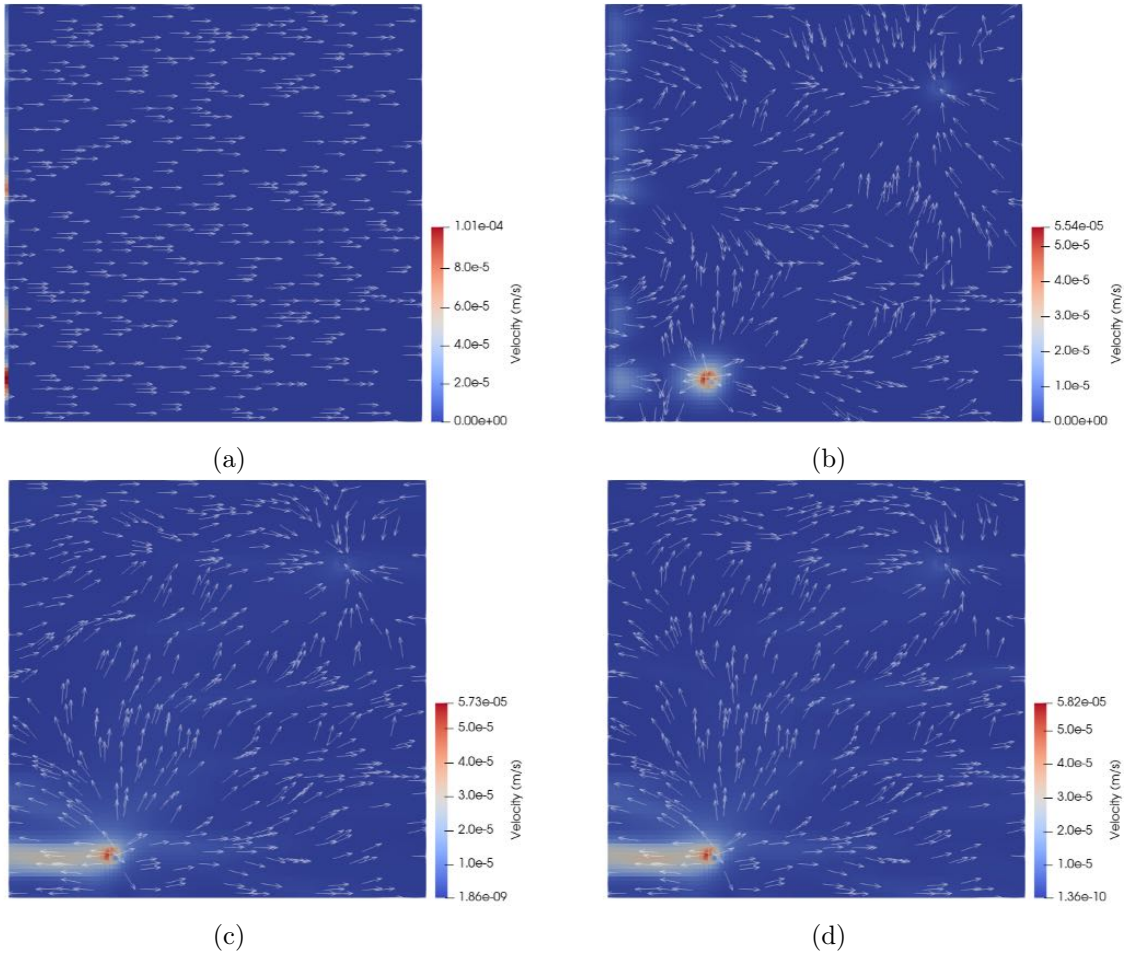


Figure 7: Velocity field at time $t = 0$ day (a), 1 day (b), 120 days (c), and 2000 days (d)

domain due to groundwater velocity from the source region; while most of contaminant flows out of the system, some is extracted by the sink regions. The upper-right sink appears to be more effective at contaminant removal than the bottom-right sink, as illustrate in Figure 8c and 8d.

5 Conclusion

In this study, the Discontinuous Galerkin (DG) method, combined with a two-stage, second-order Alexander scheme, is used to solve a coupled system of the advection-diffusion equation and the transient groundwater flow equation. The numerical solution for the concentration function $c(x, y, t)$ is obtained through an implementation in the DUNE library. Moreover, the inclusion of the transient term in the groundwater flow equation, along with the use of randomized hydraulic conductivity values, contributes to a more realistic simulation. The accuracy of the simulation is evaluated using the discrete maximum norm and discrete L^2 norm is computed following to the approximation in the second order. The Pump and Treat (P&T) technique is applied to the model for remediating contaminant groundwater. The simulations show that a strategically

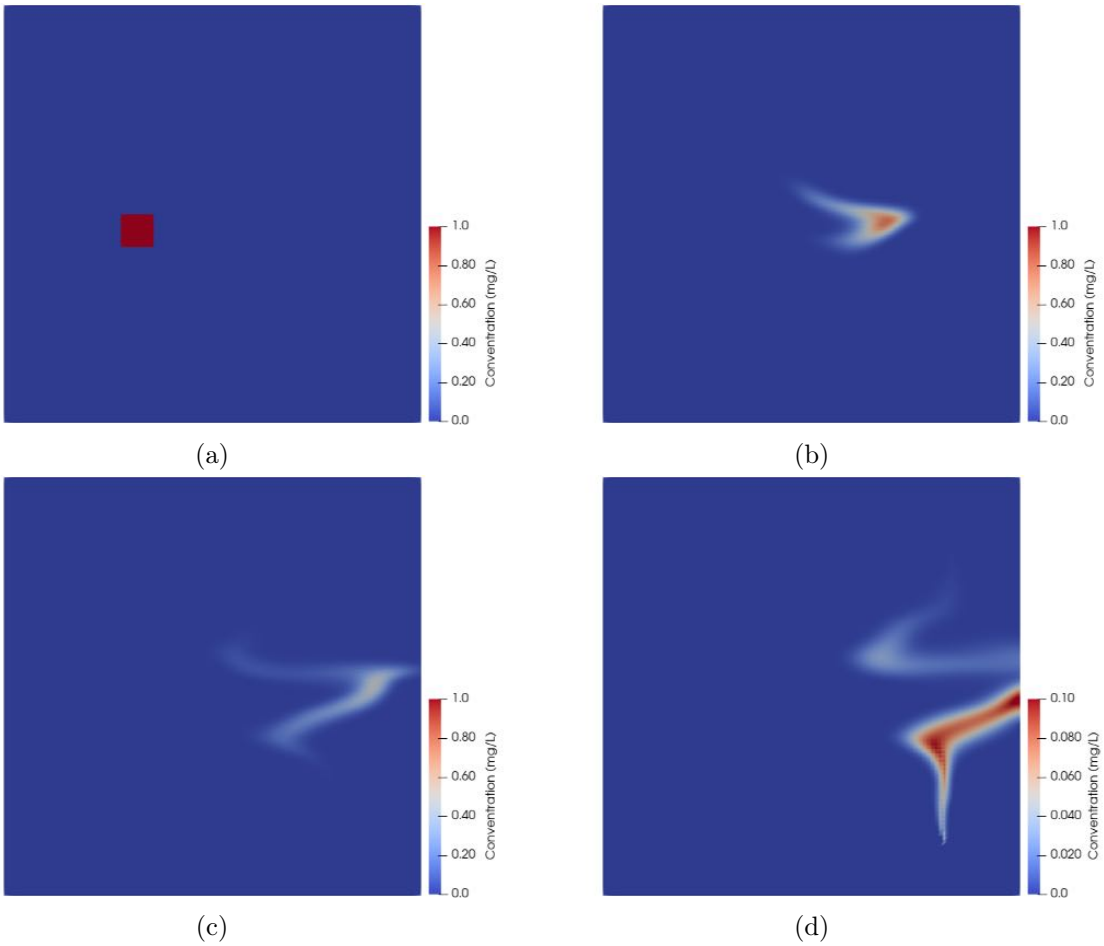


Figure 8: (a) Contaminant concentration at time $t = 0$ day, (b) 500 days, (c) 1000 days, and (d) 2000 days

placed groundwater source can drive the contaminant plume towards a groundwater sink, leading to significant contaminant removal and a reduction in overall concentration within the system.

This simulation can help decision-makers predict the movement of the contaminant and design efficient groundwater well locations and optimal pumping rates to minimize operational costs and timelines. Future studies could explore using the real-world data or coupling the model with optimization algorithms.

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments. The first author would like to acknowledge the Development and Promotion of Science and Technology (DPST) for the financial support.

References

[1] D. A. Leslie, M. B. Peter, R. R. Michael, The Quality of Our Nation's Waters: Water Quality in Principal Aquifers of the United States, 1991-2010, 2015.

[2] G. T. Yeh, and W. S. David. FEMWASTE: a Finite-Element Model of Waste Transport through Saturated-Unsaturatedporous Media, Oak Ridge National Laboratory, 1981.

[3] K. Morrakot, Crank-Nicolson finite element for 2-D groundwater flow, advection-dispersion and interphase mass transfer: I. Model development. International Journal of Numerical Analysis and Modeling, 3 (2) (2012), 109–125.

[4] S. Hassan, Z. Lahcen, O. Abdellatif, C. Erick. Modelling of groundwater flow in heterogeneous porous media by finite element method. Hydrological Processes, 26 (4) (2012), 558–569.

[5] V. Dolejší, and F. Miloslav. Discontinuous Galerkin Method: Analysis and Applications to Compressible Flow, Springer, 2015.

[6] W. F. Mary, S. Shuyu, E. Owen, R. Beatrice. Discontinuous Galerkin method for modeling flow and reactive transport in porous media. Analysis and Simulation of Multifield Problems, 12 (2003), 37–56.

[7] D. Clint, S. Shuyu, & W. F. Mary. Compatible algorithms for coupled flow and transport. Computer Methods in Applied Mechanics and Engineering, 193 (2004), 2565–2580.

[8] J. P. Timothy. Discontinuous Galerkin (DG) methods for variable density groundwater flow and solute transport, (Ph.D Dissertation), University of Texas at Austin, USA, 2012.

[9] S. Rajandrea, and M. D. Antonio. Groundwater Engineering: A Technical Approach to Hydrogeology, Contaminant Transport and Groundwater Remediation, Springer, 2019.

[10] S. M. V. Sharief, T. I. Eldho, & A. K. Rastogi. Optimal pumping policy for aquifer decontamination by pump and treat method using genetic algorithm. ISH Journal of Hydraulic Engineering, 14(2) (2008), 1–17.

[11] H. Ahmadi, F. Kilanehei, M. Nazari-Sharabian. Impact of pumping rate on contaminant transport in groundwater–A numerical study. Hydrology, 8(3) (2021), 103.

[12] I. Jonathan, Groundwater Modeling by the Finite Element Method, Wiley, 1989.

[13] A. Roger, Diagonally implicit Runge-Kutta methods for stiff O.D.E.'S. SIAM Journal of Numerical Analysis, 14 (6) (1977), 1006–1021.

[14] B. Peter, B. Markus, D. Andreas, et al, The Dune framework: Basic concepts and recent developments. Computers & Mathematics with Applications, 81 (2021), 75–112.

[15] D. M. Braja, Principles of Geotechnical Engineering, 7th Edition, Cengage Learning, 2010.

[16] D. A. Patrick and S. W. Franklin, Physical and Chemical Hydrogeology, 2nd Edition, John Wiley & Sons, 1990.

A Large Neighborhood Search for Two-picker Routing Problem with Picker Blocking Consideration

Panupong Daengpradap[1,*] and Panupong Vichitkunakorn[1,2,†]

[1]Division of Computational Science, Faculty of Science,
Prince of Songkla University, Songkhla, Thailand

[2]Research Center in Mathematics and Statistics with Applications,
Prince of Songkla University, Songkhla, Thailand

Abstract

In this paper, we consider the two-picker routing problem with picker blocking consideration in a single block warehouse with narrow aisles. We aim to find routes of two pickers that minimize total travel time while avoiding collision as much as possible. Our method is based on the large neighborhood search together with a penalty when two pickers are near each other.

Keywords: picker routing problem, picker blocking, large neighborhood search, metaheuristics, penalty

2020 MSC: 90B06, 90B20, 90B80, 65K99

1 Introduction

Order picking is an important process in warehousing. It was estimated to account for up to 55% of the total warehouse operating expense [6]. To optimize the order picking process, we can study picker routing optimization which aims to find a travel path to minimize the travel time. Recently, metaheuristics have been used extensively in picker routing problems. Some examples are the Large Neighborhood Search (LNS) [11], the Adaptive Large Neighborhood Search (ALNS) [14], tabu search method [15], etc.

When two or more pickers work together in the same warehouse whose aisles are very narrow, picker blocking can occur. In this case, metaheuristics can also be used to solve the picker routing problem such as the Ant Colony Optimization (ACO) and the Genetic Algorithm (GA) [3-5]. However, the past research aims to find only collision-free paths. In addition, some kinds of collision resolution policies are used, for example, a picker is forced to wait outside the aisle until the aisle is available. In this paper, we introduce an LNS that gives penalty when pickers are near each other. This is under the expectation that, when the penalty is high enough, our method will produce a route with a very small number of collisions.

Speaker (panupong.daeng@gmail.com)
Corresponding author (panupong.v@psu.ac.th)

2 Problem description

We are interested in a two-picker routing problem where both pickers collaboratively pick all items in each order. Our problem is a joint assigning and routing problem. Firstly, the items in an order are separated into two groups and are assigned to two pickers. Then, each picker plans their own route. In particular, the picker chooses a sequence of items to be picked. We call this sequence a picklist. Then, a picker starts from the depot and travels to all item locations, according to the picklist. When the pick is completed, the picker must bring all the items back to the depot. We consider a simple case of a single block warehouse consisting of two cross aisles, the front cross aisle and the back cross aisle, for pickers to move between aisles. See Figure 1 for an example. In the figure, each square represents a storage location, while a shaded square represents a location of an item in the customer order. Pickers can enter an aisle from any end, traverse in both directions, and change direction when inside the aisle. We assume that the picking aisles are narrow enough so that pickers can pick items from both sides of the aisle without an additional side-to-side move.

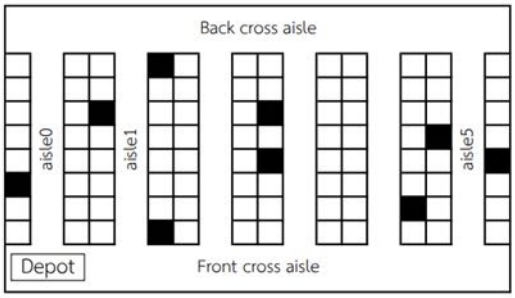


Figure 1: A single block layout warehouse

This travel time inside an aisle will be only in the direction along the aisle. In addition, pickers cannot pass each other in picking aisles and cross aisles. The depot is located at the front cross aisle of the first picking aisle. For each order, the picking process is finished when both pickers have obtained all assigned items and return to the depot. So, for each picking order, we define a makespan to be the length of time when the last picker comes back to the depot. In other words, the makespan is the maximum of the travel time between the two pickers.

Due to the narrow aisles, picker blocking can occur in multiple situations. 1) Two pickers travel toward each other in an aisle in the opposite direction. See subfigures (a) and (b) in Figure 2. 2) When one picker is picking an item, while the other want to travel pass through the location. See subfigure (c) in Figure 2.

For each picking order, the objective is to assign items to both pickers and design a route for each picker so that the makespan is minimum while avoiding picker blocking. However, it is hard to find the optimal routes for two pickers that avoid blocking. In this paper, we propose a penalty-based Large Neighborhood Search (LNS) algorithm to solve the problem. The idea is to introduce a penalty when two routes intersect. We then include the penalty into the objective function so that the obtained solution has the number of blockings as low as possible.

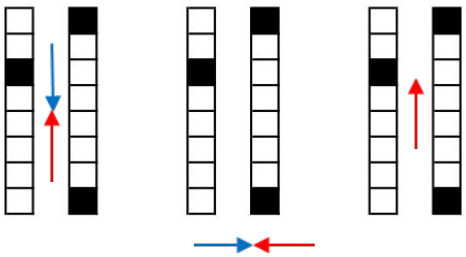


Figure 2: Picker blocking in various situations

3 Solution method

In this section, we aim to find a route for each picker to avoid a collision. However, it is not easy to find the collision-free routes in a general situation. We propose a solution method which introduces a penalty when two pickers collide.

3.1 Problem instance

We consider a single block warehouse with m picking aisles and n storage locations on each side of a picking aisle. We assume that each storage location has one meter (m) width. We also assume for simplicity that each aisle is 1 m wide, and the travel speed of the pickers is 1 m/s. The distance between two adjacent picking aisles is w meters, which we set $w = 3$ in this paper. The number of items in each order is k. From the setting, the time for a picker to move between two consecutive picking locations is 1 second, and the time for a picker to move from the front cross aisle to the back cross aisle is $n + 1$ seconds.

3.2 Item coordinate and travel time between two item locations

We represent an item location by a coordinate (x, y) for $x \in \{0, 1, \dots, m-1\}$ and $y \in \{1, 2, \dots, n\}$. The depot is at the coordinate $(0, 0)$. Let $(x, 0)$ and $(x, n+1)$ denote the front end and the back end of the picking aisle x at the front and back cross aisles respectively. We let $C = \{0, 1, \dots, m-1\} \times \{0, 1, \dots, n+1\}$ be the set of all coordinates in the warehouse including cross aisles. To find a path with the shortest time (and hence the shortest distance) between two locations (x_1, y_1) and (x_2, y_2), we need to compare two routes. See Figure 3. The first route travels through the front cross aisle. Its travel time is $t_1 = y_1 + y_2 + w \cdot |x_1 - x_2|$. The second route travels through the back cross aisle. Its travel time is $t_2 = 2(n+1) - (y_1 + y_2) + w \cdot |x_1 - x_2|$. Hence, the minimum travel time from (x_1, y_1) to (x_2, y_2) is $t\big((x_1, y_1), (x_2, y_2)\big) = \min(t_1, t_2)$.

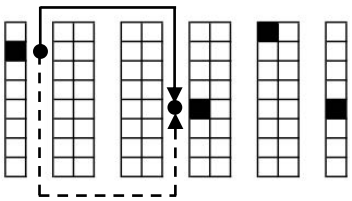


Figure 3: Two routes between two locations.

3.3 Penalized makespan

Let the customer order be $I = \{i_1, i_2, \dots, i_k\}$. A solution to our problem is $S = (S_1, S_2)$ where $S_1 = (i_1^{(1)}, i_2^{(1)}, \dots, i_j^{(1)})$ and $S_2 = (i_{j+1}^{(2)}, i_{j+2}^{(2)} \dots, i_k^{(2)})$ s.t. $(i_1^{(1)}, i_2^{(1)} \dots, i_j^{(1)}, i_{j+1}^{(2)}, \dots, i_k^{(2)})$ is a permutation of I and $1 \leq j \leq k-1$. In this case, Picker 1 needs to pick j items, while Picker 2 needs to pick $k-j$ items. We call S_1 and $S_2$the picklist of Picker 1 and 2, respectively. For example, Picker 1 will start from the depot, travels to the location $i_1^{(1)}$, then to the location $i_2^{(1)}$, and so on until $i_j^{(1)}$, then travels back to the depot.

We will define a penalty value when there is blocking between two pickers. We first write each picklist as a list of all locations at each second the picker will visit. Let

$$L^{(1)} = \left\{\ell_0^{(1)}, \ell_1^{(1)}, \dots, \ell_{T_1}^{(1)}\right\}$$

be a sequence of coordinates in C obtained from the route of Picker 1. The coordinate $\ell_t^{(1)}$ is the coordinate at time t (Seconds) where t runs from 0 to T_1. Similarly, we let

$$L^{(2)} = \left\{\ell_0^{(2)}, \ell_1^{(2)}, \dots, \ell_{T_2}^{(2)}\right\}$$

be a sequence of coordinates of Picker 2. In this case, the makespan is $\max(T_1, T_2)$. We define the penalty

$$p = p\left(L^{(1)}, L^{(2)}\right) = \alpha \cdot c\left(L^{(1)}, L^{(2)}\right)$$

where α is the penalty coefficient (Which we can vary) and

$$c\left(L^{(1)}, L^{(2)}\right) = \left|\left\{t \mid \left\|\ell_t^{(1)} - \ell_t^{(2)}\right\| \leq 1\right\}\right|$$

is the number of times (seconds) both pickers are at the same location or at the adjacent locations. The norm which we are using here is the Euclidean norm: $\|(x, y)\|^2 = x^2 + y^2$. We then define the penalized makespan

$$T^* = T_\alpha(S) = \max(T_1, T_2) + p.$$

Note our penalized makespan is different from the delayed makespan which takes into account waiting time or picker travel speed reduction from collision. Instead, we aim to find collision-free routes and use penalty as an intermediate method to achieve that solution. The penalized makespan calculation is separated into two parts. A travel time of each picker without blocking consideration is computed. Then, the penalty value is computed from the given routes.

For example, we consider a problem instance where $m = 3, n = 8$ *and* $k = 5$ as shown in the figure below. We set the penalty coefficient $\alpha = 4$. Let the customer order I be $I = \{1, 2, 3, 4, 5\}$. We consider a solution $S = (S_1, S_2)$ where the pick lists of Picker 1 and Picker 2 are $S_1 = \{1,4,3\}$ and $S_2 = \{2,5\}$, respectively. An example of pick routes of both pickers is shown in Figure 4. Both pickers start and end at the depot, where they start at the same time. We note that the distance between two adjacent aisles is set as $w = 3$. The picker travel speed is fixed as 1 m/s.

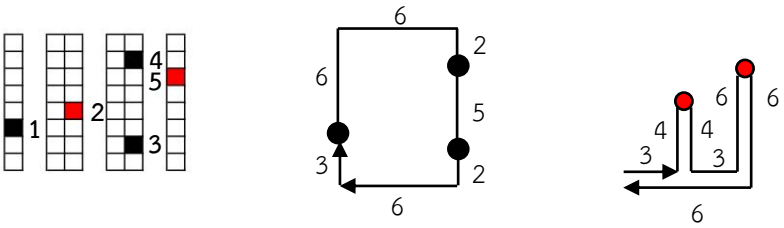


Figure 4: A route of two pickers.

The first picker travel time is $T_1 = 30$, and the second picker travel time is $T_2 = 32$. The makespan is equal to 32. After considering the list of coordinates of both routes, we get $c\left(L^{(1)}, L^{(2)}\right) = |\{0, 19\}| = 2$. Hence, penalized makespan $T^*(S) = 32 + (4 \cdot 2) = 40$.

3.4 The large neighborhood search (LNS)

Large Neighborhood Search (LNS) is a metaheuristic optimization technique introduced by [11]. It is designed to solve large-scale problems and often yields improved solutions compared to existing methods. The process begins with the creation of an initial solution, either randomly or using a heuristic method [15]. It will be treated as the current solution. In each iteration, we apply two main operations to the current solution: the destroy operation and the repair operation, which are indicated by destroy operators and repair operators, respectively. In our case, a destroy operator removes items from a picklist in the current solution, creating an incomplete solution (as there are items not in any picklists). Then, a repair operator reconstructs the solution by reinserting the removed items, maybe in their original positions or new positions in the pick sequence. It can also be inserted to the picklist of the other picker. The penalized makespan of the resulting solution, generated through the destroy and repair process, is then calculated. If the penalized makespan of the new solution satisfies an acceptance criteria, it is accepted as the current solution. This iterative process continues until a termination condition is met. Among all the solutions that have been considered, the solution with the minimum penalized makespan is chosen as the output of our LNS.

3.4.1 Destroy-Repair operators

As mentioned above, the principle components of LNS are the choices of destroy and repair operators. In this paper, we consider the following two destroy operators and two repair operators. Let r be a positive integer such that $r \leq k$.

1) Random-Destroy operator, $RD(r)$: We randomly remove r items from the current solution.

2) Best-Destroy operator, $BD(r)$: For each item in the current solution, we compute the penalized makespan of the solution which results from removing the item from the solution. We choose a solution which has the minimum penalized makespan. This process continues until r items have been removed.

3) Random-Repair operator, $RR(r)$: From the previous destroy operator, we get r removed items and an incomplete solution. We then randomly insert r items to the solution.

4) Best-Repair operator, $BR(r)$: For each item that has been removed from the destroy operation, we insert the item to the incomplete solution then compute the penalized makespan. After that, we choose a solution which has the minimum penalized makespan. This process continues until all r items have been inserted.

3.4.2 Flow chart

In Figure 5, we summarize our LNS. An initial solution is established randomly. Let ρ represent the maximum removal percentage. Hence, the number of items to be removed by the destroy operators satisfies $1 \leq r \leq \rho k$. For the acceptance criteria, we define β to be the accepting threshold. The current solution S is updated when a newly created solution S' satisfies $T^*(S') < (1+\beta) \cdot T^*(S)$. The algorithm stops after N iterations. Then, it returns S_B, the best solution found.

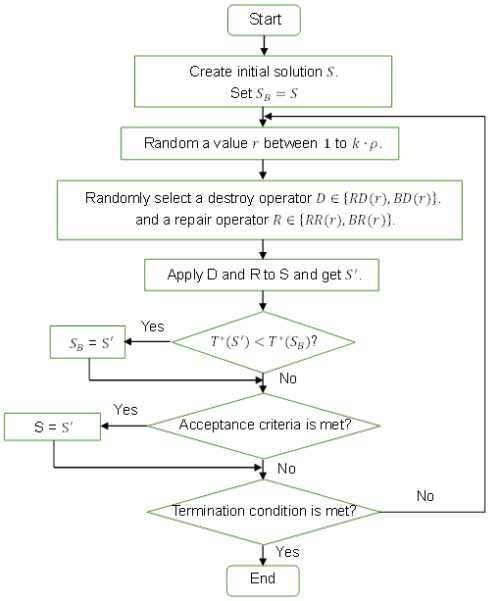


Figure 5: An algorithm of the LNS.

4 Main Results

This section, we calculate the penalized makespan and computing time under various conditions. We first tune the two parameters: the accepting threshold, the removal percentage, using the penalized makespan and computing time in Section 4.1. We then use their tuned values throughout the section. After that, we show the performance of the proposed algorithm in Section 4.2. In Section 4.3, we study relationship between penalty and other variables including penalized makespan, the number of collisions.

In our experiment, the values of the three parameters: the number of items in an order, the number of aisles and the aisle length, are from [10]. The number of items is 10 or 30, the number of aisles is 7 or 15, and the aisle length is 10 or 30. There are in total 8 scenarios. The accepting threshold varies among 5, 10 and 15. The distance between adjacent aisles is three meters. Since the pickers travel with stable velocity of one meter per second, it is equivalent to 3 seconds. All experiments were done using $N = 1000$ for the LNS and were done on a PC with an Intel Core i3-12100 12th generation (3.30 GHz) processor and 16 GB of RAM.

4.1 Parameter tuning

We set $\alpha = 5$. In each one of the eight scenarios, we consider three values of the accepting threshold $\beta = 5\%, 10\%$ and 15%, and four values of the removal percentage $\rho = 10\%, 20\%, 30\%$ and 40%. We compare the resulting penalized makespan from the algorithm to the minimum penalized makespan in the same scenario. We record it as a percentage difference to the minimum value, $PD_{\beta,\rho}$, which is calculated by:

$$PD_{\beta,\rho} = \frac{T^*_{\beta,\rho} - T^*_{min}}{T^*_{min}},$$

where $T^*_{\beta,\rho}$ is the resulting penalized makespan in the scenario (β, ρ) and $T^*_{min} = \min_{\beta,\rho} T^*_{\beta,\rho}$.

		Warehouse Layouts							
Number of aisles		7	7	15	15	7	7	15	15
Number of cells		10	10	10	10	30	30	30	30
Number of items		10	30	10	30	10	30	10	30
β (%)	ρ (%)	S1	S2	S3	S4	S5	S6	S7	S8
5	10	5.22	19.87	2.87	12.53	4.31	18.77	3.31	14.21
	20	2.63	10.42	1.47	5.90	2.31	8.54	1.64	5.55
	30	1.72	4.45	0.69	3.07	1.56	3.31	0.90	2.18
	40	0.98	1.46	0.31	0.89	0.72	1.04	0.45	0.99
10	10	3.87	17.30	2.38	11.48	3.33	14.69	2.54	12.95
	20	2.19	7.87	1.35	4.90	1.69	6.21	1.18	5.14
	30	0.79	4.03	0.86	2.18	0.99	2.11	0.56	1.62
	40	0.63	0.77	0.37	**0.00**	0.15	0.51	0.10	0.37
15	10	3.14	15.51	2.45	12.69	3.00	14.93	1.87	14.33
	20	2.16	7.09	1.24	4.54	1.38	4.80	1.16	4.21
	30	1.32	2.94	0.51	1.81	0.44	1.99	0.57	1.53
	40	**0.00**	**0.00**	**0.00**	0.24	**0.00**	**0.00**	**0.00**	**0.00**

Figure 6: Percentage difference to the minimum makespan $PD_{\beta,\rho}$.

From Figure 6, the setting $\beta = 15\%$ and $\rho = 40\%$ gives the smallest percentage difference in most scenarios. We notice that the setting $\beta = 15\%$ and $\rho = 20\%$ gives a similar performance. Their percentage differences differ by only at most 5% except in the scenario 2. However, the computing time is significantly lower, especially in a big scenario such as scenario 8 (55.42 minutes to 35.67 minutes). So, we pick the setting $\beta = 15\%$ and $\rho = 20\%$ for further experiments.

4.2 Performance

From Figure 7, we show the relationship between the number of iterations used in the LNS and the average penalized makespan (in the second), by varying three penalty values $\alpha = 0, 5 \text{ and } 10$. The average penalized makespan is calculated from the eight scenarios. We see that the solutions from the LNS begin to converse at around 200 iterations, in all three values of α. Even though we use 1000 iterations in all of our experiments, Figure 7 shows that we still can reduce the number of iterations and get almost the same result.

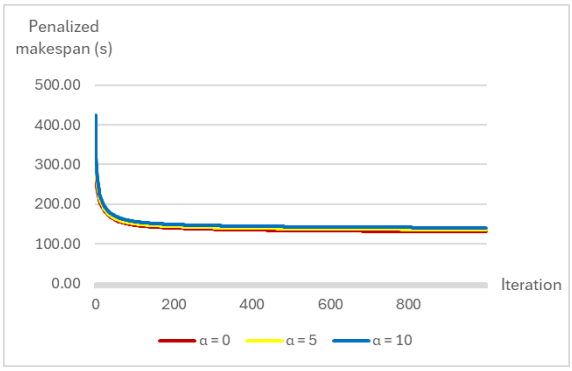


Figure 7: The number of iterations and the average penalized makespan.

4.3 Relationship between penalty value to other variables

From Figure 8 (left), we compare the penalized makespan of different penalty values. Since the penalized makespan when $\alpha = 0$ is always the smallest, we will consider the gap of penalized makespan to the minimum $(at\ \alpha = 0)$. We found that when the penalty value increases, the gap will also increase linearly, reaching 8% when $\alpha = 10$. In other words, a unit increase of α gives rise to around 1% increase of the penalized makespan. Since α is the coefficient of the blocking time, we can infer that the contribution of blocking is around 1% of the total makespan.

Next, we want to study the reduction of collision of pickers when α is varied. In Figure 8 (right), we begin from the case when $\alpha = 0$, where there is no collision reduction. When $\alpha = 1$, the number of collisions is reduced by 85.87%. When α increases further, the reduction also increases and reaches 91.87% when $\alpha = 10$. This indicates that our LNS helps reduce the collision by a very large scale, which satisfies our original goal where no collision can happen.

The reduction begins to stabilize when $\alpha = 4$, with the reduction of 91.26%. So, it is reasonably to set α to be at least 4 when aiming to prevent collision.

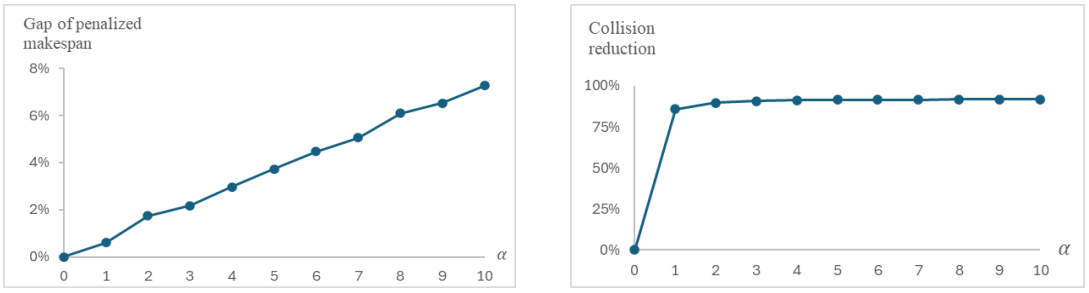


Figure 8: Gaps of penalized makespan to the minimum case when $\alpha = 0$ (left) and the collision reduction for different penalties (right)

5 Discussion

This paper proposed a solution method for the two-picker routing problem avoiding the picker collision as much as possible. Instead of a collision-free algorithm, we introduced a LNS with picker blocking penalty. We then studied the effects of penalty value to the penalized makespan and the number of iterations. The results show that the penalized makespan increases by 1% given a unit increase of α. The number of collisions is reduced by 85.87% when using $\alpha = 1$ instead of $\alpha = 0$, and is reduced by 91.26% - 91.87% when using $\alpha \geq 4$.

For future research, we suggest a study in various directions including a case of more than two pickers, a case of a larger warehouse, or a larger order size. Furthermore, our solution method reduces picker blocking by imposing a penalty at the end of the picker tour construction. This means the blocking can still happen. It is interesting to see a solution method for collision-free picker routing.

Acknowledgment. This work was supported by the PSU-TUYF Charitable Trust Fund, Print of Songkhla University, Contrast no. 2-2564-03.

References

[1] C. G. Petersen, "An evaluation of order picking routeing policies," Int. J. Oper. Prod. Manag., vol. 17, no. 11, pp. 1098–1111, 1997.

[2] C. G. Petersen, "The impact of routing and storage policies on warehouse efficiency," Int. J. Oper. Prod. Manag., vol. 19, no. 10, pp. 1053 – 1064, 1999.

[3] F. Chen, et al., "An ACO-based online routing method for multiple order pickers with congestion consideration in warehouse," J. Intell. Manuf., vol. 27, pp. 389 – 408 , 2016.

[4] F. Chen, et al., "An ant colony optimization routing algorithm for two order pickers with congestion consideration," Comput. Ind. Eng., vol. 66, no. 1, pp. 77 – 85 , 2013.

[5] F. Chen, G. Xu, and Y. Wei, "Heuristic routing methods in multiple-block warehouses with ultra-narrow aisles and access restriction," Int. J. Prod. Res., vol. 57, no. 1, pp. 228 – 249 , 2019.

[6] J. A. Tompkins, et al., Facilities Planning, John Wiley & Sons, 2010.

[7] J. C.-H. Pan and P.-H. Shih, "Evaluation of the throughput of a multiple-picker order picking system with congestion consideration," Comput. Ind. Eng., vol. 55, no. 2, pp. 379 – 389, 2008.

[8] J. C.-H. Pan, P.-H. Shih, and M.-H. Wu, "Storage assignment problem with travel distance and blocking considerations for a picker-to-part order picking system," Comput. Ind. Eng., vol. 62, no. 2, pp. 527 – 535, 2012.

[9] J. D. C. Little, et al., "An algorithm for the traveling salesman problem," Oper. Res., vol. 11, no. 6, pp. 972 – 989, 1963.

[10] K. J. Roodbergen and R. de Koster, "Routing methods for warehouses with multiple cross aisles," Int. J. Prod. Res., vol. 39, no. 9, pp. 1865 – 1883, 2001.

[11] P. Shaw, "Using constraint programming and local search methods to solve vehicle routing problems," in Proc. Int. Conf. Principles Pract. Constraint Program., Berlin, Heidelberg: Springer, 1998, pp. 417 – 431.

[12] R. de Koster, T. Le-Duc, and K. J. Roodbergen, "Design and control of warehouse order picking: A literature review," Eur. J. Oper. Res., vol. 182, no. 2, pp. 481 – 501, 2007.

[13] R. Elbert, et al., "Agent-based analysis of picker blocking in manual order picking systems: Effects of routing combinations on throughput time," in Proc. 2015 Winter Simul. Conf. (WSC), IEEE, 2015.

[14] S. Ropke and D. Pisinger, "An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows," Transp. Sci., vol. 40, no. 4, pp. 455 – 472 , 2006.

[15] T. Phannikul and S. Sindhuchao, "A customized tabu search for the vehicle routing problem with simultaneous pick-up and delivery," Sci. Technol. Asia, vol. 2010, pp. 70 – 82, 2010.

Analysis of a Conservative Finite Difference Method for Burgers-type Equation

Keerati Buchatip[1,*] and Nattapol Ploymaklam[2,†]

[1]Graduate Master Degree Program in Applied Mathematics, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand

[2]Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand

Abstract

We are interested in designing a finite difference method for solving Burgers-type equations describing various phenomena such as shock wave theory, mathematical modeling of turbulent fluids, and continuous stochastic processes. Examples of Burgers-type equations include Burgers' equation, Burgers-Huxley equation, and Burgers-Fisher equation. In this work, we use the Burgers-Fisher equation as a case study. The equation describes phenomena in fluid dynamics, heat conduction, shock wave formation, thermodynamics, fiber optics, etc. The analytical solution of this equation preserves energy, so we propose a second-order implicit finite difference scheme that conserves energy in a discrete sense. Numerical experiments are performed to verify the accuracy with respect to temporal and spatial variables. The results also confirm the discrete conservation of energy.

Keywords: Burgers-type equation, Burgers-Fisher equation, Finite difference, Discrete conservation

2020 MSC: 65M06, 65M12

1 Introduction

Partial differential equations (PDEs) are fundamental in modeling natural and engineering phenomena including heat transfer, fluid mechanics, quantum theory, electromagnetism, and plasma physics. Among nonlinear PDEs, Burgers-type equations describe models in shock wave theory, turbulent fluid modeling, and stochastic processes [1]. Examples of Burgers-type equations include Burgers' equation, Burgers-Huxley equation, and Burgers-Fisher equation [2].

The Burgers-Fisher equation [3–5] is a crucial model widely applied in fluid dynamics, heat conduction, shock wave formation, thermodynamics, and fiber optics. We mainly consider the following initial-boundary value problem

$$u_t - u_{xx} + uu_x = \lambda u(1-u), \quad 0 < x < L, \quad 0 < t \leq T, \tag{1.1}$$

$$u(x,0) = \varphi(x), \quad 0 \leq x \leq L, \tag{1.2}$$

$$u(0,t) = u(L,t) = 0, \quad 0 < t \leq T, \tag{1.3}$$

*Speaker (keerati_buchatip@cmu.ac.th)

†Corresponding author (nattapol.p@cmu.ac.th)

where the coefficient λ is a positive constant, $\varphi(x)$ is a given function satisfying $\varphi(0) = \varphi(L) = 0$.

We will briefly review the two basic components of the Burgers-Fisher equation, the Burgers equation, and the Fisher equation to better understand thier structures. The Burgers' equation is a fundamental PDE in fluid dynamics, nonlinear acoustics, and traffic flow. H. Bateman introduced the equation in 1915 [6, 7] and was later studied by J.M. Burgers in 1948 [8]. It is given by

$$u_t + uu_x = \nu u_{xx}, \tag{1.4}$$

where $\nu > 0$ is the viscosity coefficient.

The Fisher equation is a reaction-diffusion model describing traveling wave solutions. It is used in ecology, physiology, and phase transitions. R.A. Fisher proposed this equation in 1937 [9]. In the same year, A. Kolmogorov, I. Petrovskii, and N. Piskunov introduced a more general reaction-diffusion form [10]:

$$u_t - u_{xx} = \lambda u(1 - u), \tag{1.5}$$

where λ is a positive constant.

Historically on the Burgers-Fisher equation, there are many works in the analytical aspects. For example, H. A. Levine, et al. [11] analyzed the Burgers-Fisher equation within the class of Convective Reaction-Diffusion equations. Subsequently, A.M. Wazwaz [2, 12] studied the exact solutions; H. Zhang, et al. [13] analyzed the traveling wave behavior; and Y. Zhou, et al. [14] examined the bifurcation behavior.

Numerical approaches based on various frameworks, such as finite difference and finite element methods, have been extensively studied in the context of solving the Burgers-Fisher equation. Among these, finite difference methods are particularly well-established; however, other numerical techniques have also been developed to address the equation. For instance, O.P. Yadav et al. [18] (2017) demonstrated the existence and uniqueness of weak solutions using Galerkin's finite element method, providing an approximation for the Burgers-Fisher equation. In the finite difference framework, L. Zhang, et al. [4] in 2014 studied the nonstandard explicit finite difference method. Following that, D.A. Hammad, and M.S. El-Azab [15] in 2015 studied the analysis and derivation of a compact fourth- and sixth-order implicit finite difference approximation schemes. Subsequently, C. Vinay, et al. [5] in 2016 studied implicit difference schemes and the Crank-Nicolson method, constructing a linear system with a tridiagonal matrix. In 2022, H. Hajinezhad [16] studied the implicit finite difference method for the Burgers-Fisher equation, leading to a nonlinear system of equations, which was then solved using the Jacobian-free Newton-Krylov method with the generalized minimal residual method (GMRES). Later, in 2024, A. Loyinmi, et al. [17] studied the weight average method for the Burgers-Fisher equation. Finally, in 2024 G. Gao, et al. [3] studied the three-level linearized finite difference method for solving the Burgers-Fisher equation, solvability, convergence, stability, and conservation properties using second-order and fourth-order compact difference schemes.

From the previous works, we observe that the two-level finite difference scheme for Burgers-Fisher equation has not yet been thoroughly explored. This approach allows us to use fewer neighboring points to approximate the solution and carry out the analysis. Therefore, we focus on designing and analyzing a two-level finite difference scheme for Burgers-Fisher equation (1.1)-(1.3) as a case study for Burger-type equations.

The rest of this paper is organized as follows. Section 2 focuses on deriving a difference scheme, exploring its analytical properties, and computing approximate solutions. The study

also investigates the discrete invariant-conserving properties of the numerical solutions by analyzing the proposed scheme in Section 3. Finally, some conclusion is given in Section 4.

2 Preliminaries

In this section, we present the properties of the analytical solution, along with notation, definitions, lemmas, and the derivation of the difference scheme.

2.1 Analytic property

Let $u \in C^{2,1}([0,L] \times [0,T])$.We start by presenting the invariant property of the solution to problem (1.1)-(1.3).

Theorem 2.1. *Let $u(x,t)$ be the solution of the problem (1.1)-(1.3). Denote*

$$E(t) = \int_0^L u^2(x,t)dx + 2\int_0^t \left[\int_0^L u_x^2(x,s)dx + \lambda \int_0^L \left[u^3(x,s) - u^2(x,s)\right] dx\right] ds,$$

Then

$$E(t) = E(0),\ 0 < t \leq T.$$

Proof. By differentiating $\frac{1}{2}E(t)$ with respect to t, we obtain

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}E(t) &= \int_0^L uu_t\ dx + \int_0^L u_x^2(x,t)dx + \lambda \int_0^L \left[u^3(x,t) - u^2(x,t)\right] dx \\
&= \int_0^L u(u_{xx} - uu_x + \lambda u(1-u))\ dx + \int_0^L u_x^2(x,t)dx + \lambda \int_0^L \left[u^3(x,t) - u^2(x,t)\right] dx \\
&= \int_0^L uu_{xx}\ dx - \int_0^L u^2u_x\ dx\ + \int_0^L u_x^2(x,t)dx \\
&= [u(x,t)u_x(x,t)]_0^L - \int_0^L u_x^2(x,t)\ dx\ - \frac{\left[u^2(x,t)u(x,t)\right]_0^L}{3} + \int_0^L u_x^2(x,t)dx \\
&= 0.
\end{aligned}$$

Therefore, the function $E(t)$ remains constant over time. That is, $E(t) = E(0)$. □

2.2 Notations and Lemmas

To solve the problem (1.1)-(1.3) numerically, we divide the domain $[0,L] \times [0,T]$. We choose positive integers M and N and let $h = L/M$ and $\tau = T/N$. We define the grid points as $x_i = ih$ for $0 \leq i \leq M$ and $t_k = k\tau$ for $0 \leq k \leq N$. The sets of grid points are $\Omega_h = \{x_i \mid 0 \leq i \leq M\}$, $\Omega_\tau = \{t_k \mid 0 \leq k \leq N\}$, and $\Omega_{h\tau} = \Omega_h \times \Omega_\tau$. For any grid function $u = \{u_i^k \mid 0 \leq i \leq M, 0 \leq k \leq N\}$ defined on $\Omega_{h\tau}$, we introduce the following notations:

$$u_i^{k+\frac{1}{2}} = \frac{1}{2}(u_i^k + u_i^{k+1}), \quad \delta_t u_i^{k+\frac{1}{2}} = \frac{1}{\tau}(u_i^{k+1} - u_i^k),$$

$$\delta_x u_{i-\frac{1}{2}}^k = \frac{1}{h}(u_i^k - u_{i-1}^k), \quad \delta_x^2 u_i^k = \frac{1}{h}\left(\delta_x u_{i+\frac{1}{2}}^k - \delta_x u_{i-\frac{1}{2}}^k\right), \quad \Delta_x u_i^k = \frac{1}{2h}(u_{i+1}^k - u_{i-1}^k).$$

As for the nonlinear term uu_x, we use the approximation introduced in [3], [21–24]

$$\psi(u,v)_i = \frac{1}{3}[u_i \Delta_x v_i + \Delta_x (uv)_i], \quad 1 \le i \le M-1.$$

The next step is the definition of the solution space as well as the definition of norm and seminorm.

Definition 2.2. Define

$$V_h = \{v | v = (v_0, v_1, \ldots, v_{M-1}, v_M)\},$$

and

$$V_h^0 = \{v | v \in V_h \ \ and \ \ v_0 = v_M = 0\}.$$

Definition 2.3. For any $u, v \in V_h$, we present the inner product as the following notations:

$$(u,v) = h\left(\frac{1}{2}u_0 v_0 + \sum_{i=1}^{M-1} u_i v_i + \frac{1}{2}u_M v_M\right), \quad \langle u, v\rangle = h\sum_{i=0}^{M-1}\left(\delta_x u_{i+\frac{1}{2}}\right)\left(\delta_x v_{i+\frac{1}{2}}\right),$$

and semi(norms).

$$\|u\| = \sqrt{(u,u)}, \quad |u|_1 = \sqrt{\langle u, u\rangle}, \quad ||u||_\infty = \max_{0\le i\le M}|u_i|.$$

Lemma 2.4. *For any grid functions $u \in V_h$ and $v \in V_h^0$, it follows that*

$$(\psi(u,v), v) = 0.$$

2.3 Derivation of the difference scheme

Define a grid function $U = \{U_i^k | 0 \le i \le M, 0 \le k \le N\}$ in $\Omega_{h\tau}$, where

$$U_i^k = u(x_i, t_k), \quad 0 \le i \le M, \quad 0 \le k \le N.$$

Consider (1.1) at the point $\left(x_i, t_{k+\frac{1}{2}}\right)$, we obtain

$$\begin{aligned} &u_t\left(x_i, t_{k+\frac{1}{2}}\right) - u_{xx}\left(x_i, t_{k+\frac{1}{2}}\right) + u\left(x_i, t_{k+\frac{1}{2}}\right) u_x\left(x_i, t_{k+\frac{1}{2}}\right) \\ &= \lambda\left[u\left(x_i, t_{k+\frac{1}{2}}\right)\left(1 - u\left(x_i, t_{k+\frac{1}{2}}\right)\right)\right], 0 \le i \le M-1, \quad 0 \le k \le N-1. \end{aligned} \tag{2.1}$$

Using Taylor expansion, we have

$$u_t\left(x_i, t_{k+\frac{1}{2}}\right) = \delta_t U_i^{k+\frac{1}{2}} + O(\tau^2), \tag{2.2}$$

$$\begin{aligned} u_{xx}\left(x_i, t_{k+\frac{1}{2}}\right) &= \frac{1}{2}\left[u_{xx}(x_i, t_k) + u_{xx}(x_i, t_{k+1})\right] + O(\tau^2) \\ &= \delta_x^2 U_i^{k+\frac{1}{2}} + O(\tau^2 + h^2), \end{aligned} \tag{2.3}$$

$$\begin{aligned} u\left(x_i, t_{k+\frac{1}{2}}\right) &= \frac{1}{3}\left[u\left(x_{i-1}, t_{k+\frac{1}{2}}\right) + u\left(x_i, t_{k+\frac{1}{2}}\right) + u\left(x_{i+1}, t_{k+\frac{1}{2}}\right)\right] + O(h^2) \\ &= \frac{1}{3}\left[U_{i-1}^{k+\frac{1}{2}} + U_i^{k+\frac{1}{2}} + U_{i+1}^{k+\frac{1}{2}}\right] + O(\tau^2 + h^2), \end{aligned} \tag{2.4}$$

$$\begin{aligned} u_x\left(x_i, t_{k+\frac{1}{2}}\right) &= \frac{1}{2}\left[u_x(x_i, t_k) + u_x(x_i, t_{k+1})\right] + O(\tau^2) \\ &= \Delta_x U_i^{k+\frac{1}{2}} + O(\tau^2 + h^2). \end{aligned} \tag{2.5}$$

Next, we consider the nonlinear term

$$\lambda\left[u\left(x_i,t_{k+\frac{1}{2}}\right)\left(1-u\left(x_i,t_{k+\frac{1}{2}}\right)\right)\right]=\lambda\left[u\left(x_i,t_{k+\frac{1}{2}}\right)-u^2\left(x_i,t_{k+\frac{1}{2}}\right)\right]. \tag{2.6}$$

Expanding this term using the Taylor series, we obtain

$$u(x_i,t_k)=u(x_i,t_{k+\frac{1}{2}})-\frac{\tau}{2}u_t(x_i,t_{k+\frac{1}{2}})+O(\tau^2)$$

and

$$u(x_i,t_{k+1})=u(x_i,t_{k+\frac{1}{2}})+\frac{\tau}{2}u_t(x_i,t_{k+\frac{1}{2}})+O(\tau^2),$$

or equivalently

$$u(x_i,t_{k+\frac{1}{2}})=u(x_i,t_k)+\frac{\tau}{2}u_t(x_i,t_{k+\frac{1}{2}})+O(\tau^2)$$

and

$$u(x_i,t_{k+\frac{1}{2}})=u(x_i,t_{k+1})-\frac{\tau}{2}u_t(x_i,t_{k+\frac{1}{2}})+O(\tau^2).$$

Thereby,

$$\begin{aligned}u^2(x_i,t_{k+\frac{1}{2}})&=\left[U_i^k+\frac{\tau}{2}u_t(x_i,t_{k+\frac{1}{2}})+O(\tau^2)\right]\left[U_i^{k+1}-\frac{\tau}{2}u_t(x_i,t_{k+\frac{1}{2}})+O(\tau^2)\right]\\&=U_i^kU_i^{k+1}+O(\tau^2).\end{aligned} \tag{2.7}$$

Substitute (2.7) into (2.6), and then substitute (2.2)-(2.6) into (2.1). We get

$$\begin{aligned}&\delta_tU_i^{k+\frac{1}{2}}-\delta_x^2U_i^{k+\frac{1}{2}}+\frac{1}{3}\left[U_{i-1}^{k+\frac{1}{2}}+U_i^{k+\frac{1}{2}}+U_{i+1}^{k+\frac{1}{2}}\right]\Delta_xU_i^{k+\frac{1}{2}}\\&\quad=\lambda\left(U_i^{k+\frac{1}{2}}-U_i^kU_i^{k+1}\right)+R_i^{k+\frac{1}{2}},\quad 1\le i\le M-1,\quad 0\le k\le N-1.\end{aligned} \tag{2.8}$$

We can show there exists a constants c_1 such that

$$|R_i^{k+\frac{1}{2}}|\le c_1(\tau^2+h^2),\quad 0\le i\le M-1,\quad 0\le k\le N-1. \tag{2.9}$$

Equation (2.8) can be written as

$$\begin{aligned}&\delta_tU_i^{k+\frac{1}{2}}-\delta_x^2U_i^{k+\frac{1}{2}}+\psi\left(U^{k+\frac{1}{2}},U^{k+\frac{1}{2}}\right)_i\\&\quad=\lambda\left(U_i^{k+\frac{1}{2}}-U_i^kU_i^{k+1}\right)+R_i^{k+\frac{1}{2}},\quad 1\le i\le M-1,\quad 0\le k\le N-1.\end{aligned} \tag{2.10}$$

Noticing the initial and boundary value conditions (1.2) and (1.3), we obtain

$$U_i^0=\varphi(x_i),\quad 1\le i\le M-1, \tag{2.11}$$

$$U_k^0=0,\ U_M^k=0,\quad 1\le k\le N. \tag{2.12}$$

Omitting the small term $R_i^{k+\frac{1}{2}}$ in (2.8) and replacing U_i^k with u_i^k, we obtain the difference scheme for problem (1.1)-(1.3) as follows:

$$\begin{aligned}&\delta_tu^{k+\frac{1}{2}}-\delta_x^2u^{k+\frac{1}{2}}+\psi\left(u^{k+\frac{1}{2}},u^{k+\frac{1}{2}}\right)_i\\&=\lambda\left(u_i^{k+\frac{1}{2}}-u_i^ku_i^{k+1}\right),\quad 1\le i\le M-1,\quad 1\le k\le N-1,\end{aligned} \tag{2.13}$$

$$u_i^0=\varphi(x_i),\quad 1\le i\le M-1, \tag{2.14}$$

$$u_k^0=0,\quad u_M^k=0,\quad 1\le k\le N. \tag{2.15}$$

Remark 2.5. The difference scheme (2.13)–(2.15) is a two-level nonlinear scheme. Define

$$u^k = (u_0^k, u_1^k, \ldots, u_M^k).$$

From (2.14) and (2.15), we know that u^0 is given. Suppose that u^k has already been computed. Then, based on (2.13) and (2.15), we obtain the following equations:

$$\frac{2}{\tau}\left(u_i^{k+\frac{1}{2}} - u_i^k\right) - \delta_x^2 u_i^{k+\frac{1}{2}} + \psi\left(u_i^{k+\frac{1}{2}}, u_i^{k+\frac{1}{2}}\right)_i = \lambda\left(u_i^{k+\frac{1}{2}} - u_i^k u_i^{k+1}\right), \quad 1 \le i \le M-1,$$
$$u_0^{k+\frac{1}{2}} = 0, \quad u_M^{k+\frac{1}{2}} = 0.$$

Now, introduce the notation

$$w_i = u_i^{k+\frac{1}{2}}, \quad 0 \le i \le M.$$

This leads to the following system:

$$\frac{2}{\tau}\left(w_i - u_i^k\right) - \delta_x^2 w_i + \psi\left(w, w\right)_i = \lambda\left(w_i - u_i^k u_i^{k+1}\right), \quad 1 \le i \le M-1, \tag{2.16}$$
$$w_0 = 0, \quad w_M = 0. \tag{2.17}$$

Finally, we express u_i^{k+1} in terms of w_i:

$$u_i^{k+1} = 2w_i - u_i^k, \quad 0 \le i \le M.$$

Remark 2.6. The nonlinear term uu_x in (1.1) can be written as $\frac{1}{3}[uu_x + (u^2)_x]$ and its discrete expression $\psi\left(U^{k+\frac{1}{2}}, U^{k+\frac{1}{2}}\right)_i$ can be construed as the discretization of the term at the point $\left(x_i, t_{k+\frac{1}{2}}\right)$.

3 Results

In this section, we present the discrete conservation property and demonstrate accuracy of the scheme through numerical tests.

3.1 Discrete Conservation

Theorem 3.1. *Let $\{u_i^k | 0 \le i \le M, 0 \le k \le N\}$ be the solution of the difference scheme (2.13)-(2.15). Define*

$$E^k = ||u^k||^2 + 2\tau \sum_{l=0}^{k-1}\left[|u^{l+\frac{1}{2}}|_1^2 + \lambda\left((u^l u^{l+1}, u^{l+\frac{1}{2}}) - ||u^{l+\frac{1}{2}}||^2\right)\right], \quad 0 \le k \le N.$$

Then we have

$$E^k = E^0, \quad 1 \le k \le N.$$

Proof. Taking the inner product of (2.13) with $u^{k+\frac{1}{2}}$ we have

$$\left(\delta_t u^{k+\frac{1}{2}}, u^{k+\frac{1}{2}}\right) - \left(\delta_x^2 u^{k+\frac{1}{2}}, u^{k+\frac{1}{2}}\right) + \left(\psi\left(u^{k+\frac{1}{2}}, u^{k+\frac{1}{2}}\right), u^{k+\frac{1}{2}}\right) - \left(\lambda\left(u_i^{k+\frac{1}{2}} - u_i^k u_i^{k+1}\right), u_i^{k+\frac{1}{2}}\right) = 0.$$

We can show that

$$\left(\delta_t u^{k+\frac{1}{2}}, u^{k+\frac{1}{2}}\right) = \frac{1}{2\tau}\left(||u^{k+1}||^2 - ||u^k||^2\right),$$
$$\left(\delta_x^2 u^{k+\frac{1}{2}}, u^{k+\frac{1}{2}}\right) = -|u^{k+\frac{1}{2}}|_1^2,$$
$$\left(\psi\left(u^{k+\frac{1}{2}}, u^{k+\frac{1}{2}}\right), u^{k+\frac{1}{2}}\right) = 0,$$
$$\left(\lambda\left(u_i^{k+\frac{1}{2}} - u_i^k u_i^{k+1}\right), u_i^{k+\frac{1}{2}}\right) = -\lambda\left(\left(u_i^k u_i^{k+1}, u_i^{k+\frac{1}{2}}\right) - ||u_i^{k+\frac{1}{2}}||^2\right).$$

Thus, we have

$$\frac{1}{2\tau}\left(||u^{k+1}||^2 - ||u^k||^2\right) + |u^{k+\frac{1}{2}}|_1^2 + \lambda\left(\left(u_i^k u_i^{k+1}, u_i^{k+\frac{1}{2}}\right) - ||u_i^{k+\frac{1}{2}}||^2\right) = 0, \quad 0 \le k \le N-1.$$

Replacing k with l and summing from $l = 0$ to $l = k-1$, we have

$$\frac{1}{2\tau}\left(||u^k||^2 - ||u^0||^2\right) + \sum_{l=0}^{k-1}\left[|u^{l+\frac{1}{2}}|_1^2 + \lambda\left(\left(u_i^l u_i^{l+1}, u_i^{l+\frac{1}{2}}\right) - ||u_i^{l+\frac{1}{2}}||^2\right)\right] = 0, \quad 1 \le k \le N.$$

That is

$$||u^k||^2 + 2\tau\sum_{l=0}^{k-1}\left[|u^{l+\frac{1}{2}}|_1^2 + \lambda\left(\left(u_i^l u_i^{l+1}, u_i^{l+\frac{1}{2}}\right) - ||u_i^{l+\frac{1}{2}}||^2\right)\right] = ||u^0||^2, \quad 0 \le k \le N.$$

We have shown that

$$E^k = ||u^k||^2 + 2\tau\sum_{l=0}^{k-1}\left[|u^{l+\frac{1}{2}}|_1^2 + \lambda\left(\left(u_i^l u_i^{l+1}, u_i^{l+\frac{1}{2}}\right) - ||u_i^{l+\frac{1}{2}}||^2\right)\right] = ||u^0||^2 = E^0, \quad 0 \le k \le N,$$

and we can conclude that $E^k = E^0, \quad 1 \le k \le N.$ This completes the proof. □

3.2 Numerical tests

For computational purposes, Newton's method is used to solve the nonlinear difference scheme (2.16). Suppose we already know $\{u_i^k \mid 0 \le i \le M\}$. Let

$$r_1 = \frac{\tau}{2h^2}, \quad r_2 = \frac{\tau}{12h}, \quad r_3 = \frac{\lambda\tau}{2} \quad \text{and} \quad w_i^{(0)} = u_i^k, \quad 0 \le i \le M.$$

The scheme is solved using the following tri-diagonal system of linear algebraic equations:

$$\left[-r_1 - r_2(w_i^{(l)} + 2w_{i-1}^{(l)})\right]\delta_{i-1}^{(l)} + \left[1 + 2r_1 + r_2(w_{i+1}^{(l)} - w_{i-1}^{(l)}) - r_3(1 - u_i^k)\right]\delta_i^{(l)}$$
$$+ \left[-r_1 + r_2(w_i^{(l)} + 2w_{i+1}^{(l)})\right]\delta_{i+1}^{(l)} = -\left[w_i^{(l)} - u_i^k - r_1(w_{i+1}^{(l)} - 2w_i^{(l)} + w_{i-1}^{(l)})\right]$$
$$- \left[r_2(w_i^{(l)}w_{i+1}^{(l)} - w_i^{(l)}w_{i-1}^{(l)} + (w_{i+1}^{(l)})^2 - (w_{i-1}^{(l)})^2) - r_3(w_i^{(l)} - 2u_i^k w_i^{(l)} + (u_i^k)^2)\right],$$

$$1 \le i \le M-1, \quad \delta_0^{(l)} = 0, \quad \delta_M^{(l)} = 0.$$

The solution at the next time step is updated by

$$w_i^{(l+1)} = w_i^{(l)} + \delta_i^{(l)},$$

until

$$\left\| \delta^{(l)} \right\|_\infty \leq \varepsilon \quad l = 0, 1, \ldots.$$

Then we have

$$u_i^{k+1} = 2w_i^{(l+1)} - u_i^k, \quad 0 \leq i \leq M.$$

In our tests, we take $\varepsilon = 1 \times 10^{-16}$.

We compute the numerical accuracy in time by varying τ with fixed h, and in space by varying h with fixed τ. The corresponding errors and convergence orders are defined as follows:

$$E_\infty(h,\tau) = \max_{0\leq k\leq N} \max_{0\leq i\leq M} |U_i^k(h,\tau) - u_i^k(h,\tau)|, \quad Ord_h = \log_2 \frac{E_\infty(2h,\tau)}{E_\infty(h,\tau)}, \quad Ord_\tau = \log_2 \frac{E_\infty(h,2\tau)}{E_\infty(h,\tau)},$$

$$G_\infty(h,\tau) = \max_{0\leq k\leq N} \max_{0\leq i\leq M} \left| u_i^k(h,\tau) - u_i^{2k}\left(\frac{h}{2},\tau\right) \right|, \quad Ord_h = \log_2 \frac{G_\infty(2h,\tau)}{G_\infty(h,\tau)},$$

$$F_\infty(h,\tau) = \max_{0\leq k\leq N} \max_{0\leq i\leq M} \left| u_i^k(h,\tau) - u_i^{2k}\left(h,\frac{\tau}{2}\right) \right|, \quad Ord_\tau = \log_2 \frac{F_\infty(h,2\tau)}{F_\infty(h,\tau)}.$$

3.2.1 Extension to the nonhomogeneous equation

Theorem 3.1 focuses on the numerical analysis of a nonlinear implicit scheme for the homogeneous Burgers-Fisher equation. To extend this approach to the nonhomogeneous case, it is necessary to formulate an extended difference scheme. Therefore, we conduct an accuracy test using a representative example involving a nonhomogeneous equation. For the nonhomogeneous Burgers-Fisher equation:

$$u_t - u_{xx} + uu_x = \lambda u(1-u) + f(x,t), \quad 0 < x < L, \quad 0 < t \leq T, \tag{3.1}$$

$$u(x,0) = \varphi(x), \quad 0 \leq x \leq L, \tag{3.2}$$

$$u(0,t) = u(L,t) = 0, \quad 0 < t \leq T, \tag{3.3}$$

we use the following difference scheme

$$\frac{2}{\tau}\left(w_i - u_i^k\right) - \delta_x^2 w_i + \psi\left(w,w\right)_i = \lambda\left(w_i - u_i^k u_i^{k+1}\right) + f\left(x_i, t_{k+\frac{1}{2}}\right), \quad 1 \leq i \leq M-1, \tag{3.4}$$

$$w_0 = 0, \quad w_M = 0. \tag{3.5}$$

We can show that the difference scheme (3.4)-(3.5) is accurate to the second order in both time and space by using concept similar to Theorem 3.1, as shown in Example 3.2.

Example 3.2. An accuracy test for the nonhomogeneous case is conducted using the nonhomogeneous example with the exact solution given by

$$u(x,t) = e^{-t}\sin(\pi x).$$

For the problem (3.1)-(3.3), take $L = 1,\ T = 1,\ \lambda = 1,\ \varphi(x) = \sin(\pi x)$. An additional term

$$f(x,t) = (\pi^2 - 2)e^{-t}\sin(\pi x) + e^{-2t}\left[\sin^2(\pi x) + \pi\sin(\pi x)\cos(\pi x)\right]$$

is included on the right-hand side of (1.1).

The schemes (3.4)-(3.5) are run to check the accuracy in space and time, respectively. The numerical errors $E_\infty(h,\tau)$ and convergence orders in Table 1 confirm that both schemes have second-order accuracy. To check the discrete conservation properties, the absolute energy error $|E^k - E^0|$ is calculated at times $T = 3, 4, 5$, as shown in Table 2. The results consistently show small energy error, confirming that our scheme follows Theorem 3.1.

$h = 1/1,000$			$\tau = 1/1,000$		
τ	$E_\infty(h,\tau)$	Ord_τ	h	$E_\infty(h,\tau)$	Ord_h
1/80	$4.014435e-04$		1/80	$3.665058e-05$	
1/160	$1.006247e-04$	2.00	1/160	$9.162977e-06$	2.00
1/320	$2.505359e-05$	2.01	1/320	$2.290810e-06$	2.00
1/640	$6.257408e-06$	2.00	1/640	$5.727028e-07$	2.00
1/1280	$1.564328e-06$	2.00	1/1280	$1.431758e-07$	2.00

Table 1: Numerical errors and convergence orders in time and space for Example 3.2.

k	Energy Error ($T = 3$)	Energy Error ($T = 4$)	Energy Error ($T = 5$)
110	3.33139908e-10	3.40439482e-13	3.33066907e-16
120	3.74730490e-10	3.61641272e-13	3.43475248e-16
130	3.79922847e-10	3.62963132e-13	3.46944695e-16
140	3.80571082e-10	3.63042929e-13	3.46944695e-16
150	3.80652010e-10	3.63049868e-13	3.46944695e-16
160	3.80662116e-10	3.63049868e-13	3.46944695e-16
170	3.80663376e-10	3.63049868e-13	3.46944695e-16
180	3.80663535e-10	3.63049868e-13	3.46944695e-16
190	3.80663553e-10	3.63049868e-13	3.46944695e-16
200	3.80663556e-10	3.63049868e-13	3.46944695e-16
210	3.80663556e-10	3.63049868e-13	3.46944695e-16
220	3.80663556e-10	3.63049868e-13	3.46944695e-16

Table 2: Energy errors at different time steps $|E^k - E(0)|$ for Example 3.2.

Example 3.3. This example is different from the previous one because it is a homogeneous equation. The main purpose of this example is to test the difference scheme (2.16)-(2.17) with an unknown exact solution. Specifically, we consider the problem (1.1)-(1.3) with the initial condition

$$\varphi = x^3(1-x)^3.$$

We use $\lambda = 1$, and the computational domain is $[0,1] \times [0,1]$.

The numerical solutions at various time levels are shown in Figure 1. The second order of accuracy in time is confirmed in Table 3, and the accuracy in space is verified in Table 4. To examine the conservation properties of (2.16)-(2.17), the absolute energy errors $|E^k - E(0)|$ for different times $T = 1, \ldots, 5$ are detailed in Table 5. The small difference between E^k and E^0 supports the theoretical discrete conservation in Theorem 3.1.

$h = 1/1,000$			$h = 1/10,000$		
τ	$F_\infty(h,\tau)$	Ord_τ	τ	$F_\infty(h,\tau)$	Ord_τ
1/80	$1.804218e-04$		1/80	$1.804283e-04$	
1/160	$5.056159e-05$	1.84	1/160	$5.056553e-05$	1.84
1/320	$1.337721e-05$	1.92	1/320	$1.337855e-05$	1.92
1/640	$3.441787e-06$	1.96	1/640	$3.442617e-06$	1.96
1/1280	$8.732061e-07$	1.98	1/1280	$8.735452e-07$	1.98

Table 3: Numerical errors and convergence orders in time for example 3.3.

$\tau = 1/1,000$			$\tau = 1/5,000$		
h	$G_\infty(h,\tau)$	Ord_h	h	$G_\infty(h,\tau)$	Ord_h
1/40	$4.624435e-06$		1/40	$4.624586e-06$	
1/80	$1.156042e-06$	2.00	1/80	$1.156080e-06$	2.00
1/160	$2.890058e-07$	2.00	1/160	$2.890154e-07$	2.00
1/320	$7.225116e-08$	2.00	1/320	$7.225355e-08$	2.00

Table 4: Numerical errors and convergence orders in space for example 3.3.

4 Discussion

This case study presents a numerical solution for the nonlinear Burgers-Fisher equation using a two-level implicit finite difference method. The results show that the proposed scheme preserves numerical conservation properties, which are both theoretically proven and supported by numerical experiments. In addition, numerical experiments confirm that the method achieves second-order accuracy in both space and time. Its compact formulation requires fewer neighboring grid points, making it efficient for solving nonlinear equations. Overall, the results show that the proposed method can accurately solve the nonlinear Burgers-Fisher equation while maintaining its conservation properties.

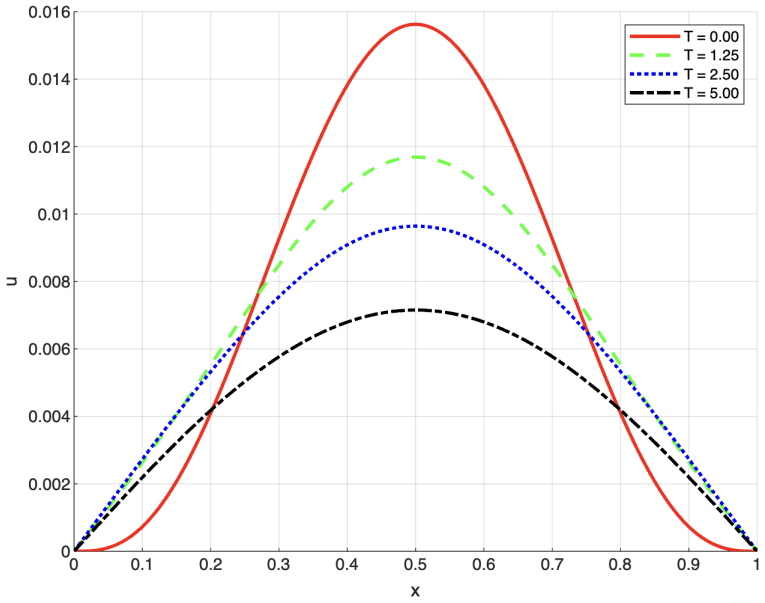


Figure 1: The numerical solution at difference time levels with $h = 1 \cdot 2^{-7}$, $\tau = 5 \cdot 2^{-7}$ of Example 3.3.

k	Energy Error ($T = 1$)	Energy Error ($T = 2$)	Energy Error ($T = 3$)	Energy Error ($T = 4$)	Energy Error ($T = 5$)
130	5.7984e-08	5.7842e-08	5.7605e-08	5.7275e-08	5.6851e-08
140	5.7984e-08	5.7842e-08	5.7605e-08	5.7275e-08	5.6851e-08
150	5.7984e-08	5.7842e-08	5.7605e-08	5.7275e-08	5.6851e-08
160	5.7984e-08	5.7842e-08	5.7605e-08	5.7275e-08	5.6851e-08
170	5.7984e-08	5.7842e-08	5.7605e-08	5.7275e-08	5.6851e-08
180	5.7984e-08	5.7842e-08	5.7605e-08	5.7275e-08	5.6851e-08
190	5.7984e-08	5.7842e-08	5.7605e-08	5.7275e-08	5.6851e-08
200	5.7984e-08	5.7842e-08	5.7605e-08	5.7275e-08	5.6851e-08

Table 5: Energy errors at different time steps $|E^k - E(0)|$ by $h = 1/1024$, $\tau = 1/256$, for Example 3.3.

Acknowledgment. This work was partially supported by the Department of Mathematics, Faculty of Science, Chiang Mai University, Thailand. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments.

References

[1] R. C. Mittal, R. Jiwari, A differential quadrature method for numerical solutions of Burgers'-type equations, Int. J. Numer. Methods Heat Fluid Flow, 22 (7) (2012) 880–895.

[2] A. M. Wazwaz, Partial Differential Equations and Solitary Waves Theory, Springer, 2009.

[3] G. Gao, B. Ge, Y. Chen, Derivation and error analysis of three-level linearized difference schemes for solving the Burgers-Fisher equation, J. Differ. Equ. Appl., 31 (1) (2024) 115–153.

[4] L. Zhang, L. Wang, X. Ding, Exact finite difference scheme and nonstandard finite difference scheme for Burgers and Burgers-Fisher equations, J. Appl. Math., 2014 (2014) 597926.

[5] V. Chandraker, A. Awasthi, S. Jayaraj, Numerical treatment of Burger-Fisher equation, Procedia Technology, 26 (2016) 1217–1225.

[6] H. Bateman, Some recent researches on the motion of fluids, Mon. Weather Rev., 43 (4) (1915) 163–170.

[7] G. B. Whitham, Linear and Nonlinear Waves (Vol. 42), John Wiley & Sons, 1974.

[8] J. M. Burgers, A Mathematical Model Illustrating the Theory of Turbulence, Adv. Appl. Mech., 1 (1948) 171–199.

[9] R. A. Fisher, The wave of advance of advantageous genes, Annals of Eugenics, 7 (4) (1937) 353–369.

[10] A. Kolmogorov, I. Petrovskii, N. Piskunov, Study of a Diffusion Equation That Is Related to the Growth of a Quality of Matter and Its Application to a Biological Problem, Moscow University Mathematics Bulletin, 1 (1937) 1-26.

[11] H. A. Levine, L. E. Payne, P. E. Sacks, B. Straughan, Analysis of a convective reaction-diffusion equation II, SIAM J. Math. Anal., 20 (1989) 133–147.

[12] A. M. Wazwaz, Analytic study on Burgers, Fisher, Huxley equations and combined forms of these equations, Appl. Math. Comput., 195 (2008) 754–761.

[13] H. Zhang, Y. Xia, P. R. N'gbo, Global existence and uniqueness of a periodic wave solution of the generalized Burgers–Fisher equation, Appl. Math. Lett., 121 (2021) 107353.

[14] Y. Zhou, Q. Liu, W. Zhang, Bounded traveling waves of the generalized Burgers–Fisher equation, Int. J. Bifurcation Chaos, 23 (3) (2013) 1350054.

[15] D. A. Hammad, M. S. El-Azab, 2N order compact finite difference scheme with collocation method for solving the generalized Burger's-Huxley and Burger's-Fisher equations, Appl. Math. Comput., 258 (2015) 296–311.

[16] H. Hajinezhad, A numerical approximation for the one-dimensional Burger–Fisher equation, Asian Research Journal of Mathematics, 18 (5) (2022) 22–30.

[17] A. C. Loyinmi, M. D. Sanyaolu, S. Gbodogbe, Exploring the efficacy of the weighted average method for solving nonlinear partial differential equations: A study on the Burger-Fisher equation, EDUCATUM Journal of Science, Mathematics and Technology, 12 (1) (2024) 60–79.

[18] O. P. Yadav, R. Jiwari, Finite element analysis and approximation of Burgers'-Fisher equation, Numer. Methods Partial Differential Equations, 33 (5) (2017) 1652–1677.

[19] K. W. Morton, D. F. Mayers, Numerical Solution of Partial Differential Equations: 2nd Edition, Cambridge University Press, 2005.

[20] L. C. Evans, Partial Differential Equations (2nd ed.), American Mathematical Society, 2010.

[21] Q. Zhang, X. Wang, Z. Sun, The pointwise estimates of a conservative difference scheme for Burgers' equation, Numer. Methods Partial Differential Equations, 36 (2020) 1611–1628.

[22] Z. Sun, Q. Zhang, G. Gao, Finite Difference Methods for Nonlinear Evolution Equations, De Gruyter, 2023.

[23] Q. Zhang, L. Liu, Convergence and stability in maximum norms of linearized fourth-order conservative compact scheme for Benjamin–Bona–Mahony–Burgers' equation, J. Sci. Comput., 87 (2021) 1–31.

[24] X. Wang, Analytical and numerical studies of the modified Kawahara equation with dual-power law nonlinearities, Numerical Algorithms, (2024) 1–33.

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

Determining Characteristics of Solution for Vehicle Routing Problem with Time Windows

Teerapong Chaochanglek[1,*] and Boonyarit Intiyot[1,†]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

The Vehicle Routing Problem with Time Windows (VRPTW) has been widely studied, with heuristic methods remaining the most effective approach for finding solutions. These heuristics are typically based on intuitive rules formulated by researchers, and their efficiency is evaluated only by empirical data. However, such heuristics do not always guarantee optimal solutions. A more systematic approach would involve identifying key characteristics of high-quality VRPTW solutions to guide heuristic development more effectively. Arnold and Sörensen previously introduced a statistical learning model to identify distinguishing characteristics of near-optimal solutions for the Vehicle Routing Problem (VRP). Their insights were used to design a guided local search heuristic, which yielded promising results. Building upon this concept, we extend their methodology to VRPTW by defining new solution metrics that incorporate time-window constraints. Using a machine learning model, we analyzed more than 8000 VRPTW instances to determine the most distinctive characteristics of high-quality solutions. Our findings indicate that the average number of intersections per customer serves as the most significant distinguishing characteristic, differing from the characteristics identified in Arnold and Sörensen's work. This work provides a deeper understanding of the solution structures in VRPTW and offers valuable guidance for heuristic development in this domain.

Keywords: Vehicle routing problem with time windows, Heuristic method, Machine learning
2020 MSC: 90B06, 90C27, 90C59

1 Introduction

The Vehicle Routing Problem with Time Windows (VRPTW) is a prominent extension of both the classic Traveling Salesman Problem (TSP) and the Vehicle Routing Problem (VRP). While traditional VRP focuses solely on minimizing transportation costs (often linked directly to travel distance or fuel consumption), VRPTW introduces an additional layer of complexity by incorporating strict time constraints. Each vehicle must service a set of customers not only while respecting vehicle capacity and minimizing cost, but also within predefined time windows. This consideration reflects real-world logistics scenarios where customers are only available at specific times, such as store opening hours or delivery slots.

*Speaker (6570064923@student.chula.ac.th)
†Corresponding author (boonyarit.i@chula.ac.th)

When time windows are ignored, the problem simplifies to VRP. However, the time-window constraints significantly increase the complexity, rendering VRPTW a well-known NP-hard problem [1]. As a result, even obtaining feasible solutions for large instances is computationally demanding. Over the years, a variety of approaches have been proposed for solving VRPTW. Exact methods, such as branch-and-price [2] and set partitioning models [3], provide optimal solutions but are limited to small or medium-sized instances due to their exponential computational time. Consequently, heuristics and metaheuristics—such as savings algorithms [1], tabu search [4], genetic algorithms [5], and ant colony optimization [6]—have become widely adopted for large-scale problems, offering high-quality solutions within reasonable time limits.

Given the intractability of exact methods for large VRPTW instances, heuristics remain the most effective approach. Heuristics are rule-based strategies that seek good (though not necessarily optimal) solutions and are often refined through empirical evaluation. However, understanding the structure and characteristics of near-optimal solutions can significantly guide these heuristics toward better results.

In this context, Arnold and Sörensen [7] proposed a statistical learning framework to identify solution features associated with near-optimal solutions for the VRP. By learning from these features, they enhanced local search heuristics to yield improved solution quality. Inspired by their methodology, this study aims to extend their approach to the more complex VRPTW by analyzing the unique features of solutions that satisfy time-window constraints. These solution metrics are used as features in a machine learning model to classify VRPTW solutions into high- and low-quality categories. To enable this, a large number of VRPTW instances and corresponding solution labels were generated and used to train and evaluate classification models.

To provide a formal foundation, the mathematical model of VRPTW [8] is presented below. Let $N = \{0, 1, \ldots, n+1\}$ be a set of nodes that includes the depot (nodes 0 and $n+1$) and n customers. Each edge (i, j) has an associated travel cost c_{ij} and travel time t_{ij}. Each customer $i \in C = \{1, \ldots, n\}$ has a demand d_i, a time window $[a_i, b_i]$, and each vehicle has capacity q. The decision variables include $x_{ijk} \in \{0, 1\}$ to indicate if vehicle k travels from i to j where $i \neq j, i \neq n+1$, and $j \neq 0$, and s_{ik} denoting the service start time of vehicle k at customer i.

The objective is to minimize the total transportation cost:

$$\min \sum_{k \in V} \sum_{i \in N} \sum_{j \in N} c_{ij} x_{ijk} \tag{1.1}$$

Subject to the following constraints:

$$\sum_{k \in V} \sum_{j \in N} x_{ijk} = 1 \qquad \forall i \in C \quad \text{(each customer is visited exactly once)} \tag{1.2}$$

$$\sum_{i \in C} d_i \sum_{j \in N} x_{ijk} \leq q \qquad \forall k \in V \quad \text{(vehicle capacity)} \tag{1.3}$$

$$\sum_{j \in N} x_{0jk} = 1 \qquad \forall k \in V \quad \text{(departure from depot)} \tag{1.4}$$

$$\sum_{i \in N} x_{ihk} - \sum_{j \in N} x_{hjk} = 0 \qquad \forall h \in C, \forall k \in V \quad \text{(flow conservation)} \tag{1.5}$$

$$\sum_{i \in N} x_{i,n+1,k} = 1 \qquad \forall k \in V \quad \text{(return to depot)} \tag{1.6}$$

$$x_{ijk}(s_{ik} + t_{ij} - s_{jk}) \leq 0 \qquad \forall i, j \in N, \forall k \in V \quad \text{(time continuity)} \tag{1.7}$$

$$a_i \leq s_{ik} \leq b_i \qquad \forall i \in N, \forall k \in V \quad \text{(time window respect)} \tag{1.8}$$

$$x_{ijk} \in \{0,1\} \qquad \forall i,j \in N, \forall k \in V \tag{1.9}$$

This model ensures that all customers are served once, within their time windows, while minimizing overall cost and respecting vehicle capacity constraints.

The remainder of this paper is structured as follows. In Section 2 we describe how to define and generate instances for VRPTW. Then, we used the data as input to a classification learner. We present the results and discuss in Section 3. In Section 4 we conclude with a summary of the results and give some future work possibilities.

2 From Problem Definition to Data Generation

Our methodology is to identify the key characteristics of VRPTW solutions. We begin by generating VRPTW instances, followed by applying heuristics to compute both near-optimal and non-optimal solutions. Next, we use these solution outcomes to calculate the instance and solution metrics. Finally, these metrics are incorporated as features in a statistical learning model to identify the most distinctive characteristics of VRPTW solutions.

2.1 Instance Generation

CVRPLIB [9] collects many VRPTW instances with the best known solution values that can be varied according to the following attributes: (1) the positioning of the depot and the customer, including the number of customers, (2) the distribution of demand, (3) the time windows of the depot and customers, and (4) the service time. So, we tried to construct a representative set of instances by randomizing the choice of attributes for a particular instance with reasonable bounds. By the definition of these bounds, we derived 4 different instance classes.

We define a set of small instances with 20-50 customers and a set of large instances with 70-100 customers. The customers are randomly located on the first quadrant of a two-dimensional space on the 100×100 grid. We also define a class for the position of the depot. The depot is either located at the center of the grid at position $(50, 50)$, or at its edge which we set at the original point $(0, 0)$. Then, we randomly define vehicle capacity so that we have between 3 and 6 routes in the small classes and 6-10 routes in the large class. Those routes and size of the instance classes are followed by the generated VRP instances of Arnold and Sörensen [7] to obtain some comparable results. Finally, the distribution of the demand is varied in a way that is reasonable with the capacity and the number of vehicles. So, we generate the data into four instance classes as presented in Table 1.

Table 1: Instance parameters for the different instance classes in the experiments.

Class	Number of Customers	Number of Instances	Depot position	Route
1	20-50	2000	Center	3-6
2	20-50	2000	Edge	3-6
3	70-100	2000	Center	6-10
4	70-100	2000	Edge	6-10

2.2 Computation of Solutions

In this section, our objective is to determine the characteristics that distinguish an optimal or near-optimal solution from a worse one. Understanding these factors is crucial for efficiently identifying better solutions, which is the primary objective of any heuristic. Therefore, analyzing and comparing solutions of varying quality is essential.

Following the work of Arnold and Sörensen, near-optimal solutions are obtained using heuristic methods designed to efficiently find high-quality solutions. Although these solutions are not guaranteed to be the absolute best, they are very high in quality. In contrast, non-optimal solutions are deliberately generated to be of lower quality than near-optimal ones. For each instance, these solutions are specifically constructed to be worse than their corresponding near-optimal counterparts. The rationale behind generating non-optimal solutions is to enable a comparative analysis of different solution qualities, helping to identify key characteristics that differentiate high-quality from lower-quality solutions.

In our approach to solve the VRPTW instances, we use Google OR-Tools [10, 11], which follows a two-step process: first, constructing an initial solution and then refining it through local search optimization.

For the non-optimal solution, we generated the initial routes using Sequential Cheapest Insertion (SCI) [1]. This greedy heuristic sequentially adds customers to the least-costly feasible position, often resulting in suboptimal solutions due to its limited foresight. To refine this solution, we applied Simulated Annealing (SA) [12, 13], an optimization technique inspired by the annealing process in metallurgy. SA allows for occasional acceptance of worse solutions to escape local optima and explore alternative routing configurations, though it lacks a strong mechanism to guide improvements effectively.

For the near-optimal solution, we employed Parallel Cheapest Insertion (PCI) [14] to construct the initial routes. Unlike SCI, PCI simultaneously considers multiple insertion possibilities, promoting better load distribution and more cost-efficient solutions. This solution is further enhanced using Guided Local Search (GLS) [15], which dynamically penalizes inefficient routing choices. By focusing the search on improving weak areas of the solution, GLS guides the optimization process toward higher-quality routes.

For each VRPTW instance, we generated and labeled two solutions: SCI with SA, representing the non-optimal category, and PCI with GLS, representing the near-optimal category. This structured comparison helps us analyze how different heuristics influence solution quality, offering insights into the trade-offs between computational efficiency and route optimization.

2.3 Solution Metrics

The notations that Arnold and Sörensen [7] define to construct the solution characteristics or the solution metrics are as follows. Let $G = (V, E)$ be a graph representing VRP where $V = \{0, 1, 2, 3, ..., N\}$ is a set of nodes (customers) with 0 being the depot and E is a set of edges. Let R denote the set of all routes r of a VRP solution where $r = \{n_1^r, n_2^r, ..., n_{N_r}^r\} \in R, n_i^r$ is the ith customer in route r for $i = 1, 2, ..., N_r$ that are visited on this route, starting and ending at the depot D, and N_r is the number of customers in the route r. Thus, a route r has the edges $\{(D, n_1^r), (n_1^r, n_2^r), ..., (n_{N_r}^r, D)\}$. The Euclidean distance between node n_1 and n_2 is denoted by $d(n_1, n_2)$, and $x(n_1)$ and $y(n_1)$ define the x and $y-$coordinate of the node n_1, respectively. Let G_r be the center of gravity of the route r where $x-$coordinate of G_r is defined

as $x(G_r) = \dfrac{\sum x(n_i^r) + x(D)}{N_r + 1}$ and the $y-$coordinate is computed accordingly. Let L_{G_r} be the line which passes through the depot and G_r. Then $dis(L_{G_r}, n_1)$ denotes the distance between this line and a customer n_1. The distance is positive if the customer is on the right side of the line and negative otherwise. Last but not least, an angle between node n_1 and n_2 with respect to the depot is denoted by $rad(n_1, n_2)$. Finally, let $I(r_1, r_2)$ be the number of intersections between routes r_1 and r_2. The solution metrics S1-S10 of the VRP solution proposed by Arnold and Sörensen [7] are shown as the following:

S1-Average number of intersections per customer

$$\frac{\sum_{i=1}^{|R|-1} \sum_{j=i+1}^{|R|} I(r_i, r_j)}{N} \tag{2.1}$$

S2-Longest distance between two connected customers per route

$$\frac{\sum_{r \in R} \max_{i \in \{1,\ldots,N_r-1\}} d(n_i^r, n_{i+1}^r)}{|R|} \tag{2.2}$$

S3-Average distance between depot to a directly-connected node

$$\frac{\sum_{r \in R} \left(d(D, n_1^r) + d(n_{N_r}^r, D)\right)}{2|R|} \tag{2.3}$$

S4-Average distance between routes (their centers of gravity)

$$\frac{\sum_{r_i \in R} \sum_{r_j \in R \setminus r_i} d(G_{r_i}, G_{r_j})}{|R| \cdot (|R| - 1)} \tag{2.4}$$

S5-Average width per route

$$\frac{\sum_{r \in R} \left(\max_{i \in \{1,\ldots,N_r\}} dis(L_{G_r}, n_i^r) - \min_{i \in \{1,\ldots,N_r\}} dis(L_{G_r}, n_i^r) \right)}{|R|} \tag{2.5}$$

S6-Average span in radian per route

$$\frac{\sum_{r \in R} \max_{i \in \{1,\ldots,N_r\}} rad(n_i^r, n_j^r)}{|R|} \tag{2.6}$$

S7-Average compactness per route, measured by width

$$\frac{\sum_{r \in R} \sum_{i=1}^{N_r} |dis(L_{G_r}, n_i^r)|}{N} \tag{2.7}$$

S8-Average compactness per route, measured by radian

$$\frac{\sum_{r\in R}\sum_{i=1}^{N_r} rad(G_r, n_i^r)}{N} \tag{2.8}$$

S9-Average depth per route

$$\frac{\sum_{r\in R}\max_{i\in\{1,\ldots,N_r\}} d(n_i^r, D)}{|R|} \tag{2.9}$$

S10-Standard deviation of the number of customers per route

$$\sqrt{\frac{\sum_{r\in R}\left(N_r - \frac{N}{|R|}\right)^2}{|R|}} \tag{2.10}$$

S11-Average total service time per route (This is an additional metric that we defined based on the service time of customers.)

$$\frac{\sum_{r\in R}\sum_{i=1}^{N_r} s_i^r}{|R|} \tag{2.11}$$

where s_i^r is the service time for customer i.

The first two metrics S1 and S2 align with well-established findings in operations research, particularly in relation to the well-known Traveling Salesman Problem. Generally, high-quality solutions tend to feature fewer intersections and shorter with less extreme edge lengths. In contrast, the impact of other metrics is less well-defined. Solution metric S3 is more exploratory, and they hypothesize that in high-quality solutions, the edges connecting the depot to customers should be relatively short. Ideally, routes should be distinctly separated (S4) and should not overlap significantly, either in width (S5) or depth (S9). As a result, they anticipate that good solutions will have narrower and less deep routes on average, ensuring a relatively greater distance between them. The width of a route can also be analyzed by examining the angular distribution of customers relative to the depot (S6). Routes that are more compact tend to have smaller variations in customer angles, a principle commonly used in the sweep heuristic. Additionally, routes can be either more linear (high compactness) or more circular (low compactness), which they quantify using S7. It is important to note that compactness and width are closely related, as more compact routes tend to be narrower. Compactness can also be assessed by measuring the variation in customer angles along a route (S8). Solution metric S10 evaluates how evenly balanced the routes are in terms of the number of served customers. Finally, S11 consider how impact about the service time for each route.

The number of routes is typically a key aspect of VRPTW solutions. However, for simplicity, we assume that the number of routes is predetermined and remains constant for both the near-optimal and non-optimal solutions. This assumption allows us to focus on analyzing the structural impact of different solutions on solution quality, as varying the number of routes could significantly alter the overall structure of the route.

2.4 Instance Metrics

In addition, Arnold and Sörensen also defined features called instance metrics, which reflect the characteristics of instances. Each instance metric is defined as follows.

I1-Number of customers

I2-Number of routes

I3-Degree of capacity utilization

I4-Average distance between each pair of customers

I5-Standard deviation of the pair-wise distance between customers

I6-Average distance from customers to the depot

I7-Standard deviation of the distance from customers to the depot

I8-Standard deviation of the radians from customers towards the depot

3 Preliminaries Results and Discussion

The dataset consists of small and large VRPTW instances, each generated with 2,000 instances per class. For every instance, two solutions are to be obtained: a non-optimal solution and a near-optimal solution. As a result, each class should have contained 4,000 solutions (or datapoints). However, our solution method did not yield both a non-optimal and a near-optimal solution for some instances due to undefined outcomes in either case. These instances are excluded from the experiment, and the remaining number of usable datapoints per class is summarized in Table 2.

Table 2: Number of usable datapoints

Class	Number of Generated Datapoints	Number of Usable Datapoints
1	4000	3930
2	4000	3884
3	4000	3712
4	4000	3964

Initially, we evaluated the performance of the features by inputting them into a support vector machine (SVM), a classification model. SVMs are well-suited for high-dimensional datasets and are known for their robustness, particularly when dealing with both linear and non-linear decision boundaries. They perform well even when the number of features is large compared to the number of samples, which aligns well with our setting involving numerous instance and solution metrics (I1–I8 and S1–S11). The average prediction accuracies obtained from 5-fold cross-validation are presented in Table 3. All reported accuracies exceed 78%, some reaching above 85%, indicating that the classifier effectively captures patterns from the solution and instance metrics. In machine learning applications that involve real-world, noisy, or multiclass

data, prediction accuracies above 70% are generally regarded acceptable. Therefore, the results obtained suggest that the features used carry substantial predictive power to differentiate between classes of instances.

Table 3: Average prediction accuracies of usable datapoints from 5-fold cross validation using all features (I1-I8 and S1-S11)

Class	Number of Usable Datapoints	Average Accuracy of Prediction (%)
1	3930	78.68
2	3884	82.00
3	3712	85.63
4	3964	79.62

Next, we assessed the predictive capability of individual solution metrics by performing a classification analysis using a decision tree model. Decision trees were selected for their interpretability and ability to reveal the contribution of individual features to prediction outcomes. For each experiment, one solution metric was combined with the instance metrics (I1–I8) as input features following Arnold and Sörensen's experiments. A decision tree classifier was trained on this feature set, and 5-fold cross-validation was employed to obtain stable mean accuracy estimates. Afterward, the model was trained on the training set and evaluated on the separate test set to report the accuracy of the test. The prediction accuracy for each solution metric where the accuracy exceeded 55% with the same range of Arnold and Sörensen are presented in Table 4. Although an accuracy threshold of 55% may appear modest in isolation, it is considered informative in the context of single-feature evaluation in a multi-class problem. In such cases, even small gains over random guessing indicate a meaningful pattern captured by the metric.

Table 4: Solution metrics with an individual prediction accuracy of greater than 55% per instance class. The highest prediction accuracies per instance class is highlighted in bold.

Class	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
1	57	**63**	61	61	59	58	58				
2	61	56	56	55	**67**	63	63				
3	71	**72**	65	66	67	60	65				
4	63		59	56	**66**	60	57				
Avg.	63	63.67	60.25	59.5	64.75	60.25	60.75				

From Table 4, the results indicate that S1, S2, and S5 consistently outperform the others in all instance classes, achieving average precisions of 63%, 63.67%, and 64.75%, respectively. In particular, S2 demonstrates the highest prediction accuracy in class 1 (63%) and class 3 (72%), while S5 achieves the highest accuracy in Class 2 (67%) and Class 4 (66%). However, the average highest prediction accuracy is S5 (average width per route), which is the same as in the Arnold and Sörensen's findings. Other metrics, such as S3, S4, S6, and S7, show moderate predictive power. In contrast, S8–S10 demonstrate limited predictive capability, with no exceeding 55%. In terms of S11, which is the metric we constructed, it does not give the prediction accuracy as our

expectation. So, S8–S11 suggest that these metrics contribute less to distinguishing near-optimal solutions. The results also highlight that instances with more routes (6–10) and central depots produce stronger prediction accuracy, indicating a clearer structural pattern in these solutions.

The findings suggest that S5 and S2 are the most critical metrics in identifying near-optimal VRPTW solutions. The superior predictive precision observed in Class 3 (70–100 customers, center depot, 6–10 routes) implies that larger problem instances tend to exhibit more structured and distinguishable solution patterns. On the other hand, instances with edge depots and fewer routes show a lower prediction accuracy, indicating greater solution variability. These insights can be leveraged to develop heuristics and machine learning models that prioritize S5 and S2, potentially leading to more effective and efficient solution methods for VRPTW.

However, the results we obtained were similar to those reported by Arnold and Sörensen, as well as those by Inbunsong and Intiyot [16], who identified key characteristics of a good solution to the non-Euclidean capacitated vehicle routing problem (CVRP). Our results also highlight the average width per route (S5) as the most distinctive solution metric which is similar to their findings. In contrast, while Arnold and Sörensen observed that the longest distance between two connected customers per route (S2) had low predictive power, our study found that both S1 and S2 metrics serve as effective indicators for VRPTW. Consistent with previous studies, the S6 and S7 metrics continued to demonstrate strong predictive capabilities. On the other hand, the S8 metric exhibited relatively low prediction accuracy.

Obviously, the discrepancies in the results can be attributed to the influence of the instances. Especially, we consider VRP with time-window constraint while Arnold and Sörensen consider only VRP. Various factors, such as the number of instances, vehicle capacity, number of vehicles, and demand distribution, likely contribute to these differences. In particular, we did not categorize the demand into distinct classes, which may also explain some of the observed variations.

4 Conclusion and Future work

In this paper, we propose a solution metric related to the duration of the service. Then, we adapt Arnold and Sörensen's works to get results that can describe what makes a VRPTW solution good. By comparing the results with the Arnold and Sörensen finding, the results are similar to the VRP. The most influent characteristic is S5 (average width per route). In addition, the average number of intersections per customer (S1) and the longest distance between two connected customers per route (S2) demonstrate strong predictive performance. Those findings hold across a wide range of instance types, indicating that the structural differences between near-optimal and non-optimal solutions are consistent across diverse problem settings. This suggests that such structural characteristics can be effectively leveraged to enhance heuristic performance by guiding the search process more intelligently. The degree to which problem-specific knowledge can be incorporated into a heuristic may depend on the design of the method itself. In our case, integrating this knowledge was straightforward—we modified the evaluation function to favor solutions with fewer widths, which led to improved results. More generally, it appears advantageous to design heuristics around known structural properties of high-quality solutions, rather than attempting to adapt such knowledge to heuristics not originally intended for the specific problem. Designing operators and evaluation strategies that exploit problem-specific insights can therefore lead to more efficient and effective search processes. However, the results cannot be directly compared between VRP (or CVRP) and VRPTW due to their

differing constraints.

For the future work, having several classes may give interesting solutions, such as the class of demand, the class of vehicle capacity, and the class of service time. The number of datapoints is also an important key for finding the distinctive the solution metrics. However, the number of datapoints presented here is quite limited because it took a long computational run-time to solve them. Nevertheless, more datapoints will be added in our future work. Moreover, other heuristic methods can be used to obtain a better datapoints. In addition, we suggest constructing more solution metrics related to VRPTW specifically to obtain more interesting results.

Acknowledgment. The authors would like to express their sincere appreciation to the referees for their thorough review and valuable suggestions that have greatly improved the quality of this manuscript. Special thanks are extended to Piyaboot Inbunsong for providing insightful advice throughout the development of this work. The computational experiments were conducted using Google Colab [17]. The authors also gratefully acknowledge the support from Chulalongkorn University for providing essential facilities, as well as financial support from the Development and Promotion of Science and Technology Talents Project (DPST).

References

[1] M. M. Solomon, Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints, Operations Research. 35(2) (1987) 254–265.

[2] S. Dabia, S. Røpke, T. V. Woensel, T. De Kok, Branch and Price for the Time-Dependent Vehicle Routing Problem with Time Windows, Transportation Science. 47(3) (2013) 380–396.

[3] M. Desrochers, J. Desrosiers, M. M. Solomon, A new optimization algorithm for the vehicle routing problem with time windows, Operations research. 40(2) (1992) 342–354.

[4] O. Bräysy, M. Gendreau, Tabu search heuristics for the vehicle routing problem with time windows, Top. 10(2) (2002) 211–237.

[5] J. Berger, M. Barkaoui, A hybrid genetic algorithm for the capacitated vehicle routing problem, Journal of the Operational Research Society, 54(12) (2003) 1254–1262.

[6] L. M. Gambardella, É. D. Taillard, G. Agazzi, MACS-VRPTW: A multiple ant colony system for vehicle routing problems with time windows, in: New Ideas in Optimization, McGraw-Hill, London, pp. 63–76.

[7] F. Arnold, K. Sörensen, What makes a VRP solution good? the generation of problem-specific knowledge for heuristics, Computers & Operations Research. 106 (2019) 280–288.

[8] B. Kallehauge, J. Larsen, O.B. Madsen, M.M. Solomon, Vehicle Routing Problem with Time Windows. In: Desaulniers, G., Desrosiers, J., Solomon, M.M. (eds) Column Generation, Springer, Boston, 2005.

[9] CVRPLIB (Capacitated Vehicle Routing Problem Library), VRPTW benchmarks.

<http://vrp.atd-lab.inf.puc-rio.br/index.php/en/>, 2014 (accessed 03.08.24).

[10] Google, OR-Tools, Version 8.4.

<https://developers.google.com/optimization>, 2021 (accessed 10.09.24).

[11] Google OR-Tools, Routing Options.

<https://developers.google.com/optimization/routing/routing_options>, 2021 (accessed 10.09.24).

[12] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by Simulated Annealing, Science. 220(4598) (1983) 671–680.

[13] I. H. Osman, Metastrategy Simulated Annealing and Tabu Search Algorithms for the Vehicle Routing Problem, Annals of Operations Research. 41(4) (1993) 421–451.

[14] M.W.P. Savelsbergh, A parallel insertion heuristic for vehicle routing with side constraints, CWI Quarterly. 3(4) (1990) 305–319.

[15] C. Voudouris, E. P. KTsang, A. Alsheddy, Guided local search, in Handbook of Metaheuristics, Boston, 2010.

[16] P. Inbunsong, B. Intiyot, Identifying the most distinctive characteristics of a good solution for non-Euclidean CVRP using machine learning, Proceedings of the Operation research network of Thailand, 16–18 March 2022, Department of Industrial and Manufacturing Engineering, Faculty of Engineering, Prince of Songkla University, Songkla, Thailand, 2022, pp. 176–182.

[17] Google Research, Google Colaboratory.

<https://colab.research.google.com/>, 2020 (accessed 03.08.24).

Two-Step-Change pivot rule for the simplex method

Rawit Chaiareekij[1,*] and Krung Sinapiromsaran[1,†]

[1]Department of Mathematics and Computer Science, Faculty of Science
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

The purpose of this paper is to design a new pivot rule of the simplex method, a widely used algorithm for solving linear programming problems. The simplex method's efficiency hinges on the choice of the entering variable, determined by the pivot rule. Dantzig's original pivot rule, while effective in many cases, can exhibit exponential behavior on certain problems. To address this problem, this paper presents two-step-change improvement from the objective value. The rule selects the entering variable with the greatest two-step improvement and the leaving variable using the minimum ratio test. Experiments are conducted to compare the performance of this pivot rule with three established pivot rules which are Dantzig, Absolute change, and the largest distance. By our experimental results with the randomly generated multi-period planning problems from small-sized to medium-sized, the two-step-change pivot rule is faster than Dantzig, Absolute change, and the largest distance pivot rule on average ranking of computational time across different problem sizes, its average ranking of computational time was 1.0, 1.2, 1.8, 1.4 and 1.0, respectively, consistently achieving the lowest average rank. On Netlib LP problems, it performs the most effective result, achieving the best average rank of 1.83 across all Netlib LP problems. Moreover, this rule can solve a family of Klee-Minty cube linear programming problems with fewer iterations than Dantzig that requires an exponential number of pivoting steps.

Keywords: Linear programming, Simplex method, Pivot rule

2020 MSC: 90B05, 90C05, 90C08

1 Introduction

Linear programming (LP) [1,3] has been one of the most dynamic areas of applied mathematics. An LP model describes the mathematical programming problem dealing with the optimal value of a linear objective function which is defined as minimizing subject to a finite number of linear equality and inequality constraints. The system of constraints will form a feasible region.

*Speaker (6570097623@student.chula.ac.th)
†Corresponding author (krung.s@chula.ac.th)

The popular used method for solving LP problems is the simplex method which was first proposed by George B. Dantzig [1] in 1963 and efficiently used for small to medium sized LP problems. This method starts at a corner point corresponding to the initial basis of the feasible region and moves along the edge of the polytope to an adjacent extreme point until an optimal solution is obtained.

The main concept of the pivot rule from the simplex method is to select the entering variable and determine the leaving variable that improves the objective value. An effective pivot rule can significantly reduce the number of iterations for reaching the optimal solution. It is well known that Dantzig's pivot rule is the first rule to select the entering variable from non-basic variables with the most negatively reduced cost. However, the performance of Dantzig's pivot rule requires an exponential number of pivoting steps on some LP problems, there is an example of the worst-case running time proposed by Klee and Minty [2] in 1972. In addition, the simplex method with Dantzig's rule cannot prevent cycling [3]. Later, many researchers have attempted to improve the simplex performance by reducing the number of iterations and the computational time. In 1977, Forrest and Goldfarb [4] presented a way to reduce the number of iterations which was called "Steepest-edge rule". Other rules followed such as the largest distance pivot rule by Pan [5] in 2008 which is based on the normalized reduced cost, Absolute change pivot rule by K. Chankong, B. Intiyot and K. Sinapiromsaran [6] in 2014 by trying to block a leaving variable that makes a little change in the objective function value as much as possible, and A unified worst case for classical simplex and policy iteration pivot rule by Y. Disser and N. Mosis [14] in 2023 which applied the Markov decision process (MDP) - based technique to classical pivot rules and obtained a unified construction that excludes several pivot rules at the same time and any combination of them.

In this paper, we propose the pivot rule called two-step-change pivot rule for the simplex method which selects the entering variable based on the greatest improvement from the reduced cost and the objective function value from the leaving variable. Herein, we consider only linear programming problems are feasible and suppose that the starting basic feasible solution x with basis B is given. Then we perform experiments and compare the average ranking of this rule with three known pivot rules which are Dantzig, absolute change and the largest distance.

2 Preliminaries

2.1 The simplex method

The simplex method starts with the linear programming (LP) model in the standard form as follows:

$$\begin{aligned} \text{Minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b, \\ & x \geq 0, \end{aligned} \tag{2.1}$$

where $c \in \mathbb{R}^n, x \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = m$, and b is a nonnegative vector in $\mathbb{R}^m$. When the LP model in the standard form is transformed to the simplex tableau, the simplex technique may be computed with ease. Without loss of generality, let $A = [B \mid N]$ where B is an $m \times m$ invertible matrix and N is an $m \times (n - m)$ matrix. Here B is called the basic matrix and N the associated non-basic matrix. Therefore, x can be decomposed into x_B and x_N where x_B is a basic variable and x_N is a non-basic variable. Also, c can be decomposed as ${c_B}^T$ and ${c_N}^T$, and denoting I_B be the current set of the indices of the basic variables and I_N be the current set of the indices of the non-basic variables, respectively. Let $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$ be the solution of the equation $Ax = b$, where $x_B = B^{-1}b$ and $x_N = 0$ is called a basic solution of the system. If $x_B \geq 0$, x is called a basic feasible solution of the system. Suppose that a basic feasible solution of the system (2.1) is

$$[(B^{-1}b)^T \quad 0^T]^T,$$

and its associated objective value is z_0. Then

$$z_0 = {c_B}^T B^{-1} b. \tag{2.2}$$

Then feasibility requires that $x_B \geq 0$ and $x_N \geq 0$. We can rewrite the system $Ax = b$ as:

$$Bx_B + Nx_N = b. \tag{2.3}$$

Then

$$x_B = \bar{b} - \sum_{j \in I_N} (y_j x_j), \tag{2.4}$$

where $\bar{b} = B^{-1}b$, $y_j = B^{-1}A_j$, and A_j denotes the jth column vector of A.
Let z denote the objective value and

$$c^T = [{c_B}^T \ {c_N}^T].$$

From (2.1), (2.2) and (2.3), we have

$$z = z_0 - \sum_{j \in I_N} (z_j - c_j) x_j$$

where $z_j = {c_B}^T B^{-1} A_j$ for each non-basic variable. The non-basic reduced cost is obtained by $z_j - c_j$. The main result exhibits that the optimal solution is achieved if the index set

$$J = \{ j \mid z_j - c_j > 0, \ j \in I_N \} = \emptyset.$$

2.2 The simplex method in tableau format

Suppose that we have a starting basic feasible solution x with basis B. The linear programming (LP) model can be represented as follows:

$$\begin{aligned} &\text{Minimize} && z \\ &\text{subject to} && z - {c_B}^T x_B - {c_N}^T x_N = 0, && (2.5) \\ & && Bx_B + Nx_N = b, && (2.3) \\ & && x_B, x_N \geq 0. \end{aligned}$$

By multiplying B^{-1} to equation (2.3), we get

$$x_B + B^{-1}Nx_N = B^{-1}b. \quad (2.6)$$

Multiplying (2.6) by ${c_B}^T$,

$${c_B}^T x_B + {c_B}^T B^{-1}Nx_N = {c_B}^T B^{-1}b. \quad (2.7)$$

Adding to equation (2.5),

$$z - {c_B}^T B^{-1}b + {c_B}^T B^{-1}Nx_N - {c_N}^T x_N = 0. \quad (2.8)$$

Then, we get

$$z + 0x_B + ({c_B}^T B^{-1}N - {c_N}^T)x_N = {c_B}^T B^{-1}b. \quad (2.9)$$

Set, $x_N = 0$, and from the equations (2.6) and (2.9), we get $x_B = B^{-1}b$ and $z = {c_B}^T B^{-1}b$. Moreover, from the equations (2.6) and (2.9), we can conveniently represent the current basic feasible solution with basis B in the following tableau.

Table 1: The initial simplex tableau.

	z	x_B	x_N	RHS	
z	1	0	${c_B}^T B^{-1}N - {c_N}^T$	${c_B}^T B^{-1}b$	*Row 0*
x_B	0	I	$B^{-1}N$	$B^{-1}b$	*Rows 1 through m*

To determine the effect of introducing the non-basic variable x_k into the basis, the direction vector y_k is computed by solving the linear system $B_{y_k} = A_{:K}$ where B is the current basis matrix and $A_{:K}$ is the column vector k^{th} of A. The expression B_{y_k} represents the matrix vector multiplication which composes of the basis B and the column vector y_k. The resulting direction vector $y_k = B^{-1}A_{:K}$ describes how the basic variables will change if x_k enters the basis. Moreover, the initial tableau reports the value of the objective function $z = {c_B}^T B^{-1}b$, the basic variables $x_B = B^{-1}b$ and the objective row ${c_B}^T B^{-1}N - {c_N}^T$, which consists of the $z_j - c_j$ values for non-basic variable x_j. If $z_j - c_j \leq 0, \forall j$ then the current basic feasible solution is optimal. Otherwise, the current tableau can be improved for some non-basic variable x_N. The

process continues iteratively until no further improvement is attainable, ultimately leading to the determination of the optimal solution.

2.3 The simplex method with Dantzig's pivot rule

The basic steps of the standard simplex method with Dantzig's pivot rule are summarized (Minimization problem) as follows:

Initialization Step: Choose the initial basic feasible solution (with the basis B and the associated non-basic N) and construct the simplex tableau as follows:

Main Step:

Step 1: Select the entering variable from the non-basic variables: By Dantzig's rule

choose x_k such that $z_k - c_k = max\{ z_j - c_j \mid j \in I_N \}$.

Step 2: If $z_k - c_k \leq 0$, then $\begin{bmatrix} x_B \\ x_N \end{bmatrix}$ is an optimal solution and the algorithm stops.

Step 3: Select the leaving variable x_r by examining y_k where $y_k = B^{-1}A_{:K}$.

I. If $y_k \leq 0$ then LP has an unbounded solution, and the algorithm stops.

II. Otherwise, determine the index r as follows:

$$argmin\left\{ \frac{\bar{b}_j}{y_{jk}} \middle| y_{jk} > 0 \text{ and } j \in \{1, \dots, m\} \right\}.$$

This is called the minimum ratio test.

Step 4: Perform the pivot operation using the entering variable and the leaving variable and repeat from step 1 until the optimal solution is found or the LP model has the unbounded optimal solution.

2.4 The simplex method with the largest distance pivot rule

The largest distance pivot rule proposed by Ping-Qi Pan in 2008 [5] which is based on the normalized reduced cost. The steps of the simplex method with the largest distance pivot rule are summarized (Minimization problem) as follows:

Initialization Step: Choose the initial basic feasible solution (with the basis B and the associated non-basic N) and construct the simplex tableau as follows:

Main Step:

Step 1: Select the entering variable from the non-basic variables: By the largest distance pivot rule, choose x_k such that $z_k - c_k = max\left\{ \frac{z_k - c_k}{\|A_{:k}\|} \middle| k \in I_N \right\}$ where $\|A_{:k}\|$ is norm of A_k.

Step 2: If $z_k - c_k \leq 0$, then $\begin{bmatrix} x_B \\ x_N \end{bmatrix}$ is an optimal solution and the algorithm stops.

Step 3: Select the leaving variable x_r by examining y_k where $y_k = B^{-1}A_{:K}$.

I. If $y_k \leq 0$ then LP has an unbounded solution, and the algorithm stops.

II. Otherwise, determine the index r as follows:

$$argmin\left\{\frac{\bar{b}_j}{y_{jk}} \;\middle|\; y_{jk} > 0 \text{ and } j \in \{1, \dots, m\}\right\}.$$

This is called the minimum ratio test.

Step 4: Perform the pivot operation using the entering variable and the leaving variable and repeat from step 1 until the optimal solution is found or the LP model has the unbounded optimal solution.

2.5 The simplex method with absolute change pivot rule

Absolute change pivot rule proposed by k. Chankong, B. Intiyot and K. Sinapiromsaran in 2014 [6]. The idea is to have the maximum improvement in the objective value function by trying to block a leaving variable that makes a little change in the objective function value as much as possible. The steps of the simplex method with absolute change pivot rule are summarized (Minimization problem) as follows:

Initialization Step: Choose the initial basic feasible solution (with the basis B and the associated non-basic N) and construct the simplex tableau as follows:

Main Step:

Step 1: Check $z_k - c_k \leq 0$ for all $j \in J$, then $\begin{bmatrix} x_B \\ x_N \end{bmatrix}$ is an optimal solution. Stop the algorithm.

Step 2: Determine the entering variable by using absolute change pivot rule.

I. Set $CI_B = I_B$. Let $J = \{j \mid z_j - c_j > 0, j \in I_N\}$.

II. Select index r such that $r = argmin\{\bar{b}_i \mid i \in I_B\}$.

III. Let $\hat{J} = \{j \in J \mid A_{rj} \leq 0\}$. If $\hat{J} \neq \emptyset$, let $J = \hat{J}$.

- If $|J| = 1$, go to (IV).
- Else, remove i from CI_B. Go to (II).

IV. Else, choose x_k by $z_k - c_k = max\{z_j - c_j \mid j \in J\}$.

Step 3: Select the leaving variable x_r by examining y_k where $y_k = B^{-1}A_{:K}$.

I. If $y_k \leq 0$ then LP has an unbounded solution, and the algorithm stops.

II. Otherwise, determine the index r as follows:

$$argmin\left\{\frac{\bar{b}_j}{y_{jk}}\middle|\, y_{jk} > 0 \text{ and } j \in \{1, \ldots, m\}\right\}.$$

This is called the minimum ratio test.

Step 4: Perform the pivot operation using the entering variable and the leaving variable and repeat from step 1 until the optimal solution is found or the LP model has the unbounded optimal solution.

3 Main Results

An effective pivot rule can significantly reduce the number of iterations for reaching the optimal solution in LP problems. However, Dantzig's pivot rule requires an exponential number of pivoting steps on some LP problems. Therefore, we propose a new pivot rule by considering two-step improvement in the change of an objective value then select the leaving variable called **two-step-change pivot rule for the simplex method.**

3.1 The simplex method with two-step-change pivot rule

Initialization step: Let $J = \{j \mid z_j - c_j < 0, j \in I_N\}$. Given a basis B, (I_B is index of basic variables and I_N is index of non-basic variables).

Main step:

Step 1. If $z_j - c_j \leq 0$ for all $j \in J$, then $\begin{bmatrix} x_B \\ x_N \end{bmatrix}$ is an optimal solution. Stop the algorithm.

Step 2. Apply two-step-change pivot rule to an LP problem.

I. For each $k \in I_N$ and $z_k - c_k > 0$,
- Solve $B_{yk} = A_{:k}$ for the direction vector y_k.
- If $y_k \leq 0$ then LP has an unbounded solution, and the algorithm stops.
- Otherwise, go to (II).

II. Determine the entering variable x_k and leaving variable x_r by computing the optimal improvement (op) to determine the best step size:

$$op = (z_k - c_k) \times min\left\{\frac{\bar{b}_i}{y_{ki}}\middle|\, y_{ki} > 0,\ i \in I_B \text{ and } k \in I_N\right\}$$

- Select the entering variable x_k with the greatest op.
- Simultaneously, identify the leaving variable x_r corresponds to the basic variable $i \in I_B$ that minimizes:

$$r = min\left\{\frac{\bar{b}_i}{y_{ki}}\middle|\, y_{ki} > 0,\ i \in I_B\right\}$$

where r is the index of the leaving variable from the minimum ratio test.

Step 3. Swap the entering variable x_k with the leaving variable x_r.

Step 4. Perform the pivot operation using the entering and the leaving variable and go to Step 1.

3.2 Illustration of the method

In this section, we illustrate the implementation of our proposed pivot rule with two sections, one from the Klee-Minty cube problem and another one is LP problems from publications which are Kuhn example, Marshall and Suurballe example, and Sierksma example.

3.2.1 Application to the Klee-Minty cube problem

A family of Klee-Minty cube or Klee-Minty polytope problems was originally specified with a parameterized system of linear inequalities, with the dimension as the parameter. The cube in two-dimensional space is a squashed square, and the cube in three-dimensional space is a squashed cube. Illustrations of the cube have been described by picture for dimensions two and three [7].

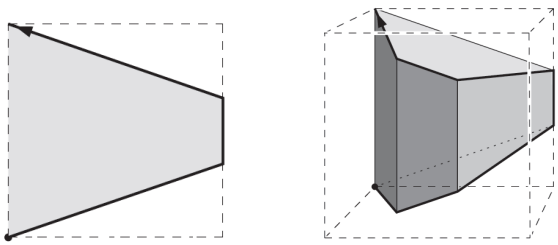


Figure 1: The Klee–Minty cube in two and three-dimensional space.

A family of Klee-Minty cube problems is given by [8].

$$\begin{aligned}
\text{Minimize} \quad & -2^{D-1}x_1 - 2^{D-2}x_2 - \cdots - 2x_{D-1} - x_D, \\
\text{Subject to} \quad & x_1 \leq 5, \\
& 4x_1 + x_2 \leq 25, \\
& 8x_1 + 4x_2 + x_3 \leq 125, \\
& \vdots \\
& 2^D x_1 + 2^{D-1}x_2 + \cdots + 4x_{D-1} + x_D \leq 5^D, \\
& x_1, \ldots, x_D \geq 0.
\end{aligned}$$

A family of Klee-Minty cube problems can be solved in $2^n - 1$ iterations by using the simplex method with Dantzig's pivot rule. We illustrate the case for $D = 3$ by Example 1.

Example 1. Klee and Minty cube problem with $D = 3$

$$\begin{aligned} \text{Minimize} \quad & -4x_1 - 2x_2 - x_3\,, \\ \text{subject to} \quad & x_1 \le 5, \\ & 4x_1 + x_2 \le 25, \\ & 8x_1 + 4x_2 + x_3 \le 125, \\ & x_1, x_2, x_3 \ge 0. \end{aligned}$$

The problem above can be written in tableau format, where x_4, x_5, x_6 are the slack variables:

	z	x_1	x_2	x_3	x_4	x_5	x_6	RHS
z	1	4	2	1	0	0	0	0
x_4	0	1	0	0	1	0	0	5
x_5	0	4	1	0	0	1	0	25
x_6	0	8	4	1	0	0	1	125

From the tableau, we can see that x_1, x_2 and x_3 can be the entering variables. Then we compute the optimal improvement (op) for the first column (x_1), the second column (x_2), and the third column (x_3), to determine the best step size as the entering variable. Following the method, we obtain the values of op for the first, the second and the third column are

$$(z_1 - c_1)\, min\left\{\frac{\bar{b}_i}{y_{1i}}\right\}, (z_2 - c_2)\, min\left\{\frac{\bar{b}_i}{y_{2i}}\right\}, (z_3 - c_3)\, min\left\{\frac{\bar{b}_i}{y_{3i}}\right\} \Bigg| \; y_{ki} > 0, i \in I_B \text{ and } k = 1, 2, 3.$$

They are $20, 50,$ and $125,$ respectively. Since the maximum value among all op is 125 corresponding to $k = 3$. As the result, x_3 is chosen to be the entering variable. From the minimum ratio test, we get x_6 as the leaving variable. After the pivot operation, we obtain the optimal solution which is shown in the tableau as follows.

	z	x_1	x_2	x_3	x_4	x_5	x_6	RHS
z	1	-4	-2	0	0	0	-1	-125
x_4	0	1	0	0	1	0	0	5
x_5	0	4	1	0	0	1	0	25
x_3	0	8	4	1	0	0	1	125

This is the optimal tableau. The optimal solution is $x_1 = 0$, $x_2 = 0$ and $x_3 = 125$. The optimal value is 125 with the number of iterations is 1. By the simplex method with Dantzig's pivot rule, the number of iterations is 7. In a similar fashion, the two-step- change pivot rule requires only one iteration for any D.

3.2.2 Some examples of cycling found in publications

The collection of examples (Examples 2-4) exhibits cycling when they are solved by the simplex method with Dantzig's rule [9]. For each problem, we note the number of iterations that forms a cycle when the problem is solved by the simplex method with Dantzig's rule.

Example 2. Kuhn example (Balinski and Tucker) [10]

$$\begin{aligned} \text{Minimize} \quad & -2x_4 - 3x_5 + x_6 + 12x_7, \\ \text{subject to} \quad & x_1 - 2x_4 - 9x_5 + x_6 + 9x_7 = 0, \\ & x_2 + \tfrac{1}{3}x_4 + x_5 - \tfrac{1}{3}x_6 - 2x_7 = 0, \\ & x_3 + 2x_4 + 3x_5 - x_6 - 12x_7 = 2, \\ & x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0. \end{aligned}$$

Solution: The optimal solution is $x_5 = -\frac{1}{6},\ x_6 = \frac{3}{2}, \text{and } x_7 = -\frac{1}{3}$; Minimum $= -2$; Cycle $= 6$. Two-step-change pivot rule converges in 1 iteration with computational time 0.0018 s.

Example 3. Marshall and Suurballe example [11]

$$\begin{aligned} \text{Minimize} \quad & -\frac{2}{5}x_5 - \frac{2}{5}x_6 + \frac{9}{5}x_7, \\ \text{subject to} \quad & x_1 + \tfrac{3}{5}x_5 - \tfrac{32}{5}x_6 + \tfrac{24}{5}x_7 = 0, \\ & x_2 + \tfrac{1}{5}x_5 - \tfrac{9}{5}x_6 + \tfrac{3}{5}x_7 = 0, \\ & x_3 + \tfrac{2}{5}x_5 - \tfrac{8}{5}x_6 + \tfrac{1}{5}x_7 = 0, \\ & x_4 + x_6 = 1, \\ & x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0. \end{aligned}$$

Solution: The optimal solution is $x_1 = 4,\ x_2 = 1,\ x_5 = 4, \text{and } x_6 = 1$; Minimum $= -2$; Cycle $= 6$. Two-step-change pivot rule converges in 2 iterations with computational time 0.0023 s.

Example 4 Sierksma example [12]

$$\begin{aligned} \text{Minimize} \quad & -3x_1 + 80x_2 - 2x_3 + 24x_4, \\ \text{subject to} \quad & x_1 - 32x_2 - 4x_3 + 36x_4 + x_5 = 0, \\ & x_1 - 24x_2 - x_3 + 6x_4 + x_6 = 0, \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0. \end{aligned}$$

This example shows that our rule may prevent cycling for the case $b_i = 0$ for all i.

Solution: Unbounded above; Cycle $= 6$. Two-step-change pivot rule converges in 4 iterations with computational time 0.0035 s.

3.3 Computational Experiment and Results

We perform the numerical runs and the computational results of simplex method with two-step-change pivot rule on 30 randomly generated multi-period planning problems and 30 standard problems from Netlib. For randomly generated multi-period planning problems, we compare various sizes ranging from small-sized (size 50) to medium-sized (size 500). Then we perform experiments and compare the average ranking of this rule with DP (Dantzig's pivot rule), ACP (Absolute change pivot rule) and LDP (the largest distance pivot rule) to ensure robust and efficient. The simplex code is implemented using Python. Dantzig [13], absolute change, the largest distance and two-step-change were implemented by Python libraries numpy. All codes were performed by using an Intel(R) Xeon(R) CPU@2.20 GHz with 13.00 GB RAM on Windows 10.

3.3.1 The randomly generated multi-period planning problems

Multi-period planning problems are optimization problems that involve making decisions over $\boldsymbol{T}$ time periods while considering costs, constraints, and future demand. These problems are commonly used in supply chain management, production planning, and inventory control. The model of multi-period planning problem is given below:

I. Decision variables

- $\boldsymbol{p_t}$: Production rate in period $\boldsymbol{t}$.
- $\boldsymbol{I_t}$: Inventory level at the end of period $\boldsymbol{t}$.

II. Objective function

$$\boldsymbol{min} \sum_{\boldsymbol{t=1}}^{\boldsymbol{T}} (\boldsymbol{c_p p_t} + \boldsymbol{c_I I_t})$$

where $\boldsymbol{c_p}$ is the cost per unit of production, and $\boldsymbol{c_I}$ is the cost per unit of inventory.

III. Constraints

Inventory must be balanced across time periods. This is ensured using inventory balance equations of the form:

$$\boldsymbol{I_{t-1} + p_t = d_t + I_t,} \quad \text{for } \boldsymbol{t = 1, 2, \dots, T}$$

- Initial inventory constraint:

$$\boldsymbol{I_0 + p_1 = d_1 + I_1}$$

This constraint ensures that the initial inventory $\boldsymbol{I_0}$ and production in the first period $\boldsymbol{p_1}$ are sufficient to meet the demand $\boldsymbol{d_1}$ and leave a remaining inventory $\boldsymbol{I_1}$.

- Ending inventory constraint:

$$\boldsymbol{I_{T-1} + p_T = d_T + I_T}$$

This constraint ensures that the inventory is carried over from period $\boldsymbol{T-1}$ and production in the last period $\boldsymbol{p_T}$ together meet the final period's demand $\boldsymbol{d_T}$ and satisfy the target ending inventory $\boldsymbol{I_T}$.

- Non-Negativity:

$$\boldsymbol{p_t, d_T \; and \; I_t \geq 0}$$

where $\boldsymbol{p_t}$: production amount in period $\boldsymbol{t}$, for $\boldsymbol{t = 1, 2, \dots, T}$.

$\boldsymbol{d_t}$: demand in period $\boldsymbol{t}$, for $\boldsymbol{t = 1, 2, \dots, T}$.

$\boldsymbol{I_t}$: inventory at the end of period $\boldsymbol{t}$.

To measure the performance of two-step-change pivot rule for solving multi-period planning linear programming problems, randomly generated multi-period planning linear programming problems are formulated as minimization problems and generated according to the following specifications:

I. The objective function

The objective function minimizes the total cost, which includes production costs and inventory holding costs. The cost coefficients are assigned as follows:

- Each production cost coefficient is independently drawn from a uniform distribution: $c_p \sim U(10,20)$.
- Each inventory holding cost coefficient is independently drawn from a uniform distribution: $c_I \sim U(5,20)$.

II. Constraints and right-hand side vector

The constraint matrix A is structured to enforce inventory balance constraints over T time periods. These constraints ensure that inventory and production satisfy demand while accounting for initial and ending inventory levels. The coefficients are assigned as follows:

- Initial inventory (I_0) is randomly selected from [10, 30].
- Ending inventory (I_T) is randomly selected from [10, 50].
- The right-hand side vector b is determined by randomly generated demand values d_t drawn from [50, 100].
- All constraints are formulated in the form $\leq$ to maintain feasibility.

III. Problem sizes and scaling

The randomly generated multi-period planning problems are simulated across varying sizes with different numbers of time periods T and decision variables N, according to the following specifications:

- Time periods: $T \in \{25, 75, 125, 200 \text{ and } 250\}$
- Decision variables: The total number of decision variables is determined by $\boldsymbol{N = 2T}$ because for each period $\boldsymbol{t = 1, 2, \ldots, T}$, they include two decision variables: $\boldsymbol{p_t}$ and $\boldsymbol{I_t}$. Even though the ending inventory $\boldsymbol{I_T}$ is randomly generated and fixed, it is used to define the boundary conditions in the inventory balance constraints. Thus, the total number of decision variables is set as $\boldsymbol{N = 2T}$, while understanding that $\boldsymbol{I_T}$ is a parameter and not subject to optimization.
- Number of constraints: $\boldsymbol{T}$

The plots of average computational time (s) comparison on multi-period planning problems with DP, ACP, LDP and 2-STEP for each problem sizes are obtained by running 30 trials per instance size, and the results are reported in Figures 2-6 below, which are shown in the log scale.

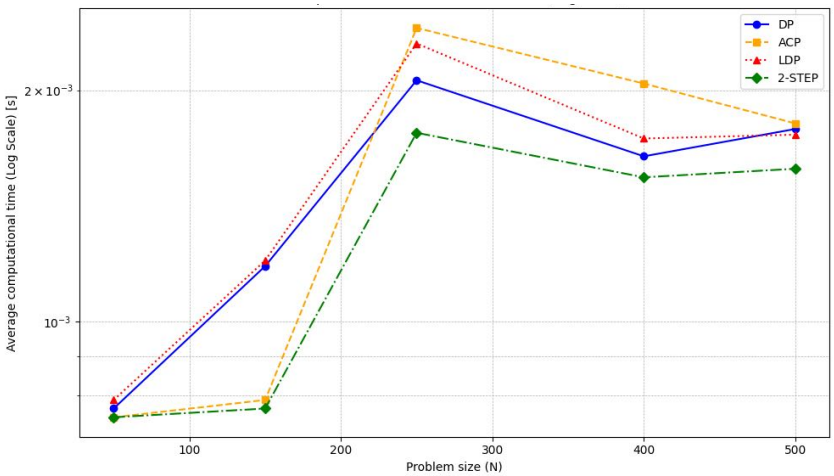


Figure 2. The average computational time (s) comparison on multi-period planning problems of $T = 25$ with DP, ACP, LDP and 2-STEP

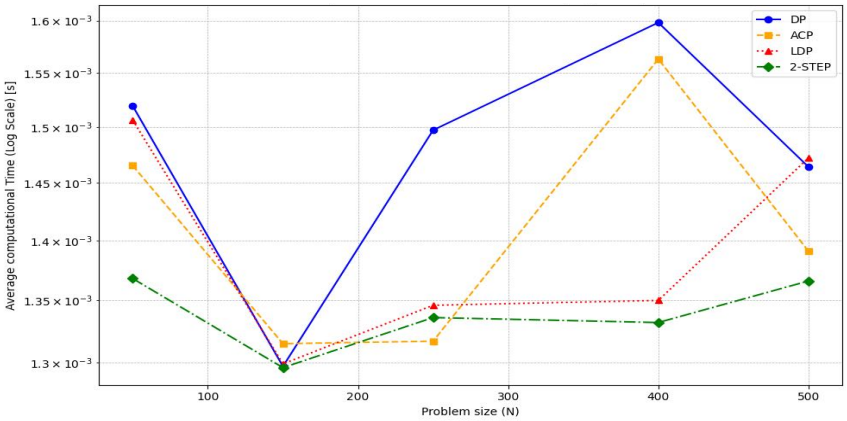


Figure 3. The average computational time (s) comparison on multi-period planning problems of $T = 75$ with DP, ACP, LDP and 2-STEP

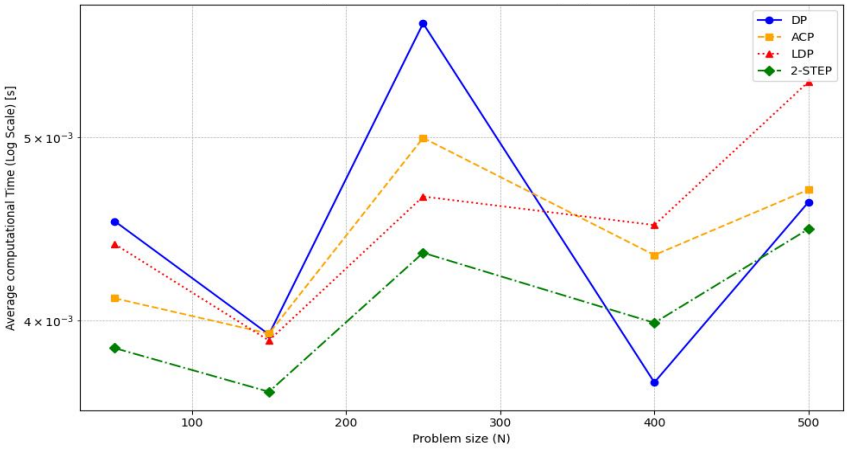


Figure 4. The average computational time (s) comparison on multi-period planning problems of $T = 125$ with DP, ACP, LDP and 2-STEP

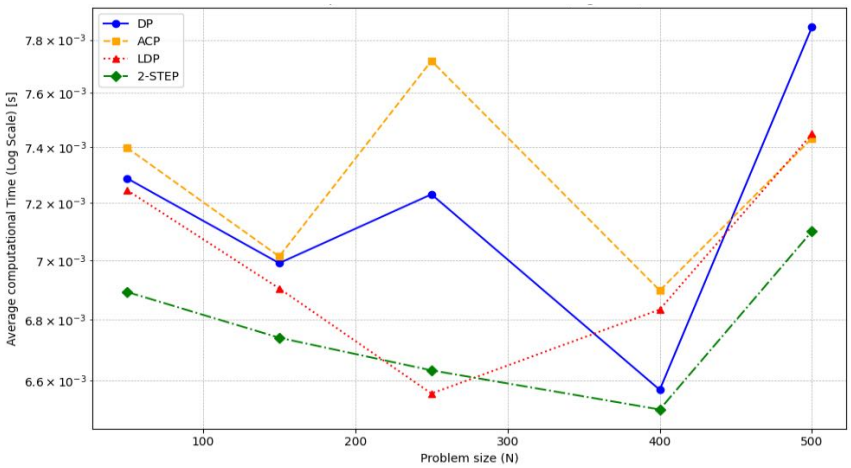


Figure 5. The average computational time (s) comparison on multi-period planning problems of $T = 200$ with DP, ACP, LDP and 2-STEP

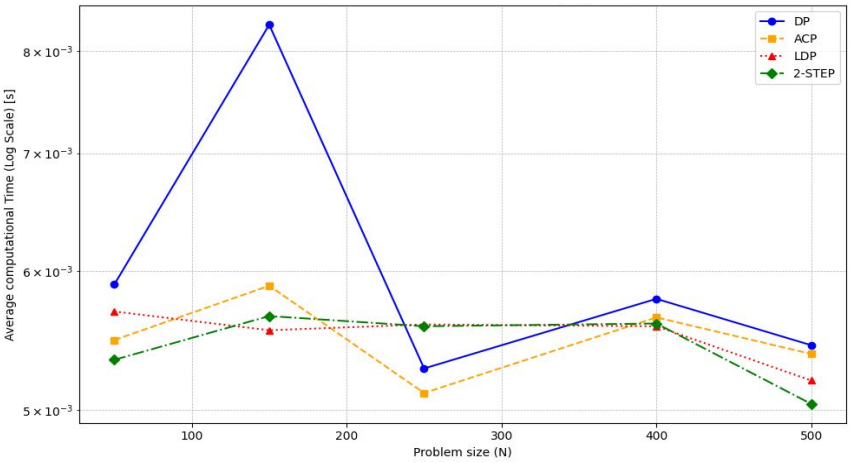


Figure 6. The average computational time (s) comparison on multi-period planning problems of $T = 250$ with DP, ACP, LDP and 2-STEP

To assess the performance of four pivot rules - Dantzig, absolute change, the largest distance and two-step-change - we evaluate the efficiency of each pivot rule by its average computational time across different problem sizes (N) and various values of (T). For each $N \in \{50, 150, 250, 400 \text{ and } 500\}$, we assign a rank (1-4) based on its computational time:

- Rank 1: The pivot rule with the lowest average computational time (best performance).
- Rank 2: The pivot rule with the second average computational time.

- Rank 3: The pivot rule with the third average computational time.
- Rank 4: The pivot rule with the highest average computational time (worst performance).

After ranking the pivot rules for each problem size, we compute the average rank across various values of T from Figures 2-6 to provide a holistic view of the performance of each method. Herein, we check for optimal solutions before they are included in the experiment.

Table 2: The average rank for each problem size comparison on multi-period planning problems across values of T with DP, ACP, LDP and 2-STEP

Time period (T)	Problem size (N)	Computational time (s) (Rank order)			
		DP	ACP	LDP	2-STEP
25	50	0.000770 (3)	0.000750 (1)	0.000790 (4)	0.000750 (1)
75		0.001519 (4)	0.001465 (2)	0.001506 (3)	0.001368 (1)
125		0.004514 (4)	0.004110 (2)	0.004390 (3)	0.003868 (1)
200		0.007288 (3)	0.007398 (4)	0.007246 (2)	0.006893 (1)
250		0.005895 (4)	0.005481 (2)	0.005690 (3)	0.005341 (1)
Average rank (size 50)		**3.6**	**2.2**	**3.0**	**1.0**
25	150	0.001180 (3)	0.000790 (2)	0.001200 (4)	0.000770 (1)
75		0.001297 (2)	0.001315 (4)	0.001299 (3)	0.001296 (1)
125		0.003933 (3)	0.003937 (4)	0.003904 (2)	0.003668 (1)
200		0.006991 (3)	0.007014 (4)	0.006905 (2)	0.006740 (1)
250		0.008283 (4)	0.005885 (3)	0.005552 (1)	0.005656 (2)
Average rank (size 150)		**3.0**	**3.4**	**2.4**	**1.2**
25	250	0.002060 (2)	0.002410 (4)	0.002300 (3)	0.001760 (1)
75		0.001497 (4)	0.001317 (1)	0.001346 (3)	0.001336 (2)
125		0.005743 (4)	0.004996 (3)	0.004652 (2)	0.004344 (1)
200		0.007231 (3)	0.007720 (4)	0.006558 (1)	0.006633 (2)
250		0.005281 (2)	0.005114 (1)	0.005594 (4)	0.005581 (3)
Average rank (size 250)		**3.0**	**2.6**	**2.6**	**1.8**

Time period (T)	Problem size (N)	Computational time (s) (Rank order)			
		DP	ACP	LDP	2-STEP
25	400	0.001640 (2)	0.002040 (4)	0.001730 (3)	0.001540 (1)
75		0.001598 (4)	0.001563 (3)	0.001350 (2)	0.001332 (1)
125		0.003710 (1)	0.004332 (3)	0.004493 (4)	0.003989 (2)
200		0.006570 (2)	0.006898 (4)	0.006834 (3)	0.006507 (1)
250		0.005785 (4)	0.005646 (3)	0.005581 (1)	0.005601 (2)
Average rank (size 400)		**2.6**	**3.4**	**2.6**	**1.4**
25	500	0.001780 (3)	0.001810 (4)	0.001750 (2)	0.001580 (1)
75		0.001464 (3)	0.001391 (2)	0.001472 (4)	0.001366 (1)
125		0.004621 (2)	0.004691 (3)	0.005350 (4)	0.004472 (1)
200		0.007849 (4)	0.007432 (2)	0.007448 (3)	0.007101 (1)
250		0.005444 (4)	0.005382 (3)	0.005198 (2)	0.005041 (1)
Average rank (size 500)		**3.2**	**2.8**	**3.0**	**1.0**

From Table 2, the 2-STEP is the most efficient method across all problem sizes, consistently achieving the best average computational times and securing the lowest average rank (close to 1.0 in all cases). For LDP, it performs moderately well, the average rank be second or third depending on the problem size. It tends to outperform DP, except in some cases where their efficiency is similar. While ACP shows inconsistent performance because it is competitive for smaller problem sizes but struggles as the problem complexity increases. It ranks last in several cases, particularly for $\boldsymbol{N = 150}$ and $\boldsymbol{N = 400}$. Lastly, DP is the worst efficient method, consistently achieving the third or fourth on average computational times and average rank.

3.3.2 The Netlib LP problems

We perform the computational results of simplex method with two-step-change pivot rule on 30 standard problems from Netlib LP test sets to compare the effectiveness of our pivot rule with other pivot rules, all methods are in the Mathematical Programming System (MPS) format.

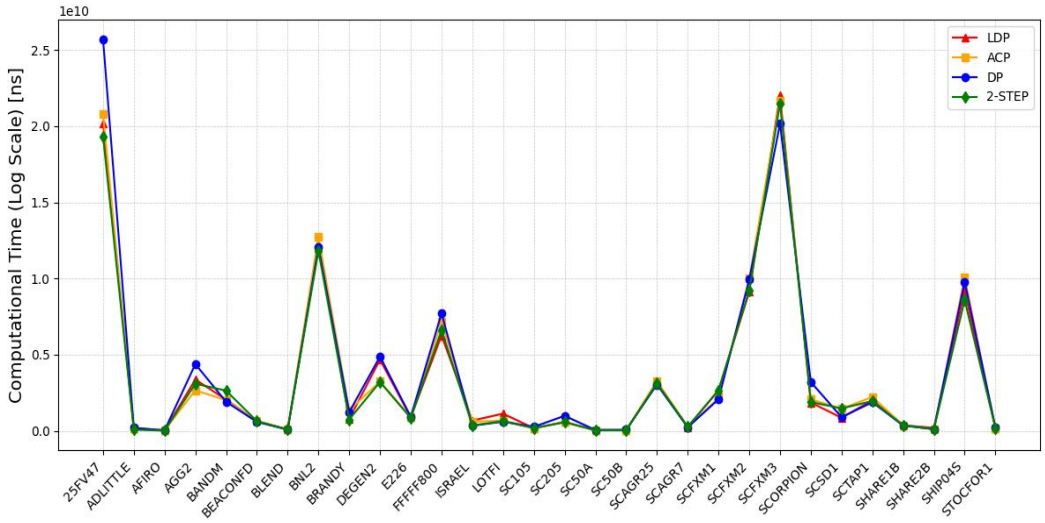


Figure 7. The comparison of DP, ACP, LDP and 2-STEP using the computational time (ns) - Netlib problems.

Figure 7 demonstrates the effectiveness of computational time from solving 30 standard test problems from real-world Netlib problems varying according to Dantzig's, absolute change, the largest distance and two-step-change pivot rule, which are shown in the nanoseconds (ns) scale. To measure the efficiency of each pivot rule, we assign a rank (1-4) based on its computational time similar to the comparison of randomly generated multi-period planning problems. After ranking the pivot rules for each Netlib LP problem, we compute the average rank across all problems from Netlib LP test sets to provide a holistic view of the efficiency of each method.

Table 3: The average rank comparison on 30 Netlib LP problems with DP, ACP, LDP and 2-STEP

No.	Problem name	Computational time (s) (Rank order)			
		DP	ACP	LDP	2-STEP
1	25FV47	25.698 (4)	20.809 (3)	20.206 (2)	19.339 (1)
2	ADLITTLE	0.205 (4)	0.161 (3)	0.143 (2)	0.084 (1)
3	AFIRO	0.038 (3)	0.046 (4)	0.029 (2)	0.020 (1)
4	AGG2	4.386 (4)	2.672 (1)	3.371 (3)	3.099 (2)
5	BANDM	1.910 (1)	2.011 (2)	2.057 (3)	2.644 (4)

No.	Problem name	Computational time (s) (Rank order)			
		DP	ACP	LDP	2-STEP
6	BEACONFD	0.593 (1)	0.656 (4)	0.633 (2)	0.638 (3)
7	BLEND	0.090 (1)	0.098 (4)	0.095 (3)	0.090 (1)
8	BNL2	12.083 (2)	12.740 (4)	12.217 (3)	11.809 (1)
9	BRANDY	1.235 (3)	1.362 (4)	0.788 (2)	0.756 (1)
10	DEGEN2	4.881 (4)	3.246 (2)	4.680 (3)	3.212 (1)
11	E226	0.928 (4)	0.895 (3)	0.884 (2)	0.853 (1)
12	FFFFF800	7.714 (4)	7.155 (3)	6.267 (1)	6.614 (2)
13	ISRAEL	0.342 (1)	0.605 (3)	0.681 (4)	0.367 (2)
14	LOTFI	0.612 (1)	0.683 (3)	1.132 (4)	0.636 (2)
15	SC105	0.268 (4)	0.168 (1)	0.177 (2)	0.187 (3)
16	SC205	0.993 (4)	0.590 (2)	0.594 (3)	0.573 (1)
17	SC50A	0.062 (4)	0.041 (2)	0.058 (3)	0.040 (1)
18	SC50B	0.075 (4)	0.049 (3)	0.044 (2)	0.041 (1)
19	SCAGR25	3.047 (1)	3.282 (4)	3.147 (3)	3.140 (2)
20	SCAGR7	0.236 (1)	0.265 (3)	0.260 (2)	0.267 (4)
21	SCFXM1	2.079 (1)	2.145 (2)	2.663 (3)	2.670 (4)
22	SCFXM2	9.932 (3)	9.995 (4)	9.133 (1)	9.222 (2)
23	SCFXM3	20.165 (1)	21.682 (3)	22.037 (4)	21.498 (2)
24	SCORPION	3.209 (4)	2.072 (3)	1.868 (1)	1.897 (2)
25	SCSD1	0.903 (2)	1.460 (3)	0.862 (1)	1.477 (4)
26	SCTAP1	1.885 (1)	2.228 (4)	2.018 (3)	1.964 (2)
27	SHARE1B	0.358 (1)	0.380 (4)	0.372 (3)	0.367 (2)
28	SHARE2B	0.117 (2)	0.118 (3)	0.199 (4)	0.110 (1)
29	SHIP04S	9.801 (3)	10.100 (4)	9.189 (2)	8.586 (1)
30	STOCFOR1	0.209 (4)	0.188 (3)	0.184 (2)	0.178 (1)
Average time (rank)		**3.801 (2.57)**	**3.596 (3.03)**	**3.533 (2.50)**	**3.412 (1.83)**

From Table 3, the 2-STEP is the most efficient method, achieving the best average rank of 1.83 across all Netlib LP problems. This means it consistently had the lowest computational time in overall of problems. Next, LDP and DP follow closely with the second average rank

2.50 and third average rank 2.57, meaning it performed slightly worse than 2-STEP but was still more efficient than ACP in most cases. While ACP performs the worst overall, with an average rank of 3.03. This indicates that ACP generally requires more computation time than the other pivot rules. However, the performance of the two-step pivot rule depends on the characteristics of different Netlib LP problems.

4 Discussion

This paper proposes a new pivot rule for solving a linear programming problem, namely the two-step-change pivot rule. The idea of this rule is to ensure that each iteration of the simplex method takes the best path to the optimal solution by considering improvement in the objective function per step of its increase. We compute the direction of changing for each entering variable to determine the best step size. The one with the greatest optimal improvement is chosen as the entering variable. This proposed rule could reduce the computational time. According to our test problems, Figures 2-7 present a summary of the average computational time and rank comparison on multi-period planning problems and Netlib LP problems with DP, ACP, LDP and 2-STEP. We conclude that the simplex method with two-step-change pivot rule has a better performance for solving some multi-period planning problems in small-sized to medium-sized, although its performance is not much different from the three known pivot rules in small-scale problems.

In future work, the two-step-change pivot rule could be further developed by considering ideas from the study "Exponential lower bounds for many pivot rules for the simplex Method" [15], which shows that some pivot rules may follow very long paths in specific cases. While the two-step-change rule aims to improve performance based on reduced cost and step size, its behavior in such worst-case scenarios is not yet clear. To reduce the risk of such behavior, one possible direction is to enhance the two-step-change rule by incorporating broader geometric information, such as the angle or direction of movement in the feasible region. This may help the pivot rule avoid inefficient paths. From the study [15], we can investigate how rule design affects path length and use those insights to adjust selection criteria or create hybrid rules. These ideas may extend to a more robust method suitable for a wider range of linear programming problems.

References

[1] G.B. Dantzig, *Linear Programming and Extensions*, Princeton University Press, Princeton, 1963.

[2] V. Klee, G. Minty, How good is the simplex algorithm?, in: *Inequalities*, Academic Press, New York, 1972.

[3] S.H. Bazara, M. Jarvis, *Linear Programming and Network Flows*, 2nd ed., John Wiley, New York, 1990.

[4] J.J. Forrest, D. Goldfarb, Steepest-edge simplex algorithm for linear programming, *Math. Program.* 57 (1992) 341–374. http://dx.doi.org/10.1007/BF01581089

[5] P.-Q. Pan, A largest-distance pivot rule for the simplex algorithm, *Eur. J. Oper. Res.* 187 (2008) 393–402.

[6] K. Chankong, B. Intiyot, K. Sinapiromsaran, Absolute change pivot rule for the simplex algorithm, in: *Proc. Int. MultiConf. Eng. Comput. Sci. 2014*, Vol. II, IMECS 2014, Hong Kong, 2014.

[7] A. Deza, E. Nematollahi, T. Terlaky, How good are interior point methods? Klee–Minty cubes tighten iteration-complexity bounds, *arXiv preprint* (May 2008).

[8] H.J. Greenberg, Klee-Minty polytope shows exponential time complexity of simplex method, *Univ. Colo. Denver* (1997).

[9] S.I. Gass, S. Vinjamuri, Cycling in linear programming problems, *Comput. Oper. Res.* 31 (2004) 303–311.

[10] M.L. Balinski, A.W. Tucker, Duality theory of linear programs: a constructive approach with applications, *SIAM Rev.* 11 (1997) 347–377.

[11] K. Marshall, J. Suurballe, A note on cycling in the simplex method, *Nav. Res. Logist. Q.* 12 (1969) 121–137.

[12] G. Sierksma, *Linear and Integer Programming*, 2nd ed., Marcel Dekker, Inc., New York, 1969.

[13] M. Lutz, *Programming Python*, 24th ed., O'Reilly Media, 2010.

[14] Y. Disser, N. Mosis, A unified worst case for classical simplex and policy iteration pivot rules, *arXiv preprint arXiv:2309.14034* (2023).

[15] A.E. Black, Exponential lower bounds for many pivot rules for the simplex method, *arXiv preprint arXiv:2403.04886* (2024).

Mathematical Model of *Helicobacter* with Liver Fluke as Reservoir

Tachin Kaewmala[1,*], Apichat Suratanee[2],
Kitiporn Plaimas[1] and Ratinan Boonklurb[1,†]

[1]Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Phayathai Rd., Pathumwan, Bangkok, 10330 Thailand.

[2]Department of Mathematics, Faculty of Applied Science, King Mongkut's University of Technology, North Bangkok, Bangkok 10800, Thailand.

Abstract

Helicobacter pylori (H. pylori) infection is one of the factors that cause gastric cancer. Liver flukes specifically *Opisthorchis viverrini* (*O. viverrini*) can also serve as a reservoir for this bacterium. Analyzing and investigating *H. pylori* and liver fluke infections would be of value in better controlling their spread. Thus, we have developed a mathematical model to explain the relationship between the proportion of *H. pylori* in humans and *O. viverrini* in the human body. The model considers six components: the average number of parasites per person, the average number of bacteria per person, the average number of parasite eggs excreted by humans, the ratio of infected snails, the average number of cercariae per snail and the ratio of infected fish. This study aims to analyze the interaction between *H. pylori* and *O. viverrini* in the host body. The model is defined by using ordinary differential equations and analyzed for the basic reproduction number, disease-free equilibrium, endemic equilibrium and their stability using the Routh–Hurwitz theorem and the Lyapunov function. Sensitivity analysis is performed for each parameter, including the basic reproduction number. The simulation results show that the bacterium grows rapidly due to its high division rate. We also found that the bacterial population increase due to the rate of bacterial production by the liver fluke, which releases the bacterium through its excretory/secretory products. Although the number of eggs released was high, the ratio of infected snails remained low due to the low consumption rate of eggs by snails.

Keywords: *O. viverrini*, *H. pylori*, basic reproduction number, disease free equilibrium point, endemic equilibrium point

2020 MSC: 92D25, 37N25

*Speaker (tachin3457@gmail.com)

†Corresponding author (ratinan.b@chula.ac.th)

1 Introduction

Helicobacter pylori (*H. pylori*) is a microaerophilic, spiral-shaped, Gram-negative bacterium that was first discovered in 1982. It is one of the most virulent human pathogens, colonizing the stomach of more than half of the global population, particularly in adults. *H. pylori* infection is a major risk factor for gastric cancer and significantly increases the likelihood of developing peptic ulcers and other gastric malignancies. Gastric cancer is ranked as the fifth most common cancer in the world, following lung, breast, colon and prostate cancers. In year 2022, it is also the third leading cause of cancer-related deaths worldwide, after lung and colon cancer, with an estimated 783,000 deaths. Gastric cancer is more prevalent in developing countries than in developed countries [1].

However, we found that this bacterium may have some relationship with the liver fluke, *Opisthorchis viverrini* (*O. viverrini*), which is a causative agent of cholangiocarcinoma. In [2], they investigated the relationship between the hamster and this bacterium by dividing the hamsters into five groups: i) Control hamsters not infected with *O. viverrini*; ii) *O. viverrini*-infected hamsters; iii) non-*O. viverrini*-infected hamsters treated with antibiotics (ABx); iv) *O. viverrini*-infected hamsters treated with ABx; and v) *O. viverrini*-infected hamsters treated with both ABx and praziquantel (PZQ). They then kept the hamsters for up to 3 months and measured the concentration of *H. pylori* bacteria in their feces. The results showed that the *H. pylori* concentration was the highest in group iii and group iv had a higher concentration than group v, which received anthelmintic medication. This led us to suspect that there might be some relationship between *O. viverrini* and *H. pylori*. They also cultured *O. viverrini* in vitro for 4 weeks and tested it for *H. pylori*. They found this bacterium in the gut of the liver fluke and also tested the excretory-secretory (ES) products of the fluke, finding *H. pylori* in the ES products as well.

It has also been reported in [3] that *H. pylori* infection with Cytotoxin-associated gene A (cagA) and liver fluke infection together are significant factors in the manifestation of biliary diseases, including cholangiocarcinoma, in liver fluke infection in the northeastern region of Thailand. Additionally, it has been reported that *H. pylori* can survive on the intestinal walls of *O. viverrini* cultured in vitro with antibiotics for up to 30 days. The ability of *H. pylori* to survive in *O. viverrini* may be due to the environment in the intestinal tract of *O. viverrini*, where the pH is around 5–6, which is different from the pH of bile, which is between 7 and 8. This may help protect *H. pylori* from the unfavorable environment in bile.

The first intermediate host of *O. viverrini* is the freshwater snail *Bithynia*. The eggs hatch into miracidia after about two months. Infected snails release cercariae, which are free-swimming. These cercariae penetrate the skin of the secondary intermediate host, the fish, and become fully infective metacercariae within 21 days. Then, when the final host, humans, eat undercooked fish, they become infected with *O. viverrini*. Once *O. viverrini* mature in the secondary host, they lay their eggs in the host's bile ducts, and these eggs are excreted into the outside world through the feces [4].

We incorporated the life cycle of *O. viverrini* along with a report stating that these parasites serve as reservoirs for *H. pylori* to create a diagram presented in Figure 1. Afterward, we will find the disease-free equilibrium and endemic equilibrium points and analyze their stability, as well as calculate the basic reproduction number of the model. Additionally, we will conduct sensitivity analysis of each parameter's effect on the variables and the basic reproduction number.

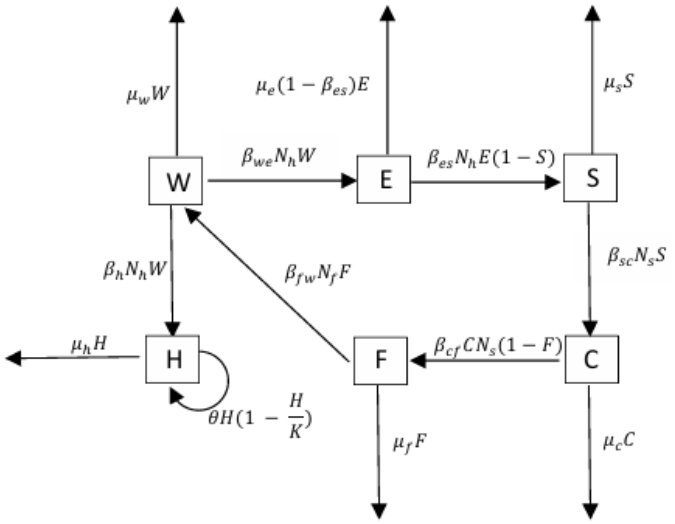


Figure 1: Diagram of our model

2 Our Model

From the diagram in Figure 1, we obtain system of ordinary differential equations (ODEs).

$$\frac{dW}{dt} = \beta_{fw}N_f F - \mu_w W, \tag{2.1}$$
$$\frac{dH}{dt} = \beta_h N_h W + \theta H(1-\frac{H}{K}) - \mu_h H, \tag{2.2}$$
$$\frac{dE}{dt} = \beta_{we}N_h W - \mu_e(1-\beta_{es})E - \beta_{es}N_h E, \tag{2.3}$$
$$\frac{dS}{dt} = \beta_{es}EN_h(1-S) - \mu_s S - \beta_{sc}N_s S, \tag{2.4}$$
$$\frac{dC}{dt} = \beta_{sc}N_s S - \mu_c C - \beta_{cf}N_s C, \tag{2.5}$$
$$\frac{dF}{dt} = \beta_{cf}CN_s(1-F) - \mu_f F - \beta_{fw}N_f F. \tag{2.6}$$

The term $\beta_h N_h W$ describes the process of producing the mean *H. pylori* per human (H) from *O. viverrini* through the fluids. $\theta H\left(1-\frac{H}{K}\right)$ describes the reproduction of *H. pylori* by themselves, while $\mu_h H$ represents the death of *H. pylori* caused by nature. The terms $\beta_{fw}N_f F$ and $\mu_w W$ explain the increasing of the mean worm per human (W) from transmission from fish to humans through consumption of infected fish and the decreasing with natural death, respectively. For the mean release eggs worm per human (E), $\beta_{we}N_h W$ describes the process of eggs produced by worm in human. For the destruction terms $\beta_{es}N_h E$ and $\mu_h(1-\beta_e s)E$ describe the situation that these eggs fall into water and snails come to eat them and the situation that these eggs are not eaten by snails with they will finally decay, respectively. The rate of change of the proportion of infected snails (S) increases by non infected snail eating eggs of parasite which is represented by $\beta_{es}EN_h(1-S)$ and one of the decreasing factors $\mu_s S$ is due to natural death of infection snails. After that in infection snail these eggs are develop into cercariae and

come out from snail into the water. The rate of change of The mean release cercariae per snail (C) depends on the number of infection snail and release rate of cercariae from snail, $\beta_s c N_s S$ and the natural death rate of cercariae, $\mu_c C$. The proportion infection fish (F) increases by cercariae penetrate under the skin of fish which is represented by $\beta_c f C N_s(1-F)$ and decreases by the natural death of infection fish, $\mu_f F$. All variables and parameters are shown in Tables 1 and 2.

Table 1: Description of model variables

Variable	**Description**
W	Mean worm per human H
H	Mean *H. pylori* per human H
E	Mean release eggs worm per human
S	Proportion infected snail
C	Mean release cercariae per snail
F	Proportion infected fish

Table 2: Model parameters and their descriptions

Parameter	**Description**	**Dimension**
N_h	Population size of human	Animals
N_f	Population size of fish	Animals
N_s	Population size of snail	Animals
β_h	Produce rate of *H. pylori* per person per *O. viverrini*	1/(Time × Animals)
β_{fw}	Transmission rate from infected fish to human per person per fish	1/(Time × Animals)
β_{we}	Produce rate of parasite egg per parasite	1/(Time × Animals)
β_{es}	Transmission rate from parasite eggs per snail	1/(Time × Animals)
β_{sc}	Rate of cercariae release from infected snails per snail	1/(Time × Animals)
β_{cf}	Transmission rate from cercariae to fishes per fishes	1/(Time × Animals)
θ	Division rate of *H. pylori*	1/Time
μ_h	Natural death rate of *H. pylori*	1/Time
μ_w	Natural death rate of worm	1/Time
μ_e	Natural death rate of eggs that cannot be hatched	1/Time
μ_s	Natural death rate of infected snail	1/Time
μ_c	Natural death rate of cercariae	1/Time
μ_f	Natural death rate of infected fish	1/Time
K	Carrying capacity	Animals

3 Model analysis

3.1 Existence and Uniqueness of solution

Observe that the right-hand-side equations in (2.1) − (2.6) are continuous. In addition their jacobian can be represented by the following matrix:

$$J = \begin{bmatrix} -\mu_w & 0 & 0 & 0 & 0 & N_f\beta_{fw} \\ N_h\beta_h & -\mu_h - \theta\left(\frac{H}{K} - 1\right) - \frac{H\theta}{K} & 0 & 0 & 0 & 0 \\ N_h\beta_{we} & 0 & \mu_e(\beta_{es} - 1) - \beta_{es}N_h & 0 & 0 & 0 \\ 0 & 0 & -N_h\beta_{es}(S-1) & -\mu_s - \beta_{sc}N_s - \beta_{es}N_hE & 0 & 0 \\ 0 & 0 & 0 & N_s\beta_{sc} & -\mu_c - N_s\beta_{cf} & 0 \\ 0 & 0 & 0 & 0 & -N_s\beta_{cf}(F-1) & -\mu_f - N_f\beta_{fw} - CN_s\beta_{cf} \end{bmatrix}.$$

it is clear that each element in J is continuous. From Theorem 3.4 on pages 125 - 126 of [5], there exists a unique solution for the system of Equations (2.1) - (2.6) with repect to non - negatives intitial conditions $W(0)$, $H(0)$, $E(0)$, $S(0)$, $C(0)$ and $F(0)$.

3.2 Basic reproduction number

Definition 3.1. [4] The basic reproduction number R_0 is the average number of new cases of an infection (or number of parasite offspring) caused by one typical infected individual (or one parasite), from one generation to the next, in a population with no pre-existing infections.

In [6], they determined R_0 by using the next-generation matrix L. This matrix represents the number of new infections or the number of adult parasites in the next generation. Then, R_0 is defined as the spectral radius of L.

To construct L, we use the idea of [6] by defindig f_r as the transmission part, describing the production of new infections, and and v_r as the transition part, describing changes in state (including removal by death or the acquisition of immunity). That is,

$$f_r = \begin{bmatrix} \beta_{fw}N_fF \\ \beta_hN_hW + \theta\left(1 - \frac{H}{K}\right)H \\ \beta_{we}N_hW \\ \beta_{es}EN_h(1-S) \\ \beta_{sc}N_sS \\ \beta_{cf}CN_s(1-F) \end{bmatrix} \text{ and } v_r = \begin{bmatrix} -\mu_wW \\ -\mu_hH \\ -\mu_e(1-\beta_{es})E - \beta_{es}N_hE \\ -\mu_sS - \beta_{sc}N_sS \\ -\mu_cC - \beta_{cf}N_sC \\ -\mu_fF - \beta_{fw}N_fF \end{bmatrix}.$$

Then, $L = -F_rV_r^{-1}$ at disease-free equilibrium point (DFE), where F_r and V_r are jacobian matrices of f_r and v_r, respectively [6]. For our problem, we have

$$F_r = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \beta_{fw}N_f \\ N_h\beta_h & \theta\left(1 - \frac{2H}{K}\right) & 0 & 0 & 0 & 0 \\ \beta_{we}N_h & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_{es}N_h(1-S) & -\beta_{es}N_hE & 0 & 0 \\ 0 & 0 & 0 & \beta_{sc}N_s & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta_{cf}N_s(1-F) & -\beta_{cf}N_sC \end{bmatrix}$$

and

$$V_r = \begin{bmatrix} -\mu_w & 0 & 0 & 0 & 0 & 0 \\ 0 & -\mu_h & 0 & 0 & 0 & 0 \\ 0 & 0 & -\mu_e(1-\beta_{es}) - \beta_{es}N_h & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mu_s - \beta_{sc}N_s & 0 & 0 \\ 0 & 0 & 0 & 0 & -\mu_c - \beta_{cf}N_s & 0 \\ 0 & 0 & 0 & 0 & 0 & -\mu_f - \beta_{fw}N_f \end{bmatrix}.$$

Thus, the next-generation matrix L at DFE is

$$L = -F_r V_r^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \frac{\beta_{fw}N_f}{\mu_f+\beta_{fw}N_f} \\ \frac{N_h\beta_h}{\mu_w} & \frac{\theta}{\mu_h} & 0 & 0 & 0 & 0 \\ \frac{\beta_{we}N_h}{\mu_w} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\beta_{es}N_h}{\mu_e+\beta_{es}N_h-\beta_{es}\mu_e} & \frac{\beta_{es}N_hE}{\mu_s+\beta_{sc}N_s} & 0 & 0 \\ 0 & 0 & 0 & \frac{\beta_{sc}N_s}{\mu_s+\beta_{sc}N_s} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\beta_{cf}N_s}{\mu_c+\beta_{cf}N_s} & 0 \end{bmatrix}.$$

Direct calculation gives the spectral radius which is the basic reproduction number in the form:

$$R_0 = \sqrt[5]{\frac{a}{beh}}, \tag{3.1}$$

where

$$\begin{aligned} a &= N_f N_h^2 N_s^2 \beta_{cf}\beta_{sc}\beta_{es}\beta_{we}\beta_{fw}, \\ b &= \mu_w(\mu_c + N_s\beta_{cf})(\mu_e + N_h\beta_{es} - \beta_{es}\mu_e), \\ e &= \mu_f + N_f\beta_{fw}, \\ h &= \mu_s + N_s\beta_{sc}, \end{aligned}$$

and a, b, e, h are positive.

3.3 Equilibrium points and their stability

3.3.1 Disease free equilibrium (DFE) point and its local stability

This point is a point where the *H. pylori* does not spread in the population [7]. Thus, we given $H = 0$ and we solve Equations (2.1) and (2.3) − (2.6). Therefore, the DFE point is given by:

$$E_0 = \{W^0, H^0, E^0, S^0, C^0, F^0\} = \{0, 0, 0, 0, 0, 0\}. \tag{3.2}$$

To establish the local stability of E_0, we first find the Jacobian matrix J at E_0 which can be written as

$$J(E_0) = \begin{bmatrix} -\mu_w & 0 & 0 & 0 & 0 & N_f\beta_{fw} \\ N_h\beta_h & -\mu_h+\theta & 0 & 0 & 0 & 0 \\ N_h\beta_{we} & 0 & \mu_e(\beta_{es}-1) - \beta_{es}N_h & 0 & 0 & 0 \\ 0 & 0 & N_h\beta_{es} & -\mu_s - \beta_{sc}N_s & 0 & 0 \\ 0 & 0 & 0 & N_s\beta_{sc} & -\mu_c - N_s\beta_{cf} & 0 \\ 0 & 0 & 0 & 0 & N_s\beta_{cf} & -\mu_f - N_f\beta_{fw} \end{bmatrix} = \begin{bmatrix} j_{11} & 0 & 0 & 0 & 0 & j_{16} \\ j_{21} & j_{22} & 0 & 0 & 0 & 0 \\ j_{31} & 0 & j_{33} & 0 & 0 & 0 \\ 0 & 0 & j_{43} & j_{44} & 0 & 0 \\ 0 & 0 & 0 & j_{54} & j_{55} & 0 \\ 0 & 0 & 0 & 0 & j_{65} & j_{66} \end{bmatrix}.$$

Next, the characteristic equation of matrix $J(E_0)$ is

$$a_0\lambda^6 + a_1\lambda^5 + a_2\lambda^4 + a_3\lambda^3 + a_4\lambda^2 + a_5\lambda + a_6 = 0 \tag{3.3}$$

where

$a_0 = 1,$
$a_1 = -j_{11} - j_{22} - j_{33} - j_{44} - j_{55} - j_{66},$
$a_2 = j_{11}j_{22} + j_{11}j_{33} + j_{11}j_{44} + j_{22}j_{33} + j_{11}j_{55} + j_{22}j_{44} + j_{11}j_{66}$
$+ j_{22}j_{55} + j_{33}j_{44} + j_{22}j_{66} + j_{33}j_{55} + j_{33}j_{66} + j_{44}j_{55} + j_{44}j_{66} + j_{55}j_{66},$
$a_3 = -j_{11}j_{22}j_{33} - j_{11}j_{22}j_{44} - j_{11}j_{22}j_{55} - j_{11}j_{33}j_{44} - j_{11}j_{22}j_{66} - j_{11}j_{33}j_{55}$
, $- j_{22}j_{33}j_{44} - j_{11}j_{33}j_{66} - j_{11}j_{44}j_{55} - j_{22}j_{33}j_{55} - j_{11}j_{44}j_{66} - j_{22}j_{33}j_{66}$
$- j_{22}j_{44}j_{55} - j_{11}j_{55}j_{66} - j_{22}j_{44}j_{66} - j_{33}j_{44}j_{55} - j_{22}j_{55}j_{66} - j_{33}j_{44}j_{66} - j_{33}j_{55}j_{66} - j_{44}j_{55}j_{66},$
$a_4 = j_{11}j_{22}j_{33}j_{44} + j_{11}j_{22}j_{33}j_{55} + j_{11}j_{22}j_{33}j_{66} + j_{11}j_{22}j_{44}j_{55} + j_{11}j_{22}j_{44}j_{66} + j_{11}j_{33}j_{44}j_{55}$
$+ j_{11}j_{22}j_{55}j_{66} + j_{11}j_{33}j_{44}j_{66} + j_{22}j_{33}j_{44}j_{55} + j_{11}j_{33}j_{55}j_{66} + j_{22}j_{33}j_{44}j_{66}$
$+ j_{11}j_{44}j_{55}j_{66} + j_{22}j_{33}j_{55}j_{66} + j_{22}j_{44}j_{55}j_{66} + j_{33}j_{44}j_{55}j_{66},$
$a_5 = -j_{11}j_{22}j_{33}j_{44}j_{55} - j_{11}j_{22}j_{33}j_{44}j_{66} - j_{11}j_{22}j_{33}j_{55}j_{66} - j_{11}j_{22}j_{44}j_{55}j_{66} - j_{11}j_{33}j_{44}j_{55}j_{66}$and
$- j_{16}j_{31}j_{43}j_{54}j_{65} - j_{22}j_{33}j_{44}j_{55}j_{66}$
$a_6 = j_{11}j_{22}j_{33}j_{44}j_{55}j_{66} + j_{16}j_{22}j_{31}j_{43}j_{54}j_{65}.$

From the Routh-Hurwitz Theorem [8,9], the number of roots for Equation (3.3) with positive real parts is equal to the number of sign changes in either one of the following Routh array,

λ^6	a_0	a_2	a_4	a_6	0	0
λ^5	a_1	a_3	a_5	0	0	0
λ^4	b_1	b_2	b_3	0	0	0
λ^3	c_1	c_2	0	0	0	0
λ^2	d_1	d_2	0	0	0	0
λ^1	e_1	0	0	0	0	0
λ^0	g_1	0	0	0	0	0

where

$$b_1 = \frac{a_1a_2 - a_0a_3}{a_1}, \quad b_2 = \frac{a_1a_4 - a_0a_5}{a_1}, \quad b_3 = \frac{a_1a_6 - a_00}{a_1}$$
$$c_1 = \frac{b_1a_3 - a_1a_2}{b_1}, \quad c_2 = \frac{b_1a_5 - a_1b_3}{b_1}$$
$$d_1 = \frac{c_1b_2 - b_1c_2}{c_1}, \quad d_2 = \frac{c_1b_3 - b_10}{c_1}$$
$$e_1 = \frac{d_1c_2 - c_1d_2}{d_1}, \quad g_1 = \frac{e_1d_2 - d_10}{e_1}.$$

Since $a_0 = 1$, a_1, b_1, c_1, d_1, e_1 and g_1 in the first column of the Routh array are positive. Then, the systems of Eqautions (2.1) - (2.6) are locally asymptomatically stable at the DFE point.

3.3.2 Endemic equilibrium point (EEP) and global stability

This point is a bacteria persist in the population [7]. From Eqautions (2.1) − (2.6), if we set all derivatives to be zero, then we obtain W^*, H^*, E^*,S^*,C^*, F^* as follow,

$$\begin{aligned}
W^* &= \frac{F^* N_f \beta_{fw}}{\mu_w}, \\
H_2^* &= \frac{K(\theta - \mu_h + \sqrt{\frac{K\mu_h^2 - 2K\mu_h\theta + K\theta^2 + 4N_h W^* \beta_h \theta}{K}})}{2\theta}, \\
E^* &= \frac{N_h W^* \beta_{we}}{\mu_e + N_h\beta_{es} - \beta_{es}\mu_e}, \\
S^* &= \frac{E^* N_h \beta_{es}}{\mu_s + N_s\beta_{sc} + E^* N_h \beta_{es}}, \\
C^* &= \frac{N_s S^* \beta_{sc}}{\mu_c + N_s\beta_{cf}} \quad \text{and} \\
F^* &= \frac{C^* N_s \beta_{cf}}{\mu_f + N_f\beta_{fw} + C^* N_s \beta_{cf}}.
\end{aligned}$$

Since actually we obtain two values of H^* which are

$$\begin{aligned}
H_1^* &= -\frac{K(\mu_h - \theta + \sqrt{\frac{K\mu_h^2 - 2K\mu_h\theta + K\theta^2 + 4N_h W^* \beta_h \theta}{K}})}{2\theta} \quad \text{and} \\
H_2^* &= \frac{K(\theta - \mu_h + \sqrt{\frac{K\mu_h^2 - 2K\mu_h\theta + K\theta^2 + 4N_h W^* \beta_h \theta}{K}})}{2\theta}.
\end{aligned}$$

However, H_1^* its negative which it not meaningful in the context of out problem. Therefore, we consider only H_2^* which is represented by H^* above. Since, W^*, E^*, S^* and C^* depend on F^*, we obtain

$$F^* = \frac{F^* ad}{F^* ad + F^* bce + bdeh}, \tag{3.4}$$

where

$$\begin{aligned}
&a = N_f N_h^2 N_s^2 \beta_{cf}\beta_{sc}\beta_{es}\beta_{we}\beta_{fw}, \quad b = \mu_w(\mu_c + N_s\beta_{cf})(\mu_e + N_h\beta_{es} - \beta_{es}\mu_e), \\
&c = N_f N_h^2 \beta_{es}\beta_{we}\beta_{fw}, \quad d = \mu_w(\mu_e + N_h\beta_{es} - \beta_{es}\mu_e), \\
&e = \mu_f + N_f\beta_{fw} \quad \textit{and} \quad h = \mu_s + N_s\beta_{sc}.
\end{aligned}$$

From Eqaution (3.4), we obtain

$$F^* = \left(\frac{a - beh}{ad + bce}\right) d \quad \text{and} \quad 0.$$

Since, we are finding EE point, we then consider only positive F^*. Therefore, we need the condition $a - beh > 0$. It is clear Eqaution (3.1) this condition holds if and only if $R_0 > 1$. In addition positivity of F^* implies positivity of W^*, E^*, S^* and C^*. Thus, our EE point which indicates that the disease is still present in the system is given by

$$X^* = \{W^*, H^*, E^*, S^*, C^*, F^*\}.$$

To establish the global stability for X^*, we use the idea of Corollary 1 on page 224 of [5]. First, we use the idea of [1] to defind the Lyapunov function as follows:

$$\begin{aligned} L(Y) = &\left(W - W^* + W^* \ln\left(\frac{W^*}{W}\right)\right) + \left(H - H^* + H^* \ln\left(\frac{H^*}{H}\right)\right) \\ &+ \left(E - E^* + E^* \ln\left(\frac{E^*}{E}\right)\right) + \left(S - S^* + S^* \ln\left(\frac{S^*}{S}\right)\right) \\ &+ \left(C - C^* + C^* \ln\left(\frac{C^*}{C}\right)\right) + \left(F - F^* + F^* \ln\left(\frac{F^*}{F}\right)\right), \end{aligned}$$

where $Y = X - X^*$ and X is a vector function of $\{W, H, E, S, C, F\}$. Observe that $L(Y)$ is positive and $L(Y) = 0$ if and only if $X = X^*$. Thus, $L(Y)$ is positive definite.

Next, the derivative of $L(Y)$ with respect to t is given by:

$$\dot{L}(Y) = \frac{dL(Y)}{dt} = \nabla L(Y) \cdot \frac{dY}{dt} = \nabla L(Y) \cdot \frac{dX}{dt} \tag{3.5}$$

$$\begin{aligned} \dot{L}(Y) = &\left(\frac{W - W^*}{W}\right)\frac{dW}{dt} + \left(\frac{H - H^*}{H}\right)\frac{dH}{dt} + \left(\frac{E - E^*}{E}\right)\frac{dE}{dt} \\ &+ \left(\frac{S - S^*}{S}\right)\frac{dS}{dt} + \left(\frac{C - C^*}{C}\right)\frac{dC}{dt} + \left(\frac{F - F^*}{F}\right)\frac{dF}{dt}. \end{aligned} \tag{3.6}$$

By replacing Eqautions (2.1) − (2.6) into Eqaution (3.6), we obtain

$$\dot{L}(Y) = P - N, \tag{3.7}$$

where

$$\begin{aligned} P = &\frac{(W - W^*)\beta_{fw} N_f F}{W} + \beta_h N_h W \cdot \frac{H - H^*}{H} + \theta(H - H^*)\left(1 - \frac{H}{K}\right) + \beta_{we} N_h W \cdot \frac{E - E^*}{E} \\ &+ \beta_{es} N_h E(1 - S) \cdot \frac{S - S^*}{S} + \beta_{sc} N_s S \cdot \frac{C - C^*}{C} + \beta_{cf} C N_s (1 - F) \cdot \frac{F - F^*}{F} \\ N = &+\mu_w (W - W^*) + \mu_h (H - H^*) + \mu_e (1 - \beta_{es})(E - E^*) + \beta_{es} N_h (E - E^*) \\ &+ \mu_s (S - S^*) + \beta_{sc} N_s (S - S^*) + \mu_c (C - C^*) + \beta_{cf} N_f F \cdot \frac{C - C^*}{C} \\ &+ \mu_f (F - F^*) + \beta_{fw} N_f (F - F^*). \end{aligned}$$

Observe that if $X = X^*$, then $\dot{L}(Y) = 0$.

Therefore, due to Corollary 1 on page 224 of [5], if $P < N$ then, $\dot{L}(Y) < 0$ which implies that X^* is globally asymptomatically stable.

4 Sensitivity analysis

The Partial Rank Correlation Coefficient (PRCC) [1] is a powerful tool for analyzing uncertainty or sensitivity. The PRCC value ranges from -1 to $+1$, with its absolute value indicating the strength of the relationship between the variables x and y. If the PRCC value is close to -1 or

+1, the linear relationship between x and y is strong. Conversely, if the PRCC value is far from −1 or +1, the relationship between x and y is weaker. A negative PRCC value suggests that x and y have opposite effects, while a positive value indicates that x and y influence the outcome in the same direction. The PRCC values for our model parameters and R_0 are presented in the diagrams shown in Figure 2.

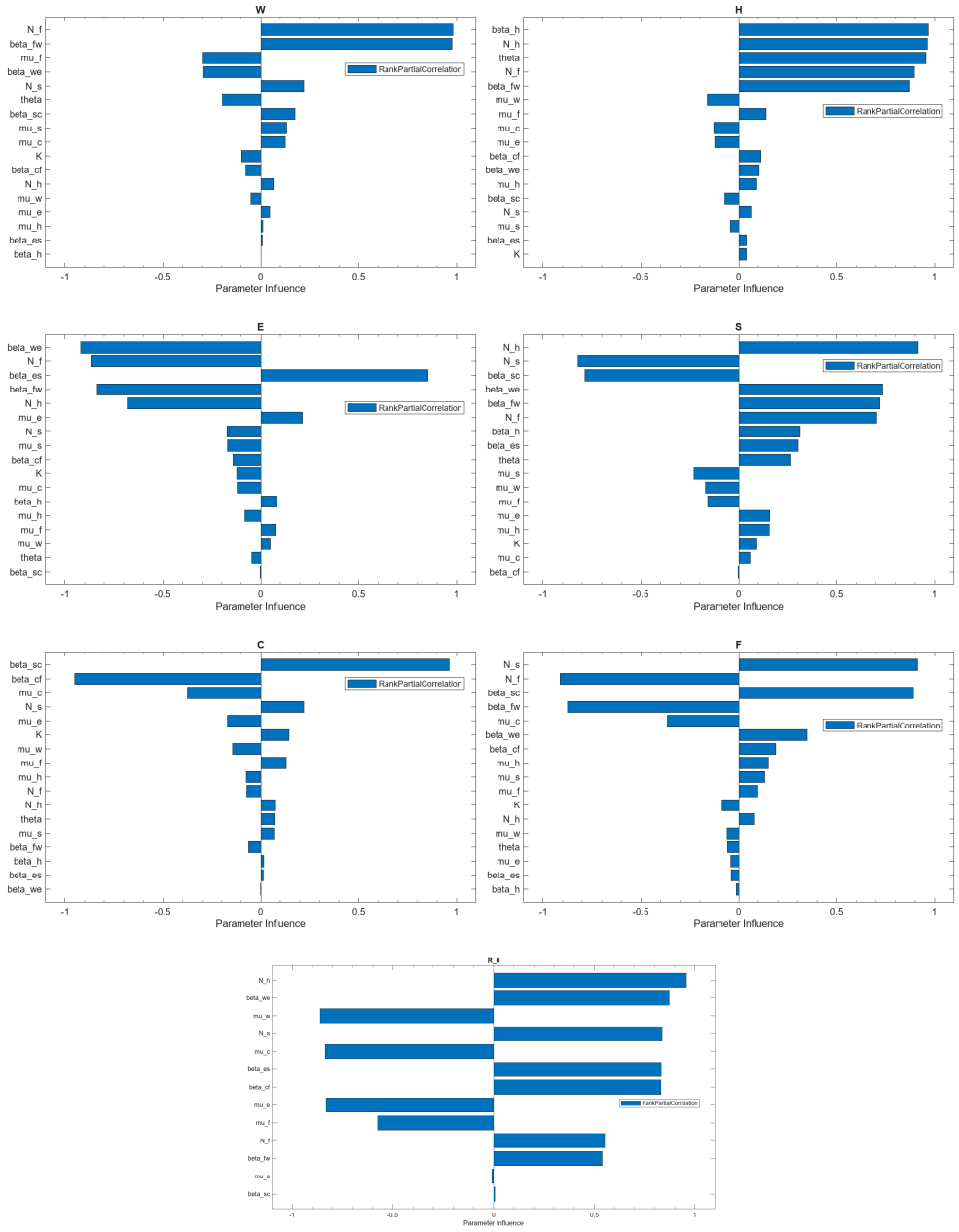


Figure 2: Show Partial Rank Correlation Coefficient (PRCC) for each parameter of all variables and R_0

In Figure 2, the parameters with the greatest impact on each variable are as follows.

- For the average number of parasites per person, the most influential parameters are N_f and β_{fw}, both having a positive effect.
- For the variable H, the parameters with the most significant positive impact are β_h, N_h, θ, N_f, and β_{fw}.
- For the variable E, the most impactful parameters are β_{we}, N_f, and β_{es}.
- For the proportion of infected snails, the key parameters are N_h, N_s, and β_{sc}.
- For the variable C, the most influential parameters are β_{sc} and β_{cf}.
- For the proportion of infected fish, the key parameters are N_f, N_s, and β_{sc}.
- For the basic reproduction number R_0, the most influential parameters are N_h, β_{we}, and μ_w.

5 Numerical simulation

In this model, we consider a closed system, assuming a small population size (as shown in Table 3) and that each liver fluke serves as a reservoir of H. pylori, producing these bacteria via secreted products (ES products). Furthermore, all liver fluke eggs produced are excreted from the body via feces. We use MATLAB R2024b to numerically solve our system of differential equations (2.1) – (2.6), with the parameters given in Table 3.

In Figure 3, the average number of *H. pylori* per person is significantly high and grows much faster compared to other variables. This rapid growth is due to their fast division rate and the fact that liver flukes can produce *H. pylori* through their ES products. Additionally, their mortality rate is low, with a lifespan of around 20–30 years, meaning their population hardly decreases in a short period.

As for the average number of liver flukes per person, their growth is relatively slow compared to *H. pylori*. This is because the transmission rate from infected fish to humans is relatively low. However, the number of eggs produced are remarkably high and increases rapidly. This is due to the high and rapid egg production of liver flukes, combined with the relatively slow consumption of these eggs by snails. As a result, the total number of liver fluke eggs remains high.

Nevertheless, the number of cercariae released is quite low compared to the number of eggs. This is because the consumption of liver fluke eggs by snails is relatively low, reducing the likelihood of these eggs hatching and developing into cercariae. Consequently, the number of infected fish gradually increases, but at a slower rate than the growth of *H. pylori*. This is likely due to the limited number of cercariae, which slows down the transmission process to fish compared to the rapid growth of *H. pylori*.

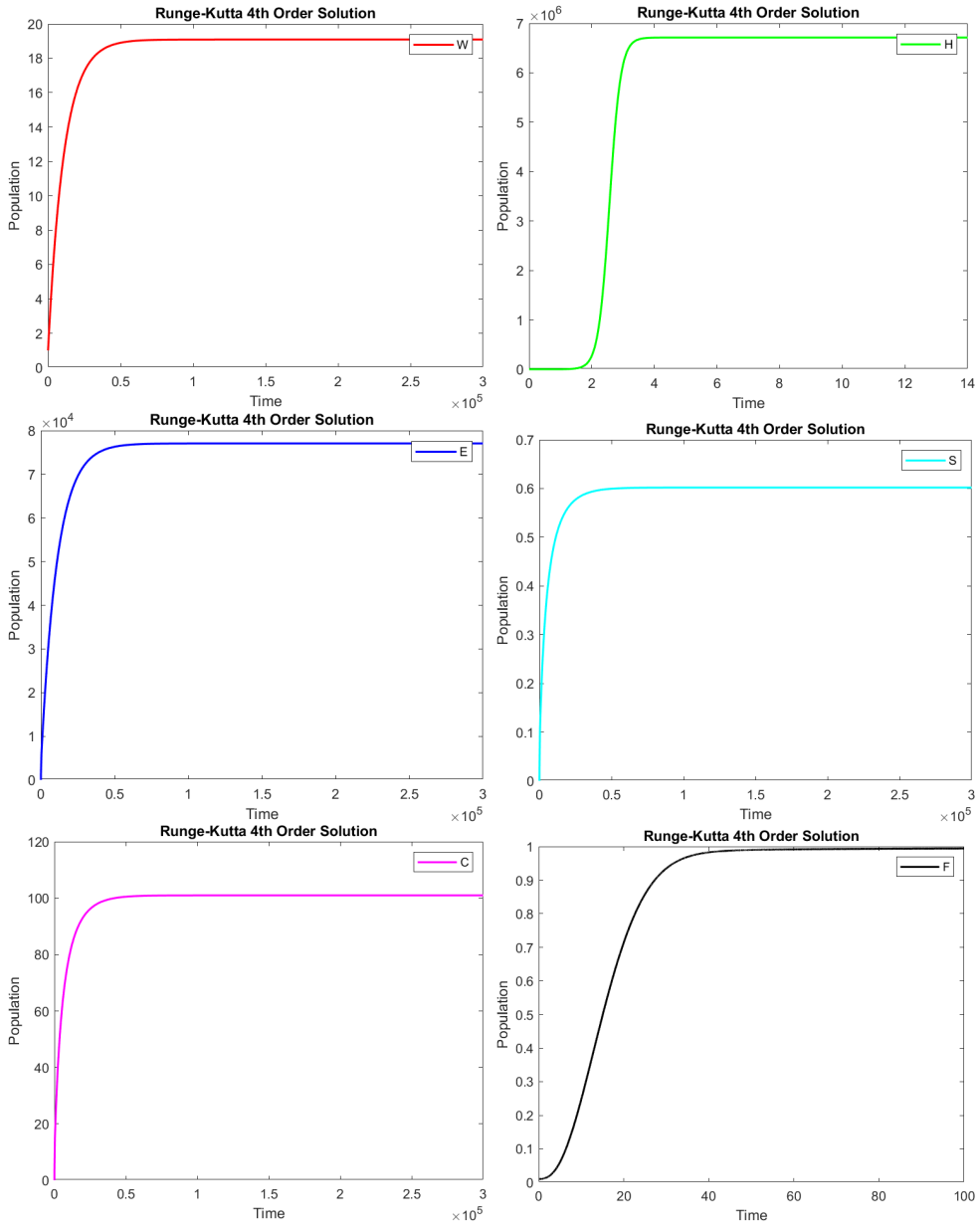


Figure 3: Numerical simulation of model

Table 3: Table of parameters and their sources

Parameter	Value	Source
N_h	100	Assume
N_f	50	Assume
N_s	50	Assume
β_h	0.169	estimated based on [2]
β_{fw}	3.4891×10^{-5}	[4]
β_{we}	0.493	estimated based on [10]
β_{es}	1.3892×10^{-5}	estimated based on [11]
β_{sc}	1.84	estimated based on [12]
β_{cf}	1.2×10^{-3}	estimated based on [13]
θ	5.5814	estimated based on [14]
μ_e	1.1×10^{-2}	[15]
μ_c	0.43	[16]
μ_h	1.0959×10^{-4}	Assume
μ_w	1.0959×10^{-4}	Assume
μ_s	0.0027	[4]
μ_f	0.0011	[4]
K	6.7132×10^{6}	[2]

6 Discussion

In this study, we examine the existence and uniqueness of the solutions to the model, as well as compute the basic reproduction number (R_0), which is a key factor in analyzing the endemic equilibrium. Additionally, we employ the Routh-Hurwitz criteria to prove the local stability of the disease-free equilibrium and use a Lyapunov function to verify the global stability of the endemic equilibrium. We also perform a sensitivity analysis using the PRCC to assess the impact of the parameters on R_0 and other state variables. The results show that β_h, N_h, θ, N_f and β_{fw} have the greatest impact on the average *H. pylori* prevalence per person. For R_0, the most influential parameters are N_h, β_{we}, and μ_w. The PRCC results indicate that N_h, β_{we}, and μ_w are the key determinants of transmission of this simulation.

Finally, numerical simulations reveal that liver flukes and their larvae exhibit slower reproduction rates compared to *H. pylori*, which can proliferate rapidly both through direct division and the presence of liver flukes. Moreover, the simulations indicate that both liver flukes and *H. pylori* can persist in the human body for decades if left untreated. The long-term presence of liver flukes in the host may increases the risk of gastric cancer or other gastrointestinal diseases if timely treatment is not administered.

References

[1] G.K. Mutua, C.G. Ngari, G.G. Muthuri, D.M. Kitavi, Mathematical modeling and simulating of *H. pylori* treatment and transmission implications on stomach cancer dynamics, Commun. Math. Biol. Neurosci., vol. 2022, (2022), http://dx.doi.org/10.28919/cmbn/7542.

[2] R. Deenonpoe, C. Chomvarin, C. Pairojkul, Y. Chamgramol, A. Loukas, P.J. Brindley, B. Sripa, The Carcinogenic Liver Fluke *Opisthorchis viverrini* is a Reservoir for Species of *Helicobacter*, Asian Pac. J. Cancer Prev. 16(5) (2015) 1751-1758. https://doi.org/10.7314/apjcp.2015.16.5.1751

[3] B. Sripa, R. Deenonpoe, P.J. Brindley, Co-infections with liver fluke and *Helicobacter* species: A paradigm change in pathogenesis of *opisthorchiasis* and *cholangiocarcinoma*, Trends Parasitol. 33 (2017) 420–426. https://doi.org/10.1016/j.parint.2016.11.016.

[4] C. Bürli, H. Harbrecht, P. Odermatt, S. Sayasone, & N. Chitnis, Mathematical analysis of the transmission dynamics of the liver fluke, *Opisthorchis viverrini*. J. Theor. Biol., 439 (2018), 181-194. https://doi.org/10.1016/j.jtbi.2017.11.020.

[5] F. Brauer, The Qualitative Theory of Ordinary Differential Equations. New York: Dover Publications, (2001).

[6] O. Diekmann, J.A.P. Heesterbeek, M.G. Roberts, The construction of next-generation matrices for compartmental epidemic models, J. R. Soc. Interface 7 (2009) 873–885. https://doi.org/10.1098/rsif.2009.0386

[7] G.T. Tilahun, O.D. Makinde, D. Malonza, Modelling and optimal control of pneumonia disease with cost-effective strategies, J. Biol. Dyn. 11 (2017) 400–426. https://doi.org/10.1080/17513758.2017.1337245

[8] K. Jetter, M.D. Buhmann, W. Haussmann, R. Schaback, J. Stöckler (Eds.), Stability Considerations, in: Topics in Multivariate Approximation and Interpolation, Elsevier, (1999), pp. 583-586.

[9] M.A. Haidekker, Stability Analysis for Linear Systems, in: Linear Feedback Controls: The Essentials, Elsevier, (2013), Chapter 10, pp. 140-143.

[10] P. Sithithaworn, R.H. Andrews, V.A. Mordvinov, M.Y. Pakharukova, M.N. Lvov, *Opisthorchis viverrini* and *Opisthorchis felineus*, in: Y. Motarjemi (Ed.), Encyclopedia of Food Safety, Elsevier (2024) 673-685, https://doi.org/10.1016/B978-0-12-822521-9.00068-X

[11] S. Prasopdee, J. Kulsantiwong, S. Piratae, P. Khampoosa, C. Thammasiri, A. Suwannatrai, T. Laha, R. Grams, A. Loukas, S. Tesana, Temperature dependence of Opisthorchis viverrini infection in first intermediate host snail, Bithynia siamensis goniomphalos, Acta Tropica 141 (2015) 112–117, https://doi.org/10.1016/j.actatropica.2013.10.011

[12] P. Phongsasakulchoti, P. Sri-Aroon, Y. Kerdpuech, Emergence of *Opisthorchis viverrini* cercariae from naturally infected Bithynia (Digoniostoma) siamensis goniomphalos, Southeast Asian J. Trop. Med. Public Health 36 (2005) 189–193, https://pubmed.ncbi.nlm.nih.gov/16438207/

[13] H.T.T. Dao, V. Dermauw, S. Gabriël, A. Suwannatrai, S. Tesana, G.T.T. Nguyen, P. Dorny, *Opisthorchis viverrini* infection in the snail and fish intermediate hosts in Central Vietnam,

Acta Tropica 170 (2017) 120–125, https://doi.org/10.1016/j.actatropica.2017.02.028

[14] X. Jiang, M.P. Doyle, Growth supplements for *Helicobacter pylori*, J. Clin. Microbiol. 38 (2000) 1984–1987, https://doi.org/10.1128/jcm.38.5.1984-1987.2000

[15] P. Echaubard, T. León, K. Suwannatrai, J. Chaiyos, C.S. Kim, F.F. Mallory, S. Kaewkes, R.C. Spear, B. Sripa, Experimental and modelling investigations of *Opisthorchis viverrini* miracidia transmission over time and across temperatures: implications for control, Int. J. Parasitol. 47 (2017) 257–270, https://doi.org/10.1016/j.ijpara.2016.10.007

[16] C. Trevisin, L. Kamber, L. Mari, N. Chitnis, J. Perez-Saez, S. Sayasone, P. Odermatt, M. Gatto, A. Rinaldo, A spatially explicit model of the dynamics of *Opisthorchis viverrini* spread, Ecohydrology (2024), http://dx.doi.org/10.1002/eco.2754

ตัวแบบคณิตศาสตร์การแพร่กระจายของเชื้อเอชไอวีและโรคเอดส์ที่มีผลจากประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พและการรักษาด้วยวิธียาต้านไวรัส ART

ฐาปณี ทองคำ[1,*] เกตุกนก หนูดี[1] อัญชุลี ณ ตะกั่วทุ่ง[1] และ กันยากร อ่อนรักษ์[1,†]

[1]สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสุราษฎร์ธานี 84000

บทคัดย่อ

การระบาดของเอชไอวีและโรคเอดส์เป็นปัญหาสาธารณสุขที่ทั่วโลกให้ความสำคัญ การระบาดยังส่งผลกระทบโดยรวมต่อสุขภาพประชาชน การพัฒนาสังคม และเศรษฐกิจ ในปัจจุบันเอชไอวีและโรคเอดส์กลายเป็นโรคเรื้อรังที่รักษาได้ โดยต้องรับประทานยาต้านไวรัสอย่างต่อเนื่อง สามารถกดปริมาณไวรัสในกระแสเลือดได้ และในกรณีผู้ไม่ติดเชื้อเอชไอวีสามารถป้องกันด้วยถุงยางอนามัยและการใช้ยาต้านเอชไอวีก่อนสัมผัสเชื้อที่เรียกว่าเพร็พ (PrEP) ซึ่งสามารถลดการแพร่กระจายเชื้อเอชไอวีไปสู่บุคคลอื่นได้ ในการวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาและวิเคราะห์ตัวแบบคณิตศาสตร์การแพร่กระจายของเชื้อเอชไอวีและโรคเอดส์ที่มีผลจากประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พและการรักษาด้วยวิธียาต้านไวรัส ART ได้พัฒนาตัวแบบคณิตศาสตร์การแพร่กระจายของเชื้อเอชไอวีและโรคเอดส์และวิเคราะห์จุดสมดุล วิเคราะห์เสถียรภาพ ค่าระดับการติดเชื้อ วิเคราะห์ดัชนีความไว และผลเชิงตัวเลข ผลการศึกษาพบว่าค่าระดับการติดเชื้อ $(\mathfrak{R}_0)$ เป็นเกณฑ์ในการควบคุมโรค ถ้า $\mathfrak{R}_0 < 1$ แล้วจุดสมดุลที่ไม่มีการแพร่ระบาดมีความเสถียรภาพเชิงเส้นกำกับ แต่ถ้า $\mathfrak{R}_0 > 1$ แล้วจุดสมดุลที่มีการแพร่ระบาดมีความเสถียรภาพเชิงเส้นกำกับ นอกจากนี้ยังพบว่าถ้าประสิทธิภาพของการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมากกว่า 96% และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART มากกว่า 90% จะสามารถควบคุมโรคได้

คำสำคัญ: เชื้อเอชไอวี, โรคเอดส์, ยาเพร็พ, การรักษาแบบ ART

2020 MSC: 92D30

*ผู้นำเสนอ (6504302001003@student.sru.ac.th)

†ผู้ประพันธ์บรรณกิจ (kanyakon.onr@sru.ac.th)

1 บทนำ

การติดเชื้อเอชไอวีและโรคเอดส์เป็นหนึ่งในโรคติดต่อทางเพศสัมพันธ์ที่องค์การอนามัยโลกให้ความสำคัญเพื่อยุติการแพร่กระจายโรค ในปัจจุบันการแพร่ระบาดยังคงเกิดขึ้นในทุกประเทศทั่วโลก จากรายงาน พบว่า ในปี 2023 ผู้ติดเชื้อเอชไอวีทั่วโลกประมาณ 39.9 ล้านคน ซึ่ง 65% ของผู้ติดเชื้ออยู่ในภูมิภาคแอฟริกา และมีผู้เสียชีวิตจากสาเหตุที่เกี่ยวข้องกับเอชไอวีประมาณ 630,000 คน และติดเชื้อเอชไอวีประมาณ 1.3 ล้านคน [1] สถานการณ์ในไทยจากรายงานของกองโรคเอดส์และโรคติดต่อทางเพศสัมพันธ์ กรมควบคุมโรค พบว่า ปี 2566 มีผู้ติดเชื่อที่ยังมีชีวิตอยู่ประมาณ 560,000 คน ในจำนวนนี้เป็นผู้ติดเชื้อรายใหม่ 9,200 คน กำลังรับยาต้านไวรัส 457,133 คน และพบผู้เสียชีวิตจากการติดเชื้อเอชไอวีประมาณ 1,100 คน [2] องค์การอนามัยโลก ได้กำหนดให้วันที่ 1 ธันวาคมของทุกปีเป็น “วันเอดส์โลก” เพื่อรณรงค์สร้างความตระหนักถึงความสำคัญของการป้องกัน โรคเอดส์ ยอมรับและเข้าใจผู้ติดเชื้อเอชไอวี และมีการกำหนดเป้าหมายในการยุติโรคเอดส์ได้ภายในปี 2573 สำหรับประเทศไทยได้กำหนดยุทธศาสตร์แห่งชาติว่าด้วยการยุติปัญหาเอดส์ พ.ศ. 2560–2573 มีเป้าหมายหลัก 3 ประการ คือ ลดการติดเชื้อเอชไอวี รายใหม่เหลือไม่เกิน 1,000 รายต่อปี ลดการเสียชีวิตในผู้ติดเชื้อเอชไอวี เหลือปีละไม่เกิน 4,000 ราย และลดการเลือกปฏิบัติอันเกี่ยวเนื่องจากเอชไอวี และเพศภาวะลงจากเดิมร้อยละ 90 [3]

วิธีป้องกันการติดเชื้อเอชไอวี นั่นคือใช้ถุงยางอนามัยทุกครั้งเมื่อมีเพศสัมพันธ์ เป็นวิธีหลักในการลดความเสี่ยงจากการติดเชื้อเอชไอวี และโรคติดต่อทางเพศสัมพันธ์อื่น ๆ หลีกเลี่ยงการใช้เข็มฉีดยาร่วมกัน หากจำเป็นต้องใช้เข็มฉีดยา ควรใช้ของใหม่ที่ผ่านการฆ่าเชื้อเท่านั้น ไม่ควรใช้เข็มซ้ำหรือใช้ร่วมกับผู้อื่น ไม่เปลี่ยนคู่นอนบ่อย และยาต้านเอชไอวีก่อนสัมผัสเชื้อที่เรียกว่า เพร็พ (PrEP) [4] โดย ยาเพร็พ (PrEP) ย่อมาจาก Pre-Exposure Prophylaxis คือยาที่ใช้ในการป้องกันเชื้อเอชไอวีก่อนการสัมผัสโรค โดยการรับประทานยาวันละหนึ่งเม็ดทุกวัน ยาเพร็พไม่สามารถป้องกันโรคติดต่อทางเพศสัมพันธ์อื่น ๆ ได้ จึงยังจำเป็นต้องใช้ถุงยางอนามัยในการป้องกันโรคติดต่อทางเพศสัมพันธ์อื่น ๆ และยาเพร็พสามารถป้องกันการติดเชื้อเอชไอวีได้ประมาณ 90 - 95% และหากผู้มารับบริการรับประทานทุกวันอย่างต่อเนื่อง และมีวินัยในการรับประทาน คือรับประทานในเวลาเดียวกันในทุก ๆ วัน และใช้ถุงยางอนามัยทุกครั้ง สามารถป้องกันโรคได้เกือบ 100% [5] ในปัจจุบันแนวทางการรักษาเอชไอวีจะเป็นการรักษาด้วยการทานยาต้านไวรัส (Antiretroviral Therapy, ART) หากผู้ป่วยพบเชื้อเร็ว-รักษาเร็ว มีวินัยการทานยาที่ดี และติดตามนัดพบแพทย์สม่ำเสมอ ทั้งนี้ในปัจจุบันยาต้านไวรัสยังไม่สามารถกำจัดเชื้อเอชไอวีให้ไปหมดไปจากร่างกายได้ ผู้มีเชื้อเอชไอวียังคงมีความจำเป็นต้องทานยาต้านไวรัสไปตลอดเพื่อทำให้ร่างกายแข็งแรงอยู่เสมอ [6]

เอชไอวี/เอดส์มีการระบาดอย่างแพร่หลายทั่วโลกจึงทำให้นักวิจัยทั่วโลกพยายามศึกษาการแพร่ระบาดของเอชไอวี/เอดส์ เพื่อกำหนดมาตรการต่างๆ อาทิเช่น Ayele (2021) [7] ได้ศึกษาการแพร่ระบาดของเอชไอวี/เอดส์ ใน Botswana ที่มีผลจากผู้ติดเชื้อที่หยุดรักษา ผลการศึกษาพบว่าถ้าผู้ติดเชื้อเข้ารับการรักษาหยุดการรักษามากกว่า 70% จะส่งผลให้เกิดการแพร่ระบาดของโรค แต่ถ้ากลุ่มผู้ติดเชื้อที่หยุดรักษากลับเข้ารับการรักษามากกว่า 80% จะสามารถควบคุมการแพร่ระบาดของโรคได้ Rana et al. (2024) [8] ได้พัฒนาและวิเคราะห์ตัวแบบคณิตศาสตร์ของ

การแพร่ระบาดของโรคเอดส์ที่มีผลจากการฟื้นตัวจากการรักษา ผลการศึกษาพบว่ากลยุทธ์การควบคุมโรคที่เหมาะสมควรรณรงค์ให้ประชากรเข้ารับการตรวจเลือดและเข้ารับการรักษาทันที เมื่อทราบผลเลือดเป็นบวก Moya et al. (2024) [9] ได้พัฒนาตัวแบบคณิตศาสตร์การแพร่เชื้อเอชไอวี/เอดส์ โดยพิจารณาการใช้โปรแกรมป้องกันก่อนการสัมผัสเชื้อด้วยยาเพร็พ ผลการศึกษาพบว่ารูปแบบการใช้ยาเพร็พที่แตกต่างกันส่งผลต่อจำนวนผู้ติดรายใหม่ ถ้ามีการใช้ยาเพร็พมากกว่า 35% ต่อวัน จะสามารถควบคุมการแพร่ระบาดของโรคได้ Adesola et al. (2024) [10] ได้พัฒนาตัวแบบคณิตศาสตร์เกี่ยวกับประสิทธิภาพและความสม่ำเสมอในการใช้ถุงยางอนามัยต่อการแพร่เชื้อเอชไอวี/เอดส์ ผลการศึกษาพบว่าการใช้ถุงยางอนามัยอย่างสม่ำเสมอในระหว่างการมีเพศสัมพันธ์นั้นสามารถป้องกันโรคได้มากกว่าการมีเพศสัมพันธ์โดยปราศจากการป้องกัน Meetei et al. (2024) [11] ได้พัฒนาตัวแบบคณิตศาสตร์การแพร่กระจายของเอชไอวี/เอดส์ โดยใช้ข้อมูลการแพร่ระบาดของของเอชไอวีแบบรายปีของประเทศปากีสถานในช่วงปี ค.ศ. 1992-2020 ผลการศึกษาพบว่าแนวทางป้องกันและควบคุมโรคในระยะยาว คือ การปรับเปลี่ยนพฤติกรรมทางเพศอย่างปลอดภัยด้วยการสวมถุงยาอนามัย และเข้ารับการรักษาโรคอย่างต่อเนื่อง และ Ogunmodimu et al. (2024) [12] ได้พัฒนาตัวแบบคณิตศาสตร์การแพร่ระบาดของเอชไอวี/เอดส์ในแอฟริกา โดยใช้เกาะเคปเวิร์ดเป็นกรณีศึกษา ในงานวิจัยนี้สนใจเกี่ยวกับการรักษาด้วยวิธี ART ที่ส่งผลให้ผู้ติดเชื้อที่เข้ารับการรักษากลายเป็นผู้ติดเชื้อที่อยู่ในสถานะไม่พบเชื้อไวรัสในร่างกาย (Undetectable = Untransmittable) ผลการศึกษาพบว่าการรักษาด้วยวิธี ART จะทำให้ระดับไวรัสของผู้ที่ติดเชื้อเอชไอวีเข้าสู่สถานะไม่พบเชื้อไวรัสในร่างกายและไม่สามารถแพร่กระจายเชื้อให้กับผู้อื่นได้ ซึ่งวิธีการนี้อาจทำให้เกาะเคปเวิร์ดบรรลุเป้าหมายของ UNAIDS ในการยุติการระบาดของโรคเอดส์ภายในปี 2030

จากที่กล่าวมาข้างต้น พบว่าเอชไอวี/เอดส์ ยังมีการแพร่ระบาดอย่างต่อเนื่อง ปัจจุบันบ่งชี้ว่ามีความจำเป็นที่จะต้องเน้นย้ำถึงการสวมถุงยางอนามัยร่วมกับการใช้ยาเพร็พ และการเข้ารับการรักษาด้วยวิธี ART เพื่อให้ผู้ติดเชื้ออยู่ในสถานะไม่พบเชื้อไวรัสในร่างกาย ดังนั้นในการศึกษาครั้งนี้ได้พัฒนาตัวแบบคณิตศาสตร์การแพร่ระบาดของเอชไอวี/เอดส์ที่มีผลจากประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พ และการรักษาด้วยวิธียาต้านไวรัส ART ผลการศึกษาที่ได้อาจเป็นแนวทางในการนำไปใช้ในการควบคุมการแพร่ระบาดของโรค และลดจำนวนผู้ติดเชื้อเอชไอวี/เอดส์ได้ในอนาคต

2 วิธีการดำเนินการวิจัย

ในการดำเนินการทำวิจัยมีวิธีการและขั้นตอนดังต่อไปนี้

2.1 ขั้นตอนศึกษาและเก็บรวบรวมข้อมูล

2.1.1 ศึกษาข้อมูลเกี่ยวกับการแพร่ระบาดของเชื้อไวรัสเอชไอวีและโรคเอดส์ ประกอบด้วย ลักษณะของเชื้อไวรัสเอชไอวีและโรคเอดส์ อาการและการแสดงอาการ สถานการณ์การแพร่ระบาด ปัจจัยที่มีผลต่อการแพร่ระบาดของเชื้อไวรัสเอชไอวี และแนวทางการป้องกันและการรักษา

2.1.2 ศึกษาเอกสารที่เกี่ยวข้องกับตัวแบบคณิตศาสตร์ ประกอบด้วย ความหมายของตัวแบบคณิตศาสตร์ ประเภทของตัวแบบคณิตศาสตร์สำหรับการระบาดของโรค รวมทั้งศึกษาทฤษฎีและบทนิยมที่เกี่ยวข้อง

2.1.3 ศึกษางานวิจัยที่เกี่ยวข้องกับเอชไอวี/เอดส์

2.2 ขั้นตอนการดำเนินงานวิจัย

2.2.1 นำข้อมูลจากการศึกษาค้นคว้าที่เก็บรวบรวมมาเกี่ยวกับเชื้อไวรัสเอชไอวีและโรคเอดส์ มาวิเคราะห์เพื่อหาปัจจัยที่ต้องการศึกษา ในการวิจัยครั้งนี้ได้ศึกษาพารามิเตอร์ที่ส่งผลต่อการแพร่ระบาดของเชื้อไวรัสเอชไอวีและโรคเอดส์ นั่นคือ ประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พและการรักษาด้วยวิธียาต้านไวรัส ART

2.2.2 พัฒนาตัวแบบคณิตศาสตร์การแพร่ระบาดของเชื้อเอชไอวีและโรคเอดส์ที่มีผลจากประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พและการรักษาด้วยวิธียาต้านไวรัส ART โดยแบ่งประชากรออกเป็น 5 กลุ่ม ประกอบด้วย ประชากรกลุ่มเสี่ยง (S) ประชากรกลุ่มติดเชื้อระยะเอชไอวี (H) ประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา (T) ประชากรกลุ่มระยะเอดส์ (A) และประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ (U) โดยแสดงรายละเอียดในหัวข้อ 3.1

3.2.3 ตรวจสอบความสมเหตุสมผลของตัวแบบคณิตศาสตร์โดยผู้เชี่ยวชาญ โดยนำตัวแบบคณิตศาสตร์ที่ผู้จัดทำสร้างขึ้น ให้ผู้เชี่ยวชาญตรวจสอบความสมเหตุสมผลของตัวแบบคณิตศาสตร์และชี้แนะข้อบกพร่องรวมทั้งข้อเสนอแนะในการปรับปรุงแก้ไข

3.2.4 วิเคราะห์ตัวแบบคณิตศาสตร์ ประกอบด้วย วิเคราะห์จุดสมดุล ค่าระดับการติดเชื้อ วิเคราะห์เสถียรภาพ วิเคราะห์ดัชนีความไว และวิเคราะห์ผลเชิงตัวเลข

2.3 ขั้นตอนสรุปผลการวิจัย

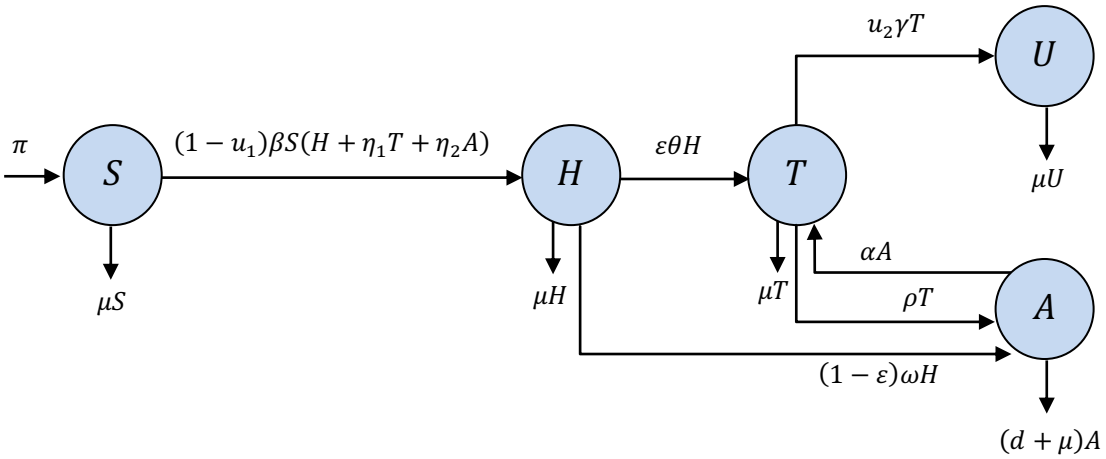
สรุปผลศึกษาจากตัวแบบคณิตศาสตร์ที่ได้พัฒนาขึ้นเพื่อเสนอแนวทางในการป้องกันและควบคุมการแพร่ระบาดของเชื้อเอชไอวีและโรคเอดส์

3 ผลการศึกษา

3.1 ผลการพัฒนาตัวแบบคณิตศาสตร์การแพร่กระจายของเชื้อเอชไอวีและโรคเอดส์

ผลการพัฒนาตัวแบบคณิตศาสตร์การแพร่กระจายของเชื้อเอชไอวีและโรคเอดส์ที่มีผลจากกประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พและการรักษาด้วยวิธียาต้านไวรัส ART โดยเสนอตัวแบบ $SHTAU$ ซึ่งแบ่งกลุ่มประกรออกเป็น 5 กลุ่ม ได้แก่ ประชากรกลุ่มเสี่ยง (S) คือกลุ่มบุคคลที่มีพฤติกรรมเสี่ยงต่อการติดเชื้อ ประชากรกลุ่มติดเชื้อระยะเอชไอวี (H) คือกลุ่มบุคคลที่ได้รับเชื้อเอชไอวีเข้าสู่ร่างกาย ไม่รู้สถานะการติดเชื้อและสามารถแพร่กระจายเชื้อได้ ประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา (T) คือกลุ่มบุคคลที่ติดเชื้อเอชไอวี รู้

สถานะการติดเชื้อและเข้ารับการรักษาด้วยยาต้านไวรัส ART และสามารถแพร่เชื้อได้ ประชากรกลุ่มระยะเอดส์ (A) คือกลุ่มบุคคลที่อยู่ในระยะสุดท้ายของการติดเชื้อเอชไอวี เสี่ยงต่อการติดเชื้อฉวยโอกาสและเสียชีวิตได้ ถ้าไม่เข้ารับการรักษาด้วยยาต้านไวรัส ART และประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ (U) คือกลุ่มบุคคลที่ติดเชื้อเอชไอวีที่เข้ารับการรักษาจนปริมาณไวรัสเอชไอวีในเลือดให้ต่ำลงจนไม่สามารถตรวจพบ และไม่สามารถแพร่เชื้อได้ โดยตัวแบบนี้ไม่พิจารณาประชากรแรกเกิดที่ติดเชื้อจากแม่สู่ลูก ซึ่งแสดงแผนภาพความสัมพันธ์ของตัวแบบคณิตศาสตร์ $SHTUA$ ที่ได้พัฒนาขึ้น ดังรูปที่ 1



รูปที่ 1: แผนภาพของตัวแบบคณิตศาสตร์ $SHTAU$ ที่พัฒนาขึ้น

จากรูปที่ 1 เป็นแผนภาพแสดงอัตราการเปลี่ยนแปลงของประชากรแต่ละกลุ่ม ณ ช่วงเวลาใดๆ โดยเริ่มจากประชากรเกิดใหม่เพิ่มขึ้นในประชากรกลุ่มเสี่ยง เมื่อประชากรกลุ่มเสี่ยงได้รับเชื้อจากประชากรกลุ่มติดเชื้อระยะเอชไอวีหรือได้รับเชื้อจากประชากรกลุ่มติดเชื้อที่อยู่ในระยะเข้ารับการรักษาหรือได้รับเชื้อจากประชากรกลุ่มติดเชื้อระยะเอดส์โดยมีการป้องกันการติดเชื้อด้วยการสวมถุงยางอนามัยร่วมกับการรับประทานยาเพร็บ ซึ่งประสิทธิภาพในการป้องกันโรคได้ประมาณ 90 - 95% ดังนั้นจะมีประชากรกลุ่มเสี่ยงกลายเป็นประชากรกลุ่มติดเชื้อระยะเอชไอวีได้ และประชากรกลุ่มติดเชื้อระยะเอชไอวีบางส่วนได้เข้ารับการตรวจคัดกรองเชื้อ เมื่อทราบสถานะการติดเชื้อจะเข้ารักษาด้วยยาต้านไวรัสและกลายเป็นประชากรกลุ่มติดเชื้อที่อยู่ในระยะเข้ารับการรักษา ส่วนประชากรกลุ่มติดเชื้อระยะเอชไอวีที่ไม่เข้ารับการตรวจคัดกรองและไม่ได้รับการรักษาจะกลายเป็นผู้ติดเชื้อระยะเอดส์และเสียชีวิตได้ ประชากรกลุ่มติดเชื้อระยะเอดส์ที่เข้ารับการรักษาด้วยยาต้านไวรัสจะกลายเป็นประชากรกลุ่มติดเชื้อที่อยู่ในระยะเข้ารับการรักษา และประชากรกลุ่มติดเชื้อที่อยู่ในระยะเข้ารับการรักษาเมื่อตรวจปริมาณไวรัสเอชไอวีในเลือดให้ต่ำลงจนไม่สามารถตรวจพบจะกลายเป็นประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ นอกจากนี้ประชากรทุกกลุ่มยังสามารถเสียชีวิตตามธรรมชาติด้วยอัตราที่เท่ากัน จากอัตราการเปลี่ยนแปลงของประชากรแต่ละกลุ่มจากรูปที่ 1 สามารถอธิบายในระบบสมการเชิงอนุพันธ์ไม่เชิงเส้นได้ดังนี้

$$\frac{dS}{dt} = \pi - (1-u_1)\beta S(H + \eta_1 T + \eta_2 A) - \mu S \quad (3.1)$$

$$\frac{dH}{dt} = (1-u_1)\beta S(H + \eta_1 T + \eta_2 A) - (\varepsilon\theta + (1-\varepsilon)\omega + \mu)H \quad (3.2)$$

$$\frac{dT}{dt} = \varepsilon\theta H + \alpha A - (u_2\gamma + \rho + \mu)T \quad (3.3)$$

$$\frac{dA}{dt} = (1-\varepsilon)\omega H + \rho T - (\alpha + d + \mu)A \quad (3.4)$$

$$\frac{dU}{dt} = u_2\gamma T - \mu U \quad (3.5)$$

โดยที่ $N(t)$ แทนผลรวมประชากร ณ เวลา t ใดๆ ซึ่ง $N = S + H + T + A + U$ และพารามิเตอร์ของระบบสมการข้างต้นมีค่าเป็นบวกทุกค่า และ $u_1, \varepsilon < 1$ และมีความหมายดังนี้ $\boldsymbol{\pi}$ คืออัตราการเกิด, β คืออัตราการติดเชื้อจากการมีเพศสัมพันธ์, u_1 คือประสิทธิภาพของการป้องกันโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ, η_1 คือสัดส่วนการติดเชื้อจากประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา, η_2 คือสัดส่วนการติดเชื้อของประชากรกลุ่มติดเชื้อระยะเอดส์, θ คืออัตราการเข้ารับการตรวจเลือดและเข้ารับการรักษา, ε คือสัดส่วนของการเข้ารับการตรวจเลือดและเข้ารับการรักษา, ω คืออัตราการเปลี่ยนสถานะจากระยะเอชไอวีเป็นระยะเอดส์, ρ คืออัตราการล้มเหลวของการรักษา, α คืออัตราของการเข้ารับการรักษาของประชากรกลุ่มติดเชื้อระยะเอดส์, d คืออัตราการเสียชีวิตจากโรค, γ คืออัตราการเปลี่ยนสถานะจากประชากรกลุ่มติดเชื้อที่อยู่ในระยะเข้ารับการรักษาเป็นประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จเนื่องจากรับประทานยาต้านไวรัส ART อย่างต่อเนื่อง, u_2 คือประสิทธิภาพของการรักษาด้วยยาด้านไวรัส ART และ μ คืออัตราการตายโดยธรรมชาติ

3.2 ผลการวิเคราะห์ตัวแบบคณิตศาสตร์

ในหัวข้อนี้จะนำเสนอผลการวิเคราะห์ตัวแบบคณิตศาสตร์เป็นการวิเคราะห์ด้วยวิธีมาตรฐาน (Standard method) [13] โดยการหาจุดสมดุล (Equilibrium point) การหาค่าระดับการติดเชื้อ (Basic reproductive number) และการวิเคราะห์เสถียรภาพเชิงเส้นกำกับ (Local asymptotically stable)

3.2.1 จุดสมดุล

ผลการหาค่าจุดสมดุลของระบบสมการ (3.1) - (3.5) โดยกำหนดให้ระบบสมการ เท่ากับ 0 ทำให้ได้จุดสมดุล 2 จุด ได้แก่ จุดสมดุลที่ไม่มีการแพร่ระบาด คือ $E_0 = \left(\frac{\pi}{\mu}, 0, 0, 0, 0\right)$ และจุดสมดุลที่มีการแพร่ระบาด คือ $E_1 = (S^*, H^*, T^*, U^*, A^*)$ โดยที่ $S^* = \frac{\pi}{\mu\Re_0}$, $T^* = \frac{(\xi_5\varepsilon\theta + \xi_4\alpha)H^*}{\xi_6}$, $A^* = \frac{(\xi_3\xi_4 + \rho\varepsilon\theta)H^*}{\xi_6}$, $U^* = \frac{u_2\gamma(\xi_5\varepsilon\theta + \xi_4\alpha)H^*}{\mu\xi_6}$ และ $H^* = \frac{\mu\xi_6(\Re_0 - 1)}{\xi_1(\xi_6 + \eta_1(\xi_5\varepsilon\theta + \xi_4\alpha) + \eta_2(\xi_3\xi_4 + \rho\varepsilon\theta))}$ เมื่อ $\Re_0 = \frac{\pi\xi_1(\xi_6 + \eta_1(\xi_5\varepsilon\theta + \xi_4\alpha) + \eta_2(\xi_3\xi_4 + \rho\varepsilon\theta))}{\mu\xi_2\xi_6}$ และกำหนดให้ $\xi_1 = (1-u_1)\beta$, $\xi_2 = \varepsilon\theta + (1-\varepsilon)\omega + \mu$, $\xi_3 = u_2\gamma + \rho + \mu$, $\xi_4 = (1-\varepsilon)\omega$, $\xi_5 = \alpha + d + \mu$ และ $\xi_6 = (u_2\gamma + \mu)\xi_5 + \rho(d + \mu)$ ซึ่งจะเห็นได้ว่า ถ้า $\Re_0 > 1$ แล้ว $S^*, H^*, T^*, A^*, U^* > 0$ นั่น

คือมีจุดสมดุลที่มีการแพร่ระบาด (E_1) ถ้า $\mathfrak{R}_0 = 1$ นั่นคือจะมีจุดสมดุลเพียงจุดสมดุลเดียว คือ จุดสมดุลที่ไม่มีการแพร่ระบาด (E_0) และถ้า $\mathfrak{R}_0 < 1$ จะไม่มีจุดสมดุลที่มีการแพร่ระบาด เนื่องจากจำนวนประชากรกลุ่มติดเชื้อระยะเอชไอวีมีค่าเป็นลบ ($H^* < 0$) และค่า $\mathfrak{R}_0$ ที่ได้จากการวิเคราะห์จุดสมดุลมีค่าเท่ากับค่าระดับการติดเชื้อที่ได้จากการหาค่าด้วยวิธีการ Next generation matrix ซึ่งแสดงวิธีการหาค่าในหัวข้อถัดไป

3.2.3 ค่าระดับการติดเชื้อ

ค่าระดับการติดเชื้อ ($\mathfrak{R}_0$) เป็นค่าเฉลี่ยของจำนวนผู้ติดเชื้อในรุ่นที่ 2 ที่ติดเชื้อจากผู้ป่วยรุ่นแรก 1 คน ในการหาค่า $\mathfrak{R}_0$ จะใช้วิธีการ Next generation matrix [14] โดยพิจารณาเฉพาะประชากลุ่มติดเชื้อที่สามารถแพร่เชื้อได้ ประกอบด้วย ประชากรกลุ่มติดเชื้อระยะเอชไอวี (H) ประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา (T) และประชากรกลุ่มระยะเอดส์ (A) ดังนั้นจะพิจารณาเฉพาะสมการที่ (3.2) - (3.4) โดยมีขั้นตอนดังนี้

ขั้นที่ 1 จัดรูประบบสมการ (3.2) - (3.4) ให้อยู่ในรูปเมทริกซ์ $\frac{dX}{dt} = F(X) - V(X)$ โดยที่ X คือเมทริกซ์ของตัวแปร $F(X)$ คือเมทริกซ์ที่ทำให้เกิดผู้ติดเชื้อรายใหม่ และ $V(X)$ คือเมทริกซ์ที่เปลี่ยนสถานะจากกลุ่มหนึ่งไปยังอีกกลุ่มหนึ่ง จะได้

$$X = \begin{bmatrix} H \\ T \\ A \end{bmatrix}, \quad F(X) = \begin{bmatrix} \xi_1 S(H + \eta_1 T + \eta_2 A) \\ 0 \\ 0 \end{bmatrix} \quad \text{และ} \quad V(X) = \begin{bmatrix} \xi_2 H \\ -\varepsilon\theta H - \alpha A + \xi_3 T \\ -\xi_4 H - \rho T + \xi_5 A \end{bmatrix}$$

ขั้นที่ 2 หาจาโคเบียนเมทริกซ์สำหรับเมทริกซ์ F และเมทริกซ์ V โดยการหาอนุพันธ์ย่อย $F(X)$ และ $V(X)$ เทียบ X จะได้เมทริกซ์ดังนี้

$$F = \begin{bmatrix} \xi_1 S & \xi_1 \eta_1 S & \xi_1 \eta_2 S \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{และ} \quad V = \begin{bmatrix} \xi_2 & 0 & 0 \\ -\varepsilon\theta & \xi_3 & -\alpha \\ -\xi_4 & -\rho & \xi_5 \end{bmatrix}$$

แทนค่า $E_0 = \left(\frac{\pi}{\mu}, 0, 0, 0, 0\right)$ ในจาโคเบียนเมทริกซ์ F และเมทริกซ์ V ตามลำดับ จะได้

$$F_{E_0} = \begin{bmatrix} \frac{\xi_1 \pi}{\mu} & \frac{\xi_1 \eta_1 \pi}{\mu} & \frac{\xi_1 \eta_2 \pi}{\mu} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{และ} \quad V_{E_0} = \begin{bmatrix} \xi_2 & 0 & 0 \\ -\varepsilon\theta & \xi_3 & -\alpha \\ -\xi_4 & -\rho & \xi_5 \end{bmatrix}$$

ขั้นที่ 3 หาค่า $\mathfrak{R}_0$ จากค่ารัศมีสเปกตรัมของ ${F_{E_0}}{V_{E_0}}^{-1}$ นั่นคือ $\mathfrak{R}_0 = \rho\left({F_{E_0}}{V_{E_0}}^{-1}\right)$ จากสมการลักษณะเฉพาะ $det\left({F_{E_0}}{V_{E_0}}^{-1} - \lambda I\right) = 0$ โดยที่ I คือเมทริกซ์เอกลักษณ์ มิติ 3x3 จะได้

$${F_{E_0}}{V_{E_0}}^{-1} = \begin{bmatrix} \frac{\pi\xi_1\left(\xi_6 + \eta_1(\xi_5\varepsilon\theta + \xi_4\alpha) + \eta_2(\xi_3\xi_4 + \rho\varepsilon\theta)\right)}{\mu\xi_2\xi_6} & \frac{\pi\xi_1(\xi_5\eta_1 + \eta_2\rho)}{\mu\xi_6} & \frac{\pi\xi_1(\eta_1\alpha + \eta_2\xi_3)}{\mu\xi_6} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

และค่าลักษณะเฉพาะของ $F_{E_0}V_{E_0}{}^{-1}$ คือ $\lambda_1 = \frac{\pi\xi_1(\xi_6+\eta_1(\xi_5\varepsilon\theta+\xi_4\alpha)+\eta_2(\xi_3\xi_4+\rho\varepsilon\theta))}{\mu\xi_2\xi_6}$ และ $\lambda_{2,3} = 0$ จะได้ว่า $\mathfrak{R}_0 = \rho(F_{E_0}V_{E_0}{}^{-1}) = max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}$ ดังนั้น

$$\mathfrak{R}_0 = \frac{\pi\xi_1(\xi_6 + \eta_1(\xi_5\varepsilon\theta + \xi_4\alpha) + \eta_2(\xi_3\xi_4 + \rho\varepsilon\theta))}{\mu\xi_2\xi_6}$$

จากค่า $\mathfrak{R}_0$ ที่ได้จากการคำนวณด้วยวิธี Next generation matrix มีค่าเท่ากับ $\mathfrak{R}_0$ ที่ได้จากการหาค่าจุดสมดุลในหัวข้อ 3.2.1 จะได้ว่า ค่า $\mathfrak{R}_0$ มีความสัมพันธ์กับจุดสมดุล ถ้า $\mathfrak{R}_0 > 1$ จะมีจุดสมดุลที่มีการแพร่ระบาด (E_1) นั่นคือ $S^*,\ H^*,\ T^*,\ A^*,\ U^* > 0$ และเมื่อแทนค่า ξ_i โดยที่ $i = 1,2,3,4,5,6$ ในค่า $\mathfrak{R}_0$ พบว่าพารามิเตอร์ทุกตัวในตัวแบบ $SHTUA$ มีผลต่อการเปลี่ยนแปลงของค่า $\mathfrak{R}_0$

3.2.4 การวิเคราะห์เสถียรภาพเชิงเส้นกำกับ

จากการหาค่าจุดสมดุลของระบบสมการ (3.1) - (3.5) พบจุดสมดุล 2 จุด คือ จุดสมดุลที่ไม่มีการแพร่ระบาด และจุดสมดุลที่มีการแพร่ระบาด ดังนั้นในหัวข้อนี้จะแสดงเงื่อนไขเสถียรภาพเชิงเส้นกำกับของจุดสมดุลทั้ง 2 จุด มีขั้นตอน ดังนี้

ในการวิเคราะห์หาเงื่อนไขเสถียรภาพเชิงเส้นกำกับของแต่ละจุดสมดุลทำได้โดยการสร้างจาโคเบียนเมทริกซ์ (J) ของสมการ (3.1) - (3.5) ผลดังนี้

$$J = \begin{bmatrix} -\xi_1(H+\eta_1T+\eta_2A)-\mu & -\xi_1S & -\xi_1\eta_1S & -\xi_1\eta_2S & 0 \\ \xi_1(H+\eta_1T+\eta_2A) & \xi_1S-\xi_2 & \xi_1\eta_1S & \xi_1\eta_2S & 0 \\ 0 & \varepsilon\theta & -\xi_3 & \alpha & 0 \\ 0 & \xi_4 & \rho & -\xi_5 & 0 \\ 0 & 0 & u_2\gamma & 0 & -\mu \end{bmatrix} \quad (3.6)$$

จากนั้นให้แทนจุดสมดุลในจาโคเบียนเมทริกซ์ (3.6) และหาค่าลักษณะเฉพาะ (λ) จากสมการลักษณะเฉพาะ $det(J-\lambda I) = 0$ โดยที่ I คือเมทริกซ์เอกลักษณ์ มิติ 5x5 และถ้าส่วนจริงของค่าลักษณะเฉพาะทุกค่าเป็นลบ จะได้ว่าจุดสมดุลดังกล่าวเสถียรภาพเชิงเส้นกำกับ และลำดับถัดไปจะแสดงเงื่อนไขการเสถียรภาพของแต่ละจุดสมดุล

ทฤษฎีบท 3.1 ถ้า $\mathfrak{R}_0 < 1$ และ $R_i < 1$ เมื่อ $i = 1,2$ แล้วจุดสมดุลที่ไม่มีการแพร่ระบาด (E_0) มีความเสถียรภาพเชิงเส้นกำกับ โดยที่ $R_1 = \frac{\xi_1\pi}{\mu(\xi_2+\xi_3+\xi_5)}$ และ $R_2 = \frac{a_1\xi_1\pi(\xi_3+\xi_5+\eta_1\varepsilon\theta+\eta_2\xi_4)+\mu a_3}{\mu}$

พิสูจน์. จากจาโคเบียนเมทริกซ์ (3.6) แทนจุดสมดุล E_0 จะได้

$$J_0 = \begin{bmatrix} -\mu & -\frac{\xi_1\pi}{\mu} & -\frac{\xi_1\eta_1\pi}{\mu} & -\frac{\xi_1\eta_2\pi}{\mu} & 0 \\ 0 & \frac{\xi_1\pi}{\mu} - \xi_2 & \frac{\xi_1\eta_1\pi}{\mu} & \frac{\xi_1\eta_2\pi}{\mu} & 0 \\ 0 & \varepsilon\theta & -\xi_3 & \alpha & 0 \\ 0 & \xi_4 & \rho & -\xi_5 & 0 \\ 0 & 0 & u_2\gamma & 0 & -\mu \end{bmatrix} \tag{3.6}$$

จากจาโคเบียนเมทริกซ์ (3.6) จะได้ค่าลักษณะเฉพาะ จาก $det(J_0 - \lambda I) = 0$ ดังนี้

$$(-\mu - \lambda)(-\mu - \lambda)(\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3) = 0 \tag{3.7}$$

โดยที่ $a_1 = (\xi_2 + \xi_3 + \xi_5)(1 - R_1)$, $a_2 = \xi_2(\xi_3 + \xi_5) + \xi_6 - \frac{\xi_1\pi(\xi_3+\xi_5+\eta_1\varepsilon\theta+\eta_2\xi_4)}{\mu}$,
$a_3 = \mu\xi_2\xi_6(1 - \Re_0)$ และ $R_1 = \frac{\xi_1\pi}{\mu(\xi_2+\xi_3+\xi_5)}$

จากสมการ (3.7) จะได้ $\lambda_{1,2} = -\mu < 0$ และ $\lambda_{3,4,5}$ เป็นคำตอบของสมการ $\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$ และจะได้ว่า $\lambda_{3,4,5}$ มีส่วนจริงเป็นลบทุกค่า เมื่อสอดคล้องกับเงื่อนไข Routh-Hurwitz สำหรับ $n = 3$ [15] ซึ่งมีผลพิจารณาดังนี้

1) $a_1 = (\xi_2 + \xi_3 + \xi_5)(1 - R_1)$ จะได้ว่า $a_1 > 0$ เมื่อ $R_1 < 1$

2) $a_3 = \xi_2\xi_6(1 - \Re_0)$ จะได้ว่า $a_3 > 0$ เมื่อ $\Re_0 < 1$

3) พิจารณา $a_1a_2 - a_3 > 0$ และจาก $a_1 > 0$ เมื่อ $R_1 < 1$ และ $a_3 > 0$ เมื่อ $\Re_0 < 1$ จะได้

$$\begin{aligned} a_1a_2 - a_3 &= a_1\left(\xi_2(\xi_3 + \xi_5) + \xi_6 - \frac{\xi_1\pi(\xi_3 + \xi_5 + \eta_1\varepsilon\theta + \eta_2\xi_4)}{\mu}\right) - a_3 \\ &= a_1(\xi_2(\xi_3 + \xi_5) + \xi_6) - \frac{a_1\xi_1\pi(\xi_3+\xi_5+\eta_1\varepsilon\theta+\eta_2\xi_4) + \mu a_3}{\mu} \\ &= a_1(\xi_2(\xi_3 + \xi_5) + \xi_6)(1 - R_2) \end{aligned}$$

จะได้ว่า $a_1a_2 - a_3 > 0$ เมื่อ $\Re_0 < 1$ และ $R_{1,2} < 1$ โดยที่ $R_2 = \frac{a_1\xi_1\pi(\xi_3+\xi_5+\eta_1\varepsilon\theta+\eta_2\xi_4) + \mu a_3}{\mu}$

ดังนั้นจึงสรุปได้ว่าค่าลักษณะเฉพาะทุกค่ามีส่วนจริงเป็นลบ ถ้า $\Re_0 < 1$ และ $R_i < 1$ เมื่อ $i = 1,2$ แล้วจุดสมดุลที่ไม่มีการแพร่ระบาดของโรค (E_0) มีความเสถียรภาพเชิงเส้นกำกับ □

ทฤษฎีบท 3.2 ถ้า $\Re_0 > 1$, $R_i > 1$ เมื่อ $i = 3,4,5$ และ $K_j > 1$ เมื่อ $j = 1,2$ แล้วจุดสมดุลที่มีการแพร่ระบาด (E_1) มีความเสถียรภาพเชิงเส้นกำกับ โดยที่ $R_3 = \frac{\mu\Re_0(\mu\Re_0+\xi_2+\xi_3+\xi_5)}{\xi_1\pi}$,
$R_4 = \frac{\mu\Re_0((\xi_2+\xi_3+\xi_5)\mu\Re_0+\xi_2(\xi_3+\xi_5)+\xi_6)}{m_1\xi_1\pi}$, $R_5 = \frac{\mu\Re_0((\xi_2(\xi_3+\xi_5)+\xi_6)\mu\Re_0+\xi_2\xi_6)}{m_2\xi_1\pi}$, $K_1 = \frac{b_1b_2}{b_3}$ และ $K_2 = \frac{G_1}{G_2}$

พิสูจน์. จากจาโคเบียนเมทริกซ์ (3.6) แทนจุดสมดุล E_1 จะได้สมการลักษณะเฉพาะ ดังนี้

$$(-\mu-\lambda)(\lambda^4+b_1\lambda^3+b_2\lambda^2+b_3\lambda+b_4)=0 \tag{3.8}$$

โดยที่ $b_1=\frac{\xi_1\pi(R_3-1)}{\mu\Re_0}$, $b_2=\frac{m_1\xi_1\pi(R_4-1)}{\mu\Re_0}$, $b_3=\frac{m_2\xi_1\pi(R_5-1)}{\mu\Re_0}$, $b_4=\mu\xi_2\xi_6(\Re_0-1)$,

$m_1=\eta_1\varepsilon\theta+\eta_2\xi_4+\xi_3+\xi_5+\mu$, $m_2=\xi_6+\mu(\xi_3+\xi_5)+(\varepsilon\theta(\xi_1+\mu\xi_5)+\alpha\xi_4)\eta_1$

$+(\varepsilon\rho\theta+(\xi_3+\mu)\xi_4)\eta_2$, $R_3=\frac{\mu\Re_0(\mu\Re_0+\xi_2+\xi_3+\xi_5)}{\xi_1\pi}$, $R_4=\frac{\mu\Re_0((\xi_2+\xi_3+\xi_5)\mu\Re_0+\xi_2(\xi_3+\xi_5)+\xi_6)}{m_1\xi_1\pi}$

และ $R_5=\frac{\mu\Re_0((\xi_2(\xi_3+\xi_5)+\xi_6)\mu\Re_0+\xi_2\xi_6)}{m_2\xi_1\pi}$

จากสมการ (3.8) จะได้ $\lambda_1=-\mu<0$ และ λ_i เมื่อ $i=2,3,4,5$ เป็นคำตอบของสมการ $\lambda^4+b_1\lambda^3+b_2\lambda^2+b_3\lambda+b_4=0$ และจะได้ว่าค่าลักษณะเฉพาะมีส่วนจริงเป็นลบทุกค่า เมื่อสอดคล้องกับเงื่อนไข Routh-Hurwitz สำหรับ $n=4$ [16] ซึ่งมีผลพิจารณาดังนี้

1) $b_1=\frac{\xi_1\pi(R_3-1)}{\mu\Re_0}>0$ เมื่อ $R_3>1$

2) $b_4=\mu\xi_2\xi_6(\Re_0-1)>0$ เมื่อ $\Re_0>1$

3) พิจารณา $b_1b_2-b_3>0$

จาก $b_1>0$, $b_2>0$ และ $b_3>0$ เมื่อ $R_3>1$, $R_4>1$ และ $R_5>1$ ตามลำดับ

จะได้ว่า $b_1b_2-b_3=b_3(K_1-1)>0$ เมื่อ $K_1>1$ โดยที่ $K_1=\frac{b_1b_2}{b_3}$

4) พิจารณา $b_1b_2b_3-\left({b_3}^2+{b_1}^2b_4\right)>0$

จาก $b_1>0$, $b_2>$, $b_3>0$ และ $b_4>0$ เมื่อ $R_3>1$, $R_4>1, R_5>1$ และ $\Re_0>1$ ตามลำดับ

ให้ $G_1=b_1b_2b_3>0$ และ $G_2={b_3}^2+{b_1}^2b_4$ จะได้ว่า $G_1-G_2=G_2(K_2-1)>0$ เมื่อ $K_2>1$ โดยที่ $K_2=\frac{G_1}{G_2}$

จากผลการการพิจารณาค่าลักษณะเฉพาะตามเงื่อนไขของ Routh-Hurwitz สำหรับ $n=4$ [15] จึงสรุปได้ว่าค่าลักษณะเฉพาะทุกค่ามีส่วนจริงเป็นลบ ถ้า $\Re_0>1$, $R_i>1$ เมื่อ $i=3,4,5$ และ $K_j>1$ เมื่อ $j=1,2$ แล้วจุดสมดุลที่มีการแพร่ระบาด (E_1) มีความเสถียรภาพเชิงเส้นกำกับ □

3.3 ผลเชิงตัวเลข

ในหัวข้อนี้จะนำเสนอผลเชิงตัวเลขเพื่อตรวจสอบเงื่อนไขความเสถียรภาพเชิงเส้นกำกับของแต่ละจุดสมดุล และทำการวิเคราะห์ค่าความไวของพารามิเตอร์ที่มีผลต่อค่าระดับการติดเชื้อ นอกจากนี้จะเสนอผลการศึกษาประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ (u_1) และประสิทธิภาพของการรักษาด้วยยาต้านไวรัส ART (u_2) ต่อการป้องกันและควบคุมโรค โดยนำค่าพารามิเตอร์ที่ได้จากการศึกษาข้อมูลการแพร่ระบาดของเชื้อเอชไอวี/โรคเอดส์และจากงานวิจัยที่เกี่ยวข้อง ดังนี้

ตารางที่ 1: ความหมายของพารามิเตอร์ในตัวแบบ $SHTUA$

พารามิเตอร์	สัญลักษณ์	ค่าพารามิเตอร์	หน่วย	อ้างอิง
$\boldsymbol{\pi}$	จำนวนคนเกิดใหม่	1,432	คนต่อวัน	[16]
$\boldsymbol{\theta}$	อัตราการเข้ารับการตรวจเลือดและเข้ารับการรักษา	1/7	ต่อวัน	[17]
$\boldsymbol{\omega}$	อัตราการเปลี่ยนสถานะจากจากกลุ่ม H เป็นกลุ่ม A	1/(10x365)	ต่อวัน	[17]
$\boldsymbol{\alpha}$	อัตราของการเข้ารับการรักษาของกลุ่ม A	1/30	ต่อวัน	[18]
$\boldsymbol{\gamma}$	อัตราการเปลี่ยนสถานะจากกลุ่ม T เป็นกลุ่ม U	1/(12x30)	ต่อวัน	[17]
$\boldsymbol{\rho}$	อัตราการล้มเหลวของการรักษา	1/(2x365)	ต่อวัน	[18]
$\boldsymbol{d}$	อัตราการเสียชีวิตจากโรค	1/(12x30)	ต่อวัน	[17]
$\boldsymbol{\mu}$	อัตราการตายโดยธรรมชาติ	1/(77x365)	ต่อวัน	[16]
$\boldsymbol{\varepsilon}$	สัดส่วนการเข้ารับการตรวจเลือดและเข้ารับการรักษา	0.02	-	[8]
$\boldsymbol{\eta_1, \eta_2}$	สัดส่วนการติดเชื้อจากกลุ่ม T และกลุ่ม A	0.30, 0.25	-	[8]
$\boldsymbol{\beta}$	อัตราการติดเชื้อจากการมีเพศสัมพันธ์	$1.2\text{x}10^{-9}$	ต่อวัน	-
$\boldsymbol{u_1}$	ประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ	0.9 - 1	-	[5]
$\boldsymbol{u_2}$	ประสิทธิภาพของการรักษาด้วยยาต้านไวรัส ART	0.861	-	[19]

3.3.1 ผลการวิเคราะห์เสถียรภาพเชิงเส้นกำกับ

สำหรับผลเชิงตัวเลขสำหรับการตรวจสอบเงื่อนไขเสถียรภาพเชิงเส้นกำกับของจุดสมดุลที่ไม่มีการแพร่ระบาดและจุดสมดุลที่มีการแพร่ระบาด จะใช้ค่าพารามิเตอร์ในตารางที่ 1 และสมมติค่าเริ่มต้นของประชากรแต่ละกลุ่ม $S(0) = 2 \times 10^7,\ H(0) = 2.35 \times 10^5,\ T(0) = 2.77 \times 10^5,\ U(0) = 1.22 \times 10^4$ และ $A(0) = 1.94 \times 10^7$ โดยกำหนดให้ u_1 แตกต่างกัน คือ $u_1 = 0.9,\ 0.92,\ 0.94$ และ 0.96

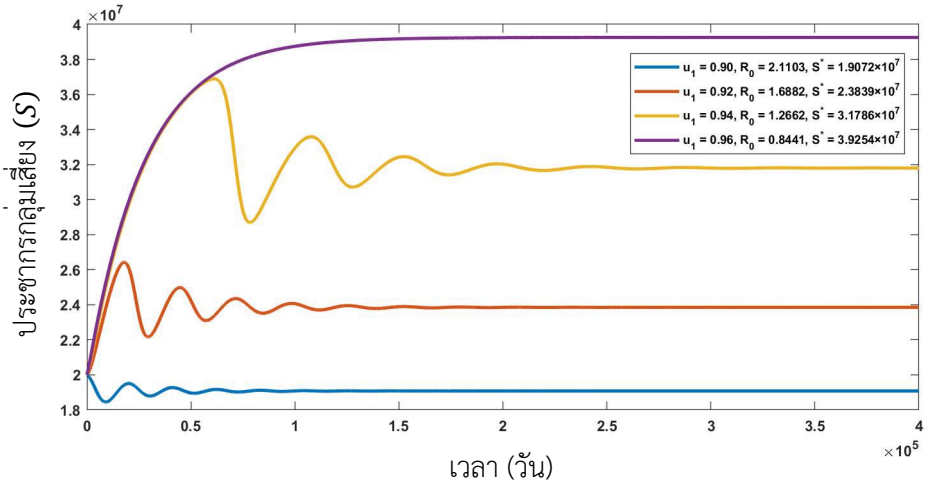
ตารางที่ 2: ผลเชิงตัวเลขของเงื่อนไขเสถียรภาพของจุดสมดุลที่มีการแพร่ระบาด

u_1	$\Re_0$	R_3	R_4	R_5	K_1	K_2	b_1	b_4	$b_1b_2 - b_3$	$b_1b_2b_3 - (b_3^2 + b_1^2b_4)$
0.9	2.1103	18.85	2.36	1.37	61.97	16.72	0.04	1.14×10^{-11}	0.5×10^{-5}	0.41×10^{-13}
0.92	1.6882	18.84	2.35	1.36	64.13	22.78	0.04	7.07×10^{-12}	0.5×10^{-5}	0.40×10^{-13}
0.94	1.2661	18.83	2.34	1.34	66.48	37.73	0.04	2.74×10^{-12}	0.5×10^{-5}	3.87×10^{-13}
0.96	0.8441	18.83	2.34	1.33	69.03	134.34	0.04	-1.60×10^{-12}	0.5×10^{-5}	3.76×10^{-13}

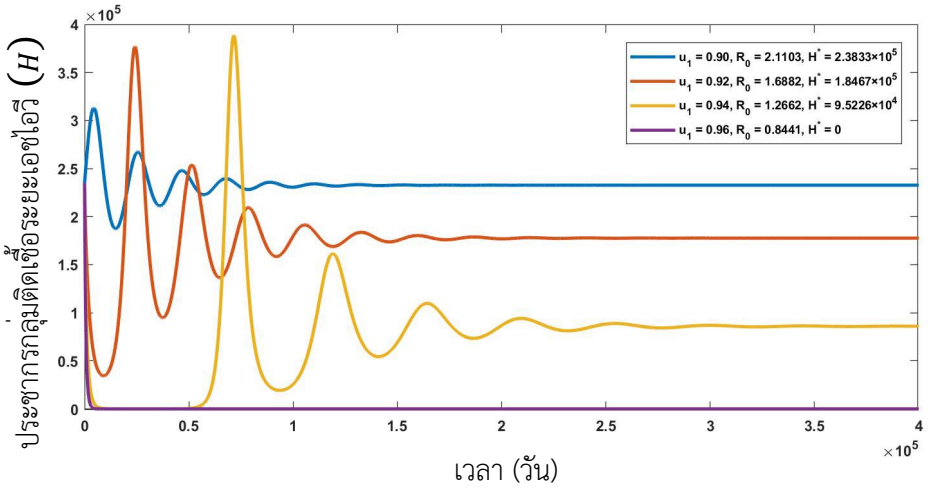
ตารางที่ 3: ผลเชิงตัวเลขของเงื่อนไขเสถียรภาพของจุดสมดุลที่ไม่มีการแพร่ระบาด

u_1	$\mathfrak{R}_0$	R_1	R_2	a_1	a_3	$a_1a_2 - a_3$
0.9	2.1103	0.112	0.75×10^{-5}	0.038	-1.14×10^{-11}	0.78×10^{-8}
0.92	1.6882	0.089	0.62×10^{-5}	0.039	-7.08×10^{-12}	0.23×10^{-5}
0.94	1.2661	0.067	0.47×10^{-5}	0.040	-2.74×10^{-12}	0.39×10^{-5}
0.96	0.8441	0.045	0.32×10^{-5}	0.041	1.60×10^{-12}	0.57×10^{-5}

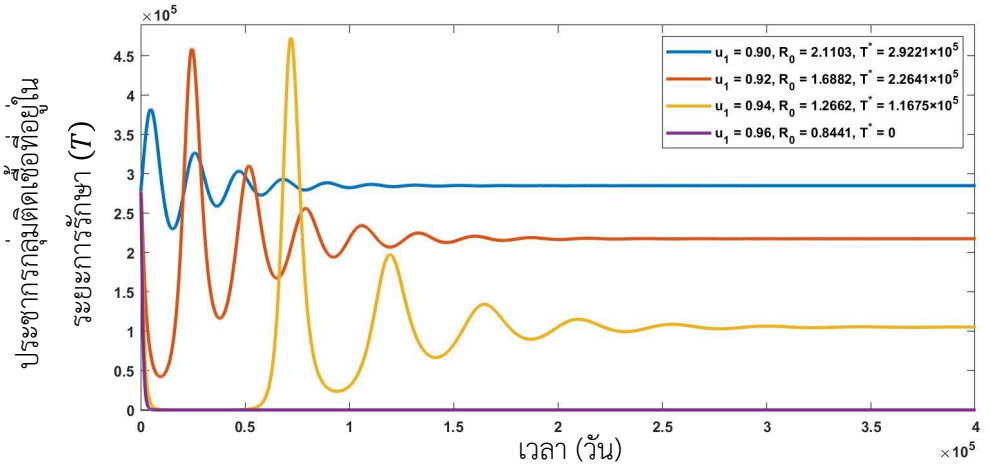
เมื่อตรวจสอบผลเชิงตัวเลขของเงื่อนไขเสถียรภาพของจุดสมดุลที่ไม่มีการแพร่ระบาด (E_0) และจุดสมดุลที่มีการแพร่ระบาด (E_1) พบว่า ตารางที่ 2 มี $u_1 = 0.9,\ 0.92$ และ 0.94 ที่ทำให้ $\mathfrak{R}_0 > 1,\ R_i > 1$ เมื่อ $i = 3,4,5$ และ $K_j > 1$ เมื่อ $j = 1,2$ ซึ่งสอดคล้องกับเงื่อนไขเสถียรภาพของจุดสมดุล E_1 นั่นคือมีการแพร่ระบาดของโรค ดังนั้นระบบจะลู่เข้าจุดสมดุลที่มีการแพร่ระบาด คือ $E_1 = (1.9072 \times 10^7,\ 2.3833 \times 10^5,\ 2.9221 \times 10^5,\ 1.2857 \times 10^4,\ 1.9641 \times 10^7)$, $E_1 = (2.3839 \times 10^7,\ 1.8467 \times 10^5,\ 2.2641 \times 10^5,\ 9.9625 \times 10^3,\ 1.5219 \times 10^7)$ และ $E_1 = (3.1786 \times 10^7,\ 9.5226 \times 10^4,\ 1.1675 \times 10^5,\ 5.1373 \times 10^3,\ 7.8478 \times 10^6)$ ตามลำดับ และตารางที่ 3 มี $u_1 = 0.96$ ที่ทำให้ $\mathfrak{R}_0 < 1$ และ $R_i < 1$ เมื่อ $i = 1,2$ ซึ่งสอดคล้องกับเงื่อนไขเสถียรภาพของจุดสมดุล E_0 ดังนั้นระบบจะลู่เข้าจุดสมดุลที่ไม่มีการแพร่ระบาด คือ $E_0 = (3.9254 \times 10^7, 0, 0, 0, 0)$ นั่นคือการแพร่ระบาดของโรคจะลดลงและหมดไปในอนาคต และแสดงผลการลู่เข้าคำตอบของแต่ละกลุ่มประชากรได้ดังรูปที่ 2 – 6



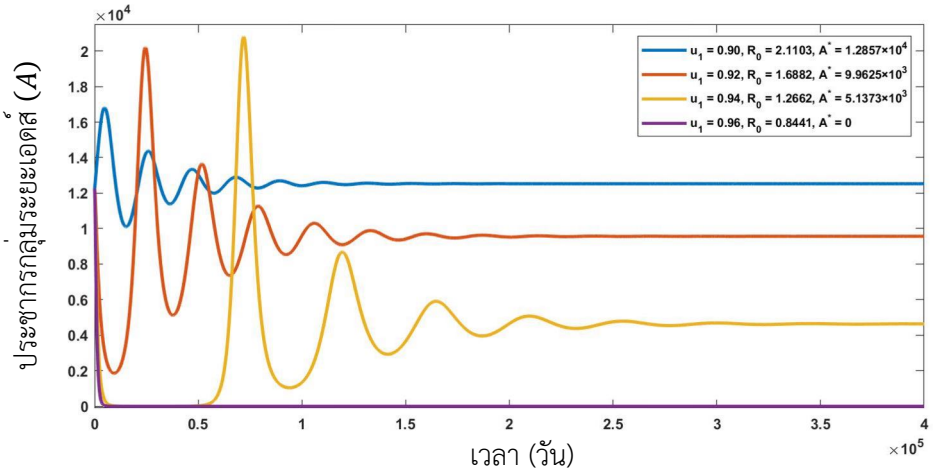
รูปที่ 2: กราฟของประชากรกลุ่มเสี่ยง (S) เมื่อประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมีค่าแตกต่างกัน เทียบกับเวลา : หน่วยเป็นวัน



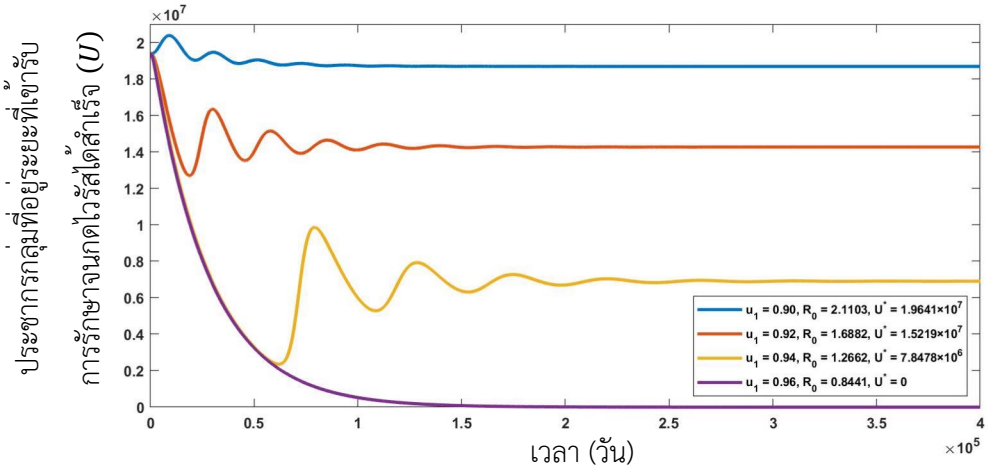
รูปที่ 3: กราฟประชากรกลุ่มติดเชื้อระยะเอชไอวี (H) เมื่อประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมีค่าแตกต่างกัน เทียบกับเวลา : หน่วยเป็นวัน



รูปที่ 4: กราฟประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา (T) เมื่อประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมีค่าแตกต่างกัน เทียบกับเวลา : หน่วยเป็นวัน



รูปที่ 5: กราฟประชากรกลุ่มระยะเอดส์ (A) เมื่อประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมีค่าแตกต่างกัน เทียบกับเวลา : หน่วยเป็นวัน



รูปที่ 6: กราฟประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ (U) เมื่อประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมีค่าแตกต่างกัน เทียบกับเวลา : หน่วยเป็นวัน

จากรูปที่ 2 - 6 แสดงพฤติกรรมการลู่เข้าคำตอบของจำนวนประชากรกลุ่มเสี่ยง (S) จำนวนประชากรกลุ่มติดเชื้อระยะเอชไอวี (H) จำนวนประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา (T) จำนวนประชากรกลุ่มระยะเอดส์ (A) และจำนวนประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ (U) เมื่อประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมีค่าแตกต่างกัน เมื่อกำหนดให้ u_1 เท่ากับ $0.9,\ 0.92$ และ 0.94 หรือประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ เท่ากับ $90\%,\ 92\%$ และ 94% พบว่ากราฟคำตอบของจำนวนประชากรกลุ่ม S มีแนวโน้มลดลง ในขณะที่จำนวนประชากรกลุ่ม $H,\ T,\ A$ และ U มีแนวโน้มเพิ่มขึ้นและเมื่อเวลาผ่านไปพบว่าทุกกลุ่มประชากรจะลู่เข้าจุดสมดุลที่มีการแพร่ระบาด และ $\Re_0 > 1$ หมายความว่าถ้าประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ เท่ากับ $90\%,\ 92\%$ และ 94% ไม่เพียงพอต่อการควบคุมการแพร่ระบาดของไวรัสเอชไอวีและโรคเอดส์ได้ และจากรูปที่ 2-6 จะพบว่าถ้ากำหนดให้ u_1 เท่ากับ 0.96 หรือประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ เท่ากับ 96% กราฟคำตอบของจำนวนประชากรกลุ่ม S มีแนวโน้มเพิ่มขึ้นและลู่เข้า $S^* = 3.9254 \times 10^7$ ในขณะที่จำนวนประชากรกลุ่ม $H,\ T,\ A$ และ U มีแนวโน้มลดลงและลู่เข้าสู่ศูนย์ หมายความว่าเมื่อเวลาผ่านไประบบจะมีเพียงประชากรกลุ่มเสี่ยง และประชากรกลุ่มติดเชื้อระยะเอชไอวี ประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา ประชากรกลุ่มระยะเอดส์ และประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ เท่ากับ ศูนย์ และพบว่า $\Re_0 < 1$ นั่นหมายความว่าถ้าประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ เท่ากับ 96% สามารถควบคุมการแพร่ระบาดของโรคได้

3.3.2 ผลวิเคราะห์ค่าดัชนีความไวของพารามิเตอร์ที่มีผลต่อค่าระดับการติดเชื้อ

การวิเคราะห์ค่าดัชนีความไวสามารถคำนวณได้โดยใช้เทคนิค Normalized Forward Sensitivity [21 - 22] ซึ่งเป็นวิธีทางคณิตศาสตร์ที่ใช้ในการตรวจสอบว่าพารามิเตอร์ตัวใดมีผลต่อค่าระดับการติดเชื้อ ซึ่งช่วยให้ทราบว่าควรใช้มาตราการควบคุมโรคแบบใดจึงจะสามารถควบคุมโรคได้มีประสิทธิภาพที่สุด สามารถคำนวณได้จากสูตร ดังนี้

$$\Upsilon_k^{\Re_0} = \frac{\partial \Re_0}{\partial k} \times \frac{k}{\Re_0}$$

โดยที่ k คือพารามิเตอร์ใดๆ ในค่าระดับการติดเชื้อ

เมื่อแทนพารามิเตอร์จากตารางที่ 1 จะได้ค่าดัชนีความไวของพารามิเตอร์ต่อค่าระดับการติดเชื้อ ดังนี้

ตารางที่ 4: ค่าดัชนีความไวของพารามิเตอร์ต่อค่าระดับการติดเชื้อ

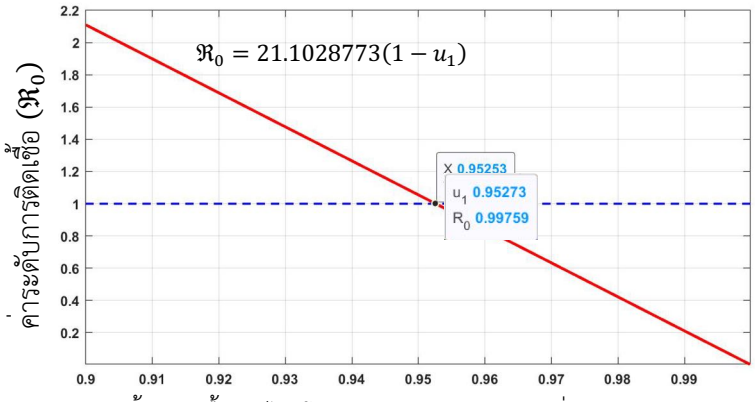
ลำดับที่	พารามิเตอร์	ค่าดัชนีความไว
1	u_1	-9.0000
2	μ	-1.0152
3	π	$+1.0000$
4	β	$+1.0000$
5	θ	-0.6510
6	ε	-0.6498
7	η_1	$+0.2663$
8	u_2	-0.2594
9	γ	-0.2594
10	ω	-0.0617
11	d	-0.0128
12	η_2	$+0.0098$
13	α	-0.0032
14	ρ	$+0.0030$

จากตารางที่ 4 ผลเชิงตัวเลขของค่าดัชนีความไวของพารามิเตอร์ต่อค่าระดับการติดเชื้อ ณ จุดสมดุลที่มีการแพร่ระบาดของโรค เมื่อเรียงลำดับความสำคัญและความสัมพันธ์ของพารามิเตอร์ต่อค่าระดับการติดเชื้อ พบว่าถ้าต้องการลดค่าระดับการติดเชื้อ ลำดับที่ 1 ควรเพิ่ม u_1 (ประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ) เช่น เพิ่มประสิทธิภาพของถุงยางอนามัยและยาแพร็พให้สามารถป้องกันโรคได้มากขึ้น หรือเพิ่มวินัยของผู้ที่รับประทานยา จะต้องมีวินัยในการรับประทานยาอย่างต่อเนื่องและตรงเวลาและควรสวมถุงยางอนามัยอย่างถูกต้องและสวมทุกครั้งที่มีเพศสัมพันธ์ ลำดับที่ 2 เพิ่ม μ (อัตราการตายตามธรรมชาติ) คือการลดระยะเวลาของการมีชีวิตดังนั้นวิธีนี้ไม่ใช่แนวทางที่ทำได้จริงในคน ลำดับที่ 3 ลด π (จำนวนคนเกิดใหม่) เช่น ในพื้นที่ที่มีการแพร่ระบาดของเอชไอวี/เอดส์ อาจจะต้องควบคุมจำนวนทารกเกิดใหม่ รวมไปถึงลดจำนวนประชากรย้ายเข้าในพื้นที่ที่มีการแพร่ระบาด ลำดับที่ 4 ลด β (อัตราการติดเชื้อจากการมีเพศสัมพันธ์) คือการลดโอกาสของการติดเชื้อหรือแพร่เชื้อเอชไอวีจากการมีเพศสัมพันธ์ เช่น ควรสวมถุงยางอนามัยทุกครั้ง รับประทานยาเพร็พอย่างต่อเนื่อง ไม่เปลี่ยนคู่นอนบ่อย ลำดับที่ 5 เพิ่ม θ (อัตราการเข้ารับการตรวจเลือดและเข้ารับการรักษา) เช่น ผู้ที่มีความเสี่ยงต่อการติดเชื้อควรเข้ารับการตรวจเลือดให้เร็ว ถ้าพบว่าติดเชื้อ ควรเข้ารับการรักษาทันที ลำดับที่ 6 เพิ่ม ε (สัดส่วนการเข้ารับการตรวจเลือดและเข้ารับการรักษา) เช่น ควรมีการรณรงค์ให้ผู้ที่มีความเสี่ยงต่อการติดเชื้อเข้ารับการตรวจเลือดและเข้ารับการรักษาให้มากขึ้น ลำดับที่ 7 ลด η_1 (สัดส่วนการติดเชื้อจากกลุ่ม T) ควรลดจำนวนการแพร่เชื้อจากคนกลุ่มที่

เข้ารับการรักษาให้น้อยลง เช่น สร้างความตระหนักในการสวมถุงยางอนามัยทุกครั้งในการมีเพศสัมพันธ์ ลำดับที่ 8 เพิ่ม u_2 (ประสิทธิภาพของการรักษาด้วยยาต้านไวรัส ART) ให้เพิ่มประสิทธิภาพของยาต้านไวรัส ART ให้ดีขึ้น ลำดับที่ 9 เพิ่ม γ (อัตราการเปลี่ยนสถานะจากกลุ่ม T เป็นกลุ่ม U) คือให้ลดระยะเวลาของการรักษาให้เร็วขึ้น โดยเพิ่มประสิทธิภาพของยาหรือพัฒนาวิธีการรักษาให้ดีขึ้น และผู้ที่เข้ารับการรักษาควรมีวินัยในการดูแลตนเองและการรับประทานยา ลำดับที่ 10 เพิ่ม ω (อัตราการเปลี่ยนสถานะจากจากกลุ่ม H เป็นกลุ่ม A) คือการลดระยะเวลาจากผู้ติดเชื้อระยะเอชไอวีกลายเป็นผู้ติดเชื้อระยะเอดส์ ซึ่งจะทำให้ผู้ติดเชื้อมีอาการรุนแรงมากขึ้น และอาจเสียชีวิตเร็วขึ้น ดังนั้นวิธีการนี้ไม่สามารถนำมาใช้ได้จริง ลำดับที่ 11 เพิ่ม d (อัตราการเสียชีวิตจากโรค) คือการทำให้ผู้ติดเชื้อที่อยู่ในระยะเอดส์เสียชีวิตเร็วขึ้น วิธีการนี้ไม่สามารถนำมาใช้ได้จริงเช่นเดียวกัน ลำดับที่ 12 ลด η_2 (สัดส่วนการติดเชื้อจากกลุ่ม A) ลำดับที่ 13 เพิ่ม α (อัตราของการเข้ารับการรักษาของกลุ่ม A) คือ ควรรณรงณ์ให้ผู้ติดเชื้อที่อยู่ในระยะเอดส์เข้ารับการรักษาให้เร็วขึ้น และลำดับที่ 14 ลด ρ (อัตราการล้มเหลวของการรักษา) คือ การลดความล้มเหลวของการรักษา โดยเพิ่มประสิทธิภาพของรักษาให้ดีมากขึ้น หรือสร้างความตระหนักให้ผู้ที่กำลังรักษา รับประทานยาอย่างต่อเนื่อง ตรงเวลา รับประทานอาหารที่มีประโยชน์ และดูแลสุขภาพร่างกาย ซึ่งจะช่วยให้การรักษามีประสิทธิภาพมากยิ่งขึ้น

3.3.3 ผลการศึกษาความสัมพันธ์ของประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พและการรักษาด้วยวิธียาต้านไวรัส ART ต่อการแพร่ระบาดของโรค

จากผลการวิเคราะห์เสถียรภาพของระบบพบว่าค่าระดับการติดเชื้อมีผลต่อการลู่เข้าของจุดสมดุล ดังนั้นในหัวข้อนี้จะนำเสนอผลการศึกษาความสัมพันธ์ของประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ (u_1) และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART (u_2) ต่อค่าระดับการติดเชื้อ และจำนวนประชากรกลุ่มติดเชื้อ ผลการศึกษาแสดงดังตารางที่ 5 - 7 และรูปที่ 7 - 9 ดังนี้



ประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ (u_1)

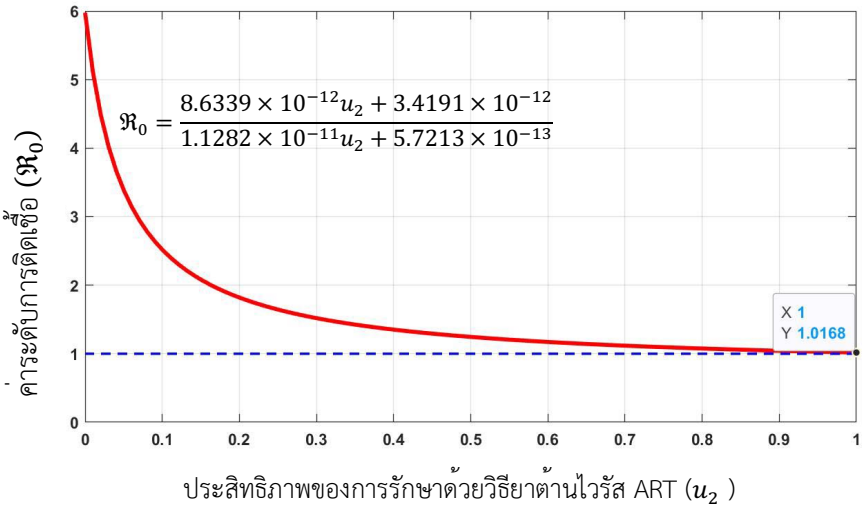
รูปที่ 7: กราฟความสัมพันธ์ระหว่างประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ (u_1) กับค่าระดับการติดเชื้อ $(\Re_0)$

จากรูปที่ 7 กราฟความสัมพันธ์ระหว่างประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ (u_1) กับค่าระดับการติดเชื้อ $(\Re_0)$ เมื่อกำหนดให้ประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART เท่ากับ 86.1% พบว่าถ้าประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พเพิ่มมากขึ้นจะส่งผลให้ค่าระดับการติดเชื้อลดลง และถ้าประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมากกว่า 0.9527 จะทำให้ค่า $\Re_0 < 1$ หมายความว่าถ้าประสิทธิภาพของการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พในการป้องกันโรคมากกว่า 95.27% จะสามารถควบคุมและลดการแพร่ระบาดของโรคได้ ซึ่งสอดคล้องกับตารางที่ 5 เมื่อกำหนด $u_1 = 0.9,\ 0.95,\ 0.9527$ และ 0.98 ตามลำดับ หมายความว่าประสิทธิภาพของการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พในการป้องกันโรค เท่ากับ $90\%,\ 95\%,\ 95.27\%$ และ 98% พบว่า ถ้าประสิทธิภาพของการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พในการป้องกันโรค เท่ากับ 90% เมื่อเวลาผ่านไป จำนวนประชากรกลุ่มติดเชื้อระยะเอชไอวี เท่ากับ $238{,}332$ คน จำนวนประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา เท่ากับ $29{,}221$ คน จำนวนประชากรกลุ่มระยะเอดส์ เท่ากับ $12{,}857$ คน และจำนวนประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ เท่ากับ $19{,}641{,}42$ คน และเมื่อเพิ่มประสิทธิภาพของการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พในการป้องกันโรค เท่ากับ 95% เมื่อเวลาผ่านไป พบว่าจำนวนประชากรกลุ่มติดเชื้อระยะเอชไอวีลดลงเหลือ เท่ากับ $23{,}674$ คน จำนวนประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา เท่ากับ $29{,}025$ คน จำนวนประชากรกลุ่มระยะเอดส์ เท่ากับ $1{,}277$ คน และจำนวนประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ เท่ากับ $1{,}951{,}033$ คน และเมื่อเพิ่มประสิทธิภาพของการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พในการป้องกันโรค เท่ากับ 95.27% พบว่าเมื่อเวลาผ่านไป จะไม่มีจำนวนผู้ติดเชื้อทั้ง 4 กลุ่ม นั่นแสดงให้เห็นว่าถ้าประสิทธิภาพของการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พในการป้องกันโรคมากกว่า 95.27% จะสามารถควบคุมและลดการแพร่ระบาดของโรคได้

ตารางที่ 5: ผลของประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พต่อค่าระดับการติดเชื้อและจำนวนประชากรกลุ่มติดเชื้อ

u_1	$\Re_0$	H^*	T^*	A^*	U^*
$\mathbf{0.90}$	2.1103	238,332	29,221	12,857	19,641,422
$\mathbf{0.95}$	1.0551	23,674	29,025	1,277	1,951,033
$\mathbf{0.9527}$	0.9982	0	0	0	0
$\mathbf{0.96}$	0.8441	0	0	0	0
$\mathbf{0.98}$	0.4221	0	0	0	0

ลำดับถัดไปจะแสดงผลการศึกษาความสัมพันธ์ของประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART (u_2) ต่อค่าระดับการติดเชื้อ และจำนวนประชากรกลุ่ม $H,\ T,\ A$ และ U



รูปที่ 8: กราฟความสัมพันธ์ระหว่างประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART (u_2) กับค่าระดับการติดเชื้อ $(\Re_0)$

ตารางที่ 6: ผลของประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART ต่อค่าระดับการติดเชื้อและจำนวนประชากรกลุ่มติดเชื้อ

u_2	$\Re_0$	H^*	T^*	A^*	U^*
0.861	1.0551	23,674	29,025	1,277	1,951,033
0.90	1.0433	18,781	22,081	977	1,551,536
0.95	1.0293	12,922	14,435	644	1,070,586
1.00	1.0168	7,484	7,962	358	621,608

จากรูปที่ 8 และตารางที่ 6 เมื่อกำหนดให้ประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ เท่ากับ 95% พบว่าเมื่อประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART เพิ่มขึ้น ส่งผลให้ค่าระดับการติดเชื้อลดลง และเมื่อกำหนด $u_2 = 0.861$ นั่นคือประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART เท่ากับ 86.1% เมื่อเวลาผ่านไป พบว่ามีจำนวนประชากรกลุ่มติดเชื้อระยะเอชไอวี เท่ากับ 23,674 คน จำนวนประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา เท่ากับ 29,025 คน จำนวนประชากรกลุ่มระยะเอดส์ เท่ากับ 1,277 คน จำนวนประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ เท่ากับ 1,951,033 คน และ $\Re_0 = 1.0551$ แสดงให้เห็นว่ามีการแพร่ระบาดของโรค และเมื่อกำหนดให้ $u_2 = 0.90,\ 0.95$ และ 1 ตามลำดับ นั่นคือ เมื่อเพิ่มประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART เป็น 90%, 95% และ 100% พบว่า $\Re_0 > 1$ และยังคงพบจำนวนประชากรกลุ่มติดเชื้อทุกกลุ่ม จะเห็นได้ว่าถ้าประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ เท่ากับ 95% และการรักษาด้วยวิธียาต้านไวรัส ART มีประสิทธิภาพถึง 100% ก็ไม่สามารถควบคุมการแพร่ระบาดของโรคได้

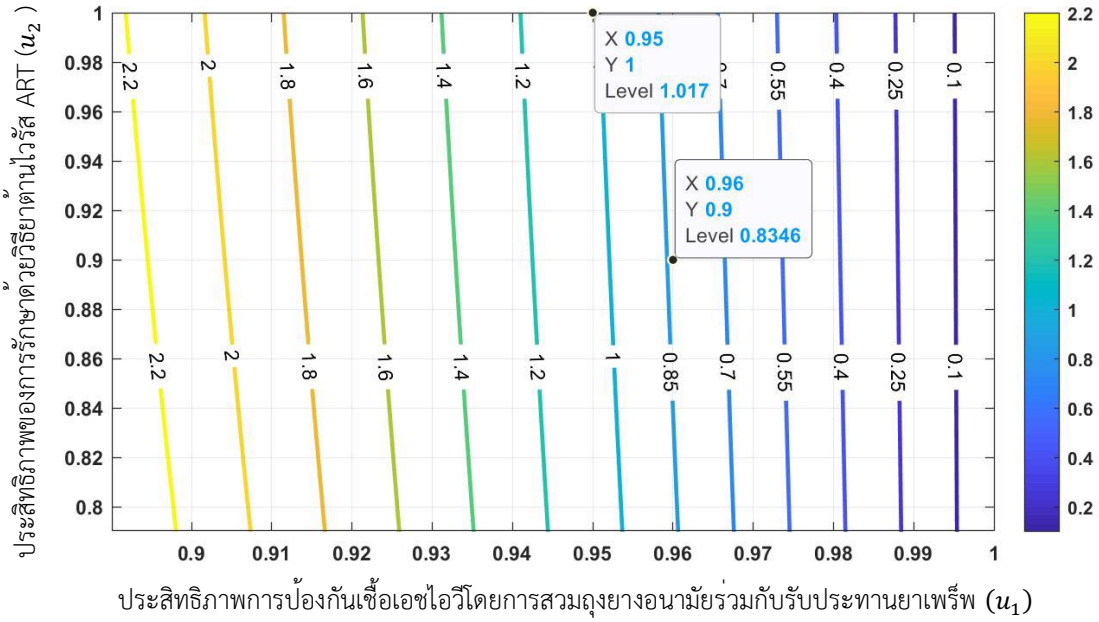
ดังนั้นลำดับถัดไปจะเสนอผลการศึกษาประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ (u_1) และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART (u_2) ต่อค่าระดับการติดเชื้อ และจำนวนประชากรกลุ่ม $H,\ T,\ A$ และ U

ตารางที่ 7: ผลของประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พต่อค่าระดับการติดเชื้อและจำนวนประชากรกลุ่มติดเชื้อ

u_1	u_2	$\mathfrak{R}_0$	H^*	T^*	A^*	U^*
0.95	0.86	1.0555	23,804	29,216	1,285	1,961,575
0.9525	0.862	1.0021	942	1,154	51	77,680
0.9525	0.8711	1.0000	2.68	3.25	0.14	222
0.96	0.90	0.8346	0	0	0	0

จากตารางที่ 7 กำหนดให้ประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART แตกต่างกัน จะเห็นได้ว่าเมื่อ $u_1 = 0.95$ และ $u_2 = 0.861$ นั่นคือเมื่อประสิทธิภาพในการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็บและประสิทธิภาพของการรักษาด้วยวิธียาต้ายไวรัส ART เท่ากับ 95% และ 86% ตามลำดับ เมื่อเวลาผ่านไป จะทำให้จำนวนประชากรกลุ่มติดเชื้อระยะเอชไอวี เท่ากับ 23,804 คน จำนวนประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษาเท่ากับ 29,216 คน จำนวนประชากรกลุ่มระยะเอดส์ เท่ากับ 1,285 คน จำนวนประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ เท่ากับ 1,961,575 คน ซึ่งไม่เพียงพอต่อการควบคุมโรค เช่นเดียวกับเมื่อเพิ่ม $u_1 = 0.9525$ และ $u_2 = 0.862$ คือเพิ่มประสิทธิภาพของการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พเป็น 95.25% และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART เป็น 86.2% พบว่าเมื่อเวลาผ่านไปจำนวนผู้ติดเชื้อทั้ง 4 กลุ่มลดลง จำนวนประชากรกลุ่มติดเชื้อระยะเอชไอวี เท่ากับ 942 คน จำนวนประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา เท่ากับ 1,154 คน จำนวนประชากรกลุ่มระยะเอดส์ เท่ากับ 51 คน จำนวนประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ เท่ากับ 77,680 คน แต่จำนวนผู้ติดเชื้อทั้ง 4 กลุ่มยังคงสูงกว่าค่าเป้าหมายที่กรมควบคุมโรคต้องการลดจำนวนผู้ติดเชื้อรายใหม่ไม่ให้เกิน 1,000 คนต่อปี และเมื่อกำหนดให้ $u_1 = 0.9525$ แต่เพิ่ม $u_2 = 0.8711$ นั่นคือเพิ่มประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART เป็น 87.11% พบว่าเมื่อเวลาผ่านไปจำนวนประชากรกลุ่มติดเชื้อระยะเอชไอวี ประมาณ $2 - 3$ คน จำนวนประชากรกลุ่มติดเชื้อที่อยู่ในระยะการรักษา ประมาณ 2 คน ไม่พบประชากรกลุ่มระยะเอดส์ และจำนวนประชากรกลุ่มที่อยู่ระยะที่เข้ารับการรักษาจนกดไวรัสได้สำเร็จ เหลือเพียง 222 คน ซึ่งจากผลเชิงตัวเลขนี้จะเห็นได้ว่าประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ เท่ากับ 95.25% ร่วมกับประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART 87.11% สามารถลดจำนวนผู้ติดเชื้อรายใหม่ได้น้อยกว่า 1,000 คนต่อปีได้ แต่ก็ยังไม่เพียงพอต่อการควบคุมโรคให้หมดไป แต่เมื่อเพิ่ม $u_1 = 0.96$ และ $u_2 = 0.9$ นั่นคือประสิทธิภาพของการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พเป็น 96% และ

ประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART เป็น 90% พบว่าเมื่อเวลาผ่านไป จะไม่มีจำนวนผู้ติดเชื้อทั้ง 4 กลุ่ม นั่นแสดงว่าถ้าประสิทธิภาพของการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พในการป้องกันโรค 96% และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART 90% สามารถควบคุมและลดการแพร่ระบาดของโรคได้ ซึ่งสอดคล้องกับรูป 9



รูปที่ 9: กราฟความสัมพันธ์ระหว่างประสิทธิภาพการป้องกันเชื้อเอชไอวีโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ (u_1) และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART (u_2) กับค่าระดับการติดเชื้อ $(\mathfrak{R}_0)$

4 สรุปผลและอภิปรายผล

การศึกษาในครั้งนี้ได้พัฒนาตัวแบบคณิตศาสตร์การแพร่ระบาดของเอชไอวี/เอดส์ที่มีผลจากประสิทธิภาพของการป้องกันด้วยถุงยางอนามัยร่วมกับการใช้ยาเพร็พ และการรักษาด้วยวิธียาต้านไวรัส ART ผลการศึกษาพบว่าค่าระดับการติดเชื้อ $(\mathfrak{R}_0)$ เป็นเกณฑ์ในการควบคุมโรค ถ้า $\mathfrak{R}_0 < 1$ แล้วจุดสมดุลที่ไม่มีการแพร่ระบาดมีความเสถียรภาพเชิงเส้นกำกับ แต่ถ้า $\mathfrak{R}_0 > 1$ แล้วจุดสมดุลที่มีการแพร่ระบาดมีความเสถียรภาพเชิงเส้นกำกับ จากผลการวิเคราะห์ค่าดัชนีความไวของค่าพารามิเตอร์ต่อค่าระดับการติดเชื้อพบว่าประสิทธิภาพการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมีผลต่อค่าระดับการติดเชื้อมากที่สุด และมีค่าดัชนีความไวเป็นบวก นั่นคือแนวทางป้องกันและควบคุมการแพร่ระบาดของเอชไอวี/เอดส์ที่ดีที่สุดคือเพิ่มประสิทธิภาพของการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ เช่น ควรใช้ถุงยางอนามัยทุกครั้งเมื่อมีเพศสัมพันธ์ รับประทานยาทุกวันอย่างต่อเนื่อง และรับประทานในเวลาเดียวกันในทุก ๆ วัน เช่นเดียวกันกับประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART มีผลต่อค่าระดับการติดเชื้อในลำดับที่แปด และมีค่าดัชนีความไวเป็นบวกเช่นเดียวกัน นั่นคือควรเพิ่ม

ประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART มากขึ้น และเมื่อตรวจสอบผลเชิงตัวเลขพบว่าถ้าประสิทธิภาพของการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พอยู่ระหว่าง 90 – 95% และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART เท่ากับ 86.1% ตามข้อมูลผลศึกษาประสิทธิภาพของการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พ [5] และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART [19] ยังพบการแพร่ระบาดของโรค อย่างไรก็ตาม ผลที่ได้จากการศึกษาพบว่า ถ้าประสิทธิภาพของการป้องกันโรคโดยการสวมถุงยางอนามัยร่วมกับรับประทานยาเพร็พมากกว่า 96% และประสิทธิภาพของการรักษาด้วยวิธียาต้านไวรัส ART มากกว่า 90% จะสามารถควบคุมโรคได้ ซึ่งสอดคล้องกับแนวทางควบคุมเอชไอวี/เอดส์ของกรมควบคุมโรค และสอดคล้องกับผลการศึกษาของ Adesola et al. (2024) [10] ได้พัฒนาตัวแบบคณิตศาสตร์เกี่ยวกับประสิทธิภาพและความสม่ำเสมอในการใช้ถุงยางอนามัยต่อการแพร่เชื้อเอชไอวี/เอดส์ ผลการศึกษาพบว่าการใช้ถุงยางอนามัยอย่างสม่ำเสมอในระหว่างการมีเพศสัมพันธ์นั้นสามารถป้องกันโรคได้มากกว่าการมีเพศสัมพันธ์โดยปราศจากการป้องกัน และสอดคล้องกับ Meetei et al. (2024) [11] ได้ศึกษาการแพร่ระบาดของของเอชไอวี/เอดส์แบบรายปีของประเทศปากีสถานในช่วงปี ค.ศ. 1992 – 2020 พบว่าแนวทางป้องกันและควบคุมโรคในระยะยาว คือ การปรับเปลี่ยนพฤติกรรมทางเพศอย่างปลอดภัยด้วยกันสวมถุงยาอนามัย และควรเข้ารับการรักษาโรคอย่างต่อเนื่อง เช่นเดียวกันกับ Ogunmodimu et al. (2024) [12] ได้ศึกษาการแพร่ระบาดของเอชไอวี/เอดส์ในแอฟริกา ผลการศึกษาพบว่าการรักษาด้วยวิธี ART จะทำให้ระดับไวรัสของผู้ที่ติดเชื้อเอชไอวี/เอดส์ เข้าสู่สถานะไม่พบเชื้อไวรัสในร่างกายและไม่สามารถแพร่กระจายเชื้อให้กับผู้อื่นได้ ทำให้การแพร่ระบาดของเอชไอวี/เอดส์ในแอฟริกาลดลงและหมดไปได้ และสอดคล้องกับข้อมูลจาก CDC (2025) [22]เกี่ยวกับการศึกษาแนวทางการป้องกันการติดเชื้อเอชไอวีด้วยยาเพร็พ ผลการศึกษาพบว่าวิธีการรับประทานยาเพร็พให้มีประสิทธิภาพในการป้องกันการติดเชื้อเอชไอวีให้มากถึง 96 - 99% คือ ควรรับประทานยาอย่างต่อเนื่องทุกวัน รับประทานในเวลาเดียวกัน เช่น หลังอาหารเช้า ร่วมกับการสวมถุงยางอนามัยอย่างถูกต้อง และสวมทุกครั้งที่มีเพศสัมพันธ์

กิตติกรรมประกาศ ขอบคุณผู้ทรงคุณวุฒิทุกท่านที่ได้ให้ข้อแนะนำและข้อเสนอแนะต่างๆ เพื่อปรับปรุงงานวิจัยนี้ และขอขอบคุณสาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสุราษฎร์ธานีที่สนับสนุนการดำเนินการวิจัยในครั้งนี้จนสำเร็จลุล่วงด้วยดี

เอกสารอ้างอิง

[1] World Health Organization, HIV and AIDS. < https://www.who.int/news-room/fact- sheets/detail /hiv-aids >, 2024 (accessed 2025.02.23).

[2] Department of Disease Control, Estimated HIV infections Thailand. < https://hivhub.ddc.moph. go.th/epidemic.php >, 2022 (accessed 2025.02.23).

[3] คณะกรรมการแห่งชาติว่าด้วยการป้องกันและแก้ไขปัญหาเอดส์, แผนปฏิบัติการยุติปัญหาเอดส์ พ.ศ.2566 – 2569, บริษัท เอ็นซี คอนเซ็ปต์ จำกัด, กรุงเทพมหานคร, น. XIV.

[4] คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี, ติด HIV อาการ สาเหตุ และการป้องกัน. < https://www.rama.mahidol.ac.th/ramachannel >, 2024 (สืบค้น 26 ม.ค. 68).

[5] ศูนย์วิจัยโรคเอดส์และโรคติดเชื้อ สภากาชาดไทย (คลินิกนิรนาม), เพร็พ (PrEP) ,เป็ป (PEP) คืออะไร. < https://th.trcarc.org/เพร็พ-prep-เป็ป-pep-คืออะไร/ >, 2562 (สืบค้น 1 ก.พ. 68).

[6] โรงพยาบาลบางปะกอก 9, HIV กับ เอดส์ ความเหมือนที่แตกต่าง. < https://bpk9internationalhospital.com/care_blog/content/ >, 2566 (สืบค้น 1 ก.พ. 68).

[7] T. K. Ayele, E. F. D. Goufo and S. Mugisha, Mathematical modeling of HIV/AIDS with optimal control: a case study in Ethiopia, Results in Physics. 26 (2021) 104263.

[8] P. S. Rana, N. Sharma and A. Priyadarshi, Mathematical modeling and optimal control of a deterministic SHATR model of HIV/AIDS with possibility of rehabilitation: a dynamic analysis, Tamkang Journal of Mathematics. 55 (3) (2024) 267-285

[9] E. M. D. Moya, R. A. Rodriguez and A. Pietrus, A mathematical model for the study of HIV/AIDS transmission with PrEP coverage increase and parameter estimation using MCMC with a Bayesian approach, Bulletin of Biomathematics. 2 (2) (2024) 218-244.

[10] O. I. Adesola, S. S. Oloruntoyin, P. M. Emmanuel, M. I. Temilade, A. G. Adeyemi, A. S. Oladele and A. A. Kareem, Mathematical Modelling and Analyzing the Dynamics of Condom Efficacy and Compliance in the Spread of HIV/AIDS, Asian Research Journal of Current Science. 6 (1) (2024) 54-65.

[11] M. Z. Meetei, M. H. DarAssi, M., A. Khan, A. N. A. Koam, E. Alzahrani and A. A. H. Ahmadini, Analysis and simulation study of the HIV/AIDS model using the real cases, PLOS ONE. 19 (6) (2024) e0304735.

[12] M. O. Ogunmodimu, E. P. Enock, A. P. Kenyatta, S. B. Affognon and F. C. Onwuegbuche, A mathematical model for the prevention of HIV/AIDS in the presence of undetectable equals untransmittable viral load. International Journal of Mathematical Sciences and Optimization: Theory and Applications. 10 (2) (2024) 36-57.

[13] ญาดา ป้องกัน, เบญญาภา น้ำขาว, อนุสรา ศิริรัตน์, อัญชุลี ณ ตะกั่วทุ่ง, กันยากร อ่อนรักษ์ และเกตุกนก หนูดี, ตัวแบบเชิงคณิตศาสตร์การแพร่ระบาดของโรคเมลิออยโดสิสที่มีผลจากการใส่รองเท้าบูท, วารสารวิทยาศาสตร์ลาดกระบัง. 33 (1) (2567) 8 – 26.

[14] Fosu G, Akweittey E and Adu-sackey, Next-generation matrices and basic reproductive numbers for all phases of the Coronavirus disease, Open J. Math. Sci. 2020 (4), 261-272.

[15] พรวิลัย ชาญกิจกรรณ์ และพันธนี พงศ์สัมพันธ์, แบบจำลองทางคณิตศาสตร์การระบาดของโรคเมลิออยโดสิสจากสัตว์สู่คนโดยพิจารณาจากผลกระทบปัจจัยปริมาณน้ำฝน, วารสารวิทยาศาสตร์ประยุกต์. 19 (1) (2563), 80 - 94.

[16] สำนักงานสถิติแห่งชาติ, สถิติประชากรทางการทะเบียนราษฎร ปี 2563 - 2567. < https://stat.bora.dopa.go.th/stat/statnew/statMONTH/statmonth/#/mainpage >, 2568 (สืบค้น10 ก.พ. 68).

[17] ศูนย์รวมข้อมูลสารสนเทศด้านเอชไอวีของประเทศไทย กรมควบคุมโรค, ระบบข้อมูลโรคเอดส์ ปี 2566, < https://hivhub.ddc.moph.go.th/report.php >, 2566 (สืบค้น 30 ม.ค. 68).

[18] อนุวัฒน์ กีระสุนทรพงษ์. ความรู้เกี่ยวกับเอชไอวีและเอดส์, < https://www.bumrungrad.com/th/ health-blog/november-2016/hiv-aids-infection-treatment >, 2567 (สืบค้น 10 ก.พ. 68).

[19] Cleveland Clinic, Antiretroviral Therapy, < https://my.clevelandclinic.org/health/ treatments/antiretroviral-therapy >, 2024 (accessed 2025.02.10)

[20] Surapol Naowarat and Kanyakon Onruk, Global Stability and Sensitivity Analysis of a Conjunctivitis Epidemic Model in Two Populations with Patient Isolation, Treatment and Insecticide Control, Journal of Applied Science and Emerging Technology (JASET). 23 (1) (2024), 1 - 10.

[21] Ngoteya, F.N. and Nkansah-Gyekye, Y. Sensitivity Analysis of Parameters in a Competition Model. Applied and Computational Mathematics, 4 (5) (2015), 363 - 368.

[22] Centers for Disease Control and Prevention. Preventing HIV with PrEP. < https://www.cdc.gov/hiv /basics/prep.html >, 2025 (accessed 2025.04.29).

Influenza Dynamics: A mathematical modeling approach

Wirawan Chinviriyasit[1,*] and Settapat Chinviriyasit[1,†]

[1]Department of Mathematics, Faculty of Science,
King Mongkut's University of Technology Thonburi, 126 Pracha Uthit Rd,
Bang Mot, Thung Khru, Bangkok, Thailand 10140

Abstract

Influenza is a significant concern among emerging infectious diseases and recurrent infections due to its frequent outbreaks worldwide, including in Thailand. Therefore, this paper aims to construct a mathematical model to determine the influential factors in increasing transmission and to suggest measures to prevent influenza. Mathematical analysis reveals that the disease-free equilibrium of the proposed model is globally asymptotically stable whenever the reproductive number denoted by $\mathcal{R}_0$ is less than or equal to unity. This result indicates that the disease is eliminated if the control measure is related to making $\mathcal{R}_0 \leq 1$. Numerical solutions are demonstrated using real data from Thailand as a case study. The findings highlight key factors influencing influenza transmission, helping to develop effective preventive measures.

Keywords: influenza H1N1, reproductive number, sensitivity analysis
2020 MSC: 00A71

1 Introduction

Influenza H1N1 is one of the seasonal flu viruses that circulate each year. This disease was first detected in people in the United States in April 2009. However, the fear surrounding the pandemic is still associated with it [1]. Based on past pandemics, the preparedness and response of public health care for an influenza pandemic crisis must be improved. To analyze control strategies on the 2009 A/H1N1 influenza epidemic, the factors that cause the disease transmission and prevention need to determine in order to get the important information to design useful influenza control programs. Mathematical modeling of the spread of influenza is one useful tool in comparing the effects of different control strategies [1–5]. Sensitivity analysis of model parameters is a method to design control strategies as well as a direction to future research. Chitnis et al. [6] and Samsuzzoha [3] have evaluated the sensitivity indices of the basic reproduction number to the parameters in the model in order to determine the relative importance of different parameters in the transmission and prevalence of malaria and influenza, respectively. Therefore, this paper aims to formulate a mathematical model of influenza H1N1 based on epidemiology of disease to gain insight the dynamical behavior of disease transmission and to explore the factors which can help guide in developing efficient intervention strategies for influenza control programs.

*Speaker (wirawon.chi@kmutt.ac.th)
†Corresponding author (settapat.chi@kmutt.ac.th)

2 Preliminaries

2.1 Model formula

From an epidemiological viewpoint, the total population at time t denoted by $N(t)$ is divided into five subpopulations, namely, susceptible individuals $S(t)$ who do not have specific immune defenses against the disease; exposed individuals $E(t)$ who infect and are in the incubation period of the disease; symptomatic individuals $I(t)$ who infect and show disease symptom; asymptomatic individuals $A(t)$ who infect and do not show symptoms of the disease; recovered individuals $R(t)$who recover from disease and have partially immune to disease, respectively. The population of susceptible individuals is increased by the recruitment of individuals into this population at a rate Π. DCD [1] reported that people infected influenza H1N1 may be able to transmit to others from 1 day before getting sick to 5 to 7 days after. Patient who do not show disease symptoms also are able to transmit. Therefore, susceptible population decrease at the rate $\frac{\beta}{N}(I + \psi_1 E + \psi_2 A)$, where β is contact rate, the modification parameters ψ_1 and ψ_2 account for the assumed reduction in transmissibility of exposed individuals and asymptomatic individuals, respectively, relative to symptomatic individuals [7]. Further, the model for the transmission dynamics of influenza H1N1 is given by the following deterministic system of nonlinear differential equations:

$$\left.\begin{aligned}
\frac{dS}{dt} &= \Pi - \frac{\beta S}{N}(I + \psi_1 E + \psi_2 A) - \mu S,\\
\frac{dE}{dt} &= \frac{\beta S}{N}(I + \psi_1 E + \psi_2 A) - (k + \mu)E,\\
\frac{dI}{dt} &= pkE - (\gamma + \mu + \alpha)I,\\
\frac{dA}{dt} &= (1 - p)\,kE - (\eta + \mu)A,\\
\frac{dR}{dt} &= \gamma I + \eta A - \mu R,
\end{aligned}\right\} \tag{2.1}$$

where Π is the recruitment of population into community (assumed susceptible), β is infection rate and μ is the natural death rate. When exposed individuals develop clinical symptoms, this population move to the symptomatic class (I) at a rate pk and move to the asymptomatic class (A) at a rate $(1 - p)k$ if they do not show clinical symptoms. When both symptomatic and asymptomatic individuals recover, they move to the R class at a rate γ and η, respectively. Since the model (2.1) monitors human populations, all the associated parameters and state variables are non-negative as described in Table 1.

2.2 Disease free equilibrium and reproductive number

It is easy to show that the disease-free equilibrium of the model (2.1) denoted E^0 is given by

$$E^0(S, E, I, A, R) = \left(\frac{\Pi}{\mu}, 0, 0, 0, 0\right) \tag{2.2}$$

According the next generation method [7, 8] and the notation used in [8], we define two

matrices:

$$F = \begin{bmatrix} \psi_1\beta & \beta & \psi_2\beta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } V = \begin{bmatrix} pk + (1-p)k + \mu & 0 & 0 \\ -pk & \gamma + \mu + \alpha & 0 \\ -(1-p)k & 0 & \eta + \mu \end{bmatrix}, \tag{2.3}$$

respectively.

Define $\mathcal{R}_0 = \rho\left(FV^{-1}\right)$, where ρ is the spectral radius. It is easy to show that

$$\mathcal{R}_0 = \frac{\beta\left(\left(\tilde{p}\,\psi_2 d_2 + p d_3\right)k + \psi_1 d_2 d_3\right)}{d_2 d_1 d_3}, \tag{2.4}$$

where $\tilde{p} = 1 - p$, $d_1 = k + \mu$, $d_2 = \gamma + \mu + \alpha$ and $d_3 = \eta + \mu$, respectively Consequently, using Theorem 2 of [8], if $\rho\left(FV^{-1}\right) < 1$, then all eigenvalues of Jacobian of the model (2.1), evaluated at E^0, have negative real parts, and for$\rho\left(FV^{-1}\right) > 1$, at least one of the eigenvalues of Jacobian has a positive real part, therefore, the following result is established.

Theorem 2.1. *The disease free equilibrium E^0 of the model* (2.1) *is locally asymptotically stable whenever $\mathcal{R}_0 < 1$, and unstable if $\mathcal{R}_0 > 1$.*

The local stability result presented in Theorem 2.1 indicates that when $\mathcal{R}_0 < 1$, the total number of infected individuals in the population can be reduced to zero. This is due to a small influx of infected individuals into the community, which will not lead to a significant outbreak of influenza if $\mathcal{R}_0 < 1$; however, an epidemic will occur if $\mathcal{R}_0 > 1$. Thus, $\mathcal{R}_0$ is called the effective reproduction number, represents the average number of secondary infections produced by a single infected individual in a fully susceptible population. It serves as a threshold parameter determining whether an infectious disease will spread within a population. To ensure that the effective control or eradication of the infection within the community at steady state is independent of the initial sizes of the sub-populations of the model, it is necessary to demonstrate that the disease-free equilibrium is globally asymptotically stable (GAS), which is analyzed below.

Theorem 2.2. *The disease free equilibrium E^0 of the model* (2.1) *is globally asymptotically stable in Ω whenever $\mathcal{R}_0 \leq 1$.*

Proof. Consider the Lyapunov function

$$V = E + \frac{d_1 d_2 \psi_2 A}{\left(\tilde{p}\,\psi_2 d_2 + p\,d_3\right)k + \psi_1 d_2 d_3} + \frac{d_1 d_3 I}{\left(\tilde{p}\,\psi_2 d_2 + p\,d_3\right)k + \psi_1 d_2 d_3}$$

Its derivative along the solutions to the system (2.1) is

$$\begin{aligned} V' &= \frac{d_1 d_2 d_3\left(A\,\psi_2 + E\,\psi_1 + I\right)}{\left(pp\,\psi_2 d_2 + p d_3\right)k + \psi_1 d_2 d_3}\left(\frac{\beta\left(\left(\tilde{p}\,\psi_2 d_2 + p d_3\right)k + \psi_1 d_2 d_3\right)S}{N\,d_1 d_2 d_3} - 1\right) \\ &\leq \frac{d_1 d_2 d_3\left(A\,\psi_2 + E\,\psi_1 + I\right)}{\left(pp\,\psi_2 d_2 + p d_3\right)k + \psi_1 d_2 d_3}\left(\mathcal{R}_0 - 1\right). \end{aligned}$$

Clearly, $V' < 0$ for $\mathcal{R}_0 < 1$. Furthermore, $V' = 0$ if and only if $E = A = I = 0$ or $\mathcal{R}_0 = 1$. Hence, V' is a Lyapunov function on Ω. Therefore, by the LaSalle's Invariance Principle [9], every solution to the equations of the model (2.1), with initial conditions in Ω, approaches E^0 as $t \to \infty$. This implies that E^0 is globally asymptotically stable in Ω. This complete the proof. □

3 Main Results

3.1 Appropriate model parameters

In this section, the model (2.1) is applied to predict the influenza H1N1 cases and then they are compared with the real cases reported by the Ministry of Public Health in Thailand, 2016-2017. In simulations, the model parameters used are given in Table 1 and the initial conditions used are $S(0) = 200,000,\ E(0) = 200,\ I(0) = 1,000,\ A(0) = 67,\ R(0) = 60,000$. The initial number of infected individuals $I(0)$ equals the real cases on May, 2016, while the others are assumed for simulation purpose. Figure 1 shows the number of infected individuals predicted by the model (2.1) and they are compared with the real cases. The accuracy of approximation is test by using statistical test called the coefficient of determination denoted by $\mathfrak{R}^2$. It is found that $\mathfrak{R}^2$= 0.9920 which confirms that the predicted cases fit well to the reported cases, see Figure 1. This study result indicates that the values of model parameters used in Table 1 are appropriate for the model (2.1) and shall be used to investigate the sensitivity analysis of reproductive numbers.

3.2 Sensitivity analysis

To develop an effective strategy for controlling and preventing influenza H1N1, it is essential to understand the relative significance of various factors contributing to disease transmission, which is directly linked to the reproductive number ($\mathcal{R}_0$). To this end, sensitivity analysis [6] is employed to evaluate how a state parameter responds to changes in a parameter by calculating sensitivity indices. The sensitivity indices of reproductive number using partial derivative (see more details in [6], therefore, is denoted by $\gamma_p^{\mathcal{R}_0}$ and is given by

$$\gamma_p^{\mathcal{R}_0} := \frac{\partial \mathcal{R}_0}{\partial p} \times \frac{p}{\mathcal{R}_0}, \tag{3.1}$$

Table 1: Parameter values for the H1N1 model.

Parameter	Definition	Value (day^{-1})
Π	Recruitment rate	480
β	Infection rate	0.173
k	The mean duration of latency	1/3
p	Transmission probability from exposed $0 \leq p \leq 1$ to symptomatic group,	0.67
γ	The recovery rate of symptomatic individuals	1/7
η	The recovery rate of asymptomatic individuals	1/7
μ	Natural death rate	0.00003634
α	Death rate from influenza H1N1	0.000306
ψ_1	Probability of transmission of infection from an exposed individual to a susceptible individual	0.6
ψ_2	Probability of transmission of infection from an asymptomatic individual to a susceptible individual	1

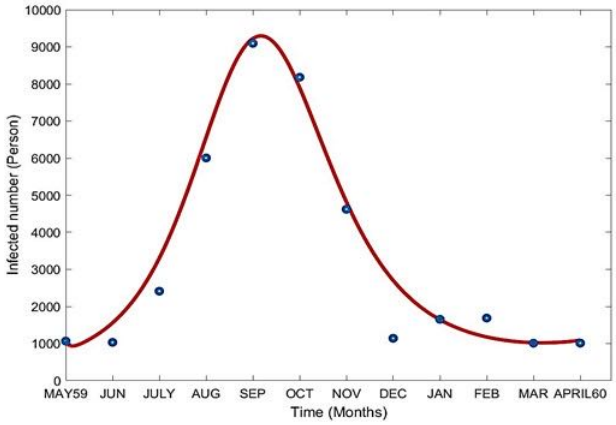


Figure 1: Comparison of predicted cases (red curve) and real cases (blue dots) reported by the Ministry of Public Health in Thailand, May 2016 - April 2017.

Table 2: Sensitivity indices of $\mathcal{R}_0$ with respect to model parameters that cause the spread of disease.

Parameter, p	Sensitivity indices, $\gamma_p^{\mathcal{R}_0}$
β	$+1$
ψ_1	$+0.26278$
ψ_2	$+0.20482$

where p denotes the model parameter in Table 1. The sensitivity indices of $\mathcal{R}_0$ to each of the five different parameters evaluated at the baseline parameter values given in Table 1 are tabulated in Tables 2-3. The negative sign means increasing (decreasing) a model parameter, the value of $\mathcal{R}_0$ would be decreased (increased) which leads to decrease (increase) infected population and vice versa in the case positive sign. The results are split into two sets: the set of model parameters that is impact on disease transmission see Table 2 and the set of model parameters that is impact on prevention see Table 3, respectively.

Table 3: Sensitivity indices of model parameters that can reduce the number of patients.

Parameter, p	Sensitivity indices, $\gamma_p^{\mathcal{R}_0}$
γ	-0.53112
η	-0.26272

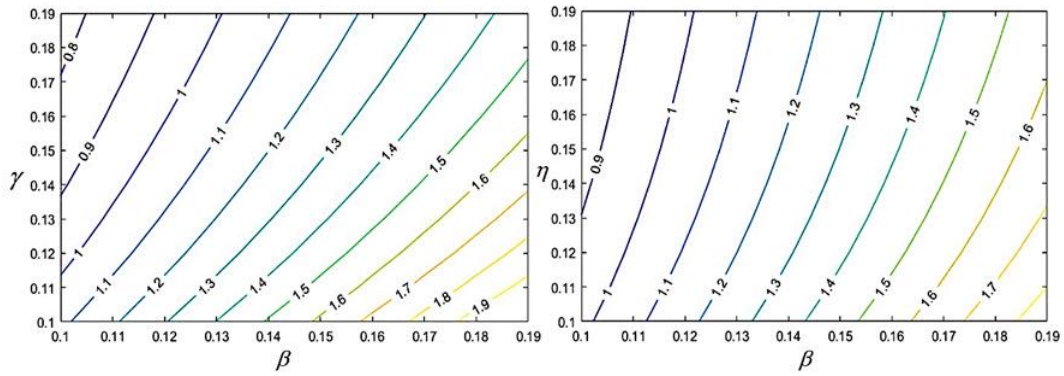
From Table 2, it can be observed that the primary cause of the influenza outbreak is direct contact with infected individuals. In contrast, contact with asymptomatic patients is a significant factor that contributes more to the outbreak than contact with individuals in the incubation period. When analyzing factors affecting disease control from Table 3, it is found that the duration of treatment for symptomatic patients impacts disease control. However, attention

should also be given to the surveillance duration of asymptomatic patients, as this can affect disease control since this group of individuals can spread the infection.

Next, an analysis will be conducted on the two control factors that influence the spread of the disease and disease control by plotting contour plots of reproduction number as the function of two control factors, as shown in Figures 2-4. The study results indicate that reducing the contact rate with patients and patients in the incubation period will reduce recovery time from illness (see Figures 2-3), thus enabling disease control. Furthermore, the study also indicates that reducing the transmission rate of asymptomatic patients will decrease the treatment time for symptomatic patients but does not affect the recovery time of asymptomatic patients (see Figure 4). These findings show that managing symptomatic patients is effective for disease control and prevention, whereas managing asymptomatic patients does not affect control but increases disease transmission. Therefore, the study suggests that an effectiveness control strategy must be addressed in surveillance measures for asymptomatic patients.

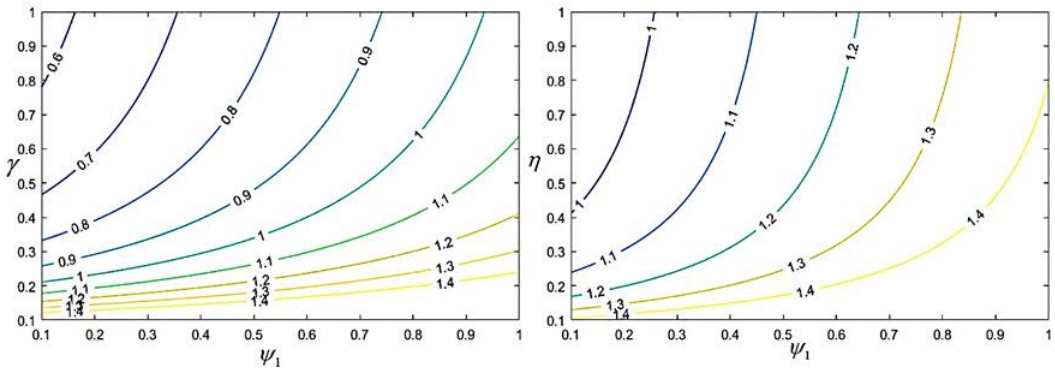
4 Discussion

The model of influenza H1N1 is constructed based on its epidemiology, and the values of all model parameters for this study are obtained by fitting the predicted data with the real cases of influenza H1N1 in Thailand from 2016 to 2017. Stability analysis of the model reveals that the disease dies out when $\mathcal{R}_0 \leq 1$ and the disease persists in the community whenever $\mathcal{R}_0 > 1$. This verifies that $\mathcal{R}_0$ is directly related to initial disease transmission. The sensitivity analysis reveals that the model parameters that cause the spread of disease are the infection rate and the probability of transmission of infection from an exposed individual and an asymptomatic individual to a susceptible individual, respectively. On the other hand, the model parameters that help in disease control are the recovery rates of symptomatic and asymptomatic individuals, respectively. The relation among risk factors is investigated. The study results reveal that decreasing the contact rate and the probability of transmission of infection from an asymptomatic individual to a susceptible individual and increasing the recovery rates of symptomatic and asymptomatic individuals will decrease the number of infectious populations. These study results suggest that the best ways to control disease are to beware of asymptomatic individuals and



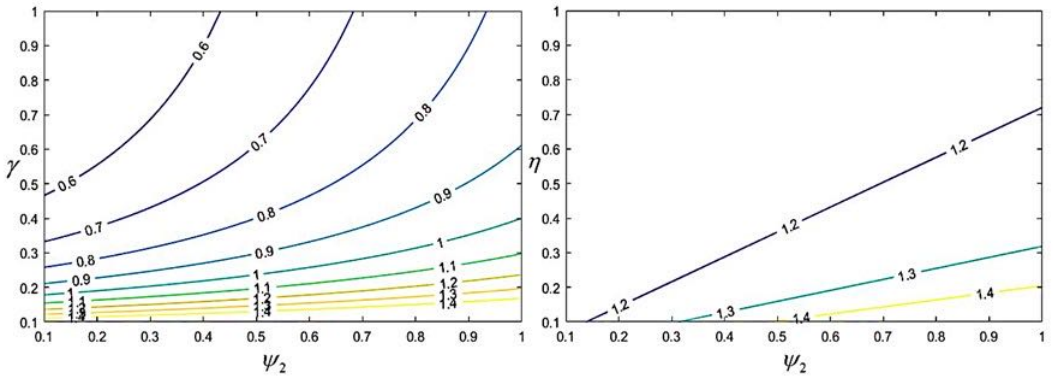
(a) (b)

Figure 2: Contour plot of reproductive number: (a) $\mathcal{R}_0(\beta, \gamma)$ is the function of β and γ; (b) $\mathcal{R}_0(\beta, \eta)$ is the function of β and η.



(a) (b)

Figure 3: Contour plot of reproductive number: (a) $\mathcal{R}_0(\psi_1, \gamma)$ is the function of ψ_1 and γ; (b) $\mathcal{R}_0(\psi_1, \eta)$ is the function of ψ_1 and η.



(a) (b)

Figure 4: Contour plot of reproductive number: (a) $\mathcal{R}_0(\psi_2, \gamma)$ is the function of ψ_2 and γ; (b) $\mathcal{R}_0(\psi_2, \eta)$ is the function of ψ_2 and η.

provide good treatments for exposed, symptomatic, including asymptomatic, individuals. The main results in this study effectively guide public policy for decision-making when influenza H1N1 is pandemic.

Acknowledgment. The authors would like to thank Department of Medical Sciences, Ministry of Public Health, the Institute of Disease Prevention and Control, Department of Disease Control, for the data of the number of influenza patients in the year 2007 - 2018.

References

[1] Center of Disease Control and Prevention (CDC), H1N1 Flu. <https://www.cdc.gov/h1n1flu/qa.htm>, 2019 (accessed 15.02.19).

[2] J. Arino, F. Brauer, P. Van Den Driessche, J. Watmough, J. Wu, Model for Influenza with vaccination and antiviral treatment, J. Theor. Biol. 253 (2008) 118–130.

[3] Md. Samsuzzoha, M. Singh, D. Lucy, Uncertainly and sensitivity analysis of the basic reproduction number of a vaccinated epidemic model of Influenza, Appl. Math. Modeling. 37 (3) (2012) 903–915.

[4] A. K. Srivastav, M. Ghosh, Modeling and analysis of the symptomatic and asymptomatic infections of swine flu with optimal control. Model. Earth Syst. Environ. 2 (177) (2016) 1–9.

[5] S. Kim, J. Lee, E. Jung, Mathematical model of transmission dynamics and optimal control strategies for 2009 A/H1N1 Influenza in the republic Korea, J. Theo. Biol. 412 (2017) 74–85.

[6] N. Chitnis, J.M. Hyman, J.M. Cushing, Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model, Bull. Math. Biol. 70 (2008) 1272–1296.

[7] O. Diekmann, J.A.P. Heesterbeek, J.A.J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, J. Math. Biol. 28 (1990) 365–382.

[8] P. van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Math. Biosci. 180 (2002) 29–48.

[9] J.P. LaSalle, The stability of dynamical systems. New Jersey: Hamilton Press, 1976.

[10] R.M. Anderson, R.M. May, Infectious Diseases of Humans, Dynamics and Control, Oxford University Press, London, New York, 1991.

พหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน

ธีรภัทร ไกรช่วย[1,*], พณิชพล รามมาก[1], อดิศักดิ์ เด็นเพ็ชรหน่อง[1]
และ ภัทราวรรณ สิงควานนท์[1,†]

สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสงขลา 90000

บทคัดย่อ

งานวิจัยนี้ศึกษาเกี่ยวกับพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน ซึ่งเป็นการต่อยอดจากงานวิจัยของ H. Arslan (2021) โดยได้ศึกษาความสัมพันธ์เวียนเกิด สูตรบิเนต ผลบวก ผลต่าง ผลคูณ และเอกลักษณ์แคสสินี ของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน นอกจากนี้ยังศึกษาผลบวก ผลต่าง และผลคูณ ระหว่างพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน กับสังยุค และระหว่างพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียนกับสังยุคในลักษณะเดียวกัน

คำสำคัญ: พหุนามเพลล์ พหุนามเพลล์-ลูคัส พหุนามเกาส์เซียนเพลล์ พหุนามเกาส์เซียนเพลล์-ลูคัส พหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน พหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน

2020 MSC: 11B37; 11B39

1 บทนำ

A. F. Horadam [1], [2] ได้แนะนำลำดับฟีโบนักชีเชิงซ้อน หลังจากนั้นผู้เขียนหลายคนก็ได้ศึกษาลำดับเกาส์เซียนฟีโบนักชี ลูคัส เพลล์ และเพลล์-ลูคัส ซึ่งลำดับเกาส์เซียนฟีโบนักชี และลูคัสถูกกำหนดโดยความสัมพันธ์เวียนเกิด $GF_{n+1} = GF_n + GF_{n-1}$ โดยที่ $GF_0 = i, GF_1 = 1$ และ $GL_{n+1} = GL_n + GL_{n-1}$ โดยที่ $GL_0 = 2 - i$, $GL_1 = 1 + 2i$ ตามลำดับ นอกจากนี้ลำดับเกาส์เซียนเพลล์ จะถูกกำหนดโดยความสัมพันธ์เวียนเกิด $GP_{n+1} = 2GP_n + GP_{n-1}$ โดยที่ $GP_0 = i, GP_1 = 1$ และลำดับเกาส์เซียนเพลล์-ลูคัส จะถูกกำหนดโดยความสัมพันธ์เวียนเกิด $GQ_{n+1} = 2GQ_n + GQ_{n-1}$ โดยที่ $GQ_0 = 2 - 2i, GQ_1 = 2 + 2i$ ในทางกลับกันลำดับพหุนามเพลล์ถูกกำหนดความสัมพันธ์เวียนเกิด $P_{n+1}(x) = 2xP_n(x) + P_{n-1}(x)$ โดยที่ $P_0(x) = 0, P_1(x) = 1$ ในทำนองเดียวกัน ลำดับพหุนามเพลล์-ลูคัส ถูกกำหนดโดย $Q_{n+1} = 2xQ_n(x)+Q_{n-1}(x)$ โดยที่ $Q_0(x) = 2, Q_1 = 2x$ นอกจากนี้สมบัติบางประการที่เกี่ยวข้องกับลำดับเหล่านี้ยังได้รับการศึกษาโดย A. F.

*ผู้นำเสนอ (644254034@parichat.skru.ac.th)
†ผู้ประพันธ์บรรณกิจ (pattarawan.pe@skru.ac.th)

Horadam และ Bro. J. M. Mahon [3] ต่อมา S. HALICI และ S. OZ [4] ได้นิยามพหุนามเกาส์เซียนเพลล์ โดยความสัมพันธ์เวียนเกิด $GP_{n+1}(x) = 2xGP_n(x) + GP_{n-1}(x)$ โดยที่ $GP_0(x) = i, GP_1(x) = 1$ T. Yagmur [5] ได้นิยามพหุนามเกาส์เซียนเพลล์-ลูคัส โดยความสัมพันธ์เวียนเกิด $GQ_{n+1}(x) = 2xGQ_n(x)+GQ_{n-1}(x)$ โดยที่ $GQ_0(x) = 2 - 2xi, GQ_1(x) = 2x + 2i$

งานวิจัยนี้ได้ขยายแนวคิดของ H. Arslan [6] ซึ่งควอเทอร์เนียนมีแนวคิดมาจากเซตของจำนวนเชิงซ้อน ได้ถูกคิดค้นโดย W.R. Hamilton ซึ่งเขียนแทนในรูปแบบ $k = k_0e_0 + k_1e_1 + k_2e_2 + k_3e_3 = (k_0, k_1, k_2, k_3)$ เมื่อ k_0, k_1, k_2 และ k_3 เป็นจำนวนจริง และ e_0, e_1, e_2 และ e_3 เป็นหน่วยควอเทอร์เนียน อีกทั้งในปี 2021 H. Arslan [6] ได้ให้บทนิยามของ ลำดับเกาส์เซียนเพลล์ควอเทอร์เนียน และลำดับเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน ดังนี้ $QGP_n = e_0GP_n + e_1GP_{n+1} + e_2GP_{n+2} + e_3GP_{n+3}$ และ $QGQ_n = e_0GQ_n + e_1GQ_{n+1} + e_2GQ_{n+2} + e_3GQ_{n+3}$ ตามลำดับ

งานวิจัยนี้มุ่งเน้นศึกษาความสัมพันธ์เวียนเกิด สูตรบิเนต ผลบวก ผลต่าง ผลคูณ และเอกลักษณ์แคสสินีของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน นอกจากนี้ยังศึกษาผลบวก ผลต่าง และผลคูณ ระหว่างพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน กับสังยุค และระหว่างพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน กับสังยุค

2 ความรู้พื้นฐาน

H. Arslan [6] ได้ให้สูตรบิเนตของลำดับเกาส์เซียนเพลล์ควอเทอร์เนียน และลำดับเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน ดังนี้

$$QGP_n = \frac{\alpha^n A - \beta^n B}{\alpha - \beta} \text{ และ } QGQ_n = \frac{\alpha^n A + \beta^n B}{2} \text{ ตามลำดับ}$$

เมื่อ $\alpha = 1 + \sqrt{2}$, $\beta = 1 - \sqrt{2}$, $\alpha + \beta = 2$, $\alpha - \beta = 2\sqrt{2}$, $\alpha\beta = -1$

โดยที่ $A = \sum_{s=0}^{3}(\alpha^s + i\alpha^{s-1})e_s$ และ $B = \sum_{s=0}^{3}(\beta^s + i\beta^{s-1})e_s$

สังยุคของลำดับเกาส์เซียนเพลล์ควอเทอร์เนียน QGP_n^* และลำดับเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน QGQ_n^* กำหนดโดย

$$\begin{aligned} QGP_n^* &= e_0GP_n - e_1GP_{n+1} - e_2GP_{n+2} - e_3GP_{n+3} \text{ และ} \\ QGQ_n^* &= e_0GQ_n - e_1GQ_{n+1} - e_2GQ_{n+2} - e_3GQ_{n+3} \text{ ตามลำดับ} \end{aligned}$$

สังยุคเชิงซ้อนของลำดับเกาส์เซียนเพลล์ควอเทอร์เนียน $\overline{QGP_n}$ และลำดับเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน $\overline{QGQ_n}$ กำหนดโดย

$$\begin{aligned} \overline{QGP_n} &= e_0\overline{GP_n} + e_1\overline{GP_{n+1}} + e_2\overline{GP_{n+2}} + e_3\overline{GP_{n+3}} \text{ และ} \\ \overline{QGQ_n} &= e_0\overline{GQ_n} + e_1\overline{GQ_{n+1}} + e_2\overline{GQ_{n+2}} + e_3\overline{GQ_{n+3}} \text{ ตามลำดับ} \end{aligned}$$

A. F. Horadam และ Bro. J. M.Mahon [3] ได้ให้เอกลักษณ์ของพหุนามเพลล์ และพหุนามเพล์-ลูคัส ดังนี้

$$P_{n+1}(x) + P_{n-1}(x) = Q_n(x) = 2xP_n(x) + 2P_{n-1}(x) \tag{2.1}$$

$$Q_{n+1}(x) + Q_{n-1}(x) = (4x^2 + 4)P_n(x) \tag{2.2}$$

$$P_n(x)Q_n(x) = P_{2n}(x) \tag{2.3}$$

$$P_{n+2}(x) + P_{n-2}(x) = (4x^2 + 2)P_n(x) \tag{2.4}$$

$$P^2_{n+2}(x) - P^2_{n-2}(x) = (8x^3 + 4x)P_{2n}(x) \tag{2.5}$$

$$Q^2_{n+1}(x) - Q^2_{n-1}(x) = 2x(4x^2 + 4)P_{2n}(x) \tag{2.6}$$

$$\alpha^n(x) = \alpha(x)P_n(x) + P_{n-1}(x) \tag{2.7}$$

$$\beta^n(x) = \beta(x)P_n(x) + P_{n-1}(x) \tag{2.8}$$

เมื่อ $\alpha(x) = x + \sqrt{1 + x^2}$, $\beta(x) = x - \sqrt{1 + x^2}$

S. HALICI และ S. OZ [4] ได้ให้ความสัมพันธ์ระหว่างพหุนามเกาส์เซียนเพลล์ กับพหุนามเพลล์ ดังนี้

$$GP_n(x) = P_n(x) + iP_{n-1}(x) \tag{2.9}$$

โดยที่ $P_0(x) = 0$ และ $P_1(x) = 1$
และพหุนามเกาส์เซียนเพล์เขียนในรูปความสัมพันธ์เวียนเกิด ดังนี้

$$GP_{n+1}(x) = 2xGP_n(x) + GP_{n-1}(x) \tag{2.10}$$

โดยที่ $GP_0(x) = i$ และ $GP_1(x) = 1$
สูตรบิเนตของพหุนามเกาส์เซียนเพลล์ คือ

$$GP_n(x) = \frac{\alpha^n(x) - \beta^n(x)}{\alpha(x) - \beta(x)} + i\frac{\alpha(x)\beta^n(x) - \alpha^n(x)\beta(x)}{\alpha(x) - \beta(x)}$$

เมื่อ $\alpha(x) = x + \sqrt{1 + x^2}$, $\beta(x) = x - \sqrt{1 + x^2}$

T. Yagmur [5] ความสัมพันธ์ระหว่างพหุนามเกาส์เซียนเพลล์-ลูคัส กับพหุนามเพลล์-ลูคัส เขียนได้ดังนี้

$$GQ_n(x) = Q_n(x) + iQ_{n-1}(x) \tag{2.11}$$

โดยที่ $Q_0(x) = 2$ และ $Q_1(x) = 2x$
และพหุนามเกาส์เซียนเพลล์-ลูคัสเขียนในรูปความสัมพันธ์เวียนเกิด ดังนี้

$$GQ_{n+1}(x) = 2xGQ_n(x) + GQ_{n-1}(x) \tag{2.12}$$

โดยที่ $GQ_0(x) = 2 - 2xi$ และ $GQ_1(x) = 2x + 2i$
สูตรบิเนตของพหุนามเกาส์เซียนเพลล์-ลูคัส คือ

$$GQ_n(x) = \alpha^n(x) + \beta^n(x) - [\beta(x)\alpha^n(x) + \alpha(x)\beta^n(x)]i$$

เมื่อ $\alpha(x) = x + \sqrt{1 + x^2}$, $\beta(x) = x - \sqrt{1 + x^2}$

3 ผลการศึกษา

ผลการวิจัยในครั้งนี้ ได้แบ่งศึกษาออกเป็น 2 หัวข้อหลัก คือ เอกลักษณ์ของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และเอกลักษณ์ของพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน โดยศึกษาความสัมพันธ์เวียนเกิด สูตรบิเนต เอกลักษณ์แคสสินี นอกจากนี้ยังได้ศึกษาผลบวก ผลต่าง ผลคูณ และเอกลักษณ์อื่น ๆ ของทั้งสองพหุนามในลักษณะเดียวกัน

3.1 เอกลักษณ์ของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน

บทนิยาม 3.1. ให้ n เป็นจำนวนเต็มที่ไม่เป็นลบ พหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน กำหนดโดย

$$QGP_n(x) = e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x)$$

เมื่อ e_0, e_1, e_2 และ e_3 เป็นหน่วยควอเทอร์เนียน โดยที่ $e_0^2 = 1, e_0e_i = e_ie_0 = e_i, i = 1, 2, 3$
$e_1^2 = e_2^2 = e_3^2 = e_1e_2e_3 = -1$ และ $GP_n(x)$ เป็นพจน์ที่ n ของพหุนามเกาส์เซียนเพลล์

ทฤษฎีบท 3.2. *ให้ n เป็นจำนวนเต็มบวก จะได้ความสัมพันธ์เวียนเกิด ดังนี้*

$$QGP_n(x) + 2xQGP_{n+1}(x) = QGP_{n+2}(x)$$

พิสูจน์. $QGP_n(x) + 2xQGP_{n+1}(x)$

$$\begin{aligned}
&= e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x) \\
&\quad +2x(e_0GP_{n+1}(x) + e_1GP_{n+2}(x) + e_2GP_{n+3}(x) + e_3GP_{n+4}(x)) \\
&= e_0(2xGP_{n+1}(x) + GP_n(x)) + e_1(2xGP_{n+2}(x) + GP_{n+1}(x)) \\
&\quad +e_2(2xGP_{n+3}(x) + GP_{n+2}(x)) + e_3(2xGP_{n+4}(x) + GP_{n+3}(x)) \\
&= e_0GP_{n+2}(x) + e_1GP_{n+3}(x) + e_2GP_{n+4}(x) + e_3GP_{n+5}(x) \\
&= QGP_{n+2}(x)
\end{aligned}$$

□

จากทฤษฎีบท 3.2 ให้ $\alpha(x)$ และ $\beta(x)$ เป็นรากของสมการ $QGP_n(x) + 2xQGP_{n+1}(x) = QGP_{n+2}(x)$ จะได้ $\alpha(x) = x + \sqrt{1+x^2}$, $\beta(x) = x - \sqrt{1+x^2}$, $\alpha(x) + \beta(x) = 2x$, $\alpha(x) - \beta(x) = 2\sqrt{x^2+1}$ และ $\alpha(x)\beta(x) = -1$

บทตั้ง 3.3. *ให้ n เป็นจำนวนเต็มบวก จะได้*

$$\begin{aligned}
\alpha(x)QGP_n(x) + QGP_{n-1}(x) &= \alpha^n(x)A \\
\beta(x)QGP_n(x) + QGP_{n-1}(x) &= \beta^n(x)B
\end{aligned}$$

โดยที่ $A = \sum_{k=0}^{3}(\alpha^k(x) + i\alpha^{k-1}(x))e_k$ *และ* $B = \sum_{k=0}^{3}(\beta^k(x) + i\beta^{k-1}(x))e_k$

พิสูจน์. $GP_n(x) = P_n(x) + iP_{n-1}(x)$ และ (2.7) : $\alpha^n(x) = \alpha(x)P_n(x) + P_{n-1}(x)$

$\alpha(x)QGP_n(x) + QGP_{n-1}(x)$

$$\begin{aligned}
&= \alpha(x)(e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x)) \\
&\quad + (e_0GP_{n-1}(x) + e_1GP_n(x) + e_2GP_{n+1}(x) + e_3GP_{n+2}(x))
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{3} e_k(\alpha(x)GP_{n+k}(x) + GP_{n+k-1}(x)) \\
&= \sum_{k=0}^{3} e_k[(\alpha(x)P_{n+k}(x) + i\alpha(x)P_{n+k-1}(x)) + (P_{n+k-1}(x) + iP_{n+k-2}(x))] \\
&= \sum_{k=0}^{3} e_k[\alpha^{n+k}(x) + i\alpha^{n+k-1}(x)] \\
&= \alpha^n(x)A
\end{aligned}
$$

และ $GP_n(x) = P_n(x) + iP_{n-1}(x)$ และ (2.8) : $\beta^n(x) = \beta(x)P_n(x) + P_{n-1}(x)$

$$
\begin{aligned}
&\beta(x)QGP_n(x) + QGP_{n-1}(x) \\
&\quad = \beta(x)(e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x)) \\
&\qquad + (e_0GP_{n-1}(x) + e_1GP_n(x) + e_2GP_{n+1}(x) + e_3GP_{n+2}(x)) \\
&\quad = \sum_{k=0}^{3} e_k(\beta(x)GP_{n+k}(x) + GP_{n+k-1}(x)) \\
&\quad = \sum_{k=0}^{3} e_k[(\beta(x)P_{n+k}(x) + P_{n+k-1}(x)) + (\beta(x)iP_{n+k-1}(x) + iP_{n+k-2}(x))] \\
&\quad = \sum_{k=0}^{3} e_k[\beta^n(x)(\beta^k(x) + i\beta^{k-1}(x))] \\
&\quad = \beta^n(x)B
\end{aligned}
$$

□

ทฤษฎีบท 3.4. *(สูตรบิเนต) ให้ n เป็นจำนวนเต็มบวก สูตรบิเนตสำหรับพหุนามเกาส์เซียนเพลล์ควอเทอร์นียน คือ*

$$QGP_n(x) = \frac{\alpha^n(x)A - \beta^n(x)B}{\alpha(x) - \beta(x)}$$

โดยที่ $A = \sum_{k=0}^{3}(\alpha^k(x) + i\alpha^{k-1}(x))e_k$ *และ* $B = \sum_{k=0}^{3}(\beta^k(x) + i\beta^{k-1}(x))e_k$

พิสูจน์. จากบทตั้ง 3.3 สามารถนำมาพิสูจน์ได้ดังนี้

$$
\begin{aligned}
\frac{\alpha^n(x)A - \beta^n(x)B}{\alpha(x) - \beta(x)} &= \frac{(\alpha(x)QGP_n(x) + QGP_{n-1}(x)) - (\beta(x)QGP_n(x) + QGP_{n-1}(x))}{\alpha(x) - \beta(x)} \\
&= QGP_n(x)
\end{aligned}
$$

□

ทฤษฎีบท 3.5. *(เอกลักษณ์แคสสินี) ให้ n เป็นจำนวนเต็มบวก เอกลักษณ์ต่อไปนี้เป็นจริง*

$$QGP_{n+1}(x)QGP_{n-1}(x) - QGP_n^2(x) = (-1)^n AB$$

โดยที่ $A = \sum_{k=0}^{3}(\alpha^k(x) + i\alpha^{k-1}(x))e_k$ *และ* $B = \sum_{k=0}^{3}(\beta^k(x) + i\beta^{k-1}(x))e_k$

พิสูจน์. ใช้ผล จากทฤษฎีบท 3.4
$QGP_{n+1}(x)QGP_{n-1}(x) - QGP_n^2(x)$

$$= (\frac{\alpha^{n+1}(x)A - \beta^{n+1}(x)B}{\alpha(x) - \beta(x)})(\frac{\alpha^{n-1}(x)A - \beta^{n-1}(x)B}{\alpha(x) - \beta(x)}) - (\frac{\alpha^n(x)A - \beta^n(x)B}{\alpha(x) - \beta(x)})^2$$

$$= \frac{1}{(\alpha(x) - \beta(x))^2}[-\alpha^n(x)\alpha(x)\beta^n(x)\beta^{-1}(x)AB - \beta^n(x)\beta(x)\alpha^n(x)\alpha^{-1}(x)BA + 2\alpha^n(x)\beta^n(x)AB]$$

$$= \frac{1}{(\alpha(x) - \beta(x))^2}(\alpha(x)\beta(x))^n AB[(-\alpha(x)\beta^{-1}(x) - \beta(x)\alpha^{-1}(x) + 2)]$$

$$= \frac{1}{(\alpha(x) - \beta(x))^2}(-1)^n AB[-\frac{\alpha(x)}{\beta(x)} - \frac{\beta(x)}{\alpha(x)} + 2]$$

$$= \frac{1}{(\alpha(x) - \beta(x))^2}(-1)^n AB[-\frac{\alpha^2(x)}{-1} - \frac{\beta^2(x)}{-1} + 2]$$

$$= \frac{1}{(\alpha(x) - \beta(x))^2}(-1)^n AB[(\alpha(x) - \beta(x))^2]$$

$$= (-1)^n AB$$ □

บทนิยาม 3.6. สังยุคของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน กำหนดโดย

$$QGP_n^*(x) = e_0GP_n(x) - e_1GP_{n+1}(x) - e_2GP_{n+2}(x) - e_3GP_{n+3}(x)$$

เมื่อ e_0, e_1, e_2 และ e_3 เป็นหน่วยควอเทอร์เนียน โดยที่ $e_0^2 = 1, e_0e_i = e_ie_0 = e_i, i = 1, 2, 3$
$e_1^2 = e_2^2 = e_3^2 = e_1e_2e_3 = -1$ และ $GP_n(x)$ เป็นพจน์ที่ n ของพหุนามเกาส์เซียนเพลล์

บทนิยาม 3.7. สังยุคเชิงซ้อนของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน กำหนดโดย

$$\overline{QGP_n(x)} = e_0\overline{GP_n(x)} + e_1\overline{GP_{n+1}(x)} + e_2\overline{GP_{n+2}(x)} + e_3\overline{GP_{n+3}(x)}$$

เมื่อ e_0, e_1, e_2 และ e_3 เป็นหน่วยควอเทอร์เนียน โดยที่ $e_0^2 = 1, e_0e_i = e_ie_0 = e_i, i = 1, 2, 3$
$e_1^2 = e_2^2 = e_3^2 = e_1e_2e_3 = -1$ และ $GP_n(x)$ เป็นพจน์ที่ n ของพหุนามเกาส์เซียนเพลล์

บทตั้ง 3.8. *สำหรับพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน $QGP_n(x)$ สอดคล้องกับเอกลักษณ์ดังต่อไปนี้*

$$QGP_n(x)QGP_n^*(x) = GP_n^2(x) + GP_{n+1}^2(x) + GP_{n+2}^2(x) + GP_{n+3}^2(x)$$

พิสูจน์. จากบทนิยาม 3.1 และบทนิยาม 3.6 จะได้ว่า

$$\begin{aligned} &QGP_n(x)QGP_n^*(x) \\ &\quad = (e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x)) \\ &\qquad (e_0GP_n(x) - e_1GP_{n+1}(x) - e_2GP_{n+2}(x) - e_3GP_{n+3}(x)) \\ &\quad = e_0^2GP_n^2(x) - e_1^2GP_{n+1}^2(x) - e_2^2GP_{n+2}^2(x) - e_3^2GP_{n+3}^2(x) \\ &\quad = GP_n^2(x) + GP_{n+1}^2(x) + GP_{n+2}^2(x) + GP_{n+3}^2(x) \end{aligned}$$ □

ทฤษฎีบท 3.9. *พหุนามเกาส์เซียนเพลล์ควอเทอร์เนียนสอดคล้องกับเอกลักษณ์ ดังต่อไปนี้*

$$QGP_n(x)QGP_n^*(x) = (x+i)(8x^2+4)P_{2n+2}(x)$$

พิสูจน์. จากบทตั้ง 3.8, (2.1), (2.3), (2.4) และ (2.5)

$$\begin{aligned}
QGP_n(x)QGP_n^*(x) &= GP_n^2(x) + GP_{n+1}^2(x) + GP_{n+2}^2(x) + GP_{n+3}^2(x) \\
&= P_{n+3}^2(x) - P_{n-1}^2(x) + 2i(P_n(x)(P_{n-1}(x) + P_{n+1}(x)) \\
&\quad + P_{n+2}(x)(P_{n+1}(x) + P_{n+3}(x))) \\
&= P_{n+3}^2(x) - P_{n-1}^2(x) + 2i(P_n(x)Q_n(x)) + (P_{n+2}(x)Q_{n+2}(x))) \\
&= P_{n+3}^2(x) - P_{n-1}^2(x) + 2i(P_{2n}(x)) + P_{2(n+2)}(x)) \\
&= (8x^3+4x)P_{2(n+1)}(x) + 2i(4x^2+2)P_{2(n+1)}(x) \\
&= (x+i)(8x^2+4)P_{2n+2}(x)
\end{aligned}$$

□

ทฤษฎีบท 3.10. *สำหรับพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน $QGP_n(x)$ สังยุคของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน $QGP_n^*(x)$ และสังยุคเชิงซ้อนของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน $\overline{QGP_n}(x)$ จะได้เอกลักษณ์ดังต่อไปนี้เป็นจริง*

$$QGP_n(x) + QGP_n^*(x) = 2e_0GP_n(x)$$

$$QGP_n(x) + \overline{QGP_n(x)} = 2\sum_{k=0}^{3} e_kP_{n+k}(x)$$

$$QGP_n^2(x) + QGP_n(x)QGP_n^*(x) = 2e_0GP_n(x)QGP_n(x)$$

พิสูจน์.

$$\begin{aligned}
QGP_n(x) + QGP_n^*(x) &= (e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x)) \\
&\quad + (e_0GP_n(x) - e_1GP_{n+1}(x) - e_2GP_{n+2}(x) - e_3GP_{n+2}(x)) \\
&= 2e_0GP_n(x)
\end{aligned}$$

$$\begin{aligned}
QGP_n(x) + \overline{QGP_n(x)} &= (e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x)) \\
&\quad + (e_0\overline{GP_n(x)} + e_1\overline{GP_{n+1}(x)} + e_2\overline{GP_{n+2}(x)} + e_3\overline{GP_{n+3}(x)}) \\
&= 2e_0P_n(x) + 2e_1P_{n+1}(x) + 2e_2P_{n+2}(x) + 2e_3P_{n+3}(x) \\
&= 2\sum_{k=0}^{3} e_kP_{n+k}(x)
\end{aligned}$$

และ

$$\begin{aligned}
QGP_n^2(x) + QGP_n(x)QGP_n^*(x) &= QGP_n(x)(QGP_n(x) + QGP_n^*(x)) \\
&= 2e_0GP_n(x)QGP_n(x)
\end{aligned}$$

□

ทฤษฎีบท 3.11. *สำหรับพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน $QGP_n(x)$ สังยุคของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน $QGP_n^*(x)$ และสังยุคเชิงซ้อนของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน $\overline{QGP_n(x)}$ จะได้เอกลักษณ์ดังต่อไปนี้เป็นจริง*

$$QGP_n(x) - QGP_n^*(x) = 2\sum_{k=1}^{3} e_kGP_{n+k}(x)$$

$$QGP_n(x) - \overline{QGP_n(x)} = 2i\sum_{k=-1}^{2} e_kP_{n+k}(x)$$

$$QGP_n^2(x) - QGP_n(x)QGP_n^*(x) = 2QGP_n(x)\sum_{k=1}^{3} e_k GP_{n+k}(x)$$

พิสูจน์. $QGP_n(x) - QGP_n^*(x) = (e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x))$
$$- (e_0GP_n(x) - e_1GP_{n+1}(x) - e_2GP_{n+2}(x) - e_3GP_{n+2}(x))$$
$$= 2\sum_{k=1}^{3} e_k GP_{n+k}(x)$$
$$QGP_n(x) - \overline{QGP_n(x)} = (e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x))$$
$$-(e_0\overline{GP_n(x)} + e_1\overline{GP_{n+1}(x)} + e_2\overline{GP_{n+2}(x)} + e_3\overline{GP_{n+3}(x)})$$
$$= 2ie_0P_{n-1}(x) + 2ie_1P_n(x) + 2ie_2P_{n+1}(x) + 2ie_3P_{n+2}(x)$$
$$= 2i\sum_{k=-1}^{2} e_k P_{n+k}(x)$$

และ

$$QGP_n^2(x) - QGP_n(x)QGP_n^*(x) = QGP_n(x)(QGP_n(x) - QGP_n^*(x))$$
$$= 2QGP_n(x)\sum_{k=1}^{3} e_k GP_{n+k}(x)$$ □

3.2 เอกลักษณ์ของพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน

บทนิยาม 3.12. ให้ n เป็นจำนวนเต็มที่ไม่เป็นลบ พหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน กำหนดโดย

$$QGQ_n(x) = e_0GQ_n(x) + e_1GQ_{n+1}(x) + e_2GQ_{n+2}(x) + e_3GQ_{n+3}(x)$$

เมื่อ e_0, e_1, e_2 และ e_3 เป็นหน่วยควอเทอร์เนียน โดยที่ $e_0^2 = 1, e_0e_i = e_ie_0 = e_i,\ i = 1, 2, 3,$
$e_1^2 = e_2^2 = e_3^2 = e_1e_2e_3 = -1$ และ $GQ_n(x)$ เป็นพจน์ที่ n ของพหุนามเกาส์เซียนเพลล์-ลูคัส

ทฤษฎีบท 3.13. *ให้ n เป็นจำนวนเต็มบวก จะได้ความสัมพันธ์เวียนเกิด ดังนี้*

$$QGQ_n(x) + 2xQGQ_{n+1}(x) = QGQ_{n+2}(x)$$

พิสูจน์. ในทำนองเดียวกันกับ ทฤษฎีบท 3.2 □

จากทฤษฎีบท 3.13 ให้ $\alpha(x)$ และ $\beta(x)$ เป็นรากของสมการ $QGQ_n(x) + 2xQGQ_{n+1}(x) = QGQ_{n+2}(x)$ จะได้ $\alpha(x) = x + \sqrt{1+x^2},\ \beta(x) = x - \sqrt{1+x^2},\ \alpha(x) + \beta(x) = 2x,\ \alpha(x) - \beta(x) = 2\sqrt{x^2+1}$ และ $\alpha(x)\beta(x) = -1$

บทตั้ง 3.14. *ให้ n เป็นจำนวนเต็มบวก จะได้*

$$2xGP_{n+1}(x) + 2GP_n(x) = GQ_{n+1}(x)$$

พิสูจน์. จาก (2.1), (2.9) และ (2.11)
$2xGP_{n+1}(x) + 2GP_n(x) = 2x(P_{n+1}(x) + iP_n(x)) + 2(P_n(x) + iP_{n-1}(x))$
$$= (2xP_{n+1}(x) + 2xiP_n(x)) + (2P_n(x) + 2iP_{n-1}(x))$$

$$\begin{aligned}
&= (2xP_{n+1}(x) + 2P_n(x)) + (2xiP_n(x) + 2iP_{n-1}(x)) \\
&= (2xP_{n+1}(x) + 2P_n(x)) + i(2xP_n(x) + 2P_{n-1}(x)) \\
&= Q_{n+1}(x) + iQ_n(x) \\
&= GQ_{n+1}(x)
\end{aligned}$$

□

ทฤษฎีบท 3.15. *ให้ n เป็นจำนวนเต็มบวก จะได้*

$$2xQGP_{n+1}(x) + 2QGP_n(x) = QGQ_{n+1}(x)$$

พิสูจน์. จากบทตั้ง 3.14

$$\begin{aligned}
&2xQGP_{n+1}(x) + 2QGP_n(x) \\
&\quad = 2x(e_0GP_{n+1}(x) + e_1GP_{n+2}(x) + e_2GP_{n+3}(x) + e_3GP_{n+4}(x)) \\
&\qquad + 2(e_0GP_n(x) + e_1GP_{n+1}(x) + e_2GP_{n+2}(x) + e_3GP_{n+3}(x)) \\
&\quad = (2xe_0GP_{n+1}(x) + 2e_0GP_n(x)) + (2xe_1GP_{n+2}(x) + 2e_1GP_{n+1}(x)) \\
&\qquad + (2xe_2GP_{n+3}(x) + 2e_2GP_{n+2}(x)) + (2xe_3GP_{n+4}(x) + 2e_3GP_{n+3}(x)) \\
&\quad = e_0(2xGP_{n+1}(x) + 2GP_n(x)) + e_1(2xGP_{n+2}(x) + 2GP_{n+1}(x)) \\
&\qquad + e_2(2xGP_{n+3}(x) + 2GP_{n+2}(x)) + e_3(2xGP_{n+4}(x) + 2GP_{n+3}(x)) \\
&\quad = e_0GQ_{n+1}(x) + e_1GQ_{n+2}(x) + e_2GQ_{n+3}(x) + e_3GQ_{n+4}(x) \\
&\quad = QGQ_{n+1}(x)
\end{aligned}$$

□

ทฤษฎีบท 3.16. *(สูตรบิเนต) ให้ n เป็นจำนวนเต็มบวก สูตรบิเนตสำหรับพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน คือ*

$$QGQ_n(x) = \alpha^n(x)A + \beta^n(x)B$$

โดยที่ $A = \sum_{k=0}^{3}(\alpha^k(x) + i\alpha^{k-1}(x))e_k$ *และ* $B = \sum_{k=0}^{3}(\beta^k(x) + i\beta^{k-1}(x))e_k$

พิสูจน์. จากบทตั้ง 3.3 ทฤษฎีบท 3.15 และ $\alpha(x) + \beta(x) = 2x$

$$\begin{aligned}
\alpha^n(x)A + \beta^n(x)B &= (\alpha(x)QGP_n(x) + QGP_{n-1}(x)) + (\beta(x)QGP_n(x)) + QGP_{n-1}(x)) \\
&= (\alpha(x)QGP_n(x) + \beta(x)QGP_n(x)) + 2QGP_{n-1}(x) \\
&= (\alpha(x) + \beta(x))QGP_n(x) + 2QGP_{n-1}(x) \\
&= 2xQGP_n(x) + 2QGP_{n-1}(x) \\
&= QGQ_n(x)
\end{aligned}$$

□

ทฤษฎีบท 3.17. *(เอกลักษณ์แคสสินี) ให้ n เป็นจำนวนเต็มบวก เอกลักษณ์ต่อไปนี้เป็นจริง*

$$QGQ_{n+1}(x)QGQ_{n-1}(x) - QGQ_n^2(x) = (-1)^{n+1}(\alpha(x) - \beta(x))^2AB$$

โดยที่ $A = \sum_{k=0}^{3}(\alpha^k(x) + i\alpha^{k-1}(x))e_k$ *และ* $B = \sum_{k=0}^{3}(\beta^k(x) + i\beta^{k-1}(x))e_k$

พิสูจน์. ใช้ผลจากทฤษฎีบท 3.16

$$\begin{aligned} QGQ_{n+1}(x)QGQ_{n-1}(x) - QGQ_n^2(x) \\ &= (\alpha^{n+1}(x)A + \beta^{n+1}(x)B)(\alpha^{n-1}(x)A + \beta^{n-1}(x)B) - (\alpha^n(x)A + \beta^n(x)B)^2 \\ &= \alpha^{n+1}(x)A\beta^{n-1}(x)B + \beta^{n+1}(x)B\alpha^{n-1}(x)A - 2\alpha^n(x)\beta^n(x)AB \\ &= (\alpha(x)\beta(x))^n AB[(\alpha^1(x)\beta^{-1}(x) + \beta^1(x)\alpha^{-1}(x) - 2)] \\ &= (-1)^n(-\alpha^2(x) - \beta^2(x) - 2)AB \\ &= (-1)^{n+1}(\alpha(x) - \beta(x))^2 AB \end{aligned}$$

□

ทฤษฎีบท 3.18. *สำหรับพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียนใด ๆ และพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียนใด ๆ จะได้*

$$\frac{QGQ_n^2(x)}{4} - (x^2+1)QGP_n^2(x) = (-1)^n AB$$

โดยที่ $A = \sum_{k=0}^{3}(\alpha^k(x) + i\alpha^{k-1}(x))e_k$ *และ* $B = \sum_{k=0}^{3}(\beta^k(x) + i\beta^{k-1}(x))e_k$

พิสูจน์. $\frac{QGQ_n^2(x)}{4} - (x^2+1)QGP_n^2(x)$

$$\begin{aligned} &= \frac{(\alpha^n(x)A + \beta^n(x)B)^2}{4} - (x^2+1)(\frac{\alpha^n(x)A - \beta^n(x)B}{\alpha(x) - \beta(x)})^2 \\ &= \frac{(\alpha^{2n}(x)A^2 + 2\alpha^n(x)\beta^n(x)AB + \beta^{2n}(x)B^2)}{4} - (x^2+1)\frac{(\alpha^{2n}(x)A^2 - 2\alpha^n(x)\beta^n(x)AB + \beta^{2n}(x)B^2}{4x^2+4} \\ &= \frac{\alpha^{2n}(x)A^2 + 2\alpha^n(x)\beta^n(x)AB + \beta^{2n}(x)B^2}{4} - \frac{\alpha^{2n}(x)A^2 - 2\alpha^n(x)\beta^n(x)AB + \beta^{2n}(x)B^2}{4} \\ &= \frac{4\alpha^n(x)\beta^n(x)AB}{4} \\ &= (-1)^n AB \end{aligned}$$

□

บทนิยาม 3.19. สังยุคของพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน กำหนดโดย

$$QGQ_n^*(x) = e_0GQ_n(x) - e_1GQ_{n+1}(x) - e_2GQ_{n+2}(x) - e_3GQ_{n+3}(x)$$

เมื่อ e_0, e_1, e_2 และ e_3 เป็นหน่วยควอเทอร์เนียน โดยที่ $e_0^2 = 1, e_0e_i = e_ie_0 = e_i$, $i = 1, 2, 3$, $e_1^2 = e_2^2 = e_3^2 = e_1e_2e_3 = -1$ และ $GQ_n(x)$ เป็นพจน์ที่ n ของพหุนามเกาส์เซียนเพลล์-ลูคัส

บทนิยาม 3.20. สังยุคเชิงซ้อนของพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน-ลูคัส กำหนดโดย

$$\overline{QGQ_n(x)} = e_0\overline{GQ_n(x)} + e_1\overline{GQ_{n+1}(x)} + e_2\overline{GQ_{n+2}(x)} + e_3\overline{GQ_{n+3}(x)}$$

เมื่อ e_0, e_1, e_2 และ e_3 เป็นหน่วยควอเทอร์เนียน โดยที่ $e_0^2 = 1, e_0e_i = e_ie_0 = e_i$, $i = 1, 2, 3$, $e_1^2 = e_2^2 = e_3^2 = e_1e_2e_3 = -1$ และ $GQ_n(x)$ เป็นพจน์ที่ n ของพหุนามเกาส์เซียนเพลล์-ลูคัส

บทตั้ง 3.21. *พหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียนสอดคล้องกับเอกลักษณ์ ดังต่อไปนี้*

$$QGQ_n(x)QGQ_n^*(x) = GQ_n^2(x) + GQ_{n+1}^2(x) + GQ_{n+2}^2(x) + GQ_{n+3}^2(x)$$

พิสูจน์. ในทำนองเดียวกันกับ บทตั้ง 3.8 □

ทฤษฎีบท 3.22. *พหุนามเกาส์เซียนเพลล์เพลล์-ลูคัสควอเทอร์เนียนสอดคล้องกับเอกลักษณ์ ดังต่อไปนี้*

$$QGQ_n(x)QGQ_n^*(x) = ((4x^2+4)(8x^3+4x) + (8ix^2+8i)(4x^2+2))P_{2n+2}(x)$$

พิสูจน์. จากบทตั้ง 3.21, (2.2), (2.3), (2.4), (2.6) และ (2.11)

$$\begin{aligned}
QGQ_n(x)QGQ_n^*(x) &= GQ_n^2(x) + GQ_{n+1}^2(x) + GQ_{n+2}^2(x) + GQ_{n+3}^2(x) \\
&= Q_{n+3}^2(x) - Q_{n-1}^2(x) + 2i(Q_n(x)(Q_{n-1}(x) + Q_{n+1}(x)) \\
&\quad + Q_{n+2}(x)(Q_{n+1}(x) + Q_{n+3}(x))) \\
&= Q_{n+3}^2(x) - Q_{n-1}^2(x) + 2i(4(x^2+1)P_{2n}(x) + 4(x^2+1)P_{2(n+2)}(x)) \\
&= (4x^2+4)(8x^3+4x)P_{2n+2}(x) + (8ix^2+8i)P_{2n}(x) \\
&\quad + (8ix^2+8i)P_{2n+4}(x) \\
&= ((4x^2+4)(8x^3+4x) + (8ix^2+8i)(4x^2+2))P_{2n+2}(x)
\end{aligned}$$ □

ทฤษฎีบท 3.23. *สำหรับพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน $QGQ_n(x)$ สังยุคของพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน $QGQ_n^*(x)$ และสังยุคเชิงซ้อนของพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน $\overline{QGQ_n(x)}$ จะได้เอกลักษณ์ดังต่อไปนี้เป็นจริง*

$$QGQ_n(x) + QGQ_n^*(x) = 2e_0GQ_n(x)$$

$$QGQ_n(x) + \overline{QGQ_n(x)} = 2\sum_{s=0}^{3} e_sQ_{n+s}(x)$$

$$QGQ_n^2(x) + QGQ_n(x)QGQ_n^*(x) = 2e_0GQ_n(x)QGQ_n(x)$$

พิสูจน์. ในทำนองเดียวกันกับ ทฤษฎีบท 3.10 □

ทฤษฎีบท 3.24. *สำหรับพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน $QGQ_n(x)$ สังยุคของพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน $QGQ_n^*(x)$ และสังยุคเชิงซ้อนของพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน $\overline{QGQ_n(x)}$ จะได้เอกลักษณ์ดังต่อไปนี้เป็นจริง*

$$QGQ_n(x) - QGQ_n^*(x) = 2\sum_{k=1}^{3} e_kGQ_{n+s}(x)$$

$$QGQ_n(x) - \overline{QGQ_n(x)} = 2i\sum_{k=-1}^{2} e_kQ_{n+k}(x)$$

$$QGQ_n^2(x) - QGQ_n(x)QGQ_n^*(x) = 2QGQ_n(x)\sum_{k=1}^{3} e_kGQ_{n+s}(x)$$

พิสูจน์. ในทำนองเดียวกันกับ ทฤษฎีบท 3.11 □

4 สรุปผลและอภิปรายผล

งานวิจัยนี้ศึกษาเรื่องพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน โดยได้ศึกษาความสัมพันธ์เวียนเกิด สูตรบิเนต ผลบวก ผลต่าง ผลคูณ เอกลักษณ์แคสสินี อีกทั้งยังได้ศึกษา ผลบวก ผล

ต่าง และผลคูณ ระหว่างพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน กับสังยุค และระหว่างพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน กับสังยุค ซึ่งผลการศึกษาปรากฏว่า สามารถเขียนความสัมพันธ์เวียนเกิดของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน ได้ดังนี้ $QGP_n(x) + 2xQGP_{n+1}(x) = QGP_{n+2}(x)$ และ $QGQ_n(x) + 2xQGQ_{n+1}(x) = QGQ_{n+2}(x)$ ตามลำดับ สามารถเขียนสูตรบิเนตของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน ได้ในรูป $QGP_n(x) = \dfrac{\alpha^n(x)A - \beta^n(x)B}{\alpha(x) - \beta(x)}$ และ $QGQ_n(x) = \alpha^n(x)A + \beta^n(x)B$ เมื่อ $\alpha(x) = x + \sqrt{1+x^2}$, $\beta(x) = x - \sqrt{1+x^2}$ ตามลำดับ ได้เอกลักษณ์แคสสินี ของพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน คือ $QGP_{n+1}(x)QGP_{n-1}(x) - QGP_n^2(x) = (-1)^n AB$ และ $QGQ_{n+1}(x)QGQ_{n-1}(x) - QGQ_n^2(x) = (-1)^{n+1}(\alpha(x) - \beta(x))^2 AB$ ตามลำดับ นอกจากนี้เราสามารถศึกษาทฤษฎีบทอื่น ๆ เกี่ยวกับพหุนามเกาส์เซียนเพลล์ควอเทอร์เนียน และพหุนามเกาส์เซียนเพลล์-ลูคัสควอเทอร์เนียน ได้ในงานวิจัยเล่มนี้

เอกสารอ้างอิง

[1] A. F. Horadam, Generalized Fibonacci Sequence, American Math. 68 (1) (1961) 455-459.

[2] A. F. Horadam, Complex Fibonacci Numbers and Fibonacci Quaternions, American Math. 70 (1) (1963) 289-291.

[3] A. F. Horadam, Bro. J. M. Mahon, M. Abbas, PELL AND PELL-LUCAS POLYNOMIALS, Pearson Education, Australia, 1985.

[4] S. Halici and S. Oz, On Gaussian Pell Polynomials and Their Some Properties, Palestine Journal of Mathematics. 7 (1) (2018) 251-256.

[5] T. Yagmur, Gaussian Pell-Lucas Polynomials, Communications in Mathematics and Applications. 10 (4) (2019) 1609-1617.

[6] H. Arslana, Gaussian Pell and Gaussian Pell-Lucas Quaternions, Filomat. 35 (5) (2021) 673-679.

สมบัติบางประการของพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์ลูคัส

อรประภา เถี้ยมแก้ว[1,*], สุตาภัทร ชุมสมบูรณ์[1], ศรัณยา เฮงสวัสดิ์[1]
และ ภัทราวรรณ สิงควานนท์[1,†]

[1]สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสงขลา 90000

บทคัดย่อ

งานวิจัยนี้ศึกษาเกี่ยวกับพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส ซึ่งเป็นการต่อยอดแนวคิดจากงานวิจัยของ E. Ozkan และ T. Alp (2022) โดยได้ศึกษาผลคูณ ผลบวก ผลต่าง และสังยุคของพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส รวมทั้งศึกษา ผลบวกย่อย ผลบวกย่อยพจน์คู่ ผลบวกย่อยพจน์คี่ และเมทริกซ์ของพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส

Keywords: พหุนามเพลล์ พหุนามเพลล์-ลูคัส พหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส

2020 MSC: 11B37, 11B83

1 บทนำ

A. Dasdemir [2] ได้ศึกษาจำนวนเพลล์ และจำนวนเพลล์-ลูคัส นิยามโดยความสัมพันธ์เวียนเกิด $P_n = 2P_{n-1} + P_{n-2}, n \geq 2$ และ $Q_n = 2Q_{n-1} + Q_{n-2}, n \geq 2$ โดยมีเงื่อนไขค่าเริ่มต้น $P_0 = 0, P_1 = 1$ และ $Q_0 = 2, Q_1 = 2$ ตามลำดับ

มีงานวิจัยจำนวนมากได้ศึกษาสมบัติของพหุนามเพลล์ และพหุนามเพลล์-ลูคัส รวมทั้งเอกลักษณ์ที่สำคัญของจำนวนดังกล่าว A. F. Horadam และ Bro. J. M. Mahon [3] ได้ศึกษาพหุนามเพลล์ และพหุนามเพลล์-ลูคัส โดยนิยามความสัมพันธ์เวียนเกิด $P_n(x) = 2P_{n-1}(x) + P_{n-2}(x), n \geq 2$ โดยมีเงื่อนไขค่าเริ่มต้น $P_0(x) = 0, P_1(x) = 1$ และ $Q_n(x) = 2Q_{n-1}(x) + Q_{n-2}(x), n \geq 2$ โดยมีเงื่อนไขค่าเริ่มต้น $Q_0 = 2, Q_1 = 2x$ ต่อมา S. Halici และ S.Oz [4] ได้นิยามพหุนามเกาส์เซียนเพลล์ โดยนิยามความสัมพันธ์เวียนเกิด $GP_{n+1}(x) = 2xGP_n(x) + GP_{n-1}(x)$ โดยมีเงื่อนไขค่าเริ่มต้น $GP_0(x) = i, GP_1(x) = 1$ และ T. Yagmur [5] ได้นิยามพหุนามเกาส์เซียนเพลล์-ลูคัส $GQ_{n+1}(x) = 2xGQ_n(x) + GQ_{n-1}(x)$ โดยมีเงื่อนไขค่าเริ่มต้น $GQ_0(x) = 2 - 2xi, GQ_1(x) = 2x + 2i$ ต่อมา H. Gokas [6] ได้ให้นิยามของจำนวนไบเกาส์เซียนเพลล์ และจำนวนไบเกาส์เซียนเพลล์-ลูคัส เมื่อ

*ผู้นำเสนอ (644254050@parichat.skru.ac.th)
†ผู้ประพันธ์บรรณกิจ (pattarawan.pe@skru.ac.th)

P_n และ Q_n เป็นพจน์ที่ n ของจำนวนเพลล์ และจำนวนเพลล์-ลูคัส จะได้ความสัมพันธ์เวียนเกิด คือ $BGP_n = 2BGP_{n-1} + BGP_{n-2}$ และ $BGQ_n = 2BGQ_{n-1} + BGQ_{n-2}$

การทำวิจัยในครั้งนี้ผู้จัดทำสนใจที่จะศึกษาพหุนามเพลล์ และพหุนามเพลล์-ลูคัส และจำนวนไบเกาส์เซียนเพลล์ และจำนวนไบเกาส์เซียนเพลล์-ลูคัส ได้ศึกษาเพิ่มเติมให้อยู่ในรูปพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส และศึกษาผลคูณ ผลต่าง ผลบวก ผลบวกย่อย ผลบวกย่อยพจน์คู่ ผลบวกย่อยพจน์คี่ เมทริกซ์ รวมถึงสังยุคที่เกี่ยวข้องกับพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส

2 ความรู้พื้นฐาน

A. Dasdemir [2] ได้ศึกษาลำดับของจำนวนเพลล์ ดังนี้

$$0, 1, 2, 5, 12, 29, 70, 169, 408, 985, 2378, ...$$

ความสัมพันธ์เวียนเกิดของจำนวนเพลล์ กำหนดโดย

$$P_n = 2P_{n-1} + P_{n-2}, n \geq 2 \text{ โดยมีเงื่อนไขค่าเริ่มต้น } P_0 = 0, P_1 = 1 \tag{2.1}$$

สูตรบิเนตของจำนวนเพลล์ คือ

$$P_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} \text{ โดยที่ } n \geq 2 \tag{2.2}$$

เมื่อ $\alpha = 1 + \sqrt{2}$, $\beta = 1 - \sqrt{2}$
นอกจากนี้ยังได้ศึกษาลำดับของจำนวนเพลล์-ลูคัส

$$2, 2, 6, 14, 34, 82, 198, 478, 1154, 2786, ...$$

ความสัมพันธ์เวียนเกิดของจำนวนเพลล์-ลูคัส กำหนดโดย

$$Q_n = 2Q_{n-1} + Q_{n-2}, n \geq 2 \text{ โดยมีเงื่อนไขค่าเริ่มต้น } Q_0 = 2, Q_1 = 2 \tag{2.3}$$

สูตรบิเนตของจำนวนเพลล์-ลูคัส คือ

$$Q_n = \alpha^n + \beta^n \text{ โดยที่ } n \geq 2 \tag{2.4}$$

เมื่อ $\alpha = 1 + \sqrt{2}$, $\beta = 1 - \sqrt{2}$

H. Gokbas [6] ได้ให้นิยามของจำนวนไบเกาส์เซียนเพลล์ BGP_n ดังนี้

$$BGP_n = P_n + iP_{n-1} + jP_{n-2} + ijP_{n-3} \tag{2.5}$$

เมื่อ i, j และ ij เป็นส่วนจินตภาพ เป็นไปตามเงื่อนไข $i^2 = -1, j^2 = -1, ij = ji$
เมื่อ P_n เป็นพจน์ที่ n ของจำนวนเพลล์ และจะได้ความสัมพันธ์เวียนเกิด คือ

$$BGP_n = 2BGP_{n-1} + BGP_{n-2} \tag{2.6}$$

สูตรบิเนตของจำนวนไบเกาส์เซียนเพลล์ คือ

$$BGP_n = \frac{\widehat{\alpha}\alpha^{n-3} - \widehat{\beta}\beta^{n-3}}{\alpha - \beta} \tag{2.7}$$

เมื่อ $\hat{\alpha} = \alpha^3 + i\alpha^2 + j\alpha + ij, \alpha = 1 + \sqrt{2}$, $\hat{\beta} = \beta^3 + i\beta^2 + j\beta + ij, \beta = 1 - \sqrt{2}$

ให้ BGP_n และ BGP_m เป็นจำนวนไบเกาส์เซียนเพลล์ใด ๆ

การบวก และการลบ นิยามโดย

$$\begin{aligned} BGP_n \pm BGP_m &= (P_n + iP_{n-1} + jP_{n-2} + ijP_{n-3}) \\ &\quad \pm (P_m + iP_{m-1} + jP_{m-2} + ijP_{m-3}) \\ &= (P_n \pm P_m) + i(P_{n-1} \pm P_{m-1}) + j(P_{n-2} \pm P_{m-2}) \\ &\quad + ij(P_{n-3} \pm P_{m-3}) \end{aligned}$$

การคูณ นิยามโดย

$$\begin{aligned} BGP_n \times BGP_m &= (P_n + iP_{n-1} + jP_{n-2} + ijP_{n-3}) \\ &\quad \times (P_m + iP_{m-1} + jP_{m-2} + ijP_{m-3}) \\ &= (P_nP_m - P_{n-1}P_{m-1} - P_{n-2}P_{m-2} + P_{n-3}P_{m-3}) \\ &\quad + i(P_nP_{m-1} + P_{n-1}P_m - P_{n-2}P_{m-3} - P_{n-3}P_{m-2}) \\ &\quad + j(P_nP_{m-2} + P_{n-2}P_m - P_{n-1}P_{m-3} - P_{n-3}P_{m-1} \\ &\quad + ij(P_nP_{m-3} + P_{n-3}P_m + P_{n-1}P_{m-2} + P_{n-2}P_{m-1}) \\ &= BGP_m \times BGP_n \end{aligned}$$

สังยุคจำนวนไบเกาส์เซียนเพลล์ กำหนด 3 แบบต่อไปนี้

$$(BGP_n)_i^* = P_n - iP_{n-1} + jP_{n-2} - ijP_{n-3} \tag{2.8}$$

$$(BGP_n)_j^* = P_n + iP_{n-1} - jP_{n-2} - ijP_{n-3} \tag{2.9}$$

$$(BGP_n)_{ij}^* = P_n - iP_{n-1} - jP_{n-2} + ijP_{n-3} \tag{2.10}$$

และจำนวนไบเกาส์เซียนเพลล์-ลูคัส นิยามดังนี้

$$BGQ_n = Q_n + iQ_{n-1} + jQ_{n-2} + ijQ_{n-3} \tag{2.11}$$

เมื่อ Q_n เป็นพจน์ที่ n ของจำนวนเพลล์-ลูคัส และจะได้ความสัมพันธ์เวียนเกิด คือ

$$BGQ_n = 2BGQ_{n-1} + BGQ_{n-2} \tag{2.12}$$

สูตรบิเนตของจำนวนไบเกาส์เซียนเพลล์-ลูคัส คือ

$$BGQ_n = \widehat{\alpha}\alpha^{n-3} + \widehat{\beta}\beta^{n-3} \tag{2.13}$$

เมื่อ $\hat{\alpha} = \alpha^3 + i\alpha^2 + j\alpha + ij, \alpha = 1 + \sqrt{2}$, $\hat{\beta} = \beta^3 + i\beta^2 + j\beta + ij, \beta = 1 - \sqrt{2}$

ให้ BGQ_n และ BGQ_m เป็นจำนวนไบเกาส์เซียนเพลล์-ลูคัสใด ๆ

การบวกและการลบ นิยามโดย

$$\begin{aligned} BGQ_n \pm BGQ_m &= (Q_n + iQ_{n-1} + jQ_{n-2} + ijQ_{n-3}) \\ &\quad \pm(Q_m + iQ_{m-1} + jQ_{m-2} + ijQ_{m-3}) \\ &= (Q_n \pm Q_m) + i(Q_{n-1} \pm Q_{m-1}) + j(Q_{n-2} \pm Q_{m-2}) \\ &\quad +ij(Q_{n-3} \pm Q_{m-3}) \end{aligned}$$

การคูณ นิยามโดย

$$\begin{aligned} BGQ_n \times BGQ_m &= (Q_n + iQ_{n-1} + jQ_{n-2} + ijQ_{n-3}) \\ &\quad \times(Q_m + iQ_{m-1} + jQ_{m-2} + ijQ_{m-3}) \\ &= (Q_nQ_m - Q_{n-1}Q_{m-1} - Q_{n-2}Q_{m-2} + Q_{n-3}Q_{m-3}) \\ &\quad +i(Q_nQ_{m-1} + Q_{n-1}Q_m - Q_{n-2}Q_{m-3} - Q_{n-3}Q_{m-2}) \\ &\quad +j(Q_nQ_{m-2} + Q_{n-2}Q_m - Q_{n-1}Q_{m-3} - Q_{n-3}Q_{m-1} \\ &\quad +ij(Q_nQ_{m-3} + Q_{n-3}Q_m + Q_{n-1}Q_{m-2} + Q_{n-2}Q_{m-1}) \\ &= BGQ_m \times BGQ_n \end{aligned}$$

สังยุคจำนวนไบเกาส์เซียนเพลล์-ลูคัส มี 3 รูปแบบดังต่อไปนี้

$$(BGQ_n)_i^* = Q_n - iP_{n-1} + jQ_{n-2} - ijQ_{n-3} \tag{2.14}$$

$$(BGQ_n)_j^* = Q_n + iP_{n-1} - jQ_{n-2} - ijQ_{n-3} \tag{2.15}$$

$$(BGQ_n)_{ij}^* = Q_n - iQ_{n-1} - jQ_{n-2} + ijQ_{n-3} \tag{2.16}$$

A. F. Horadam และ Bro. J. M. Mohon [3] ได้ศึกษาพหุนามเพลล์โดยให้นิยามความสัมพันธ์เวียนเกิดของพหุนามเพลล์ กำหนดโดย

$$P_n(x) = 2xP_{n-1}(x) + P_{n-2}(x)\,, n \geq 2 \tag{2.17}$$

โดยมีเงื่อนไขค่าเริ่มต้น $P_0(x) = 0$, $P_1(x) = 1$

สูตรบิเนตของพหุนามเพลล์ คือ

$$P_n(x) = \frac{\alpha^n(x) - \beta^n(x)}{\alpha(x) - \beta(x)}\,, n \geq 2 \tag{2.18}$$

เมื่อ $\alpha = x + \sqrt{x^2+1}$ และ $\beta = x - \sqrt{x^2+1}$
และความสัมพันธ์เวียนเกิดของพหุนามเพลล์-ลูคัส กำหนดโดย

$$Q_n(x) = 2xQ_{n-1}(x) + Q_{n-2}(x)\,,\ n \geq 2 \tag{2.19}$$

โดยมีเงื่อนไขค่าเริ่มต้น $Q_0(x) = 2,\ Q_1(x) = 2x$
สูตรบิเนตของพหุนามเพลล์-ลูคัส คือ

$$Q_n(x) = \alpha^n(x) + \beta^n(x)\,,\ n \geq 2 \tag{2.20}$$

เมื่อ $\alpha = x + \sqrt{x^2+1}$ และ $\beta = x - \sqrt{x^2+1}$
จะได้เอกลักษณ์ของพหุนามเพลล์ และพหุนามเพลล์-ลูคัส ดังนี้

$$P_{-n}(x) = (-1)^{n+1}P_n(x) \tag{2.21}$$

$$Q_{-n}(x) = (-1)^n Q_n(x) \tag{2.22}$$

$$P_{2n+1}(x) = P_{n+1}^2(x) + P_n^2(x) \tag{2.23}$$

$$Q_n(x) = P_{n+1}(x) - P_{n-1}(x) \tag{2.24}$$

$$Q_{n+1}(x) + Q_{n-1}(x) = 4(x^2+1)P_n(x) \tag{2.25}$$

$$P_{n+2}(x) = 2xP_{n+1}(x) + P_n(x) \tag{2.26}$$

$$Q_{n+2}(x) = 2xQ_{n+1}(x) + Q_n(x) \tag{2.27}$$

$$2xP_n(x) + 2P_{n-1}(x) = Q_n(x) \tag{2.28}$$

$$P_n(x)Q_n(x) = P_{2n}(x) \tag{2.29}$$

$$\sum_{k=1}^{n} P_k(x) = \frac{P_{n+1}(x) + P_n(x) - 1}{2x} \tag{2.30}$$

$$\sum_{k=1}^{n} P_{2k-1}(x) = \frac{P_{2n}(x)}{2x} \tag{2.31}$$

$$\sum_{k=1}^{n} P_{2k}(x) = \frac{P_{2n+1}(x) - 1}{2x} \tag{2.32}$$

$$\sum_{k=1}^{n} Q_k(x) = \frac{Q_{n+1}(x) + Q_n(x) - 2x - 2}{2x} \tag{2.33}$$

$$\sum_{k=1}^{n} Q_{2k}(x) = \frac{Q_{2n+1}(x) - 2}{2x} \tag{2.34}$$

$$\sum_{k=1}^{n} Q_{2k-1}(x) = \frac{Q_{2n}(x) - 2}{2x} \tag{2.35}$$

$$Q_{n+r}^2(x) - Q_{n-r}^2(x) = 4(x^2+1)P_{2n}(x)P_{2r}(x) \tag{2.36}$$

$$P_n^2(x) - P_{n+r}(x)P_{n-r}(x) = (-1)^{n-r}P_r^2(x) \tag{2.37}$$

$$Q_n^2(x) - Q_{n+r}(x)Q_{n-r}(x) = (-1)^{n-r+1}4(x^2+1)P_r^2(x) \tag{2.38}$$

$$P_{n+h}(x)P_{n+k}(x) - P_n(x)P_{n+h+k}(x) = (-1)^n P_h(x)P_k(x) \tag{2.39}$$

$$P_{n+r}(x) + P_{n-r}(x) = \begin{cases} P_n(x)Q_r(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ Q_n(x)P_r(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases} \tag{2.40}$$

$$Q_{n+r}(x) + Q_{n-r}(x) = \begin{cases} Q_n(x)Q_r(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ 4(x^2+1)P_n(x)P_r(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases} \tag{2.41}$$

$$P_{n+r}(x) - P_{n-r}(x) = \begin{cases} Q_n(x)P_r(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ P_n(x)P_r(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases} \tag{2.42}$$

$$Q_{n+r}(x) - Q_{n-r}(x) = \begin{cases} 4(x^2+1)P_n(x)P_r(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ Q_n(x)Q_r(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases} \tag{2.43}$$

E. Ozkan และ T. Alp [1] ได้ให้นิยามพหุนามไบเกาส์เซียนเพลล์ $BGP_n(x)$ นิยามโดย

$$BGP_n(x) = P_n(x) + iP_{n-1}(x) + jP_{n-2}(x) + ijP_{n-3}(x) \tag{2.44}$$

เมื่อ $P_n(x)$ เป็นพจน์ที่ n ของพหุนามเพลล์ และจะได้ความสัมพันธ์เวียนเกิดของพหุนามไบเกาส์เซียนเพลล์ ดังนี้

$$BGP_n(x) = 2xBGP_{n-1}(x) + BGP_{n-2}(x) \tag{2.45}$$

สูตรบิเนตของพหุนามไบเกาส์เซียนเพลล์ คือ

$$BGP_n(x) = \frac{\widehat{\alpha}\alpha^{n-3}(x) - \widehat{\beta}\beta^{n-3}(x)}{\alpha(x) - \beta(x)} \tag{2.46}$$

และพหุนามไบเกาส์เซียนเพลล์-ลูคัส $BGQ_n(x)$ นิยามโดย

$$BGQ_n(x) = Q_n(x) + iQ_{n-1}(x) + jQ_{n-2}(x) + ijQ_{n-3}(x) \tag{2.47}$$

เมื่อ $Q_n(x)$ เป็นพจน์ทั่วไปของพหุนามเพลล์-ลูคัส และจะได้ความสัมพันธ์เวียนเกิดของพหุนามไบเกาส์เซียนเพลล์-ลูคัส

$$BGQ_n(x) = 2xBGQ_{n-1}(x) + BGQ_{n-2}(x) \tag{2.48}$$

สูตรบิเนตของพหุนามไบเกาส์เซียนเพลล์-ลูคัส คือ

$$BGQ_n(x) = \widehat{\alpha}\alpha^{n-3}(x) + \widehat{\beta}\beta^{n-3}(x) \tag{2.49}$$

3 ผลการศึกษา

ในบทนี้จะกล่าวถึงสมบัติบางประการของพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส รวมถึงผลบวก ผลต่าง ผลคูณ ผลบวกย่อย ผลบวกย่อยพจน์คู่ ผลบวกย่อยพจน์คี่ เมทริกซ์ และสังยุคต่าง ๆ ของพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส

3.1 สมบัติบางประการของพหุนามไบเกาส์เซียนเพลล์

จากพหุนามไบเกาส์เซียนเพลล์ ซึ่งกำหนดโดย

$$BGP_n(x) = P_n(x) + iP_{n-1}(x) + jP_{n-2}(x) + ijP_{n-3}(x)$$

สังยุคของพหุนามไบเกาส์เซียนเพลล์ มี 3 แบบดังต่อไปนี้

$$(BGP_n(x))_i^* = P_n(x) - iP_{n-1}(x) + jP_{n-2}(x) - ijP_{n-3}(x) \tag{3.1}$$

$$(BGP_n(x))_j^* = P_n(x) + iP_{n-1}(x) - jP_{n-2}(x) - ijP_{n-3}(x) \tag{3.2}$$

$$(BGP_n(x))_{ij}^* = P_n(x) - iP_{n-1}(x) - jP_{n-2}(x) + ijP_{n-3}(x) \tag{3.3}$$

ทฤษฎีบท 3.1. *ให้ $BGP_n(x)$ และ $BGP_m(x)$ เป็นพหุนามไบเกาส์เซียนเพลล์ใด ๆ จะได้สังยุคของผลคูณของสองจำนวน ดังนี้*

1) $(BGP_n(x) \times BGP_m(x))_i^* = (BGP_n(x))i^* \times (BGP_m(x))_i^*$
2) $(BGP_n(x) \times BGP_m(x))_j^* = (BGP_n(x))_j^* \times (BGP_m(x))_j^*$
3) $(BGP_n(x) \times BGP_m(x))_{ij}^* = (BGP_n(x))_{ij}^* \times (BGP_m(x))_{ij}^*$

พิสูจน์. 1) พิจารณา

$$\begin{aligned}
&(BGP_n(x) \times BGP_m(x))_i^* \\
&= [(P_n(x) + iP_{n-1}(x) + jP_{n-2}(x) + ijP_{n-3}(x)) \times (P_m(x) \\
&\quad +iP_{m-1}(x) + jP_{m-2}(x) + ijP_{m-3}(x))]_i^* \\
&= [P_n(x)P_m(x) + iP_n(x)P_{m-1}(x) + jP_n(x)P_{m-2}(x) + ijP_n(x)P_{m-3}(x) \\
&\quad +iP_{n-1}(x)P_m(x) + i^2P_{n-1}(x)P_{m-1}(x) + ijP_{n-1}(x)P_{m-2}(x) \\
&\quad +i^2jP_{n-1}(x)P_{m-3}(x) + jP_{n-2}(x)P_m(x) + ijP_{n-2}(x)P_{m-1}(x) \\
&\quad +j^2P_{n-2}(x)P_{m-2}(x) + ij^2P_{n-2}(x)P_{m-3}(x) + ijP_{n-3}(x)P_m(x) \\
&\quad +i^2jP_{n-3}(x)P_{m-1}(x) + ij^2P_{n-3}(x)P_{m-2}(x) + i^2j^2P_{n-3}(x)P_{m-3}(x)]_i^* \\
&= [P_n(x)P_m(x) + iP_n(x)P_{m-1}(x) + jP_n(x)P_{m-2}(x) + ijP_n(x)P_{m-3}(x) \\
&\quad +iP_{n-1}(x)P_m(x) - P_{n-1}(x)P_{m-1}(x) + ijP_{n-1}(x)P_{m-2}(x) \\
&\quad -jP_{n-1}(x)P_{m-3}(x) + jP_{n-2}(x)P_m(x) + ijP_{n-2}(x)P_{m-1}(x) \\
&\quad -P_{n-2}(x)P_{m-2}(x) - iP_{n-2}(x)P_{m-3}(x) + ijP_{n-3}(x)P_m(x) \\
&\quad -jP_{n-3}(x)P_{m-1}(x) - iP_{n-3}(x)P_{m-2}(x) + P_{n-3}(x)P_{m-3}(x)]_i^* \\
&= [(P_n(x)P_m(x) - P_{n-1}(x)P_{m-1}(x) - P_{n-2}(x)P_{m-2}(x) + P_{n-3}(x)P_{m-3}(x)) \\
&\quad +i(P_n(x)P_{m-1}(x) + P_{n-1}(x)P_m(x) - P_{n-2}(x)P_{m-3}(x) - P_{n-3}(x)P_{m-2}(x)) \\
&\quad +j(P_n(x)P_{m-2}(x) - P_{n-1}(x)P_{m-3}(x) + P_{n-2}(x)P_m(x) - P_{n-3}(x)P_{m-1}(x)) \\
&\quad +ij(P_n(x)P_{m-3}(x) + P_{n-1}(x)P_{m-2}(x) + P_{n-2}(x)P_{m-1}(x) + P_{n-3}(x)P_m(x))]_i^* \\
&= (P_n(x)P_m(x) - P_{n-1}(x)P_{m-1}(x) - P_{n-2}(x)P_{m-2}(x) + P_{n-3}(x)P_{m-3}(x)) \\
&\quad -i(P_n(x)P_{m-1}(x) + P_{n-1}(x)P_m(x) - P_{n-2}(x)P_{m-3}(x) - P_{n-3}(x)P_{m-2}(x)) \\
&\quad +j(P_n(x)P_{m-2}(x) - P_{n-1}(x)P_{m-3}(x) + P_{n-2}(x)P_m(x) - P_{n-3}(x)P_{m-1}(x)) \\
&\quad -ij(P_n(x)P_{m-3}(x) + P_{n-1}(x)P_{m-2}(x) + P_{n-2}(x)P_{m-1}(x) + P_{n-3}(x)P_m(x))
\end{aligned}$$

$$
\begin{aligned}
&= (P_n(x) - iP_{n-1}(x) + jP_{n-2}(x) - ijP_{n-3}(x)) \times (P_m(x) - iP_{m-1}(x) \\
&\quad +jP_{m-2}(x) - ijP_{m-3}(x)) \\
&= (BGP_n(x))_i^* \times (BGP_m(x))_i^*
\end{aligned}
$$

การพิสูจน์ 2) และ 3) สามารถพิสูจน์ได้ในทำนองเดียวกัน □

ทฤษฎีบท 3.2. *ให้ $BGP_n(x)$ เป็นพหุนามไบเกาส์เซียนเพลล์ใด ๆ และ $(BGP_n(x))_i^*$, $(BGP_n(x))_j^*$ และ $(BGP_n(x))_{ij}^*$ เป็นสังยุค 3 แบบของจำนวนไบเกาส์เซียนเพลล์ จะได้*

1) $BGP_n(x) - (BGP_n(x))_i^* = 2i(P_{n-1}(x) + jP_{n-3}(x))$
2) $BGP_n(x) + (BGP_n(x))_i^* = 2(P_n(x) + jP_{n-3}(x))$
3) $BGP_n(x) - (BGP_n(x))_j^* = 2j(P_{n-2}(x) + iP_{n-3}(x))$
4) $BGP_n(x) + (BGP_n(x))_j^* = 2(P_n(x) + iP_{n-1}(x))$
5) $BGP_n(x) - (BGP_n(x))_{ij}^* = 2(iP_{n-1}(x) + jP_{n-2}(x))$
6) $BGP_n(x) + (BGP_n(x))_{ij}^* = 2(P_n(x) + ijP_{n-3}(x))$

พิสูจน์. 1) พิจารณา

$$
\begin{aligned}
&BGP_n(x) - (BGP_n(x))_i^* \\
&= (P_n(x) + iP_{n-1}(x) + jP_{n-2}(x) + ijP_{n-3}(x)) - (P_n(x) - iP_{n-1}(x) \\
&\quad +jP_{n-2}(x) - ijP_{n-3}(x)) \\
&= P_n(x) + iP_{n-1}(x) + jP_{n-2}(x) + ijP_{n-3}(x) - P_n(x) + iP_{n-1}(x) \\
&\quad -jP_{n-2}(x) + ijP_{n-3}(x) \\
&= 2iP_{n-1}(x) + 2ijP_{n-3}(x)) \\
&= 2i(P_{n-1}(x) + jP_{n-3}(x))
\end{aligned}
$$

2) พิจารณา

$$
\begin{aligned}
&BGP_n(x) + (BGP_n(x))_i^* \\
&= (P_n(x) + iP_{n-1}(x) + jP_{n-2}(x) + ijP_{n-3}(x)) + (P_n(x) - iP_{n-1}(x) \\
&\quad +jP_{n-2}(x) - ijP_{n-3}(x)) \\
&= 2P_n(x) + 2jP_{n-2}(x) \\
&= 2(P_n(x) + jP_{n-2}(x))
\end{aligned}
$$

การพิสูจน์ 3), 4), 5) และ 6) สามารถพิสูจน์ได้ในทำนองเดียวกัน □

ทฤษฎีบท 3.3. *ให้ $BGP_n(x)$ เป็นพหุนามไบเกาส์เซียนเพลล์ใด ๆ จะได้*

1) $\sum_{k=1}^{n} BGP_k(x) = \frac{1}{2x}(BGP_{n+1}(x) + BGP_n(x) - 4(x^2 - x + 1))$
2) $\sum_{k=1}^{n} BGP_{2k-1}(x) = \frac{1}{2x}(BGP_{2n}(x) - 2(2x^2 - x + 1))$
3) $\sum_{k=1}^{n} BGP_{2k}(x) = \frac{1}{2x}(BGP_{2n+1}(x) + 2(x - 1))$

พิสูจน์. 1) จากเอกลักษณ์ (2.21) และ (2.31)

$$\begin{aligned}
\sum_{k=1}^{n} BGP_k(x) &= \sum_{k=1}^{n} (P_k(x) + iP_{k-1}(x) + jP_{k-2}(x) + ijP_{k-3}(x)) \\
&= \sum_{k=1}^{n} P_k(x) + i\sum_{k=1}^{n} P_{k-1}(x) + j\sum_{k=1}^{n} P_{k-2}(x) + ij\sum_{k=1}^{n} P_{k-3}(x) \\
&= \sum_{k=1}^{n} P_k(x) + i\sum_{k=0}^{n-1} P_k(x) + j\sum_{k=-1}^{n-2} P_k(x) + ij\sum_{k=-2}^{n-3} P_k(x) \\
&= \sum_{k=1}^{n} P_k(x) + i\left(P_0(x) + \sum_{k=1}^{n-1} P_k(x)\right) + j\left(P_{-1}(x) + P_0(x) + \sum_{k=1}^{n-2} P_k(x)\right) \\
&\quad +ij\left(P_{-2}(x) + P_{-1}(x) + P_0(x) + \sum_{k=1}^{n-3} P_k(x)\right) \\
&= \sum_{k=1}^{n} P_k(x) + i\left(P_0(x) + \sum_{k=1}^{n-1} P_k(x)\right) \\
&\quad +j\left((-1)^{1+1}P_1(x) + P_0(x) + \sum_{k=1}^{n-2} P_k(x)\right) \\
&\quad +ij\left((-1)^{2+1}P_2(x) + (-1)^{1+1}P_1(x) + P_0(x) + \sum_{k=1}^{n-3} P_k(x)\right) \\
&= \left(\frac{P_{n+1}(x) + P_n(x) - 1}{2x}\right) + i\left(+0 + \frac{P_n(x) + P_{n-1}(x) - 1}{2x}\right) \\
&\quad +j\left(1 + 0 + \frac{P_{n-1}(x) + P_{n-2}(x) - 1}{2x}\right) \\
&\quad +ij\left(-2x + 1 + 0 + \frac{P_{n-2}(x) + P_{n-3}(x) - 1}{2x}\right) \\
&= \left(\frac{P_{n+1}(x) + P_n(x) - 1}{2x}\right) + i\left(\frac{P_n(x) + P_{n-1}(x) - 1}{2x}\right) \\
&\quad +j\left(\frac{P_{n-1}(x) + P_{n-2}(x) + 2x - 1}{2x}\right) \\
&\quad +ij\left(\frac{P_{n-2}(x) + P_{n-3}(x) - 4x^2 + 2x - 1}{2x}\right) \\
&= \left(\frac{P_{n+1}(x) + iP_n(x) + jP_{n-1}(x) + ijP_{n-2}(x)}{2x}\right) \\
&\quad + \left(\frac{P_n(x) + iP_{n-1}(x) + jP_{n-2}(x) + ijP_{n-3}(x)}{2x}\right) \\
&\quad + \left(\frac{-4x^2 + 4x - 4}{2x}\right) \\
&= \frac{1}{2x}(BGP_{n+1}(x) + BGP_n(x) - 4(x^2 - x + 1))
\end{aligned}$$

การพิสูจน์ 2) และ 3) สามารถพิสูจน์ได้ในทำนองเดียวกัน □

ทฤษฎีบท 3.4. *ให้ $BGP_n(x)$ เป็นพหุนามไบเกาส์เซียนเพลล์ใด ๆ จะได้*

$$\begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix}^n \begin{bmatrix} BGP_0(x) \\ BGP_1(x) \end{bmatrix} = \begin{bmatrix} BGP_n(x) \\ BGP_{n+1}(x) \end{bmatrix}$$

พิสูจน์. พิจารณา

ให้ $P(n)$ แทนข้อความ $\begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix}^n \begin{bmatrix} BGP_0(x) \\ BGP_1(x) \end{bmatrix} = \begin{bmatrix} BGP_n(x) \\ BGP_{n+1}(x) \end{bmatrix}$

ขั้นฐาน ให้ $n = 1$ พิจารณา

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix}^1 \begin{bmatrix} BGP_0(x) \\ BGP_1(x) \end{bmatrix} &= \begin{bmatrix} BGP_1(x) \\ BGP_0(x) + 2xBGP_1(x) \end{bmatrix} \\ &= \begin{bmatrix} BGP_1(x) \\ BGP_2(x) \end{bmatrix} \end{aligned}$$

ดังนั้น $P(1)$ เป็นจริง

ขั้นอุปนัย สมมติให้ $P(k)$ เป็นจริง นั่นคือ $\begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix}^k \begin{bmatrix} BGP_0(x) \\ BGP_1(x) \end{bmatrix} = \begin{bmatrix} BGP_k(x) \\ BGP_{k+1}(x) \end{bmatrix}$

จะแสดงว่า $P(k+1)$ เป็นจริง พิจารณา

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix}^{k+1} \begin{bmatrix} BGP_0(x) \\ BGP_1(x) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix}^k \begin{bmatrix} BGP_0(x) \\ BGP_1(x) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix} \begin{bmatrix} BGP_k(x) \\ BGP_{k+1}(x) \end{bmatrix} \\ &= \begin{bmatrix} BGP_{k+1}(x) \\ BGP_k(x) + 2xBGP_{k+1}(x) \end{bmatrix} \\ &= \begin{bmatrix} BGP_{k+1}(x) \\ BGP_{k+2}(x) \end{bmatrix} \end{aligned}$$

ดังนั้น $P(k+1)$ เป็นจริง

นั่นคือ $\begin{bmatrix} 0 & x \\ x & 2x \end{bmatrix}^n \begin{bmatrix} BGP_0(x) \\ BGP_1(x) \end{bmatrix} = \begin{bmatrix} BGP_n(x) \\ BGP_{n+1}(x) \end{bmatrix}$ □

ทฤษฎีบท 3.5. *ให้ n และ r เป็นจำนวนเต็มที่ไม่เป็นลบ จะได้*

$$1)\ BGP_{n+r}(x) + BGP_{n-r}(x) = \begin{cases} Q_r(x)BGP_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ P_r(x)BGQ_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases}$$

$$2)\ BGP_{n+r}(x) - BGP_{n-r}(x) = \begin{cases} P_r(x)BGQ_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ Q_r(x)BGP_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases}$$

พิสูจน์. แบ่งพิจารณาออกเป็น 2 กรณี ดังนี้
กรณีที่ 1 เมื่อ r เป็นจำนวนเต็มคู่
พิจารณา $BGP_{n+r}(x) + BGP_{n-r}(x)$

$$\begin{aligned}
&= (P_{n+r}(x) + iP_{n+r-1}(x) + jP_{n+r-2}(x) + ijP_{n+r-3}(x)) + (P_{n-r}(x) + iP_{n-r-1}(x) \\
&\quad +jP_{n-r-2}(x) + ijP_{n-r-3}(x)) \\
&= (P_{n+r}(x) + P_{n-r}(x)) + i(P_{(n-1)+r}(x) + P_{(n-1)-r}(x)) \\
&\quad + j(P_{(n-2)+r}(x) + P_{(n-2)-r}(x)) + ij(P_{(n-3)+r}(x) + P_{(n-3)-r}(x)) \\
&= (P_n(x)Q_r(x)) + i(P_{n-1}(x)Q_r(x)) + j(P_{n-2}(x)Q_r(x)) \\
&\quad + ij(P_{n-3}(x)Q_r(x)) \\
&= Q_r(x)(P_n(x) + iP_{n-1}(x) + jP_{n-2}(x) + ijP_{n-3}(x)) \\
&= Q_r(x)BGP_n(x)
\end{aligned}$$

กรณีที่ 2 เมื่อ r เป็นจำนวนเต็มคี่
พิจารณา $BGP_{n+r}(x) + BGP_{n-r}(x)$

$$\begin{aligned}
&= (P_{n+r}(x) + iP_{n+r-1}(x) + jP_{n+r-2}(x) + ijP_{n+r-3}(x)) \\
&\quad +(P_{n-r}(x) + iP_{n-r-1}(x) + jP_{n-r-2}(x) + ijP_{n-r-3}(x)) \\
&= (P_{n+r}(x) + P_{n-r}(x)) + i(P_{(n-1)+r}(x) + P_{(n-1)-r}(x)) \\
&\quad +j(P_{(n-2)+r}(x) + P_{(n-2)-r}(x)) + ij(P_{(n-3)+r}(x) + P_{(n-3)-r}(x)) \\
&= (Q_n(x)P_r(x)) + i(Q_{n-1}(x)P_r(x)) + j(Q_{n-2}(x)P_r(x)) \\
&\quad + ij(Q_{n-3}(x)P_r(x)) \\
&= P_r(x)(Q_n(x) + iQ_{n-1}(x) + jQ_{n-2}(x) + ijQ_{n-3}(x)) \\
&= P_r(x)BGQ_n(x)
\end{aligned}$$

การพิสูจน์ 2) สามารถพิสูจน์ได้ในทำนองเดียวกัน □

บทแทรก 3.6. *ให้ n เป็นจำนวนเต็มที่ไม่เป็นลบ จะได้*

$$\begin{aligned}
1)\ BGP_{n+1}(x) + BGP_{n-1}(x) &= BGQ_n(x) \\
2)\ BGP_{n+1}(x) - BGP_{n-1}(x) &= 2xBGP_n(x)
\end{aligned}$$

พิสูจน์. จากทฤษฎีบท 3.5 พิจารณาเมื่อ $r = 1$ จะได้ $BGP_{n+1}(x) + BGP_{n-1}(x) = P_1(x)BGQ_n(x)$
การพิสูจน์ 2) สามารถพิสูจน์ได้ในทำนองเดียวกัน □

3.2 สมบัติบางประการของพหุนามไบเกาส์เซียนเพลล์-ลูคัส

ในหัวข้อนี้เราศึกษาพหุนามไบเกาส์เซียนเพลล์-ลูคัสในลักษณะเดียวกันกับหัวข้อ 3.1 จากพหุนามไบเกาส์เซียนเพลล์-ลูคัส กำหนดโดย

$$BGQ_n(x) = Q_n(x) + iQ_{n-1}(x) + jQ_{n-2}(x) + ijQ_{n-3}(x)$$

สังยุคของพหุนามไบเกาส์เซียนเพลล์-ลูคัส มี 3 แบบดังต่อไปนี้

$$(BGQ_n(x))_i^* = Q_n(x) - iQ_{n-1}(x) + jQ_{n-2}(x) - ijQ_{n-3}(x) \tag{3.4}$$

$$(BGQ_n(x))_j^* = Q_n(x) + iQ_{n-1}(x) - jQ_{n-2}(x) - ijQ_{n-3}(x) \tag{3.5}$$

$$(BGQ_n(x))_{ij}^* = Q_n(x) - iQ_{n-1}(x) - jQ_{n-2}(x) + ijQ_{n-3}(x) \tag{3.6}$$

ทฤษฎีบท 3.7. *ให้ $BGQ_n(x)$ และ $BGQ_m(x)$ เป็นพหุนามไบเกาส์เซียนเพลล์-ลูคัสใด ๆ จะได้สังยุคของผลคูณของสองจำนวนใด ๆ ดังนี้*

1) $(BGQ_n(x) \times BGQ_m(x))_i^* = (BGQ_n(x))i^* \times (BGQ_m(x))_i^*$
2) $(BGQ_n(x) \times BGQ_m(x))_j^* = (BGQ_n(x))_j^* \times (BGQ_m(x))_j^*$
3) $(BGQ_n(x) \times BGQ_m(x))_{ij}^* = (BGQ_n(x))_{ij}^* \times (BGQ_m(x))_{ij}^*$

พิสูจน์. ในทำนองเดียวกันกับทฤษฎีบท 3.1 □

ทฤษฎีบท 3.8. *ให้ $BGQ_n(x)$ เป็นพหุนามไบเกาส์เซียนเพลล์ใด ๆ และ $(BGQ_n(x))_i^*$, $(BGQ_n(x))_j^*$ และ $(BGQ_n(x))_{ij}^*$ เป็นสังยุค 3 แบบของจำนวนไบเกาส์เซียนเพลล์ จะได้*

1) $BGQ_n(x) - (BGQ_n(x))_i^* = 2i(Q_{n-1}(x) + jQ_{n-3}(x))$
2) $BGQ_n(x) + (BGQ_n(x))_i^* = 2(Q_n(x) + jQ_{n-2}(x))$
3) $BGQ_n(x) - (BGQ_n(x))_j^* = 2j(Q_{n-2}(x) + iQ_{n-3}(x))$
4) $BGQ_n(x) + (BGQ_n(x))_j^* = 2(Q_n(x) + iQ_{n-1}(x))$
5) $BGQ_n(x) - (BGQ_n(x))_{ij}^* = 2(iQ_{n-1}(x) + jQ_{n-2}(x))$
6) $BGQ_n(x) + (BGQ_n(x))_{ij}^* = 2(Q_n(x) + ijQ_{n-3}(x))$

พิสูจน์. ในทำนองเดียวกันกับทฤษฎีบท 3.2 □

ทฤษฎีบท 3.9. *ให้ $BGQ_n(x)$ เป็นพหุนามไบเกาส์เซียนเพลล์-ลูคัสใด ๆ จะได้*

1) $\sum_{k=1}^{n} BGQ_k(x) = \frac{1}{2x}(BGQ_{n+1}(x) + BGQ_n(x) + 8(x-1)(x^4+1))$
2) $\sum_{k=1}^{n} BGQ_{2k-1}(x) = \frac{1}{2x}(BGQ_{2n}(x) + 4(2x-1)(x^2+1))$
3) $\sum_{k=1}^{n} BGQ_{2k}(x) = \frac{1}{2x}(BGQ_{2n+1}(x) - 4(x^2+1))$

พิสูจน์. ในทำนองเดียวกันกับทฤษฎีบท 3.3 □

ทฤษฎีบท 3.10. *ให้ $BGQ_n(x)$ เป็นพหุนามไบเกาส์เซียนเพลล์-ลูคัสและ $n \geq 1$ เป็นจำนวนเต็ม จะได้*

$$\begin{bmatrix} 0 & 1 \\ 1 & 2x \end{bmatrix}^n \begin{bmatrix} BGQ_0(x) \\ BGQ_1(x) \end{bmatrix} = \begin{bmatrix} BGQ_n(x) \\ BGQ_{n+1}(x) \end{bmatrix}$$

พิสูจน์. ในทำนองเดียวกันกับทฤษฎีบท 3.4 โดยใช้อุปนัยเชิงคณิตศาสตร์ □

ทฤษฎีบท 3.11. *ให้ n และ r เป็นจำนวนเต็มที่ไม่เป็นลบ จะได้*

$$1)\ BGQ_{n+r}(x)+BGQ_{n-r}(x)=\begin{cases}Q_r(x)BGQ_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่}\\ 4(x^2+1)P_r(x)BGP_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่}\end{cases}$$

$$2)\ BGQ_{n+r}(x)-BGQ_{n-r}(x)=\begin{cases}4(x^2+1)P_r(x)BGP_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่}\\ Q_r(x)BGQ_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่}\end{cases}$$

พิสูจน์. ในทำนองเดียวกันกับทฤษฎีบท 3.5 □

บทแทรก 3.12. *ให้ n เป็นจำนวนเต็มที่ไม่เป็นลบ จะได้*

$$\begin{aligned}1)\ BGQ_{n+1}(x)+BGQ_{n-1}(x) &= 4(x^2+1)BGP_n(x)\\ 2)\ BGQ_{n+1}(x)-BGQ_{n-1}(x) &= 2xBGQ_n(x)\end{aligned}$$

พิสูจน์. 1) จากทฤษฎีบท 3.11 พิจารณาเมื่อ $r=1$ จะได้
$BGQ_{n+1}(x)+BGQ_{n-1}(x)=4(x^2+1)P_1(x)BGP_n(x)$
การพิสูจน์ 2) สามารถพิสูจน์ได้ในทำนองเดียวกัน □

4 สรุปผลและอภิปรายผล

งานวิจัยนี้ศึกษาสมบัติบางประการของพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส โดยได้ศึกษาผลคูณ ผลบวก และผลต่างระหว่างพหุนามไบเกาส์เซียนเพลล์ กับสังยุคของพหุนามไบเกาส์เซียนเพลล์ และศึกษาผลคูณ ผลบวก และผลต่างระหว่างพหุนามไบเกาส์เซียนเพลล์-ลูคัส กับสังยุคของพหุนามไบเกาส์เซียนเพลล์-ลูคัส ในลักษณะเดียวกัน นอกจากนี้ยังได้ศึกษาผลบวก ผลต่าง ผลบวกย่อย ผลบวกย่อยพจน์คู่ ผลบวกย่อยพจน์คี่ และเมทริกซ์ของพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัส ซึ่งผลการศึกษาปรากฏว่า สามารถเขียนผลบวกย่อย ผลบวกย่อยพจน์คี่ และพจน์คู่ของพหุนามไบเกาส์เซียนเพลล์ ได้ดังนี้

$\sum_{k=1}^{n} BGP_k(x) = \frac{1}{2x}(BGP_{n+1}(x)+BGP_n(x)-4(x^2-x+1))$

$\sum_{k=1}^{n} BGP_{2k-1}(x) = \frac{1}{2x}(BGP_{2n}(x)-2(2x^2-x+1))$ และ

$\sum_{k=1}^{n} BGP_{2k}(x) = \frac{1}{2x}(BGP_{2n+1}(x)+2(x-1))$ ตามลำดับ

สามารถเขียนผลบวกย่อย ผลบวกย่อยพจน์คี่ และพจน์คู่ของพหุนามไบเกาส์เซียนเพลล์-ลูคัส ได้ดังนี้

$\sum_{k=1}^{n} BGQ_k(x) = \frac{1}{2x}(BGQ_{n+1}(x)+BGQ_n(x)+8(x-1)(x^4+1))$

$\sum_{k=1}^{n} BGQ_{2k-1}(x) = \frac{1}{2x}(BGQ_{2n}(x)+4(2x-1)(x^2+1))$ และ

$\sum_{k=1}^{n} BGQ_{2k}(x) = \frac{1}{2x}(BGQ_{2n+1}(x) - 4(x^2+1))$ ตามลำดับ

สามารถเขียนผลบวก และผลต่าง ของพหุนามไบเกาส์เซียนเพลล์ ได้ดังนี้

$$BGP_{n+r}(x) + BGP_{n-r}(x) = \begin{cases} Q_r(x)BGP_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ P_r(x)BGQ_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases}$$

$$BGP_{n+r}(x) - BGP_{n-r}(x) = \begin{cases} P_r(x)BGQ_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ Q_r(x)BGP_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases}$$

สามารถเขียนผลบวก และผลต่างของพหุนามไบเกาส์เซียนเพลล์-ลูคัส ได้ดังนี้

$$BGQ_{n+r}(x) + BGQ_{n-r}(x) = \begin{cases} Q_r(x)BGQ_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ 4(x^2+1)P_r(x)BGP_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases}$$

$$BGQ_{n+r}(x) - BGQ_{n-r}(x) = \begin{cases} 4(x^2+1)P_r(x)BGP_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคู่} \\ Q_r(x)BGQ_n(x) & \text{เมื่อ } r \text{ เป็นจำนวนคี่} \end{cases}$$

นอกจากนี้เราสามารถศึกษาทฤษฎีบทอื่น ๆ เกี่ยวกับพหุนามไบเกาส์เซียนเพลล์ และพหุนามไบเกาส์เซียนเพลล์-ลูคัสได้ในงานวิจัยนี้

เอกสารอ้างอิง

[1] E. Ozkan and T. Alp, biGaussian Pell and Pell-Lucas polynomials, Arts and Sciences 53 (3) (2022) 17–25.

[2] A. Dasdemir, On the Pell, Pell-Lucas and Modified Pell numbers by Matrix Method, Arts and Sciences 64 (5) (2011) 3173–3181.

[3] A. F. Horadam and Bro. J. M. Mahon, Pell and Pell-Lucas polynomials, Australia (1983) .

[4] S. Halici and S. Oz, On Gaussian Pell polynomials and Their some properties, Palestine Journal of Mathematics. 7 (1) (2018) 251–256.

[5] T. Yagmur, Gaussian Pell-Lucas polynomials, Communications in Mathematics and Applications. 10 (4) (2019) 1609–1617.

[6] H. Gokbas, A note on biGaussian Pell and Pell-Lucas numbers, Sciences and Arts 56 (3) (2021) 669–680.

Optimal Parameters for an Investing Strategy based on Relative Strength Index (RSI) Indicator

Kamonwan Konghom[1,*], Weenakorn Ieosanurak[1] and Watcharin Klongdee[1,†]

[1]Department of Mathematics, Faculty of Science
Khon Kaen University, Khon Kaen 40002, Thailand

Abstract

Confidence in the global economy is declining in a world filled with turmoil, such as wars and economic problems. Investors are turning to safe assets like gold, which helps reduce the risks of inflation and recession. Gold investment has become increasingly popular, with investors using various strategies and indicators. One is the Relative Strength Index (RSI), an easy-to-use tool that helps capture trading opportunities through trend analysis and market signals. This research, therefore, studied the effectiveness of RSI in various forms to determine the maximum profit outcomes. Analyzed four different RSI strategies, including the original version (TRSI) and the improved versions (IRSI-1, IRSI-2, IRSI-3), which include delays, take-profit, and stop-loss. The parameters of TRSI are (B_{rsi}, S_{rsi}). The parameters of IRSI-1 are $(B_{rsi}, S_{rsi}, tp, sl)$. The parameters of IRSI-2 are $(B_{rsi}, S_{rsi}, delay)$. The parameters of IRSI-3 are $(B_{rsi}, S_{rsi}, tp, sl, delay)$. The parameters $B_{rsi}, S_{rsi}, tp, sl, delay$ are the buy signal of the RSI, sell signal of the RSI, take-profit, stop-loss and delay, respectively. The delay is a number of days ago. This study used daily data from the gold market between 2020-2023. The data is divided into two parts: training data from 2020-2022 and tested data in 2023. The results from the training data showed that the traditional RSI yielded the highest profit of 1742.9, IRSI-1 the highest profit of 1014.8, IRSI-2 the highest profit of 1743.9 and IRSI-3 the highest profit of 1011.9. We found that TRSI and IRSI-2 provided the highest profits while the results from the training data were the same.

Keywords: RSI, Investing Strategy, Delay, Take-profit and stop-loss, Optimization

2020 MSC: 91G10, 91B84, 62P05

1 Introduction

In today's chaotic world, events such as wars and economic crises are occurring like never before and are unlikely to end anytime soon. These events have impacted global economic confidence. Increasing investors' demand for safe-haven assets, such as gold, is valuable during a crisis. Many investors are considering investing in a gold portfolio. Gold helps prevent deficits and economic

*Speaker (Kamonwan.ko@kkumail.com)
†Corresponding author (kwatch@kku.ac.th)

downturns, making it popular as a safe-haven asset. Investors are focusing on gold, each with their unique trading strategies. This research will explore the power of the Relative Strength Index indicator. (RSI). These tools are among the most popular and user-friendly trend indicators. They indicate price strength, allowing for assessing overbought or oversold conditions. They can identify price reversal points, helping investors track trading momentum and develop trading ideas through the Relative Strength Index (RSI), which measures overbought or oversold levels from 1 to 100. The default setting for the RSI is 14, which means tracking the value over the past 14 periods on the chart, whether in days or minutes. RSI is a versatile indicator and is used in various ways, such as Hari and Dewi [1] aim to assist the community by analyzing stock trends using the RSI indicator and Moving Average Convergence Divergence (MACD) to determine the appropriate timing for buying or selling. Chong, Eung, and Liao [2] Considered the unexpected returns in the London market. He found that RSI(21,50), which uses the RSI indicator calculated with historical price data over 21 candles and the value 50 as a criterion for determining the direction and trend of the stock market, has consistently generated significant abnormal returns in the Milan Comit General Index and the S&P/TSX Composite Index. Additionally, the RSI(14,30/70) rule has also generated profits in the Dow Jones Industrial Index, using 14 historical candles for calculation, with 30/70 as the criterion for determining whether a stock is overbought or oversold Panigrahi1, Vachhani1 and, Chaudhury [3] proposed the use of RSI as a momentum indicator in situations of overbought and oversold conditions by assessing two return trends: 50-50 where RSI movements are near the 50 level, indicating a balanced market or unclear trend (Sideways Market), and 60-40 indicating a bullish market trend with buying pressure greater than selling pressure. With NIFTY 50, they found that the 50-50 strategy provided higher long-term returns, while the 60-40 strategy offered better short-term returns. Bartkus [4] Compare the performance of the classic RSI method with the adjusted formula to increase speed using RSI at eight different hourly intervals (H) as follows: H8, H10, H12, H14, H16, H18, H20, and H22. However, only one of them yielded positive results, including the analysis. Phuong [5] used RSI to measure investor sentiment on Vietnam's stock market; it was found that investor emotions play a significant role in explaining stock market returns. Cervelló-Royo et al. [6] analyzed the integration of TP and SL with the MACD system and found that appropriate threshold levels can improve profitability and reduce drawdowns. Similarly, QuantInsti [7] proposed a flexible framework for optimizing exit conditions using TP and SL as adaptive parameters, showing that this approach can enhance performance in dynamic market conditions. Additionally, Vezeris et al. [8] examined various TP and SL strategies within a MACD-based automated trading system across multiple asset classes, concluding that certain configurations, such as a sliding ATR-based stop-loss, yielded superior results. Cartea and Sánchez-Betancourt [9] modeled optimal trade execution with stochastic delay, demonstrating how delayed actions could lead to better timing and reduced slippage. Moallemi and Saglam [10] further emphasized that execution latency can significantly affect profitability in high-frequency trading environments, making delay a critical component in algorithmic strategy design. Furthermore, Hsieh et al. [11] investigated the impact of execution delay on Kelly-based stock

trading strategies, revealing that delays can influence the comparative performance of high-frequency trading versus buy-and-hold approaches. This work has added delay because the stock market often behaves based on past data. Therefore, we introduced a delay, the number of days, to wait for buying and selling decisions. We also observed market price trends, including take-profit and stop-loss, which are acceptable profits and minimal losses, respectively, to reduce potential investment risks. We presented it in four parts: a literature review discussing Simple Moving Average(SMA), Exponential Moving Average (EMA), and RSI; a methodology explaining the process of improving RSI, including TRSI, by analyzing and finding the most suitable parameters for trading financial assets using RSI (Relative Strength Index) and trading signals to maximize profits from parameter adjustments; IRSI-1, which develops the traditional RSI strategy to be more efficient by adding conditions to prevent losses (stop -loss, sl) and appropriate profits (take-profit, tp) to help reduce risks and increase the chances of achieving profit targets; IRSI-2, which uses the RSI indicator to create buy/sell signals in the market by adding a delay in response to signals, helping to avoid noise and better capture opportunities during volatile periods; and IRSI-3, which improves RSI trading by combining conditions to enhance efficiency and reduce risk, focusing on position management by considering buy-sell signals along with conditions controlling losses and profits, including using a delay in opening or closing positions to find the most suitable relationship; and results and conclusions.

2 Preliminaries

This section deals with literary theory and related research.

Let p_k be the value at k by $k = 1,2,3,\ldots$ for convenience, we give P a set of asset price sequences of $p_1, p_2, \ldots, p_k, \ldots$

2.1 Simple Moving Averages

Simple moving average (SMA) is a mathematical value commonly used in statistical and financial calculations. It involves putting a set of values, such as stock prices, together over a given period, adding a plus to the amount of value in a set. The result is an average that represents a smooth trend over that time. SMA is often used in technical analysis of financial markets to identify trends and trade decisions. The selection of intervals for SMA may affect sensitivity to price movements. More extended periods make the average smoother, and shorter periods respond better to recent price changes. [3]

Calculate from

$$SMA(t+1;n) = \frac{(p_t + p_{t-1} + \cdots + p_{t-n+1})}{n}. \quad (1)$$

Where $SMA(t+1;n)$ is the SMA calculated at the next time, $t+1$. $p_t, p_{t-1}, \ldots, p_{t-n+1}$ are the prices at different time periods. n is the number of time periods.

2.2 Exponential Moving Averages

The exponential moving average (EMA) differs from the simple moving averages (SMA). The SMA calculates the average price information, while the EMA weighs more on current data. Due to its unique calculation, the EMA tracks prices more closely than the relevant SMA. EMA responds more effectively to recent price changes. In trading, it is more sensitive to price movements and can see trends faster than the SMA.
Calculate from

$$\mathrm{EMA}(t+1;\ n, P) = \left(\frac{2}{n+1}\right) \times\ p_t +\ \left(1\ -\frac{2}{n+1}\right) \times \mathrm{EMA}(t;\ n, P). \tag{2}$$

Where $\mathrm{EMA}(t+1; n)$ is the EMA calculated at the next time, $t+1$. p_t is the price at time t. n is the number of time periods. P is a set of asset price sequences $p_t, p_{t-1}, \ldots, p_{t-n+1}$.

2.3 Relative Strength indicator (RSI)

2.3.1 Up Change and Down Change, UC and DC

Down Change (DC) refers to the change in the closing price on a day when the closing price is lower than the previous day, and Up Change (UC) refers to the change in the closing price on the current day when the closing price is higher than the previous day's closing price. The value of UC is calculated only when the closing price on the current day is higher than the previous day.

For closing dates higher than the previous day, DC will be zero, and UC will be the closing day deducted by closing the preceding day. For closing days lower than the prior day, UC is zero, and DC is the closure day deduced by the closing price of that day. The day's cost is equal to the previous year's price for closure days. Both UC and DC are zero.[12]
Calculate from

$$\mathrm{UC}_t = \begin{cases} p_t - p_{t-1}, & p_t - p_{t-1} \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

$$\mathrm{DC}_t = \begin{cases} p_{t-1} - p_t, & p_{t-1} - p_t \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Where p_t is the closed price at time t and t is a period in time.

2.3.2 Relative Strength Indicator (RSI)

The RSI indicates the strength of a market that uses over-purchase or over-sales marks by representing a range between 0 and 100. A value higher than 70 usually indicates an over-purchasing status. While a value lower than 30 often indicates too-selling status, the RSI also applies to trends.

Calculate from

$$RS_t = \frac{EMA(t;n,UC_t)}{EMA(t;n,DC_t)}. \quad (5)$$

$$RSI(t;n,P) = 100 - \frac{100}{1+RS_t}. \quad (6)$$

Where n is the number of EMA retrograde days and t is the time of interest, $t = n + 1, n + 2, \ldots$

Figure 1 (top) shows RSI calculated Equation (6) for the gold market from 2008 to 2010 is shown on the vertical axis, and the dates are displayed on the horizontal axis. Figure 1 (bottom) shows the date values on the horizontal axis and the asset prices on the vertical axis. By green points indicating buy signal, red points indicating sell signal, and purple points indicating closing buy or sell positions.

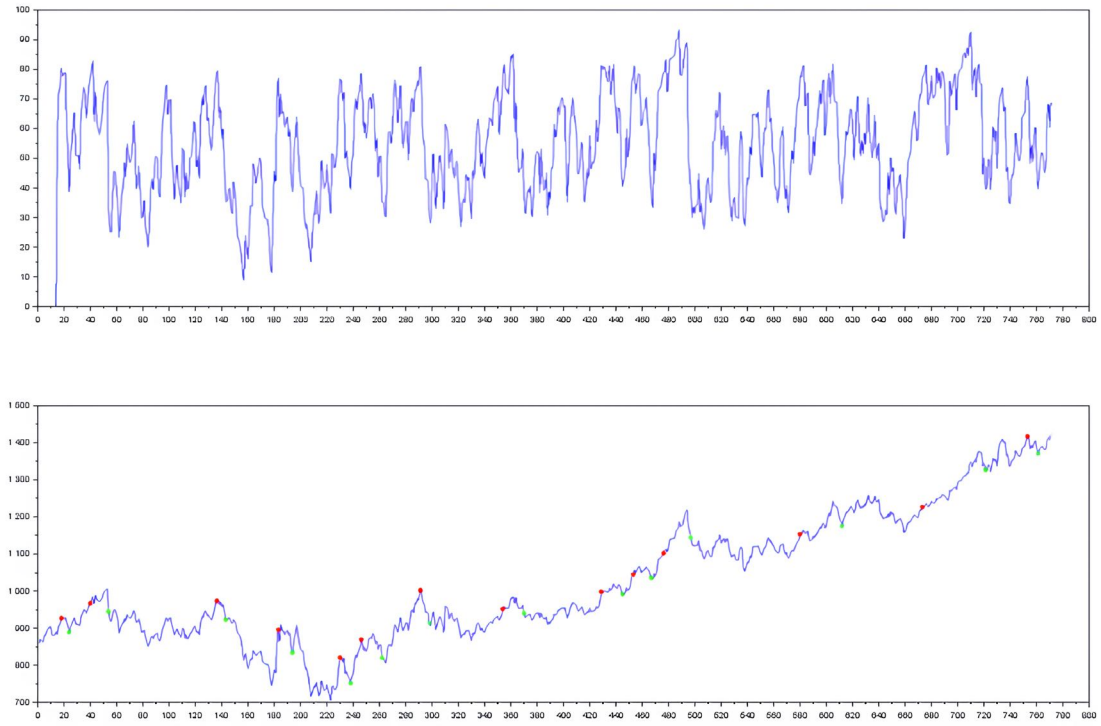


Figure 1. An example of the RSI graph in the gold market from 2008-2010, with green points indicating buying, red points indicating selling, and purple points indicating closing buy or sell positions.

2.4 Profitability Formula

$$\text{Profit} = \text{Close price} - \text{Open price},$$

where close price is the price of an asset at the closing time of the period of interest. Open price is the price of the asset at the opening time of that period.

3 Main Results

3.1 Methodology

The purpose of this research is to maximize profitability from the correlation indicator. In this section, we studied the profitability of different trading strategies using traditional correlations, using daily close price of gold per ounce data (USD per Troy Ounce) from Investing website from January 1, 2008, to December 31, 2022.

Let's start by defining some key terms. Let p_k represent the price of an asset at k, where k can be any positive integer for convenience. We'll use P to denote the sequence of $p_1, p_2,$ and so on.

Assumption.

1. Do not consider the transaction fees.
2. No orders, more than one order.
3. In the case of an outstanding order at the final time, we will not consider the order.

Definition 1. Let P be the current asset price, sl be the acceptable loss rate $(0 < sl < 1)$, and tp be the desired profit rate $(0 < tp < 1)$. The stop-loss price (P_{sl}) and take-profit (P_{tp}) are defined as follows:

$$P_{sl} = \begin{cases} P(1+sl), & \text{for sell position,} \\ P(1-sl), & \text{for buy position,} \end{cases}$$

and

$$P_{tp} = \begin{cases} P(1-tp), & \text{for sell position,} \\ P(1+tp), & \text{for buy position.} \end{cases}$$

Relative Strength Indicator Strategy (RSI Strategy)

Let P be a set of asset price sequences, n be the number of periods of interest, and t be the specific time of interest. For the RSI strategy calculated from Equation (6), buy and sell signals as followed

RSI Signal Rule.

Buy signals can be obtained from:

$$RSI_t\,(P; n, t) \leq B_{rsi}.$$

The selling signal comes from:

$$RSI_t\,(P; n, t) \geq S_{rsi}.$$

When B_{rsi} is a parameter for determining buy signals (meaning the point at which an asset is being oversold) S_{rsi}, is a parameter for deciding sell signals (meaning the point at which an asset is being overbought), with both B_{rsi} and S_{rsi} ranging from 0 to 100, generally $B_{rsi} < S_{rsi}$.

Traditional RSI Strategy (TRSI)

This strategy will consist of the following steps;

1. Calculate the values of UC_t and DC_t.
2. Calculate the EMA value.
3. Calculate the RSI using the Equation (6).
4. Consider each set of parameters (B_{rsi}, S_{rsi}) to evaluate profitability, where $B_{rsi} \in \{0.01,0.02,0.03,\dots,49.99,50\}$ and $S_{rsi} \in \{50.01,50.02,50.03,\dots,99.99,100\}$. We will analyze trading signals based on RSI signal rule. If a sell signal is detected, we will close the buy position and open a sell position for the asset; if there are no positions, we will open a sell position. Similarly, if a buy signal is detected, we will close the sell position and open a buy position for the asset; if there are no positions, we will open a buy position. We will then record the parameter values along with the profits obtained for each set of parameters and compare the profits of each parameter set to find.

Improved RSI Strategy with TP and SL (IRSI-1)

In this topic, we propose an improved RSI strategy that will differ from the traditional RSI because it includes conditions to prevent losses (stop loss, sl) and achieve acceptable profits (take profit, tp) when considering buy and sell signals. "Buy" will represent buy signals, "Sell" will represent sell signals, "sl" will represent acceptable losses, and "tp" will represent desired profits. This method is called IRSI-1. The values for sl and tp value is in the range of 0-0.1%. The parameters of IRSI-1 are $(B_{rsi}, S_{rsi}, tp, sl)$. This strategy will consist of the following steps:

1. Calculate the values of UC_t and DC_t.
2. Calculate the EMA value.
3. Calculate the RSI using the Equation (6).
4. Consider each set of parameters $(B_{rsi}, S_{rsi}, tp, sl)$ to evaluate profitability, where $B_{rsi} \in \{0.01,0.02,0.03,\dots,49.99,50\}, S_{rsi} \in \{50.01,50.02,50.03,\dots,99.99,100\}, sl \in (0,0.1)$ and $tp \in (0,0.1)$. Trading signals will be assessed based on the RSI signal rule provided. If a sell signal is detected, the buy position will be closed, and a sell position will be opened. The sell position will be closed if the asset price reaches the P_{sl} value or drops to the desired P_{tp}. If there are no open positions, a sell position will be initiated. Similarly, if a buy signal is detected, the sell position will be closed, and a buy position will be opened. The buy position will be closed if the asset price reaches the desired P_{tp} or falls to the acceptable P_{sl}. If there are no open positions, a buy position will be initiated. The parameters and the resulting profits for each set of parameters will be recorded, and the

profits from each parameter set will be compared to identify the set that yields the highest profitability for this strategy.

Improved RSI Strategy with Delay (IRSI-2)

For this RSI strategy, there will be a delay in considering buy and sell signals. This method is called IRSI-2. The delay is the time lag, which can range from 1, 2, 3, ... to 8 days. The parameters of IRSI-2 are $(B_{rsi}, S_{rsi}, \text{delay})$. This strategy will consist of the following four steps.

1. Calculate the values of UC_t and DC_t.
2. Calculate the EMA value.
3. Calculate the RSI using the Equation (6).
4. Consider each set of parameters $(B_{rsi}, S_{rsi}, \text{delay})$ to evaluate profitability, where $B_{rsi} \in \{0.01, 0.02, \dots, 49.99, 50\}, S_{rsi} \in \{50.01, 50.02, \dots, 99.99, 100\}$ and $\text{delay} \in \{1, 2, \dots, 8\}$. Trading signals will be assessed based on the RSI signal rule provided. If a sell signal is detected, we will wait for delay days, then close the buy position and open a sell position. If there are no positions, we will wait for delay days to open a sell position. Similarly, if a buy signal is detected, we will wait for delay days, then close the sell position and open a buy position. We will wait for days to open a buy position if there are no positions. We will also record the parameter values along with the profits obtained for each set of parameters and compare the earnings of each set of parameters to find the set that maximizes the profitability of this strategy.

Improved RSI Strategy with TP, SL, and Delay (IRSI-3)

For this RSI strategy, there will be conditions to prevent losses (sl), achieve an acceptable profit (tp), and delay in considering buy and sell signals. A sl will represent an acceptable loss, and tp will represent the desired profit. Delay will indicate the time lag in consideration, with values ranging from 1, 2, 3, ..., to 8 days, and both sl and tp value is in the range of 0-0.1%. This method is called IRSI-3. The parameters of IRSI-3 are $(B_{rsi}, S_{rsi}, \text{tp}, \text{sl}, \text{delay})$. This strategy will consist of the following four steps.

1. Calculate the values of UC_t and DC_t.
2. Calculate the EMA value.
3. Calculate the RSI using the Equation (6).
4. Consider each set of parameters $(B_{rsi}, S_{rsi}, \text{tp}, \text{sl}, \text{delay})$ to evaluate profitability, where $B_{rsi} \in \{0.01, 0.02, 0.03, \dots, 49.99, 50\}, S_{rsi} \in \{50.01, 50.02, 50.03, \dots, 99.99, 100\}$, $\text{sl} \in (0, 0.1), \text{tp} \in (0, 0.1)$ and $\text{delay} \in \{1, 2, \dots, 8\}$. Based on RSI signal rule, if a sell signal is detected, we will wait for delay days, close the buy position, and open a sell position. If, during that time, the asset price rises to P_{sl} or falls to the desired P_{tp}, we will proceed to close the sell position. However, if there are no positions, we will wait delay days to open a sell position. Similarly, if a buy signal is detected, we will wait for delay days, close the sell position, and open a buy position for the asset. If, during that time, the asset price

rises to the desired P_{tp} or falls to an acceptable P_{sl}, we will proceed to close the buy position. If there are no positions, we will wait for delay days to open a buy position and then record the parameter values, and the profits obtained for each set of parameters. We will compare the profits of each set of parameters to find the set that maximizes the profitability of this strategy.

To evaluate the effectiveness of each strategy, the dataset was divided into two parts: training data from 2020 to 2022 and testing data in 2023. The training data was used to determine the optimal parameters for each strategy, including $B_{rsi}, S_{rsi}, tp, sl,$ and $delay$. These parameters were selected based on the highest profitability obtained from the training data. Once the optimal parameters were identified, they were applied to the testing data in 2023 without further modification, allowing us to assess the out-of-sample performance of each strategy. This approach helps ensure that the evaluation reflects a realistic trading scenario and avoids overfitting to historical data.

3.2 Experiment Results

From the study using gold price data from 2020-2023, we divided the data into two parts: training data from 2020-2022 (777 entries) and tested data in 2023 (257 entries). It was found that different types of RSI yielded varying profits as follows: TRSI, IRSI-1, IRSI-2, and IRSI-3. From training data, the maximum profits of each presented strategy are $1742.9\ (B_{rsi} = 21.06, S_{rsi} = 83.96)$, $1014.8\ (B_{rsi} = 54.69, S_{rsi} = 0, tp = 0.04, sl = 0.09)$, $1743.9\ (B_{rsi} = 22.80, S_{rsi} = 83.16, delay = 3)$, and $1011.9\ (B_{rsi} = 54.70, S_{rsi} = 53.17, tp = 0, sl = 0.004, delay = 0)$ respectively. We found that TRSI and IRSI-2 provided the highest profits, which were significantly different from IRSI-1 and IRSI-3. As a result, the profits in the 2023 testing were similar, unlike IRSI-1 and IRSI-3, which incurred losses. The addition of delay serves as a factor in considering trading signals, and possibly due to the large number of orders, the profits decreased. However, the added delay is beneficial as it helps us consider signals. The results from the training data provided profits close to the highest ones, leading to the highest profit in tested data, as shown in Tables 1 and 2.

From Table 1, the TRSI and IRSI-2 strategies provided the highest profits during the training period (2020–2022), with values of 1742.9 and 1743.9 respectively. TRSI generated only 10 orders while IRSI-2 executed 11 orders, indicating effective signal filtering. In contrast, IRSI-1 and IRSI-3, which include take-profit and stop-loss conditions, had significantly more orders (131 and 132) and yielded lower profits of 1014.8 and 1011.9. This suggests that strategies without TP/SL may offer higher return potential during the training period.

Table 1. The profit results for each RSI strategy during the training period from 2020 to 2022.

Strategy	Year	Profit	Parameters	Number of orders
TRSI (B_{rsi}, S_{rsi})	2020-2022	1742.9	(21.06,83.96)	10
IRSI-1 $(B_{rsi}, S_{rsi}, tp, sl)$	2020-2022	1014.8	(54.69,0,0.04,0.09)	131
IRSI-2 $(B_{rsi}, S_{rsi}, delay)$	2020-2022	1743.9	(22.80,83.16,3)	11
IRSI-3 $(B_{rsi}, S_{rsi}, tp, sl, delay)$	2020-2022	1011.9	(54.70,53.17,0,0.004,0)	132

Table 2. The profit results for each RSI strategy in 2023 during testing.

Strategy	Year	Profit	Parameters	Number of orders
TRSI (B_{rsi}, S_{rsi})	2023	169.7	(21.06,83.96)	2
IRSI-1 $(B_{rsi}, S_{rsi}, tp, sl)$	2023	-62.3	(54.69,0,0.04,0.09)	29
IRSI-2 $(B_{rsi}, S_{rsi}, delay)$	2023	152.4	(22.80,83.16,3)	2
IRSI-3 $(B_{rsi}, S_{rsi}, tp, sl, delay)$	2023	-110.5	(54.70,53.17,0,0.004,0)	30

From Table 2, the TRSI and IRSI-2 strategies also showed strong performance during the testing period in 2023, with profits of 169.7 and 152.4, respectively. These strategies maintained consistency with their training performance. On the other hand, IRSI-1 and IRSI-3 incurred losses of -62.3 and -110.5, despite generating 29 and 30 trades, respectively. This result further emphasizes the robustness of TRSI and IRSI-2 in real-market scenarios.

Table 3. The profit/loss results for the TRSI strategy training period from 2020 to 2022.

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Oder	Sell Order	Buy Order	Sell Order	
37	1648.8		✓			0
53	1486.5	✓			✓	162.3
145	1890		✓	✓		403.5
237	1780.9	✓			✓	109.1
261	1954.4		✓	✓		173.5
303	1700.7	✓			✓	253.7
361	1901.2		✓	✓		200.5
377	1774.8	✓			✓	126.4
556	1926.3		✓	✓		151.5
647	1763.9				✓	162.4
					Total	1742.9

From Table 3, the TRSI strategy with training data shows that the first signal encountered was a sell signal on day 37 when the asset price was 1648.8. Over time, a buy signal was encountered on day 53 when the asset price was 1486.5, leading to the closing of the sell order and opening of a buy order, resulting in a profit of 162.3. Then, a sell signal was encountered on day 145 when the asset price was 1890, leading to the closing of the buy order and opening of a sell order, with a profit of 403.5. Finally, on day 647, a sell signal was encountered when the asset price was 1763.9, leading to the closing of the sell order with a profit of 162.4. At the end of the trading period, the total profit was 1742.9, corresponding to Figure 2, top graph is RSI and bottom graph is asset price which shows a sell position marked by a red dot and a buy position marked by green dot.

Figure 2 illustrates the trading performance of the TRSI strategy applied to gold prices from 2020 to 2022. It includes buy signals (green dots), sell signals (red dots), and closing positions (purple dots). This visualization helps demonstrate how the strategy executed trades based on traditional RSI signals without the use of stop-loss, take-profit, or delay conditions.

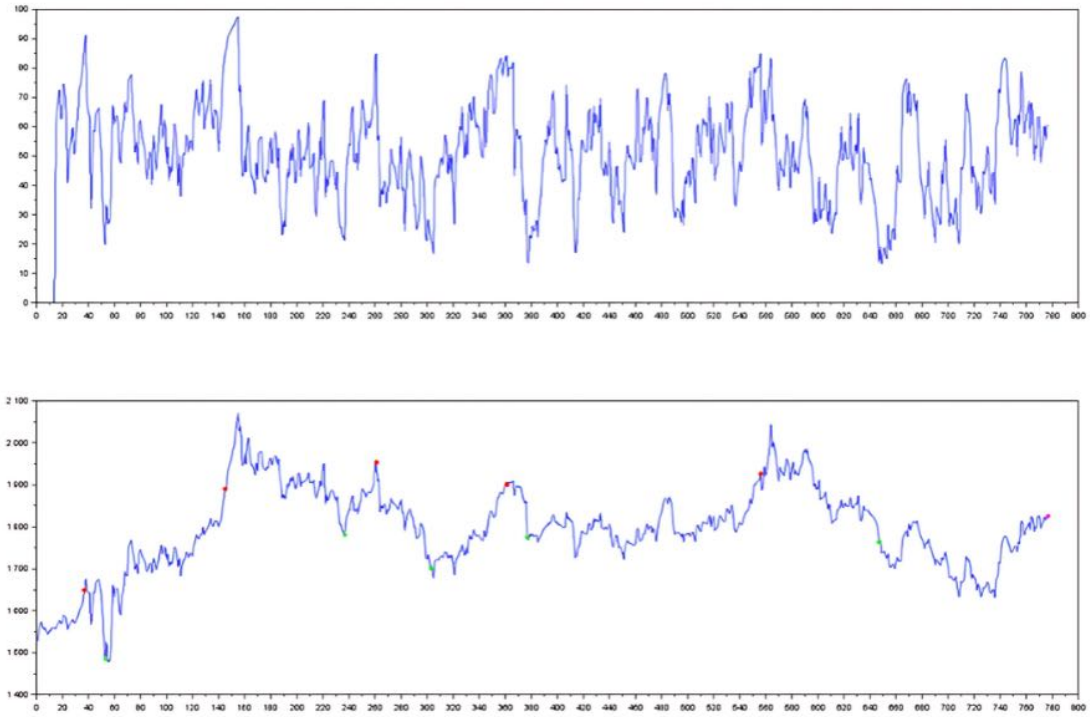


Figure 2. The price of TRSI in the gold market from 2020 to 2022, along with appropriate buy and sell points. Green dots represent buy points and sell points with red dots, and purple dots indicate the positions and times for closing trades.

Table 4. The profit/loss results for the IRSI-2 strategy training period from 2020 to 2022.

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
40	1643.1		✓			0
56	1479.3	✓			✓	163.8
148	1944.6		✓	✓		465.3
238	1818.9	✓			✓	125.7
263	1913.6		✓	✓		94.7
302	1715.8	✓			✓	197.8
360	1898		✓	✓		182.2
380	1777.4	✓			✓	120.6
559	1943.8		✓	✓		166.4
650	1739.7	✓			✓	204.1
746	1763			✓		23.3
					Total	1743.9

From Table 4, the IRSI-2 strategy incorporates a delay parameter, which represents the number of days the system waits after receiving a buy or sell signal before executing the trade. In Table 4, this delay is set to 3 days, meaning that the actual trade is executed three days after the signal is generated. The purpose of this delay is to reduce the influence of short-term market noise and avoid false signals, which can lead to premature or suboptimal trades. From the results, it can be observed that despite fewer trades (only 11 orders), the profits remain high, comparable to the traditional RSI (TRSI). This indicates that the delay helps in filtering out less reliable signals and improves the timing of trades, thereby enhancing the overall profitability and reducing overtrading.

Table 5. The profit/loss results for the TRSI strategy tested period 2023.

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
32	1845.3	✓				0
214	2015			✓		169.7
					Total	169.7

From Table 5, the TRSI strategy with tested data shows that the first signal encountered was a buy signal on day 32 when the asset price was 1845.3. Over time, on day 214, a sell signal was encountered when the asset price was 2015, leading to the closing of the buy order with a profit of 169.7. At the end of the trading period, the total profit was 169.7.

Table 6. The profit/loss results for the IRSI-2 strategy tested period 2023.

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
35	1842.5	✓				0
211	1994.9			✓		152.4
					Total	152.4

From Table 6, the IRSI-2 strategy with tested data shows that the first signal encountered was a buy signal on day 35 when the asset price was 1842.5. Over time, on day 211, a sell signal was encountered when the asset price was 1994.9, leading to the closing of the buy order with a profit of 152.4. At the end of the trading period, the total profit was 152.4.

Table 7. The profit/loss results for the IRSI-1 strategy training period from 2020 to 2022.

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
15	1556.7		✓			0
24	1555.5	✓			✓	1.2
27	1573.4		✓	✓		17.9
29	1570.1	✓			✓	3.3
31	1578.8		✓	✓		8.7
42	1566.7	✓			✓	12.1
44	1644.4		✓	✓		77.7
50	1642.3	✓			✓	2.1
59	1660.8		✓	✓		18.5
64	1596.6	✓			✓	64.2
66	1637.7		✓	✓		41.1
76	1698.8	✓			✓	-61.1
77	1711.2		✓	✓		12.4
78	1687.8	✓			✓	23.4
79	1738.3		✓	✓		50.5
82	1723.8	✓			✓	14.5
90	1725.8		✓	✓		2
91	1713.9	✓			✓	11.9
95	1740.9		✓	✓		27
100	1721.9	✓			✓	19
106	1751.7		✓	✓		29.8
108	1734	✓			✓	17.7
115	1739.8		✓	✓		5.8
117	1727.2	✓			✓	12.6
118	1736.5		✓	✓		9.3
120	1731.1	✓			✓	5.4
121	1753		✓	✓		21.9
140	1800.3	✓			✓	-47.3
141	1810		✓	✓		9.7
158	1946.3	✓			✓	-136.3
162	1998.7		✓	✓		52.4
164	1970.3	✓			✓	28.4
171	1974.9		✓	✓		4.6
174	1944.7	✓			✓	30.2

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
180	1964.3		✓	✓		19.6
181	1947.9	✓			✓	16.4
182	1963.7		✓	✓		15.8
185	1949.9	✓			✓	13.8
201	1926.2		✓	✓		-23.7
203	1894.6	✓			✓	31.6
209	1929.5		✓	✓		34.9
210	1904.6	✓			✓	24.9
218	1910.4		✓	✓		5.8
219	1896.2	✓			✓	14.2
220	1946.8		✓	✓		50.6
222	1854.4	✓			✓	92.4
242	1866		✓	✓		11.6
244	1838.5	✓			✓	27.5
248	1855.3		✓	✓		16.8
264	1835.4	✓			✓	19.9
280	1863.9		✓	✓		28.5
281	1833.4	✓			✓	30.5
314	1741.7		✓	✓		-91.7
315	1738.1	✓			✓	3.6
325	1743		✓	✓		4.9
329	1732.7	✓			✓	10.3
330	1747.6		✓	✓		14.9
331	1736.3	✓			✓	11.3
332	1766.8		✓	✓		30.5
342	1768.3	✓			✓	-1.5
344	1791.8		✓	✓		23.5
345	1776	✓			✓	15.8
346	1784.3		✓	✓		8.3
367	1873.3	✓			✓	-89
368	1892		✓	✓		18.7
373	1879.6	✓			✓	12.4
391	1802.1		✓	✓		-77.5
392	1800.2	✓			✓	1.9
393	1810.6		✓	✓		10.4
399	1809.2	✓			✓	1.4

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
407	1835.8		✓	✓		26.6
410	1814.1	✓			✓	21.7
419	1789.8		✓	✓		-24.3
420	1787.8	✓			✓	2
424	1806.3		✓	✓		18.5
426	1791	✓			✓	15.3
428	1819.5		✓	✓		28.5
434	1798.5	✓			✓	21
440	1807.1		✓	✓		8.6
441	1794.8	✓			✓	12.3
461	1794.7		✓	✓		-0.1
463	1768.3	✓			✓	26.4
466	1784.9		✓	✓		16.6
473	1783.9	✓			✓	1
474	1795.8		✓	✓		11.9
475	1789.4	✓			✓	6.4
478	1816.8		✓	✓		27.4
489	1806.3	✓			✓	10.5
507	1798.2		✓	✓		-8.1
509	1794.6	✓			✓	3.6
511	1802.2		✓	✓		7.6
518	1800.1	✓			✓	2.1
519	1814.6		✓	✓		14.5
521	1789.2	✓			✓	25.4
524	1818.5		✓	✓		29.3
527	1816.5	✓			✓	2
530	1843.2		✓	✓		26.7
535	1829.7	✓			✓	13.5
543	1821.8		✓	✓		-7.9
557	1887.6	✓			✓	-65.8
558	1900.7		✓	✓		13.1
568	1960.8	✓			✓	-60.1
576	1962.2		✓	✓		1.4
577	1954.2	✓			✓	8
581	1954		✓	✓		-0.2
582	1923.7	✓			✓	30.3

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
588	1948.2		✓	✓		24.5
593	1959	✓			✓	-10.8
618	1865.4		✓	✓		-93.6
619	1846.3	✓			✓	19.1
621	1857.3		✓	✓		11
622	1848.4	✓			✓	8.9
625	1871.4		✓	✓		23
626	1850.2	✓			✓	21.2
631	1875.5		✓	✓		25.3
632	1831.8	✓			✓	43.7
665	1769.2		✓	✓		-62.6
678	1789.7	✓			✓	-20.5
698	1740.6		✓	✓		-49.1
699	1717.4	✓			✓	23.2
713	1702		✓	✓		-15.4
718	1675.2	✓			✓	26.8
737	1676.6		✓	✓		1.4
748	1739.6	✓			✓	-63
750	1745.6		✓	✓		6
753	1740.3	✓			✓	5.3
754	1763.7		✓	✓		23.4
766	1787.8	✓			✓	-24.1
769	1825.4		✓	✓		37.6
771	1795.3	✓			✓	30.1
774	1823.1			✓		27.8
Total						1014.8

From Table 7, the IRSI-1 strategy incorporates take-profit (tp) and stop-loss (sl) conditions to manage risk and secure gains. The first signal encountered was a buy signal on day 15 when the asset price was 1556.7. A sell signal occurred on day 24 at the price of 1555.5, leading to a small loss due to the stop-loss condition being triggered. Later, a buy signal appeared on day 27 at 1573.4, which closed on the same day with a profit of 17.9 due to reaching the take-profit target. Throughout the training period, the strategy continuously opened and closed positions whenever the take-profit or stop-loss thresholds were hit. For example, on day 44, a buy position at 1566.7 was closed at 1644.4, yielding a profit of 77.7. On day 76, however, the strategy faced a loss of 61.1 due to stop-loss conditions. The strategy resulted in a total of 131 trades during the period, generating a total profit of 1014.8. Despite the lower average profit per

trade compared to TRSI and IRSI-2, the risk control through tp and sl helped to stabilize performance by minimizing large losses.

Figure 3 illustrates the trading performance of the IRSI-1 strategy applied to gold prices from 2020 to 2022. It includes buy signals (green dots), sell signals (red dots), and closing positions (purple dots). Additionally, transparent circles highlight the points where take-profit or stop-loss conditions were triggered.

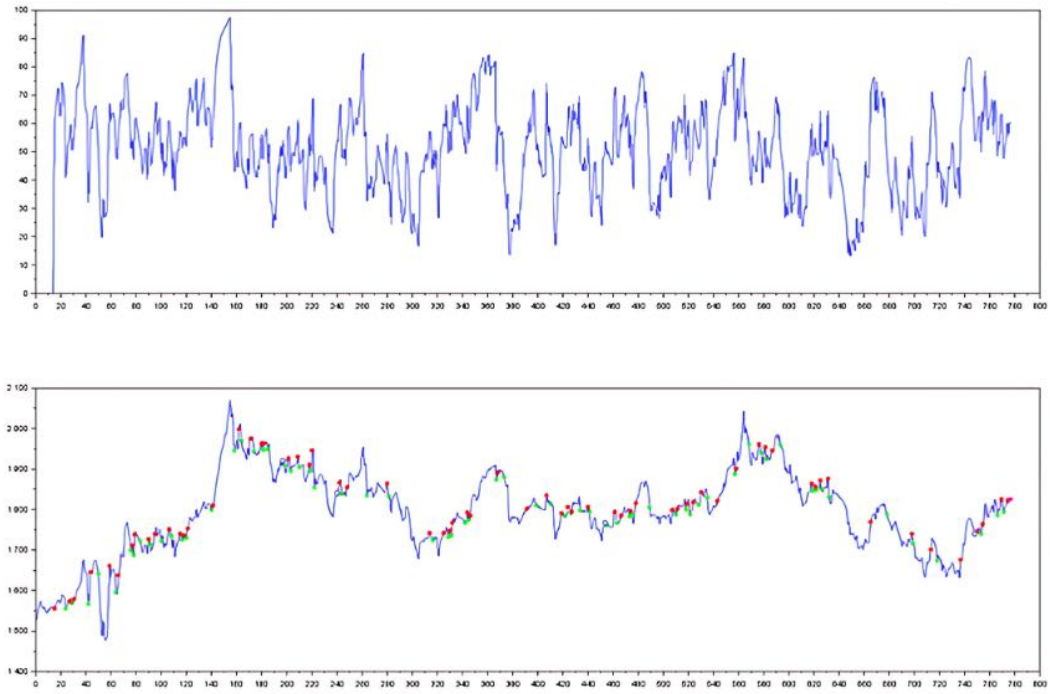


Figure 3. The price of IRSI-1 in the gold market from 2020 to 2022, along with appropriate buy and sell points. The green points indicate buy and sell points with red dots; the purple points indicate positions and times for closing trades and the positions of take profit or stop loss with transparent circles.

Table 8. The profit/loss results for the IRSI-1 strategy training period from 2020 to 2022.

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
15	1556.7		✓			0
24	1555.5	✓			✓	1.2
27	1573.4		✓	✓		17.9
29	1570.1	✓			✓	3.3

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
31	1578.8		✓	✓		8.7
42	1566.7	✓			✓	12.1
44	1644.4		✓	✓		77.7
50	1642.3	✓			✓	2.1
59	1660.8		✓	✓		18.5
64	1596.6	✓			✓	64.2
66	1637.7		✓	✓		41.1
76	1698.8	✓			✓	-61.1
77	1711.2		✓	✓		12.4
78	1687.8	✓			✓	23.4
79	1738.3		✓	✓		50.5
82	1723.8	✓			✓	14.5
90	1725.8		✓	✓		2
91	1713.9	✓			✓	11.9
95	1740.9		✓	✓		27
100	1721.9	✓			✓	19
106	1751.7		✓	✓		29.8
108	1734	✓			✓	17.7
115	1739.8		✓	✓		5.8
117	1727.2	✓			✓	12.6
118	1736.5		✓	✓		9.3
120	1731.1	✓			✓	5.4
121	1753		✓	✓		21.9
140	1800.3	✓			✓	-47.3
141	1810		✓	✓		9.7
158	1946.3	✓			✓	-136.3
162	1998.7		✓	✓		52.4
164	1970.3	✓			✓	28.4
171	1974.9		✓	✓		4.6
174	1944.7	✓			✓	30.2
180	1964.3		✓	✓		19.6
181	1947.9	✓			✓	16.4
182	1963.7		✓	✓		15.8
185	1949.9	✓			✓	13.8
201	1926.2		✓	✓		-23.7
203	1894.6	✓			✓	31.6

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
209	1929.5		✓	✓		34.9
210	1904.6	✓			✓	24.9
218	1910.4		✓	✓		5.8
219	1896.2	✓			✓	14.2
220	1946.8		✓	✓		50.6
222	1854.4	✓			✓	92.4
242	1866		✓	✓		11.6
244	1838.5	✓			✓	27.5
248	1855.3		✓	✓		16.8
264	1835.4	✓			✓	19.9
280	1863.9		✓	✓		28.5
281	1833.4	✓			✓	30.5
314	1741.7		✓	✓		-91.7
315	1738.1	✓			✓	3.6
325	1743		✓	✓		4.9
329	1732.7	✓			✓	10.3
330	1747.6		✓	✓		14.9
331	1736.3	✓			✓	11.3
332	1766.8		✓	✓		30.5
342	1768.3	✓			✓	-1.5
344	1791.8		✓	✓		23.5
345	1776	✓			✓	15.8
346	1784.3		✓	✓		8.3
367	1873.3	✓			✓	-89
368	1892		✓	✓		18.7
373	1879.6	✓			✓	12.4
391	1802.1		✓	✓		-77.5
392	1800.2	✓			✓	1.9
393	1810.6		✓	✓		10.4
399	1809.2	✓			✓	1.4
407	1835.8		✓	✓		26.6
410	1814.1	✓			✓	21.7
419	1789.8		✓	✓		-24.3
420	1787.8	✓			✓	2
424	1806.3		✓	✓		18.5
426	1791	✓			✓	15.3

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
428	1819.5		✓	✓		28.5
434	1798.5	✓			✓	21
440	1807.1		✓	✓		8.6
441	1794.8	✓			✓	12.3
461	1794.7		✓	✓		-0.1
463	1768.3	✓			✓	26.4
466	1784.9		✓	✓		16.6
473	1783.9	✓			✓	1
474	1795.8		✓	✓		11.9
475	1789.4	✓			✓	6.4
478	1816.8		✓	✓		27.4
489	1806.3	✓			✓	10.5
507	1798.2		✓	✓		-8.1
509	1794.6	✓			✓	3.6
511	1802.2		✓	✓		7.6
518	1800.1	✓			✓	2.1
519	1814.6		✓	✓		14.5
521	1789.2	✓			✓	25.4
524	1818.5		✓	✓		29.3
527	1816.5	✓			✓	2
530	1843.2		✓	✓		26.7
535	1829.7	✓			✓	13.5
543	1821.8		✓	✓		-7.9
557	1887.6	✓			✓	-65.8
558	1900.7		✓	✓		13.1
568	1960.8	✓			✓	-60.1
576	1962.2		✓	✓		1.4
577	1954.2	✓			✓	8
581	1954		✓	✓		-0.2
582	1923.7	✓			✓	30.3
588	1948.2		✓	✓		24.5
593	1959	✓			✓	-10.8
618	1865.4		✓	✓		-93.6
619	1846.3	✓			✓	19.1
621	1857.3		✓	✓		11
622	1848.4	✓			✓	8.9

Date	Price (USD per Troy Ounce)	Open		Close		Profit/Loss
		Buy Order	Sell Order	Buy Order	Sell Order	
625	1871.4		✓	✓		23
626	1850.2	✓			✓	21.2
631	1875.5		✓	✓		25.3
632	1831.8	✓			✓	43.7
665	1769.2		✓	✓		-62.6
678	1789.7	✓			✓	-20.5
698	1740.6		✓	✓		-49.1
699	1717.4	✓			✓	23.2
713	1702		✓	✓		-15.4
718	1675.2	✓			✓	26.8
737	1676.6		✓	✓		1.4
748	1739.6	✓			✓	-63
750	1745.6		✓	✓		6
753	1740.3	✓			✓	5.3
754	1763.7		✓	✓		23.4
766	1787.8	✓			✓	-24.1
769	1825.4		✓	✓		37.6
771	1795.3	✓			✓	30.1
774	1823.1		✓	✓		27.8
776	1826			✓		-2.9
Total						1011.9

From Table 8, the IRSI-3 strategy combines take-profit (tp), stop-loss (sl), and delay parameters. In this case, the optimal parameters were tp = 0, sl = 0.004, and delay = 0. The first trading signal was a buy on day 15 at a price of 1556.7. A sell signal appeared on day 24 at a price of 1555.5, resulting in a small loss of 1.2 due to the stop-loss being triggered. This pattern of frequent small gains and losses continued throughout the training period. For instance, a buy order was opened on day 44 at a price of 1566.7 and closed on day 1644.4 with a profit of 77.7. However, the tight stop-loss value also led to premature exits, such as the trade on day 314 that resulted in a loss of 91.7. Despite executing 132 trades during the training period—similar to IRSI-1—the total profit amounted to 1011.9. The zero value for both tp and delay indicates that the strategy exited trades very quickly based solely on minimal loss conditions. While this helped prevent significant losses, it also limited the ability to capture longer profitable trends.

Figure 4 shows the trading performance of the IRSI-2 strategy during the 2020 to 2022 training period. The chart presents buy points (green dots) and sell points (red dots) determined by the RSI signals, incorporating a delay parameter before executing trades. This delay mechanism helps filter out short-term noise, enabling more precise timing for entry and exit points.

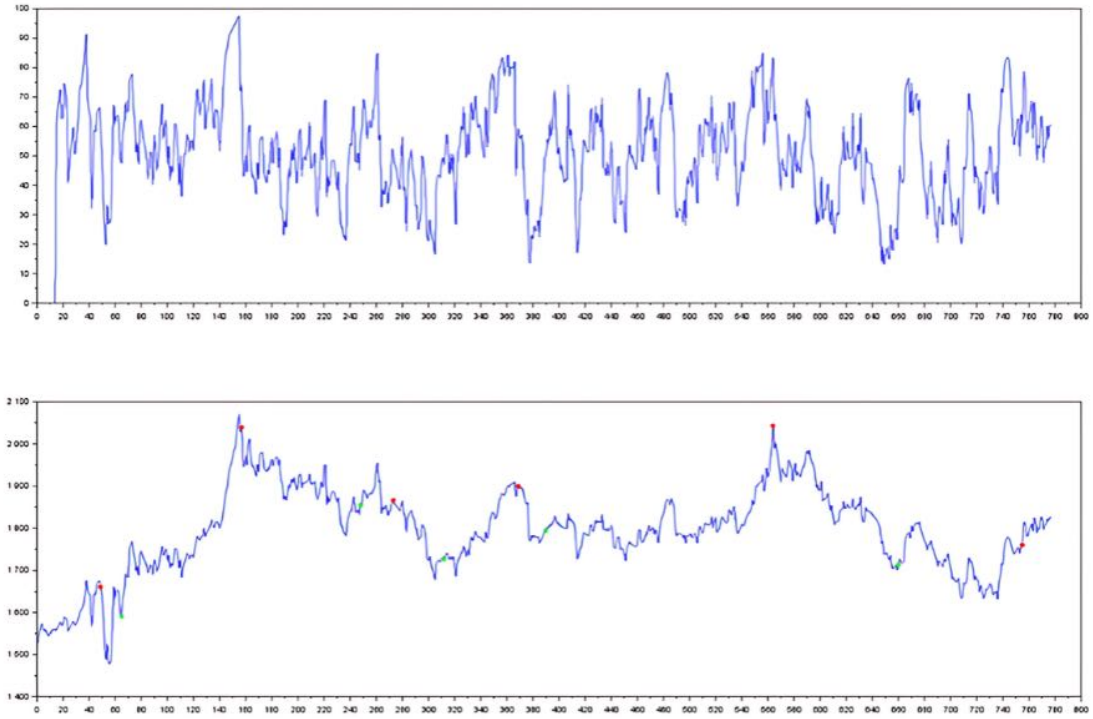


Figure 4. The price of IRSI-2 in the gold market from 2020 to 2022, along with appropriate buy and sell points. Green dots represent buy points and sell points with red dots.

Figure 5 demonstrates the performance of the IRSI-3 strategy during the 2020 to 2022 training period. The chart shows the buy (green dots) and sell (red dots) points based on the strategy's parameters, combining take-profit, stop-loss, and delay features. This visualization complements the summarized performance data discussed earlier.

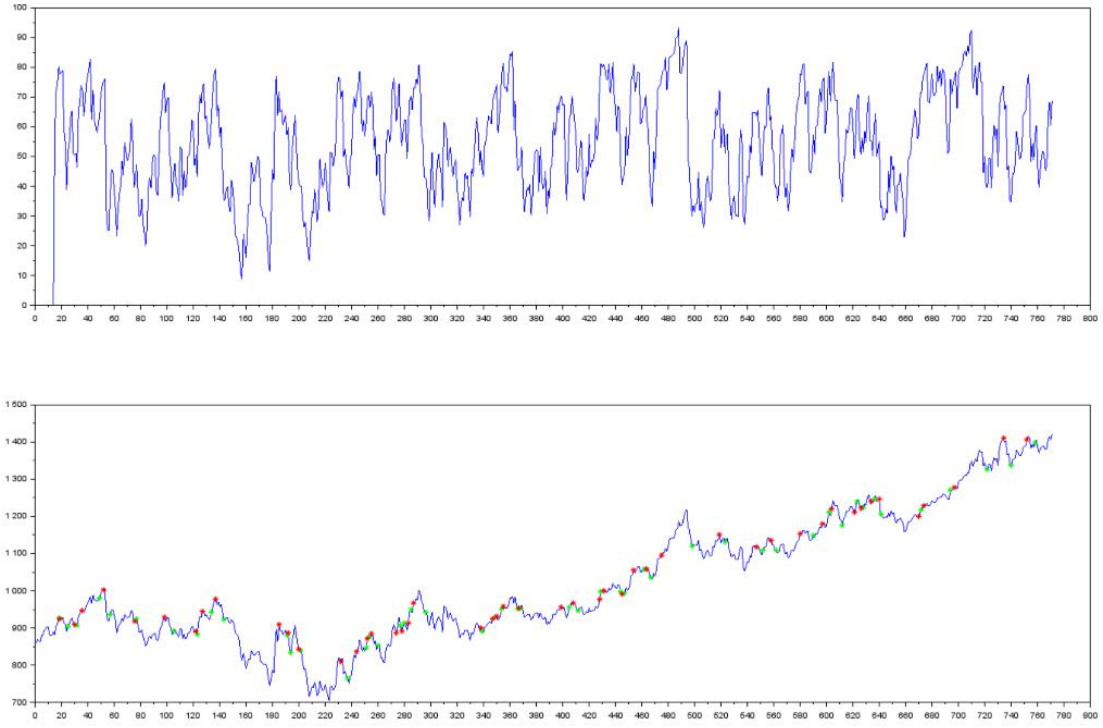


Figure 5. The price of IRSI-3 in the gold market from 2020 to 2022, along with appropriate buy and sell points. Green dots represent buy points and sell points with red dots.

4 Discussion

This study presents a method to improve the Relative Strength Index (RSI) with fixed returns and investment risk in the gold market. In the first part, we examined the profitability of four different trading strategies based on the traditional Relative Strength Index (TRSI) for investing in the gold market. According to the experimental results, the maximum profit of each strategy was TRSI with appropriate parameters, IRSI-1 with proper parameters, IRSI-2 with appropriate parameters, and IRSI-3 with appropriate parameters, with incomes of 1742.9, 1014.8, 1743.9, and 1011.9 respectively. The experimental results showed that TRSI and IRSI-2 yielded very similar returns for the training data, with profits of 1742.9 and 1743.9, respectively. During the testing period, the TRSI and IRSI-2 strategies also achieved similar profits of 169.7 and 152.4. This suggests that both strategies are comparably effective, especially when not constrained by take-profit and stop-loss conditions.IRSI-1 is a traditional RSI strategy that includes the addition of take-profit and stop-loss, which will control the risk to prevent excessive losses and achieve acceptable profits, IRSI-2 is a traditional strategy that adds a delay to the trading decision

process to prevent buying or selling slowly. IRSI-3 is a traditional RSI strategy and IRSI-3 is a strategy that combines IRSI-1 and IRSI-2. Since IRSI-1 and IRSI-3 strategies have take-profit and stop-loss into the conditions from the traditional model, it might not make much profit and not much loss because it's controlled. So, the TRSI and IRSI-2 are not controlled by take-profit and stop-loss, these two strategies have higher profits. For future research, it would be beneficial to extend the scope of these RSI-based strategies beyond the gold market. Specifically, future studies could explore the application of the proposed methods to other asset classes, such as stocks, cryptocurrencies, or foreign exchange markets, to test the robustness and adaptability of the models. Additionally, integrating machine learning techniques, such as classification models or reinforcement learning algorithms, could further enhance signal prediction accuracy and optimize decision-making processes. This hybrid approach may uncover deeper patterns in market behavior and potentially improve profitability and risk management.

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments.

References

[1] Yulius Hari, L. P. D. (2018). "Forecasting System Approach for Stock Trading with Relative Strength Index and Moving Average Indicator." Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 10: 25-29.

[2] T. T.-L. Chong and W.-K. Ng. (2014). “Revisiting the Performance of MACD and RSI Oscillators,” Journal Risk Financial Manag, vol. 7 no. 1, pp. 1-12.

[3] Ashok Kumar Panigrahi, K. V. a. S. K. C. (2021). " Trend identification with the relative strength index (RSI) technical indicator –A conceptual study." Journal of Management Research and Analysis 8(4): 159-169.

[4] Bartkus, C. (2018). "Using the Relative Strength Index for Active Investments in the Foreign Exchange Market." 23: 9-19.

[5] Phuong, L. C. M. (2021). "Investor sentiment by relative strength index and stock return: Empirical evidence on Vietnam's stock market " Journal Growing Science 7: 451-456.

[6] Cervelló-Royo, R., Guijarro, F., & Michniuk, K. (2018). Take profit and stop loss trading strategies comparison in combination with an MACD trading system. Journal of Risk and Financial Management, 11(3), 56.

[7] QuantInsti. (2023). Optimizing Exit Conditions Using a Variable Take Profit and Stop Loss. Retrieved from https://blog.quantinsti.com/optimizing-exit-conditions-project-sim-yee-kai

[8] Vezeris, D., Kyrgos, T., & Kourouklis, D. (2018). Take profit and stop loss trading strategies comparison in combination with an MACD trading system. Journal of Risk and Financial Management, 11(3), 56. https://doi.org/10.3390/jrfm11030056

[9] Cartea, Á., & Sánchez-Betancourt, M. (2022). Optimal execution with stochastic delay. Journal of Computational Finance, Springer. https://doi.org/10.1007/s00780-022-00491-w

[10] Moallemi, C. C., & Saglam, M. (2013). The cost of latency in high-frequency trading. Columbia Business School. Retrieved from https://business.columbia.edu/sites/default/files-efs/pubfiles/25476/Moallemi_latency.pdf

[11] Hsieh, C.-H., Barmish, B. R., & Gubner, J. A. (2019). The impact of execution delay on Kelly-based stock trading: High-frequency versus buy and hold. arXiv preprint arXiv:1907.08771. https://arxiv.org/abs/1907.08771

[12] Supama, R. (2024). An Improved Relative Strength Index Strategy with Fixed Return and Its Value at Risk of Investment in Forex and Gold Markets. Master thesis in Science, Graduate School, Khon Kaen University.

The Mixed Poisson Distribution with a New Two-Parameter Weighted Exponential Distribution for Overdispersed Count Data and Its Application in Motor Insurance

Jiramet Hengcharoensuk[1,*] and Adisak Moumeesri[1,†]

[1]Department of Statistics, Faculty of Science,
Silpakorn University, Nakhon Pathom 73000, Thailand

Abstract

Count data frequently exhibit overdispersion and characteristics of right-skewness, rendering the conventional Poisson distribution insufficient. This article proposes a new discrete probability distribution called the Poisson-new two parameter weighted exponential (Poisson-NTWE) distribution. This distribution integrates the Poisson distribution with the flexibility of the new two-parameter weighted exponential (NTWE) distribution, providing a better fit for overdispersed count data. The statistical properties of the proposed distribution, including the moment generating function, probability generating function, expected value, variance, skewness, and kurtosis, are determined. Parameter estimation is performed using maximum likelihood estimation (MLE). In addition, a Bayesian approach is employed to estimate an individual risk parameter using historical data (claim frequency records), which are subsequently used to determine fair premiums under the bonus-malus system (BMS). The proposed distribution is evaluated using actual insurance datasets, with comparative analyses against standard Poisson, mixed Poisson exponential (geometric), and negative binomial distributions. The findings indicate that the Poisson-NTWE distribution provides a superior model fit for overdispersed insurance claims data, offering improvements in premium calculation.

Keywords: Mixed Poisson Distribution, Motor Insurance, Overdispersion

2020 MSC: 60E05, 97M30, 97K50

1 Introduction

A probability distribution is a concept in statistical and mathematical theory that describes the characteristics of data. In many fields, probability distributions are applied to create models that assist in data analysis and decision-making for various purposes. For example,

*Speaker (hengcharoensuk_j@silpakorn.edu)
†Corresponding author (moumeesri_a@silpakorn.edu)

they are used to describe the characteristics of response variables when developing regression models, see [1-3]. In various industries, they are employed to describe or analyze the operation of systems, aiding in planning and controlling production quality, or even to create risk models in insurance and finance.

In the insurance and financial industries, probability distributions are used to describe the characteristics of potential risk events, such as the number of claims or the number of defaulters, etc. This information is then used to create risk models for further analysis and strategic planning, see [4]. Count models are employed to describe the characteristics of risk events or losses that may occur within a specified range, often modeled using the Poisson probability distribution when the data are discrete. However, in cases where the data exhibit overdispersion, with variance greater than the mean, the Poisson distribution's assumption that the mean equals the variance [5] is violated, making it an inappropriate model. This is particularly common in risk data, such as those in the financial and insurance, where the data often exhibit overdispersion and a right-skewed distribution [6-8] with excess zero counts and slow convergence toward zero (heavy-tailed). To address these issues, various techniques have been developed, such as the zero-inflated model, generalized family distribution, mixing distribution, transmuted family distribution, and other approaches.

In this article, a new distribution is introduced using a mixing distribution technique that combines the Poisson distribution and the new two parameters weighted exponential (NTWE) distribution, see [9], to model count data with a heavy right-tail. The distribution obtained by this method provides flexibility for modeling [10]. The properties of the distribution are investigated and tested for suitability using real-world data from industries such as insurance and biology. This proposed distribution will be used to model premium calculation under the bonus-malus system (BMS). A Bayesian approach is employed to derive the posterior distribution of the risk parameter (λ), where the prior distribution of λ is assumed to follow the NTWE distribution. The posterior distribution is then used to estimate the individual risk parameter $\hat{\lambda}$ under a squared error loss function, by computing the expected value of the posterior distribution based on driving history data from the insurance portfolio. The objective is to provide a fairer premium for the insured.

2 Materials and Methods

The weighted exponential (WE) distribution was introduced by Gupta and Kundu (2009) [11] as an extension of the standard exponential distribution, incorporating additional weighted parameters. The objective was to refine the distribution, making it more suitable for modeling reliability and risk-related data, particularly in finance, industrial engineering, and biology. Later, in 2017, Mahdavi and Jabbari [12] proposed the exponential weighted exponential (EWE) distribution, a three-parameter model designed to enhance flexibility and efficiency. Building on this concept, Chesneau et al. (2021) [9] developed a new distribution known as a new two-parameter weighted exponential (NTWE) distribution. This distribution includes two parameters: α (rate parameter) and β (shape parameter). The probability density function (PDF) of a random variable Y following the NTWE distribution is given as follows:

$$f(y;\alpha,\beta)=\frac{\alpha(\beta+1)^2}{\beta^2}e^{-\alpha y}\left(1-(1+\alpha\beta y)e^{-\alpha\beta y}\right) \tag{2.1}$$

where $\alpha,\beta>0$ and $y>0$.

2.1 Mixed Poisson with New Two-Parameter Weighted Exponential (Poisson-NTWE) Distribution

2.1.1 Mixing Distribution

The use of mixing distributions is common in modeling count data to address violations of the Poisson distribution's assumptions, especially when the mean and variance are not equal or when outliers exist. Over time, the complexity and diversity of real-world data have increased, making mixing models an attractive solution in various fields, including insurance, biomedicine, biology, and other disciplines where count data are commonly encountered (Cancho et al., 2011 [13], Shanker and Fesshaye, 2016 [14], Wongrin and Bodhisuwan, 2016 [15], Adetunji and Sabri, 2023 [16], Ieosanurak and Moumeesri, 2025 [17]).

Definition 2.1 Let X have a probability distribution function $f(x,\lambda)$ with a parameter λ. If the parameter λ takes different values, it is also considered a random variable, denoted as Λ. Therefore, the probability distribution function is given by the conditional probability function $f_{X|\Lambda}(x|\lambda)$, and the marginal probability distribution function of X is obtained by integrating the joint distribution function $f(x,\lambda)$ over all possible values of λ. The marginal probability distribution function of X is expressed as follows:

$$f_X(x)=\int_{-\infty}^{\infty}f(x,\lambda)d\lambda=\int_{-\infty}^{\infty}f_{X|\Lambda}(x|\lambda)\pi_\Lambda(\lambda)\,d\lambda. \tag{2.2}$$

Definition 2.2 Let X be a random variable following a Poisson distribution with parameter λ. Its probability mass function (PMF) is given by

$$f_{X|\Lambda}(x|\lambda)=\frac{e^{-\lambda}\lambda^x}{x!} \tag{2.3}$$

where $\lambda>0$ and $=0,1,2,\ldots$. The expected value and variance of X are λ.

We assume that the parameter λ is a random variable following an NTWE distribution [9] with parameters α and β. Its probability density function (PDF) is given by

$$\pi_\Lambda(\lambda)=\frac{\alpha(\beta+1)^2}{\beta^2}e^{-\alpha\lambda}\left(1-(1+\alpha\beta\lambda)e^{-\alpha\beta\lambda}\right) \tag{2.4}$$

where $\alpha,\beta>0$ and $\lambda>0$. The expected value of Λ is

$$E(\Lambda)=\frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{\alpha^2}-\frac{1}{(\alpha+\alpha\beta)^2}-\frac{2\alpha\beta}{(\alpha+\alpha\beta)^3}\right).$$

Proposition 2.3 *Let X be a random variable following a Poisson distribution with parameter λ, and let Λ be a random variable following an NTWE distribution. The unconditional probability mass function (PMF) of X, which is called a Poisson-NTWE distribution, is given by*

$$f_X(x) = \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha)^{x+1}} - \frac{1}{(1+\alpha+\alpha\beta)^{x+1}} - \frac{\alpha\beta(x+1)}{(1+\alpha+\alpha\beta)^{x+2}}\right) \tag{2.5}$$

where $\alpha, \beta > 0$ and $= 0,1,2, \ldots$.

Proof:

$$\begin{aligned}
f_X(x) &= \int_0^\infty f_{X|\Lambda}(x|\lambda)\pi_\Lambda(\lambda)\,\mathrm{d}\lambda \\
&= \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}\cdot\frac{\alpha(\beta+1)^2}{\beta^2}e^{-\alpha\lambda}\left(1-(1+\alpha\beta\lambda)e^{-\alpha\beta\lambda}\right)\mathrm{d}\lambda \\
&= \frac{\alpha(\beta+1)^2}{\beta^2 x!}\left(\int_0^\infty e^{-\lambda(1+\alpha)}\lambda^x\mathrm{d}\lambda - \int_0^\infty e^{-\lambda(1+\alpha+\alpha\beta)}\lambda^x\mathrm{d}\lambda - \alpha\beta\int_0^\infty e^{-\lambda(1+\alpha+\alpha\beta)}\lambda^{x+1}\mathrm{d}\lambda\right) \\
&= \frac{\alpha(\beta+1)^2}{\beta^2 x!}\left(\frac{\Gamma(x+1)}{(1+\alpha)^{x+1}} - \frac{\Gamma(x+1)}{(1+\alpha+\alpha\beta)^{x+1}} - \frac{\alpha\beta\Gamma(x+2)}{(1+\alpha+\alpha\beta)^{x+2}}\right) \\
&= \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha)^{x+1}} - \frac{1}{(1+\alpha+\alpha\beta)^{x+1}} - \frac{\alpha\beta(x+1)}{(1+\alpha+\alpha\beta)^{x+2}}\right).
\end{aligned}$$

□

The shape of the Poisson-NTWE distribution is shown in Figure 1, which illustrates that it can describe data with excess zeros and a right skew, often used for count data. When β increases while α remains fixed, the distribution becomes more concentrated at zero, making it heavily zero-inflated and reducing the probability of other values. In contrast, increasing α while keeping β constant spreads the distribution out, allowing for higher values while still maintaining its right-skewed shape. Essentially, α controls the degree of zero inflation, while β determines the overall spread of the distribution.

2.1.2 Statistical Properties of the Poisson-NTWE distribution

This subsection outlines various properties of the distribution, such as the probability generating function, moment generating function, and moments about the origin, along with the mean, variance, skewness, and kurtosis.

2.1.2.1 Probability Generating Function

The probability generating function (PGF) is an efficient tool for analyzing a discrete random variable X, where the variable t serves as a generating variable used to construct functions that describe the distributional properties of X. It simplifies the computation of moments, handling sums of independent random variables, and the analysis of complex systems, making it a valuable tool in both theoretical and applied contexts.

Proposition 2.4 Let X be a random variable with a Poisson-NTWE distribution. The PGF can be expressed as follows:

$$P_X(t) = \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha-t)} - \frac{1}{(1+\alpha+\alpha\beta-t)} - \frac{\alpha\beta}{(1+\alpha+\alpha\beta-t)^2}\right). \tag{2.6}$$

Proof:

$$P_X(t) = E(t^X) = E[E(t^X|\Lambda)]$$

$$= \int_0^\infty \sum_{x=0}^\infty t^x \frac{e^{-\lambda}\lambda^x}{x!} \cdot \pi_\Lambda(\lambda)\,\mathrm{d}\lambda$$

$$= \int_0^\infty e^{\lambda(t-1)} \cdot \frac{\alpha(\beta+1)^2}{\beta^2} e^{-\alpha\lambda}\left(1-(1+\alpha\beta\lambda)e^{-\alpha\beta\lambda}\right)\mathrm{d}\lambda$$

$$= \frac{\alpha(\beta+1)^2}{\beta^2}\int_0^\infty e^{-\lambda(1+\alpha-t)}\left(1-(1+\alpha\beta\lambda)e^{-\alpha\beta\lambda}\right)\mathrm{d}\lambda$$

$$= \frac{\alpha(\beta+1)^2}{\beta^2}\left(\int_0^\infty e^{-\lambda(1+\alpha-t)}\mathrm{d}\lambda - \int_0^\infty e^{-\lambda(1+\alpha+\alpha\beta-t)}\,\mathrm{d}\lambda - \alpha\beta\int_0^\infty e^{-\lambda(1+\alpha+\alpha\beta-t)}\lambda\,\mathrm{d}\lambda\right)$$

$$= \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha-t)} - \frac{1}{(1+\alpha+\alpha\beta-t)} - \frac{\alpha\beta}{(1+\alpha+\alpha\beta-t)^2}\right).$$

2.1.2.2 Moment Generating Function

The moment generating function (MGF) simplifies the calculation of moments (mean, variance, skewness, etc.) and it provides a specific characterization of the distribution. It plays an important role in the limit theorem and is an implement for comprehension the convergence of random variables. Moreover, the MGF is useful for transforming random variables and deriving compound distributions, making it essential for modeling more complex systems or processes.

Proposition 2.5 Let X be a random variable with a Poisson-NTWE distribution. Its MGF can be expressed as follows:

$$M_X(t) = \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha-e^t)} - \frac{1}{(1+\alpha+\alpha\beta-e^t)} - \frac{\alpha\beta}{(1+\alpha+\alpha\beta-e^t)^2}\right). \tag{2.7}$$

Proof:

$$M_X(t) = E(e^{tX}) = E[E(e^{tX}|\Lambda)]$$

$$= \int_0^\infty \sum_{x=0}^\infty e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} \cdot \pi_\Lambda(\lambda)\,\mathrm{d}\lambda$$

$$= \int_0^\infty e^{\lambda(e^t-1)} \cdot \frac{\alpha(\beta+1)^2}{\beta^2} e^{-\alpha\lambda}\left(1-(1+\alpha\beta\lambda)e^{-\alpha\beta\lambda}\right)\mathrm{d}\lambda$$

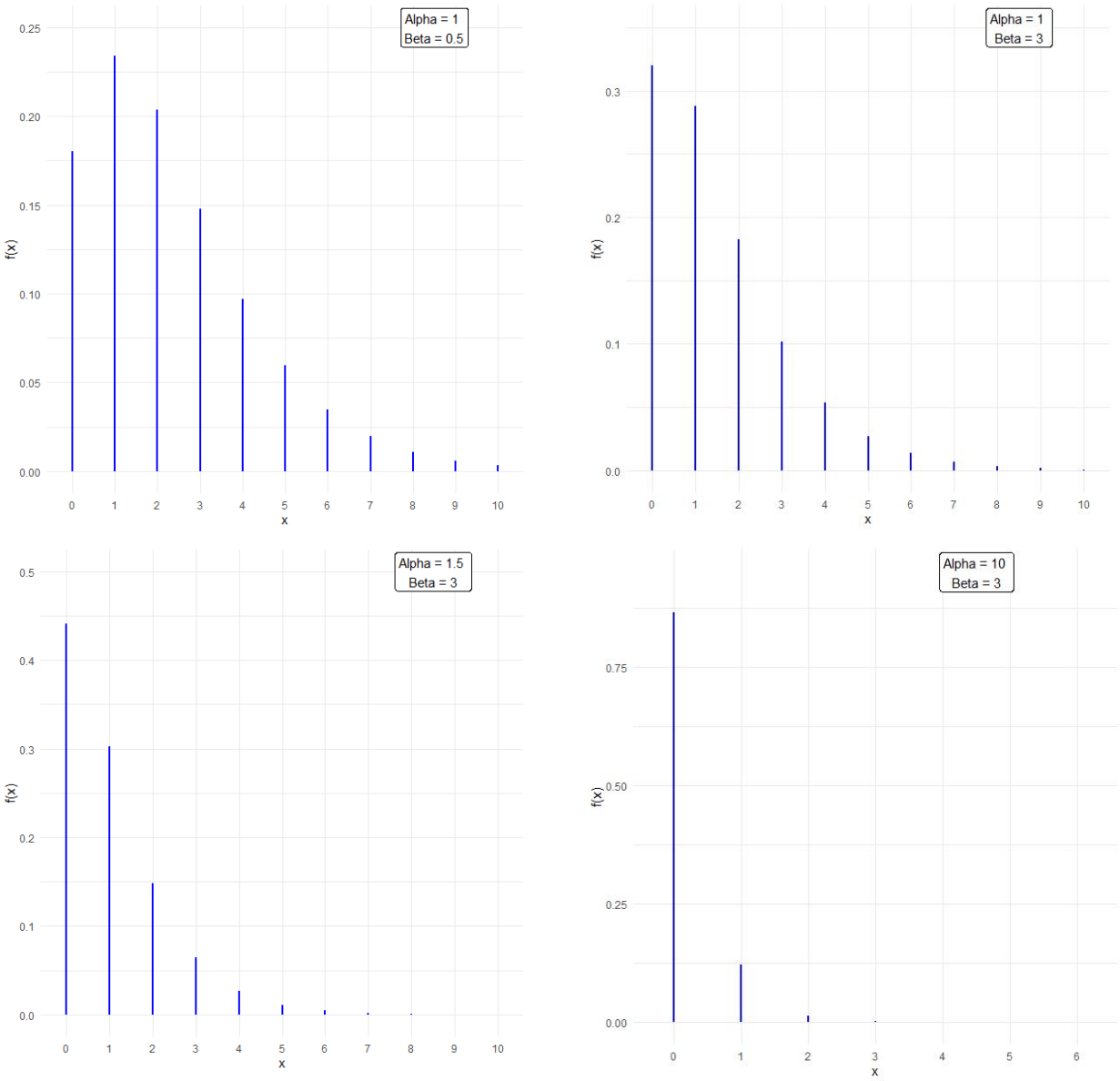


Figure 1: Shapes of the PMF of the Poisson-NTWE distribution

$$= \frac{\alpha(\beta+1)^2}{\beta^2}\int_0^\infty e^{-\lambda(1+\alpha-e^t)}\left(1-(1+\alpha\beta\lambda)e^{-\alpha\beta\lambda}\right)\mathrm{d}\lambda$$
$$= \frac{\alpha(\beta+1)^2}{\beta^2}\left(\int_0^\infty e^{-\lambda(1+\alpha-e^t)}\mathrm{d}\lambda - \int_0^\infty e^{-\lambda(1+\alpha+\alpha\beta-e^t)}\,\mathrm{d}\lambda - \alpha\beta\int_0^\infty e^{-\lambda(1+\alpha+\alpha\beta-e^t)}\lambda\,\mathrm{d}\lambda\right)$$
$$= \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha-e^t)} - \frac{1}{(1+\alpha+\alpha\beta-e^t)} - \frac{\alpha\beta}{(1+\alpha+\alpha\beta-e^t)^2}\right). \qquad \square$$

Using the relation $E(X^r) = \frac{\partial^r}{\partial t^r}M_X(t)|_{t=0}$, where $r = 1,2,\ldots$, the first four moments about the origin are given in the equations (2.8)-(2.11), respectively.

Talking $r = 1$, the first moment about the origin is shown as follows

$$M'_X(t)|_{t=0} = E(X) = E[E(X|\Lambda)]$$
$$= \frac{\partial}{\partial t}\left[\frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha-e^t)} - \frac{1}{(1+\alpha+\alpha\beta-e^t)} - \frac{\alpha\beta}{(1+\alpha+\alpha\beta-e^t)^2}\right)\right]$$
$$= \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{e^t}{(1+\alpha-e^t)^2} - \frac{e^t}{(1+\alpha+\alpha\beta-e^t)^2} - \frac{2\alpha\beta e^t}{(1+\alpha+\alpha\beta-e^t)^3}\right).$$

Then, at $t = 0$:

$$E(X) = \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{\alpha^2} - \frac{1}{(\alpha+\alpha\beta)^2} - \frac{2\alpha\beta}{(\alpha+\alpha\beta)^3}\right). \qquad (2.8)$$

Talking $r = 2$, The second moment about the origin is shown as follows:

$$M''_X(t)|_{t=0} = E(X^2) = E[E(X^2|\Lambda)]$$
$$= \frac{\alpha(\beta+1)^2}{\beta^2}\left[\frac{\partial}{\partial t^2}\left(\frac{1}{(1+\alpha-e^t)} - \frac{1}{(1+\alpha+\alpha\beta-e^t)} - \frac{\alpha\beta}{(1+\alpha+\alpha\beta-e^t)^2}\right)\right]$$
$$= \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{e^t}{(1+\alpha-e^t)^2} - \frac{e^t}{(1+\alpha+\alpha\beta-e^t)^2} + \frac{2e^t}{(1+\alpha-e^t)^3}\right.$$
$$\left. - \frac{2e^{2t}+2\alpha\beta e^t}{(1+\alpha+\alpha\beta-e^t)^3} - \frac{6\alpha\beta e^{2t}}{(1+\alpha+\alpha\beta-e^t)^4}\right).$$

Then, at $t = 0$:

$$E(X^2) = \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{\alpha^2} + \frac{2}{\alpha^3} - \frac{1}{(\alpha+\alpha\beta)^2} - \frac{2(\alpha\beta+1)}{(\alpha+\alpha\beta)^3} - \frac{6\alpha\beta}{(\alpha+\alpha\beta)^4}\right). \qquad (2.9)$$

Similarly, the third and fourth moments about origin can be found from the equation (2.10) and (2.11) by taking the third and fourth derivatives and substituting $t = 0$, respectively.

$$E(X^3) = \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{\alpha^2}+\frac{4}{\alpha^3}+\frac{6}{\alpha^4}-\frac{1}{(\alpha+\alpha\beta)^2}-\frac{2(\alpha\beta+3)}{(\alpha+\alpha\beta)^3}-\frac{6(2\alpha\beta+1)}{(\alpha+\alpha\beta)^4} - \frac{24\alpha\beta}{(\alpha+\alpha\beta)^5}\right) \tag{2.10}$$

$$E(X^4) = \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{\alpha^2}+\frac{10}{\alpha^3}+\frac{30}{\alpha^4}+\frac{24}{\alpha^5}-\frac{1}{(\alpha+\alpha\beta)^2}-\frac{2(\alpha\beta+5)}{(\alpha+\alpha\beta)^3}-\frac{6(5\alpha\beta+6)}{(\alpha+\alpha\beta)^4} - \frac{24(5\alpha\beta+1)}{(\alpha+\alpha\beta)^5}-\frac{120\alpha\beta}{(\alpha+\alpha\beta)^6}\right) \tag{2.11}$$

2.1.2.4 Mean and Variance

The mean of the random variable X with a Poisson-NTWE distribution is given by

$$\mu = E(X) = \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{\alpha^2}-\frac{1}{(\alpha+\alpha\beta)^2}-\frac{2\alpha\beta}{(\alpha+\alpha\beta)^3}\right). \tag{2.12}$$

Based on the relationship between the first and second moments about the origin, as given in the equation (2.8) and (2.9), the variance of $X \sim Poisson - NTWE(\alpha, \beta)$ is expressed as follows:

$$\begin{aligned} Var(X) &= E(X^2) - [E(X)]^2 \\ &= \frac{\alpha(\beta+1)^2}{\beta^2}\left[\left(\frac{1}{\alpha^2}+\frac{2}{\alpha^3}-\frac{1}{(\alpha+\alpha\beta)^2}-\frac{2(\alpha\beta+1)}{(\alpha+\alpha\beta)^3}-\frac{6\alpha\beta}{(\alpha+\alpha\beta)^4}\right) \right. \\ &\quad \left. - \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{\alpha^2}-\frac{1}{(\alpha+\alpha\beta)^2}-\frac{2\alpha\beta}{(\alpha+\alpha\beta)^3}\right)^2\right]. \end{aligned} \tag{2.13}$$

2.1.2.5 Skewness and Kurtosis

Skewness and kurtosis provide valuable insights into the shape of a distribution beyond what is captured by the mean and variance. In a data set, they are particularly useful for understanding the probability of extreme values or outliers. Based on the relationship between the four moments about the origin, as shown in equations (2.14) and (2.15), skewness and kurtosis can be determined as follows, respectively.

$$S_k(X) = \frac{E(X^3) - 3E(X^2)E(X) + 2[E(X)]^3}{[Var(X)]^{\frac{3}{2}}} \tag{2.14}$$

$$Kurt(X) = \frac{E(X^4) - 4E(X^3)E(X) + 6E(X^2)[E(X)]^2 - 3[E(X)]^4}{[Var(X)]^2} \tag{2.15}$$

2.1.3 Parameter Estimation

In the parameter estimation section, this article employs the maximum likelihood estimation (MLE) method, which selects parameter values that maximize the likelihood of observing the

given random sample $x_1, x_2, \ldots, x_n$. As the large sample size, the estimator obtained through this method exhibits desirable properties, including asymptotically unbiased, minimum variance, sufficiency, consistency, and the invariance property.

Let $x = (x_1, x_2, \ldots, x_n)$ be a random sample of size n from the Poisson-NTWE distribution, assumed to be independently and identically distributed (i.i.d.), with the PMF given in equation (2.5). The likelihood function is then expressed as

$$L(\alpha, \beta|x) = \prod_{i=1}^{n} f(x_i; \alpha, \beta) = \prod_{i=1}^{n} \frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha)^{x_i+1}} - \frac{1}{(1+\alpha+\alpha\beta)^{x_i+1}} - \frac{\alpha\beta(x_i+1)}{(1+\alpha+\alpha\beta)^{x_i+2}}\right).$$

Taking the natural logarithm of both sides, we obtain

$$\ln L(\alpha, \beta|x) = \sum_{i=1}^{n} \ln\left[\frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{(1+\alpha)^{x_i+1}} - \frac{1}{(1+\alpha+\alpha\beta)^{x_i+1}} - \frac{\alpha\beta(x_i+1)}{(1+\alpha+\alpha\beta)^{x_i+2}}\right)\right].$$

Since the maximum likelihood estimates do not have a closed-form solution, the parameters α and β are estimated numerically using the "optim" function in R.

2.2 Bayesian Method

Bayesian theory has been widely applied in insurance premium calculations, as premiums are derived from historical data, such as past claims and policyholder profiles [18-19]. This approach is fundamental in actuarial mathematics and is recognized as a reliable tool for premium determination under reliability theory. The Bayesian method is particularly important in estimating individual risk, as risk levels are inherently volatile and depend on available personal data. By incorporating Bayesian inference, the accuracy of risk assessment is enhanced, ensuring that premiums are appropriately adjusted to reflect policyholder-specific information. Consequently, this method contributes to a fairer and more consistent premium structure.

In this study, the insurance premiums are determined by the number of claims. It is assumed that the number of claims denoted as X follows the Poisson distribution with parameter λ which represents the average number of claims per a policyholder. To account for individual risk heterogeneity, λ is modeled as a random variable following the NTWE distribution, serving as the prior distribution. The Bayesian method is then applied to derive the posterior distribution of λ, and its expected value is used to estimate the premium charged to policyholders under observation.

2.2.1 Posterior Distribution for Claim Modeling

Consider the sample data $x = (x_1, x_2, \dots, x_t)$, where x_i represents the policyholder's claim frequency in year i. The specified period is denoted by t, and $N = \sum_{i=1}^{t} x_i$ represents the policyholder's total claim frequency during period t. The likelihood function is given by

$$L(\lambda; x_1, x_2, \dots, x_t) = f(x_1, x_2, \dots, x_t|\lambda) = \prod_{x=1}^{t} \frac{e^{-\lambda}\lambda^x}{x!} \propto e^{-\lambda t}\lambda^N.$$

The prior distribution is given by

$$\pi(\lambda) = \frac{\alpha(\beta+1)^2}{\beta^2} e^{-\alpha\lambda}\left(1 - (1+\alpha\beta\lambda)e^{-\alpha\beta\lambda}\right) \propto e^{-\alpha\lambda}\left(1 - (1+\alpha\beta\lambda)e^{-\alpha\beta\lambda}\right).$$

By Bayes' theorem, the posterior distribution for a policyholder, given a claim history $x_1, x_2, \dots, x_t$, can be determined. The posterior distribution is directly proportional to the likelihood function multiplied by the prior distribution, which can be expressed as follows:

$$\begin{aligned} g^*(\lambda|x_1, x_2, \dots, x_t) &\propto f(x_1, x_2, \dots, x_t|\lambda)\pi(\lambda) \\ &\propto e^{-\lambda t}\lambda^N e^{-\alpha\lambda} \cdot (1 - (1+\alpha\beta\lambda)e^{-\alpha\beta\lambda} \\ &\propto \left(\lambda^N e^{-\lambda(t+\alpha)}\right) - \left(\lambda^N e^{-\lambda(t+\alpha+\alpha\beta}\right) - \left(\alpha\beta\lambda^{N+1}e^{-\lambda(t+\alpha+\alpha\beta}\right). \end{aligned}$$

Given

$$\int_0^\infty g^*(\lambda|x_1, x_2, \dots, x_t)\,\mathrm{d}\lambda \propto \int_0^\infty \left(\lambda^N e^{-\lambda(t+\alpha)}\right) - \left(\lambda^N e^{-\lambda(t+\alpha+\alpha\beta}\right) - \left(\alpha\beta\lambda^{N+1}e^{-\lambda(t+\alpha+\alpha\beta}\right)\mathrm{d}\lambda.$$

This implies that

$$\int_0^\infty g^*(\lambda|x_1, x_2, \dots, x_t)\,\mathrm{d}\lambda \propto \int_0^\infty C\left[\left(\lambda^N e^{-\lambda(t+\alpha)}\right) - \left(\lambda^N e^{-\lambda(t+\alpha+\alpha\beta}\right) - \left(\alpha\beta\lambda^{N+1}e^{-\lambda(t+\alpha+\alpha\beta}\right)\right]\mathrm{d}\lambda = 1,$$

where C is a constant. We have

$$C = \frac{1}{\Gamma(N+1)\left(\frac{1}{(t+\alpha)^{N+1}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+1}} - \frac{\alpha\beta(N+1)}{(t+\alpha+\alpha\beta)^{N+2}}\right)}. \tag{2.16}$$

Therefore, the posterior distribution function for the Poisson-NTWE distribution claim frequency is represented as:

$$g^*(\lambda|x_1, x_2, \dots, x_t) = \frac{\left(\lambda^N e^{-\lambda(t+\alpha)}\right) - \left(\lambda^N e^{-\lambda(t+\alpha+\alpha\beta}\right) - \left(\alpha\beta\lambda^{N+1}e^{-\lambda(t+\alpha+\alpha\beta}\right)}{\Gamma(N+1)\left(\frac{1}{(t+\alpha)^{N+1}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+1}} - \frac{\alpha\beta(N+1)}{(t+\alpha+\alpha\beta)^{N+2}}\right)}. \tag{2.17}$$

2.2.2 Insurance Premium Calculation

Currently, various methods exist for calculating insurance premiums. In this article, the net premium calculation principle is used for premium calculation. The principle of premium determination is based on the Bayesian framework, beginning with the derivation of the posterior distribution of the risk parameter (λ), assuming that the prior distribution of λ follows the NTWE distribution. The risk parameter is then estimated using the Bayes estimator under a quadratic error loss function, $L(\lambda,\hat{\lambda}) = (\hat{\lambda}-\lambda)^2$, which measures the loss incurred by estimating the true value of λ with $\hat{\lambda}$. The Bayes estimator $\hat{\lambda}$ is chosen to minimize the expected loss $E[L(\lambda,\hat{\lambda})]$ under the squared error loss function, the Bayes estimator is equivalent to the expected value of the parameter λ under the posterior distribution. Net premiums represent the expected frequency of claims that policyholders are likely to claim. Let $\hat{\lambda}_{t+1}$ denote the average frequency of claims for policyholders with a claim history of $x_1, x_2, \dots, x_n$. This value is derived from the expected value of the posterior distribution function of the Poisson-NTWE distribution as follows:

$$\hat{\lambda}_{t+1} = E(\lambda|x_1,x_2,\dots,x_n) = \int_0^\infty \lambda \cdot g^*(\lambda|x_1,x_2,\dots,x_t)\,d\lambda$$

$$= \frac{1}{\Gamma(N+1)\left(\frac{1}{(t+\alpha)^{N+1}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+1}} - \frac{\alpha\beta(N+1)}{(t+\alpha+\alpha\beta)^{N+2}}\right)}\left(\int_0^\infty \lambda^{N+1}e^{-\lambda(t+\alpha)}\mathrm{d}\lambda - \int_0^\infty \lambda^{N+1}e^{-\lambda(t+\alpha+\alpha\beta)}\,\mathrm{d}\lambda - \int_0^\infty \alpha\beta\lambda^{N+2}e^{-\lambda(t+\alpha+\alpha\beta)}\,\mathrm{d}\lambda\right)$$

$$= \frac{1}{\Gamma(N+1)\left(\frac{1}{(t+\alpha)^{N+1}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+1}} - \frac{\alpha\beta(N+1)}{(t+\alpha+\alpha\beta)^{N+2}}\right)}\left(\frac{\Gamma(N+2)}{(t+\alpha)^{N+2}} - \frac{\Gamma(N+2)}{(t+\alpha+\alpha\beta)^{N+2}} - \frac{\alpha\beta\Gamma(N+3)}{(t+\alpha+\alpha\beta)^{N+3}}\right)$$

$$= \left(\frac{N+1}{\left(\frac{1}{(t+\alpha)^{N+1}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+1}} - \frac{\alpha\beta(N+1)}{(t+\alpha+\alpha\beta)^{N+2}}\right)}\right)\left(\frac{1}{(t+\alpha)^{N+2}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+2}} - \frac{\alpha\beta(N+2)}{(t+\alpha+\alpha\beta)^{N+3}}\right).$$

Assuming that at $t = 0$, the initial premium payable is 100, the premium for the next period, $t+1$, can be calculated using the following formula:

$$\text{Premium}_{t+1} = \frac{\hat{\lambda}_{t+1}}{\hat{\lambda}_0} \times 100\% = \frac{E(\lambda|x_1,x_2,\dots,x_n)}{E(\lambda)} \times 100\%$$

$$\text{Premium}_{t+1} = \frac{\left(\frac{N+1}{\left(\frac{1}{(t+\alpha)^{N+1}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+1}} - \frac{\alpha\beta(N+1)}{(t+\alpha+\alpha\beta)^{N+2}}\right)}\right)\left(\frac{1}{(t+\alpha)^{N+2}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+2}} - \frac{\alpha\beta(\mathrm{N}+2)}{(t+\alpha+\alpha\beta)^{N+3}}\right)}{\left(\frac{1}{\left(\frac{1}{\alpha} - \frac{1}{(\alpha+\alpha\beta)} - \frac{\alpha\beta}{(\alpha+\alpha\beta)^2}\right)}\right)\left(\frac{1}{\alpha^2} - \frac{1}{(\alpha+\alpha\beta)^2} - \frac{2\alpha\beta}{(\alpha+\alpha\beta)^3}\right)} \times 100\%$$

$$= \frac{\left(\frac{N+1}{\left(\frac{1}{(t+\alpha)^{N+1}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+1}} - \frac{\alpha\beta(N+1)}{(t+\alpha+\alpha\beta)^{N+2}}\right)}\right)\left(\frac{1}{(t+\alpha)^{N+2}} - \frac{1}{(t+\alpha+\alpha\beta)^{N+2}} - \frac{\alpha\beta(N+2)}{(t+\alpha+\alpha\beta)^{N+3}}\right)}{\frac{\alpha(\beta+1)^2}{\beta^2}\left(\frac{1}{\alpha^2} - \frac{1}{(\alpha+\alpha\beta)^2} - \frac{2\alpha\beta}{(\alpha+\alpha\beta)^3}\right)} \times 100\% \quad (2.18)$$

where $\alpha > 0, \beta > 0$, and $N = \sum_{i=1}^{t} x_i$.

3 Numerical Application

After formulating the Poisson-NTWE distribution and analyzing its properties, the next step is to estimate its parameters numerically and validate it using real data. Furthermore, the distribution will be applied to the calculation of insurance premiums within a bonus-malus system for motor insurance.

3.1 Datasets

The proposed new distribution is evaluated for its performance on three different count datasets. The first dataset consists of 67,856 observations on the number of claims reported by an Australian company during 2004–2005 [20]. The second dataset contains the number of claims filed by policyholders at a Turkish insurance company between 2012 and 2014 [21]. The third dataset records the number of epileptic seizures observed [22]. Additionally, descriptive statistics for each dataset, including the percentage of zeros, dispersion index, skewness, and kurtosis, are presented in Table 1.

Table 1: Summary of observations and descriptive statistics for datasets

Observation	Dataset 1	Dataset 2	Dataset 3
0	63232	8544	126
1	4333	1796	80
2	271	370	59
3	18	81	42
4	2	23	24
5			8
6			5
7			4
8			3

Observation	Dataset 1	Dataset 2	Dataset 3
Percentage of zeros	93.19	79.01	35.90
Dispersion index	1.06	1.26	1.87
Skewness	4.07	2.56	1.31
Kurtosis	21.50	10.71	4.71

3.2 Model Fitting

In this section, the Poisson-NTWE distribution, with parameters estimated using the MLE, will be evaluated for its effectiveness in describing the characteristics of real data. A statistical method will be employed to assess the obtained distribution of count data using the chi-square goodness of fit test. Additionally, we compare the effectiveness of the proposed distribution with baseline distributions, including the Poisson distribution, negative binomial (NB), and mixed Poisson exponential (MPE), as shown in Tables 2–4.

The chi-square goodness of fit test determines how well an observed distribution matches an expected one. It compares observed frequencies with expected frequencies predicted by a distribution. This statistical method helps assess if a distribution fits the data, especially for count data consisting of discrete random variables. In this study, we apply the test to evaluate if the proposed distribution captures the data's characteristics.

Testing the hypothesis of the chi-square test:

H_0: The data follows a specified distribution.

H_1: The data does not follow a specified distribution.

The chi-square test is used to compare the observed frequencies (O_i) with the expected frequencies (E_i) in each group to assess the fit of the data with the specified distribution. The test statistic can be calculated using the following equation:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{(k-p-1)} \quad Null\ hypothesis\ true$$

where k is the number of groups or levels of a variable, and p is the number of estimated parameters.

Tables 2, 3, and 4 assess the appropriateness of the distributions for the three datasets using the chi-square test statistic and compare them with the Poisson, NB, and MPE distributions. Additionally, model selection criteria, including the Akaike Information Criterion (AIC) [15] and Bayesian Information Criterion (BIC), are presented.

Table 2: Comparison of observed and expected frequencies for dataset 1

Number of Claims	Observed frequency	Expected frequency			
		Poisson	NB	MPE	Poisson-NTWE
0	63232	69091.61	63232.52	63255.39	63232.61
1	4333	4593.07	4329.00	4288.69	4330.02
2	271	167.19	276.17	290.77	274.84
3	18	4.06	17.19	19.71	17.36
4	2	0.07	1.13	1.43	1.17
Estimated parameters by MLE		$\hat{\lambda} = 0.0728$	$\hat{\alpha} = 1.1582$ $\hat{\beta} = 0.0591$	$\hat{\alpha} = 13.7494$	$\hat{\alpha} = 14.8321$ $\hat{\beta} = 24.3034$
χ^2		177.66	0.26	1.87	**0.17**
Df.		2	1	2	1
p-value		<0.01	0.61	0.39	**0.68**
AIC		36205	36103.36	36102.89	36103.24
BIC		36214.13	36121.61	36112.02	36121.49

Table 3: Comparison of observed and expected frequencies for dataset 2

Number of Claims	Observed frequency	Expected frequency			
		Poisson	NB	MPE	Poisson-NTWE
0	8544	8292.25	8543.00	8545.15	8543.71
1	1796	2201.76	1795.99	1792.83	1795.06
2	370	292.31	375.79	376.15	375.75
3	81	25.87	78.51	78.92	78.65
4	23	1.81	20.71	20.95	20.82
Estimated parameters by MLE		$\hat{\lambda} = 0.266$	$\hat{\alpha} = 1.0095$ $\hat{\beta} = 0.2083$	$\hat{\alpha} = 3.7663$	$\hat{\alpha} = 3.7773$ $\hat{\beta} = 679.7360$
χ^2		468.19	0.42	0.36	0.39
Df.		3	2	3	2
p-value		<0.01	0.81	0.95	0.82
AIC		14308.32	14063.43	14061.44	14063.43
BIC		14315.61	14078.01	14068.73	14078.01

Table 4: Comparison of observed and expected frequencies for dataset 3

Number of epileptic seizures	Observed frequency	Expected frequency			
		Poisson	NB	MPE	Poisson-NTWE
0	126	74.93	120.23	137.95	119.42
1	80	115.71	93.00	83.73	95.82
2	59	89.34	59.17	50.82	58.87
3	42	45.99	34.94	30.85	33.76
4	24	17.75	19.83	18.72	19.00
5	8	5.46	10.99	11.37	10.64
6	5	1.41	5.98	6.90	5.95
7	4	0.31	3.22	4.19	3.33
8	3	0.06	1.72	2.47	1.86
Estimated Parameters by MLE		$\hat{\lambda} = 1.5442$	$\hat{\alpha} = 1.5503$ $\hat{\beta} = 0.4990$	$\hat{\alpha} = 0.6475$	$\hat{\alpha} = 0.7892$ $\hat{\beta} = 8.1411$
χ^2		80.91	**5.37**	10.82	7.14
Df.		4	5	6	5
p-value		<0.01	**0.37**	0.09	0.21
AIC		1274.09	**1193.88**	1198.79	1196.21
BIC		1277.95	**1201.61**	1202.65	1203.93

From Tables 2, 3, and 4, the results indicate that the Poisson distribution has a p-value less than 0.05, leading to the rejection of the null hypothesis. Therefore, the Poisson distribution is not suitable for any of the three datasets. In contrast, the NB, MPE, and Poisson-NTWE distributions have p-values greater than 0.05, indicating that the null hypothesis cannot be rejected. Thus, these three distributions are considered appropriate for all three count datasets exhibiting overdispersion at a significance level of 0.05.

The results indicate that the Poisson-NTWE distribution is well-suited for count data exhibiting overdispersion, particularly for dataset 1 and dataset 2, which consist of insurance data. Additionally, it provides the best fit for heavy-tailed data (strong right skewness) or data that converges slowly to zero, both of which are typical characteristics of insurance data. As shown in Table 1, dataset 1 exhibits the greatest skewness among the three datasets. The Poisson-NTWE distribution best captures the characteristics of dataset 1, as evidenced by the lowest chi-square test statistic. Therefore, in the next subsection, this distribution is applied to calculate premiums under the bonus-malus system.

3.3 Insurance Premium Pricing

In this section, the calculation of premiums under the bonus-malus system is based on an equation (2.18). The premium rates presented in Table 5 are derived using the Poisson-NTWE distribution.

Table 5 presents premium rates calculated using the Poisson-NTWE distribution, structured based on the time (t) and the number of claims. The first column represents t, which indicates the underwriting period or renewal year. The remaining columns display the premium rates corresponding to different claim counts, ranging from 0 to 4.

Observing Table 5, the premium for policyholders with zero claims decreases over time, starting at 100.00 and gradually reducing to 70.15 by t=7. Setting the initial premium to 100 implies that all insured individuals begin by paying the full premium rate specified by the insurance company. For instance, based on Table 5, if the average initial premium is set at 20,000 baht at the beginning of the policy year ($t = 0$), every policyholder must pay 100% of this amount, i.e., the full 20,000 baht. In the first year, if the policyholders do not file any claims, they receive a premium discount of 5.87% based on the initial premium, reducing the payable premium to 18,826 baht. If there are no claims in the second year, the premium is further reduced by 11.05% of the initial premium, resulting in a payment of 17,790 baht. This pattern continues, and by the 7th year, if the policyholders have made no claims throughout, the total discount reaches 29.85% of the initial premium, lowering the premium to 14,030 baht.

This represents a form of discount or experience-based rating for policyholders with no claims. Conversely, for those who have filed claims, premiums increase proportionally with the number of claims and may decrease in subsequent years if no additional claims are made. For example, if the policyholders file one claim in the first year, the premium will increase by 74.49% relative to the initial premium. If another claim is filed in the second year, the premium will rise to 145.07% of the initial premium. However, if no claim is made in the third year, a discount is applied, reducing the premium to 131.34% of the initial premium.

This premium adjustment model illustrates the mechanism of the bonus-malus system, which adjusts premiums through decreases or increases based on claim history. This system incentivizes safe driving behavior by offering progressive premium reductions to low-risk policyholders with no claims.

For comparison purposes, we calculate the premium rate for the MPE distribution, as shown in Table 6. Both model results in Table 5 and Table 6 exhibit similar trends in premium rates, but the MPE model in Table 6 is stricter in penalizing policyholders when they file a claim and offers slightly higher discounts for claim-free policyholders. On the other hand, the Poisson-NTWE model in Table 5 applies more gradual premium adjustments, which may make it more suitable for policyholder retention strategies.

Table 5: Premium rate based on the Poisson-NTWE distribution

t	Number of Claims				
	0	1	2	3	4
0	100.00				
1	94.13	174.49	260.54	347.29	434.11
2	88.95	164.22	245.07	326.66	408.32
3	84.35	155.11	231.34	308.34	385.42
4	80.24	146.96	219.07	291.97	364.95
5	76.54	139.64	208.04	277.25	346.55
6	73.19	133.03	198.06	263.94	329.92
7	70.15	127.03	189.01	251.85	314.80

Table 6: Premium rate based on the MPE distribution

t	Number of Claims				
	0	1	2	3	4
0	100.00				
1	93.22	186.44	279.66	372.88	466.10
2	87.30	174.60	261.90	349.20	436.51
3	82.09	164.18	246.27	328.36	410.44
4	77.46	154.93	232.39	309.86	387.32
5	73.33	146.66	220.00	293.33	366.66
6	69.62	139.24	208.86	278.48	348.10
7	66.26	132.53	198.79	265.06	331.32

4 Discussion

In this study, a new distribution called Poisson-NTWE is proposed as an efficient modeling approach for overdispersed count data that cannot be adequately described by the standard Poisson distribution. The results indicate that mixed Poisson distributions, including the Poisson-Exponential distribution, the Negative Binomial distribution, and the proposed Poisson-NTWE distribution, effectively capture the characteristics of data that deviate from the assumptions of the Poisson distribution. The Chi-square goodness of fit test, along with AIC and BIC results presented in Tables 2–4, confirm that the Poisson-NTWE distribution provides a good fit for all three overdispersed datasets at a significance level of 0.05.

The Poisson-NTWE distribution demonstrates superior performance in modeling insurance data, particularly for datasets exhibiting heavy-tailed and substantial right-skewness dispersion, which are common in reliability and risk-related data. This is evident in dataset 1, as shown in

Table 2, where the Poisson-NTWE distribution yields the lowest Chi-square test statistic compared to other distributions. This result indicates that the Poisson-NTWE distribution is suitable for capturing the characteristics of insurance data. However, the Poisson-NTWE distribution may have limitations when applied to dispersion datasets with low skewness or a small number of repeated zeros. In such cases, its performance may not differ from other mixed Poisson distributions, as observed in Table 3. Specifically, the mixed Poisson exponential distribution proves to be effective in modeling data with substantial dispersion when the data does not exhibit strong skewness. Building on the strengths of the proposed distribution, this study applies it to calculate insurance premiums under the BMS. The Bayesian method is utilized to determine premiums that align with each policyholder's driving history, ensuring a fair pricing structure. The premium payable to the insurance company is determined based on the number of claims made and the duration of the policy.

5 Conclusion

This study introduces the Poisson-new two parameter weighted exponential (Poisson-NTWE) distribution as an effective mixed distribution for overdispersed count data that the Poisson distribution cannot be adequately described. The Chi-square goodness of fit test confirmed that the Poisson-NTWE distribution provides a better fit for insurance datasets exhibiting overdispersion and heavy-tailed distributions.

The Poisson-NTWE distribution was effective for modeling claim frequency data in motor insurance, particularly for datasets with right-skewness. Furthermore, this study applied the Bayesian method to improve insurance premium calculations under the bonus-malus system (BMS), ensuring a fairer pricing structure based on policyholder risk.

Future research can extend this work by incorporating claim severity modeling into the BMS premium calculation framework. This enhancement would allow for a more comprehensive assessment of policyholder risk, ensuring that premium adjustments account not only for claim frequency but also for variations in claim severity.

Acknowledgment. The authors would like to thank the editor and referees for their reviews and valuable suggestions, which greatly improved the quality of this paper. This research received financial support from Faculty of Science, Silpakorn University, grant number SRIF-PRG-2567-10.

References

[1] Castillo, E., & Fernández-Canteli, A. (2001). A general regression model for lifetime evaluation and prediction. *International Journal of Fracture*, *107*, 117-137.

[2] Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment*, *91*(2), 121-136.

[3] Ismail, N., & Jemain, A. A. (2005). Generalized Poisson regression: An alternative for risk classification. *Jurnal Teknologi (Sciences & Engineering)*, 39-54.

[4] Bening, V. & Korolev, V. (2002). Generalized Poisson Models and their Applications in Insurance and Finance. Berlin, Boston: De Gruyter.

[5] Gschlößl, S., & Czado, C. (2008). Modelling count data with overdispersion and spatial effects. *Statistical papers*, *49*, 531-552.

[6] Yip, K. C., & Yau, K. K. (2005). On modeling claim frequency data in general insurance with extra zeros. *Insurance: Mathematics and Economics*, *36*(2), 153-163.

[7] Omari, C.O., Nyambura, S.G. & Mwangi, J.M.W. (2018). Modeling the Frequency and Severity of Auto Insurance Claims Using Statistical Distributions. *Journal of Mathematical Finance, 8*, 137-160.

[8] Sabri, S. R. M., & Adetunji, A. A. (2024). On the Poisson-transmuted exponential distribution and its application to frequency of claim in actuarial science. *Statistics in Transition new series*, *25*(2), 103-120.

[9] Chesneau, C., Tomy, L., & Jose, M. (2024). On a New Two-parameter Weighted Exponential Distribution and Its Application to the COVID-19 and Censored Data. *Thailand Statistician*, *22*(1), 17-30.

[10] Karlis, D., & Xekalaki, E. (2005). Mixed poisson distributions. *International Statistical Review/Revue Internationale de Statistique*, 35-58.

[11] Gupta, R. D., & Kundu, D. (2009). A new class of weighted exponential distributions. *Statistics*, *43*(6), 621-634.

[12] Mahdavi, A., & Jabbari, L. (2017). An extended weighted exponential distribution. *Journal of Modern Applied Statistical Methods*, *16*(1), 296-307.

[13] Cancho, V. G., Louzada-Neto, F., & Barriga, G. D. (2011). The Poisson-exponential lifetime distribution. *Computational Statistics & Data Analysis*, *55*(1), 677-686.

[14] Shanker, R., & Fesshaye, H. (2016). On Poisson-Sujatha distribution and its applications to model count data from biological sciences. *Biometrics & Biostatistics International Journal*, *3*(4), 1-7.

[15] Wongrin, W., & Bodhisuwan, W. (2016). The Poisson-generalised Lindley distribution and its applications. *Songklanakarin Journal of Science & Technology*, *38*(6), 645-656.

[16] Adetunji, A. A., & Sabri, S. R. M. (2023). On the Poisson transmuted Ailamujia distribution with applications to dispersed and skewed count data. *Journal of Statistics & Management Systems*, *26*(4), 929-943.

[17] Ieosanurak, W., & Moumeesri, A. (2025). An Adaptive Quadratic Transmuted Exponential Distribution: Unbounded Weight Parameter and Mixed Poisson Model Integrations. *American Journal of Mathematical and Management Sciences*, 1-22.

[18] Hernawati, L., Purnaba, I. G. P., & Mangku, I. W. (2017). Comparing bonus malus premiums of two types claim which is assessed using Bayesian method. *Applied Mathematical Sciences*, *11*(34), 1687-1694.

[19] Moumeesri, A., & Pongsart, T. (2022). Bonus-malus premiums based on claim frequency and the size of claims. *Risks*, *10*(9), 181.

[20] De Jong, P., & Heller, G. Z. (2008). *Generalized linear models for insurance data*. Cambridge University Press.

[21] Sarul, L. S., & Sahin, S. (2015). An application of claim frequency data using zero inflated and hurdle models in general insurance. *Journal of Business Economics and Finance*, *4*(4), 732-743.

[22] Chakraborty, S. (2010). On some distributional properties of the family of weighted generalized Poisson distribution. *Communications in Statistics—Theory and Methods*, *39*(15), 2767-2788.

[23] Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716-723.

การวิเคราะห์การถดถอยเชิงเส้นพหุคูณราคาปาล์มน้ำมันประเทศไทยปี พ.ศ. 2558 – 2566

ทรงกรด มัคอาน[1,*] อัลฟาตีฮะห์ แมเราะ[1] และ ปิยธิดา บุญสนอง[1,†]

[1]สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสงขลา 90000

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) สร้างสมการพยากรณ์ราคาปาล์มน้ำมัน 2) ศึกษาปัจจัยที่ส่งผลต่อราคาปาล์มน้ำมัน ใช้ข้อมูลรายเดือนตั้งแต่มกราคม 2558 ถึงธันวาคม 2566 รวม 108 เดือน ใช้การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ เพื่อประมาณค่าสัมประสิทธิ์ของตัวแปรอิสระกับตัวแปรตาม ด้วยวิธีกำลังสองน้อยที่สุด มีตัวแปรทั้งหมด 6 ตัว ได้แก่ ราคาปาล์มน้ำมัน ราคาน้ำมันปาล์มดิบ ปริมาณส่งออกน้ำมันปาล์ม ปริมาณการขายผลผลิตปาล์มน้ำมัน ปริมาณส่งออกสินค้าเกษตร และอัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา และเลือกสมการที่ดีที่สุดจาก 4 วิธี คือ วิธีพิจารณาทุกรูปแบบที่เป็นไปได้ วิธีเพิ่มตัวแปรอิสระ วิธีลดตัวแปรอิสระ และวิธีเพิ่มตัวแปรอิสระแบบขั้นตอน

ผลการวิจัยผู้วิจัยเลือกวิธีเพิ่มตัวแปรและวิธีเพิ่มตัวแปรอิสระแบบขั้นตอน ที่ให้ผลลัพธ์เหมือนกัน คือ ปัจจัยที่ส่งผลต่อราคาปาล์มน้ำมัน (Y) คือ ราคาน้ำมันปาล์มดิบ (X1) กับปริมาณส่งออกสินค้าเกษตร (X4) มีความสัมพันธ์กับราคาปาล์มน้ำมันไปในทิศทางเชิงบวก และปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) มีความสัมพันธ์กับราคาปาล์มน้ำมันในทิศทางตรงกันข้าม ที่ระดับนัยสำคัญทางสถิติ 0.05 มีตัวแบบพยากรณ์ คือ

$$\hat{\sqrt{Y}} = 0.914+0.042X1-0.002X3+0.002X4$$

ซึ่งค่า R^2 เท่ากับ 0.952 กล่าวคือ ปัจจัยที่เกี่ยวข้องนี้มีความสามารถในการอธิบายความผันผวนของราคาปาล์มน้ำมันได้ร้อยละ 95.2

คำสำคัญ: ราคาปาล์มน้ำมันในประเทศไทย, การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

2020 MSC: 62J05, 62P20

*ผู้นำเสนอ (644254012@parichat.skru.ac.th)

†ผู้ประพันธ์บรรณกิจ (piyatida.bo@skru.ac.th)

1 บทนำ

ปาล์มน้ำมันถือเป็นพืชเศรษฐกิจที่สำคัญของประเทศไทย โดยเริ่มปลูกในประเทศตั้งแต่ปี พ.ศ. 2472 ที่สถานีทดลองยางคอหงส์ในจังหวัดสงขลา และสถานีกสิกรรมพลิ้วในจังหวัดจันทบุรีในฐานะพืชประดับ ปาล์มน้ำมันเป็นพืชที่มีศักยภาพในการใช้ประโยชน์อย่างหลากหลาย ทั้งการแปรรูปโดยตรงและการนำไปเป็นส่วนผสมในผลิตภัณฑ์ต่างๆ นอกจากนี้ยังเป็นพืชน้ำมันที่มีศักยภาพในการแข่งขันสูงกว่าพืชน้ำมันชนิดอื่น เนื่องจากให้ผลผลิตสูงกว่าพืชน้ำมันอื่น ๆ มีต้นทุนการผลิตและราคาต่ำ จึงทำให้ปาล์มน้ำมันมีส่วนแบ่งการผลิตที่สูงในอุตสาหกรรมน้ำมัน ส่งผลให้เกษตรกรมีการเปลี่ยนแปลงจากการผลิตเพื่อยังชีพไปสู่การผลิตเพื่อการค้าที่เน้นปริมาณ คุณภาพและการผลิตมากขึ้น [1] และอุตสาหกรรมน้ำมันปาล์มมีแนวโน้มขยายตัว โดยในช่วงปี 2568-2569 อุปทานมีทิศทางเพิ่มขึ้นโดยได้ปัจจัยหนุนจาก (1) ผลผลิตต่อไร่ที่สูงขึ้นตามสภาพอากาศที่เอื้ออำนวยโดยเฉพาะภาคใต้ (2) เปอร์เซ็นต์น้ำมันที่สูงจากการเก็บเกี่ยวผลผลิตในช่วงอายุของต้นปาล์มที่เหมาะสม รวมทั้ง (3) แรงจูงใจในการเก็บเกี่ยวจากราคาผลปาล์มที่สูง [2] สวนปาล์มน้ำมันในประเทศไทยส่วนใหญ่ตั้งอยู่ในภาคใต้และภาคตะวันออก โดยปัจจุบันมีการขยายพื้นที่เพาะปลูกไปยังพื้นที่อื่น ๆ มากขึ้น แม้ว่าประเทศไทยจะสามารถผลิตน้ำมันปาล์มได้เพียงพอต่อความต้องการของผู้ประกอบการในประเทศ แต่ผลผลิตส่วนใหญ่ยังมาจากสวนขนาดเล็กที่ดำเนินการโดยเกษตรกรและผู้ประกอบการรายย่อย ปัจจุบันพันธุ์ปาล์มน้ำมันในประเทศไทยยังไม่ได้รับการพัฒนามากนัก และมีการจัดการและนโยบายที่ไม่ดี ทำให้ต้นทุนการผลิตสูงและประสิทธิภาพในการผลิตต่ำกว่าผู้ผลิตหลักในอินโดนีเซียและมาเลเซียซึ่งมีศักยภาพในการผลิตสูงกว่า [3]

จากข้อมูลสำนักงานเศรษกิจการเกษตร สถานการณ์ปาล์มน้ำมัน ปี 2568 (ข้อมูลพยากรณ์ ณ พฤศจิกายน 2567) คาดว่า เนื้อที่ให้ผล 6.439 ล้านไร่ ผลผลิต 18.901 ล้านตัน ผลผลิตต่อเนื้อที่ให้ผล 2,935 กิโลกรัมต่อไร่ ด้านราคาผลปาล์มน้ำมันทั้งทะลาย ณ มกราคม 2568 เฉลี่ยอยู่ที่ 8.28 บาทต่อกิโลกรัม ยังอยู่ในระดับสูง [3] ทำให้เกษตรกรผู้ปลูกปาล์มน้ำมันต้องเผชิญกับความไม่เสถียรของราคาผลผลิต เนื่องจากผลผลิตปาล์มน้ำมันมีความผันผวนขึ้นอยู่กับหลายปัจจัย เช่น ปริมาณผลผลิตที่อาจลดลงหรือเพิ่มขึ้นตามสภาพภูมิอากาศและการจัดการสวน รวมถึงต้นทุนการผลิตที่สูงขึ้น เช่น ค่าปุ๋ย ค่าคนงาน เป็นต้น ทำให้เกษตรกรต้องใช้เงินลงทุนมากขึ้นเพื่อให้ได้ผลผลิตที่คาดหวัง หากราคาผลปาล์มน้ำมันในตลาดไม่สามารถปรับตัวตามต้นทุนที่เพิ่มขึ้นได้ จะทำให้เกษตรกรไม่สามารถทำกำไรจากการขายผลผลิตได้ หรือแม้กระทั่งขาดทุน ซึ่งทำให้เกษตรกรเสี่ยงต่อรายได้จากการทำสวนปาล์มน้ำมัน ส่งผลกระทบต่อการดำเนินชีวิตของพวกเขา ซึ่งการเปลี่ยนแปลงของปริมาณผลผลิตปาล์มน้ำมันมีความสัมพันธ์ทิศทางเดียวกับราคาน้ำมันปาล์มดิบ [4-5] และราคาน้ำมันปาล์มดิบมีการเคลื่อนไหวตลอดเวลาและขึ้นอยู่กับปัจจัยหลายอย่าง การเข้าใจปัจจัยที่มีผลต่อราคาน้ำมันปาล์มจะช่วยให้สามารถทำนายทิศทางของราคาได้อย่างแม่นยำมากขึ้น [6] จะเห็นได้ว่าราคาปาล์มน้ำมันมีผลต่อเสถียรภาพทางเศรษฐกิจของประเทศ การปลูกปาล์มน้ำมันยังเป็นอาชีพของเกษตรกรจำนวนมาก การผันผวนของราคาจะส่งผลกระทบต่อความมั่นคงรายได้ของเกษตรกรและอาจทำให้เกิดปัญหาทางสังคมและเศรษฐกิจในระดับท้องถิ่น รัฐบาลจึงต้องมีนโยบายควบคุมและสนับสนุนเพื่อรักษาเสถียรภาพราคาและป้องกันผลกระทบต่อทั้งเกษตรกรและเศรษฐกิจโดยรวม [3]

ดังนั้น จากเหตุผลดังกล่าวข้างต้นการเปลี่ยนแปลงราคาปาล์มน้ำมันในประเทศไทย มีเหตุมาจากปัจจัยต่าง ๆ ผู้วิจัยได้เห็นถึงความสำคัญของราคาปาล์มน้ำมันในประเทศไทย จึงได้ศึกษาเกี่ยวกับการหาตัวแบบถดถอยเชิงเส้นพหุคูณ เพื่อพยากรณ์ปัจจัยที่ส่งผลต่อราคาปาล์มน้ำมันในประเทศไทยซึ่งคาดว่าจะเป็นประโยชน์ต่อเกษตรกรปาล์มน้ำมันและเป็นฐานข้อมูลในการปรับปรุงพัฒนาเกษตรกรในประเทศไทย

วัตุประสงค์การวิจัย

1) เพื่อพยากรณ์ราคาปาล์มน้ำมันในประเทศไทยโดยการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

2) เพื่อศึกษาปัจจัยที่ส่งผลต่อราคาปาล์มน้ำมันในปี พ.ศ. 2558 – 2566

ระเบียบวิธีวิจัย

ประชากรและกลุ่มตัวอย่าง เป็นข้อมูลทุติยภูมิ (Secondary Data) เป็นข้อมูลที่เกี่ยวกับราคาปาล์มน้ำมัน และปัจจัยที่มีผลกระทบต่อราคาปาล์มน้ำมัน ได้แก่ ราคาน้ำมันปาล์มดิบ ปริมาณส่งออกน้ำมันปาล์ม ปริมาณการขายผลผลิตปาล์มน้ำมัน ปริมาณส่งออกสินค้าเกษตร อัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา ทำการรวบรวมข้อมูลแต่ละเดือน โดยใช้สถิติข้อมูลรายเดือนตั้งแต่เดือนมกราคม 2558 ถึงเดือนธันวาคม 2566 รวม จำนวน 108 เดือน

การเก็บรวบรวมข้อมูล

ตารางที่ 1: แสดงข้อมูลตัวแปรอิสระและตัวแปรตาม

	ตัวแปร	นิยาม	ที่มา
Y	ราคาปาล์มน้ำมัน (บาท : กิโลกรัม)	ราคาผลปาล์มน้ำมันทั้งทะลายของประเทศไทย	จากเว็บไซต์ สำนักงานเศรษฐกิจการเกษตร https://www.shorturl.asia/T5oEa
X1	ราคาน้ำมันปาล์มดิบ (บาท : กิโลกรัม)	ราคาน้ำมันปาล์มดิบขายส่งถึงโรงกลั่นของประเทศไทย	จากเว็บไซต์ กรมการค้าภายใน https://www.shorturl.asia/TQL0b
X2	ปริมาณส่งออกน้ำมันปาล์ม (ร้อยตัน : เดือน)	ปริมาณน้ำมันปาล์มของประเทศไทยส่งออกไปต่างประเทศ	จากเว็บไซต์ สำนักงานเศรษฐกิจการเกษตร https://impexpth.oae.go.th/export
X3	ปริมาณการขายผลผลิตปาล์มน้ำมัน (หมื่นตัน : เดือน)	ปริมาณการขายผลผลิตปาล์มน้ำมันของประเทศไทย	จากเว็บไซต์ สำนักงานเศรษฐกิจการเกษตร https://shorturl.asia/Yn5dt
X4	ปริมาณส่งออกสินค้าเกษตร (ตัน : เดือน)	การส่งออกสินค้าเกษตรทั้งหมดไปยังต่างประเทศ	จากเว็บไซต์ สำนักงานเศรษฐกิจการเกษตร https://www.shorturl.asia/Yn5dt

X5	อัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา (บาท : ดอลลาร์สหรัฐอเมริกา)	-	จากเว็บไซต์ investing.com https://shorturl.asia/dQrbs

2 ความรู้พื้นฐาน

2.1 การวิเคราะห์สถิติเชิงพรรณนา

การวิเคราะห์สถิติเชิงพรรณนา (Descriptive Statistics) เป็นการวิเคราะห์ขั้นพื้นฐานที่ทำให้เห็นภาพรวมของข้อมูลและความสัมพันธ์ระหว่างข้อมูล ซึ่งจะแสดงผลการวิเคราะห์ในรูปแบบตารางประกอบคำอธิบายโดยใช้ค่าสถิติ คือ ค่าเฉลี่ย ค่าต่ำสุด ค่าสูงสุด และส่วนเบี่ยงเบนมาตรฐาน (S.D.) [7]

2.2 การวิเคราะห์การถดถอย

การวิเคราะห์การถดถอยเป็นวิธีการทางสถิติใช้เพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป ประกอบด้วยตัวแปรที่เราทราบค่า เรียกว่า ตัวประมาณการหรือตัวแปรต้นและตัวแปรที่เราต้องการทราบค่า เรียกว่า ตัวตอบสนองหรือตัวแปรตามซึ่งความสัมพันธ์ของตัวแปรทั้งสองจะอยู่ในรูปแบบใด ๆ เช่น เส้นตรง เส้นโค้ง เป็นต้น [7]

2.2.1 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (multiple regression analysis)

เป็นการศึกษาระหว่างตัวแปรอิสระที่ทำหน้าที่พยากรณ์ตั้งแต่ 2 ตัวขึ้นไปกับตัวแปรตาม 1 ตัว ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณนั้นจะต้องหาค่าสัมประสิทธิ์สหสัมพันธ์เชิงพหุคูณ เพื่อให้ทราบถึงความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามว่ามีความสัมพันธ์กันอย่างไร สำหรับการวิเคราะห์การถดถอยเชิงเส้นพหุคูณจะต้องหาสมการถดถอยเพื่อใช้ในการพยากรณ์ของตัวแปรตามและหาค่าคลาดเคลื่อนมาตรฐาน รวมทั้งหาค่าสหสัมพันธ์พหุคูณ เพื่อหาความสัมพันธ์เชิงเส้นตรงที่เป็นไปได้สูงสุดระหว่างตัวแปรอิสระกับตัวแปรตาม ความสัมพันธ์ของตัวแปรอิสระและตัวแปรตามอยู่ในรูปของสมการ ดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + e$$

โดยที่

Y คือ ตัวแปรตาม

X_i คือ ตัวแปรอิสระ สำหรับ $i = 1,2,3,\ldots,k$

β_0 คือ ระยะตัดแกน Y เมื่อกำหนด $X_1 = X_2 = \ldots = X_k = 0$

β_i คือ สัมประสิทธิ์ความถดถอยเชิงส่วน $i = 1,2,3,\ldots,k$

e คือ ค่าคลาดเคลื่อน

โดยที่ค่า β_i เป็นค่าที่แสดงถึงการเปลี่ยนแปลง ของตัวแปรตาม Y โดยเฉลี่ยเมื่อตัวแปรอิสระ X_i เปลี่ยนไป 1 หน่วย โดยที่ตัวแปรอิสระ X ตัวอื่น ๆ มีค่าคงที่

2.3 เงื่อนไขของการวิเคราะห์การถดถอยเชิงเส้นพหุคูณมี 4 ข้อ ดังนี้

2.3.1 ค่าคลาดเคลื่อนมีการแจกแจงปกติ คือ ในการทดสอบว่าข้อมูลมีการแจกแจงปกติหรือไม่ สามารถทำการทดสอบได้โดยพิจารณาว่าข้อมูลอยู่ในระดับใด ถ้าข้อมูลอยู่ในระดับอันตรภาคและอัตราส่วนให้เลือกใช้สถิติทดสอบโคลโมโกรอฟ-สไมรนอฟ (Kolmogorov Smirnov Test)

2.3.2 ความสัมพันธ์ของค่าคลาดเคลื่อนเป็นอิสระกัน คือ การตรวจสอบความเป็นอิสระต่อกันของค่าคลาดเคลื่อน และสามารถตรวจสอบโดยวิธีการลงจุดค่าความคลาดเคลื่อนกับเวลา จุดกระจายแบบไม่มีแบบแผนเรียกว่า No Autocorrelation

2.3.3 ค่าแปรปรวนของความคลาดเคลื่อนเป็นค่าคงที่ คือ การตรวจสอบด้วยวิธีการลงจุดค่าความคลาดเคลื่อน กับค่าประมาณของ Y จุดกระจายอยู่รอบ ๆ ศูนย์ แบบไม่มีแบบแผนไม่ปรากฏรูปแบบการกระจายที่คล้ายลำโพงบาน

2.3.4 ตัวแปรอิสระแต่ละตัวต้องเป็นอิสระกัน มิฉะนั้นจะเรียกว่า เกิดปัญหา Multicollinearity คือ ตัวแปรอิสระเหล่านั้นจะต้องไม่มีความสัมพันธ์กันเอง การตรวจสอบตัวแปรอิสระมีความสัมพันธ์กัน จะวิเคราะห์จากผลของค่า VIF และ ค่า Tolerance [7]

2.4 การเลือกสมการถดถอยพหุคูณที่ดีที่สุด

2.4.3 วิธีลดตัวแปรอิสระ (Backward elimination procedure)

การเลือกสมการถดถอยเชิงเส้นโดยวิธีลดตัวแปรอิสระ เป็นการกำจัดตัวแปรอิสระที่คิดว่าไม่สำคัญ ออกไปจากสมการทีละตัวจนกระทั่งเหลือตัวแปรอิสระที่มีความสำคัญต่อตัวแปรตาม โดยมีขั้นตอนดังนี้

1. คำนวณสมการถดถอยเชิงเส้นที่มีตัวแปรอิสระทั้งหมดทุกตัว

2. ใช้การทดสอบเอฟบางส่วน (partial F–test) กับตัวแปรอิสระทุกตัวทีละตัว โดยถือว่าตัวแปรอิสระตัวที่ทดสอบเป็นตัวสุดท้ายที่ใส่เข้าไปในสมการถดถอยเชิงเส้น

3. นำค่าเอฟบางส่วนที่ได้ที่มีค่าต่ำสุด เปรียบเทียบกับค่าวิกฤตที่ระดับนัยสำคัญที่กำหนดซึ่งเป็นค่าเอฟที่เปิดได้จากตาราง ถ้ามีค่าต่ำกว่าค่าวิกฤตแสดงว่าตัวแปรอิสระตัวนั้น ไม่สามารถอธิบายความผันแปรของตัวแปรตามให้ตัดตัวแปรอิสระนั้นออกจากสมการถดถอยเชิงเส้น

4. แล้วนำตัวแปรอิสระที่เหลือสร้างสมการถดถอยเชิงเส้นแล้วดำเนินการเช่นเดิมจากข้อ 1-3 จนกระทั้งได้ค่าเอฟบางส่วนที่มีค่าต่ำที่สุดที่มีค่ามากกว่าค่าวิกฤตที่ระดับนัยสำคัญ ที่กำหนดแสดงว่าตัวแปรอิสระตัวนั้นสามารถอธิบายความผันแปรของตัวแปรตาม ตัวแปรอิสระที่ได้ เป็นตัวแปรอิสระที่ควรอยู่ในสมการถดถอยเชิงเส้น [7]

2.4.4 วิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise Selection)

วิธีเพิ่มตัวแปรอิสระแบบขั้นตอน เป็นวิธีการคัดเลือกตัวแปรอิสระเช่นเดียวกับวิธีเพิ่มตัวแปรอิสระแต่ทำการทดสอบความมีนัยสำคัญทางสถิติด้วยการทดสอบเอฟโดยลำดับทุกขั้นตอน [9]

2.5 การทดสอบ F (F-test) และ การทดสอบ t (t-test)

2.5.1 การทดสอบ F ใช้เพื่อพิจารณาความเหมาะสมของสมการถดถอยเชิงพหุ (Multiple Regression) โดยมีวัตถุประสงค์เพื่อตรวจสอบว่าตัวแปรอิสระทั้งหมดรวมกัน มีอิทธิพลอย่างมีนัยสำคัญทางสถิติในการพยากรณ์ค่าของตัวแปรตามหรือไม่ กล่าวคือ เป็นการทดสอบสมมติฐานร่วมของตัวแปรอิสระทุกตัวในแบบจำลองถดถอย โดยพิจารณาค่าp-value ของสถิติ F หากค่า p-value น้อยกว่าระดับนัยสำคัญที่กำหนดไว้ จะสามารถปฏิเสธสมมติฐานได้ซึ่งหมายความว่า แบบจำลองถดถอยโดยรวมนั้นมี นัยสำคัญทางสถิติกล่าวคือ ตัวแปรอิสระอย่างน้อยหนึ่งตัว มีอิทธิพลต่อค่าของตัวแปรตาม Y

2.5.2 การทดสอบ t การทดสอบ t ใช้เพื่อทดสอบความสำคัญของตัวแปรอิสระแต่ละตัวภายในสมการถดถอย โดยมีวัตถุประสงค์เพื่อตรวจสอบว่า ตัวแปรอิสระแต่ละตัวมีผลต่อการเปลี่ยนแปลงของตัวแปรตามหรือไม่ เมื่อควบคุมตัวแปรอื่น ๆ ไว้คงที่ การพิจารณาดำเนินการโดยพิจารณาค่า p-value ที่สัมพันธ์กับแต่ละค่าสัมประสิทธิ์ของตัวแปรอิสระ หากค่า p-value ของตัวแปรใดน้อยกว่าระดับนัยสำคัญที่กำหนด แสดงว่าตัวแปรอิสระนั้นมีอิทธิพลต่อตัวแปรตาม Y อย่างมีนัยสำคัญทางสถิติ ซึ่งหมายความว่าสามารถปฏิเสธสมมติฐานนั่นคือ ค่าสัมประสิทธิ์ของตัวแปรอิสระนั้นเท่ากับศูนย์ แสดงว่าตัวแปรอิสระนั้นไม่มีอิทธิผลต่อตัวแปร Y

2.6 การตรวจสอบความสัมพันธ์ของตัวแปรอิสระ (Multicollinearity)

ในการศึกษาหาความสัมพันธ์ระหว่างตัวแปรหลาย ๆ ตัว มีข้อกำหนดว่าตัวแปรอิสระเหล่านั้นจะต้องไม่มีความสัมพันธ์กันเองการตรวจสอบตัวแปรอิสระมีความสัมพันธ์กันจะวิเคราะห์จากผลของค่า VIF (Variance Inflation Factor) และค่า Tolerance

2.6.1 ค่า VIF (Variance Inflation Factor)

- ถ้าค่า VIF (X_i) มีค่ามาก นั่นคือ ตัวแปรอิสระแต่ละตัวมีความสัมพันธ์กันมาก ทำให้เกิดปัญหา Multicollinearity

- ถ้าค่า VIF (X_i) มีค่าเป็น 1 หรือใกล้ 1 นั่นคือ ตัวแปรอิสระมีความสันพันธ์กันน้อยมาก ทำให้ไม่เกิดปัญหา Multicollinearity

- ถ้าค่า VIF (X_i) มากกว่า 10 นั่นคือ ตัวแปรอิสระแต่ละตัวมีความสัมพันธ์กันมาก ทำให้เกิดปัญหา Multicollinearity

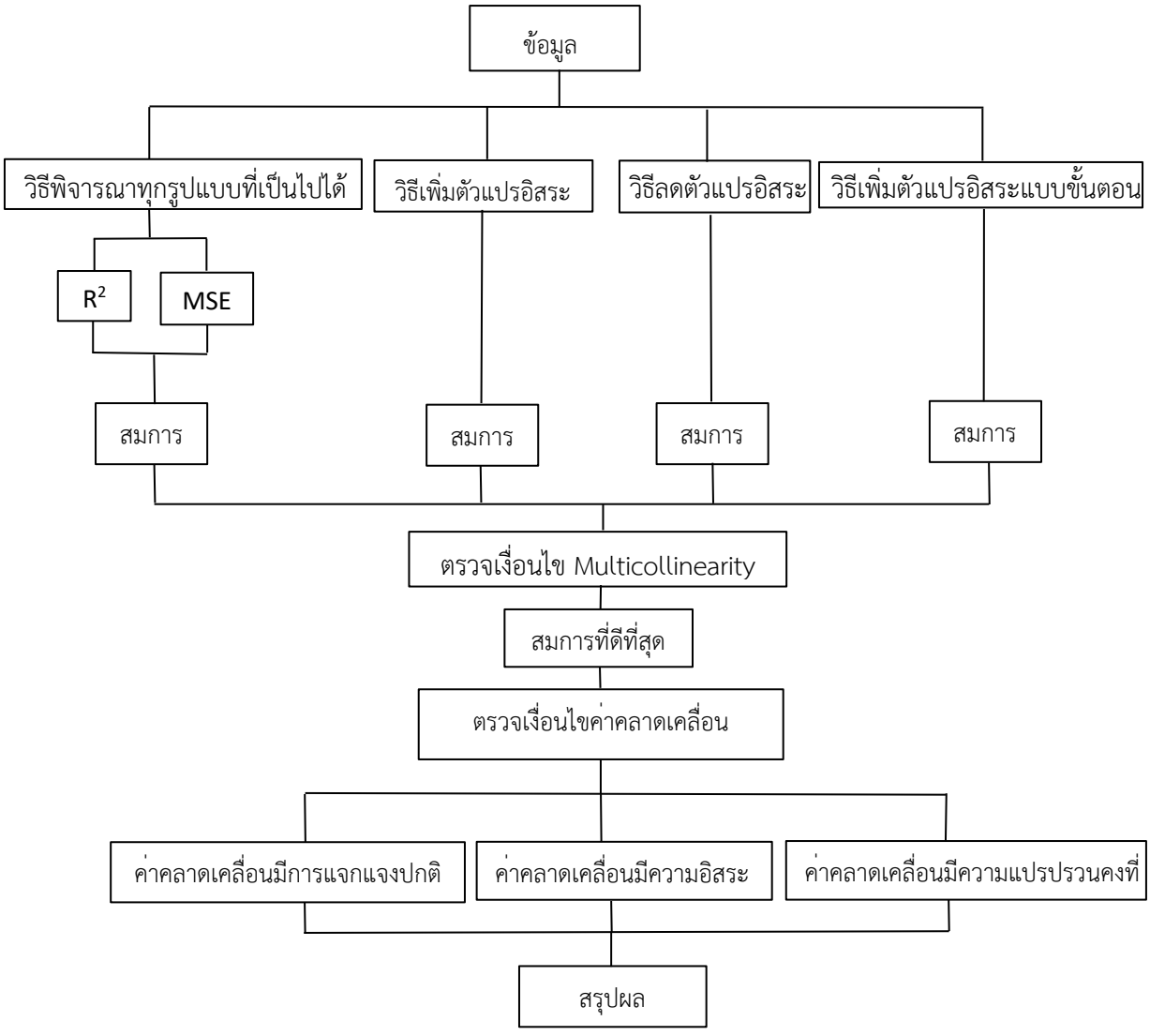
2.6.2 ค่า Tolerance

- Tolerance ใช้วัดความเป็นอิสระของตัวแปรอิสระทุกตัว ถ้าค่า Tolerance มีค่าเข้าใกล้ 1 แสดงว่าตัวแปรอิสระนั้นมีความสัมพันธ์กับตัวแปรอิสระอื่น ๆ ในระดับต่ำ แต่ถ้ามีค่าเข้าใกล้ 0 แสดงว่าตัวแปรอิสระนั้นมีความสัมพันธ์กับตัวแปรอิสระอื่น ๆ ในระดับสูงและถ้ามีค่าต่ำกว่า 0.1 แสดงว่ามีปัญหาเกี่ยวกับ Multicollinearity ขั้นรุนแรง

2.6.3 Kolmogorov-Smirnov test

- การทดสอบ Kolmogorov-Smirnov (K-S test) เป็นวิธีการทางสถิติที่พัฒนาโดย Andrey Kolmogorov และ Nikolai Smirnov เพื่อทดสอบว่าตัวอย่างมาจากประชากรที่มีการแจกแจงตามที่กำหนดหรือไม่ ในที่นี้คือการแจกแจงแบบปกติ (Normal Distribution) ซึ่งเป็นข้อตกลงเบื้องต้นที่สำคัญในการใช้สถิติเชิงพารามิเตอร์

2.7 flowchart



3 ผลการศึกษา

3.1 การศึกษาสถิติพรรณนา

ตารางที่ 2: แสดงค่าเฉลี่ย ค่าต่ำสุด ค่าสูงสุด และค่าส่วนเบี่ยงเบนมาตรฐาน

Variables	Mean	Minimum	Maximum	Std. Dev.
Y	2.16000	1.09000	3.21000	0.42000
X1	29.21000	14.76000	56.52000	9.28000
X2	393.56000	10.55000	2461.70000	485.42000
X3	130.15000	65.56000	201.81000	31.84000
X4	120.60000	89.26000	187.15000	20.24000
X5	33.84000	30.20000	38.47000	1.98000

จากตารางที่ 2 เป็นการแสดงค่าสถิติพรรณนา จากการเก็บรวบรวมข้อมูล ตัวแปรที่ศึกษา (Y) และตัวแปรอิสระต่าง ๆ (X_i) ข้อมูลเป็นรายเดือนตั้งแต่เดือน มกราคม พ.ศ. 2558 ถึงเดือนธันวาคม พ.ศ. 2566 รวมระยะเวลา 9 ปี หรือขนาดข้อมูลเท่ากับ 108 เดือน ซึ่งสามารถสรุปผลได้ดังนี้

ราคาปาล์มน้ำมัน (Y) มีราคาเฉลี่ยอยู่ที่ 2.16000 บาทต่อกิโลกรัม มีราคาต่ำสุดอยู่ที่ 1.09000 บาทต่อกิโลกรัม มีราคาสูงสุดอยู่ที่ 3.21 บาทต่อกิโลกรัม และมีส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.42000

ราคาน้ำมันปาล์มดิบ (X1) มีราคาเฉลี่ยอยู่ที่ 29.21000 บาทต่อกิโลกรัม มีราคาต่ำสุดอยู่ที่ 14.76000 บาทต่อกิโลกรัม มีราคาสูงสุดอยู่ที่ 56.52000 บาทต่อกิโลกรัม และมีส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 9.28000

ปริมาณส่งออกน้ำมันปาล์ม (X2) มีปริมาณการส่งออกเฉลี่ยอยู่ที่ 393.56000 ร้อยตันต่อเดือน มีปริมาณการส่งออกต่ำสุดอยู่ที่ 10.55000 ร้อยตันต่อเดือน มีปริมาณการส่งออกสูงสุดอยู่ที่ 2461.70000 ร้อยตันต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 485.42000

ปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) มีปริมาณการขายเฉลี่ยอยู่ที่ 130.15000 หมื่นตันต่อเดือน มีปริมาณการการขายต่ำสุดอยู่ที่ 65.56000 หมื่นตันต่อเดือน มีปริมาณการการขายสูงสุดอยู่ที่ 201.81000 หมื่นตันต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 31.84000

ปริมาณส่งออกสินค้าเกษตร (X4) มีปริมาณการส่งออกเฉลี่ยอยู่ที่ 120.60000 ตันต่อเดือน มีปริมาณการส่งออกต่ำสุดอยู่ที่ 89.26000 ตันต่อเดือน มีปริมาณการส่งออกสูงสุดอยู่ที่ 187.15000 ตันต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 20.24000

อัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา (X5) มีอัตราแลกเปลี่ยนเฉลี่ยอยู่ที่ 33.84000 บาทต่อดอลลาร์สหรัฐอเมริกา มีอัตราแลกเปลี่ยนต่ำสุดอยู่ที่ 30.20000 บาทต่อดอลลาร์สหรัฐอเมริกา มีอัตราแลกเปลี่ยนสูงสุดอยู่ที่ 38.47000 บาทต่อดอลลาร์สหรัฐอเมริกา และมีส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 1.98000

3.2 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณทั้ง 4 วิธี

การตรวจสอบเงื่อนไขเบื้องต้น Y มีการแจกแจงปกติ ในการวิเคราะห์การถดถอยพหุคูณ ตัวแปรที่ศึกษาต่อราคาปาล์มน้ำมัน (Y) ทำการทดสอบการแจกแจงปกติด้วย Kolmogorov-Smirnov ได้ว่าไม่มีการแจกแจงปกติ จึงแปลงเป็น $\sqrt{Y}$ แล้วทำการทดสอบ Kolmogorov-Smirnov สรุปได้ว่ามีการแจกแจงปกติ [10]

3.2.1 วิธีพิจารณาทุกรูปแบบที่เป็นไปได้ (All possible regression)

ตารางที่ 3: ผลการวิเคราะห์การถดถอยพหุคูณโดยวิธีพิจารณาทุกรูปแบบที่เป็นไปได้ (* p-value < 0.05)

ตัวแปรอิสระในสมการ	ค่าสัมประสิทธิ์การตัดสินใจ (R^2)	MSE
X1*, X2*, X3*, X4, X5*	0.94394	0.13078
X1X2*, **X1X3***, X1X4*, X1X5*, X2X3*, X2X4*, X2X5*, X3X4*, X3X5*, X4X5*	0.94808	0.08983
X1X2X3*, X1X2X4*, X1X2X5*, **X1X3X4***, X1X3X5*, X1X4X5*, X2X3X4*, X2X3X5*, X2X4X5*, X3X4X5*,	0.95161	0.05630
X1X2X3X4*, **X1X2X3X5***, X1X2X4X5*, X1X3X4X5*, X2X3X4X5*	0.95301	0.02938
X1X2X3X4X5*	0.95381	0.00852

การเลือกตัวแปรอิสระโดยการใช้สมการถดถอยเชิงเส้นที่เป็นไปได้ทุกรูปแบบมาพิจารณา ซึ่งสมการถดถอยเชิงเส้นที่เป็นไปได้ทั้งหมด 32 สมการ พิจารณาค่า R^2 โดยทำการจัดกลุ่มที่มี 1 ตัวแปร 2 ตัวแปร 3 ตัวแปร 4 ตัวแปร และ 5 ตัวแปร ในแต่ละกลุ่มเลือกสมการที่มีค่า R^2 สูงที่สุด และค่า p-value ของ F-test น้อยกว่า α = 0.05 โดยในกลุ่มที่มีตัวแปรอิสระ 1 ตัว ค่า R^2 สูงที่สุดคือ X1 มีค่าเท่ากับ 0.94394 กลุ่มตัวแปรอิสระ 2 ตัว ค่า R2 สูงที่สุดคือ X1X3 มีค่าเท่ากับ 0.94808 กลุ่มตัวแปรอิสระ 3 ตัว ค่า R2 สูงที่สุดคือ X1X3X4 มีค่าเท่ากับ 0.95161 กลุ่มตัวแปรอิสระ 4 ตัว ค่า R2 สูงที่สุดคือ X1X2X3X5 มีค่าเท่ากับ 0.95301 กลุ่มตัวแปรอิสระ 5 ตัว ค่า R2 สูงที่สุดคือ X1X2X3X4X5 มีค่าเท่ากับ 0.95381 แล้วคำนวณค่า MSE ในแต่ละกลุ่มที่ได้คัดเลือกสมการที่ดีที่สุด จากตารางที่ 3 กลุ่มตัวแปรอิสระที่มีค่าสัมประสิทธิ์การตัดสินใจ (R^2) สูงสุด และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำสุด คือ **X1X2X3X4X5**

ค่า F-statistic (F-test) มีค่า เท่ากับ 427.27190 หรือมีค่า Prob (F-statistic) ที่คำนวณได้มีค่าเท่ากับ 0 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 ซึ่งเป็นไปตามสมมติฐานที่ว่าตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กัน โดยตัวแปรอิสระอย่างน้อย 1 ตัว ที่มีผลกระทบต่อราคาปาล์มน้ำมัน (Y)

ตารางที่ 4: ผลการวิเคราะห์การถดถอยพหุคูณด้วยวิธีพิจารณาทุกรูปแบบที่เป็นไปได้

	Coefficients	Standard Error	t Stat	P-value
Intercept	0.75486	0.17321	4.35807	3.14E-05
X1	0.04161	0.00118	35.28018	5.74E-59
X2	0.00004	2.71E-05	1.51120	0.13383
X3	-0.00159	0.00042	-3.76757	0.00028
X4	0.00097	0.00073	1.33095	0.18617
X5	0.00775	0.00519	1.49298	0.13853

จากตารางที่ 4 จะได้สมการถดถอยเชิงเส้นพหุคูณคือ

$\hat{Y}$ = 0.75486+0.04161X1+0.00004X2-0.00159X3+0.00097X4+0.00775X5 และเมื่อพิจารณาจาก T-statistic (T-test) พบว่า ราคาน้ำมันปาล์มดิบ (X1) และปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) มีค่า Prob (T-statistic) ที่คำนวณได้น้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 นั่นคือ ราคาน้ำมันปาล์มดิบ (X1) และปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) มีความสัมพันธ์กับราคาปาล์มน้ำมัน แต่ปริมาณส่งออกน้ำมันปาล์ม (X2) ปริมาณส่งออกสินค้าเกษตร (X4) และอัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา (X5) ไม่มีอิทธิพลต่อ Y ที่ระดับนัยสำคัญทางสถิติ α = 0.05

3.2.2 วิธีเพิ่มตัวแปรอิสระ (Forward selection procedure)

ค่า F-statistic (F-test) มีค่า เท่ากับ 681.73400 หรือมีค่า Prob (F-statistic) ที่คำนวณได้มีค่าเท่ากับ 0 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 ซึ่งเป็นไปตามสมมติฐานที่ว่าตัวแปรตามและตัวแปรอิสระ มีความสัมพันธ์กัน โดยตัวแปรอิสระอย่างน้อย 1 ตัว ที่มีผลกระทบต่อราคาปาล์มน้ำมัน (Y)

ตารางที่ 5: ผลการวิเคราะห์ค่าสัมประสิทธิ์ถดถอยด้วยวิธีเพิ่มตัวแปรอิสระ

Model		B	Std. Error	t	P-value
1	(Constant)	0.87600	0.03200	27.49800	0. 00000
	X1	0.04400	0.00100	42.24600	0.00000
2	(Constant)	0.99100	0.05000	19.68600	0.00000
	X1	0.04400	0.00100	43.46200	0.00000
	X3	-0.00100	0.00000	-2.89400	0.00500
3	(Constant)	0.91400	0.05600	16.23200	0.00000
	X1	0.04200	0.00100	36.34000	0.00000
	X3	-0.00200	0.00000	-4.05600	0.00000
	X4	0.00200	0.00100	2.75500	0.00700

ตารางที่ 5 การวิเคราะห์ด้วยวิธีเพิ่มตัวแปรอิสระ จะได้สมการถดถอยเชิงเส้นพหุคูณคือ

$\hat{\sqrt{Y}}$ = 0.91400+0.04200X1-0.00200X3+0.00200X4

ใน Model ที่ 1 จะเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมากที่สุด ใน Model ที่ 1 เลือก X1 เข้าสมการทดสอบ F-test และจะทำการคัดเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามอันดับถัดไปเข้าสู่สมการ จาก Model ที่ 2 เลือก X3 เพิ่ม เมื่อมี X1 อยู่ในสมการแล้วทดสอบ F-test และจะทำไปจนกว่าจะไม่มีตัวแปรอิสระเข้าไปในสมการได้อีก จาก Model ที่ 3 เลือก X4 เพิ่ม เมื่อมี X1 และ X3 อยู่ในสมการแล้วทดสอบ F-test

จาก Model ที่ 3 เมื่อพิจารณาจาก T-statistic (T-test) พบว่า ราคาน้ำมันปาล์มดิบ (X1) ปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) และปริมาณส่งออกสินค้าเกษตร (X4) มีค่า Prob (T-statistic) ที่คำนวณได้น้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 นั่นคือ ราคาน้ำมันปาล์มดิบ (X1) ปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) และปริมาณส่งออกสินค้าเกษตร (X4) มีอิทธิพลต่อราคาปาล์มน้ำมัน (Y) ที่ระดับนัยสำคัญทางสถิติ α = 0.05

3.2.3 วิธีลดตัวแปรอิสระ (Backward elimination procedure)

ค่า F-statistic (F-test) มีค่า เท่ากับ 522.24100 หรือมีค่า Prob (F-statistic) ที่คำนวณได้มีค่าเท่ากับ 0 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 ซึ่งเป็นไปตามสมมติฐานที่ว่าตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กัน โดยตัวแปรอิสระอย่างน้อย 1 ตัว ที่มีผลกระทบต่อราคาปาล์มน้ำมัน (Y)

ตารางที่ 6: ผลการวิเคราะห์ค่าสัมประสิทธิ์ถดถอยด้วยวิธีลดตัวแปรอิสระ

Model		B	Std. Error	t	P-value
1	(Constant)	0.75500	0.17300	4.35800	0.00000
	X1	0.04200	0.00100	35.28000	0.00000
	X2	4.091E-5	0.00000	1.51100	0.13400
	X3	-0.00200	0.00000	-3.76800	0.00000
	X4	0.00100	0.00100	1.33100	0.18600
	X5	0.00800	0.00500	1.49300	0.13900
2	(Constant)	0.72300	0.17200	4.19900	0.00000
	X1	0.04200	0.00100	38.57900	0.00000
	X2	5.087E-5	0.00000	1.94800	0.05400
	X3	-0.00100	0.00000	-3.55300	0.00100
	X5	0.01000	0.00500	2.19100	0.03100

จากตารางที่ 6 การวิเคราะห์ด้วยวิธีลดตัวแปรอิสระ จะได้สมการคือ

$\hat{\sqrt{Y}}$ = 0.723+0.042X1+5.087E-5X2-0.001X3+0.010X5

จะทำการเลือกตัวแปรอิสระทุกตัวเข้าในสมการ ใน Model ที่ 1 เลือก X5 X3 X1 X2 และ X4 ทดสอบ F-test แล้วจึงคัดเลือกตัวแปรอิสระออกทีละตัว โดยตัวแปรอิสระที่มีความสัมพันธ์น้อยที่สุดจะถูกคัดออก ใน Model ที่ 2 เลือก X5 X3 X1 และ X2 แล้วทดสอบ F-test

จาก Model 2 พิจารณาจาก T-statistic (T-test) พบว่า ราคาน้ำมันปาล์มดิบ (X1) ปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) และอัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา (X5) มีค่า Prob (T-statistic) ที่คำนวณได้น้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 นั่นคือ ราคาน้ำมันปาล์มดิบ (X1) ปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) และอัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา (X5) มีอิทธิพลต่อราคาปาล์มน้ำมัน (Y) ที่ระดับนัยสำคัญทางสถิติ α = 0.05 แต่ปริมาณส่งออกน้ำมันปาล์ม (X2) ไม่มีอิทธิพลต่อราคาปาล์มน้ำมัน (Y) ที่ระดับนัยสำคัญทางสถิติ α = 0.05

3.2.4 วิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise Selection)

ค่า F-statistic (F-test) มีค่า เท่ากับ 681.73400 หรือมีค่า Prob (F-statistic) ที่คำนวณได้มีค่าเท่ากับ 0 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 ซึ่งเป็นไปตามสมมติฐานที่ว่าตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กัน โดยตัวแปรอิสระอย่างน้อย 1 ตัว ที่มีผลกระทบต่อราคาปาล์มน้ำมัน (Y)

ตารางที่ 7: ผลการวิเคราะห์ค่าสัมประสิทธิ์ถดถอยด้วยวิธีวิธีเพิ่มตัวแปรอิสระแบบขั้นตอน

Model		B	Std. Error	t	P-value
1	(Constant)	0.87600	0.03200	27.49800	0.00000
	X1	0.04400	0.00100	42.24600	0.00000
2	(Constant)	0.99100	0.05000	19.68600	0.00000
	X1	0.04400	0.00100	43.46200	0.00000
	X3	-0.00100	0.00000	-2.89400	0.00500
3	(Constant)	0.91400	0.05600	16.23200	0.00000
	X1	0.04200	0.00100	36.34000	0.00000
	X3	-0.00200	0.00000	-4.05600	0.00000
	X4	0.00200	0.00100	2.75500	0.00700

จากตารางที่ 7 การวิเคราะห์ด้วยวิธี เพิ่มตัวแปรอิสระแบบขั้นตอน จะได้สมการคือ

$\hat{\sqrt{Y}}$ = 0.914+0.042X1-0.002X3+0.002X4 จะ

เหมือนกับวิธีเพิ่มตัวแปรอิสระ ใน Model ที่ 1 จะเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมากที่สุด ใน Model ที่ 1 เลือก X1 เข้าสมการทดสอบ F-test และจะทำการคัดเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามอันดับถัดไปเข้าสู่สมการ จาก Model ที่ 2 เลือก X3 เพิ่ม เมื่อมี X1 อยู่ในสมการแล้วทดสอบ F-test และจะ

ทำไปจนกว่าจะไม่มีตัวแปรอิสระเข้าไปในสมการได้อีก จาก Model ที่ 3 เลือก X4 เพิ่ม เมื่อมี X1 และ X3 อยู่ในสมการแล้วทดสอบ F-test

จาก Model 3 พิจารณาจาก T-statistic (T-test) พบว่า ราคาน้ำมันปาล์มดิบ (X1) ปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) และปริมาณส่งออกสินค้าเกษตร (X4) มีค่า Prob (T-statistic) ที่คำนวณได้น้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 นั่นคือราคาน้ำมันปาล์มดิบ (X1) ปริมาณการขายผลผลิตปาล์มน้ำมัน (X3) และปริมาณส่งออกสินค้าเกษตร (X4) มีอิทธิพลต่อราคาปาล์มน้ำมัน (Y) ที่ระดับนัยสำคัญทางสถิติ α = 0.05

3.3 การเลือกสมการถดถอยเชิงเส้นพหุคูณที่ดีที่สุด

3.3.1 การเปรียบเทียบค่า R^2 ของสมการถดถอยเชิงเส้นพหุคูณทั้ง 4 วิธี

ตารางที่ 8: เปรียบเทียบค่า R^2 และสมการถดถอยเชิงเส้นพหุคูณทั้ง 4 วิธี

วิธี	R^2	สมการถดถอยเชิงเส้นพหุคูณ
All possible	0.953	Y = 0.75486+0.04161X1+0.00004X2-0.00159X3+0.00097X4+0.00775X5
Stepwise	0.952	Y = 0.914+0.042X1-0.002X3+0.002X4
Forward	0.952	Y = 0.914+0.042X1-0.002X3+0.002X4
Backward	0.953	Y = 0.723+0.042X1+5.087E-5X2-0.001X3+0.010X5

จากตารางที่ 8 สมการถดถอยเชิงเส้นพหุคูณทั้ง 4 วิธี พบว่า ค่าสัมประสิทธิ์การตัดสินใจ (R^2) แต่ละวิธีมีค่าไม่แตกต่างกันมากนัก คือ 0.95 และพบว่าทั้ง 4 วิธี มีผลลัพธ์ 3 แบบ คือ วิธีพิจารณาทุกรูปแบบที่เป็นไปได้ (All possible) มีตัวแปรอิสระ 5 ตัวแปร วิธีลดตัวแปรอิสระ (Backward) มีตัวแปรอิสระ 4 ตัวแปร และอีก 2 วิธี คือ วิธีเพิ่มตัวแปรอิสระ (Forward) และวิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise) มีสมการรูปแบบเดียวกันที่มีตัวแปรอิสระ 3 ตัว

3.3.2 การตรวจสอบคุณสมบัติของปัญหา Multicollinearity สมการถดถอยเชิงเส้นพหุคูณ

ตารางที่ 9: การตรวจสอบคุณสมบัติของปัญหา Multicollinearity ของสมการถดถอยเชิงเส้นพหุคูณทั้ง 4 วิธี

ตัวแปร	วิธี					
	Stepwise/ Forward		All possible regression		Backward	
	Tolerance	VIF	Tolerance	VIF	Tolerance	VIF
X1	0.71000	1.40800	0.66500	1.50300	0.77900	1.28400
X2	-	-	0.46100	2.16800	0.49900	2.00300
X3	0.58500	1.70900	0.44100	2.27000	0.57200	1.74900
X4	0.48800	2.05100	0.36600	2.73600	-	-
X5	-	-	0.75200	1.32900	0.89200	1.12100

จากตารางที่ 9 วิธีเพิ่มตัวแปร (Forward) และ วิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise) ค่า Tolerance เข้าใกล้ 0 และค่า VIF น้อย และเมื่อเพิ่มตัวแปร X5 จากวิธีพิจารณาทุกรูปแบบที่เป็นไปได้ (All possible) และวิธีลดตัวแปรอิสระ (Backward) ทำให้ ค่า Tolerance เพิ่มขึ้น

ดังนั้นสมการถดถอยเชิงเส้นพหุคูณที่ดีที่สุดเลือกจากวิธี Stepwise และ Forward เนื่องจาก จากการทดสอบค่า t-test ทุกตัวแปรอิสระ(X1 X3 และ X4) มีอิทธิพลต่อ Y และ R^2 มีค่าสูง และเมื่อตรวจสอบค่า VIF และค่า Tolerance พบว่า หากตัวแปรอิสระแต่ละตัวมี VIF น้อยกว่า 10 และ Tolerance เข้าใกล้ 1 จะสามารถสรุปได้ว่า ไม่มีปัญหา Multicollinearity และสามารถใช้ตัวแปรนั้นในการวิเคราะห์การถดถอยได้ สมการคือ

$\hat{Y}$ = 0.914+0.042X1-0.002X3+0.002X4

3.4 การทดสอบคุณสมบัติของค่าคลาดเคลื่อน

นำสมการ $\hat{Y}$ = 0.914+0.042X1-0.002X3+0.002X4 ทดสอบคุณสมบัติของค่าคลาดเคลื่อน

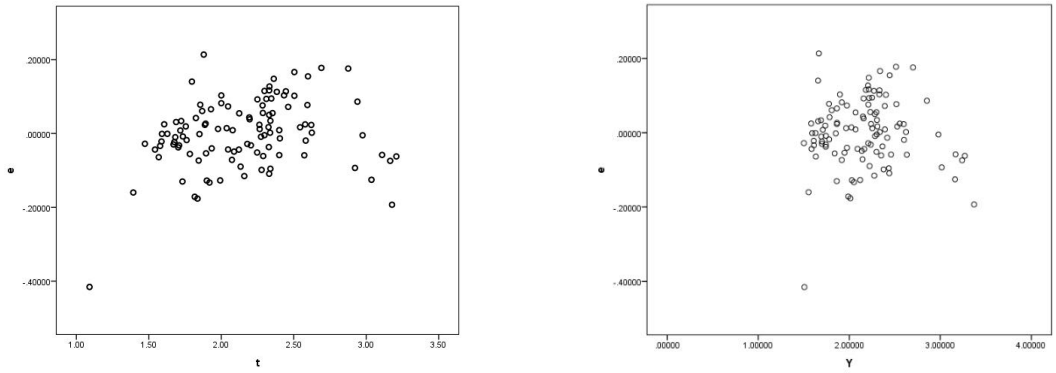
3.4.1 ค่าคลาดเคลื่อนมีการแจกแจงปกติ

ตารางที่ 10: ผลการทดสอบการแจกแจงปกติของค่าคลาดเคลื่อน โดยใช้สถิติทดสอบโครโมโกรอฟ-สไมรนอฟ

	Kolmogorov-Smirnov		
	Statistic	df	P-value
Residual	0.06700	108	0.20000

จากตารางที่ 10 พบว่า ค่า Kolmogorov-Smirnov test มีค่าเท่ากับ 0.067 หรือมีค่า P-value ที่คำนวณได้มีค่าเท่ากับ 0.2 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญทางสถิติ α = 0.05 ซึ่งเป็นไปตามสมมติฐานที่ว่า ค่าคลาดเคลื่อนมีการแจกแจงปกติ

3.4.2 ค่าคลาดเคลื่อนมีความอิสระกันและมีความแปรปรวนคงที่



(a) (b)

รูปที่ 1 : (a) กราฟแสดงค่าคลาดเคลื่อน (e) กับ เวลา (t) และ (b) กราฟแสดงค่าคลาดเคลื่อน (e) กับ ค่าพยากรณ์ ($\hat{Y}$)

รูปที่ 1 (a) พบว่าจุดกระจายตัวแบบไม่เป็นแบบแผน ดังนั้นค่าคลาดเคลื่อนมีความเป็นอิสระกันเรียกว่าไม่เกิดปัญหา Autocorrelation และรูปที่ 1 (b) พบว่าค่าคลาดเคลื่อนอยู่ในช่วง -2 ถึง 2 และมีค่ากระจายแบบสุ่ม จึงถือว่าค่าความแปรปรวนของค่าคลาดเคลื่อนคงที่

4 สรุปผล อภิปรายผล ข้อจำกัด และข้อเสนอแนะของงานวิจัย

4.1 สรุปผล

จากผลการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ โดยวิธีเพิ่มตัวแปร (Forward) และ วิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise) พบว่า ปัจจัยที่ส่งผลต่อราคาปาล์มน้ำ คือ ราคาน้ำมันปาล์มดิบ กับปริมาณส่งออกสินค้าเกษตร มีความสัมพันธ์กับราคาปาล์มน้ำมันไปในทิศทางเดียวกัน และปริมาณการขายผลผลิตปาล์มน้ำมัน มีความสัมพันธ์กับราคาปาล์มน้ำมันไปในทิศทางตรงกันข้าม และปริมาณส่งออกน้ำมันปาล์มกับอัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา ไม่มีความสัมพันธ์กับราคาปาล์มน้ำมัน อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญทางสถิติ α = 0.05 โดยมีค่าสัมประสิทธิ์การตัดสินใจ (R^2) เท่ากับ 0.952 กล่าวคือปัจจัยที่เกี่ยวข้องกับสมการนี้มีความสามารถในการอธิบายความผันผวนของราคาปาล์มน้ำมันได้ร้อยละ 95.2 ส่วนอีกร้อยละ 4.8 เกิดจากปัจจัยอื่น ๆ ที่ไม่ได้ นำมาวิเคราะห์ ตัวแบบพยากรณ์ คือ

$$\sqrt{\hat{Y}} = 0.914+0.042X1-0.002X3+0.002X4$$

4.2 อภิปรายผล

การศึกษาปัจจัยที่ส่งผลต่อราคาปาล์มน้ำมันประเทศไทยปี พ.ศ. 2558- 2566 กล่าวได้ว่าปัจจัยที่ส่งผลต่อราคาปาล์มน้ำมัน ได้แก่ ราคาน้ำมันปาล์มดิบ ปริมาณการขายผลผลิตปาล์มน้ำมัน และปริมาณส่งออกสินค้าเกษตร โดยที่ราคาน้ำมันปาล์มดิบ และปริมาณส่งออกสินค้าเกษตร มีความสัมพันธ์ในทิศทางบวกกับราคาปาล์มน้ำมัน และปริมาณการขายผลผลิตปาล์มน้ำมัน มีความสัมพันธ์ในทิศทางผกผันกับราคาปาล์มน้ำมัน อย่างมีนัยสำคัญทางสถิติ 0.05 ข้อมูลงานวิจัยนี้สอดคล้องกับการศึกษาของ วาริสวริน ชาญสุไชย [11] ที่ศึกษาปัจจัยที่มีผลกระทบต่อราคาน้ำมันปาล์มบริสุทธิ์ในประเทศไทย และอุทิศ ฉัตรสิรภพ [12] คือ ปัจจัยที่ส่งผลต่อราคาน้ำมันปาล์มบริสุทธิ์ในประเทศไทยปี 2560 ในทิศทางเดียวกัน คือ ราคาผลปาล์มทะลายคุณภาพน้ำมัน 17% น้ำมันปาล์มดิบในประเทศไทย ไขปาล์มบริสุทธิ์/สเตียรีนในประเทศไทย น้ำมันปาล์มบริสุทธิ์ที่ส่งออกประเทศมาเลเซีย และสต๊อกน้ำมันปาล์มบริสุทธิ์ภายในประเทศ และ วิชัย รุ่งเรืองอนันต์ [4] ได้ศึกษากลไกการวิเคราะห์สถานการณ์ของราคาปาล์มน้ำมัน พบว่า ปัจจัยที่มีผลต่อเสถียรภาพของราคาน้ำมันปาล์มดิบ ประกอบด้วย ปริมาณการผลผลิตปาล์มน้ำมัน ราคาน้ำมันปาล์มดิบในตลาดมาเลเซีย กระบี่ สุราษฎร์ธานี และชุมพร ราคาไบโอดีเซล ราคาน้ำมันปาล์มบริสุทธิ์ของไทย และสอดคล้องกับ นอกจากนี้การศึกษาของ เอกกร ภุมมารินทร์ [5] กล่าวว่าราคาน้ำมันปาล์มดิบในภาคใต้ของประเทศไทย มีความสัมพันธ์กับราคาผลปาล์มน้ำมันทั้งทะลายของประเทศไทย และราคาน้ำมันดีเซล

แม้ว่าการส่งออกน้ำมันปาล์มจะมีมูลค่าสูง แต่การเปลี่ยนแปลงของปริมาณการส่งออกกลับไม่มีผลโดยตรงหรือนัยสำคัญต่อราคาผลปาล์มน้ำมันในประเทศไทย เนื่องจากโครงสร้างตลาดภายในประเทศแยกขาดระหว่างต้นน้ำกับปลายน้ำ โดยราคาที่เกษตรกรได้รับมักถูกกำหนดโดยโรงงานแปรรูป ซึ่งอ้างอิงจากต้นทุนวัตถุดิบภายในประเทศและอุปสงค์ในประเทศ มากกว่าจากตลาดส่งออก [14] รายงานวิจัยจากมหาวิทยาลัยสงขลานครินทร์ยังชี้ให้เห็นว่า

ราคาผลปาล์มสดที่เกษตรกรขายได้มักถูกกำหนดโดยพ่อค้าคนกลาง ซึ่งจะส่งต่อไปยังโรงงานแปรรูปที่เป็นผู้กำหนดราคาซื้ออีกทอดหนึ่ง [15] ส่งผลให้กลไกราคาในระดับเกษตรกรได้รับอิทธิพลจากตลาดภายในและนโยบายของรัฐมากกว่าปัจจัยภายนอก เช่น อัตราแลกเปลี่ยนเงินตราต่างประเทศ การศึกษาด้วยแบบจำลองเชิงปริมาณพบว่า อัตราแลกเปลี่ยนเงินบาทไม่มีนัยสำคัญทางสถิติต่อราคาผลปาล์มน้ำมัน [16] แม้ความเคลื่อนไหวของอัตราแลกเปลี่ยนจะมีผลต่อราคาสินค้าส่งออกในระดับปลายน้ำ เช่น น้ำมันปาล์ม แต่ก็ไม่มีผลต่อราคาวัตถุดิบในระดับเกษตรกร [17] อย่างไรก็ตาม การส่งออกสินค้าเกษตรโดยรวมที่เพิ่มขึ้นมีแนวโน้มส่งผลให้อุปสงค์ภายในประเทศต่อวัตถุดิบเพิ่มขึ้น ซึ่งสามารถผลักดันให้ราคาผลปาล์มน้ำมันปรับตัวสูงขึ้นได้ [18] สะท้อนว่าราคาผลปาล์มน้ำมันในประเทศไทยตอบสนองต่อภาวะเศรษฐกิจการเกษตรโดยรวม มากกว่าปัจจัยเฉพาะตัวของสินค้า โครงสร้างตลาดไทยที่กระจุกอำนาจอยู่ในมือของผู้แปรรูปและผู้ส่งออกยังคงเป็นอุปสรรคต่อการถ่ายทอดผลประโยชน์จากตลาดโลกมายังเกษตรกร [14] จึงควรส่งเสริมการรวมกลุ่มเกษตรกรผ่านสหกรณ์ที่เข้มแข็ง พร้อมทั้งพัฒนาระบบราคาที่อิงกับตลาดสากล เพื่อสร้างความเป็นธรรมและเพิ่มขีดความสามารถในการแข่งขันในตลาดโลก

4.3 ข้อจำกัด

1. งานวิจัยนี้ไม่ได้นำตัวแปรต้นทุนการผลิต (ราคาปุ๋ย ค่าแรง) สภาพภูมิอากาศ (ปริมาณน้ำฝน อุณหภูมิ) ราคาน้ำมันดิบตลาดโลก ราคาพืชน้ำมันคู่แข่ง (ถั่วเหลือง) สต็อกน้ำมัน ปาล์ม นโยบายภาครัฐ (ราคาประกัน การส่งเสริมไบโอดีเซล) มาวิเคราะห์

2. งานวิจัยนี้ข้อมูลเป็นอนุกรมเวลา นำการวิเคราะห์ถดถอยเชิงเส้นพหุคูณมาใช้มีความเสี่ยงที่จะเกิดปัญหาความสัมพันธ์ปลอม (Spurious Correlation)

3. ข้อจำกัดของข้อมูลรายเดือน อาจไม่สามารถจับความผันผวนระยะสั้นได้ ช่วงเวลาที่ศึกษา (พ.ศ. 2558-2566) ผลอาจไม่สามารถใช้ได้กับช่วงเวลาอื่น

4.4 ข้อเสนอแนะ

1. ศึกษาตัวแปรอิสระเพิ่มเติม ข้อมูลด้านภูมิอากาศ

2. ทดสอบ Unit Root/Cointegration เพื่อลดความเสี่ยงจาก Spurious Regression ซึ่งเป็นข้อควรระวังใน การนำผลไปใช้และเสนอให้เป็นแนวทางวิจัยในอนาคตแทน

กิตติกรรมประกาศ

การศึกษาวิจัยครั้งนี้สำเร็จลุล่วงไปด้วยดีต้องขอขอบคุณสำนักงานเศรษฐกิจการเกษตร กรมการค้าภายใน และ investing.com ที่ให้ข้อมูลการศึกษาในการทำวิจัยครั้งนี้ ได้แก่ ราคาปาล์มน้ำมัน ราคาน้ำมันปาล์มดิบ ปริมาณส่งออกน้ำมันปาล์ม ปริมาณการขายผลผลิตปาล์มน้ำมัน ปริมาณส่งออกสินค้าเกษตร และอัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐอเมริกา

เอกสารอ้างอิง

[1] ศรราม แก้วตาทิพย์, การศึกษาศักยภาพการผลิตปาล์มน้ำมันของเกษตรกรในตำบลสวนหลวง อำเภอเฉลิมพระเกียรติ จังหวัดนครศรีธรรมราช, (วิทยานิพนธ์ปริญญามหาบัณฑิต), มหาวิทยาลัยทักษิณ, 2563.

[2] ชัยวัฒน์ สาวเจริญสุข, แนวโน้มธุรกิจ/อุตสาหกรรม ปี 2568-2570: อุตสาหกรรมน้ำมันปาล์ม. ออนไลน์. สืบค้นเมื่อ 25 มีนาคม 2568. แหล่งที่มา : https://www.krungsri.com/ th/research/ industry/industry-outlook/agriculture/palm-oil/IO/Oil-palm-industry-2022-2024

[3] สำนักงานเศรษกิจการเกษตร สืบค้นเมื่อ 1 มีนาคม 2568. แหล่งที่มา : https://www.oae.go.th/

[4] วิชัย รุ่งเรืองอนันต์, กลไกในการวิเคราะห์สถานการณ์ของราคาปาล์มน้ำมัน, สำนักงานกองทุนสนับสนุนการวิจัย (สกว.), กรุงเทพมหานคร, 2558.

[5] เอกกร ภุมมารินทร์, ลลิตา จันทรวงศ์ไพศาล, ปัจจัยที่ส่งผลกระทบต่อราคาน้ำมันปาล์มดิบของในภาคใต้ประเทศไทย, การประชุมทางวิชาการและนำเสนอผลงานวิชาการระดับชาติ UTCC, 2561.

[6] E. Pummarinth, L. Jantarawongpaisan, Factors effecting the palm oil price in southern of Thailand, UTCC Academic Day Proceeding. (2018) 852-864.

[7] กัลยา วานิชย์บัญชา, ฐิตา วานิชย์บัญชา, การใช้ SPSS for Windows ในการวิเคราะห์ข้อมูล, ศูนย์หนังสือแห่งจุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร, 2561.

[8] ทรงศิริ แต้สมบัติ, การวิเคราะห์การถดถอย, พิมพ์ครั้งที่ 3, กรุงเทพมหานคร, มหาวิทยาลัยเกษตรศาสตร์, 2548.

[9] มัลลิกา บุนนาค, สถิติเพื่อการวิจัยและตัดสินใจ, โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร, 2555.

[10] S. Mushtak, A.K., Multivariate Statistical Analysis. British Journal of Science, 6(1), (2012) 55-66.

[11] วาริสวริน ชาญสุไชย, ปัจจัยที่มีผลต่อราคาปาล์มน้ำมันของไทย, มหาวิทยาลัยราชภัฏสวนสุนันทา, กรุงเทพมหานคร, 2555.

[12] อุทิศ ฉัตรสิรภพ, ปัจจัยที่ส่งผลต่อราคาน้ำมันปาล์มในประเทศไทยปี 2560. วารสารการวิจัยการบริหารการพัฒนา. 9 (1), (2560) 79-85.

[13] สำนักงานเศรษฐกิจการเกษตร สืบค้นเมื่อ 1 พฤษภาคม 2568. แหล่งที่มา : https://www.oae.go.th/

[14] ภณัชษ์ฉัตร คำนวณศิลป์, ฉัตรแก้ว ธนินทรานนท์, รพิพรรณี วรรณเพชร, ญาณินี ศรีสำราญ, ปกรณ์ ลิ้มโยธิน, เจตน์สฤษฎ์ สังขพันธ์, ปัจจัยเสี่ยงของการเปลี่ยนแปลงราคาน้ำมันปาล์มดิบในประเทศไทย: การทบทวนวรรณกรรม. การประชุมหาดใหญ่วิชาการระดับชาติและนานาชาติ ครั้งที่ 14, 2566.

[15] ปองพชร ธาราสุข, อภิญญา รัตนไชย, รุ่งรัตน์ แซ่หยาง, การวิเคราะห์ต้นทุนและผลตอบแทนทางเศรษฐกิจของเกษตรกรผู้ปลูกปาล์มน้ำมันรายย่อย, คณะทรัพยากรธรรมชาติ, มหาวิทยาลัยสงขลานครินทร์, 2562.

[16] เจตน์สฤษฎ์ สังขพันธ์, ภณัชษ์ฉัตร คำนวณศิลป์, ปกรณ์ ลิ้มโยธิน, ปัจจัยที่ส่งผลต่อการเปลี่ยนแปลงราคาน้ำมันปาล์มดิบในประเทศไทย, Journal of Roi Kaensarn Academi, 9(7), (2567): 995–1013.

[17] World Bank, Commodity Markets Outlook – October 2023, Washington, D.C.: World Bank Publications, 2023.

[18] สำนักงานเศรษฐกิจการเกษตร, สืบค้นเมื่อ 1 พฤษภาคม 2568 จาก https://www.oae.go.th

การเปรียบเทียบประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวนเมื่อมีค่านอกเกณฑ์ในชุดข้อมูล

วนิดา พงษ์ศักดิ์ชาติ[1,*]

[1]ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา 20131

บทคัดย่อ

การทดสอบความเท่ากันของความแปรปรวนเป็นวิธีเชิงสถิติที่สำคัญที่ถูกนำมาใช้ในการตรวจสอบข้อสมมุติของการวิเคราะห์ความแปรปรวนซึ่งมีอยู่ด้วยกันหลายวิธี แต่การทดสอบส่วนใหญ่มักมีประสิทธิภาพดีเมื่อข้อมูลที่ศึกษามีการแจกแจงปรกติ ในการศึกษานี้จึงต้องการศึกษาการทดสอบความเท่ากันของความแปรปรวนของ ประชากร 3 กลุ่ม จำนวน 7 ชนิด เมื่อมีค่านอกเกณฑ์ร้อยละ 10 ในชุดข้อมูลแต่ละกลุ่ม เพื่อหาการทดสอบที่มีประสิทธิภาพสำหรับข้อมูลที่มีค่านอกเกณฑ์ ในการศึกษาได้ใช้การจำลองมอนติคาร์โล ศึกษาในกรณีของค่านอกเกณฑ์ 2 ชนิด คือค่านอกเกณฑ์ไม่รุนแรง และค่านอกเกณฑ์สุดขีด ขนาดตัวอย่างทั้งกรณีที่เท่ากันทุกกลุ่มและแตกต่างกันบางกลุ่ม และอัตราส่วนความแปรปรวนแตกต่างกันหลายระดับ ผลการศึกษาพบว่าการทดสอบเลวีนแบบไม่อิงพารามิเตอร์เป็นการทดสอบที่มีประสิทธิภาพมากที่สุดเมื่อมีค่านอกเกณฑ์ในชุดข้อมูล

คำสำคัญ: การทดสอบความเท่ากันของความแปรปรวน, ค่านอกเกณฑ์, ความสามารถในการควบคุมความผิดพลาดแบบที่ 1, ค่าประมาณกำลังการทดสอบ

2020 MSC: 62F03, 62F35

1 บทนำ

การวิเคราะห์ความแปรปรวน (Analysis of variance) เป็นวิธีที่ใช้ในการหาค่าสถิติทดสอบเอฟ (F test) เพื่อทดสอบเปรียบเทียบค่าเฉลี่ยของประชากรตั้งแต่ 3 กลุ่มขึ้นไป โดยมีข้อสมมุติ (Assumptions) 3 ข้อ คือ (1) ข้อมูลเป็นอิสระกัน (2) ข้อมูลแต่ละกลุ่มมาจากประชากรที่มีการแจกแจงปรกติ และ (3) ข้อมูลทุกกลุ่มมีความแปรปรวนเท่ากัน (Homogeneity of variance) ในกรณีที่ข้อมูลแต่ละกลุ่มหรือบางกลุ่มมีความแปรปรวนแตกต่างกันจะมีผลกระทบอย่างมากต่อการวิเคราะห์ความแปรปรวน โดยอาจทำให้ผลการวิเคราะห์ที่ได้ไม่ถูกต้อง [1]

จากงานวิจัยที่ผ่านมาได้มีผู้นำเสนอและศึกษาเกี่ยวกับการทดสอบความเท่ากันของความแปรปรวนของประชากรอยู่หลายวิธี เช่น การทดสอบบาร์ตเลตต์ (Bartlett test) การทดสอบเลวีนที่อิงค่าเฉลี่ย (Mean-based Levene test) การทดสอบเลวีนที่อิงค่ามัธยฐาน (Median-based Levene test) การทดสอบเลวีนที่อิงค่าเฉลี่ยแบบตัดปลาย

*ผู้นำเสนอ (vanida@buu.ac.th)

(Trimmed mean Levene test) การทดสอบเลวีนแบบไม่อิงพารามิเตอร์ (Non-parametric Levene test) การทดสอบค็อกแครน (Cochran test) การทดสอบ O'Brien (O'Brien test) การทดสอบ Fligner-Killeen (Fligner-Killeen test) โดยได้ศึกษาตามสถานการณ์ต่าง ๆ เช่น กรณีที่ข้อมูลมาจากประชากรที่มีการแจกแจงล็อกนอร์มอล (Log-normal distribution) [2] ข้อมูลมาจากประชากรที่มีการแจกแจงสมมาตรและไม่สมมาตรแบบต่าง ๆ [3] และข้อมูลมาจากประชากรที่มีความเบ้และความโด่งต่างกัน [4]

นอกจากข้อมูลที่นำมาศึกษาจะมีการแจกแจงความน่าจะเป็นแบบต่าง ๆ แล้ว ยังมีข้อมูลอีกลักษณะหนึ่งคือข้อมูลมีมีค่านอกเกณฑ์ (outlier) ในชุดข้อมูล โดยค่านอกเกณฑ์จะมีผลกระทบที่ทำให้การวิเคราะห์เชิงสถิติไม่ว่าจะเป็นการทดสอบสมมุติฐานหรือการประมาณค่าจะเกิดความผิดพลาดได้ง่าย ดังนั้นในการวิเคราะห์ข้อมูลที่มีค่านอกเกณฑ์ในชุดข้อมูลจึงต้องใช้ความระมัดระวังในการจัดการค่านอกเกณฑ์หรือพิจารณาเลือกวิธีเชิงสถิติที่เหมาะสมกับข้อมูลที่มีค่านอกเกณฑ์ [5]

จากงานวิจัยที่ผ่านมาจะเห็นได้ว่ามีผู้ศึกษาประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวนของประชากรเมื่อข้อมูลมีการแจกแจงความน่าจะเป็นชนิดต่าง ๆ แต่ยังไม่มีการศึกษาในกรณีที่ข้อมูลมีค่านอกเกณฑ์ในชุดข้อมูล ดังนั้นในการศึกษานี้จึงมีวัตถุประสงค์ที่จะศึกษาและเปรียบเทียบวิธีการทดสอบความเท่ากันของความแปรปรวนของประชากร 3 กลุ่ม เมื่อมีค่านอกเกณฑ์ในชุดข้อมูล เพื่อหาการทดสอบที่มีประสิทธิภาพและเหมาะสมที่จะนำมาใช้กับข้อมูลประเภทนี้ โดยการทดสอบที่นำมาศึกษามี 7 วิธี ได้แก่ การทดสอบบาร์ตเลตต์ การทดสอบเลวีนที่อิงตามค่าเฉลี่ย การทดสอบเลวีนที่อิงตามค่ามัธยฐาน การทดสอบเลวีนแบบไม่อิงพารามิเตอร์ การทดสอบค็อกแครน การทดสอบ Fligner-Killeen และการทดสอบ O'Brien

2 ความรู้พื้นฐาน

2.1 การทดสอบความเท่ากันของความแปรปรวน

การทดสอบความเท่ากันของความแปรปรวนมีสมมุติฐานที่ทดสอบคือ

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2,$$
$$H_a : \text{มี } \sigma_j^2 \text{ แตกต่างกันอย่างน้อย 2 ค่า}$$

เมื่อ k คือจำนวนกลุ่มของประชากร และ $j = 1, 2, \ldots, k$ โดยในการศึกษานี้ได้นำการทดสอบที่นิยมใช้มาศึกษาจำนวน 7 วิธี ได้แก่ การทดสอบบาร์ตเลตต์ การทดสอบเลวีนที่อิงตามค่าเฉลี่ย การทดสอบเลวีนที่อิงตามค่ามัธยฐาน การทดสอบเลวีนแบบไม่อิงพารามิเตอร์ การทดสอบค็อกแครน การทดสอบ Fligner-Killeen และการทดสอบ O'Brien

2.1.1 การทดสอบบาร์ตเลตต์

ในปี 1937 Bartlett ได้นำเสนอการทดสอบบาร์ตเลตต์ที่ใช้การทดสอบไคกำลังสอง (chi-square test) ในการทดสอบความเท่ากันของความแปรปรวน [6] โดยมีสถิติทดสอบคือ

$$B = \frac{(N-k)\log\left[\frac{\sum_{j=1}^{k}(n_j-1)S_j^2}{N-k}\right] - \sum_{j=1}^{k}(n_j-1)\log(S_j^2)}{1+\frac{\left[\sum_{j=1}^{k}\frac{1}{n_j-1}\right]-\frac{1}{N-k}}{3(k-1)}} \tag{2.1}$$

เมื่อ N คือจำนวนข้อมูลทั้งหมด n_j คือจำนวนข้อมูลในกลุ่มที่ j k คือจำนวนกลุ่ม และ S_j^2 คือความแปรปรวนของข้อมูลในกลุ่มที่ j

ปฏิเสธสมมุติฐานหลักเมื่อ $B > \chi^2_{\alpha,k-1}$ โดยที่ $\chi^2_{\alpha,k-1}$ คือควอนไทล์ที่ α ของการแจกแจงไคกำลังสองที่องศาเสรี $k-1$

2.1.2 การทดสอบเลวีนที่อิงค่าเฉลี่ย

ในปี 1960 Levene ได้นำเสนอการทดสอบเลวีนที่อิงค่าเฉลี่ยโดยใช้ค่าสัมบูรณ์ของส่วนเหลือ (residual) $Z_{ij} = |y_{ij} - \bar{y}_{.j}|$ ในการวิเคราะห์ความแปรปรวน ซึ่งเป็นการแปลงรูปจากการทดสอบค่าเฉลี่ยเป็นการทดสอบความแปรปรวน ทำให้ได้การทดสอบที่มีความแกร่งต่อข้อสมมุติเกี่ยวกับการแจกแจงปรกติของข้อมูล [6] โดยมีสถิติทดสอบคือ

$$L = \frac{(N-k)\sum_{j=1}^{k} n_j(\bar{Z}_{.j} - \bar{Z}_{..})^2}{(k-1)\sum_{j=1}^{k}\sum_{i=1}^{n_j}(Z_{ij} - \bar{Z}_{.j})^2} \tag{2.2}$$

เมื่อ y_{ij} คือค่าสังเกตที่ i ในกลุ่มที่ j $\bar{y}_{.j}$ คือค่าเฉลี่ยของข้อมูลในกลุ่มที่ j $\bar{Z}_{.j}$ คือค่าเฉลี่ยของ Z_{ij} ในกลุ่มที่ j และ $\bar{Z}_{..}$ คือค่าเฉลี่ยรวมของข้อมูลมั้งหมด

ปฏิเสธสมมุติฐานหลักเมื่อ $L > F_{\alpha,k-1,N-k}$ โดยที่ $F_{\alpha,k-1,N-k}$ คือควอนไทล์ที่ α ของการแจกแจงเอฟที่องศาเสรี $k-1$ และ $N-k$

2.1.3 การทดสอบเลวีนที่อิงค่ามัธยฐาน

การทดสอบเลวีนที่อิงค่ามัธยฐานหรือการทดสอบ Brown-Forsyte นำเสนอโดย Brown และ Forsythe ในปี 1974 โดยการทดสอบนี้ใช้หลักการเดียวกันกับการทดสอบเลวีนที่อิงค่าเฉลี่ยแต่ใช้ค่ามัธยฐานแทนค่าเฉลี่ยนั่นคือ $Z_{ij} = |y_{ij} - \tilde{y}_{.j}|$ [6] มีสถิติทดสอบคือ

$$BF = \frac{(N-k)\sum_{j=1}^{k} n_j(\bar{Z}_{.j} - \bar{Z}_{..})^2}{(k-1)\sum_{j=1}^{k}\sum_{i=1}^{n_j}(Z_{ij} - \bar{Z}_{.j})^2} \tag{2.3}$$

เมื่อ y_{ij} คือค่าสังเกตที่ i ในกลุ่มที่ j $\tilde{y}_{.j}$ คือค่าเมัธยฐานของข้อมูลในกลุ่มที่ j $\bar{Z}_{.j}$ คือค่าเฉลี่ยของ Z_{ij} ในกลุ่มที่ j และ $\bar{Z}_{..}$ คือค่าเฉลี่ยรวมของ Z_{ij} ทั้งหมด

ปฏิเสธสมมุติฐานหลักเมื่อ $BF > F_{\alpha,k-1,N-k}$ โดยที่ $F_{\alpha,k-1,N-k}$ คือควอนไทล์ที่ α ของการแจกแจงเอฟที่องศาเสรี $k-1$ และ $N-k$

2.1.4 การทดสอบเลวีนแบบไม่อิงพารามิเตอร์

Nordstokke และ Zumbo ได้นำเสนอการทดสอบเลวีนแบบไม่อิงพารามิเตอร์ในปี 2010 [7] ซึ่งปรับปรุงมาจากการทดสอบเลวีนที่อิงค่าเฉลี่ย โดยใช้การแปลงรูป y_{ij} ให้เป็นลำดับที่ R_{ij} และ $Z_{ij} = |R_{ij} - \bar{R}_{.j}|$ เมื่อ $\bar{R}_{.j}$ คือค่าเฉลี่ยของลำดับที่ในกลุ่มที่ j และมีสถิติทดสอบรูปแบบเดียวกับการทดสอบเลวีนที่อิงค่าเฉลี่ย คือ

$$NL = \frac{(N-k)\sum_{j=1}^{k} n_j(\bar{Z}_{.j} - \bar{Z}_{..})^2}{(k-1)\sum_{j=1}^{k}\sum_{i=1}^{n_j}(Z_{ij} - \bar{Z}_{.j})^2} \tag{2.4}$$

ปฏิเสธสมมุติฐานหลักเมื่อ $NL > F_{\alpha,k-1,N-k}$ โดยที่ $F_{\alpha,k-1,N-k}$ คือควอนไทล์ที่ α ของการแจกแจงเอฟที่องศาเสรี $k-1$ และ $N-k$

2.1.5 การทดสอบค็อกแครน

การทดสอบค็อกแครนเป็นการทดสอบที่มีวิธีการคำนวณค่าสถิติทดสอบที่ง่ายกว่าการทดสอบบาร์ตเลตต์ [8] มีสถิติทดสอบคือ

$$C = \frac{\text{largest } S_j^2}{\sum_{j=1}^{k} S_j^2} \tag{2.5}$$

เมื่อ S_j^2 คือความแปรปรวนของข้อมูลในกลุ่มที่ j และ largest S_j^2 คือค่า S_j^2 ที่มากที่สุด

ปฏิเสธสมมุติฐานหลักเมื่อ $C > \frac{1}{1+\frac{k-1}{F_{\alpha/k,n-1,(k-1)(n-1)}}}$ โดยที่ $F_{\alpha/k,n-1,(k-1)(n-1)}$ คือควอนไทล์ที่ α/k ของการแจกแจงเอฟที่องศาเสรี $n-1$ และ $(k-1)(n-1)$

2.1.6 การทดสอบ Fligner-Killeen

การทดสอบ Fligner-Killeen ได้รับการนำเสนอโดย Fligner และ Killeen ในปี 1976 [9] จากนั้น Conover, Johnson [10] ได้นำการทดสอบมาศึกษาและพบว่าการทดสอบนี้เป็นการทดสอบที่มีความแกร่งเมื่อข้อมูลไม่มีการแจกแจงปรกติ

การหาค่าสถิติทดสอบนี้ทำได้โดยนำข้อมูล y_{ij} มาแปลงรูปเป็น $Z_{ij} = |y_{ij} - \bar{y}_{.j}|$ และหาลำดับที่ของ Z_{ij} กำหนดให้เป็น R_{ij} จากนั้นนำมาหาค่าสกอร์ $a_{N,R_{ij}}$ จาก $a_{N,R_{ij}} = \Phi^{-1}(0.5 + 0.5R_{ij}/(N+1))$ และได้สถิติทดสอบคือ

$$FK = \frac{\sum_{j=1}^{k} n_j(\bar{A}_j - \bar{a})^2}{V^2} \tag{2.6}$$

เมื่อ $\bar{A}_j$ คือค่าเฉลี่ยของสกอร์ $a_{N,R_{ij}}$ ในกลุ่มที่ j $\bar{a}$ คือค่าเฉลี่ยรวมของ $a_{N,R_{ij}}$ และ V^2 ความแปรปรวนของ $a_{N,R_{ij}}$

ปฏิเสธสมมุติฐานหลักเมื่อ $FK > \chi^2_{\alpha,k-1}$ โดยที่ $\chi^2_{\alpha,k-1}$ คือควอนไทล์ที่ α ของการแจกแจงไคกำลังสองที่องศาเสรี $k-1$

2.1.7 การทดสอบ O'Brien

ในปี 1979 O'Brien ได้นำเสนอการทดสอบ O'Brien ที่ใช้การแปลงรูปข้อมูล y_{ij} ให้เป็น V_{ij} [11] ดังนี้

$$V_{ij} = \frac{(n_j - 1.5)n_j(y_{ij} - \bar{y}_j)^2 - 0.5S_j^2(n_j - 1)}{(n_j - 1)(n_j - 2)} \tag{2.7}$$

และสถิติทดสอบของการทดสอบนี้คือ

$$OB = \frac{(N-k)\sum_{j=1}^{k} n_j(\bar{V}_j - \bar{V})^2}{(k-1)\sum_{j=1}^{k}\sum_{i=1}^{n_j}(V_{ij} - \bar{V}_j)^2} \tag{2.8}$$

เมื่อ $\bar{V}_j$ ค่าเฉลี่ยของ V_{ij} ในกลุ่มที่ j และ $\bar{V}$ คือค่าเฉลี่ยรวมของ V_{ij}

ปฏิเสธสมมุติฐานหลักเมื่อ $OB > \chi^2_{\alpha,k-1}$ โดยที่ $\chi^2_{\alpha,k-1}$ คือควอนไทล์ที่ α ของการแจกแจงไคกำลังสองที่องศาเสรี $k-1$

2.2 ค่านอกเกณฑ์

ค่านอกเกณฑ์คือจุดข้อมูล (data point) ที่มีค่าห่างหรือแตกต่างจากจุดข้อมูลอื่นในชุดข้อมูลมาก ๆ และอาจส่งผลเสียต่อการวิเคราะห์เชิงสถิติได้หลายประการ ได้แก่ 1. ทำให้ความแปรปรวนของความคลาดเคลื่อนเพิ่มขึ้น และลดกำลังการทดสอบของการทดสอบสมมุติฐานเชิงสถิติ 2. หากค่านอกเกณฑ์ไม่ได้เกิดขึ้นโดยสุ่ม ค่านอกเกณฑ์นี้สามารถทำให้ชุดข้อมูลสูญเสียลักษณะของการแจกแจงปรกติไปได้ และทำให้ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 และ 2 เปลี่ยนแปลงไปจากที่ควรเป็น 3. ค่านอกเกณฑ์อาจทำให้เกิดการประมาณค่าที่เอนเอียงได้ [5] ในกรณีที่มีค่านอกเกณฑ์ในชุดข้อมูล โดยปรกติจะไม่สามารถตัดค่านี้ออกจากชุดข้อมูลได้เนื่องจากค่านี้เป็นค่าที่แท้จริงของมูลไม่ได้เกิดจากการผิดพลาดในการเก็บหรือบันทึกข้อมูล และมักเป็นค่าที่ให้สารสำคัญเกี่ยวกับข้อมูลชุดนี้

ในการวิเคราะห์ความแปรปรวน การที่มีค่านอกเกณฑ์ในชุดข้อมูลอาจส่งผลให้เกิดการละเมิดข้อสมมุติของการวิเคราะห์ความแปรปรวน นั่นคือข้อมูลไม่มีการแจกแจงปรกติและความแปรปรวนแตกต่างกัน ดังนั้น ในการวิเคราะห์ข้อมูลจึงต้องมีการจัดการเกี่ยวกับค่านอกเกณฑ์นี้ วิธีหนึ่งก็คือการหาวิธีเชิงสถิติที่มีความแกร่งต่อค่านอกเกณฑ์ในชุดข้อมูล นั่นคือวิธีเชิงสถิติที่ยังคงมีประสิทธิภาพและเชื่อถือได้

3 วิธีการศึกษา

การศึกษานี้ใช้การจำลองมอนติคาร์โล (Monte Carlo simulation) และโปรแกรม R โดยมีรายละเอียดดังนี้

3.1 สถานการณ์ที่ศึกษา

กำหนดสถานการณ์ที่ศึกษาดังนี้

1. จำนวนกลุ่มหรือจำนวนประชากรที่ศึกษา 3 กลุ่ม ($k = 3$)
2. ขนาดตัวอย่างในแต่ละกลุ่ม (n_1, n_2, n_3) มีทั้งกรณีที่เท่ากันคือ (10,10,10) (20,20,20) (30,30,30) (50,50,50) และ (100,100,100) และกรณีที่ขนาดตัวอย่างต่างกัน คือ (10,10,20) (10,10,40) (10,10,60) (20,20,40) (20,20,60) (20,20,80) (30,30,60) และ (50,50,100)
3. อัตราส่วนความแปรปรวน $(\sigma_1^2 : \sigma_2^2 : \sigma_3^2)$ คือ (1:1:1) (1:1:2) (1:1:4) และ (1:2:4) ในกรณีที่ขนาดตัวอย่างเท่ากันทั้ง 3 กลุ่ม และ (1:1:1) (1:1:2) (2:1:1) (1:1:4) (4:1:1) (1:2:4) และ (4:2:1) ในกรณีที่ขนาดตัวอย่างแตกต่างกัน

4. การแจกแจงของประชากรที่ศึกษา คือการแจกแจงปรกติมาตรฐาน (N(0,1))

5. ค่านอกเกณฑ์ 2 ชนิด คือ 1. ค่านอกเกณฑ์ไม่รุนแรง (mild outlier) และ 2. ค่านอกเกณฑ์สุดขีด (extreme outlier) จำนวนร้อยละ 10 ของขนาดตัวอย่างแต่ละกลุ่ม

6. ระดับนัยสำคัญ (α) เท่ากับ 0.05

7. ทำซ้ำในแต่ละสถานการณ์ 10,000 ครั้ง

8. เกณฑ์ในการพิจารณาประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวน 2 ชนิด คือ

   - ความสามารถในการควบคุมความผิดพลาดแบบที่ 1 (ความแปรปรวนของประชากรเท่ากันทุกกลุ่ม) ซึ่งจะพิจารณาจากค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ($\hat{\alpha}$) โดยใช้เกณฑ์ของ Bradley [12] คือ เมื่อ $\alpha = 0.05$ แล้ว $\hat{\alpha}$ ต้องมีค่าระหว่าง (0.025, 0.075) จึงจะถือว่าการทดสอบนั้นสามารถควบคุมความผิดพลาดแบบที่ 1 ได้
   - ค่าประมาณกำลังการทดสอบ $(\widehat{1-\beta})$ (ความแปรปรวนของประชากรแตกต่างกันอย่างน้อย 1 กลุ่ม) จะพิจารณาเฉพาะการทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้เท่านั้น และการทดสอบที่มีค่าประมาณกำลังการทดสอบมากที่สุดจะถือว่าเป็นการทดสอบที่มีประสิทธิภาพที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ

3.2 ขั้นตอนการศึกษา

ในแต่ละสถานการณ์มีขั้นตอนการศึกษา ดังนี้

1. สร้างข้อมูลตัวอย่างแต่ละกลุ่มจากการแจกแจงปรกติมาตรฐานจำนวน $n_j - r_j$ เมื่อ n_j คือขนาดตัวอย่างกลุ่มที่ j และ $r_j = 0.10 \times n_j$ คือจำนวนค่านอกเกณฑ์ในกลุ่มที่ j เมื่อ $j = 1, 2, 3$

2. สร้างค่านอกเกณฑ์มีขั้นตอนดังนี้

   (a) นำข้อมูล $n_j - r_j$ ของแต่ละกลุ่มมาหาค่าควอร์ไทล์ที่ 1 3 และพิสัยควอร์ไทล์ (Q_1, Q_3, IQR)

   (b) สุ่มค่านอกเกณฑ์จำนวน r_j โดย

   - หากเป็นค่านอกเกณฑ์ไม่รุนแรงจะสุ่มค่านอกเกณฑ์จาก $[Q3+1.55\times IQR, Q_3+3\times IQR]$ หรือ $[Q1 - 3 \times IQR, Q_1 - 1.55 \times IQR]$
   - หากเป็นค่านอกเกณฑ์รุนแรงจะสุ่มค่านอกเกณฑ์จาก $[Q3 + 3.5 \times IQR, Q_3 + 7 \times IQR]$ หรือ $[Q1 - 7 \times IQR, Q_1 - 3.5 \times IQR]$

   (c) นำค่านอกเกณฑ์ r_j ค่า ของแต่ละกลุ่มไปรวมกับข้อมูล $n_j - r_j$ จะได้ข้อมูลจำนวน n_j ค่า ในแต่ละกลุ่ม จากนั้นใช้ Box plot ตรวจสอบว่าค่านอกเกณฑ์ที่สร้างขึ้นเป็นค่านอกเกณฑ์ในชุดข้อมูลหรือไม่ หากไม่จะทำขั้นตอนการสร้างค่านอกเกณฑ์ใหม่จนกว่าจะได้ค่านอกเกณฑ์ตามที่ต้องการ

3. นำข้อมูลทั้ง 3 กลุ่ม มาทดสอบความเท่ากันของความแปรปรวนด้วยการทดสอบ 7 วิธี คือ การทดสอบบาร์ตเลตต์ การทดสอบเลวีนที่อิงค่าเฉลี่ย การทดสอบเลวีนที่อิงค่ามัธยฐาน การทดสอบเลวีนแบบไม่อิงพารามิเตอร์ การทดสอบค็อกแครน การทดสอบ Fligner-Killeen และการทดสอบ O'Brien

4. ทำซ้ำขั้นตอนที่ 1 - 3 จำนวน 10,000 ครั้ง บันทึกจำนวนครั้งที่ปฏิเสธสมมุติฐานหลัก

5. คำนวณค่า $\hat{\alpha}$ ในกรณีที่ความแปรปรวนของประชากร 3 กลุ่ม เท่ากัน และ $\widehat{1-\beta}$ ในกรณีที่ความแปรปรวนของประชากรแตกต่างกันอย่างน้อย 1 กลุ่ม จาก

$$\hat{\alpha} = \frac{\text{จำนวนครั้งที่ปฏิเสธ } H_0 | H_0 \text{ เป็นจริง}}{10,000} \tag{3.1}$$

และ

$$\widehat{1-\beta} = \frac{\text{จำนวนครั้งที่ปฏิเสธ } H_0 | H_0 \text{ เป็นเท็จ}}{10,000} \tag{3.2}$$

6. พิจารณาประสิทธิภาพของการทดสอบทั้ง 7 วิธี

4 ผลการศึกษา

การศึกษาประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวนทั้ง 7 วิธี ได้ผลการศึกษาดังนี้

4.1 ความสามารถในการควบคุมความผิดพลาดแบบที่ 1

ตารางที่ 1 และ 2 แสดงค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในกรณีของค่านอกเกณฑ์ไม่รุนแรงและค่านอกเกณฑ์สุดขีด เมื่อพิจารณาตามเกณฑ์ของ Bradley พบว่า

ในกรณีของค่านอกเกณฑ์ไม่รุนแรงการทดสอบเลวีนแบบไม่อิงพารามิเตอร์ (NL) สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ที่ทุกขนาดตัวอย่าง และการทดสอบ Fligner-Killeen สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ที่ขนาดตัวอย่าง (100,100,100) และ (50,50,100) ส่วนการทดสอบ O'Brien สามารถควบคุมความผิดพลาดแบบที่ 1 ได้เฉพาะที่ขนาดตัวอย่าง (100,100,100) เท่านั้น ในขณะที่การทดสอบชนิดอื่นไม่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้เลย

ในกรณีของค่านอกเกณฑ์สุดขีด เมื่อขนาดตัวอย่างเท่ากันทุกกลุ่มมีเพียงการทดสอบ Fligner-Killeen ที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ที่ขนาดตัวอย่าง (100,100,100) และเมื่อขนาดตัวอย่างแตกต่างกันบางกลุ่มมีเพียงการทดสอบเลวีนแบบไม่อิงพารามิเตอร์ที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ที่ขนาดตัวอย่าง (10,10,20) (10,10,40) (10,10,60) (20,20,60) และ (20,20,80) ในขณะที่การทดสอบอื่น ๆ ไม่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้เลย

ตารางที่ 1: ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในกรณีค่านอกเกณฑ์ไม่รุนแรง

$\sigma_1^2 : \sigma_2^2 : \sigma_3^2$	n_1, n_2, n_3	B	L	BF	NL	FK	C	OB
1:1:1	10,10,10	0.1442	0.0127	0.0047	0.0344*	0.0177	0.1335	0.0004
	20,20,20	0.1613	0.0188	0.0087	0.0284*	0.0152	0.1427	0.0013
	30,30,30	0.1756	0.0210	0.0144	0.0275*	0.0171	0.1585	0.0055
	50,50,50	0.1959	0.0225	0.0176	0.0275*	0.0233	0.1742	0.0160
	100,100,100	0.1928	0.0237	0.0202	0.0284*	0.0418*	0.1690	0.0285*
	10,10,20	0.1444	0.0145	0.0050	0.0305*	0.0149	0.1618	0.0004
	10,10,40	0.1443	0.0160	0.0084	0.0376*	0.0178	0.2511	0.0063
	10,10,60	0.1532	0.0186	0.0108	0.0465*	0.0237	0.3270	0.0192
	20,20,40	0.1664	0.0186	0.0117	0.0283*	0.0152	0.1742	0.0056
	20,20,60	0.1734	0.0171	0.0120	0.0303*	0.0170	0.2241	0.0109
	20,20,80	0.1689	0.0185	0.0121	0.0292*	0.0173	0.2629	0.0177
	30,30,60	0.1842	0.0214	0.0156	0.0270*	0.0186	0.1941	0.0142
	50,50,100	0.1936	0.0205	0.0181	0.0271*	0.0261*	0.2006	0.0241

* หมายถึงการทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

ตารางที่ 2: ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในกรณีค่านอกเกณฑ์สุดขีด

$\sigma_1^2 : \sigma_2^2 : \sigma_3^2$	n_1, n_2, n_3	B	L	BF	NL	FK	C	OB
1:1:1	10,10,10	0.2960	0.0010	0.0000	0.0243	0.0052	0.2590	0.0000
	20,20,20	0.3227	0.0017	0.0001	0.0220	0.0057	0.2778	0.0000
	30,30,30	0.3350	0.0025	0.0005	0.0209	0.0096	0.2831	0.0001
	50,50,50	0.3634	0.0014	0.0006	0.0206	0.0132	0.3073	0.0017
	100,100,100	0.3574	0.0015	0.0009	0.0236	0.0262*	0.3106	0.0063
	10,10,20	0.3148	0.0013	0.0000	0.0258*	0.0053	0.3033	0.0000
	10,10,40	0.3185	0.0034	0.0001	0.0303*	0.0088	0.4062	0.0000
	10,10,60	0.3251	0.0063	0.0007	0.0331*	0.0093	0.5016	0.0038
	20,20,40	0.3245	0.0027	0.0000	0.0221	0.0088	0.3134	0.0002
	20,20,60	0.3383	0.0033	0.0001	0.0255*	0.0093	0.3647	0.0017
	20,20,80	0.3413	0.0040	0.0005	0.0262*	0.0102	0.4206	0.0051
	30,30,60	0.3525	0.0025	0.0006	0.0224	0.0101	0.3334	0.0015
	50,50,100	0.3619	0.0023	0.0009	0.0222	0.0176	0.3451	0.0061

* หมายถึงการทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

4.2 ค่าประมาณกำลังการทดสอบ

ในการพิจารณาค่าประมาณกำลังการทดสอบจะพิจารณาเฉพาะในกรณีที่การทดสอบนั้นสามารถควบคุมความผิดพลาดแบบที่ 1 ได้เท่านั้น ตารางที่ 3 และ 4 แสดงค่าประมาณกำลังการทดสอบในกรณีของค่านอกเกณฑ์ไม่รุนแรง และตารางที่ 5 และ 6 แสดงค่าประมาณกำลังการทดสอบในกรณีของค่านอกเกณฑ์สุดขีด ได้ผลการศึกษาดังนี้

สำหรับค่านอกเกณฑ์ไม่รุนแรงเมื่อตัวอย่างมีขนาดเท่ากันทุกกลุ่ม จากตารางที่ 3 จะเห็นได้ว่าการทดสอบ Fligner-Killeen มีค่าประมาณกำลังการทดสอบสูงที่สุดที่ขนาดตัวอย่าง (100,100,100) ส่วนที่ขนาดตัวอย่างอื่น ๆ แสดงค่า

ประมาณกำลังการทดสอบเฉพาะของการทดสอบเลวีนแบบไม่อิงพารามิเตอร์เท่านั้น เนื่องจากเป็นการทดสอบชนิดเดียวที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ ในกรณีขนาดตัวอย่างแตกต่างกันบางกลุ่ม ที่ขนาดตัวอย่าง (50,50,100) การทดสอบ Fligner-Killeen มีค่าประมาณกำลังการทดสอบสูงที่สุด และมีค่าเพิ่มขึ้นเมื่ออัตราส่วนความแปรปรวนเพิ่มขึ้น ดังแสดงในตารางที่ 4

ตารางที่ 3: ค่าประมาณกำลังการทดสอบในกรณีค่านอกเกณฑ์ไม่รุนแรงและขนาดตัวอย่างเท่ากันทุกกลุ่ม

$\sigma_1^2 : \sigma_2^2 : \sigma_3^2$	n_1, n_2, n_3	NL	FK	OB
1:1:2	10,10,10	0.0731	-	-
	20,20,20	0.1258	-	-
	30,30,30	0.1896	-	-
	50,50,50	0.3328	-	-
	100,100,100	0.6516	0.7871*	0.7328
1:1:4	10,10,10	0.2216	-	-
	20,20,20	0.4651	-	-
	30,30,30	0.6804	-	-
	50,50,50	0.9144	-	-
	100,100,100	0.9987	0.9997*	0.9995
1:2:4	10,10,10	0.1649	-	-
	20,20,20	0.3553	-	-
	30,30,30	0.5570	-	-
	50,50,50	0.8293	-	-
	100,100,100	0.9938	0.9984*	0.9957

* หมายถึงการทดสอบที่มีค่าประมาณกำลังการทดสอบมากที่สุด

ตารางที่ 4: ค่าประมาณกำลังการทดสอบในกรณีค่านอกเกณฑ์ไม่รุนแรงและขนาดตัวอย่างต่างกันบางกลุ่ม

$\sigma_1^2 : \sigma_2^2 : \sigma_3^2$	n_1, n_2, n_3	NL	FK
1:1:2	50,50,100	0.5457	0.6184*
2:1:1	50,50,100	0.3589	0.4849*
1:1:4	50,50,100	0.9941	0.9983*
4:1:1	50,50,100	0.9301	0.9828*
1:2:4	50,50,100	0.9569	0.9749*
4:2:1	50,50,100	0.9253	0.9719*

* หมายถึงการทดสอบที่มีค่าประมาณกำลังการทดสอบมากที่สุด

สำหรับกรณีของค่านอกเกณฑ์สุดขีด ตารางที่ 5 แสดงค่าประมาณกำลังการทดสอบของการทดสอบ Fligner-Killeen ที่ขนาดตัวอย่าง (100,100,100) จะเห็นได้ว่าค่าประมาณกำลังการทดสอบเพิ่มขึ้นเมื่ออัตราส่วนความแปรปรวนเพิ่มขึ้น ส่วนในตารางที่ 6 แสดงค่าประมาณกำลังการทดสอบของการทดอสอบเลวีนแบบไม่อิงพารามิเตอร์ เมื่อขนาดตัวอย่างแตกต่างกันบางกลุ่ม จะเห็นได้ว่าค่าประมาณกำลังการทดสอบจะเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นและอัตราส่วนความแปรปรวนมีความแตกต่างกันมากขึ้น

ตารางที่ 5: ค่าประมาณกำลังการทดสอบ Fligner-Killeen ในกรณีค่านอกเกณฑ์สุดขีดและขนาดตัวอย่างเท่ากันทุกกลุ่ม

n_1, n_2, n_3	$\sigma_1^2 : \sigma_2^2 : \sigma_3^2$		
	1:1:2	1:1:4	1:2:4
100,100,100	0.7222	0.9996	0.9964

ตารางที่ 6: ค่าประมาณกำลังการทดสอบของการทดสอบเลวีนไม่อิงพารามิเตอร์ในกรณีค่านอกเกณฑ์สุดขีดและขนาดตัวอย่างแตกต่างกันบางกลุ่ม

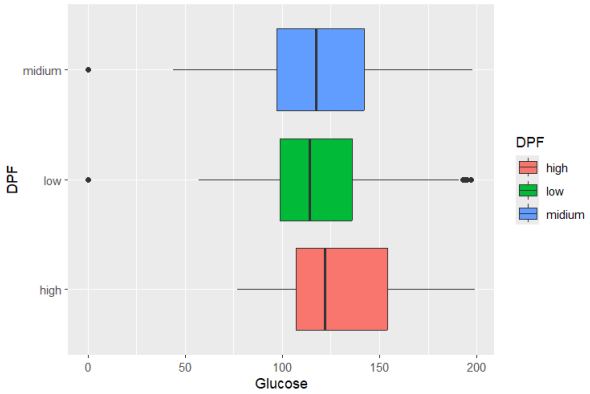
n_1, n_2, n_3	$\sigma_1^2 : \sigma_2^2 : \sigma_3^2$					
	1:1:2	2:1:1	1:1:4	4:1:1	1:2:4	4:2:1
10,10,20	0.0945	0.0525	0.3259	0.1771	0.2144	0.1473
10,10,40	0.1432	0.0555	0.5122	0.1832	0.3294	0.1822
10,10,60	0.1677	0.0558	0.6071	0.1934	0.3970	0.2027
20,20,60	0.2351	0.1105	0.8039	0.4480	0.5990	0.4456
20,20,80	0.2847	0.1168	0.8663	0.4526	0.6687	0.4889

5 ตัวอย่างข้อมูลจริง

ข้อมูลที่นำมาศึกษาเป็นข้อมูลโรคเบาหวานของชาวอินเดียนเผ่าพิมา (Pima Indians) จำนวน 768 คน ซึ่งเดิมมาจากสถาบันโรคเบาหวานและระบบย่อยอาหารและโรคไตแห่งชาติ (The National Institute of Diabetes and Digestive and Kidney Diseases) และสามารถดาวน์โหลดได้จาก https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database [13] ข้อมูลชุดนี้ประกอบด้วยตัวแปรเกี่ยวกับคุณลักษณะต่าง ๆ จำนวน 9 ตัวแปร ในการศึกษาได้นำมาศึกษา 2 ตัวแปร คือน้ำตาลในเลือด (glucose) และค่าฟังก์ชันสายพันธุกรรมเบาหวาน (diabetes pedigree function, DPF) โดยได้ทำการแบ่งค่า DPF เป็น 3 ระดับคือ 1. ต่ำ (≤ 0.300) 2. กลาง (0.301 - 0.800) 3. สูง (> 0.800)

นำข้อมูลมาหาค่าสถิติเบื้องต้นได้ดังตารางที่ 7 และแสดงการกระจายของข้อมูลด้วยแผนภาพกล่อง (box plot) ดังรูปที่ 1 ซึ่งจากตารางที่ 7 และแผนภาพกล่องจะเห็นได้ว่าข้อมูลชุดนี้มีขนาดใหญ่และมีค่านอกเกณฑ์อยู่ในชุดข้อมูล จากนั้นใช้การทดสอบความเท่ากันของความแปรปรวนทั้ง 7 วิธี ทดสอบความแปรปรวนของน้ำตาลในเลือดของคนที่มีค่า DPF แตกต่างกัน 3 กลุ่ม และได้ผลการทดสอบดังตารางที่ 8

จากตารางที่ 8 พบว่าการทดสอบบาร์ตเลตต์ การทดสอบเลวีนที่อิงค่ามัธยฐาน การทดสอบค็อกแครน และการทดสอบ O'Brien มีค่าพีสูงกว่าระดับนัยสำคัญ 0.05 จึงสรุปได้ว่าความแปรปรวนของประชากรไม่แตกต่างกัน ในขณะที่การทดสอบเลวีนที่อิงค่าเฉลี่ย การทดสอบเลวีนแบบไม่อิงพารามิเตอร์ และการทดสอบ Fligner-Killeen มีค่าพีน้อยกว่าระดับนัยสำคัญ 0.05 จึงสรุปได้ว่าความแปรปรวนของประชากรแตกต่างกัน



รูปที่ 1: แผนภาพกล่องของน้ำตาลในเลือดจำแนกตามค่า DPF

ตารางที่ 7: ค่าสถิติของน้ำตาลในเลือดจำแนกตาม DPF

DPF	จำนวนข้อมูล	ค่าเฉลี่ย	มัธยฐาน	ค่าเบี่ยงเบนมาตรฐาน
ต่ำ	305	118.016	114.0	29.898
กลาง	358	120.969	117.5	33.344
สูง	105	129.000	122.0	31.919

ตารางที่ 8: ผลการทดสอบความเท่ากันของความแปรปรวนของน้ำตาลในเลือด

	การทดสอบ						
	B	L	BF	NL	FK	C	OB
ค่าสถิติทดสอบ	3.870	3.011	2.931	3.225	6.370	0.368	1.466
ค่าพี	0.144	0.050*	0.054	0.040*	0.041*	0.237	0.231

* หมายถึงค่าพีน้อยกว่าหรือเท่ากับระดับนัยสำคัญ 0.05

6 สรุปผลและอภิปรายผล

การศึกษานี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบการทดสอบความเท่ากันของความแปรปรวนทั้งหมด 7 วิธี โดยพิจารณาจากความสามารถในการควบคุมความผิดพลาดแบบที่ 1 และค่าประมาณของกำลังการทดสอบ พบว่า ในกรณีของค่านอกเกณฑ์ไม่รุนแรงการทดสอบเลวีนไม่อิงพารามิเตอร์เป็นการทดสอบที่มีประสิทธิภาพมากที่สุดเมื่อขนาดตัวอย่างเท่ากันทุกกลุ่ม เนื่องจากสามารถควบคุมความผิดพลาดแบบที่ 1 ที่ทุกขนาดตัวอย่าง อย่างไรก็ตามเมื่อตัวอย่างมีขนาดใหญ่ (100,100,100) การทดสอบ Fligner-Killeen เป็นการทดสอบที่มีประสิทธิภาพมากที่สุด ในกรณีที่ขนาดตัวอย่างแตกต่างกันบางกลุ่ม การทดสอบเลวีนแบบไม่อิงพารามิเตอร์และการทดสอบ Fligner-Killeen จะมีประสิทธิภาพเมื่อตัวอย่างมีขนาดใหญ่เท่านั้น โดยการทดสอบ Fligner-Killeen จะให้ค่าประมาณกำลังการทดสอบสูงกว่าการทดสอบเลวีนแบบไม่อิงพารามิเตอร์ และมีค่าเพิ่มขึ้นเมื่ออัตราส่วนความแปรปรวรแตกต่างกันมากขึ้น

ในกรณีของค่านอกเกณฑ์สุดขีดมีเพียงการทดสอบ Fligner-Killeen ที่มีประสิทธิภาพเมื่อตัวอย่างมีขนาดเท่ากัน ในขณะที่การทดสอบเลวีนแบบไม่อิงพารามิเตอร์จะมีประสิทธิภาพเมื่อขนาดตัวอย่างแตกต่างกันบางกลุ่ม และค่าประมาณ

กำลังการทดสอบจะเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นและอัตราส่วนความแปรปรวนแตกต่างกันเพิ่มขึ้น

จากผลการศึกษาที่ได้จะเห็นได้ว่าในกรณีที่มีค่านอกเกณฑ์ในชุดข้อมูล การทดสอบที่มีความแกร่งต่อค่านอกเกณฑ์มากที่สุดจากทั้งหมด 7 วิธี ที่นำมาศึกษา คือการทดสอบเลวีนแบบไม่อิงพารามิเตอร์ เนื่องจากมีประสิทธิภาพดีทั้งในกรณีของค่านอกเกณฑ์ไม่รุนแรงและค่านอกเกณฑ์สุดขีด

เมื่อพิจารณาจากผลการศึกษาจากข้อมูลจริงจะเห็นได้ว่าการทดสอบเลวีนแบบไม่อิงพารามิเตอร์ และการทดสอบ Fligner-Killeen ให้ผลการทดสอบที่เหมือนกัน ซึ่งสอดคล้องกับผลการศึกษาจากการจำลองมอนติคาโลที่การทดสอบทั้งสองวิธีนี้จะมีประสิทธิภาพที่สุดเมื่อตัวอย่างมีขนาดใหญ่และมีค่านอกเกณฑ์ในชุดข้อมูล

เอกสารอ้างอิง

[1] M. Bland, An Introduction to Medical Statistics (4th ed.), Oxford University Press, Oxford, 2015.

[2] K. Sritan, B. Phuenaree, A comparison of efficiency for homogeneity of variance tests under log-normal distribution, Asian J. Appl. Sci. 9 (4) (2021) 254–259.

[3] A. Yonar, H. Yonar, M.Demirsöz, M. A. Tekindal, A comparative analysis for homogeneity of variance tests, J. Sci. Arts. 24 (2) (2024) 305–328.

[4] K. N. Baydili, D. Siğirli, Comparison of variance homogeneity tests for different distributions, Turkiye Klinikleri J. Biostat. 9 (3) (2017) 197–212.

[5] J. W. Osborne, A. Overbay, The power of outliers (and why researchers should always check for them), Pract. Assess. Res. Eval. 9 (6) (2004) 1–8.

[6] Y. Wang, P. Rodriguez de Gil, Y. Chen, J. Kromrey, E. Kim, T. Pham, D. Nguyen, J. Romano, Comparing the Performance of Approaches for Testing the Homogeneity of Variance Assumption in One-Factor, Educ. Psychol. Meas. 77 (2) (2016) 305–329.

[7] D. W. Nordstokke, B. D. Zumbo, A new nonparametric Levene test for equal variances, Psicologica. 31 (2010) 401–430.

[8] T. Vorapongsathorn, S. Taejaroenkul, C. Viwatwongkasem, A comparison of type I error and power of Bartlett's test, Levene's test and Cochran's test under violation of assumptions, Songklanakarin J. Sci. Technol. 26 (4) (2004) 537–547.

[9] M. A. Fligner, T. J. Killeen, Distribution-Free Two-Sample Tests for Scale, J. Amer. Statist. Assoc., 71 (1976) 210–213.

[10] W. J. Conover, A. E. Johnson, M. M. Johnson, A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data, Technometrics, 23 (4) (1981) 351–361.

[11] K. Soikliew, A. Araveeporn, Modifications of Levene's and O'Brien's Tests for Testing the Homogeneity of Variance Based on Median and Trimmed Mean, THAIL. STATIST., 16 (2) (2018) 106–128.

[12] J. V. Bradley, Robustness?, Br. J. Math. Stat. Psychol., 31 (1978) 144–152.

[13] UCI Machine Learning Repository, Pima Indians Diabetes Database. <https://www.kaggle.com/uciml/pima-indians-diabetes-database>, .

แอปพลิเคชันช่วยเลือกแผนประกันชีวิตและโรคร้ายแรง

ศุภณัฐ แหยมดี[1], ตรีทศ กิจวรรณ[1,*], อมรรัตน์ อ้วนไตร[1,†]
และ อัชฌา เฉลยจิตร์[1]

[1]โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย ลพบุรี 15120

บทคัดย่อ

ปัจจุบันคนหันมาลงทุนประกันชีวิตและประกันสุขภาพด้านโรคร้ายแรงมากขึ้น นำมาสู่แนวคิดที่จะสร้างแอปพลิเคชันเกี่ยวกับการวางแผนในการจ่ายเบี้ยประกันที่เหมาะสมกับผู้ที่สนใจจะทำประกันชีวิต และประกันโรคร้ายแรง โดยนำความรู้เรื่องการคิดเงินรวม ดอกเบี้ยเชิงเดี่ยว และดอกเบี้ยทบต้น มาหาสูตรคำนวณทุนประกันหรือจำนวนเงินเอาประกัน จากนั้นผู้จัดทำใช้เทคโนโลยีการพัฒนาแอปพลิเคชันด้วยการออกแบบอัลกอริทึม และ user interface ในเว็บไซต์สร้างแอปพลิเคชัน Thunkable แล้วนำแอปพลิเคชัน My Plan Insurance ที่พัฒนานี้ไปใช้บนระบบปฏิบัติการ android และบนเว็บแอปพลิเคชัน เพื่อตรวจสอบความพึงพอใจ(user experience) ของครูและบุคลากรสายสนับสนุนของโรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย ลพบุรี จำนวน 87 คน พบว่า ผลการศึกษาความพึงพอใจของผู้ใช้งานแอปพลิเคชัน My Plan Insurance ในภาพรวมอยู่ในระดับดีมากที่สุด อย่างมีนัยสำคัญที่ $\alpha = 0.05$ มีผลเฉลี่ยรวม $\bar{x}$ เท่ากับ 4.58 , S.D. เท่ากับ 0.51 เมื่อพิจารณารายข้อ ผู้ใช้งานคิดเห็นว่ารูปแบบของแอปพลิเคชันมีความน่าสนใจ อยู่ในระดับดีมากที่สุด ผู้ใช้งานคิดเห็นว่าข้อมูลที่ได้รับจากการประมวลผลมีความน่าเชื่อถือ ประมวลผลถูกต้อง อยู่ในระดับดีมากที่สุด ผู้ใช้งานคิดเห็นว่าข้อมูลมีความชัดเจน กระชับเข้าใจง่าย อยู่ในระดับดีมาก ผู้ใช้งานคิดเห็นว่าแอปพลิเคชันให้ข้อมูลที่เป็นประโยชน์ในการเลือกวางแผนประกันชีวิตและประกันโรคร้ายแรง อยู่ในระดับดีมากที่สุด ผู้ใช้งานคิดเห็นว่าการจัดหมวดหมู่ของข้อมูลในแอปพลิเคชันง่ายต่อการใช้งาน อยู่ในระดับดีมากที่สุด ผู้ใช้งานคิดเห็นว่าความรวดเร็วในการใช้งานแอปพลิเคชัน อยู่ในระดับดีมาก และผู้ใช้งานคิดเห็นว่าแอปพลิเคชันใช้งานง่าย ไม่ซับซ้อน อยู่ในระดับดีมาก

คำสำคัญ: ประกันชีวิต, ประกันสุขภาพ , แอปพลิเคชัน , ดอกเบี้ย , การเงิน

2020 MSC: 68-04, 40-00

*ผู้นำเสนอ (trithotkitchawan@gmail.com)

†ผู้ประพันธ์บรรณกิจ (amornrat@pccl.ac.th)

1 บทนำ

1.1 ที่มาของการวิจัย

ความเข้าใจในเรื่องของประกันสุขภาพกลายเป็นสิ่งสำคัญมากสำหรับทุกคน เนื่องจากการดูแลสุขภาพเป็นปัจจัยหลักที่ส่งผลต่อคุณภาพชีวิตและค่าใช้จ่ายในอนาคต การเลือกแผนประกันสุขภาพที่เหมาะสมกับสถานะทางการเงินและความเสี่ยงด้านสุขภาพ [1](Arrow, 1963, p.941-973) ของแต่ละคนจึงเป็นเรื่องที่ท้าทาย อย่างไรก็ตาม ผู้คนจำนวนมากยังขาดเครื่องมือหรือระบบที่ช่วยให้การเลือกแผนประกันสุขภาพเป็นเรื่องง่ายและตรงกับความต้องการจริงๆ [2](Brown, R. S., 2016, p.150-162)

ผลการสำรวจในไตรมาส 1 ปี 2567 พบว่า จากผลการสำรวจประชาชนอายุ 6 ปีขึ้นไป ประมาณ 65.9 ล้านคน พบว่า มีผู้ใช้อินเทอร์เน็ต 59.0 ล้านคน (89.5%) มีผู้ใช้โทรศัพท์มือถือ 62.9 ล้านคน (95.4%) และผู้มีโทรศัพท์มือถือ 58.1 ล้านคน (88.1%) [3](สำนักงานสถิติแห่งชาติ, 2024, p.จ) ในปัจจุบันอุปกรณ์พกพา ได้แก่ โทรศัพท์มือถือ และคอมพิวเตอร์แท็บเล็ต เป็นกลไกสำคัญที่ทำให้ผู้ใช้ได้เข้าถึงอินเทอร์เน็ตได้ โดยอุปกรณ์พกพาได้มีการพัฒนาให้มีความก้าวหน้า ทั้งทางด้าน Hardware (ตัวเครื่อง) และ Software หรือ เรียกว่าแอปพลิเคชันบนมือถือ ซึ่งให้ผู้ใช้สามารถ ใช้งานแอปพลิเคชันได้อย่างมีประสิทธิภาพ น่าสนใจ ตอบสนองความต้องการ ยืดหยุ่นและมีความหลากหลายมากขึ้น ในปัจจุบัน การใช้งานแอปพลิเคชันบนมือถือมีแนวโน้มเพิ่มขึ้นแบบก้าวกระโดด ส่งผลให้มีการพัฒนาแอปพลิเคชันบนมือถือให้มีประสิทธิภาพมากยิ่งขึ้น[4](วิสิทธิ บุญชุม,2020,p.33)

1.2 วัตถุประสงค์และประโยชน์

งานวิจัยนี้มีวัตถุประสงค์หลักเพื่อพัฒนาแอปพลิเคชันช่วยวางแผนประกันสุขภาพและโรคร้ายแรง โดยมีคำถามวิจัยดังนี้:

1. สูตรที่ใช้คำนวณทุนประกันชีวิตตามประเภทเงินเดือนของผู้ใช้งานมีรูปแบบทั่วไปอะไร?
2. เบี้ยประกันชีวิตที่ผู้ประกันตนต้องจ่ายที่คำนวณจากอัตรามรณะต้องจ่ายเท่าไร?
3. ผู้ใช้มีความพึงพอใจจากการใช้แอปพลิเคชันช่วยตัดสินใจในการเลือกประกันสุขภาพและโรคร้ายแรงหรือไม่?

ผลการวิจัยนี้จะเป็นประโยชน์ทั้งในวงการศึกษา วงการคณิตศาสตร์การเงิน วงการประกันสุขภาพ โดยแสดงให้เห็นถึงการประยุกต์ความรู้ทางการเงินและความรู้ทางวิทยาการคอมพิวเตอร์ในระดับชั้นมัธยมศึกษา โดยให้ข้อมูลเชิงประจักษ์เกี่ยวกับชิ้นงานแอปพลิเคชัน และความพึงพอใจของผู้ใช้งาน เป็นแนวทางในการพัฒนากับประกันชีวิตประเภทอื่น และสามารถนำไปประยุกต์ใช้ในการพัฒนาแอปพลิชันทางการเงินในบริบทอื่นๆ

2 ความรู้พื้นฐาน

2.1 ความหมายของประกันชีวิต

ความหมายของประกันชีวิตความหมายของ [5] (สำนักงานราชบัณฑิตยสภา, 2555) ประกันชีวิต หมายถึง สัญญาว่าจะจ่ายเงิน เมื่อบุคคลหรือกลุ่มบุคคลที่ทำประกันเสียชีวิต และอาจรวมการประกันอื่น ๆ เพิ่มเติมไว้ด้วย เช่น การประกันอุบัติเหตุและสูญเสียอวัยวะ การประกันกรณีทุพพลภาพ การประกันสุขภาพ ผู้ที่เกี่ยวข้องกับการประกันชีวิต ได้แก่ ผู้รับประกัน ผู้เอาประกัน และ ผู้รับประโยชน์ ผู้รับประกัน คือผู้ที่ต้องจ่ายเงินทดแทนเป็นจำนวนเงินตามที่สัญญาไว้ให้ผู้รับประโยชน์ หากผู้เอาประกันเสียชีวิตในระยะเวลาของการประกัน ส่วนผู้เอาประกันต้องระบุชื่อผู้รับประโยชน์ไว้ในสัญญาประกันชีวิต หรือกรมธรรม์ประกันชีวิต และต้องจ่ายเบี้ยประกันให้ผู้รับประกันอย่างสม่ำเสมอ การประกันชีวิตในปัจจุบันมีหลายรูปแบบ บางแบบเป็นเหมือนการลงทุนหรือการสะสมทรัพย์ ซึ่งผู้เอาประกันได้รับเงินปันผลและได้รับเงินที่สะสมคืนเมื่อครบกำหนดการประกันชีวิต

2.2 โครงสร้างประกันสุขภาพ

การซื้อประกันสุขภาพเป็นสัญญาเพิ่มเติมซึ่งต้องมีประกันชีวิตเป็นสัญญาหลักก่อนจึงจะซื้อประกันสุขภาพเพิ่มเติมเข้าไปได้ ผู้จัดทำเน้นไปที่สัญญาหลักเป็นสัญญาตลอดชีพ ประเภท 90/10, 90/15 และ 90/20
[6] (K WEALTH, 2567) ช่วงอายุที่สำรวจ 25-60 ปี และจำนวนเงินเอาประกันเป็น 300,000 บาท 500,000 บาท 1,000,000 บาท และ 1,500,000 บาทและ ประกันสุขภาพด้านโรคร้ายแรง

2.3 การคำนวณทุนประกัน

ทุนประกัน = รายได้ต่อปีคูณจำนวนปีที่ยังสามารถทำงานได้ – มูลค่าของจำนวนทรัพย์สินที่แปลงเป็นเงินสดได้

2.4 การคำนวณเบี้ยประกันชีวิต

การประกันตลอดชีพ ผู้ประกันต้องชำระเบี้ยประกันเป็นงวดๆ ไปตลอดชีพ การคำนวณเบี้ยประกันอัตราคงที่จึงต้องหามูลค่าปัจจุบันของเงินได้ประจำตลอดชีพ 1 บาท การคำนวณต้องคิดมูลค่าเงินปัจจุบันของเงิน 1 บาทก่อน ในเวลาของการจ่ายเบี้ยประกัน โดยใช้ตาราง Excel (ถ้าใช้อัตราดอกเบี้ย 3%) [7](ศ.ชูเกียรติ ประมูลผล)

2.5 การคำนวณเบี้ยประกันโรคร้ายแรง

ผู้จัดทำเลือกตัวแทนบริษัท AIA เป็นตัวแทนอัตราเบี้ยประกัน

$$\text{เบี้ยประกันรายปี} = \frac{\text{ทุนประกัน} \times \text{อัตราเบี้ยประกันที่จ่ายต่อ 1000}}{1{,}000}$$

2.6 เว็บไซต์ Thunkable ใช้สร้างแอปพลิเคชัน My plan insurance

[8] วโรดม คำแผ่นชัย (2561) Thunkable เป็นเครื่องมือสร้างโมบายแอปพลิเคชัน เพื่อติดตั้งบนสมาร์ตโฟนที่ใช้ระบบปฏิบัติการ Android, iOS โดยเครื่องมือที่ใช้ในการสร้างนั้น นอกจากเครื่องมือพื้นฐานแล้ว ยังมีการเชื่อมต่อไปยังผลิตภัณฑ์จาก Google, Twitter และ Microsoft โดยชุดคำสั่งหลังจากที่ออกแบบหน้าจอด้วยเครื่องมือต่าง ๆ Thunkable คือเว็บไซต์ที่ให้ เราสามารถสร้างโมบายแอปพลิเคชัน สวยๆ ใช้งานได้ และมีประโยชน์ ตามแนวคิด “Thunkable enables anyone to create beautiful and powerful mobile apps” ผู้สร้างโมบาย

แอปพลิเคชัน สามารถแสดงถึงแนวคิดการเขียนโปรแกรมเป็นบล็อก ได้โดยง่าย ด้วยภาษาอังกฤษอย่าง แต่ต้องมีการฝึกฝนเครื่องมือและชุดคำสั่งให้คล่องและปฏิบัติการด้าน โปรแกรมทดลองแยกออกเป็นส่วน ๆ

2.6 สมมติฐานการวิจัย

ผู้ใช้แอปพลิเคชันมีระดับความพึงพอใจภาพรวมมีคะแนนเฉลี่ยระดับมากขึ้นไป หรือ 4 ขึ้นไป ที่ระดับนัยสำคัญ 0.05

2.7 การวิเคราะห์ทางสถิติ

กำหนด

สมมติฐานว่าง (H_0) : $m \leq 4$

สมมติฐานทางเลือก (H_1): $m > 4$

ระดับนัยสำคัญ (Level of Significance - α) = 0.05

คำนวณค่าสถิติทดสอบ (Calculate Test Statistic) สำหรับประชาการ 1 กลุ่ม t-Test

สูตรการคำนวณค่าสถิติ t คือ

$$t = \frac{x - m_0}{S.D./\sqrt{n}}$$

3 ผลการศึกษา

3.1 สูตรที่ใช้คำนวณทุนประกันชีวิต

3.1.1 <u>กรณี เงินเดือนคงที่ตลอดทุกปี</u>

ทุนประกันที่ต้องการรับจากการทำสัญญาประกันชีวิต จะประมาณได้ดังนี้

ผู้ใช้ต้องระบุ อายุที่เกษียณ ให้เป็น AR

อายุที่เริ่มทำประกันเป็น AB

รายได้ต่อเดือนคงที่เป็น REV

มูลค่าทรัพย์สินที่แปลงเป็นเงินสดได้ เช่น เงินออมในกระปุก เงินที่ฝากไว้ธนาคาร หรือทรัพย์สินส่วนตัวที่สามารถขายได้ เช่น บ้าน รถ ที่ดิน ทองคำ เพชร เครื่องประดับ พระเครื่อง หนังสือ เสื้อผ้า กระเป๋า ของใช้มือสอง เป็น PRO

ได้สูตรว่า

ทุนประกัน = **{** (อายุที่เกษียณ – อายุเริ่มทำประกัน) × (รายได้ต่อเดือนคงที่×12) **}** – มูลค่าของจำนวนทรัพย์สินที่แปลงเป็นเงินสดได้) = **{** (AR - AB) × (REV×12) **}** – PRO

3.1.2 <u>กรณี เงินเดือนขึ้น r% ต่องวดภายใน 1 ปี</u>

ทุนประกันที่ต้องการรับจากการทำสัญญาประกันชีวิต จะประมาณได้ดังนี้

ผู้ใช้ต้องระบุ อัตราขึ้นเงินเดือน r% ต่องวดภายใน 1 ปี ให้เป็น r

อายุที่เกษียณ ให้เป็น AR

อายุที่เริ่มทำประกัน ให้เป็น AB

รายได้ปัจจุบัน ให้เป็น REVP

จำนวนงวดใน 1 ปี ให้เป็น PER

มูลค่าทรัพย์สินที่แปลงเป็นเงินสดได้ เช่น เงินออมในกระปุก เงินที่ฝากไว้ธนาคาร หรือทรัพย์สินส่วนตัวที่สามารถขายได้ เช่น บ้าน รถ ที่ดิน ทองคำ เพชร เครื่องประดับ พระเครื่อง หนังสือ เสื้อผ้า กระเป๋า ของใช้มือสอง เป็น PRO

ให้ เมื่อ t = จำนวนงวด = (อายุที่เกษียณ – อายุเริ่มทำประกัน)×จำนวนงวดใน 1 ปี

ดังนั้น t = (AR – AB) × PER

ได้สูตรว่า

$$\text{ทุนประกัน} = \{ (REVP\times \frac{12}{PER}) + (REVP(1+r\%)\times \frac{12}{PER}) + (REVP(1+r\%)^2\times \frac{12}{PER}) + (REVP(1+r\%)^3\times \frac{12}{PER}) + (REVP(1+r\%)^4\times \frac{12}{PER}) + (REVP(1+r\%)^5\times \frac{12}{PER}) + \cdots + (REVP(1+r\%)^t \times \frac{12}{PER}) \} - PRO$$

3.2 เบี้ยประกันชีวิตตลอดชีพ

ผู้จัดทำเลือกทุนประกัน หรือจำนวนเงินเอาประกันตลอดชีพวงเงิน 300,000 บาท 500,000 บาท 1,000,000 บาท หรือ 1,500,000 บาท โดยใช้ตารางมณรณะ

คำนวณเบี้ยประกันคงที่ : (ใช้ตารางมรณะเพศชายและหญิง 2560) [9] (สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย) เช่น

เพศชาย				
อายุ	l(x)	d(x)	e0(x)	1000q(x)
25	982,945	1,412	50.28	1.4368

เพศหญิง				
อายุ	l(x)	d(x)	e0(x)	1000q(x)
25	992,104	436	55.29	0.4392

l(x) หมายถึง จำนวนคนผู้ชาย/ผู้หญิงที่มีอายุอยู่ ณ อายุ 25 ปี $e^0(x)$ หมายถึง อายุขัยเฉลี่ยที่เหลืออยู่

d(x) หมายถึง จำนวนคนที่เสียชีวิตในช่วงอายุ 25 ก่อนถึงอายุ 26 ปี

และ 1000q(x) ความน่าจะเป็นที่คนที่มีอายุ 25 ปี จะเสียชีวิตก่อนที่จะมีอายุครบ 26 ปี ใน 1000 คน

การประกันตลอดชีพ ผู้ประกันต้องชำระเบี้ยประกันเป็นงวดๆ ไปตลอดชีพ การคำนวณเบี้ยประกันอัตราคงที่จึงต้องหามูลค่าปัจจุบันของเงินได้ประจำตลอดชีพ 1 บาท

การคำนวณต้องคิดมูลค่าเงินปัจจุบันของเงิน 1 บาทก่อน ในเวลาของการจ่ายเบี้ยประกัน โดยใช้ Google sheet (ผู้จัดทำสมมติใช้ค่าอัตราดอกเบี้ย 3%) [7](ศ.ชูเกียรติ ประมูลผล)

ตารางที่ 1 : ตารางแสดงมูลค่าปัจจุบันของเงิน 1 บาทเมื่อเวลาผ่านไป

ต้นปีที่	มูลค่าปัจจุบันของเงิน 1 บาท	ต้นปีที่	มูลค่าปัจจุบันของเงิน 1 บาท	ต้นปีที่	มูลค่าปัจจุบันของเงิน 1 บาท
1	1	25	0.491934	49	0.241999
2	0.970874	26	0.477606	50	0.23495
3	0.942596	27	0.463695	51	0.228107
4	0.915142	28	0.450189	52	0.221463
5	0.888487	29	0.437077	53	0.215013
6	0.862609	30	0.424346	54	0.20875
7	0.837484	31	0.411987	55	0.20267
8	0.813092	32	0.399987	56	0.196767
9	0.789409	33	0.388337	57	0.191036
10	0.766417	34	0.377026	58	0.185472
11	0.744094	35	0.366045	59	0.18007
12	0.722421	36	0.355383	60	0.174825
13	0.70138	37	0.345032	61	0.169733
14	0.680951	38	0.334983	62	0.164789
15	0.661118	39	0.325226	63	0.15999
16	0.641862	40	0.315754	64	0.15533
17	0.623167	41	0.306557	65	0.150806
18	0.605016	42	0.297628	66	0.146413
19	0.587395	43	0.288959	67	0.142149
20	0.570286	44	0.280543	68	0.138009
21	0.553676	45	0.272372	69	0.133989
22	0.537549	46	0.264439	70	0.130086
23	0.521893	47	0.256737	71	0.126297
24	0.506692	48	0.249259	72	0.122619

ต้นปีที่	มูลค่าปัจจุบันของเงิน 1 บาท	ต้นปีที่	มูลค่าปัจจุบันของเงิน 1 บาท
73	0.119047	87	0.078704
74	0.11558	88	0.076412
75	0.112214	89	0.074186
76	0.108945	90	0.072026
77	0.105772	91	0.069928
78	0.102691	92	0.067891
79	0.0997	93	0.065914
80	0.096796	94	0.063994
81	0.093977	95	0.06213
82	0.09124	96	0.06032
84	0.086002	98	0.056858
85	0.083497	99	0.055202
86	0.081065		

ผู้จัดทำใช้ตารางคำนวณเบี้ยประกันคงที่ของคนที่อายุ 25 ปี บน Google sheet (https://docs.google.com/spreadsheets/d/1VYMQZL5A3RM_UY7hKn_rPAR0YIK8erGg_qnOUbQjy5O/edit?usp=sharing) และเปรียบเทียบเบี้ยประกันชีวิตตลอดชีพ เลือกจ่าย 10 ปี, 15 ปี และ 20 ปี คุ้มครองถึงอายุ 90 ได้ดังนี้

ตารางที่ 2: เบี้ยประกันตลอดชีพ

	300,000 จ่ายปีละ		500,000 จ่ายปีละ		1,000,000 จ่ายปีละ		1,500,000 จ่ายปีละ	
จ่าย	เพศชาย	เพศหญิง	เพศชาย	เพศหญิง	เพศชาย	เพศหญิง	เพศชาย	เพศหญิง
10 ปี	7,104	5,761	11,840	9,601	23,681	19,203	35,521	28,804
15 ปี	4,736	3,841	7,894	6,401	15,787	12,802	23,681	19,203
20 ปี	3,552	2,880	5,920	4,801	11,840	9,601	17,760	14,402

จากตารางการคำนวณใน Google sheet ผู้เอาประกันสามารถพิจารณาเลือกแผนประกันชีวิตของตนเองตามวงเงินที่เหมาะสมได้

3.3 เบี้ยประกันโรคร้ายแรง เจอ จ่าย จบ

ผู้จัดทำศึกษาเบี้ยประกันโรคร้ายแรง เจอ จ่าย จบ เลือกคุ้มครองตามความต้องการ แต่ละบริษัทมีเบี้ยประกันที่แตกต่างกัน ผู้จัดทำศึกษาข้อมูลค่าใช้จ่ายต่างๆ จากการเข้ารักษาตัว ตามเงื่อนไขดังนี้

1. ประเภทของโรงพยาบาล
   1.1 โรงพยาบาลรัฐบาล ตั้งแต่ 70,000 บาทขึ้นไป
   1.2 โรงพยาบาลเอกชน ตั้งแต่ 1,500,000 บาทขึ้นไป
2. อายุผู้เอาประกันภัยอยู่ในช่วง 25-60 ปี หรือมากกว่า 60 ปี
3. การจ่ายและความคุ้มครอง
   3.1 จ่ายปีต่อปี
   3.2 จ่าย 10 ปี คุ้มครอง 15 ปี
   3.3 จ่าย 5 ปี คุ้มครอง 5 ปี
4. เจอ จ่าย
   4.1 จบ
   4.2 เคลมซ้ำได้
5. ผู้ทำประกันควรคำนึงถึงหลักการออม คือ ออม 1 ส่วน ใช้ 3 ส่วน ซึ่ง 3 ส่วนดังกล่าวส่วนนึงควรลงทุนทำประกัน โดยทำประกันสัญญาหลักและสัญญาเพิ่มเติมเกี่ยวกับสุขภาพ เช่นประกันภัยโรคร้ายแรง เป็นต้น
6. จากการคำนวณข้อมูล 2 บริษัทเปรียบเทียบกันคือ AIA CI PLUS กับ Big 3 บริษัท SCB Protect และใช้สูตรข้างล่างนี้ พบว่า อัตราเบี้ยประกันแตกต่างกัน

$$\text{เบี้ยประกันรายปี} = \frac{\text{ทุนประกัน} \times \text{อัตราเบี้ยประกันที่จ่ายต่อ 1000}}{1{,}000}$$

ผู้จัดทำจึงเลือกยกตัวอย่างการใช้แบบแผนการคำนวณเบี้ยประกันรายปีกรมธรรม์ประกันโรคร้ายแรง AIA CI PLUS ของบริษัท AIA เป็นตัวแทน มาคำนวณเบี้ยประกันที่ต้องจ่ายของเราเอง ซึ่งเป็นข้อมูลจากตัวแทนมาเป็นตารางข้างล่างนี้

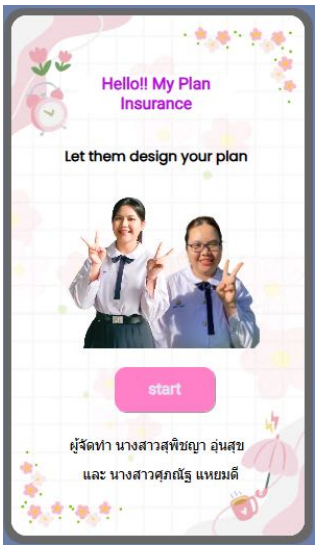
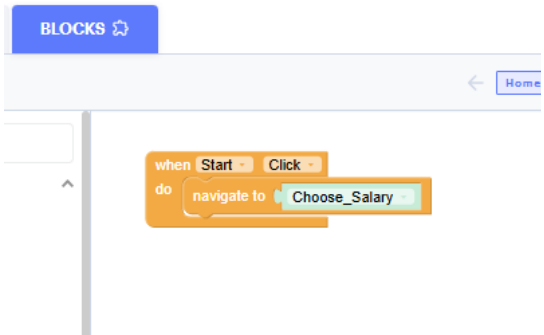
ตารางที่ 3 : อัตราเบี้ยประกันภัยมาตรฐานรายปี ต่อจำนวนเงินเอาประกันภัย 1,000 บาท

อายุ(ปี)	ชาย	หญิง
15 วัน-15 ปี	2.62	2.27
16-20	2.76	2.06

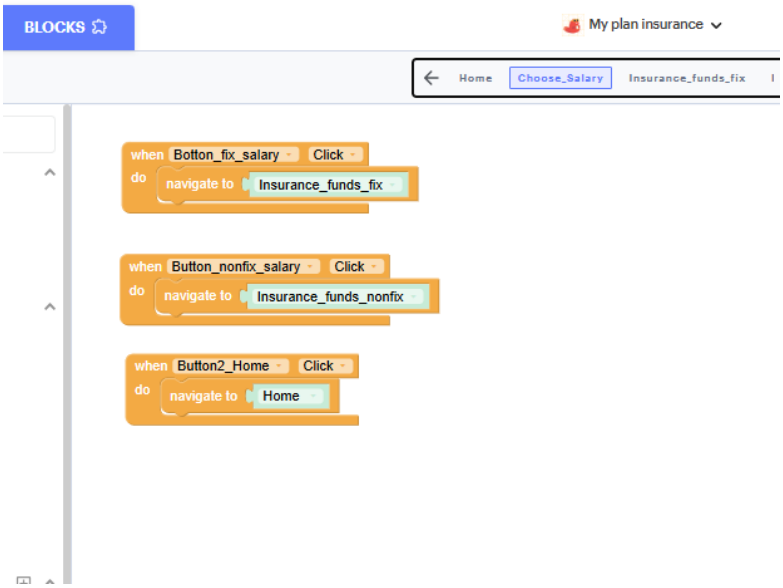
อายุ(ปี)	ชาย	หญิง
21-25	2.89	2.15
26-30	3.16	2.55
31-35	3.84	3.53
36-40	5.23	4.60
41-45	10.19	6.38
46-50	15.92	10.17
51-55	25.38	15.92
56-60	42.31	26.09
61-65	61.04	35.81
66-70	81.76	51.71
71-75 (ปีต่ออายุ)	112.64	73.52
76-79 (ปีต่ออายุ)	142.38	104.55

3.4 Workflow การทำงานของ My plan insurance

1. หน้า Home ของแอพพลิเคชัน My Plan Insurance กดปุ่ม Start

User Interface	Block คำสั่ง
	

2. กดเลือกเงินเดือนที่ได้รับ แบบคงที่ทุกเดือน และแบบขึ้นเป็นเปอร์เซ็นเท่าๆ กันต่องวด ใน 1 ปี

User Interface	Block คำสั่ง
	

3. ถ้า Click ปุ่มเงินเดือนคงที่ จะลิ้งค์มาหน้าคำนวณทุนประกันชีวิต
อ่านคำสำคัญที่ต้องรู้ในช่องบรรยายก่อน

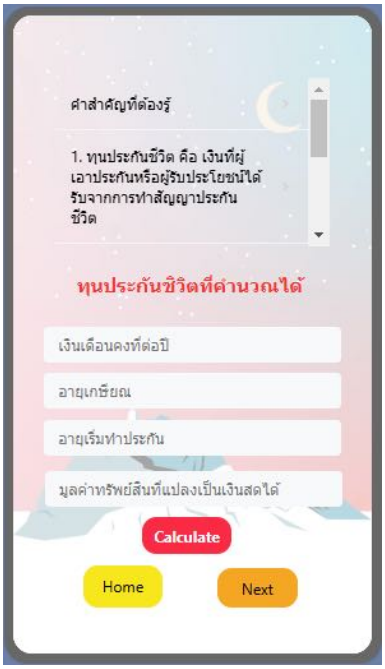
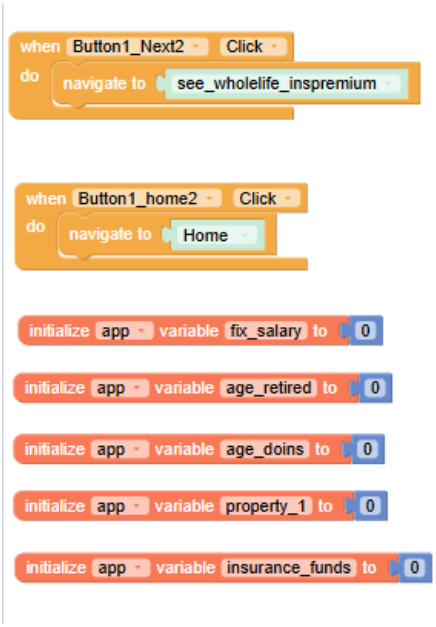
3.1 กรอกเงินเดือนคงที่ต่อปี

3.2 กรอกอายุเกษียณ (ไม่เกิน 60 ปี)

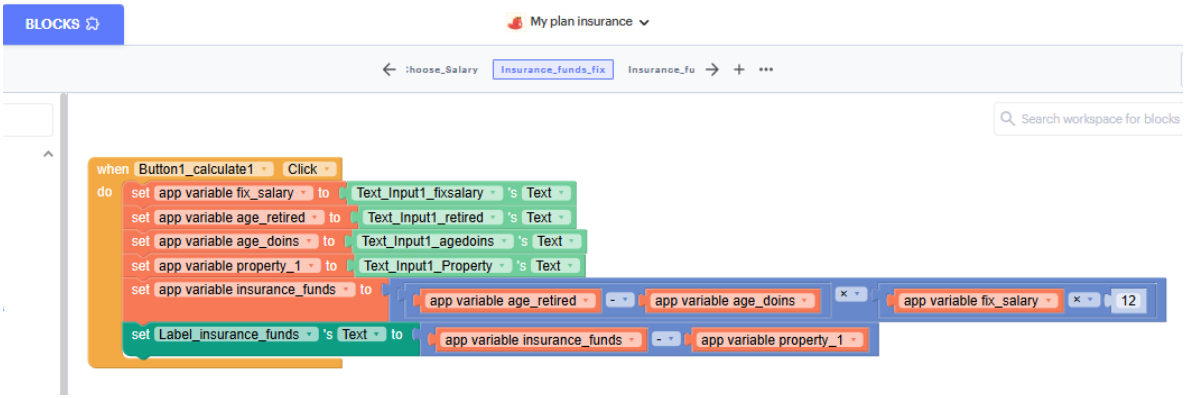
3.3 กรอกอายุเริ่มทำประกันชีวิต (ไม่ต่ำกว่า 25 ปี และให้จำไว้กรอกหน้าถัดไป)

3.4 กรอกมูลค่าสินทรัพย์ที่แปลงเป็นเงินสดได้ ถ้าไม่ทราบให้ใส่ 0 (ศูนย์)

3.5 แล้วกดปุ่มคำนวณ (Calculate) จะปรากฏทุนประกันชีวิตที่คำนวณได้ แล้ว click ปุ่ม Next หรือ กลับไปหน้า แรก Click ปุ่ม Home

User Interface	Block คำสั่ง
	กำหนดตัวแปร 

บล็อกคำสั่ง การคำนวณ



4. ถ้ากดปุ่มเลือก เงินเดือนไม่คงที่ จะลิ้งค์มาหน้าคำนวณทุนประกันชีวิต

4.1 กรอกเงินเดือนปัจจุบัน

4.2 กรอกจำนวนครั้งที่ขึ้นเงินเดือนต่อปี เป็นเลขที่ 12 หารลงตัว

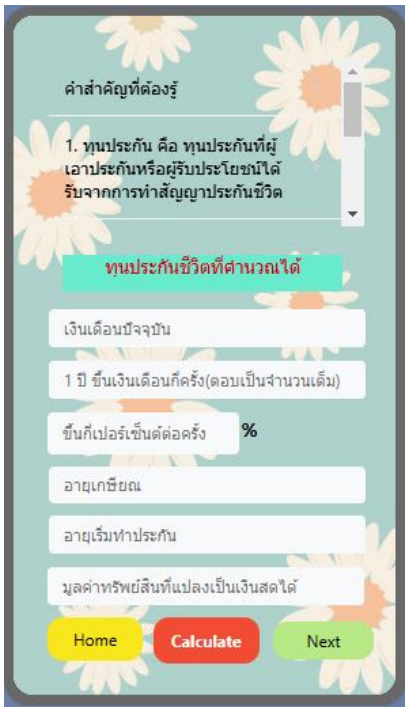
4.3 กรอกเปอร์เซ็นที่ขึ้นต่อครั้ง

4.4 กรอกอายุเกษียณ (ไม่เกิน 60 ปี)

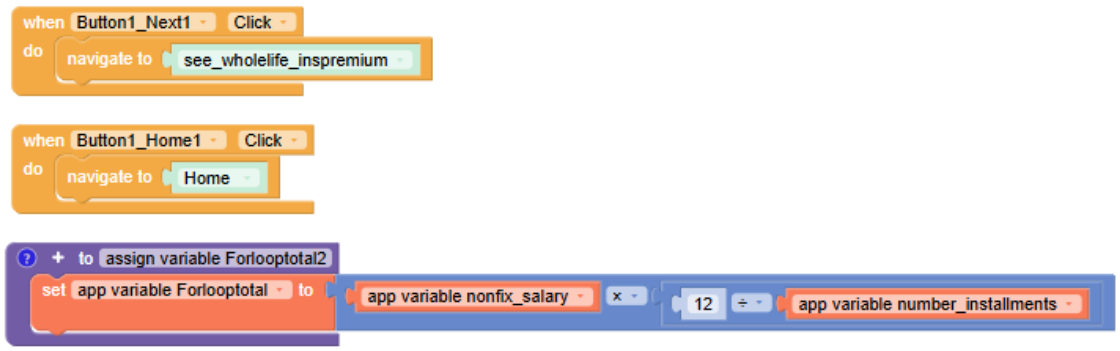
4.5 กรอกอายุเริ่มทำประกัน (ไม่ต่ำกว่า 25 ปี และให้จำไว้กรอกหน้าถัดไป)

4.6 กรอกมูลค่าสินทรัพย์ที่แปลงเป็นเงินสดได้ ถ้าไม่ทราบให้ใส่ 0 (ศูนย์)

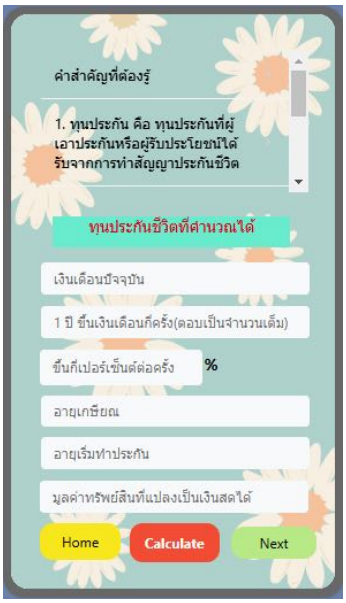
4.7 แล้วกดปุ่มคำนวณ (Calculate) จะปรากฏทุนประกันชีวิตที่คำนวณได้ แล้ว click ปุ่ม Next หรือ กลับไปหน้า แรก Click ปุ่ม Home

User Interface	Block คำสั่ง
	**บล็อกคำสั่ง กำหนดตัวแปร** 

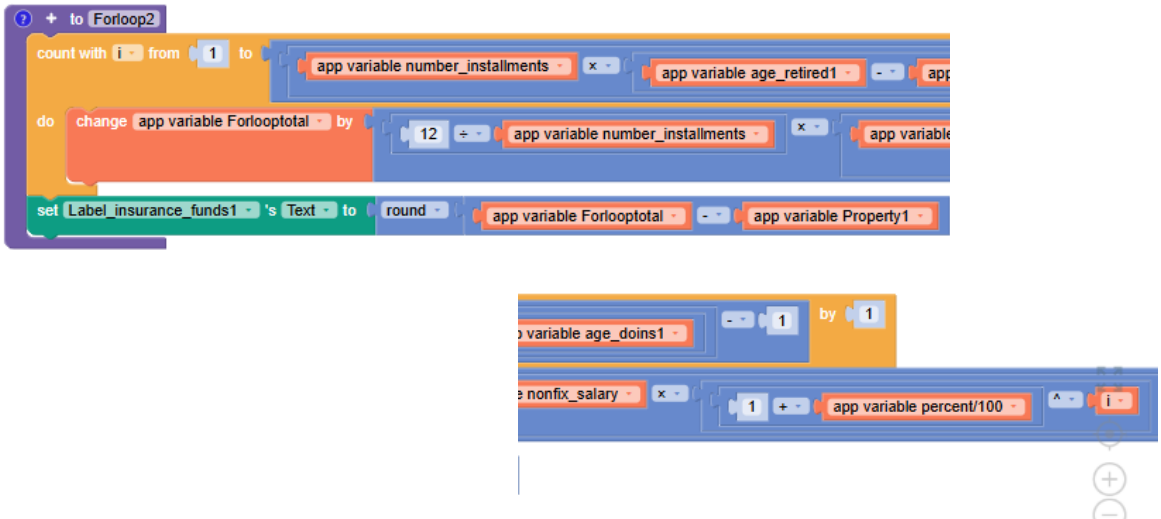
บล็อกคำสั่ง การคำนวณ

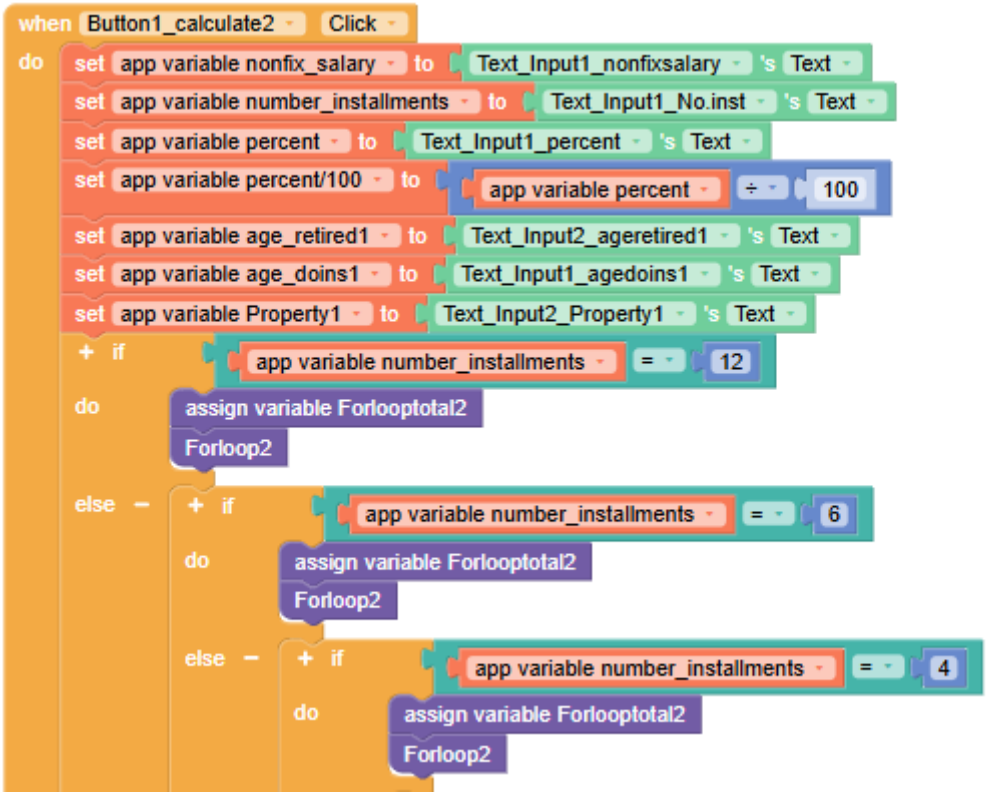


User Interface

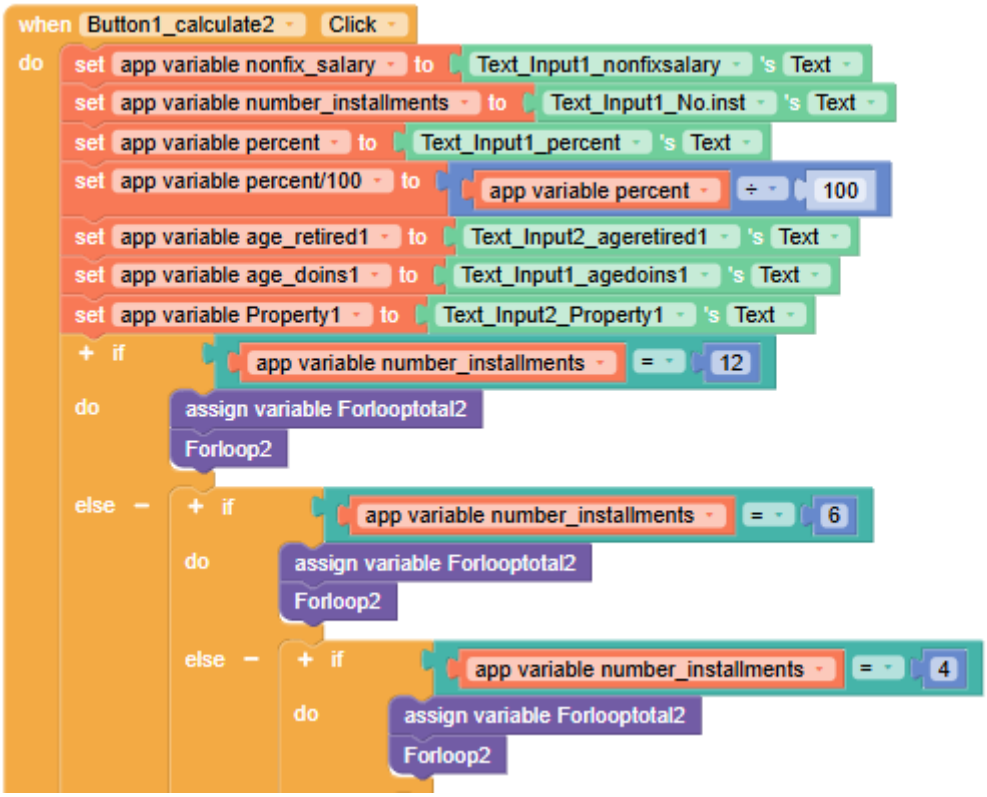


Block คำสั่ง การคำนวณ (ต่อ)

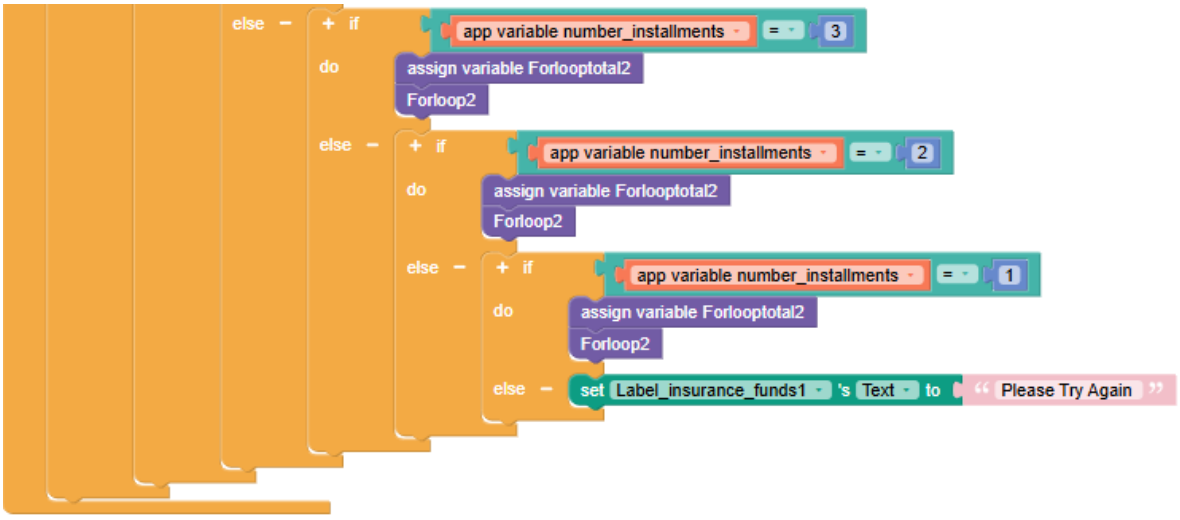



when Button1_calculate2 Click
do
set app variable nonfix_salary to Text_Input1_nonfixsalary 's Text
set app variable number_installments to Text_Input1_No.inst 's Text
set app variable percent to Text_Input1_percent 's Text
set app variable percent/100 to app variable percent ÷ 100
set app variable age_retired1 to Text_Input2_ageretired1 's Text
set app variable age_doins1 to Text_Input1_agedoins1 's Text
set app variable Property1 to Text_Input2_Property1 's Text
+ if app variable number_installments = 12
do
assign variable Forlooptotal2
Forloop2
else − + if app variable number_installments = 6
do
assign variable Forlooptotal2
Forloop2
else − + if app variable number_installments = 4
do
assign variable Forlooptotal2
Forloop2

Block คำสั่ง การคำนวณ (ต่อ)



Block คำสั่ง For Loop

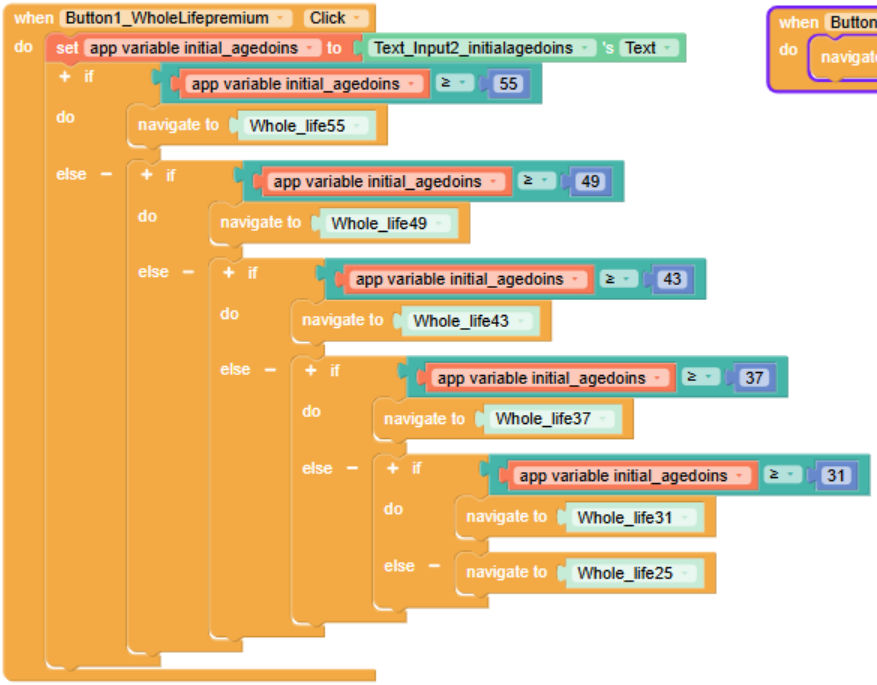


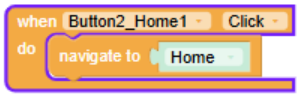
5. กรอกอายุเริ่มทำประกันที่ระบุจากหน้าที่แล้ว แล้ว Click ปุ่มเบี้ยประกันชีวิตตลอดชีพ หรือ Click ปุ่ม Home กลับหน้าแรก

User Interface	Block คำสั่ง
	บล็อกคำสั่ง กำหนดตัวแปร 

บล็อกคำสั่งหน้าลิ้งค์ไปหน้า sheet ค่าเบี้ยประกันชีวิตตลอดชีพ

ในช่วงอายุ 25-30 ปี ,31-36 ปี , 37-42 ปี , 43-48 ปี,49-54 ปี และ 55-60 ปี





6. **ตัวอย่าง หน้า Sheet** แสดงเบี้ยประกันอายุ 25-30 ปี

ให้ผู้ใช้แอพพลเคชัน เลื่อนแล้วเลือกเบี้ยประกันที่สนใจ และเลือกทุนประกันชีวิตและค่าเบี้ยประกัน**(ต่อปี)**และจดจำไว้ที่เพื่อกรอกในหน้าถัดไป

<u>หมายเหตุ</u>

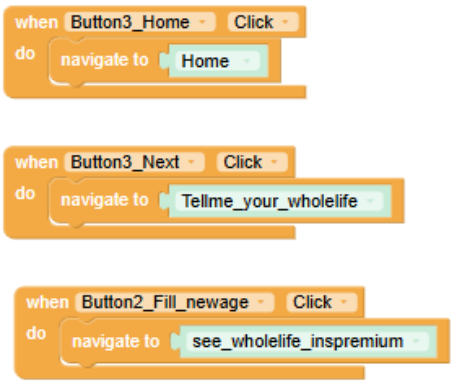
1) ทุนประกันชีวิต มี 3 ประเภท คือ 300,000 บาท 500,000 บาท 1,000,000 บาท และ 1,500,000 บาท

2) จำนวนระยะเวลาชำระเบี้ยคงที่มี 3 ประเภท คือ 10 ปี 15 ปี และ 20 ปี

6.1 Click ปุ่ม Fill age new ถ้าต้องการกรอกอายุเริ่มทำประกันใหม่

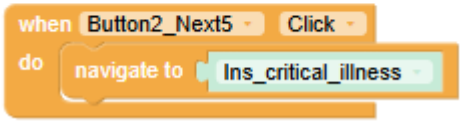
6.2 Click ปุ่ม Home กลับหน้าแรก

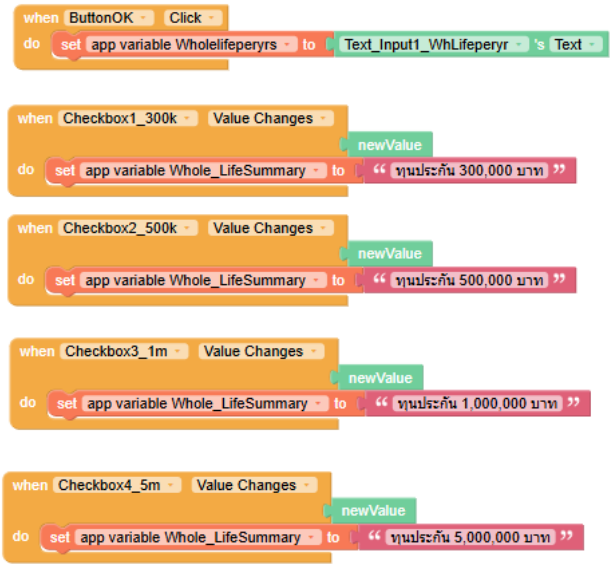
6.3 Click ปุ่ม Next ไปหน้าถัดไป

User Interface	Block คำสั่ง
	

7. ตอนที่ 1 เลือกทุนประกันชีวิตจากหน้า sheet โดยทำเครื่องหมายถูกเพียงช่องเดียวเท่านั้น
   ตอนที่ 2 กรอกเบี้ยประกันชีวิต(ต่อปี) ที่คุณเลือกจากหน้า sheet
   7.1 Click ปุ่ม It's Ok ทุกครั้ง หลังกรอกตอนที่ 1 และ 2 แล้ว
   7.2 Click ปุ่ม Next ไปหน้าถัดไป หรือ Click ปุ่ม Home กลับไปหน้าแรก

<table>
<tr><th>User Interface</th><th>Block คำสั่ง</th></tr>
<tr><td> 
</td><td>บล็อกคำสั่ง ไปหน้าถัดไป และ กลับหน้า Home
 

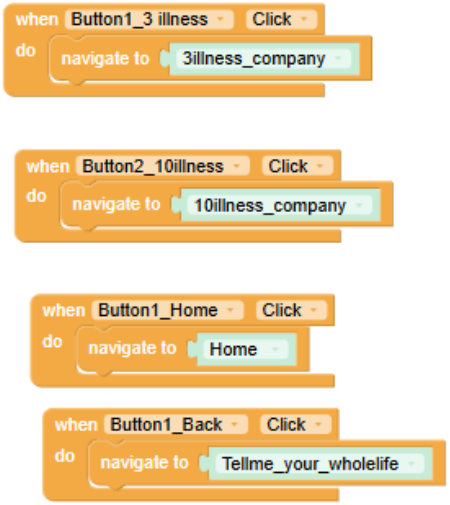
บล็อกคำสั่ง กำหนดตัวแปร
 
</td></tr>
<tr><td colspan="2">บล็อกคำสั่ง เก็บค่าที่ระบุไว้ใน ป้ายข้อความ
 
</td></tr>
</table>

8. เลือกเบี้ยประกันโรคร้ายแรง

8.1 อ่านคำศัพท์ที่ควรรู้

8.2 Click ปุ่ม เลือกคุ้มครอง 3 โรคร้ายแรง หรือ Click ปุ่ม เลือกคุ้มครองโรคร้ายแรงตั้งแต่ 10 โรคขึ้นไป ดูโรคร้ายแรงเพิ่มเติมที่ สำนักงาน คปภ. https://oiceservice.oic.or.th/document/Law/file/07170/07170_551981d3adc6632025b928a326fd0ea4_1.pdf

8.3 Click ปุ่ม Back กลับไปหน้าก่อนหน้านี้ หรือ Click ปุ่ม Home กลับไปหน้าแรก

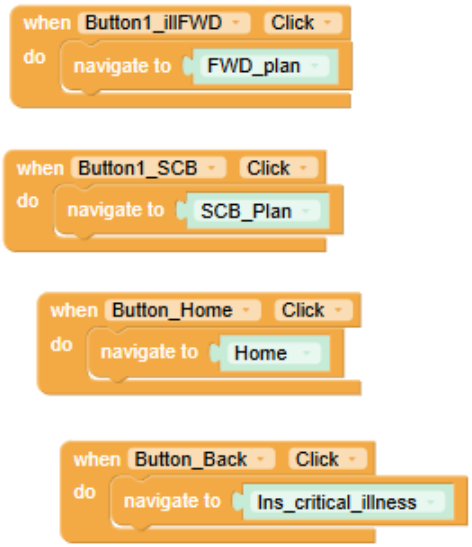
User Interface	Block คำสั่ง
	

9. ถ้า Click ปุ่มเลือกคุ้มครอง 3 โรคร้ายแรง มาหน้านี้

9.1 Click ปุ่มเลือกบริษัทให้ความคุ้มครองโรคร้ายแรงแบบเจอจ่ายจบ

9.2 ดูเบี้ยประกันของแต่ละบริษัท ใน sheet และจดจำค่าเบี้ยประกันโรคร้ายแรงเพื่อกรอกหน้าถัดไป

9.3 Click ปุ่ม Back กลับไปหน้าก่อนหน้านี้ หรือ Click ปุ่ม Home กลับไปหน้าแรก

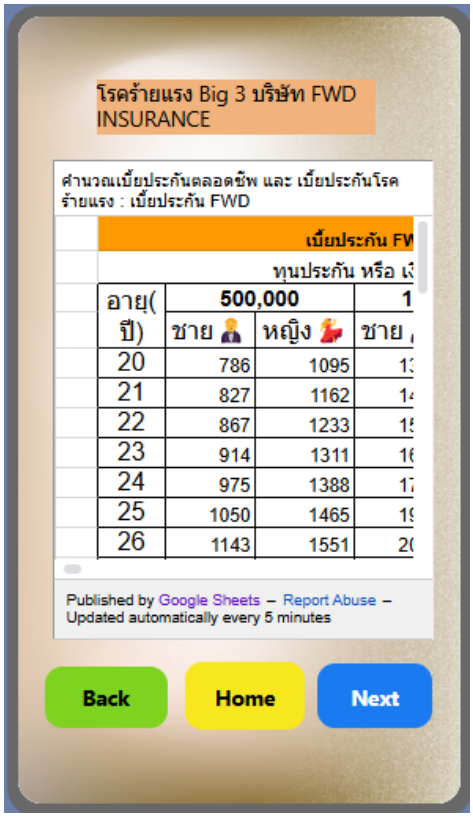
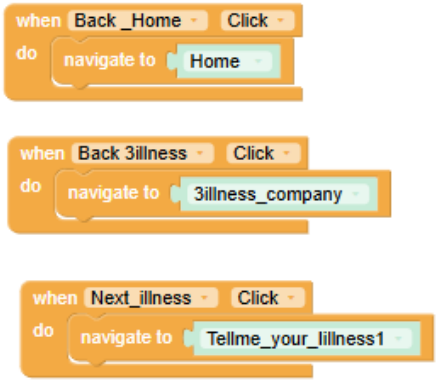
User Interface	Block คำสั่ง
	

10. ตัวอย่างหน้า sheet เบี้ยประกันคุ้มครอง 3 โรคร้ายแรง ของบริษัท FWD

10.1 Click ปุ่ม Home กลับหน้าแรก

10.2 Click ปุ่ม Back กลับหน้าก่อนหน้านี้

10.3 Click ปุ่ม Next ไปหน้าถัดไป

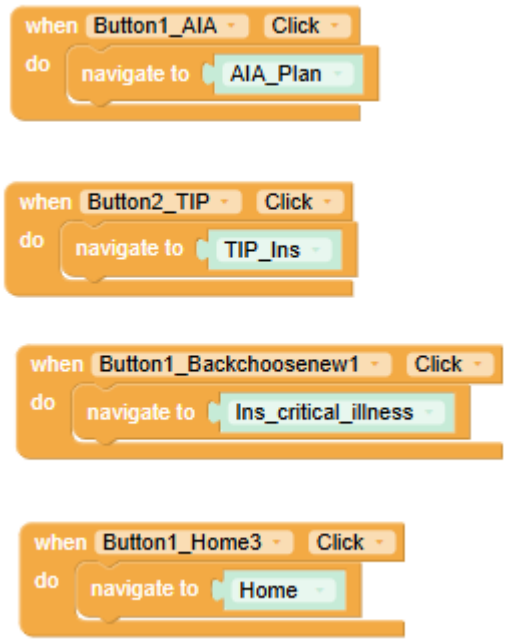
User Interface	Block คำสั่ง
	

11. ถ้า Click ปุ่มเลือกคุ้มครองตั้งแต่ 10 โรคร้ายแรง มาหน้านี้

11.1 Click ปุ่มเลือกบริษัทให้ความคุ้มครองโรคร้ายแรงแบบเจอจ่ายจบ

11.2 ดูเบี้ยประกันของแต่ละบริษัท ใน sheet และจดจำค่าเบี้ยประกันโรคร้ายแรงเพื่อกรอกหน้าถัดไป

11.3 Click ปุ่ม Back กลับไปหน้าก่อนหน้านี้ หรือ Click ปุ่ม Home กลับไปหน้าแรก

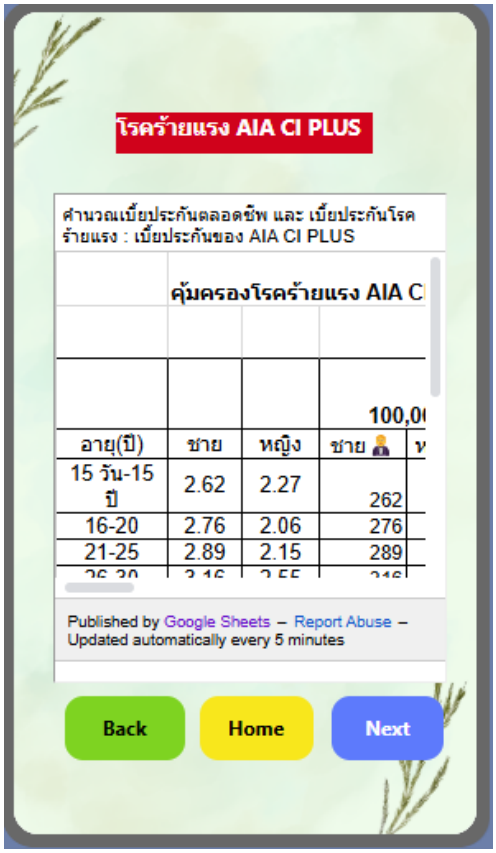
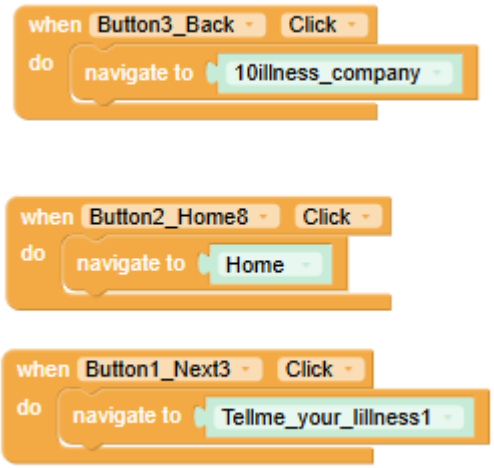
User Interface	Block คำสั่ง
	

12. ตัวอย่างหน้า sheet เบี้ยประกันคุ้มครอง ตั้งแต่ 10 โรคร้ายแรง AIA Plus

12.1 Click ปุ่ม Home กลับหน้าแรก

12.2 Click ปุ่ม Back กลับหน้าก่อนหน้านี้

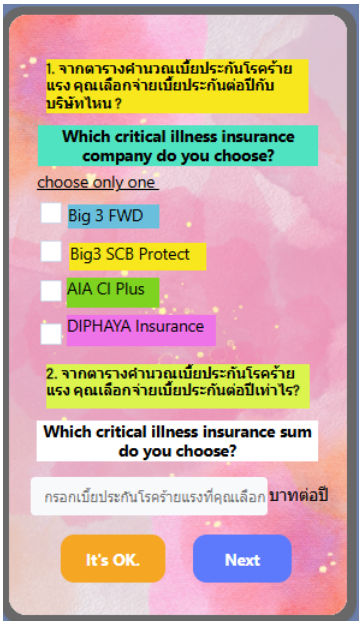
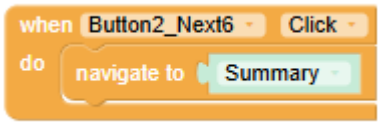
12.3 Click ปุ่ม Next ไปหน้าถัดไป

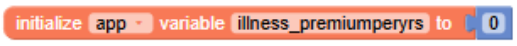
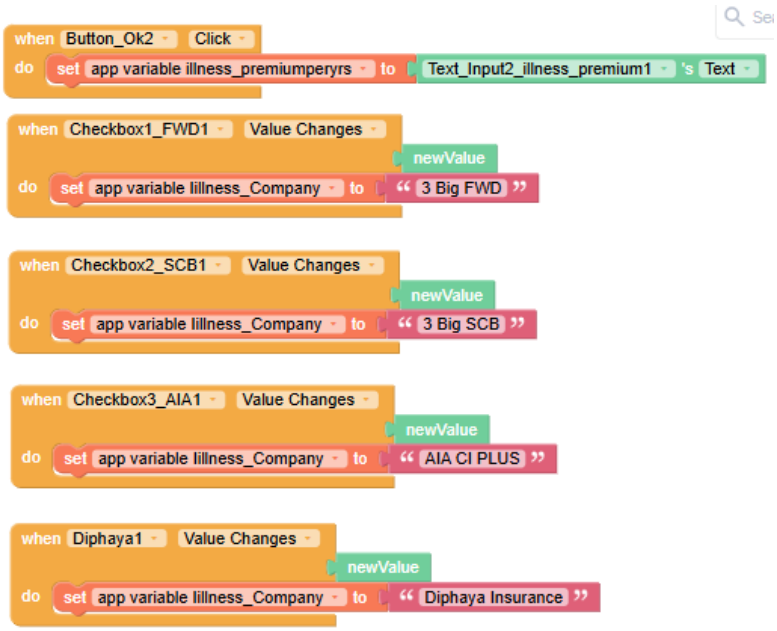
User Interface	Block คำสั่ง
	

13. ตอนที่ 1 เลือกบริษัทประกันโรคร้ายแรงจากหน้า sheet โดยทำเครื่องหมายถูกเพียงช่องเดียวเท่านั้น
ตอนที่ 2 กรอกเบี้ยประกันโรคร้ายแรง**(ต่อปี)** ที่คุณเลือกจากหน้า sheet

13.1 Click ปุ่ม It's Ok ทุกครั้ง หลังกรอกตอนที่ 1 และ 2 แล้ว

13.2 Click ปุ่ม Next ไปหน้าถัดไป

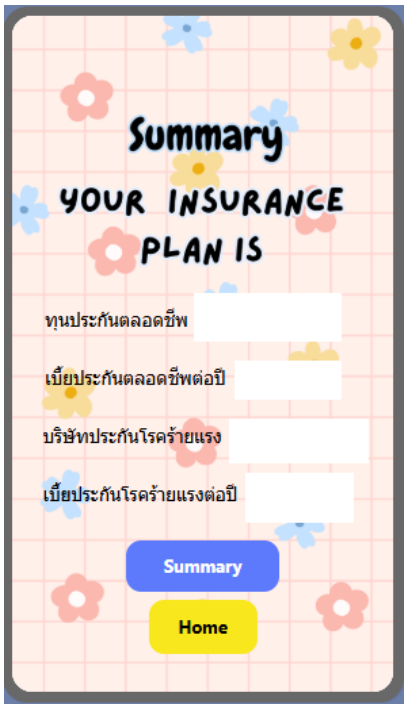
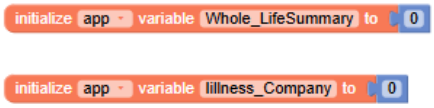
<table>
<tr><th>User Interface</th><th>Block คำสั่ง</th></tr>
<tr><td> 
</td><td>บล็อกคำสั่ง ไปหน้าถัดไป
 

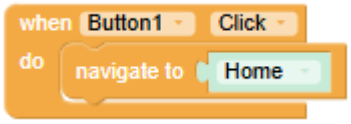
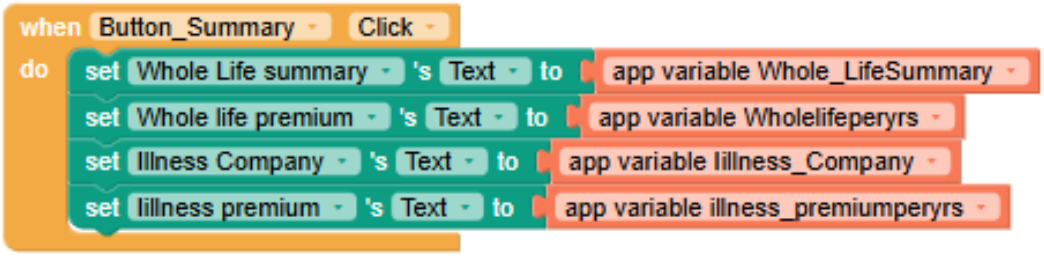
บล็อกคำสั่ง กำหนดตัวแปร
 
</td></tr>
<tr><td colspan="2">บล็อกคำสั่ง เก็บค่าที่ระบุไว้ใน ป้ายข้อความ
 
</td></tr>
</table>

14. หน้าสรุปแผนประกันชีวิตและประกันโรคร้ายแรงของคุณ

14.1 Click ปุ่ม Summary จะปรากฏค่าในช่องป้ายต่างๆ ที่คุณเลือกไว้

14.2 Click ปุ่ม Home กลับหน้าแรก

<table>
<tr><th>User Interface</th><th>Block คำสั่ง</th></tr>
<tr><td> 
</td><td>บล็อกคำสั่ง เก็บค่าตัวแปร
 

บล็อกคำสั่ง กลับหน้า Home
 
</td></tr>
<tr><td colspan="2">บล็อกคำสั่ง แสดงค่าในช่องป้ายต่างๆ
 
</td></tr>
</table>

Thank you for choosing us to plan for you

3.5 การทดสอบและเก็บข้อมูล

1. สร้างแบบสอบถาม วัดความพึงพอใจในระดับ 1–5

2. หาคุณภาพของแบบสอบถามโดยผู้เชี่ยวชาญ 3 ท่าน (IOC)

3. ทดสอบกับกลุ่มเป้าหมาย 87 คน

4.นำข้อมูลมาวิเคราะห์โดยใช้ T-test

5. สรุปผลจากค่าวิกฤต (Critical Value Method)

ตารางที่ 2: ผลการศึกษาความคิดเห็นของผู้ใช้งานแอปพลิเคชัน My Plan Insurance 87 คน

ที่	ประเด็นที่ประเมิน	$\overline{X}$	$S.D.$	แปลผล
1	รูปแบบของแอปพลิเคชันมีความน่าสนใจ	4.74	0.44	มากที่สุด
2	แอปพลิเคชันใช้งานง่าย ไม่ซับซ้อน	4.49	0.56	มาก
3	การจัดหมวดหมู่ของข้อมูลในแอปพลิเคชันง่ายต่อการใช้งาน	4.54	0.54	มากที่สุด
4	ข้อมูลที่ได้รับจากการประมวลผลมีความน่าเชื่อถือประมวลผลถูกต้อง	4.64	0.48	มากที่สุด
5	ความรวดเร็วในการใช้งานแอปพลิเคชัน	4.49	0.52	มาก
6	ข้อมูลมีความชัดเจน กระชับเข้าใจง่าย	4.63	0.48	มากที่สุด
7	แอปพลิเคชันให้ข้อมูลที่เป็นประโยชน์ในการเลือกวางแผนประกันชีวิตและประกันโรคร้ายแรง	4.54	0.50	มากที่สุด
8	ความพึงพอใจแอปพลิเคชันโดยภาพรวม	4.53	0.50	มากที่สุด
	ผลเฉลี่ยรวม	4.58	0.51	มากที่สุด

ตารางผลการศึกษาความคิดเห็นของผู้ใช้งานแอปพลิเคชัน My Plan Insurance ในภาพรวมอยู่ในระดับดีมากที่สุด มีผลเฉลี่ยรวม $\bar{x}$ เท่ากับ 4.58 , S.D. เท่ากับ 0.51 และมีข้อเสนอแนะจากผู้ใช้งานแอปพลิเคชัน ว่าควรให้คู่มือคำแนะนำเพื่อจะได้ทราบถึงขั้นตอนและวิธีการที่ในการใช้แอปพลิเคชัน

ตารางที่ 3 ผลการวิเคราะห์ค่าดัชนีความสอดคล้องของแบบสอบถามคิดเห็นจากผู้เชี่ยวชาญ

รายการประเมิน	การประเมิน			IOC	แปลผล
	ผู้เชี่ยวชาญ 1	ผู้เชี่ยวชาญ 2	ผู้เชี่ยวชาญ 3		
1. รูปแบบของแอปพลิเคชันมีความน่าสนใจ	+1	+1	+1	1	ใช้ได้
2. แอปพลิเคชันใช้งานง่าย ไม่ซับซ้อน	+1	0	+1	0.67	ใช้ได้
3. การจัดหมวดหมู่ของข้อมูลในแอปพลิเคชันง่ายต่อการใช้งาน	+1	0	+1	0.67	ใช้ได้
4. ข้อมูลที่ได้รับจากการประมวลผลมีความน่าเชื่อถือประมวลผลถูกต้อง	+1	+1	+1	1	ใช้ได้
5. ความรวดเร็วในการใช้งานแอปพลิเคชัน	+1	0	+1	0.67	ใช้ได้
6. ข้อมูลมีความชัดเจนกระชับเข้าใจง่าย	+1	+1	+1	1	ใช้ได้
7.แอปพลิเคชันให้ข้อมูลที่เป็นประโยชน์ในการเลือกวางแผนประกันชีวิตและประกันโรคร้ายแรง	+1	+1	0	0.67	ใช้ได้
8. ความพึงพอใจแอปพลิเคชันโดยภาพรวม	+1	+1	+1	1	ใช้ได้

3.6 การวิเคราะห์ทางสถิติ

หลังจากได้ข้อมูลผลเฉลี่ยรวม $\bar{x}$ และส่วนเบี่ยงเบนมาตรฐาน S.D. จะทำการวิเคราะห์ทางสถิติทดสอบค่าเฉลี่ยและตรวจสอบความมีนัยสำคัญมีขั้นตอนดังนี้:

ประชากร $n = 87$

องศาแห่งความอิสระ (degree of freedom) $df = n - 1 = 86$

คำนวณค่าสถิติทดสอบ (t-test) จากสูตร $t = \dfrac{\bar{x} - m_0}{S.D. / \sqrt{n}}$

$$t = \frac{4.58 - 4}{0.51 / \sqrt{87}} = \frac{0.58}{0.51 / 9.327} = \frac{0.58}{0.0547} \overset{a}{=} 10.61$$

ค่าวิกฤตสำหรับการทดสอบที่ $a = 0.05$ และ $df = 86$ คือ $t_{0.05,86} \overset{a}{=} 10.61$

เนื่องจากค่าสถิติทดสอบ t ที่คำนวณได้ (10.61) มีค่ามากกว่าค่าวิกฤต (1.662) ซึ่งหมายความว่าค่าสถิติทดสอบตกอยู่ในบริเวณวิกฤต (Critical Region)

ดังนั้น จึง ปฏิเสธสมมติฐานว่าง (H_0)

4 สรุปผลและอภิปรายผล

4.1 สรุปผล

ผู้จัดทำได้สูตรการคำนวณทุนประกัน ทั้งชนิดเงินเดือนคงที่ และเงินเดือนไม่คงที่มีอัตราขึ้นเงินเดือนเป็นเปอร์เซ็นต์ต่องวดที่ถูกต้อง โดยคำนวณจากความรู้พื้นฐานเรื่องเงินรวม การคิดดอกเบี้ย และการทดสอบด้วยการกำหนดค่าตัวเลข

ผู้ใช้งานแอปพลิเคชัน My plan insurance มีความพึงพอใจในภาพรวม จากการทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 เพียงพอที่จะสรุปได้ว่า ค่าเฉลี่ยความพึงพอใจภาพรวมของผู้ใช้บริการแอปพลิเคชัน มีค่ามากกว่า 4 อย่างมีนัยสำคัญทางสถิติ หรือในระดับดีมากที่สุด ซึ่งสนับสนุนเป้าหมายของคุณที่ต้องการให้ค่าเฉลี่ยไม่ต่ำกว่า 4 เมื่อพิจารณารายข้อ พบว่ารูปแบบของแอปพลิเคชันมีความน่าสนใจ มาเป็นลำดับแรก รองลงมาคือ ข้อมูลที่ได้รับจากการประมวลผลมีความน่าเชื่อถือประมวลผลถูกต้อง ข้อมูลมีความชัดเจน กระชับเข้าใจง่าย แอปพลิเคชันให้ข้อมูลที่เป็นประโยชน์ในการเลือกวางแผนประกันชีวิตและประกันโรคร้ายแรง การจัดหมวดหมู่ของข้อมูลในแอปพลิเคชันง่ายต่อการใช้งาน ความรวดเร็วในการใช้งานแอปพลิเคชัน แอปพลิเคชันใช้งานง่าย ไม่ซับซ้อน ตามลำดับ

ข้อจำกัดของปัญหาในการพัฒนาแอปพลิเคชัน My plan insurance คือใช้งานได้บนระบบปฏิบัติการ android และบน web application เท่านั้นยังไม่สามารถทำงานได้บนระบบปฏิบัติการ IOS

การคำนวณเบี้ยประกันภัยที่แท้จริงอาจมีความซับซ้อนมากกว่า ซึ่งรวมถึงปัจจัยด้านสุขภาพ และการประเมินความเสี่ยงเฉพาะบุคคล ซึ่งอยู่นอกขอบเขตของแอปพลิเคชันนี้ การคำนวณเบี้ยประกันชีวิตเป็นการประมาณการเนื่องจาไม่สามารถเข้าถึงข้อมูลเบี้ยประกันจากบริษัทได้โดยตรง

4.2 อภิปรายผล

การนำความรู้ทางการเงินมาใช้คำนวณเพื่อให้ได้ สูตรการคำนวณทุนประกัน ทั้งชนิดเงินเดือนคงที่ และเงินเดือนไม่คงที่มีอัตราขึ้นเงินเดือนเป็นเปอร์เซ็นต์ต่องวด สร้างตารางคำนวณ excel เพื่อคำนวณหาเบี้ยประกันชีวิตตลอดชีพและเบี้ยประกันโรคร้ายแรง แล้วนำมาพัฒนาแอปพลิเคชัน My plan insurance สอดคล้องกับงานวิจัยของ [10] Lav Ishan.(2023), ที่ใช้ความรู้ทางการเงินมาสร้าง Mobile Application ในการซื้อประกันชีวิต นอกจากนี้การเลือกใช้ Thunkable ในการพัฒนาแอปพลิเคชันยังช่วยให้การออกแบบ UX/UI สามารถทำได้ง่าย ใช้งานเข้าถึงได้ง่ายบนทั้งระบบปฏิบัติการ Android และ Web Application ซึ่งแตกต่างจากงานวิจัย [11] ภัทรานิษฐ์ แตงร่ม.(2565), ที่พัฒนาบนแพลตฟอร์มเฉพาะ ภาษา PHP และ JavaScript ซึ่งต้องใช้ทักษะการเขียนโค้ดที่สูงกว่า Thunkable และอาจใช้เวลาในการพัฒนา UI ผู้ใช้ที่ซับซ้อนกว่า

จากแบบประเมินความพึงพอใจรายข้อที่กล่าวว่าแอปพลิเคชัน My plan insurance ให้ข้อมูลที่เป็นประโยชน์ในการเลือกวางแผนประกันชีวิตและประกันโรคร้ายแรงในระดับดีมากที่สุด ทั้งนี้อาจเพราะ ความเข้าถึงข้อมูลของการจ่ายเบี้ยประกันในรูปแบบที่เป็น User Interface มีการแสดงผลแบบตอบกลับทันที และมีการสรุปให้เห็นเบี้ยประกันที่ผู้ใช้เลือกเพื่อคาดการณ์หรือประเมินค่าเบี้ยประกันที่ต้องจ่ายได้ สอดคล้องกับงานวิจัยของ [12] สุทัศน์ รุ่งระวิวรรณ. (2564), ที่ผลการประเมินความพึงพอใจรายข้อที่ว่าแอปพลิเคชันให้ข้อมูลที่เป็นประโยชน์ จากการพัฒนาแอปพลิเคชันการบริหารบุคลากรคณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏยะลา (APP FMS) อยู่ในระดับมาก

จากแบบประเมินความพึงพอใจของแอปพลิเคชัน My plan insurance มีความพึงพอใจของแอปพลิเคชันในภาพรวมในระดับดีมากที่สุด ทั้งนี้อาจเป็นเพราะว่า รูปแบบแอปพลิเคชันมีความน่าสนใจ เรื่องประกันชีวิตเป็นเรื่องใกล้ตัวและมีประโยชน์ช่วยวางแผนการจ่ายเบี้ยประกันเบื้องต้นได้ สอดคล้องกับงานวิจัยของ [13] ภาสกร ธนะศิระธรรม. (2566), มีผลการหาความพึงพอใจการใช้งานแอปพลิเคชันข้อมูลผู้ป่วยโรคหอบหืด พบว่ามีค่าเฉลี่ย ($\bar{X}$) เท่ากับ 4.26 และส่วนเบี่ยงเบนมาตรฐาน ($S.D.$) เท่ากับ 0.60 อยู่ในระดับความพึงพอใจมาก

4.3 แนวทางสำหรับวิจัยในอนาคต

ศึกษาและพัฒนาแอปพลิเคชันประเภทประกันชีวิตแบบอื่น เพื่อให้มีความสมบูรณ์ครบทุกประเภท ซึ่งจะเป็นประโยชน์ต่อผู้ใช้งานและผู้สนใจมากขึ้น

ศึกษาและพัฒนาแอปพลิเคชันทางการเงินเรื่องอื่น เช่น การคิดค่ารายงวดทบต้นทุกๆ ต้นเดือน การคิดค่ารายงวดทบต้นทุกๆ ปลายเดือน การคิดค่ารายงวดในการซื้อสินค้าแบบผ่อนชำระ เป็นต้น

กิตติกรรมประกาศ (optional) ผู้แต่งขอขอบคุณผู้ทรงคุณวุฒิทุกท่านที่ได้ให้ข้อคิดเห็นและข้อเสนอแนะต่าง ๆ เพื่อปรับปรุงบทความวิจัยนี้

1. ธัญณิชา ทองอยู่ , 2. อมรรัตน์ อ้วนไตร และ 3. อัชฌา เฉลยจิตร์

เอกสารอ้างอิง

[1] K. J. Arrow, Uncertainty and the welfare economics of medical care. American Economic Review, 53(5) (1963).

[2] R. S. Brown, The challenge of choosing the right health insurance: Tools for consumers. Journal of Health Economics, 45 (2016).

[3] สำนักงานสถิติแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม, การสำรวจการมีการใช้เทคโนโลยีสารสนเทศและการสื่อสารในครัวเรือน พ.ศ.2567(ไตรมาส1) (Online).สำนักงานสถิติแห่งชาติ, หน้า จ.

[4] Visit Boonchom, The Development of Android Application for Disseminating Thai Cultural Heritage of the Lower Southern Provinces of Thailand: Phatthalung province. Thaksin University Journal, 23(3) (2020) (33).

[5] สำนักราชบัณฑิตยสภา.(2555). บทวิทยุรายการ “รู้ รัก ภาษาไทย”ออกอากาศทางสถานีวิทยุกระจายเสียงแห่งประเทศไทย เมื่อวันที่ ๒๐ กุมภาพันธ์ พ.ศ. ๒๕๕๕ เวลา ๗.๐๐-๗.๓๐ น. <http://legacy.orst.go.th/>, 2567 (สืบค้น 10 ต.ค. 67).

[6] K WEALTH, ประกันชีวิตที่ทุกคนควรมี, <https://www.kasikornbank.com/th/kwealth/Pages/a502-t4-hyb-insurance-must-have-kgth.aspx>, 2567 (สืบค้น 10 ต.ค. 67).

[7] ศ.ชูเกียรติ ประมูลผล. หนังสือเรียนการประกันชีวิต (Life Insurace), http://old-book.ru.ac.th/e-book/indexstart.htm, (สืบค้น 18 ต.ค. 67).

[8] วโรดม คำแผ่นชัย. สร้าง Mobile App ง่ายๆ ทั้ง iOS และ Android แค่ลากวางด้วย Thunkable Part 1, <https://medium.com/altotech/สร้าง-mobile-app-ง่ายๆ-ทั้ง-ios-และ-android-แค่ลากวางด้วย-thunkable-part-1-9285d221752e>, 2561 (สืบค้น 18 ต.ค. 67).

[9] สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย, คำสั่งนายทะเบียนที่ 57/2560 เรื่อง ให้ใช้ตารางมรณะไทยปี 2560, https://www.oic.or.th/th, 2561 (สืบค้น 9 ต.ค. 67).

[10] Lav Ishan, The use of mobile application to buy insurance: an AHP-based study. International Journal of Operational Research (IJOR), Vol. 46, No. 3, 2023.

[11] ภัทรานิษฐ์ แตงร่ม, กลยุทธ์การพัฒนาแอปพลิเคชันเพื่อประสิทธิภาพในการเข้าถึงข้อมูลการออมเงินกรณีชรา ภาพของผู้ประกันตนในกองทุนประกันสังคม: พิษณุโลก. มหาวิทยาลัยนเรศวร, 2565

[12] สุทัศน์ รุ่งระวิวรรณ, การพัฒนาแอปพลิเคชั่นบนอุปกรณ์เคลื่อนที่สำหรับการบริหารบุคลากร: ร้อยเอ็ด. Journal of Roi Kaensarn Academi, ปีที่ 6 ฉบับที่ 3, 2564.

[13] ภาสกร ธนะศิระธรรม, การพัฒนาแอปพลิเคชันข้อมูลผู้ป่วยโรคหอบหืด: มหาสารคาม.วารสารวิชาการ “การจัดการเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม”, ปีที่ 10 ฉบับที่ 2, 2566.

Logistic Regression for Mitigating Subpopulation Effects in Familial DNA Search

Akaraphon Jantaraphum[1,*] and Monchai Kooakachai[1,†]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

In hypothesis testing for familial DNA searches, the likelihood ratio is a commonly used statistic that requires allele frequency distributions from the relevant population as input. However, since the population background of a forensic sample is generally unknown, the appropriate allele frequency distribution is uncertain. Additionally, when accounting for population substructure, multiple allele frequency distributions are available, making it unclear which one should be applied. To address this issue, several variations of likelihood ratio statistics have been developed, including those based on average allele frequencies (LRLAF), as well as the maximum (LRMAX), minimum (LRMIN), or average likelihood ratios (LRAVG), along with Naive Bayes classification (LRNB). In this study, we introduce a family of likelihood ratio-based statistics incorporating logistic regression classification (LRMLR). DNA samples are classified into subpopulations using logistic regression, and the likelihood ratio framework is then applied based on the classification results. We evaluate the performance of LRMLR alongside existing methods, using Thai population data as a case study. Our results show that LRMLR provides slightly higher power for parent-child and full-sibling tests compared to the prominent ones, such as LRNB and LRMIN, with no significant differences across classification methods. This offers a viable alternative for conducting familial DNA searches in real-world scenarios, delivering comparable power to existing approaches.

Keywords: Familial DNA search, Population substructure, Likelihood ratio, Logistic regression
2020 MSC: 62P10, 62J12

1 Introduction

In forensic investigations, genetic material or DNA is one of the most common types of evidence collected from crime scenes. It can be extracted from various biological sources such as hair, blood, sweat, or saliva. Once DNA profiles from the crime scene are obtained, they are compared to those stored in forensic databases, which often contain profiles of arrestees or convicted offenders. When an exact match is found, where genotypes align at all core loci (specific positions on the chromosome where DNA testing occurs), it strongly suggests that the sample belongs to the identified individual. However, in cases where no perfect match is available, investigators may

*Speaker (6570141523@student.chula.ac.th)
†Corresponding author (Monchai.K@chula.ac.th)

still find partial matches, offering potential leads in a case. By further examining shared alleles and using mathematical techniques, we can determine the likelihood of a familial relationship between the individual who left the DNA sample at the crime scene and someone in the DNA database. This process is carried out through hypothesis testing and is known as a familial DNA search. This approach allows law enforcement to trace individuals who may be linked to a crime, even if they are not directly present in criminal databases, thereby expanding investigative possibilities and increasing the chances of solving complex cases.

Various strategies for familial searching have been proposed in the literature, ranging from simple methods that count the number of shared alleles to more advanced probabilistic approaches [1–5]. Many studies, including the one by Curran and Buckleton in New Zealand [6], have stated that the likelihood ratio approach outperforms the traditional method of counting matching alleles. Therefore, this article focuses on an approach that uses the likelihood ratio as the basis for calculations. Familial searching is one of three indirect methods used to identify potential sources of forensic biological evidence. The other two are Y-chromosome short tandem repeat (Y-STR) database searches and investigative genetic genealogy (IGG). A discussion of the strengths and limitations, along with an overview of each method, is available in [7]. Despite its usefulness, familial searching raises ethical concerns, especially about the overrepresentation of nonwhite individuals in forensic databases and the risk of invasive investigations into suspects' relatives. Different states and countries have varying policies—some permit or prohibit it entirely, while others regulate follow-ups on accidental partial matches to address these concerns [8].

In statistical hypothesis testing for familial DNA searches, the likelihood ratio (LR) statistic is defined as the ratio of the probabilities of observing a pair of DNA profiles under two distinct relationship hypotheses. These hypotheses are as follows:

H_0 : The two individuals are unrelated.

H_1 : The two individuals share a specific genetic relationship.

The LR can be expressed as

$$\frac{P(\text{seeing two given DNA profiles} \mid H_1)}{P(\text{seeing two given DNA profiles} \mid H_0)}.$$

These probabilities in the numerator and denominator can only be calculated if the allele frequency distribution of the sample is known (which we will explain in the next section), but such information is often unavailable or uncertain. If everyone in the population had the same genetic background, meaning their genetic data came from a single allele frequency distribution, there would be no issue. However, this assumption is frequently unrealistic in forensic contexts, as numerous studies suggest that populations are better represented by multiple genetic substructures, such as racial or ancestral groups, which can greatly affect allele frequencies [9–11]. Ignoring these variations can lead to inaccurate likelihood ratio calculations and false conclusions, potentially causing miscarriages of justice [12].

To address the issue of population substructure, researchers have proposed methods that adjust the likelihood ratio formula when allele frequency distributions are available for each subpopulation. These methods can be categorized into two groups. The first group relies on arithmetic techniques to combine likelihood ratios in different ways, such as averaging available allele frequency distributions (LRLAF) or averaging, maximizing, or minimizing likelihood ratios

at the ratio level (LRAVG, LRMAX, LRMIN). The second group utilizes advanced machine learning techniques to first predict the correct subpopulation and then apply the likelihood ratio framework, such as using a Naive Bayes classifier for classification (LRNB). All of these methods will be discussed in the next section, along with their history and performance as reported in previous literature.

In this study, we introduce a new family of likelihood ratio-based statistics to address population substructure and compare it with existing methods. Our approach utilizes logistic regression, where we first predict the correct subpopulation using logistic regression and then apply the likelihood ratio framework based on the predicted classification. This method is referred to as LRMLR.

The rest of the paper is structured as follows. Section 2 presents the methodology, detailing the mathematical calculations for the likelihood ratio and covering both existing and proposed statistics. We also provide a simulation study algorithm to compare the performance of these statistics. The corresponding simulation results are presented in Section 3, followed by the discussion and conclusions in Section 4.

2 Methodology

DNA profiles consist of genotypes across m positions on the human chromosome. These m positions, carefully selected by biologists, are referred to as loci (singular: locus). The number of loci used varies depending on the region where the analysis is conducted. For instance, in Thailand, familial DNA matching is based on 15 loci [13]. These 15 loci are as follows: FGA, TH01, TPOX, CSF1PO, vWA, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, and D21S11.

Each person's genotype at a specific locus is represented as a pair of numbers. One number represents the short tandem repeats (STRs) inherited from their mother, while the other number represents the STRs inherited from their father. Each number is called an allele, and when paired together, they form a genotype. Table 1 presents an example of two DNA profiles across four loci: D3S1358, D7S820, D8S1179, and D13S317. It can be observed that the two profiles share at least one allele at each locus considered. This raises the question of performing a statistical hypothesis test to determine whether the two profiles exhibit a parent-child relationship, as a father and son always share one allele (identical by descent) at each locus.

Table 1: An example of two DNA profiles

DNA profiles	Locus			
	D3S1358	D7S820	D8S1179	D13S317
The first individual	(13,14)	(11,12)	(17,18)	(13,13)
The second individual	(13,14)	(12,14)	(18,18)	(12,13)

Denote the DNA profiles for two individuals by $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. The genetic relatedness between two individuals can be described using the measure of relatedness $\boldsymbol{\theta}$ as explained in [14]. Specifically, the authors described how familial relationships can be characterized by the probabilities of alleles at a locus being identical by descent (IBD), where alleles are considered IBD if they are inherited as copies of the same ancestral allele. For any two individuals, the relatedness at

a specific locus is represented as a 3-tuple $\boldsymbol{\theta} = (z_0, z_1, z_2)$ where z_0 denotes the probability that neither allele of $\boldsymbol{X}_1$ is IBD with any allele of $\boldsymbol{X}_2$, z_1 represents the probability that exactly one allele of $\boldsymbol{X}_1$ is IBD with an allele of $\boldsymbol{X}_2$, and z_2 indicates the probability that both alleles of $\boldsymbol{X}_1$ are IBD with those of $\boldsymbol{X}_2$. Thus, the hypotheses for testing a specific relationship can be rewritten as the following:

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 = (1, 0, 0) \quad \text{versus} \quad H_1 : \boldsymbol{\theta} = \boldsymbol{\theta}_1 = (0, 1, 0)$$

for testing parent-child relationship, and

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 = (1, 0, 0) \quad \text{versus} \quad H_1 : \boldsymbol{\theta} = \boldsymbol{\theta}_1 = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$$

for testing full-sibling relationship. The test statistic used is the likelihood ratio (LR), defined by

$$\text{LR} = \frac{P(\boldsymbol{X}_1, \boldsymbol{X}_2 \mid H_1)}{P(\boldsymbol{X}_1, \boldsymbol{X}_2 \mid H_0)}.$$

If LR > 1, it indicates a greater likelihood that the two DNA profiles $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ were generated under H_1 compared to H_0. In particular, higher LR provides stronger evidence to reject the null hypothesis and support the specific relationship of interest.

To calculate the probabilities in the numerator and denominator of LR, it is necessary to know the allele frequency distributions for each individual. Let $\boldsymbol{f}_{\boldsymbol{X}_1}$ and $\boldsymbol{f}_{\boldsymbol{X}_2}$ denote the allele frequency distributions for the populations associated with $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively. These distributions specify the frequencies of each allele, allowing the probabilities to be fully determined. In general, we express the formula for LR as

$$\text{LR} = \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{\boldsymbol{X}_1}, \boldsymbol{f}_{\boldsymbol{X}_2})}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{\boldsymbol{X}_1}, \boldsymbol{f}_{\boldsymbol{X}_2})} \tag{2.1}$$

where L denotes the likelihood function.

To evaluate the numerator and denominator in equation (2.1), we use Table 2, obtained from [15], which displays seven possible genotype combinations for two individuals, along with their corresponding probabilities for each scenario. Probabilities for each locus are calculated individually and then multiplied together, under the assumption that the loci are independent. This assumption holds if the loci are not genetically linked and are in linkage equilibrium. In fact, some core loci do not fully meet the assumption of independence, such as vWA and D12S391, which are located on the same chromosome about 6.3 megabases (Mb) apart, and genetic markers should ideally be separated by more than 50 Mb to ensure independence [16]. Although independence between vWA and D12S391 may not always hold at the individual level, no significant correlation has been observed at the population level [17, 18]. For simplicity, we assume independence between loci in our calculations.

Example 2.1. We will illustrate the calculation of the LR using a simple example with three independent loci: D8S1179, CSF1PO, and TH01. Table 3 presents the DNA profiles of two individuals, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$.

The goal is to test if $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ have a parent-child relationship. Consider the D8S1179 locus, where assuming the allele frequencies for alleles 10 and 11 are 0.1906 and 0.1089, respectively,

Table 2: Probabilities of the seven possible combinations of genotypes for multi-allelic loci in pairs of individuals, conditional on their relationship, as functions of alleles frequencies. p_i represents the allele frequency of allele i.

Genotypes	Unrelated $\boldsymbol{\theta} = (1, 0, 0)$	Parent-child $\boldsymbol{\theta} = (0, 1, 0)$	Full-siblings $\boldsymbol{\theta} = (1/4, 1/2, 1/4)$
AA, AA	p_A^4	p_A^3	$p_A^2(1+p_A)^2/4$
AA, AB	$4p_A^3p_B$	$2p_A^2p_B$	$p_A^2p_B(1+p_A)$
AA, BB	$2p_A^2p_B^2$	0	$p_A^2p_B^2/2$
AB, AB	$4p_A^2p_B^2$	$p_Ap_B(p_A+p_B)$	$p_Ap_B(2p_Ap_B+p_A+p_B+1)/2$
AA, BC	$4p_A^2p_Bp_C$	0	$p_A^2p_Bp_C$
AB, AC	$8p_A^2p_Bp_C$	$2p_Ap_Bp_C$	$p_Ap_Bp_C(2p_A+1)$
AB, CD	$8p_Ap_Bp_Cp_D$	0	$2p_Ap_Bp_Cp_D$

Table 3: DNA profiles of two individuals, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$.

DNA profiles	Locus		
	D8S1179	CSF1PO	TH01
$\boldsymbol{X}_1$	(10,10)	(12,13)	(7,8)
$\boldsymbol{X}_2$	(10,11)	(10,12)	(7,8)

the probability of observing the genotypes (10,10) and (10,11) under the parent-child hypothesis is $2(0.1906)^2(0.1089) \approx 0.007912$. In contrast, assuming the individuals are unrelated, the probability of observing these genotypes is $4(0.1906)^3(0.1089) \approx 0.003016$. This yields a likelihood ratio at the D8S1179 locus of $0.007912/0.003016 \approx 2.6233$. These probabilities can be computed separately for each locus, and under the assumption of independent loci, the overall likelihood ratio can be obtained by multiplying the individual ratios. Assuming the allele frequencies for alleles 10, 12, and 13 at the CSF1PO locus are 0.2252, 0.354, and 0.0817, respectively, and the allele frequencies for alleles 7 and 8 at the TH01 locus are 0.3243 and 0.0545, respectively, the final likelihood ratio is 9.9264. This indicates that these two profiles are approximately 9.9264 times more likely to originate from a parent-child relationship than from unrelated individuals.

In the simplest case of a population without substructure, $\boldsymbol{f}_{\boldsymbol{X}_1}$ and $\boldsymbol{f}_{\boldsymbol{X}_2}$ are identical and represent the allele frequency distribution of the entire population. In this scenario, the likelihood ratio test is straightforward, and the calculation can be carried out without any issues. However, if population substructure exists, meaning there are multiple allele frequency distributions, it becomes crucial to determine which allele frequency distributions to use for $\boldsymbol{f}_{\boldsymbol{X}_1}$ and $\boldsymbol{f}_{\boldsymbol{X}_2}$ in the calculation. For instance, in the Thai population, the probabilities of observing allele 10 at the D8S1179 locus differ across subgroups: 0.1906 in the north, 0.1480 in the northeast, 0.1906 in the central, and 0.1469 in the south. Since we don't know which subpopulation the sample belongs to, this creates ambiguity in the calculation.

The parameters $\boldsymbol{f}_{\boldsymbol{X}_1}$ and $\boldsymbol{f}_{\boldsymbol{X}_2}$ are called nuisance parameters, as they are parameters that are not of direct interest but need to be accounted for in the analysis to avoid bias or errors in

the estimation of the parameters of primary interest.

2.1 Existing Test Statistics

Consider a population partitioned into $R > 1$ distinct subpopulations, each characterized by its own allele frequency distribution, $\boldsymbol{f}_i$, corresponding to the ith subpopulation. Below, we review existing methods in the literature for conducting familial DNA tests to assess the genetic relationship between the two DNA profiles $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$.

The first test statistic is derived using the likelihood ratio framework, substituting $\boldsymbol{f}_{\boldsymbol{X}_1}$ and $\boldsymbol{f}_{\boldsymbol{X}_2}$ with $\boldsymbol{f}_{local}$, which is defined by

$$\boldsymbol{f}_{local} = p_1\boldsymbol{f}_1 + p_2\boldsymbol{f}_2 + \cdots + p_R\boldsymbol{f}_R$$

where p_i is the relative size of the ith subpopulation with respect to the entire population. This $\boldsymbol{f}_{local}$ is called the local allele frequency and the corresponding likelihood ratio test statistic that uses $\boldsymbol{f}_{local}$ will be referred as LRLAF. This test statistic was firstly mentioned in [19] as the authors claimed that it is effective in the Colorado prison population. It can be expressed mathematically as

$$\text{LRLAF} = \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{local})}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{local})}.$$

Observe that the LRLAF approach operates by averaging directly at the allele frequency level. However, a slight variation involves shifting the averaging process to the likelihood ratio level instead, resulting in a new statistic called LRAVG. In this method, likelihood ratios are repeatedly computed using the single-population framework applied to various locally available allele frequency distributions. The final step is to calculate a weighted average of these likelihood ratios. The LRAVG statistic can be mathematically expressed as follows:

$$\text{LRAVG} = p_1\frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_1)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_1)} + p_2\frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_2)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_2)} + \cdots + p_R\frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_R)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_R)}.$$

Note that the LRAVG statistic was also mentioned in the article [19]. The performance of LRAVG for the Colorado prison population, measured by power and Type I error, is substantially lower than that of LRLAF. There is a possibility of adjusting the weights. For example, if we prefer to use the uniform weights, the weight $1/R$ can be utilized. In this article, we restrict our weight to be p_i, which is defined as the the proportion of ith subpopulation.

If, in the final step, we choose to take either the maximum or minimum instead of calculating a weighted average, the resulting statistics are referred to as LRMAX and LRMIN, respectively. The definitions of these two statistics are as follows:

$$\text{LRMAX} = \max\left\{\frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_1)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_1)}, \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_2)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_2)}, \ldots, \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_R)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_R)}\right\},$$
$$\text{LRMIN} = \min\left\{\frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_1)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_1)}, \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_2)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_2)}, \ldots, \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_R)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_R)}\right\}.$$

A brief note on the history of these two statistics is worth mentioning here. The Denver Crime Lab has been using LRMAX in practice for familial DNA testing since 2008 and has successfully solved numerous cases by identifying offenders with this method. However, the effectiveness of LRMAX has been called into question, as [19] showed that its power is 10% and

5% lower than that of LRLAF and LRAVG, respectively, within the Colorado prison population. Later, Ge and Budowle (2012) [20] showed that using the minimum approach would make the likelihood ratio distribution closer to the truth, based on their simulation study on the global population. This is a strong starting point to support the use of LRMIN. In addition to their discoveries on the likelihood ratio level, Thitiyan and Jumnongwut (2022) [21] confirmed that LRMIN is indeed a better statistic than LRMAX in the Thai population, as they demonstrated through a comparison in terms of power.

Observe that all previous methods rely on the arithmetic combination of available pieces of information, such as averages, maxima, and minima. An improvement incorporating machine learning techniques was introduced by Laoiam and Kooakachai (2023) [22]. They proposed a new approach that includes a classification step to conduct hypothesis tests for parent-child and full-sibling comparisons. In particular, they introduced a test statistic based on Naive Bayes classification, which simultaneously classifies two DNA profiles to identify the subpopulation to which they belong. The likelihood ratio statistic is then applied within a single-population framework, using the classification result. This statistic, referred to as LRNB, is mathematically defined as

$$\text{LRNB} := \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{NB})}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{NB})}$$

where $\boldsymbol{f}_{NB}$ represents the subpopulation predicted from the Naive Bayes classification. According to [22], LRNB demonstrated the highest power, outperforming all methods in the arithmetic-based group for full-sibling tests in the Thai population.

2.2 Proposed Test Statistics

We aim to make a slight adaptation to the LRNB approach by using another well-known technique: logistic regression. To begin, we outline two methods for representing DNA profiles as explanatory variables in logistic regression. A DNA profile with m loci can be generally represented in the following format:

$$\{(x_{1a}, x_{1b}), (x_{2a}, x_{2b}), (x_{3a}, x_{3b}), \ldots, (x_{ma}, x_{mb})\}$$

where x_{ij} denotes the allele observed at locus $i \in \{1, 2, \ldots, m\}$ and $j \in \{a, b\}$ is the maternal or paternal origin of the allele. The target variable in the proposed logistic regression model represents the subpopulation to which an individual belongs. The predictors can be organized in one of two distinct formats:

Method I Treat each allele count x_{ij} as a categorical variable and use them directly as predictors in the regression model. This approach results in $2m$ categorical variables.

Method II Represent the genotype at each locus as a single categorical variable. Specifically, define $u_i = (x_{ia}, x_{ib})$ as the predictor for locus i. While this method reduces the number of predictors to m, each variable contains a greater number of categorical levels compared to Method I.

As an output, logistic regression returns the log-odd values, which we can convert them to the probabilities of each DNA profile belonging to different subpopulations. In notation, for a DNA profile $\boldsymbol{X}$, we denote

$$P_M(Y = i \mid \boldsymbol{X})$$

as the probability that $\boldsymbol{X}$ is classified into the ith subpopulation, where $i = 1, \ldots, R$, using logistic regression Method M, where $M \in \{\text{I}, \text{II}\}$.

It is important to note that logistic regression can produce different results depending on the number of predictors included, as demonstrated in several studies, such as [23]. This variability can affect the outcomes of our study. In our case, we work with allele frequencies, which can be viewed in two ways: as an allele or as a genotype. This introduces variation. When we treat the predictor as a genotype, the number of predictors used is half that of treating it as an allele. This difference is one of the primary reasons for considering variations of logistic regression, and it is essential to approach the problem with multiple variations for a comprehensive analysis.

Another factor contributing to the variations is that our sample consists of two individuals, giving us the option to classify them simultaneously or separately. If we opt for separate classification, we encounter an additional challenge: What if the results from separate classifications do not match? What method should we use to conclude the statistic? These issues lead us to define five different variations, which we will outline below as LRMLRa, LRMLRb, LRMLRc, LRMLRd, and LRMLRe. While more variations may exist, we have selected these five candidates for our study. We now explain five variations based on logistic regression classification, outlined as follows.

- For the first variation, given DNA profiles $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, we separately calculate $P_{\text{I}}(Y = i \mid \boldsymbol{X}_1)$ and $P_{\text{I}}(Y = i \mid \boldsymbol{X}_1)$. Note that all calculations are performed using logistic regression with Method I. We then use the multiplication rule to determine the probability of both DNA profiles belonging to each subpopulation, denoted by $P_{\text{I}}(Y = i \mid \boldsymbol{X}_1, \boldsymbol{X}_2)$. That is, we have
$$P_{\text{I}}(Y = i \mid \boldsymbol{X}_1, \boldsymbol{X}_2) = P_{\text{I}}(Y = i \mid \boldsymbol{X}_1) P_{\text{I}}(Y = i \mid \boldsymbol{X}_2)$$
for $i = 1, 2 \ldots, R$. The subpopulation corresponding to the highest probability is then selected. Following this, we apply the likelihood ratio statistic within a single population framework, based on the results from the classification. This statistic, referred to as LRMLRa, can be expressed mathematically as:
$$\text{LRMLRa} := \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}$$
where $r = \underset{i \in \{1, \ldots, R\}}{\arg\max}\, P_{\text{I}}(Y = i \mid \boldsymbol{X}_1, \boldsymbol{X}_2)$.

- Second, we follow the same process as in the first variation, but this time we use logistic regression with Method II instead of Method I. The corresponding likelihood ratio statistic is denoted as LRMLRb and expressed as:
$$\text{LRMLRb} := \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}$$
where $r = \underset{i \in \{1, \ldots, R\}}{\arg\max}\, P_{\text{II}}(Y = i \mid \boldsymbol{X}_1, \boldsymbol{X}_2) = \underset{i \in \{1, \ldots, R\}}{\arg\max}\, \{P_{\text{II}}(Y = i \mid \boldsymbol{X}_1) P_{\text{II}}(Y = i \mid \boldsymbol{X}_2)\}$.

- Next, given two DNA profiles $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, we classify the subpopulation to which the first DNA profile belongs using logistic regression Method I. The same process is then applied to the second DNA profile. If the two DNA profiles are classified into the same subpopulation, the likelihood ratio statistic under a single-population framework, based on the

classification result, is used. However, if the two DNA profiles are classified into different subpopulations, the likelihood ratio is computed using the weighted average of allele frequencies from the predicted subpopulation results for the two DNA profiles. Specifically, if the first DNA profile is classified into the ith subpopulation and the second DNA profile is classified into the jth subpopulation, $(i \neq j)$, the weighted average of allele frequencies, denoted as $\boldsymbol{f}_{ij}$, is defined by

$$\boldsymbol{f}_{ij} := \frac{p_i \boldsymbol{f}_i + p_j \boldsymbol{f}_j}{p_i + p_j},$$

where $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$ are the allele frequencies for the ith and j^{th} subpopulations, respectively, and p_i and p_j are their corresponding proportions. The corresponding LR statistic is denoted as LRMLRc and can be expressed mathematically as:

$$\text{LRMLRc} := \begin{cases} \dfrac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}, & \text{if } r = s, \\ \dfrac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{rs})}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{rs})}, & \text{if } r \neq s, \end{cases}$$

where $r = \underset{i \in \{1,\ldots,R\}}{\arg\max}\, P_{\text{I}}(Y = i \mid \boldsymbol{X}_1)$ and $s = \underset{j \in \{1,\ldots,R\}}{\arg\max}\, P_{\text{I}}(Y = j \mid \boldsymbol{X}_2)$.

- Similarly, the fourth variation uses the same definition as LRMLRc, except that we use logistic regression with Method II instead of Method I. The LR statistic for this variation is denoted as LRMLRd and can be expressed mathematically as:

$$\text{LRMLRd} := \begin{cases} \dfrac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}, & \text{if } r = s, \\ \dfrac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{rs})}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_{rs})}, & \text{if } r \neq s, \end{cases}$$

where $r = \underset{i \in \{1,\ldots,R\}}{\arg\max}\, P_{\text{II}}(Y = i \mid \boldsymbol{X}_1)$ and $s = \underset{j \in \{1,\ldots,R\}}{\arg\max}\, P_{\text{II}}(Y = j \mid \boldsymbol{X}_2)$.

- Lastly, given two DNA profiles $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, each with m loci, represented as:

$$\{(x_{1a}, x_{1b}), (x_{2a}, x_{2b}), \ldots, (x_{ma}, x_{mb})\} \text{ and } \{(y_{1a}, y_{1b}), (y_{2a}, y_{2b}), \ldots, (y_{ma}, y_{mb})\}.$$

It is possible to conduct logistic regression to classify both DNA profiles into a single subpopulation simultaneously, as described below:

Method III We use the genotypes from both individuals as categorical predictor variables in the logistic regression model, resulting in a total of $2m$ predictors.

Using this method, both individuals are classified simultaneously by selecting the subpopulation with the highest probability. The likelihood ratio statistic, referred to as LRMLRe, is then calculated based on the classification results and is defined as:

$$\text{LRMLRe} := \frac{L(\boldsymbol{\theta}_1 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}{L(\boldsymbol{\theta}_0 \mid \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{f}_r)}$$

where $r = \underset{i \in \{1,\ldots,R\}}{\arg\max}\, P_{\text{III}}(Y = i \mid \boldsymbol{X}_1, \boldsymbol{X}_2)$.

2.3 Simulation Study for Power Comparison

To evaluate the performance of the existing statistics in comparison to the proposed statistic, we assess them based on their statistical power. The power of each statistic will be estimated through simulations. It is important to note that, to the best of our knowledge, the actual distribution of each likelihood ratio statistic cannot be derived, and using the chi-square distribution as an approximation for the likelihood ratio is not valid in this case, as the sample size is only two.

The Thai population was chosen as a case study because previous archaeological and linguistic studies have provided evidence supporting the idea that Thailand has a complex genetic structure [24–26]. For simplicity, we assume four subpopulations in Thailand based on the main regions where Thai people reside: North, Northeastern, Central, and South. The allele frequency distributions for each subpopulation are derived from Shotivaranon et al. [13], in which DNA was extracted from 929 unrelated Thai individuals at $m = 15$ loci, including FGA, TH01, TPOX, CSF1PO, vWA, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, and D21S11. The sample sizes for the North, Northeastern, Central, and South subgroups are 202, 304, 212, and 211, respectively. Note that details on age and gender composition, as well as information on how the sample was recruited and the inclusion/exclusion criteria, are not publicly available due to ethical considerations.

Recall that power is defined as the probability of correctly rejecting the null hypothesis when the alternative hypothesis is true. To calculate this, we first generate 1,000,000 pairs under the null hypothesis to establish a rejection criterion, and then simulate 1,000,000 pairs under the alternative hypothesis to estimate the power. The simulation steps followed in this study, which are the same as those presented in [19, 22], are outlined as follows:

1. Simulate a pair of unrelated DNA profiles using 15 loci. The subpopulation of each individual is randomly assigned based on a multinomial distribution with success probabilities $p_{NO} = 0.11083$, $p_{NE} = 0.36944$, $p_{CT} = 0.35383$, $p_{SO} = 0.16590$. These probabilities are estimated from the Thai population database [27].

2. Calculate the likelihood ratio statistics for the pair generated in Step 1 using all different formulas in consideration: LRLAF, LRAVG, LRMIN, LRMAX, LRNB, LRMLRa, LRMLRb, LRMLRc, LRMLRd and LRMLRe. Save the resulting values.

3. Repeat Steps 1 and 2 multiple times, for instance, $B = 1,000,000$ times. This will generate the null distribution of likelihood ratios.

4. Estimate a threshold c from the null distribution obtained in Step 3. Specifically, we determine c such that $P(\text{LR} > c) = \alpha$, where α represents the specified false positive rate.

5. Generate a pair of DNA profiles based on the relationship specified in the alternative hypothesis (using 15 loci). For each pair, the subpopulation information of the first individual is randomly selected from a multinomial distribution, as described in Step 1, while the subpopulation information of the second individual is set to match the first, since they are genealogically related.

6. Calculate the likelihood ratio statistics for the pair drawn in Step 5 using all different formulas in consideration: LRLAF, LRAVG, LRMIN, LRMAX, LRNB, LRMLRa, LRMLRb, LRMLRc, LRMLRd and LRMLRe. Save the resulting values.

7. Repeat Steps 5 and 6 multiple times (for instance, $B = 1,000,000$ times) to obtain the alternative distribution of likelihood ratios based on the generated samples.

8. Estimate the power by calculating the ratio of the number of pairs generated in Step 7 with statistics exceeding the threshold c (found in Step 4) to the total number of simulated pairs.

Once the power for each statistic is obtained, we compare them at the same false positive rate α.

3 Main Results

This section presents the power comparison results from the simulation study from three perspectives: assessing the power of all methods at a single point, evaluating their power across a reasonable range, and conducting pairwise comparisons to detect differences in performance between two specific test statistics.

3.1 Comparison at a Single Point

The performance of each statistic can be compared using power; however, power is also influenced by two factors: the threshold for rejection and the false positive rate (FPR). To ensure a fair and valid comparison, it is important to evaluate the power of all candidates at the same FPR. This typically results in different null thresholds being applied for each statistic. According to [19], the Denver Crime Lab utilized a null threshold of 100,000 with the LRMAX statistic for both parent-child and full-sibling tests, which can serve as a starting point for determining our FPR.

Under the null hypothesis assumption, we found that 16 Thai parent-child pairs out of 1 million generated pairs had an LRMAX exceeding 100,000, resulting in an estimated FPR of 1.6×10^{-5}. Similarly, 17 Thai full-sibling pairs out of 1 million exceeded this threshold, leading to an estimated FPR of 1.7×10^{-5}.

In the parent-child test, the rejection region for each LR statistic is $[c, \infty)$ where c satisfies the condition $P(\text{LR statistic} > c) = 1.6 \times 10^{-5}$. The values c obtained from the simulation are provided in the Table 4. Once the rejection region is determined, power is estimated by calculating the proportion of parent-child pairs with an LR statistic greater than or equal to c. For example, the parent-child null threshold of LRLAF at $\alpha = 1.6 \times 10^{-5}$ is 18,193.7. We found that 740,380 Thai parent-child pairs out of 1 million generated pairs had the LRLAF statistic exceed 18,193.7, yielding an estimated power of 0.74038. The null thresholds and power for full-sibling tests are defined similarly.

Table 5 shows the power for all LR statistics at the estimated FPR α. For the parent-child tests, LRMLRd has the highest power (0.7552), while LRMAX has the lowest (0.6109). However, it is observed that 8 out of 10 statistics have parent-child power around $74-75\%$. The differences are not crystal-clear, and it is unclear if they are significant. A similar pattern is observed for the full-sibling power. The most noticeable gap appears between the leading group (8 out of 10 statistics) and LRAVG, followed by LRMAX. To assess whether the difference within the leading group is significant, we can calculate the 95% confidence interval (CI) using Bootstrap resampling with the algorithm provided below:

Table 4: The null thresholds c for parent-child (PC) full-sibling (SB) tests.

Statistics	PC thresholds c at FPR 1.6×10^{-5}	SB thresholds c at FPR 1.7×10^{-5}
LRLAF	18193.7	13796.4
LRAVG	41610.6	38974.5
LRMAX	127599.0	100826.0
LRMIN	10109.0	7275.3
LRNB	11337.0	8005.0
LRMLRa	12082.2	9066.6
LRMLRb	12082.2	9066.6
LRMLRc	11337.0	9066.6
LRMLRd	11337.0	8354.5
LRMLRe	11337.0	8674.3

1. Randomly select 1 million pairs of DNA profiles from the alternative distribution in the simulated data with replacement. Note that the same pair may be selected multiple times due to the sampling with replacement.

2. Estimate the power by calculating the ratio of the number of pairs from Step 1 with LR statistics greater than the threshold c to the total number of simulated pairs (1,000,000 pairs).

3. Repeat Steps 1 and 2 multiple times (for example, $B = 10{,}000$ times) to obtain B different power estimates.

4. Sort the B power estimates from Step 3 and find the values at the 2.5^{th} percentile and the 97.5^{th} percentile. These two values form the 95% bootstrap percentile CI for the power estimate.

With the help of the confidence interval, we can clearly determine if the difference is significant. Specifically, if the two CIs do not overlap, it may indicate that the difference is significant.

3.2 Comparison across a Reasonable Range

To assess the consistency of the power ranking across all LR statistics, power is also estimated for FPR values ranging from 0 to 10^{-4}, with an increment step of 10^{-6}. Figure 1 on the left presents a plot of power versus FPR for the parent-child test, with a dashed vertical line marking the estimated FPR of 1.6×10^{-5}. Each line represents a different statistic, with labels provided in the legend. Figure 1 on the right shows the same plot for the full-sibling test, with a dashed vertical line marking the estimated FPR of 1.7×10^{-5}. The ranking pattern among the LR statistics appears to be consistent across the range of considered FPRs. Specifically, the leading group (8 out of 10 statistics) continues to perform well, with no clear-cut significance differences. The bottom two statistics, LRAVG and LRMAX, are clearly defined as the least performing.

Table 5: Comparison of parent-child (PC) power at FPR 1.6×10^{-5} and full-sibling (SB) power at FPR 1.7×10^{-5}.

Statistics	PC Power	95% CI	SB Power	95% CI
LRLAF	0.7404	(0.7395, 0.7412)	0.5096	(0.5087, 0.5106)
LRAVG	0.6782	(0.6773, 0.6791)	0.4634	(0.4624, 0.4644)
LRMAX	0.6109	(0.6100, 0.6118)	0.4436	(0.4426, 0.4446)
LRMIN	0.7425	(0.7416, 0.7433)	0.5127	(0.5116, 0.5137)
LRNB	0.7496	(0.7487, 0.7504)	0.5163	(0.5153, 0.5173)
LRMLRa	0.7485	(0.7476, 0.7494)	0.5105	(0.5096, 0.5115)
LRMLRb	0.7484	(0.7476, 0.7493)	0.5107	(0.5097, 0.5115)
LRMLRc	0.7551	(0.7542, 0.7559)	0.5096	(0.5086, 0.5106)
LRMLRd	0.7552	(0.7544, 0.7561)	0.5164	(0.5154, 0.5173)
LRMLRe	0.7523	(0.7515, 0.7533)	0.5119	(0.5109, 0.5129)

To further assess the performance of the LR statistics, we introduce an additional criterion called the partial area under the curve (pAUC). This is defined similarly to the regular AUC in typical statistics, but with the domain of the x-axis restricted to the range of interest. Since the power quickly converges to 1, using the full AUC to calculate the entire area under the curve may not be ideal, as the results would be identical in the first four digits. Instead, pAUC is more helpful for comparison. In this particular context, we compute pAUC starting from $\alpha = 0$ to 0.0002. The corresponding results are shown in Table 6, where we also include a column displaying the proportion of pAUC at 0.0002.

Table 6: Comparison of pAUC of parent-child (PC) power and full-sibling (SB) power versus FPR within range $\alpha \in [0, 0.0002]$.

Statistics	PC pAUC ($\times 10^{-4}$)	Proportion (PC pAUC/0.0002)	SB pAUC ($\times 10^{-4}$)	Proportion (SB pAUC/0.0002)
LRLAF	1.8383	0.9192	1.2992	0.6496
LRAVG	1.8009	0.9005	1.2500	0.6250
LRMAX	1.7598	0.8799	1.1769	0.5885
LRMIN	1.8422	0.9211	1.3158	0.6579
LRNB	1.8450	0.9225	1.3184	0.6592
LRMLRa	1.8466	0.9233	1.3190	0.6595
LRMLRb	1.8461	0.9231	1.3195	0.6597
LRMLRc	1.8468	0.9234	1.3188	0.6594
LRMLRd	1.8466	0.9233	1.3184	0.6592
LRMLRe	1.8461	0.9231	1.3200	0.6600

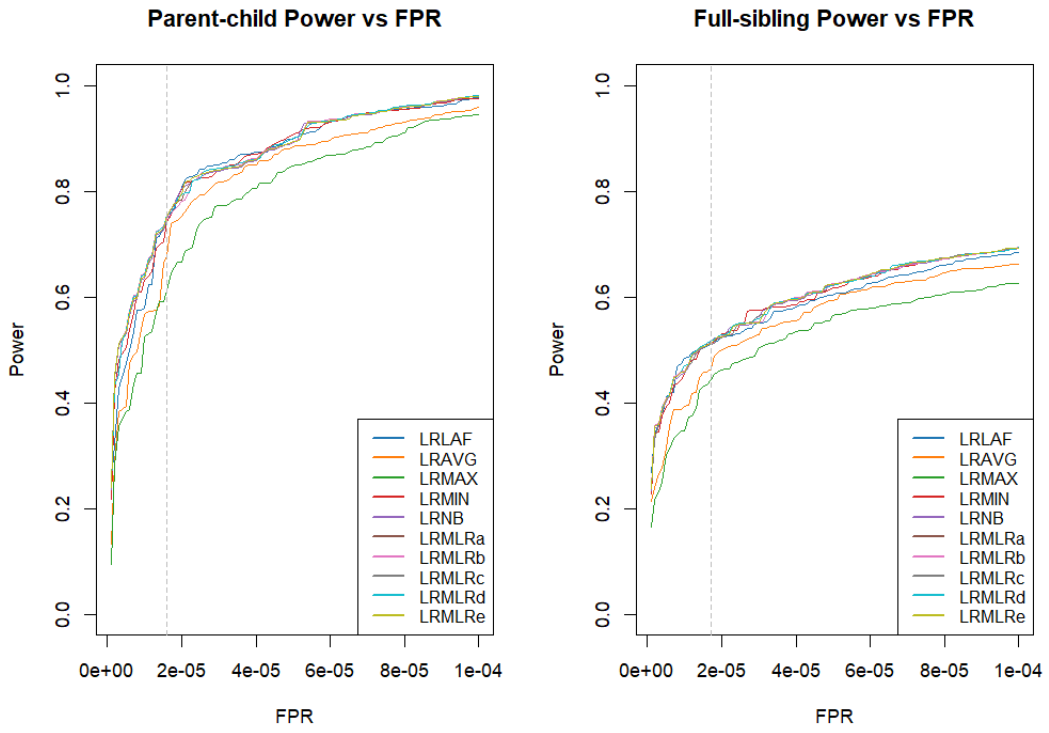


Figure 1: Plots of power versus FPR for ten different LR statistics in parent-child and full-sibling tests.

From Table 6, LRMLRc exhibits the highest pAUC for the parent-child test (1.8468×10^{-4}) corresponding to a proportion of 0.9234, while LRMLRe shows the highest pAUC for the full-sibling test (1.3200×10^{-4}), with a proportion of 0.6600. The ranking of statistics is almost consistent with the comparison found in Section 3.1. Specifically, the bottom two statistics remain the same. However, LRLAF appears to deviate slightly from the leading group, leaving the remaining 7 statistics within the leading group, with no clear distinction in pAUC.

3.3 Pairwise Comparison in pAUC

Based on Table 6, a number of statistics performed comparably, prompting us to investigate whether there are significant differences in performance among the LR statistics. In this subsection, we aim to construct a 95% confidence interval (CI) for the difference in pAUC between two selected LR statistics using the Bootstrap resampling method. If the CI includes zero, we conclude that there is no difference in pAUC between the two LR statistics, suggesting comparable performance. Conversely, if the CI is entirely above or below zero, it indicates a significant difference. The relationships of interest we focus on are parent-child and full-sibling. The algorithm for constructing a 95% CI for the difference in pAUC between two selected LR statistics is provided below:

1. Pick a pair of statistics that we want to construct a confidence interval for.
2. Select 200 values of $\alpha \in [0, 0.0002]$ with replacement, denoted as $\alpha_1, \alpha_2, \ldots, \alpha_{200}$, where each α can possibly have the same value.
3. Extract the power for the first statistic in consideration corresponding to each α chosen in Step 2, then calculate the pAUC for this statistic based on these 200 values. Repeat for the other selected LR statistic.
4. Use the output from Step 3 to calculate the difference in pAUC between the two LR statistics.
5. Repeat Steps 1 to 4 multiple times (for example, 10,000 times) to obtain a distribution of differences in pAUC between the two LR statistics.
6. Sort the collected differences from Step 6 and identify the values at the 2.5^{th} percentile and the 97.5^{th} percentile. These two values form the 95% confidence interval for the difference in pAUC between the two selected LR statistics.

Figure 2 shows the 95% CIs for the differences in pAUC between parent-child power and full-sibling power across all pairs of statistics considered. The figures highlight that some LR statistics exhibit significant differences in performance, indicated by red CIs. For the parent-child relationship, LRMAX shows a significant difference compared to all other LR statistics, as does LRAVG. In contrast, there is no significant difference among LRMIN, LRNB, and all variations of LRMLR. The results for the full-sibling relationship follow the same pattern.

4 Conclusion and Discussion

This study provides a comprehensive evaluation of ten likelihood ratio statistics used to assess parent-child and full-sibling relationships: LRLAF, LRAVG, LRMAX, LRMIN, LRNB, LRMLRa, LRMLRb, LRMLRc, LRMLRd, and LRMLRe. The performance of these statistics is compared based on estimated false positive rates and their ranges, with the aid of partial AUC analysis. Overall, it is recommended to use LR statistics derived from classification methods (LRNB and LRMLR), along with LRLAF and LRMIN, for analyzing Thai parent-child and Thai full-sibling relationships, as they all exhibit high and comparable power as shown in Table 5 and Figure 1.

The partial AUC (pAUC) results offer further insights into the discriminative power of these LR statistics under low FPR ($\alpha \in [0, 0.0002]$), which is particularly important in forensic and genetic applications. As shown in Table 6, we observed a weak indication that LRLAF performs less effectively compared to the leading group, which includes LRMIN, LRNB, and all variations of LRMLR.

Furthermore, we utilized the Bootstrap resampling method to assess significant differences in pAUC among these LR statistics. Significance was determined by confidence intervals that do not include zero. The results, presented in Figure 2, suggest that there is no significant difference between LRLAF, LRMIN, LRNB, and all variations of LRMLR regarding pAUC for both Thai parent-child and Thai full-sibling tests. This confirms that our proposed method, LRMLR, performs just as well as those found in previous literature. While no single method emerged as definitively superior, the availability of multiple effective approaches provides flexibility in

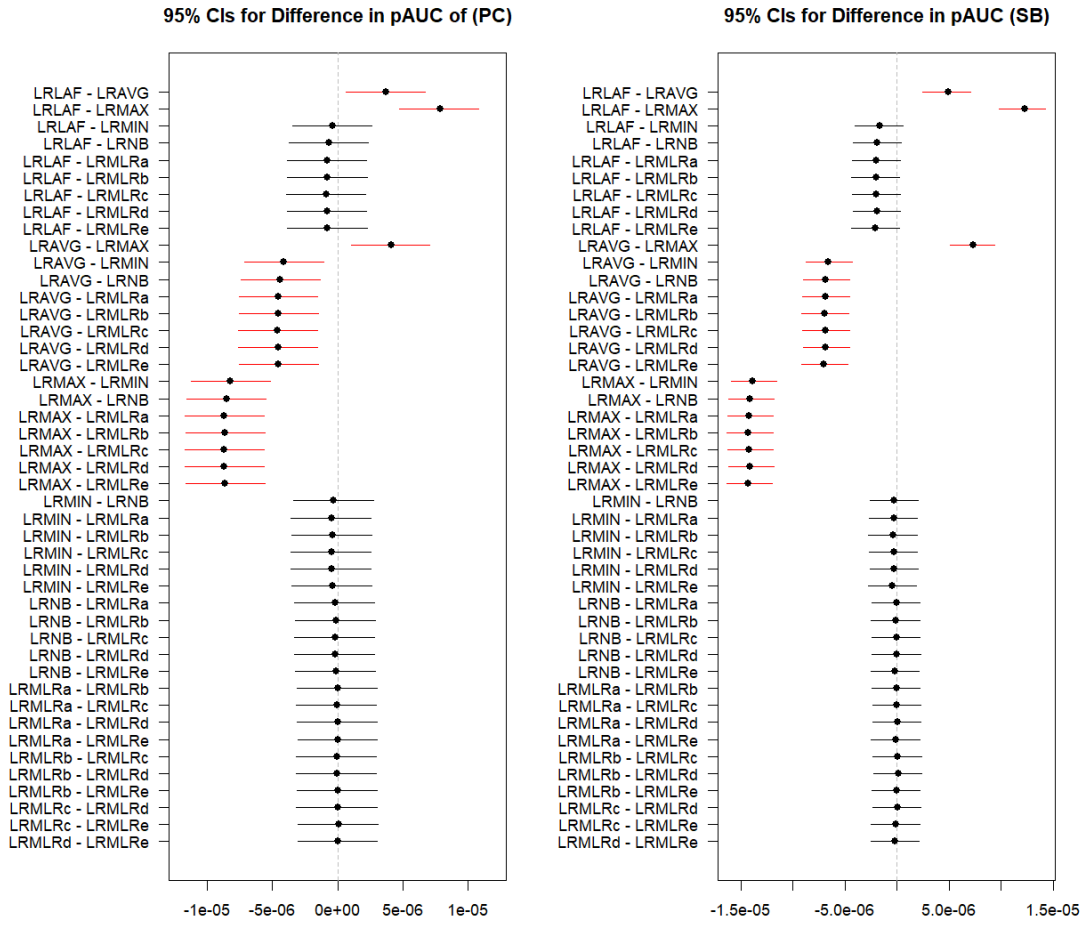


Figure 2: 95% CIs, derived using Bootstrap resampling, for the difference in pAUC of power versus FPR between two selected LR statistics for the test of a parent-child (PC) and full-sibling (SB) relationship.

practical applications, allowing practitioners to choose the most suitable method based on factors such as data availability or computational resources.

The LR statistics, LRNB and LRMLR, are derived from the classification algorithms Naive Bayes and logistic regression, respectively, to assign DNA profiles to specific subpopulations prior to computing the likelihood ratios. Both are supervised learning techniques suitable for binary or multi-class classification tasks and are used to estimate class membership probabilities. Naive Bayes is a generative model that assumes conditional independence among features given the class, while logistic regression is a discriminative model that estimates class probabilities directly, without relying on such assumptions. Logistic regression is generally more robust when dealing with correlated features and offers more interpretable coefficients, whereas Naive Bayes is computationally efficient and performs well with smaller or high-dimensional datasets.

In our context, we recommend using LRNB when computational simplicity is a priority, as it classifies both individuals simultaneously. However, if separate classification of each individual

is preferred, LRMLR is more appropriate, albeit with slightly higher computational cost. Notably, both LRNB and LRMLR yield comparable performance, with no statistically significant difference observed between them.

In concluding this paper, we suggest several directions for future research. While the LRMLR statistics have demonstrated their effectiveness, there may be other statistics that offer superior power. Currently, there is no established theory that identifies the most universally powerful statistic for this context. One possible avenue for further investigation is to apply other machine learning or classification methods to first identify the subpopulation a DNA profile belongs to, and then use a likelihood ratio specific to that subpopulation. Advanced techniques, such as Decision Trees, Random Forests, Support Vector Machines, and Neural Networks, could be useful but would need to be adapted for this context. Although we have found that the Naive Bayes classifier and various logistic regression methods yield similar results, exploring these advanced techniques could potentially improve performance in relatedness testing. Additionally, our analysis was based on a Thai population with regional substructure. A valuable next step would be to replicate these tests in different populations with distinct substructures. Comparing results across diverse populations would help us better understand the limitations of our findings and how subpopulation structure may impact relatedness testing in different genetic contexts.

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments.

References

[1] D.J. Balding, M. Krawczak, J.S. Buckleton, J.M. Curran, Decision-making in familial database searching: KI alone or not alone?, Forensic Sci. Int. Genet. 7 (1) (2013) 52–54.

[2] F.R. Bieber, C.H. Brenner, D. Lazer, Finding criminals through DNA of their relatives, Science 312 (5778) (2006) 1315–1316.

[3] J. Ge, R. Chakraborty, A. Eisenberg, B. Budowle, Comparisons of familial DNA database searching strategies, J. Forensic Sci. 56 (6) (2011) 1448–1456.

[4] M. Kruijver, R. Meester, K. Slooten, Optimal strategies for familial searching, Forensic Sci. Int. Genet. 13 (2014) 90–103.

[5] K. Slooten, R. Meester, Probabilistic strategies for familial DNA searching, J. R. Stat. Soc. C Appl. Stat. 63 (3) (2014) 361–384.

[6] J.M. Curran, J.S. Buckleton, Effectiveness of familial searches, Sci. Justice 48 (4) (2008) 164–167.

[7] J. Ge, B. Budowle, Forensic investigation approaches of searching relatives in DNA databases, J. Forensic Sci. 66 (2) (2021) 430–443.

[8] S.H. Katsanis, Pedigrees and perpetrators: uses of DNA and genealogy in forensic investigations, Annu. Rev. Genomics Hum. Genet. 21 (1) (2020) 535–564.

[9] C. Tian, P.K. Gregersen, M.F. Seldin, Accounting for ancestry: population substructure and genome-wide association studies, Hum. Mol. Genet. 17 (R2) (2008) R143–R150.

[10] G.A. Satten, W.D. Flanders, Q. Yang, Accounting for unmeasured population substructure in case-control studies of genetic association using a novel latent-class model, Am. J. Hum. Genet. 68 (2) (2001) 466–477.

[11] S.J. Walsh, R.J. Mitchell, F. Torpy, J.S. Buckleton, Use of subpopulation data in Australian forensic DNA casework, *Forensic Sci. Int. Genet.* 1 (3-4) (2007) 238–246.

[12] D.J. Balding, R.A. Nichols, DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands, Forensic Sci. Int. 64 (2-3) (1994) 125–140.

[13] J. Shotivaranon, T. Chirachariyavej, N. Leetrakool, B. Rerkamnuaychoke, DNA database of populations from different parts in the Kingdom of Thailand, Forensic Sci. Int. Genet. 4 (1) (2009) e37–e38.

[14] M. Lynch, Genetics and analysis of quantitative traits, Sinauer, 1998.

[15] S. Presciuttini, C. Toni, E. Tempestini, S. Verdiani, L. Casarino, I. Spinetti, F.D. Stefano, R. Domenici, J.E. Bailey-Wilson, Inferring relationships between pairs of individuals from locus heterozygosities, BMC Genet. 3 (2002) 1–11.

[16] K.L. O'Connor, C.R. Hill, P.M. Vallone, J.M. Butler, Linkage disequilibrium analysis of D12S391 and vWA in US population and paternity samples, Forensic Sci. Int. Genet. 5 (5) (2011) 538–540.

[17] B. Budowle, J. Ge, R. Chakraborty, A.J. Eisenberg, R. Green, J. Mulero, R. Lagace, L. Hennessy, Population genetic analyses of the NGM STR loci, Int. J. Legal Med. 125 (2011) 101–109.

[18] J. Ge, A. Eisenberg, B. Budowle, Developing criteria and data to determine best options for expanding the core CODIS loci, Investig. Genet. 3 (2012) 1–14.

[19] M. Kooakachai, G. LaBerge, S.A. Santorico, A new framework to test parent-child and full sibling relationships with population substructure, Forensic Sci. Int. 305 (2019) 110012.

[20] J. Ge, B. Budowle, Kinship index variations among populations and thresholds for familial searching, PLoS One 7 (5) (2012) e37474.

[21] C. Thitiyan, P. Jumnongwut, A statistical power comparison in familial DNA searching: an application to Thailand, Undergraduate Senior Project, Department of Mathematics and Computational Science, Faculty of Science, Chulalongkorn University, (2022).

[22] C. Laoiam, M. Kooakachai, Developing test statistics for familial DNA searches through classification ideas, Proceedings of the 27th Annual Meeting in Mathematics 2023 and International Conference in Number Theory and Applications 2023, AMM2023 & ICNA2023, 31 May–2 June 2023, Department of Mathematics, Faculty of Science, Kasetsart University, Bangkok, Thailand (2023) 141–152

[23] E.Y. Boateng, D.A. Abaye, et al., A review of the logistic regression model with emphasis on medical research, J. Data Anal. Inf. Process. 7 (4) (2019) 190.

[24] W. Kutanan, T. Kitpipit, S. Phetpeng, P. Thanakiatkrai, Forensic STR loci reveal common genetic ancestry of the Thai-Malay Muslims and Thai Buddhists in the deep Southern region of Thailand, J. Hum. Genet. 59 (12) (2014) 675–681.

[25] W. Kutanan, K. Kaen, K. Kaen, K. Kaen, M. Chiang, et al., Mitochondrial DNA-HVR1 variation reveals genetic heterogeneity in Thai-Isan peoples from the lower region of Northeastern Thailand, Adv. Anthropol. 4 (1) (2014) 7.

[26] W. Kutanan, D. Liu, J. Kampuansai, M. Srikummool, S. Srithawong, R. Shoocongdej, S. Sangkhano, S. Ruangchai, P. Pittayaporn, L. Arias, et al., Reconstructing the human genetic history of mainland Southeast Asia: insights from genome-wide data from Thailand and Laos, Mol. Biol. Evol. 38 (8) (2021) 3459–3477.

[27] National Statistical Office, Ministry of Digital Economy and Society, Thailand, Statistical yearbook Thailand 2023, <https://www.nso.go.th/public/e-book/Statistical-Yearbook/SYB-2023/>, 2003 (accessed 24.12.21).

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

Analysis of Liquidity Provider Returns in Automated Market Makers Under Independent Black-Scholes Dynamics

Nuttanicha Jirakupt [1,*] and Parkpoom Phetpradap [1]

[1]Department of Mathematics, Faculty of Science,
Chiang Mai University, Chiang Mai 50200, Thailand

Abstract

Yield farming is an investment strategy within decentralized finance (DeFi), where investors provide liquidity to automated market makers (AMMs) in cryptocurrency markets to earn rewards, which are influenced by market volatility and trading fees. Liquidity providers (LPs) play a critical role in DeFi by supplying liquidity to decentralized systems; however, their returns are highly sensitive to market conditions. Popular platforms for crypto yield farming include Uniswap, Aave, Compound, PancakeSwap, and Binance.

This study investigates the theoretical return on investment of LPs in pair trading within constant product AMMs, assuming the underlying assets follow Black-Scholes dynamics. The findings emphasize the importance of market conditions in determining profitability and offering a framework for understanding and mitigating financial risks in DeFi.

Keywords: Automated Market Maker, Decentralized Exchange, Cryptocurrency, Mathematical Finance, Black-Scholes

2020 MSC: 60J70, 91G10, 91B82

1 Introduction

For over a decade, *cryptocurrency* [1] has become a symbol of the financial industrial revolution since the introduction of Bitcoin, a first cryptocurrency allowing online payments to be sent directly from one party to another without going through a financial institution, as described by Nakamoto (2008) [2], giving rise to a complex ecosystem of digital finance. At its core, cryptocurrency is digital money that operates independently of governments and banks, using blockchain technology to ensure secure and transparent transactions. This financial system is called *decentralized finance*.

Decentralized Finance (DeFi) has emerged as a revolutionary concept within this ecosystem. DeFi provides financial instruments and services through smart contracts on programmable, permissionless blockchains, reducing the need for traditional intermediaries like banks or exchanges [3]. This approach has opened up new possibilities for lending, borrowing, trading, and earning interest on digital assets.

A key component of DeFi is the *Decentralized Exchange (DEX)*, a peer-to-peer marketplace where users can trade cryptocurrencies without a centralized intermediary [4], such as Uniswap

*Speaker and corresponding author (nuttanicha_j@cmu.ac.th)

[5] on the Ethereum. DEXs often utilize *Automated Market Makers (AMMs)*, which are protocols that use smart contracts and liquidity pools to enable cryptocurrency trading. Instead of using an order book [6], an electronic list of buy and sell orders for a security or other instrument organized by price level in traditional finance, more recent attempts at establishing DEXs have revolved around the use of automated market makers (AMMs). An AMM automates this by allowing traders to place orders with the AMM, which then algorithmically and automatically provides a price. It is worth reiterating how this process is different from the order book system that requires matches between price and volume orders provided by buyers and sellers.

One prevalent type of AMM is *Constant Product Market Maker (CPMM)*, introduced by Uniswap (2018) [7], which uses the mathematical equation xy=k to maintain a balance between the reserves of token X (x) and token Y (y) in the liquidity pool, where k is a constant. CPMMs offer efficient liquidity provision, lower impermanent loss, and better token price stability compared to traditional finance. It is important to note that CPMMs can be subject to impermanent loss [8], which arises from changes in the price ratio of the deposited assets, potentially resulting in a lower return for *liquidity providers* (those who allocate liquidity to a liquidity pool) compared to simply holding the tokens. Other AMMs employ more complex algorithms depending on the platform's intended use case. For instance, Hybrid AMMs, as shown by Curve Finance which launched in early 2020 [9], emerged as an evolution to address some of the limitations of earlier AMM models like CPMMs. These newer models combine CPMM and Constant Sum Market Maker (CSMM) features to maintain more stable prices, and reduce impermanent loss for liquidity providers with the same amount of capital compared to CPMMs where liquidity is spread across a wider price spectrum.

Yield farming, another innovative DeFi concept, allows users to earn rewards by providing liquidity to various protocols. Yield farmers can deposit their assets into liquidity pools and receive returns in the form of transaction fees, token incentives, or interest on lending deposits [10]. This mechanism has created new opportunities for passive income generation in the cryptocurrency space. However, it is crucial to be aware of the associated risks.

There are several studies and researches that focus on pricing and valuing the cryptocurrency to deepen understand the behavior of their price. For example, in 2020, Clark [11] analyzed the payoff structure of CPMMs, such as those employed by Uniswap. His findings reveal that the replicating portfolio in such markets is structurally short volatility, effectively resembling an options-selling strategy. This characteristic explains the necessity of positive transaction fees to incentivize liquidity providers. Clark further suggests that in financial environments lacking traditional futures and options markets, this payoff structure could serve as a foundation for constructing such derivative markets.

Additionally, Najnudel *et al.* [12] in 2024 developed a model to examine how arbitrage activities influence the pricing mechanisms of AMMs. Assuming that the reference market price follows a geometric Brownian motion, the authors establish that AMM prices are confined within upper and lower bounds, defined as constant multiples of this reference price. Utilizing local time and excursion theory, they derive analytical results, including a time-changed representation of the AMM price process and insights into its long-term behavior. This study advances the theoretical understanding of AMM price dynamics, particularly concerning arbitrage strategies and fee structures. Overall, existing research extensively analyzes cryptocurrency pricing and AMM mechanics. However, the return for liquidity providers remains largely unexplored.

This research analyzes the return on investment for liquidity providers (LPs) in DEXs that utilize two token AMMs with a constant product formula. The study aims to determine the

return of investment for these LPs, assuming that the underlying asset's price from an external reference market follows the independent geometric Brownian motion and that price dynamics within the DEX liquidity pool are driven by an arbitrage strategy. Relevant theorems and basic knowledge of AMMs and yield farming, including calculation examples, are introduced in Section 2. Section 3 then provides the notations and model assumptions for clarity. The main results, detailing the liquidity provider return, are presented in Section 4, followed by the conclusion of this study in Section 5.

2 Preliminaries

2.1 Stochastic Processes and Stochastic Differential Equation

Definition 2.1. (Wiener process. [13]) A Wiener process $W_t = \{W(t) : t > 0\}$ starting from $W(0) = w$, say, is a real-valued Gaussian (Normal) process with the following properties:

(a) W_t has independent increments,

(b) $W_{s+t} - W_s$ is distributed as $N(0, t)$ for all $s, t \geq 0$,

(c) the sample paths of W_t are continuous.

Definition 2.2. (Stochastic process in differential form [14]) A stochastic differential form is a formal expression of the infinitesimal dynamics of a stochastic process, typically written as

$$dX_t = f(t, X_t)\, dt + g(t, X_t)\, dW_t,$$

where $f(t, X_t)$ is the drift coefficient, $g(t, X_t)$ is the diffusion coefficient, and W_t denotes a standard Brownian motion.

Definition 2.3. (Itô's process. [13, 14]) Let a stochastic process of the form

$$X_t = X_0 + \int_0^t \mu(t, X_s)ds + \int_0^t \sigma(t, X_s)dW_s \tag{2.1}$$

be an Itô process. Then, an Itô process of the form (2.1) can be written in the shorter differential form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t.$$

Definition 2.4. (Geometric Brownian motion [15]) The process X_t follows a geometric Brownian motion with the parameters μ and σ^2 if and only if

$$dX_t = \mu X_t dt + \sigma X_t dW_t.$$

Corollary 2.5. *(Solution of geometric Brownian motion [15]) Let X_t follows a geometric Brownian motion given by Definition 2.4. Then, the exact solution of this equation is given by*

$$X_t = X_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t}.$$

Definition 2.6. (Lognormal distribution [14]) The random variable X is said to be a *lognormal* random variable with parameters μ and σ^2 if $Y = \ln X$ is a normal random variable with mean μ and variance σ^2.

Corollary 2.7. *(Expectation and variance of lognormal distribution [14]) Let X and Y be random variables that satisfy Definition 2.6. Then, the expectation and the variance of X are*

$$E(X) = e^{\mu+\frac{\sigma^2}{2}} \quad \text{and} \quad var(X) = e^{2\mu+\sigma^2}(e^{\sigma^2}-1).$$

Corollary 2.8. *(Expectation and variance of geometric Brownian motion [14]) Let X_t be a geometric Brownian motion, as mentioned in Definition 2.4. Then, X_t follows lognormal distribution with parameters μ, and σ^2. Additionally, the expectation and variance of X_t are*

$$E(X_t) = X_0 e^{\mu t} \quad \text{and} \quad var(X_t) = X_0 e^{2\mu t}(e^{\sigma^2 t}-1).$$

Theorem 2.9. *(Itô's lemma [13, 14]) Let X_t be an Itô process given by Definition 2.3. Let $f(t, X_t) \in C^2([0,\infty)\times\mathbb{R})$. Then $Y_t = f(t, X_t)$ is also an Itô process, given by*

$$dY_t = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial X_t}\mu(t, X_t) + \frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}\sigma^2(t, X_t)\right)dt + \frac{\partial f}{\partial X_t t}\sigma(t, X_t)dW_t.$$

Theorem 2.10. *(The multi-dimensional Itô's lemma [13, 14]).*
Let $\boldsymbol{X_t}$ be a multi-dimensional Itô process such that

$$d\mathbf{X_t} = \boldsymbol{\mu}dt + \boldsymbol{\Sigma}d\mathbf{W_t}$$

where $\mathbf{X_t} = \begin{bmatrix} X_t^1 \\ X_t^2 \\ \vdots \\ X_t^n \end{bmatrix}$, $\boldsymbol{\mu} = \begin{bmatrix} \mu_1(t, X_t^1) \\ \mu_2(t, X_t^2) \\ \vdots \\ \mu_n(t, X_t^n) \end{bmatrix}$, $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1m} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nm} \end{bmatrix}$, *and* $d\mathbf{W_t} = \begin{bmatrix} dW_t^1 \\ dW_t^2 \\ \vdots \\ dW_t^m \end{bmatrix}$.

Let $\boldsymbol{Y_t} = f(t, \mathbf{X_t}) = (f_1(t, X_t^1), \ldots, f_p(t, X_t^p))$ be a C^2 map from $[0,\infty)\times\mathbb{R}^n$ into $\mathbb{R}^p$. Itô's lemma then states that

$$d\boldsymbol{Y_t} = \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}dX_t^{(i)} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\partial^2 f}{\partial x_i \partial x_j}dX_t^{(i)}dX_t^{(j)}.$$

This study focus on 2-dimensional system. Let us consider independent Itô processes X_t and Y_t such that

$$dX_t = \mu_1(t, X_t)dt + \sigma_1(t, X_t)dW_t^1,$$
$$\text{and } dY_t = \mu_2(t, Y_t)dt + \sigma_2(t, Y_t)dW_t^2.$$

Let $Z_t = f(t, X_t, Y_t)$. Then, the 2-dimensional Itô's lemma is given by

$$dZ_t = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{\partial f}{\partial y}dY_t + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\sigma_1^2 + \frac{\partial^2 f}{\partial y^2}\sigma_2^2\right)dt. \tag{2.2}$$

2.2 Automated Market Maker

As mentioned earlier, AMMs use algorithmic mechanisms to facilitate the trading of digital assets. Unlike traditional financial markets that use an order book system to match the prices offered by buyers and sellers, AMMs aim to maintain liquidity in the DeFi ecosystem through *liquidity pools*. They allow users, called *liquidity providers (LPs)*, to contribute assets to these pools [10]. The liquidity providers gain the profit not only by speculating but also by receiving

the reward by providing the liquidity. However, there exist some associated risks to get loss. Several types of AMMs exist in DEXs, with one fundamental type being the *Constant Product Market Maker (CPMM)*. CPMMs hold significant importance as they were among the earliest and most influential AMM designs, demonstrating the feasibility of decentralized trading without order books. Their simplicity and effectiveness in providing continuous liquidity for a wide range of assets laid the groundwork for the increase of DeFi and continue to be widely adopted across numerous DEXs. In a CPMM, the product of quantity of two tokens in a liquidity pool is a constant. The first practical example of CPMM was represented by Uniswap (2018) [7], the largest decentralized exchange operating on the Ethereum blockchain.

Assumption 2.11. Let the liquidity pool in decentralized exchange A (DEX-A) contains tokens X and Y and the initial reserves of token X and token Y are x_0 and y_0, respectively. Then, the constant product is given by

$$x_t y_t = k \tag{2.3}$$

where x_t and y_t the reserves of token X and token Y in the liquidity pool. For specifically, an initial condition provides

$$x_0 y_0 = k.$$

The exchange rate of Y in terms of X at time t, p_t, is defined by

$$p_t = \frac{x_t}{y_t}. \tag{2.4}$$

The exchange rate of X is determined by the quantities ratio of the token pair implying that you can exchange 1 unit of token Y for $\frac{x_t}{y_t}$ units of token X. Next, we shall provide some examples to aid understanding.

Example 2.12. Let us consider a liquidity pool with two tokens: Token A and Token B. Initially, the pool contains 1,000 units of token A and 10,000 units of token B.

1. At the initial state, the constant product k is $1,000 \times 10,000 = 10,000,000$.
2. Mr.C wants to exchange 1,000 units of token B for token A from the liquidity pool
3. Adjusting the Pool:
   - After the purchase, the pool has 11,000 units of token B.
   - To maintain the constant product ($k = 10,000,000$), we adjust the amount of token B. The new amount of token B is calculated as: $\frac{k}{y} = \frac{10,000,000}{11,000} = 909.09$
4. At time 1, the pool contains 909.09 units of token A and 11,000 units of token B. Then, Mr.C gain $1,000 - 909.09 = 90.91$ units of token A.

This adjustment ensures that the product of the quantities of token A and token B remains constant at 10,000,000, maintaining the balance of the pool according to the AMM's formula. Moreover, if Mr.C gains 90.91 units of token A from exchanging 1,000 units of token B for token A at time 1.

The following example will illustrate how the level of liquidity in a pool directly affects the exchange rate between the tokens.

Example 2.13. Let the pool contain 1,000,000 units of token A and 10,000,000 units of token B.

1. At the initial state, the constant product k is $1,000,000 \times 10,000,000 = 1 \times 10^{13}$.
2. Mr.C wants to exchange 1,000 units of token B for token A from the liquidity pool
3. Adjusting the Pool:
   - After the purchase, the pool has 10,001,000 units of token B.
   - To maintain the constant product ($k = 10^{13}$), we adjust the amount of token A. The new amount of token A is calculated as: $\frac{k}{y} = \frac{10^{13}}{10,001,000} = 999,900.01$ units of token A
4. At time 1, the pool contains 999,900.01 units of token A and 10,001,000 units of token B. Then, Mr.C gains $1,000,000 - 999,900.01 = 99.99$ A.

Even though 1 unit of token A is worth 10 units of token B, Mr.C got very different amounts of token A when exchanging 1,000 units of token B in two scenarios. With more liquidity in Example 2.13, he got 99.99 units of token A (nearly 100), which is the right amount. With less liquidity in Example 2.12, he got only 90.91 units of token A. This demonstrates that a higher liquidity helps the exchange rate better reflect its real value. Then, developers need to facilitate crowdfunding to establish a high-liquidity pool. They incentivizes users to add liquidity to trading pools by rewarding providers with the fees generated when other users trade with those pools. The investment which users provide an equal-value token pair to a liquidity pool and earn a profit is called *yield farming*. AMM ensures that the product of the quantities remains constant, automatically adjusting prices based on trading activity. However, AMMs are susceptible to impermanent loss [8], which occurs when the value of assets in a pool diverges from their external market value due to price fluctuations. This difference can lead to losses for liquidity providers if they withdraw their assets during unfavorable price conditions. High market volatility makes impermanent loss worse. However, strategies like diversification and new AMM designs can help reduce how much it affects liquidity providers.

2.3 Yield Farming

Yield farming, also referred to as liquidity mining, constitutes a prominent practice within the decentralized finance (DeFi) ecosystem. This mechanism involves the allocation of digital assets by users into DeFi protocols, with the objective of accruing rewards. These rewards, often denominated in the protocol's governance token, serve as an incentivization strategy designed to attract and maintain liquidity [10].

The process of yield farming, specifically focusing on liquidity provision and fee-sharing mechanisms, can be explained as follows. This process is diagrammatically illustrated in Figure 1.

1. Digital Wallet Establishment: Participants, designated as liquidity providers (LPs), initiate the process by establishing a secure digital wallet for the storage of their digital assets.

2. Asset Acquisition and Transactional Costs: LPs subsequently acquire the requisite digital tokens through fiat currency conversion, incurring transaction fees (rates) and blockchain-specific gas fees. The magnitude of gas fees is contingent upon the underlying blockchain infrastructure supporting the DeFi protocol.

3. Liquidity Pool Contribution: The provision of liquidity necessitates the deposition of two distinct digital assets every transactions (6 seconds), ensuring the maintenance of *equivalent asset valuation*. This balanced contribution is crucial for the stability and functionality of the liquidity pool.

4. Reward Distribution Mechanism: In consideration for their liquidity provision, LPs receive rewards, which, for the scope of this analysis, will be exclusively defined as a proportional share of the transactional fees generated within the liquidity pool. These rewards are systematically distributed to the LP's digital wallet at predetermined temporal intervals.

5. Asset Withdrawal Flexibility: LPs retain the prerogative to withdraw their deposited assets at their discretion, thereby affording a degree of liquidity and flexibility.

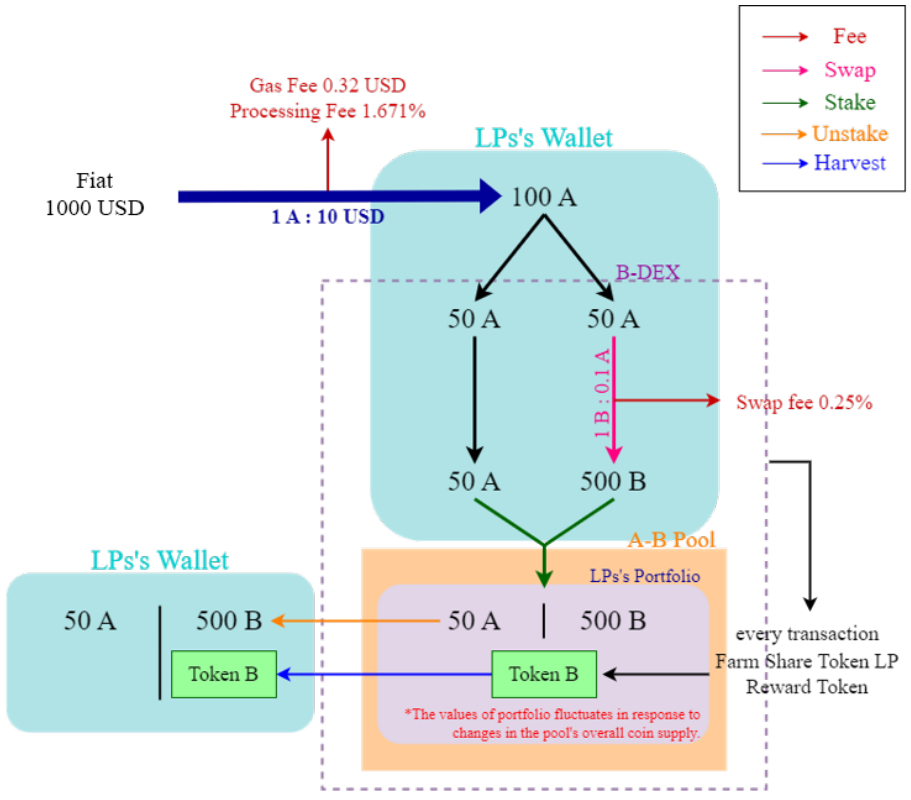


Figure 1: Example of Yield Farming

Example 2.14. Let us consider a liquidity pool with two tokens: Token A and Token B. Initially, the pool contains 1,000,000 units of token A and 10,000,000 units of token B. Mr.Bob needs to deposit 1,000 USD into a yield farming platform

1. At the initial state, exchange rate of token A is 10 USD and exchange rate of token B is 0.1 units of token A. Mr.Bob exchanges 1,000 USD for 100 units of token A and pays fee (gas fee and processing fee).

2. Mr.Bob splits his capital into 2 parts with equivalent value.

   - The first part is 50 units of token A which he holds.
   - For the second part, he exchanges 50 units of token A to $\frac{50}{0.1} = 500$ units of token B. He pays transaction fee in this step.

3. Mr.Bob deposits 50 A:500 B into A-B pool and he pays transaction fee. After this step, he will earn fee share as token B every transactions. Then, his portfolio is 50 A:500 B which portfolio value is 100 units of token A.

4. Assume that the exchange rate of token B in pool shifts to $\frac{900}{11,111.11} = 0.081$ units of token A after the first transaction ($t = 1$).

   - Since the exchange rate changes, the value of token B is $500 \times 0.081 = 40.5$ units of token A which is differ from the value of token A.
   - Due to an equivalent asset valuation in pool, the amount of token A and token B changes to $\frac{1}{2}(50 + 40.5) = 45.25$ units of token A and $\frac{1}{2}(50 + 40.5) \div 0.081 = 558.64$ units of token B.
   - new portfolio value is 90.5 units of token A.

5. Assume that the exchange rate of token B in pool shifts to 0.091 units of token A after the second transaction ($t = 2$).

   - The value of token B is $558.64 \times 0.091 = 50.84$ unita of token A which is differ from the value of token A (45.25 units of token A).
   - Similarly, the amount of token A and token B changes to $\frac{1}{2}(45.25 + 50.84) = 48.04$ units of token A and $\frac{1}{2}(50 + 50.84) \div 0.091 = 527.95$ units of token B.
   - new portfolio value is 96.08 units of token A.

   Totally, the change of his portfolio will represent in a following table.

t	**Exchange rate** p_t **(A)**	**Liquidity Providers**		
		Token A	**Token B**	**VoP (A)**
0	0.10	50	500	100
1	0.081	45.25	558.64	90.5
2	0.091	48.04	527.95	96.08

6. After the second transaction, Mr.Bob withdraws 48.04 A:527.95 B and pays fee. By exchanging token B for token A, he receives roughly 96.08 units of token A. The actual amount of A he gets is a little lower than that, but he also earns some extra reward tokens.

3 Notations and Model Assumption

3.1 Notations

This section provide a list of the notations introduced and used in the study. In this study, we consider a two-market system consisting of a reference market (Decentralized Exchange A, or DEX-A) and an external market (Decentralized Exchange E, or DEX-E), both of which maintain liquidity pools for tokens A and B. It consists of three parts of notation as follow:

1. General Notations
   - X_t, Y_t: Exchange rates of token A and token B (in terms of fiat) in DEX-E
   - Z_t: Exchange rate of token A in terms of token B in DEX-E
   - V_{A_t}, V_{B_t}: Reserves of tokens A and B in DEX-A at time t
   - k: Constant product in the CPMM model
   - S_t: Amount of token A held by the LP in the liquidity pool at time t
   - P_t: Liquidity provider's portfolio value at time t
   - R_t: Return from providing liquidity at time t
2. Stochastic Processes Notations
   - W_{X_t}, W_{Y_t}: Independent standard Wiener processes
   - μ, σ: Drift and volatility terms which are constant
3. Fee and Cost Parameters
   - ϕ: Transaction fee by DEX-A
   - ϕ_p: Processing fee
   - g: Gas fee

3.2 Model Assumption

The following assumptions are established:

1. Constant Product Market Maker (CPMM) in DEX-A: DEX-A operates under a CPMM model as mentioned in Assumption 2.11. To avoid confusion with the exchange rate in DEX-E, the reserves of token A and token B in DEX-A are V_{A_t} and V_{B_t}. Then, The CPMM mechanism ensures that the product of these reserves remains constant
   $$V_{A_t} V_{B_t} = k.$$
   The initial reserves V_{A_0} and V_{B_0} determine the initial constant $k = V_{A_0} V_{B_0}$.
2. Exchange Rates in DEX-E: The exchange rates of token A and token B, denominated in a base currency or another token, in DEX-E are denoted by X and Y, respectively. These rates follow independent Black-Scholes models.
3. Exchange Rates in DEX-A: The exchange rate of token B in terms of token A in DEX-A adjusts to follow the rate in DEX-E with a lag of 1 time unit..

4 Main Results

4.1 Dynamics and Properties of Z_t

To begin with, we observe the dynamics of the exchange rate in DEX-E which determine the exchange rate in DEX-A. The probability distribution of exchange rates of token B in term of token A within DEX-E can be expressed by Lemma 4.1.

Lemma 4.1. *Let X_t and Y_t be two independent geometric Brownian motions. That is*

$$dX_t = \mu_X X_t dt + \sigma_X X_t dW_{X,t}, \quad \text{and} \quad dY_t = \mu_Y Y_t dt + \sigma_Y Y_t dW_{Y,t}$$

which $W_{X,t}$ and $W_{Y,t}$ are independent Wiener processes. Let $Z_t = \dfrac{X_t}{Y_t}$. Then,

$$dZ_t = \left(\mu_X - \mu_Y + \sigma_Y^2\right) Z_t dt + \sqrt{\sigma_X^2 + \sigma_Y^2} Z_t dW_{Z,t}. \tag{4.1}$$

where $W_{Z,t} = \dfrac{\sigma_X W_{X,t} - \sigma_Y W_{Y,t}}{\sqrt{\sigma_X^2 + \sigma_Y^2}}$ which is a standard Wiener process. The exact solution of Equation 4.1 is

$$Z_t = Z_0 e^{[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)]t + \sqrt{\sigma_X^2 + \sigma_Y^2} W_{Z,t}}.$$

Proof. Let X_t and Y_t be geometric Brownian motions with parameters μ_X, μ_Y, σ_X, and σ_Y and independent. Then,

$$dX_t = \mu_X X_t dt + \sigma_X X_t dW_{X,t}, \quad \text{and} \quad dY_t = \mu_Y Y_t dt + \sigma_Y Y_t dW_{Y,t},$$

such that $W_{X,t}$, and $W_{Y,t}$ are standard Wiener process.

Note that $Z_t = f(X_t, Y_t) = \dfrac{X_t}{Y_t}$. We then have $\dfrac{\partial f}{\partial t} = 0$, $\dfrac{\partial f}{\partial X_t} = \dfrac{1}{Y_t}$, $\dfrac{\partial f}{\partial Y_t} = -\dfrac{X_t}{Y_t^2}$, $\dfrac{\partial^2 f}{\partial X_t^2} = 0$, and $\dfrac{\partial^2 f}{\partial Y_t^2} = \dfrac{2X_t}{Y_t^3}$.

By Equation (2.2), we obtain

$$\begin{aligned} dZ_t =& \frac{1}{Y_t}(\mu_X X_t dt + \sigma_X X_t dW_{X,t}) - \frac{X_t}{Y_t^2}(\mu_Y Y_t dt + \sigma_Y Y_t dW_{Y,t}) + \frac{1}{2}\left(\frac{2X_t}{Y_t^3}(\sigma_Y Y_t)^2\right) dt. \\ =& \left(\frac{1}{Y_t}\mu_X X_t - \frac{X_t}{Y_t^2}\mu_Y Y_t + \frac{1}{2}\left(\frac{2X_t}{Y_t^3}\sigma_Y^2 Y_t^2\right)\right) dt + \frac{1}{Y_t}\sigma_X X_t dW_{X,t} - \frac{X_t}{Y_t^2}\sigma_Y Y_t dW_{Y,t} \\ =& \left(Z_t\mu_X - Z_t\mu_Y + Z_t\sigma_Y^2\right) dt + Z_t\sigma_X dW_{X,t} - Z_t\sigma_Y dW_{Y,t} \end{aligned}$$

For simplicity, the form of $\sigma_X dW_{X,t} - \sigma_Y dW_{Y,t}$ can be reduced.

Let $W_{Z,t} = \dfrac{\sigma_X W_{X,t} - \sigma_Y W_{Y,t}}{\sqrt{\sigma_X^2 + \sigma_Y^2}}$. Note that

$$E(W_{Z,t}) = E\left(\frac{\sigma_X W_{X,t} - \sigma_Y W_{Y,t}}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right) = \frac{\sigma_X E\left(W_{X,t}\right) - \sigma_Y E\left(W_{Y,t}\right)}{\sqrt{\sigma_X^2 + \sigma_Y^2}} = 0$$

and $var(W_{Z,t}) = var\left(\dfrac{\sigma_X W_{X,t} - \sigma_Y W_{Y,t}}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right) = \dfrac{\sigma_X^2 var\left(W_{X,t}\right) + \sigma_Y^2 var\left(W_{Y,t}\right)}{\sigma_X^2 + \sigma_Y^2} = \dfrac{\sigma_X^2 t + \sigma_Y^2 t}{\sigma_X^2 + \sigma_Y^2} = t.$

This ensures that $W_{Z,t}$ is also a standard Wiener process, $W_{Z,t} \sim N(0,t)$. Then, $\sigma_X dW_{X,t} - \sigma_Y dW_{Y,t} = \sqrt{\sigma_X^2 + \sigma_Y^2} dW_{Z,t}$.

Thus, dZ_t can be rewritten as

$$dZ_t = Z_t\left(\mu_X - \mu_Y + \sigma_Y^2\right)dt + Z_t\sqrt{\sigma_X^2 + \sigma_Y^2}dW_{Z,t}.$$

Observe that Z_t follows a geometric Brownian motion with the parameters $\mu_X - \mu_Y + \sigma_Y^2$ and $\sigma_X^2 + \sigma_Y^2$, by Definition 2.4. Additionally, by Corollary 2.5, the exact solution of Z_t is

$$Z_t = Z_0 e^{[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)]t + \sqrt{\sigma_X^2 + \sigma_Y^2}W_{Z,t}}. \qquad \square$$

Another important result is that Z_t follow *lognormal* distribution which will be derived in the next theorem.

Theorem 4.2. *Let Z_t be denoted by Lemma 4.1 and let $U_t = \ln Z_t$. Then, U_t is a normal distribution and Z_t is lognormal distribution with parameters $\ln Z_0 + \left[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)\right]t$ and $\sigma_X^2 + \sigma_Y^2$.*

Proof. Let $U_t = f(t, Z_t) = \ln Z_t$. Note that $\mu(t, Z_t) = Z_t(\mu_X - \mu_Y + \sigma_Y^2)$, $\sigma(t, W_{Z,t}) = Z_t\sqrt{\sigma_X^2 + \sigma_Y^2}$, $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial Z_t t} = \frac{1}{Z_t}$, and $\frac{\partial^2 f}{\partial Z_t^2} = -\frac{1}{Z_t^2}$. By Theorem 2.9, we obtain

$$\begin{aligned} dU_t &= \left(\frac{1}{Z_t}(Z_t(\mu_X - \mu_Y + \sigma_Y^2)) + \frac{1}{2}\left(-\frac{1}{Z_t^2}\right)Z_t^2(\sigma_X^2 + \sigma_Y^2)\right)dt + \frac{1}{Z_t}\left(Z_t\sqrt{\sigma_X^2 + \sigma_Y^2}\right)dW_{Z,t} \\ &= \left(\mu_X - \mu_Y + \sigma_Y^2 - \frac{1}{2}(\sigma_X^2 + \sigma_Y^2)\right)dt + \sqrt{\sigma_X^2 + \sigma_Y^2}dW_{Z,t} \\ &= (\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2))dt + \sqrt{\sigma_X^2 + \sigma_Y^2}dW_{Z,t}. \end{aligned}$$

Integrating the above expression gives

$$\begin{aligned} \int_0^t dU_s &= \int_0^t [\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)]ds + \int_0^t \sqrt{\sigma_X^2 + \sigma_Y^2}dW_{Z,s} \\ \ln Z_t - lnZ_0 &= [\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)]t + \sqrt{\sigma_X^2 + \sigma_Y^2}dW_{Z,t} \\ \ln Z_t &= lnZ_0 + [\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)]t + \sqrt{\sigma_X^2 + \sigma_Y^2}dW_{Z,t}. \end{aligned}$$

Observe that U_t is normal distribution. Then, by the definition 2.6, Z_t should be lognormal distribution with parameters $\ln Z_0 + \left[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)\right]t$ and $\sigma_X^2 + \sigma_Y^2$. $\square$

The expectation and variance of Z_t can be expressed by using Theorem 4.2 and Corollary 2.7

Corollary 4.3. *Let Z_t be denoted by Lemma 2.7. Then, the expectation and variance of Z_t are*

$$E(Z_t) = Z_0 e^{(\mu_X - \mu_Y + \sigma_Y^2)t}, \quad \text{and} \quad var(Z_t) = Z_0^2 e^{2(\mu_X - \mu_Y + \sigma_Y^2)t}(e^{(\sigma_X^2 + \sigma_Y^2)t} - 1).$$

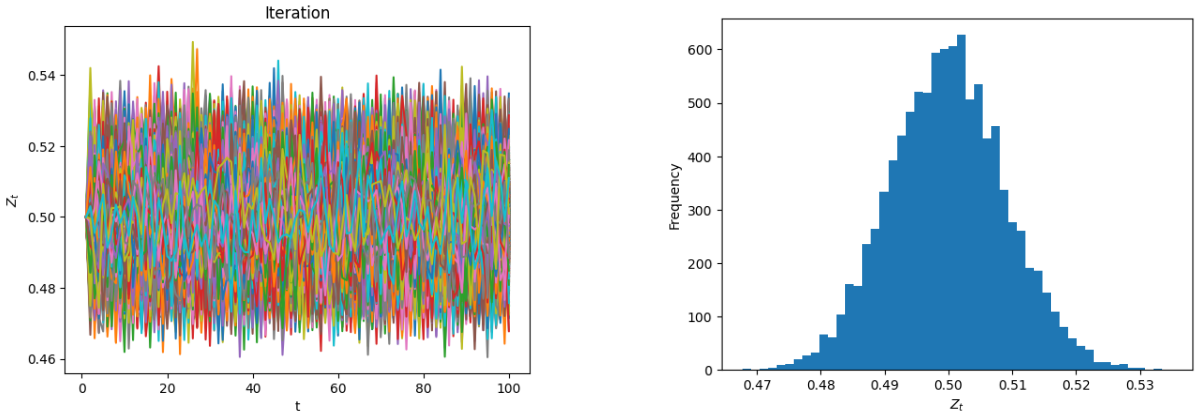
Proof. By Theorem 4.2 and Corollary 2.7. $\square$

Example 4.4 shows the application of Corollary 4.3 using the concept of Monte Carlo simulation.

Example 4.4. Consider the situation that

$$dX_t = X_t(0.05dt + 0.10dW_{X,t}), \qquad dY_t = Y_t(0.06dt + 0.15dW_{Y,t})$$

where the initial value are $X_0 = 10$, and $Y_0 = 20$. We then apply Monte Carlo simulation with 10,000 trials to this set up with $t = 1$. These are the results of simulation of $Z_1 = \frac{X_1}{Y_1}$. Figure 2a presents multiple realizations of the process Z_1 over 100 time steps. Each colored line corresponds to a different simulation (or iteration). Figure 2b displays the histogram of the simulated values of the process across multiple iterations. The distribution is approximately normal, centered around 0.5.



(a) Simulation of Z_1 (b) Histogram

Figure 2: Monte Carlo Simulation of Z_1 with 10,000 trials

The expectation and variance of the simulation are 0.500038 and 8.1×10^{-5}, which satisfy Corollary 4.3 such that

$$E(Z_1) = \frac{1}{2}e^{0.05-0.06+0.15^2} = 0.506289$$

$$\text{and} \quad var(Z_1) = Z_0^2 e^{2(0.05-0.06+0.15^2)}(e^{(0.10^2+0.15^2)t} - 1) = 8.1 \times 10^{-5}.$$

4.2 Token Volume within Liquidity Pool

Let us consider a scenario involving token trading across two markets: an external market (DEX-E) and a reference market (DEX-A). We will observe the change in the quantities of tokens A and B within DEX-A, assuming that the exchange rate of token B in terms of token A in DEX-A adjusts to follow the rate in DEX-E with a lag of one time unit. Moreover, we can expect DEX-A to earn fees from arbitrageurs. Let the quantities of tokens A and B within DEX-A at any time t be defined by the following lemma.

Lemma 4.5. *Let X_0, and Y_0 denote the initial exchange rates, denominated in a base currency or another token, of token A and token B in an external market. Furthermore, let the initial reserve quantity of token A and token B in the reference market be denoted by V_{A_0}, and V_{B_0}, respectively, which satisfy the constant product market maker. Then, under the arbitrage opportunity,*

- *the reserve of token A and token B in the reference market at any time t is denoted by*

$$V_{A_t} = \sqrt{\frac{kX_t}{Y_t}} = \sqrt{kZ_t}, \quad \textit{and} \quad V_{B_t} = \sqrt{\frac{kY_t}{X_t}} = \sqrt{\frac{k}{Z_t}}.$$

- *changes in the reserve of token A and token B can be written as*

$$\Delta V_{A_{t-1}} = \sqrt{kZ_t} - \sqrt{kZ_{t-1}}, \quad \textit{and} \quad \Delta V_{B_{t-1}} = \sqrt{\frac{k}{Z_t}} - \sqrt{\frac{k}{Z_{t-1}}}.$$

Proof. Due to the concept of a constant product market maker, it can be rewritten as

$$V_{B_t} = \frac{k}{V_{A_t}}. \tag{4.2}$$

Consider, without loss of generality,

$$\begin{aligned}
\frac{X_t}{Y_t} &= \frac{V_{A_t}}{V_{B_t}} \\
\frac{X_t}{Y_t} &= \frac{V_{A_t}^2}{k} \\
V_{A_t}^2 &= k\frac{X_t}{Y_t} \\
V_{A_t} &= \pm\sqrt{\frac{kX_t}{Y_t}} = \pm\sqrt{kZ_t}.
\end{aligned}$$

Subsequently, we shall consider the positive and negative parts separately. If $V_{A_t} = \sqrt{kZ_t}$, then substituting V_{A_t} in the equation (4.2) gives

$$V_{B_t} = \frac{k}{\sqrt{kZ_t}} = \sqrt{\frac{k}{Z_t}}.$$

Similarly, if $V_{A_t} = -\sqrt{kZ_t}$, then we obtain that

$$V_{B_t} = \frac{k}{-\sqrt{kZ_t}} = -\sqrt{\frac{k}{Z_t}}. \tag{4.3}$$

Observe the negative root in an equation (4.3) would imply negative token reserves which is impossible. Thus, only the positive terms is mathematically meaningful. This ensures the liquidity pool maintains a viable state with positive reserves after arbitrageurs align prices.

Since $V_{A_t} = V_{A_{t-1}} + \Delta V_{A_{t-1}}$ and $V_{B_t} = V_{B_{t-1}} + \Delta V_{B_{t-1}}$, the change in reserves at any time $t = 1, 2, 3, \ldots$ can be written as

$$\Delta V_{A_{t-1}} = \sqrt{kZ_t} - \sqrt{kZ_{t-1}}, \quad \text{and} \quad \Delta V_{B_{t-1}} = \sqrt{\frac{k}{Z_t}} - \sqrt{\frac{k}{Z_{t-1}}}. \qquad \square$$

We can prove that V_{A_t} follow *lognormal* distribution which will be derived in the next theorem.

Theorem 4.6. *V_{A_t} follows a lognormal distribution with parameters $\mu_{V_{A_t}} = \ln V_0 + \frac{1}{2}\left[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)\right] t$, and $\sigma_{V_{A_t}} = \frac{1}{4}\sqrt{\sigma_X^2 + \sigma_Y^2}$ if $V_{A_t} = \sqrt{kZ_t}$ and Z_t is given by Theorem 4.2.*

Proof. Since $\ln Z_t \sim N\left(\ln Z_0 + \left[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)\right] t, \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$,

$$\begin{aligned}\ln V_t \sim & \ln\sqrt{kZ_t} \sim \frac{1}{2}\ln k + \frac{1}{2}\ln Z_t \\ \sim & N\left(\frac{1}{2}\ln k + \frac{1}{2}\ln Z_0 + \frac{1}{2}\left[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)\right] t, \frac{1}{4}\sqrt{\sigma_X^2 + \sigma_Y^2}\right) \\ \sim & N\left(\ln V_0 + \frac{1}{2}\left[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)\right] t, \frac{1}{4}\sqrt{\sigma_X^2 + \sigma_Y^2}\right).\end{aligned}$$

Hence, $V_t \sim LN\left(\ln V_0 + \frac{1}{2}\left[\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)\right] t, \frac{1}{4}\sqrt{\sigma_X^2 + \sigma_Y^2}\right)$. □

Additionally, the expectation and variance of V_{A_t} is mentioned in the next corollary.

Corollary 4.7. *Let V_{A_t} follow Lemma 4.5. Then $E(V_{A_t}) = V_0 e^{\frac{1}{2}(\mu_X - \mu_Y)t + \frac{1}{8}(5\sigma_Y^2 - 3\sigma_X^2)t}$, and $var(V_{A_t}) = V_0^2 e^{(\mu_X - \mu_Y + \frac{1}{4}(3\sigma_Y^2 - \sigma_X^2))t}(e^{\frac{1}{4}(\sigma_X^2 + \sigma_Y^2)t} - 1)$.*

Proof. By Corollary 2.7,

$$\begin{aligned}E(\sqrt{Z_t}) =& e^{\frac{1}{2}\left[\ln Z_0 + [\mu_X - \mu_Y + \frac{1}{2}(\sigma_Y^2 - \sigma_X^2)]t\right] + \frac{1}{8}(\sigma_X^2 + \sigma_Y^2)t} = \sqrt{Z_0} e^{\frac{1}{2}(\mu_X - \mu_Y)t + \frac{1}{8}(5\sigma_Y^2 - 3\sigma_X^2)t}, \\ \text{and } var(\sqrt{Z_t}) =& Z_0 e^{(\mu_X - \mu_Y + \frac{1}{4}(3\sigma_Y^2 - \sigma_X^2))t}(e^{\frac{1}{4}(\sigma_X^2 + \sigma_Y^2)t} - 1).\end{aligned}$$

Thus, $E(V_{A_t}) = \sqrt{kZ_0} e^{\frac{1}{2}(\mu_X - \mu_Y)t + \frac{1}{8}(5\sigma_Y^2 - 3\sigma_X^2)t} = V_0 e^{\frac{1}{2}(\mu_X - \mu_Y)t + \frac{1}{8}(5\sigma_Y^2 - 3\sigma_X^2)t}$, and $var(V_{A_t}) = kZ_0 e^{(\mu_X - \mu_Y + \frac{1}{4}(3\sigma_Y^2 - \sigma_X^2))t}(e^{\frac{1}{4}(\sigma_X^2 + \sigma_Y^2)t} - 1) = V_0^2 e^{(\mu_X - \mu_Y + \frac{1}{4}(3\sigma_Y^2 - \sigma_X^2))t}(e^{\frac{1}{4}(\sigma_X^2 + \sigma_Y^2)t} - 1)$. □

4.3 Portfolio Value of Liquidity Provider

In this section, we analyze the actions of a liquidity provider and the evolution of their portfolio value under yield farming mechanics.

Theorem 4.8. *Let P_t denote the liquidity provider's portfolio value at time t. Assume that the LP initiates their investment with token A given by P_0. If Z_t is given by Lemma 4.1, then, the liquidity provider's portfolio value at time t under Assumption 2.11 and yield farming is given by*

$$P_t = \frac{P_0}{2^{t-1}} \prod_{i=2}^{t} \left(1 + \frac{Z_{i-1}}{Z_{i-2}}\right).$$

Proof. Let us begin by considering the interval $t \in [0, 1)$. The LP initiates their investment with token A given by P_0.

Subsequently, consider the interval $t \in [1, 2)$. At this stage, the LP allocates token A and token B of equal value into the liquidity pool, denoted as S_t. The tokens are divided into two parts, considering the fees associated with swapping and staking, $S_1 = \frac{P_0}{2}$.

Thus, the amount of tokens A and B is S_1 tokens and $\frac{S_1}{Z_0}$ tokens, respectively. The LP immediately stakes the pair of token A and B in the A-B token pool. Hence, the portfolio value after allocating the tokens in the pool is

$$P_1 = 2S_1.$$

In the case that $t \in [2, 3)$, and due to the yield farming condition that the pair of tokens provided in the liquidity pool must be of equal value at every t, and given that the LP allocates their asset with S_1 A tokens and $\frac{S_1}{Z_0}$ B tokens at the previous period, and the exchange rate at time $t = 2$ is Z_1, the amount of token A is

$$S_2 = \frac{1}{2}\left(S_1 + \frac{Z_1}{Z_0}S_1\right) = \frac{1}{2}\left(1 + \frac{Z_1}{Z_0}\right)S_1 = \frac{Z_0 + Z_1}{2Z_0}S_1.$$

Hence, the portfolio value after equalizing the pair of tokens is

$$P_2 = 2S_2 = \frac{Z_0 + Z_1}{Z_0}S_1.$$

In the case that $t \in [3, 4)$, the amount of token A in this period changes to

$$S_3 = \frac{1}{2}\left(S_2 + \frac{Z_2}{Z_1}S_2\right) = \frac{Z_1 + Z_2}{2Z_1}S_2 = \frac{Z_1 + Z_2}{2Z_1}\left(\frac{Z_0 + Z_1}{2Z_0}S_1\right) = \frac{(Z_0 + Z_1)(Z_1 + Z_2)}{2^2 Z_0 Z_1}S_1.$$

Thus, the portfolio value at this period is given by

$$P_3 = 2S_3 = \frac{(Z_0 + Z_1)(Z_1 + Z_2)}{2Z_0 Z_1}S_1.$$

For simplicity, during the period when the LP's tokens are in the pool, $t \in [2, T)$, the amount of token A at any time t can be written as

$$S_t = \frac{S_1}{2^{t-1}}\prod_{i=2}^{t}\left(\frac{Z_{i-2} + Z_{i-1}}{Z_{i-2}}\right) = \frac{S_1}{2^{t-1}}\prod_{i=2}^{t}\left(1 + \frac{Z_{i-1}}{Z_{i-2}}\right)$$

and the portfolio value at any time t is given by

$$P_t = \frac{S_1}{2^{t-2}}\prod_{i=2}^{t}\left(1 + \frac{Z_{i-1}}{Z_{i-2}}\right) = \frac{P_0}{2^{t-1}}\prod_{i=2}^{t}\left(1 + \frac{Z_{i-1}}{Z_{i-2}}\right). \qquad \square$$

Finally, the LP withdraws their tokens from the pool at time T. Then, the expectation and variance of final LP's portfolio value is derive in the next corollary.

Corollary 4.9. *The expectation and variance of P_t given by Theorem 4.8 are*

$$E(P_t) = \frac{P_0}{2^{t-1}}\left(1 + e^{(\mu_X - \mu_Y + \sigma_Y^2)}\right)^{t-1},$$

$$\text{and } var(P_t) = \frac{P_0^2}{4^{t-1}}\left[\left(1 + 2e^{\mu_X - \mu_Y + \sigma_Y^2} + e^{2(\mu_X - \mu_Y + \sigma_X^2 + 3\sigma_Y^2)}\right)^{t-1} - \left(1 + e^{\mu_X - \mu_Y + \sigma_Y^2}\right)^{2(t-1)}\right].$$

Proof. Let us consider

$$\begin{aligned}
E(P_t) =& E\left[\frac{P_0}{2^{t-1}}\prod_{i=2}^{t}\left(1+\frac{Z_{i-1}}{Z_{i-2}}\right)\right]\\
=& \frac{P_0}{2^{t-1}}\prod_{i=2}^{t}\left(1+\frac{E(Z_{i-1})}{E(Z_{i-2})}\right)\\
=& \frac{P_0}{2^{t-1}}\prod_{i=2}^{t}\left(1+e^{(\mu_X-\mu_Y+\sigma_Y^2)}\right)\\
=& \frac{P_0}{2^{t-1}}\left(1+e^{(\mu_X-\mu_Y+\sigma_Y^2)}\right)^{t-1}
\end{aligned}$$

$$\begin{aligned}
\text{and } var(P_t) =& var\left[\frac{P_0}{2^{t-1}}\prod_{i=2}^{t}\left(1+\frac{Z_{i-1}}{Z_{i-2}}\right)\right]\\
=& \frac{P_0^2}{4^{t-1}} var\left[\prod_{i=2}^{t}\left(1+\frac{Z_{i-1}}{Z_{i-2}}\right)\right]\\
=& \frac{P_0^2}{4^{t-1}}\left(\prod_{i=2}^{t}E\left[\left(1+\frac{Z_{i-1}}{Z_{i-2}}\right)^2\right]-\left[\prod_{i=2}^{t}E\left(1+\frac{Z_{i-1}}{Z_{i-2}}\right)\right]^2\right)\\
=& \frac{P_0^2}{4^{t-1}}\left(\prod_{i=2}^{t}\left[1+2E\left(\frac{Z_{i-1}}{Z_{i-2}}\right)+E\left(\frac{Z_{i-1}^2}{Z_{i-2}^2}\right)\right]-\left[\prod_{i=2}^{t}\left(1+E\left(\frac{Z_{i-1}}{Z_{i-2}}\right)\right)\right]^2\right)\\
=& \frac{P_0^2}{4^{t-1}}\left(\prod_{i=2}^{t}\left(1+2e^{\mu_X-\mu_Y+\sigma_Y^2}+e^{2(\mu_X-\mu_Y+\sigma_X^2+3\sigma_Y^2)}\right)-\left[\prod_{i=2}^{t}\left(1+e^{\mu_X-\mu_Y+\sigma_Y^2}\right)\right]^2\right)\\
=& \frac{P_0^2}{4^{t-1}}\left[\left(1+2e^{\mu_X-\mu_Y+\sigma_Y^2}+e^{2(\mu_X-\mu_Y+\sigma_X^2+3\sigma_Y^2)}\right)^{t-1}-\left(1+e^{\mu_X-\mu_Y+\sigma_Y^2}\right)^{2(t-1)}\right].
\end{aligned}$$

□

4.4 Portfolio Value of Liquidity Provider Including Fee

According to Example 2.14, LPs incur fees five times: an initial processing and gas fee ($\phi_p P_0+g$), a transaction fee when exchanging fiat for token A (ϕP_0), a transaction fee when exchanging token A for token B ($\phi\frac{P_1}{2}$), a transaction fee when staking the token pair into the pool (ϕP_1), and a transaction fee when withdrawing the token pair from the pool (ϕP_{T-1}). Then, the fee, the LP pay, is denoted by

$$f = (\phi_p P_0 + g) + \phi P_0 + \phi\frac{P_1}{2} + \phi P_1 + \phi P_{T-1} = \phi_p P_0 + g + \phi(P_0 + \frac{3P_1}{2} + P_{T-1}).$$

Hence, the final LP's portfolio value is given by

$$P_T = P_{T-1} - f = \frac{P_0}{2^{T-2}}\left[\prod_{i=2}^{T-1}\left(1+\frac{Z_{i-1}}{Z_{i-2}}\right)\right] - f.$$

4.5 Liquidity Provider Returns

The theoretical return of LPs, taking into account their portfolio value including fees, will be discussed in this section. The result is derived in the following theorem.

Theorem 4.10. *Let R_T be a return of liquidity provider from $t = 0$ to $t = T$. Then, the return of this investment is given by*

$$R_T = \frac{1}{2^{T-2}} \left[\prod_{i=2}^{T-1} \left(1 + \frac{Z_{i-1}}{Z_{i-2}} \right) \right] - 1 - \frac{f}{P_0}.$$

Proof. Let R be a return of liquidity provider from $t = 0$ to $t = T$. Then, the return of this investment is given by

$$\begin{aligned} R_T =& \frac{P_T - P_0}{P_0} \\ =& \frac{\frac{P_0}{2^{T-2}} \left[\prod_{i=2}^{T-1} \left(1 + \frac{Z_{i-1}}{Z_{i-2}} \right) \right] - P_0 - f}{P_0} \\ \therefore R_T =& \frac{1}{2^{T-2}} \left[\prod_{i=2}^{T-1} \left(1 + \frac{Z_{i-1}}{Z_{i-2}} \right) \right] - 1 - \frac{f}{P_0}. \end{aligned}$$

□

Additionally, the expected return and its variance is shown in the later corollary.

Corollary 4.11. *Let R_T be denoted by Theorem 4.10. Then, the expectation and variance of R_T is defined by*

$$E(R_T) = \left(\frac{1 + e^{\mu_X - \mu_Y + \sigma_Y^2}}{2} \right)^{T-2} - \frac{f}{P_0} - 1,$$

and $var(R_T) = \frac{1}{4^{T-2}} \left(\left(1 + 2e^{\mu_X - \mu_Y + \sigma_Y^2} + e^{2(\mu_X - \mu_Y + \sigma_X^2 + 3\sigma_Y^2)} \right)^{T-2} - \left(1 + e^{\mu_X - \mu_Y + \sigma_Y^2} \right)^{2(T-2)} \right).$

Proof. Consider,

$$\begin{aligned} E(R_T) =& E\left(\frac{P_T - P_0}{P_0} \right) \\ =& \frac{E(P_T)}{P_0} - 1 \\ =& \frac{1}{P_0} E\left(\frac{P_0}{2^{T-2}} \left[\prod_{i=2}^{T-1} \left(1 + \frac{Z_{i-1}}{Z_{i-2}} \right) \right] - f \right) - 1 \\ =& \frac{1}{P_0} \left(\frac{P_0}{2^{T-2}} \left[\prod_{i=2}^{T-1} E\left(1 + \frac{Z_{i-1}}{Z_{i-2}} \right) \right] - f \right) - 1 \\ =& \frac{1}{P_0} \left(\frac{P_0}{2^{T-2}} \left[\prod_{i=2}^{T-1} \left(1 + e^{\mu_X - \mu_Y + \sigma_Y^2} \right) \right] - f \right) - 1 \\ \therefore E(R_T) =& \left(\frac{1 + e^{\mu_X - \mu_Y + \sigma_Y^2}}{2} \right)^{T-2} - \frac{f}{P_0} - 1, \end{aligned}$$

and

$$\begin{aligned}
var(R_T) =& var\left(\frac{P_T - P_0}{P_0}\right) \\
=& \frac{var(P_T)}{P_0^2} \\
=& \frac{1}{P_0^2} var\left(\frac{P_0}{2^{T-2}}\left[\prod_{i=2}^{T-1}\left(1+\frac{Z_{i-1}}{Z_{i-2}}\right)\right] - f\right) \\
=& \frac{1}{P_0^2}\left(\frac{P_0^2}{4^{T-2}} var\left[\prod_{i=2}^{T-1}\left(1+\frac{Z_{i-1}}{Z_{i-2}}\right)\right]\right) \\
=& \frac{1}{P_0^2}\left(\frac{P_0^2}{4^{T-2}}\left[\left(1+2e^{\mu_X-\mu_Y+\sigma_Y^2}+e^{2(\mu_X-\mu_Y+\sigma_X^2+3\sigma_Y^2)}\right)^{T-2} - \left(1+e^{\mu_X-\mu_Y+\sigma_Y^2}\right)^{2(T-2)}\right]\right) \\
=& \frac{1}{4^{T-2}}\left(\left(1+2e^{\mu_X-\mu_Y+\sigma_Y^2}+e^{2(\mu_X-\mu_Y+\sigma_X^2+3\sigma_Y^2)}\right)^{T-2} - \left(1+e^{\mu_X-\mu_Y+\sigma_Y^2}\right)^{2(T-2)}\right). \qquad \square
\end{aligned}$$

5 Conclusion

This study examines the returns and risks associated with liquidity provision in Automated Market Makers (AMMs) under the assumption that asset prices follow independent Black-Scholes dynamics. By analyzing liquidity provider (LP) returns in a constant product market makers (CPMMs), we quantify the impact of key parameters such as volatility and drift on LP profitability.

Through stochastic modeling, we demonstrate that fee structures and arbitrage dynamics significantly affect LP outcomes. These insights provide a framework for managing financial risks in DeFi yield farming strategies. Future research could extend this work by incorporating correlated asset dynamics, alternative AMM mechanisms, or real-world transaction data for empirical validation.

Acknowledgment. The authors The authors sincerely thank to the referees for their careful reading of the manuscript and their insightful comments. We also acknowledge the support from department of Mathematics, faculty of Science, Chiang Mai University. Their contributions have greatly improved this study.

References

[1] S. Aggarwal, N. Kuma, Chapter Twelve - Cryptocurrencies, in: The Blockchain Technology for Secure and Smart Applications across Industry Verticals, Advances in Computers, Elsevier, 121 (2021) 227–266.

[2] S. Nakamoto. Bitcoin: a peer-to-peer electronic cash system. <`http://bitcoin.org/bitcoin.pdf`>, (accessed 15.04.25).

[3] F. Schär. Decentralized Finance: On Blockchain- and Smart Contract-Based Financial Markets. Federal Reserve Bank of St. Louis Review, Second Quarter 2021, 103(2), pp. 153–74.

[4] C. Harvey, J. Hasbrouck, F. Saleh. The Evolution of Decentralized Exchange: Risks, Benefits, and Oversight, Wharton Initiative on Financial Policy and Regulation White Paper, The University of Pennsylvania, 2024 https://wifpr.wharton.upenn.edu/wp-content/uploads/2024/07/WIFPR-Decentralized-Exchange-Harvey-Hasbrouck-and-Saleh.pdf.

[5] Uniswap. <https://app.uniswap.org/>, (accessed 15.04.25).

[6] W. Kenton, What Is an Order Book? Definition, How It Works, and Key Parts. <https://www.investopedia.com/terms/o/order-book.asp>, (accessed 15.04.25).

[7] A. Hayden., N. Zinsmeister., D. Robinson, Uniswap v2 Core. <https://uniswap.org/whitepaper.pdf>, (accessed 17.01.25).

[8] H. Matthias, D. Helmut. Impermanent Loss Conditions: An Analysis of Decentralized Exchange Platforms. (2024). 10.48550/arXiv.2401.07689.

[9] E. Michael, Automatic market-making with dynamic peg. <https://classic.curve.fi/files/crypto-pools-paper.pdf>, (accessed 12.04.25).

[10] Ch. Bhardwaj, How to build a DeFi yield farming app?. <https://appinventiv.com/blog/how-to-build-defi-yield-farming-application/>, (accessed 12.04.25).

[11] J. Clark, The replicating portfolio of a constant product market. <http://dx.doi.org/10.2139/ssrn.3550601>, (accessed 12.04.25).

[12] J. Najnudel., Sh.-N. Tung., K. Yamazaki., J.-Y. Yen. An arbitrage driven price dynamics of Automated Market Makers in the presence of fees. Frontiers of Mathematical Finance. 3. 560-571. 10.3934/fmf.2024018.

[13] G. Grimmett., D. Stirzaker, Probability and Random Process. the United States. Oxford University Press Inc., New York, 2001.

[14] B. Okesendal, Stochastic differential equations : an introduction with applications. Springer Berlin, Heidelberg, 1945.

[15] J.F.L. Gall, Brownian Motion, Martingales, and Stochastic Calculus, Graduate Texts in Mathematics, Springer International Publishing, Switzerland, 2016.

[16] Uniswap, Understanding returns | UNisWaP. (n.d.). <https://docs.uniswap.org/contracts/v2/concepts/advanced-topics/understanding-returns.>, (accessed 22.01.25).

SL-vine Copula Model

Tanabodee Monggon[1,*], Songkiat Sumetkijakan[1,†],
Tippawan Santiwipanont[1] and Monchai Kooakachai[1]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

As available multivariate copulas are too rigid to accommodate high-dimensional data and diverse dependency patterns, vine copulas, characterized by a sequence of connected trees, offer the flexibility to model dependence among many variables by an iterative structure of bivariate copulas, enabling the construction of tailored dependency models for complex datasets. In terms of unique vine structures, D-vines work well for sequential relationships, whereas C-vines encapsulate hierarchical dependencies around an essential variable. A mix of both formats is often seen in real-world data. To bridge this gap, we examine the hybrid vine copula known as the stem-and-leaf (SL) vine, which combines the key features of C and D vines. We define the SL-vine structure formally, demonstrate its compliance with regular vine (R-vine) tree sequence requirements, and establish its joint density function. We then illustrate the SL-vine's performance using the abalone and capital-market datasets, focusing on modeling relationships between weight and size measurements.

Keywords: Multivariate dependence, Vine copula, SL-vine
2020 MSC: 62H05, 62H20, 62H22

1 Introduction

Multivariate dependence modeling is a cornerstone of statistical and applied research, enabling the understanding of relationships among multiple variables. While classical approaches, such as multivariate Gaussian models, often fall short due to their rigid assumptions, copula-based methods have emerged as powerful tools for capturing complex dependencies. Vine copulas, introduced by Joe [1] and further developed by Bedford and Cooke [2], offer unparalleled flexibility by decomposing high-dimensional copulas into a cascade of pair-copulas. This decomposition allows for tailored modeling of dependencies, enhancing both interpretability and accuracy.

Among the existing vine structures, C-vines excel in capturing hierarchical dependencies centered around a dominant variable, while D-vines are adept at modeling sequential dependencies. However, many real-world datasets exhibit a combination of these structures, with some variables showing hierarchical dependence and others reflecting sequential relationships. To address

*Speaker (6670187523@student.chula.ac.th)
†Corresponding author (Songkiat.S@chula.ac.th)

this, we propose a stem-and-leaf (SL) vine structure—a type of vine copula that integrates the strengths of both C- and D-vines.

This paper first reviews fundamental concepts in copulas and vine copulas, as detailed in Section 2. Section 3 then introduces the SL-vine model, beginning with its formal definition and demonstrating its adherence to the three core requirements of regular vine (R-vine) tree sequences. In addition, we present the distribution of the SL-vine copula. Section 4 focuses on the application of the SL-vine copula model. It begins with a structured model selection process followed by an in-depth exploration of the results, using the abalone and capital-market datasets as a case study. Finally, Section 5 provides a comprehensive summary and areas of future research.

2 Theoretical background

A function $C\colon [0,1]^d \to [0,1]$ is called a d-copula, if it represents the joint distribution of a random vector $\mathbf{U} = (U_1, \ldots, U_d)$ where each component U_k follows a standard uniform distribution, i.e. $\mathbb{P}(U_k \leq u_k) = u_k$, where $u_k \in [0,1]$ for each $k \in \{1, \ldots, d\}$. The main benefit of copulas, which has led to their popularity in many fields, lies in their ability to connect the distributions of d random variables with their joint distribution. This property is a consequence of Sklar's theorem [3], which provides a better understanding and flexibility in modeling the relationship among random variables in the following way. Let $X_1, \ldots, X_d$ be continuous random variables with joint distribution function F, and respective marginal distribution functions $F_1, \ldots, F_d$. Then there exists a unique copula C such that

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \quad \text{for all } x_1, \ldots, x_d \in \mathbb{R}. \tag{2.1}$$

Such a C is called the copula of $X_1, \ldots, X_d$. If $X_1, \ldots, X_d$ are continuous random variables with joint density function f, and marginal density functions $f_1, \ldots, f_d$ respectively, then

$$f(x_1, \ldots, x_d) = c(F_1(x_1), \ldots, F_d(x_d)) \prod_{i=1}^{d} f_i(x_i) \quad \text{for all } x_1, \ldots, x_d \in \mathbb{R}, \tag{2.2}$$

where $c := \partial_1 \ldots \partial_d C$ is the density of C.

However, equations (2.1) and (2.2) involve potentially complex multivariate copula distributions and densities. The density of a d-dimensional copula can be expressed as a decomposition into $d(d-1)/2$ bivariate copula densities, as Joe [1] showed. Since this decomposition is not unique, it provides flexibility in choosing among various possible representations. Bedford and Cooke [2] introduced a valuable graphical tool called a *regular vine* (R-vine) tree structure to identify and organize all valid factorizations. An R-vine of d random variables consists of $d-1$ connected trees, denoted by $\mathcal{V} = (T_1, ..., T_{d-1})$, where the edges of T_k become the nodes of T_{k+1}. In this structure, each node in a tree represents a specific random variable from the dataset, while each edge between two nodes represents a bivariate copula. The set of bivariate copula families, referred to as pair copulas, is denoted by $\mathcal{B}(\mathcal{V})$, while their corresponding parameters set is represented as $\Theta(\mathcal{B}(\mathcal{V}))$. For a rigorous definition, a vine tree sequence or tree structure $\mathcal{V} = (T_1, ..., T_{d-1})$ with $T_k = (N_k, E_k)$ for $1 \leq k \leq d-1$, must satisfy the definition below.

Definition 2.1. The collection of trees $\mathcal{V} = (T_1, ..., T_{d-1})$ is called an R-vine tree sequence on d elements if

(i) T_1 is a connected tree with the node set $N_1 = \{1, \ldots, d\}$ and an edge set E_1.

(ii) For $k \geq 2$, T_k is a connected tree with the node set $N_k = E_{k-1}$ and an edge set E_k.

(iii) For $k \geq 2$, if two nodes of the tree T_k are connected by an edge, then the two nodes, which are edges of T_{k-1}, have a common node (proximity condition).

A vine tree sequence, which is composed entirely of paths, is known as a *drawable vine* (D-vine), while a tree sequence made entirely of stars is referred to as a *canonical vine* (C-vine), as introduced by Bedford and Cooke [2]. To illustrate, we consider pair copula decomposition of an arbitrary three dimensional density $f(x_1, x_2, x_3)$ as

$$\begin{aligned} f(x_1, x_2, x_3) = c_{1,3|2}\big(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2) \mid x_2\big) \times c_{2,3}\big(F_2(x_2), F_3(x_3)\big) \\ \times c_{1,2}\big(F_1(x_1), F_2(x_2)\big) f_3(x_3) f_2(x_2) f_1(x_1). \end{aligned} \tag{2.3}$$

For any index set $D \subset \{1, \ldots, d\}$ and an index $j \notin D$, $F_{j|D}$ and $f_{j|D}$ denote the conditional distribution or density function of X_j given $\mathbf{X}_D$, respectively, where $\mathbf{X}_D$ denotes the random vector $(X_i)_{i \in D}$ whose observed value $(x_i)_{i \in D}$ is denoted by $\mathbf{x}_D$. The copula density $c_{1,3|2}(\cdot, \cdot \mid x_2)$ denotes the copula density associated with the conditional distribution of (X_1, X_3) given $X_2 = x_2$. Equation (2.3) is represented as a D-vine tree structure in Figure 1.

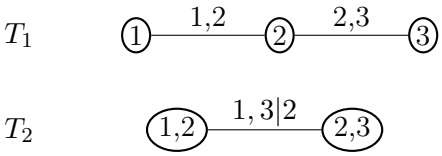


Figure 1: D-vine tree structure illustrating the construction of the three pair copulas $c_{23} - c_{12} - c_{13|2}$

By Definition 2.1, D-vine tree structure corresponding to equation (2.3) consists of nodes $N_1 = \{1, 2, 3\}$ and edges $E_1 = \{\{1, 2\}, \{2, 3\}\}$ in tree T_1. In tree T_2, the nodes are $N_2 = \{\{1, 2\}, \{2, 3\}\}$ and the edges are $E_2 = \{\{\{1, 2\}, \{2, 3\}\}\}$ in tree T_2. However, the proximity condition in Definition 2.1 makes the element in E_2 more complex. To improve clarity, we define the complete union for each edge $e = \{a, b\} \in E_k$ as follows:

$$\mathcal{U}_e = \begin{cases} e_1 & \text{if } k = 1, \\ \{i \in N_1 \mid \exists (e_1, \ldots, e_{k-1}) \in E_1 \times \cdots \times E_{k-1} \text{ such that } i \in e_1 \in \cdots \in e\} & \text{if } k \geq 2. \end{cases}$$

The set $D_e := \mathcal{U}_a \cap \mathcal{U}_b$ is called the conditioning set of e, and the conditioned sets, $\mathcal{C}_{e,a}$ and $\mathcal{C}_{e,b}$, are defined as

$$\mathcal{C}_{e,a} = \mathcal{U}_a \setminus D_e := a_e \text{ and } \mathcal{C}_{e,b} = \mathcal{U}_b \setminus D_e := b_e.$$

Moreover, the set $\mathcal{C}_{e,a}$ and $\mathcal{C}_{e,b}$ were shown to be singletons by Kurowicka and Cooke [4]. Therefore, the notation for an edge $e = (\mathcal{C}_{e,a}, \mathcal{C}_{e,b} \mid D_e)$ is commonly simplified to $e = (a_e, b_e \mid D_e)$. For instance, consider the edge $e = \{a = \{1, 2\}, b = \{2, 3\}\}$, where $\mathcal{U}_a = \{1, 2\}$ and $\mathcal{U}_b = \{2, 3\}$, leading to $D_e = \{2\}$. Consequently, we have $\mathcal{C}_{e,a} = \{1\}$ and $\mathcal{C}_{e,b} = \{3\}$, which simplifies the notation to $e = (1, 3 \mid 2)$. In summary, For $k \geq 2$, each edge in the sequence that corresponds to bivariate copula family, is denoted by $c_{a_e, b_e|D_e} \in \mathcal{B}(\mathcal{V})$, with its corresponding parameter set

$\theta_{a_e,b_e|D_e} \in \Theta(\mathcal{B}(\mathcal{V}))$. The copula density $c_{a_e,b_e|D_e}$ corresponds to the copula density associated with the conditional distribution of (X_{a_e}, X_{b_e}) given $\mathbf{X}_{D_e}$. Equivalently, $c_{a_e,b_e|D_e}$ is the copula density of the random vector

$$(U_{a_e|D_e}, U_{b_e|D_e}) = \big((F_{a_e|D_e}(X_{a_e}|\mathbf{X}_{D_e}), F_{b_e|D_e}(X_{b_e}|\mathbf{X}_{D_e})) \mid \mathbf{X}_{D_e}\big).$$

The graphical structure of regular vines offers a natural framework for decomposing the joint density function, as shown by Bedford and Cooke [2]. Pair-copulas determined by the vine tree sequence are used to accomplish this decomposition in the following manner:

$$f(x_1,\ldots,x_d) = \left[\prod_{j=1}^{d-1}\prod_{e\in E_k} c_{a_e,b_e|D_e}\left(F_{a_e|D_e}(x_{a_e}|\mathbf{x}_{D_e}), F_{b_e|D_e}(x_{b_e}|\mathbf{x}_{D_e}) \mid \mathbf{x}_{D_e}\right)\right] \times \left[\prod_{k=1}^{d} f_k(x_k)\right]. \tag{2.4}$$

Any density function of regular vine copula can be decomposed using equation (2.4). Nevertheless, the individual pair copulas $c_{a_e,b_e|D_e}$ in equation (2.4) are based on the specific values of $\mathbf{x}_{D_e}$. This dependency reflects the varying conditional relationships between a_e and b_e for different conditioning values of $\mathbf{x}_{D_e}$.

For tractability, modeling multivariate dependence via vine copulas hinges on the "simplifying assumption," stating that all conditional copulas in the model do not depend on the conditioning random variables (more in Haff [5] and Stöber [6]). That is, a distribution F satisfies the simplifying assumption with respect to the vine structure $\mathcal{V}$, if equation (2.4) holds with

$$c_{a_e,b_e|D_e}(\cdot \mid \mathbf{x}_{D_e}) = c_{a_e,b_e|D_e}(\cdot) \text{ for all } \mathbf{x}_{D_e} \in \mathbb{R}^d \text{ and } e \in \mathcal{V}.$$

Consequently, we can simplify equation (2.4) to:

$$f(x_1,\ldots,x_d) = \left[\prod_{j=1}^{d-1}\prod_{e\in E_j} c_{a_e,b_e|D_e}\right] \times \left[\prod_{k=1}^{d} f_k(x_k)\right], \tag{2.5}$$

where $c_{a_e,b_e|D_e} := c_{a_e,b_e|D_e}\left(F_{a_e|D_e}(x_{a_e}|\mathbf{x}_{D_e}), F_{b_e|D_e}(x_{b_e}|\mathbf{x}_{D_e})\right)$. This approach is motivated by the need for tractability in higher dimensions and can be further validated through empirical testing (Kurz and Spanhel [7] and Derumigny and Fermanian [8]). Within this framework, the method is termed pair copula constructions for modeling multivariate densities.

3 SL-vine copula model

In this section, we propose a new vine tree structure based on the concept of regular vines, called the stem-and-leaf (SL) vine. Specifically, we will first define how a set of tree sequences on d elements is referred to as an SL-tree. We will then demonstrate that, with this definition, the SL-vine also satisfies the R-vine tree sequence. Next, the SL-vine copula distribution will be derived. We conclude the section by discussing the uniqueness of SL-vine and R-vine trees. Unlike the SL-vine, where the structure is determined after the first step, R-vine trees generally allow for multiple ways to form a connected tree in the second step. In these cases, we can select

the path based on certain criteria, such as using the strongest dependency first. To address this, a criterion such as Kendall's Tau will also be introduced.

Even though SL-vine is defined below only for even number of random variables, SL-vine for $2m-1$ random variables can be defined as $(T_2, T_3, T_4, \ldots, T_{2m-1})$ where $(T_1, T_2, T_3, \ldots, T_{2m-1})$ is the R-vine tree sequence in the definition.

Definition 3.1. In the kth tree of the sequence, the nodes forming a path are called *stems*, collected in S_k, while the remaining nodes attached to it are referred to as *leaves*, collected in L_k. For any $2m$ random variables, $m \in \mathbb{N}$, we define the $2m-1$ trees of the SL-vine tree sequence as follows:

T_1 with the node set $N_1 = S_1 \cup L_1 = \{1, 2, \ldots, m\} \cup \{2m, \ldots, m+1\}$ and the edge set
$E_1 = \{(1,2), (2,3), \ldots, (m, m+1)\} \cup \{(2, 2m), (3, 2m-1), \ldots, (m, m+2)\}$.
T_2 with $N_2 = S_2 \cup L_2 = \{(1,2), \ldots, (m, m+1)\} \cup \{(2, 2m), \ldots, (m, m+2)\}$ and

$$\begin{aligned} E_2 &= \{((1,2),(2,3)), ((2,3),(3,4)), \ldots, ((m-1,m),(m,m+1))\} \\ &\quad \cup \{((2,3),(2,2m)), \ldots, ((m,m+1),(m,m+2))\} \\ &= \{(1,3) \mid 2, (2,4) \mid 3, \ldots, (m-1, m+1) \mid m\} \\ &\quad \cup \{(3,2m) \mid 2, (4, 2m-1) \mid 3, \ldots, (m+1, m+2) \mid m\}. \end{aligned}$$

For $3 \le k \le 2m-2$. If $k = 2l+1$ for some $1 \le l \le m-2$, then

$$T_k \text{ with } N_k = S_k \cup L_k = \left\{A_k^1, A_k^2, \ldots, A_k^{m-l}\right\} \cup \left\{B_k^1, B_k^2, \ldots, B_k^{m-l}\right\} \text{ and}$$

$$\begin{aligned} E_k &= \left\{(A_k^1, A_k^2), (A_k^2, A_k^3), \ldots, (A_k^{m-l-1}, A_k^{m-l}), (A_k^{m-l}, B_k^{m-l})\right\} \\ &\quad \cup \left\{(A_k^2, B_k^1), (A_k^3, B_k^2), \ldots, (A_k^{m-l}, B_k^{m-l-1})\right\}. \end{aligned}$$

If $k = 2l+1$ for some $2 \le l \le m-1$, then

$$T_k \text{ with } N_k = S_k \cup L_k = \left\{A_k^1, A_k^2, \ldots, A_k^{m-l}, A_k^{m-l+1}\right\} \cup \left\{B_k^1, B_k^2, \ldots, B_k^{m-l}\right\} \text{ and}$$

$$\begin{aligned} E_k &= \left\{(A_k^1, A_k^2), (A_k^2, A_k^3), \ldots, (A_k^{m-l}, A_k^{m-l+1})\right\} \\ &\quad \cup \left\{(A_k^2, B_k^1), (A_k^3, B_k^2), \ldots, (A_k^{m-l}, B_k^{m-l-1}), (A_k^{m-l+1}, B_k^{m-l})\right\}. \end{aligned}$$

$$T_{2m-1} \text{ with } N_{2m-1} = \left\{A_{2m-1}^1, B_{2m-1}^1\right\} \text{ and } E_{2m-1} = \left\{(A_{2m-1}^1, B_{2m-1}^1)\right\},$$

where $A_2^i := (i, i+1) \mid \varnothing := (i, i+1)$ and $B_2^i := (i+1, 2m-i+1) \mid \varnothing := (i+1, 2m-i+1)$, and $A_k^i := (i, k-1+i) \mid (i+1 : k-2+i)$ and $B_k^i := (k-1+i, 2m-i+1) \mid (i+1 : k-2+i)$ for $k \ge 3$, $(x : y)$ denotes $(x, x+1, \ldots, y)$.

Figure 2 illustrates the newly proposed SL-vine tree sequence. Furthermore, it satisfies the R-vine tree sequence conditions stated in Definition 2.1, ensuring its validity as a regular vine tree sequence. A detailed proof of this is presented below.

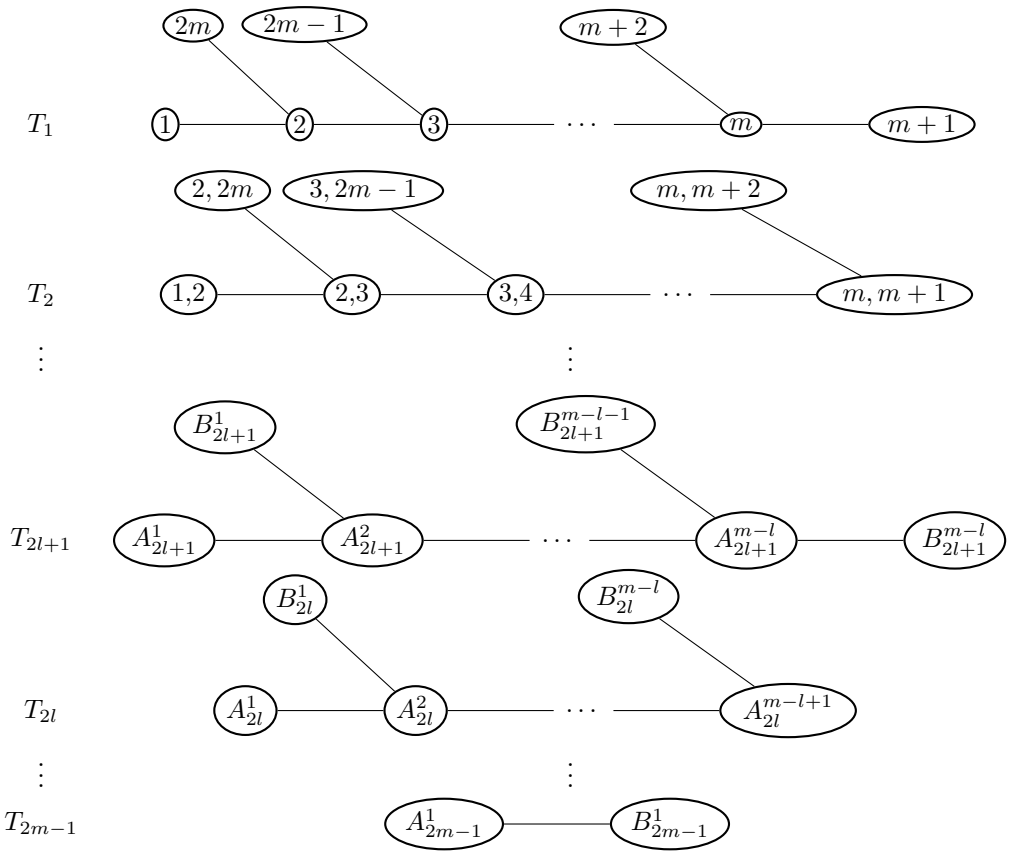


Figure 2: Vine tree structures of an SL-vine

Proposition 3.2. *The SL-vine tree sequence, as defined in Definition 3.1, satisfies the R-vine tree sequence conditions (i)-(iii) outlined in Definition 2.1.*

Proof. To begin with, we prove that an SL-vine tree sequence, $\{T_1, \ldots, T_{2m-1}\}$, satisfies conditions (i)-(iii) from Definition 2.1. The first condition (i) follows directly from the definition of T_1. The second condition requires that $N_k = E_{k-1}$ for all $k \geq 2$. For $k = 2$, this follows immediately from Definition 3.1. To establish the result for $k \geq 3$, it suffices to show that $A_k^i = (A_{k-1}^i, A_{k-1}^{i+1})$ and $B_k^i = (A_{k-1}^{i+1}, B_{k-1}^i)$. For the former case, since

$$A_{k-1}^i = (i, k-2+i) \mid (i+1 : k-3+i) \text{ and}$$
$$A_{k-1}^{i+1} = (i+1, k-1+i) \mid (i+2 : k-2+i),$$

the conditioned and conditioning of $(A_{k-1}^i, A_{k-1}^{i+1})$ are $(i, k-1+i)$ and $(i+1 : k-2+i)$, respectively. That is, $A_k^i = (A_{k-1}^i, A_{k-1}^{i+1})$. For the latter case, the conditioned and conditioning of $(A_{k-1}^{i+1}, B_{k-1}^i)$, where

$$A_{k-1}^{i+1} = (i+1, k-1+i) \mid (i+2 : k-2+i) \text{ and}$$
$$B_{k-1}^i = (k-2+i, 2m-i+1) \mid (i+1 : k-3+i),$$

are $(k-i+1, 2m-i+1)$ and $(i+1 : k-2+i)$, respectively. Consequently, $B_k^i = (A_{k-1}^{i+1}, B_{k-1}^i)$. Thus, by Definition 3.1, it follows that $N_k = E_{k-1}$ for all k in the SL-vine tree sequence. The

final requirement, known as the proximity condition, stipulates that for $k \geq 2$, two nodes in T_k can be connected only if their corresponding edges in the preceding tree T_{k-1} share a common node. For $k = 2$, it is trivial. To prove the statement for $k \geq 3$, we divide into two cases.
Case I $k = 2l + 1$ for some $1 \leq l \leq m - 2$. Let $(e_1, e_2) \in E_k$.

Case 1.1 $e_1 = A_k^i$, $e_2 = A_k^{i+1}$ for some $i = 2, 3, \ldots, m - l - 1$.
In T_{k-1}, the corresponding edges of e_1 and e_2 are $(A_{k-1}^i, A_{k-1}^{i+1})$ and $(A_{k-1}^{i+1}, A_{k-1}^{i+2})$, respectively, which have the common node A_{k-1}^{i+1}.

Case 1.2 $e_1 = A_k^{m-l}$, $e_2 = B_k^{m-l}$.
In T_{k-1}, the corresponding edges of e_1 and e_2 are $(A_{k-1}^{m-l}, A_{k-1}^{m-l+1})$ and $(A_{k-1}^{m-l+1}, B_{k-1}^{m-l})$, respectively, which have the common node A_{k-1}^{m-l+1}.

Case 1.3 $e_1 = A_k^i$, $e_2 = B_k^{i-1}$ for some $i = 2, 3, \ldots, m - l$.
In T_{k-1}, the corresponding edges of e_1 and e_2 are $(A_{k-1}^i, A_{k-1}^{i+1})$ and (A_{k-1}^i, B_{k-1}^i), respectively, which have the common node A_{k-1}^i.

Case II $k = 2l$ for some $2 \leq l \leq m - 1$. Let $(e_1, e_2) \in E_k$.

Case 2.1 $e_1 = A_k^i$, $e_2 = A_k^{i+1}$ for some $i = 2, 3, \ldots, m - l - 1$.
In T_{k-1}, the corresponding edges of e_1 and e_2 are $(A_{k-1}^i, A_{k-1}^{i+1})$ and $(A_{k-1}^{i+1}, A_{k-1}^{i+2})$, respectively, which have the common node A_{k-1}^{i+1}.

Case 2.2 $e_1 = A_k^{m-l}$, $e_2 = A_k^{m-l+1}$.
In T_{k-1}, the corresponding edges of e_1 and e_2 are $(A_{k-1}^{m-l}, A_{k-1}^{m-l+1})$ and $(A_{k-1}^{m-l+1}, B_{k-1}^{m-l})$, respectively, which have the common node A_{k-1}^{m-l+1}.

Case 2.3 $e_1 = A_k^i$, $e_2 = B_k^i$ for some $i = 2, 3, \ldots, m - l + 1$.
In T_{k-1}, the corresponding edges of e_1 and e_2 are $(A_{k-1}^i, A_{k-1}^{i+1})$ and (A_{k-1}^i, B_{k-1}^i), respectively, which have the common node A_{k-1}^i. □

Consequently, the SL-vine tree sequence fulfills the R-vine tree sequence conditions outlined in Definition 2.1. A regular vine copula constructed from an SL-vine tree sequence, along with a set of bivariate copulas $\mathcal{B}(\mathcal{V})$ and their corresponding pair-copula parameters $\theta(\mathcal{B}(\mathcal{V}))$, is termed an SL-vine copula, denoted by $\mathcal{SL}$. We construct an SL-vine distribution for a $2m$-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_{2m})$ in the following theorem.

Theorem 3.3. *The joint density $f(x_1, \ldots, x_{2m})$ of an SL-vine copula is given as*

$$\begin{aligned} f(x_1, \ldots, x_{2m}) = & \left[\prod_{\ell=1}^{m-1}\prod_{j=1}^{m-\ell-1} c_{A_{2\ell+2}^j}\right] \times \left[\prod_{\ell=1}^{m-1}\prod_{j=1}^{m-\ell-1} c_{B_{2\ell+2}^j}\right] \times \left[\prod_{\ell=1}^{m-1} c_{A_{2\ell+1}^{m-\ell}, B_{2\ell+1}^{m-\ell}}\right] \\ & \times \left[\prod_{\ell=2}^{m-1}\prod_{j=1}^{m-\ell} c_{A_{2\ell+1}^j}\right] \times \left[\prod_{\ell=2}^{m-1}\prod_{j=1}^{m-\ell} c_{B_{2\ell+1}^j}\right] \times \left[\prod_{i=1}^{m-1} c_{(i,i+2)|(i+1)}\right] \\ & \times \left[\prod_{i=m}^{2m-2} c_{(2m-i+1,i+2)|(2m-i)}\right] \times \left[\prod_{i=1}^{m} c_{i,i+1}\right] \\ & \times \left[\prod_{i=m+1}^{2m-1} c_{(2m-i+1,i+1)}\right] \times \left[\prod_{i=1}^{2m} f_i(x_i)\right], \end{aligned} \tag{3.1}$$

where A_k^i and B_k^i are defined in Definition 3.1.

Proof. The joint density $f(x_1, \ldots, x_{2m})$, derived from the SL-vine tree sequence as described in Definition 3.1, can be expressed in equation (2.5) as

$$
\begin{aligned}
f(x_1, \ldots, x_{2m}) = & \overbrace{\prod_{\ell=1}^{m-1} \left[\left(\prod_{j=1}^{m-\ell-1} c_{A_{2\ell+1}^j, A_{2\ell+1}^{j+1}} \right) \left(\prod_{j=1}^{m-\ell-1} c_{A_{2\ell+1}^{j+1}, B_{2\ell+1}^j} \right) c_{A_{2\ell+1}^{m-\ell}, B_{2\ell+1}^{m-\ell}} \right]}^{T_{2l+1}} \\
& \times \overbrace{\prod_{\ell=2}^{m-1} \left[\left(\prod_{j=1}^{m-\ell} c_{A_{2\ell}^j, A_{2\ell}^{j+1}} \right) \left(\prod_{j=1}^{m-\ell} c_{A_{2\ell}^{j+1}, B_{2\ell}^j} \right) \right]}^{T_{2l}} \times \left[\prod_{i=1}^{2m} f_i(x_i) \right] \\
& \times \overbrace{\left[\left(\prod_{i=1}^{m-1} c_{(i,i+2)|(i+1)} \right) \left(\prod_{i=m}^{2m-2} c_{(2m-i+1,i+2)|(2m-i)} \right) \right]}^{T_2} \\
& \times \overbrace{\left[\left(\prod_{i=1}^{m} c_{i,i+1} \right) \left(\prod_{i=m+1}^{2m-1} c_{(2m-i+1,i+1)} \right) \right]}^{T_1}.
\end{aligned}
$$

Since $A_k^i = (A_{k-1}^i, A_{k-1}^{i+1})$ and $B_k^i = (A_{k-1}^{i+1}, B_{k-1}^i)$, equation (3.1) holds. □

From the definition of an SL-vine, if the first tree T_1 is determined, then the kth tree T_k is uniquely defined for $k = 2, 3, \ldots, 2m-1$. However, it is possible to increase the flexibility of trees $T_2, T_3, \ldots, T_{2m-1}$ and still follow the same stem-and-leaf structure with more freedom in choosing edges. For example, the Figure 3 below illustrates all possible edges in constructing T_2 for the R-vine case with $m = 3$, given that T_1 is determined from the SL-vine. To form connected trees, we must select and remove edges in ways that ensures the graph remains connected and acyclic. Thus, all possible connected trees, which have a structure akin to T_2 in Figure 2, are presented in Figure 4. Similarly to T_2, we obtain that T_3, T_4, and T_5 maintain a structure resembling an SL-vine. Furthermore, instead of predefining T_1, we could adopt an algorithmic approach to potentially improve the efficiency of variable selection. For example, the Traveling Salesman Algorithm could be used to select stems, followed by matching them with leaves using the Hungarian Algorithm.

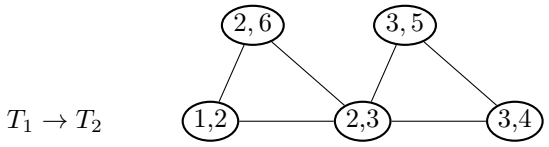


Figure 3: All possible edges in constructing T_2 for the R-vine case with $m = 3$

Recall that the vine copula is a tool for analyzing and modeling the dependence structure of multivariate random variables. When multiple choices of connected trees are available, it makes more sense to prioritize selecting the strongest dependencies first. We incorporate this principle in constructing a vine tree structure. To describe the strength of dependence between random

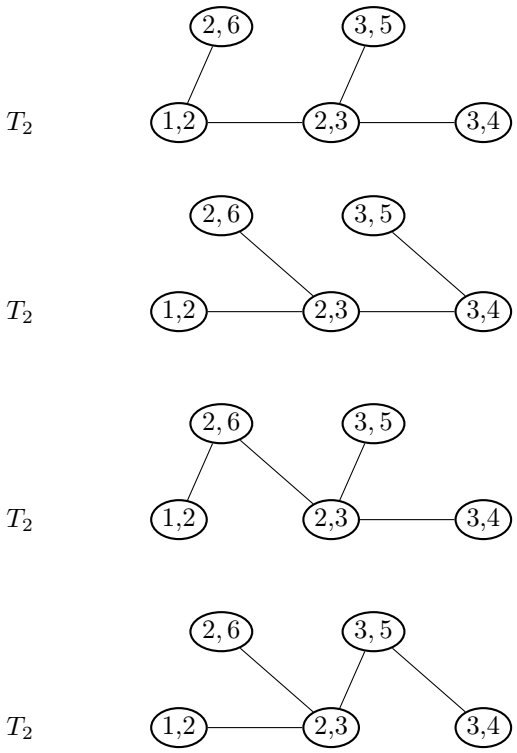


Figure 4: All possible connected trees corresponding to T_2

variables, **measures of concordance** are commonly used. One of the most common measures is **Kendall's Tau**. Let (X, Y) be a continuous random vector. The Kendall's Tau is defined by

$$\tau_{(X,Y)} \equiv \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) > 0) - \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) < 0),$$

where (X_1, Y_1) and (X_2, Y_2) are independent random vectors both with the same joint distribution as (X, Y), denoted by F. Note that the value of $\tau_{(X,Y)}$ always lies within the interval $[-1, 1]$. Moreover, Kendall's Tau can be expressed in terms of the copula c of X and Y as follows:

$$\tau(C) = 4 \int_0^1 \int_0^1 C(u_1, u_2) c(u_1, u_2) du_1 du_2 - 1.$$

As such, we select and remove the edge with the smallest absolute value of Kendall's τ while ensuring the graph remains connected and acyclic.

4 Application

4.1 Model Selection

In this context, the model consists of three key components: structure (trees), copula families, and copula parameters. In this study, we restrict the structure to our proposed method, the SL-vine decomposition. For each edge in the model, it is essential to determine the copula family and its corresponding parameters. The optimal copula families and their parameters are optimized separately for each edge. Common criteria used for selecting the copula family are

the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) with the formula provided below:

$$\mathrm{AIC}(q) = -2\left[\ln \mathcal{L}\left(\hat{\vec{\theta}}\right) - q\right] \text{ and } \mathrm{BIC}(q) = -2\left[\ln \mathcal{L}\left(\hat{\vec{\theta}}\right) - \frac{q}{2}\ln n\right],$$

where $\hat{\vec{\theta}}$ is the parameter vector that maximizes the log-likelihood, $\ln \mathcal{L}\left(\hat{\vec{\theta}}\right)$, q is the number of parameters, and n is the number of data points.

The AIC and BIC calculations rely on the log-likelihood function, which is derived from the SL-vine pair-copula construction. The copula family that minimizes these criteria is selected as the optimal choice. Once the copula family is determined, its corresponding parameters are estimated using the maximum likelihood estimation (MLE) method, as outlined in [9].

4.2 Data and Result

4.2.1 Abalone Dataset

To demonstrate the process of model selection, we utilize the abalone dataset from the University of California–Irvine repository (`https://doi.org/10.24432/C55C7W`) and apply our model selection methodology to it. This dataset provides detailed information on marine snails known as abalones, whose shells exhibit a characteristic spiral structure. It includes various measurements, including diameter, height, and different weight categories (whole, shucked, viscera, and shell), along with the number of rings, which serves as an indicator of age.

In this study, we focus exclusively on female abalone shells, using a dataset of 1,307 observations. We consider the following variables: whole weight (whole), shucked weight (shuck), viscera weight (vis), shell weight (shell), diameter (dia), length (len), height (hei), and rings. These features are selected to facilitate the application of our proposed model selection approach within a meaningful biological context. Summary statistics are provided in Table 1.

Table 1: Summary statistics of selected variables

Variable	Mean	SD	Median	Min	Max	Skew	Kurtosis
Whole	1.05	0.43	1.04	0.08	2.66	0.37	0.05
Shucked	0.45	0.20	0.44	0.03	1.49	0.55	0.65
Viscera	0.23	0.10	0.22	0.02	0.59	0.39	-0.10
Shell	0.30	0.13	0.30	0.03	0.98	0.69	1.40
Diameter	0.45	0.07	0.47	0.20	0.65	-0.51	0.15
Length	0.58	0.09	0.59	0.28	0.81	-0.53	0.10
Height	0.16	0.04	0.16	0.01	1.13	10.90	265.19
Rings	11.13	3.10	10.00	5.00	29.00	1.47	3.10

In the preprocessing step, non-parametric marginals, estimated using the empirical distribution function, are used to determine the marginal distribution functions and generate pseudo-copula data. Next, we apply the Dißmann algorithm [10], which relies on the absolute value of Kendall's Tau, along with the sequence of tree structure in Definition 3.1, to fit the SL-vine copula models to the pseudo-data. The resulting vine tree structures are shown in Figure 5,

while the selected pair-copula families and their parameter estimates are summarized in Table 2.

To provide a comparative benchmark, we also fit R-, C-, D-, and Gaussian vine copula models. Unlike the multivariate Gaussian model (also known as the standard multivariate normal distribution), the Gaussian vine copula is more general, as it allows for non-Gaussian marginal distributions. As shown in Table 3, since R-, C-, D-, Gaussian, and SL-vines have different joint density functions, their corresponding log-likelihood values also differ. In this table, although the Gaussian vine shares the same structure as the R-vine, it differs in that its pair copulas are strictly defined by the Gaussian copula family. Furthermore, the SL-vine copula model demonstrates the best performance in terms of AIC and BIC among all considered models. All computations were performed in R relying on the R package `VineCopula` and `rvinecopulib` of Nagler and Vatter [11].

Table 2: Sequential parameter estimates, chosen copula families, and inferred dependence measures for eight variables in the female abalone dataset

Tree	**Edge**[a]	**Family**	**Parameter(s)**	**Kendall's τ**
1	(2,3)	Survival Gumbel	2.46	0.59
1	(2,1)	Survival Gumbel	6.42	0.84
1	(4,2)	Tawn type 2 (180°)	(5.72, 0.97)	0.81
1	(4,8)	Tawn type 1 (180°)	(1.81, 0.32)	0.20
1	(5,4)	Tawn type 1 (180°)	(5.91, 0.98)	0.82
1	(5,6)	Survival Gumbel	3.64	0.73
1	(7,5)	Survival BB6	(1.60, 2.16)	0.65
2	(4,3 | 2)	Gaussian	0.38	0.25
2	(4,1 | 2)	Survival BB1	(0.27, 1.24)	0.29
2	(5,2 | 4)	Tawn type 1 (270°)	(-1.30, 0.18)	-0.08
2	(5,8 | 4)	Gaussian	-0.55	-0.37
2	(7,4 | 5)	Student's t	(0.88, 4.11)	0.68
2	(7,6 | 5)	Student's t	(0.47, 11.29)	0.31
3	(5,3 | 4,2)	Frank	-1.84	-0.20
3	(5,1 | 4,2)	Survival Gumbel	1.10	0.09
3	(7,2 | 5,4)	BB1	(0.09, 1.14)	0.16
3	(7,8 | 5,4)	BB1	(0.06, 1.11)	0.13
3	(6,4 | 7,5)	Student's t	(0.68, 5.78)	0.48
4	(7,3 | 5,4,2)	Student's t	(0.19, 22.11)	0.12
4	(7,1 | 5,4,2)	Tawn type 1 (270°)	(-1.25, 0.06)	-0.03
4	(6,2 | 7,5,4)	Gaussian	0.08	0.05
4	(6,8 | 7,5,4)	BB8 (90°)	(-1.35, -0.77)	-0.07
5	(6,3 | 7,5,4,2)	Survival BB8	(1.10, 0.99)	0.05
5	(6,1 | 7,5,4,2)	Frank	0.89	0.10
5	(8,2 | 6,7,5,4)	Tawn type 1 (180°)	(1.51, 0.01)	0.01
6	(1,3 | 6,7,5,4,2)	Frank	-0.62	-0.07
6	(8,1 | 6,7,5,4,2)	Student's t	(-0.09, 15.51)	-0.06
7	(8,3 | 1,6,7,5,4,2)	Frank	(0.41,0.00)	0.05

[a]Variables are len (1), dia (2), hei (3), whole (4), shuck (5), vis (6), shell (7), ring (8). **Abbreviations:** BB, biparameter bivariate.

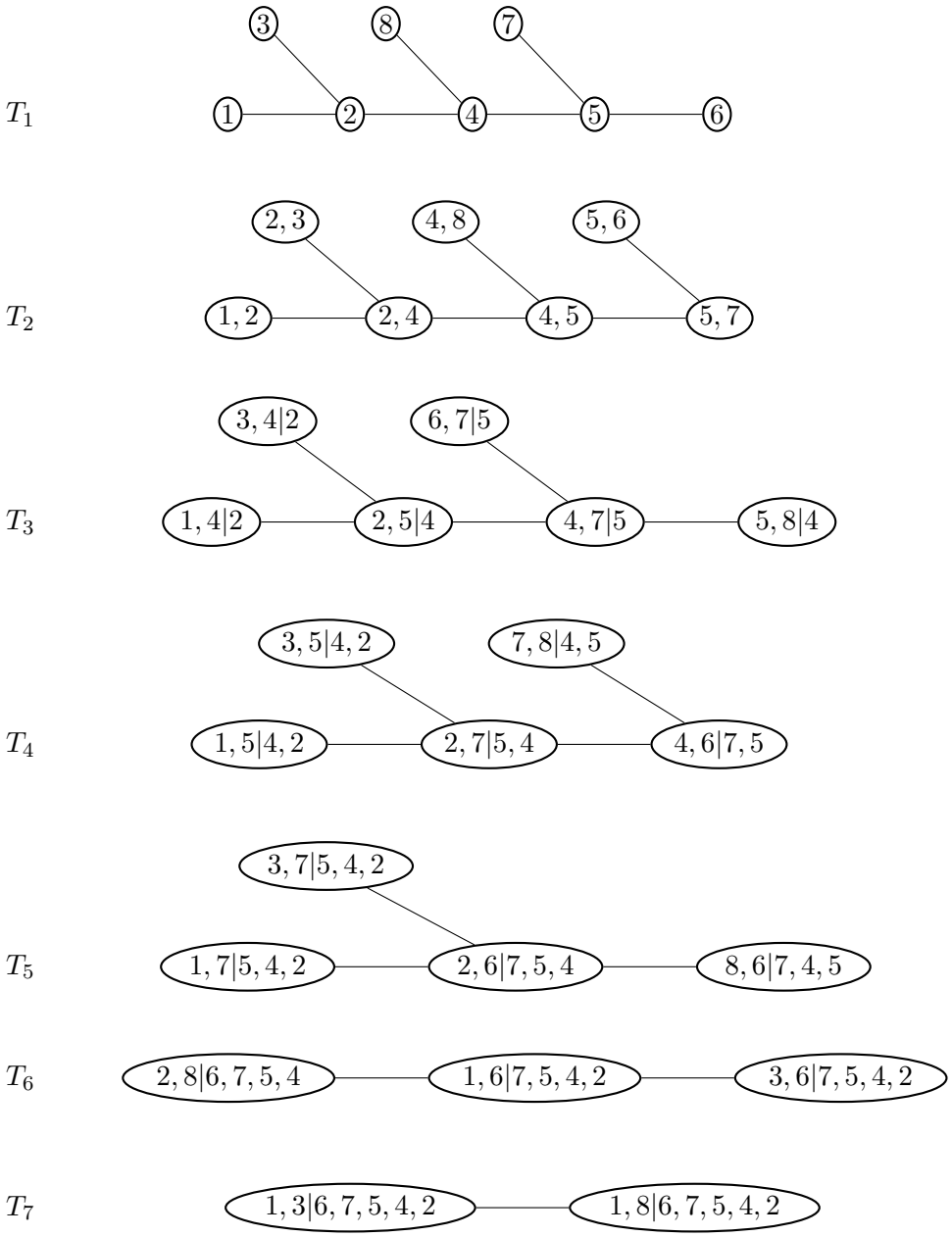


Figure 5: SL-vine tree structures for variables len (1), dia (2), hei (3), whole (4), shuck (5), vis (6), shell (7), and rings (8) from the female abalone data set

4.2.2 Capital-Market dataset

Although SL-vine demonstrates the highest overall performance on the abalone dataset compared to other copula models, there are certain scenarios in which R-vine outperforms it. In this case, we utilize the capital-market dataset [12], comprising 14 variables reflecting capital market activity across three levels: domestic (Thailand), regional (Asia/Emerging markets), and international (United States). The data consist of 14 time series, recorded over 3,389 trading days

Table 3: Comparison of copula models the abalone dataset based on Log-Likelihood, AIC, and BIC.

Copula model	Log-likelihood	AIC	BIC
SL-vine copula	$10,444.72$	$-20,799.44$	$-20,566.54$
C-vine copula	$10,421.85$	$-20,747.70$	$-20,499.27$
R-vine copula	$10,411.47$	$-20,730.93$	$-20,492.86$
D-vine copula	$9,905.04$	$-19,720.08$	$-19,487.18$
Gaussian vine copula	$9,363.33$	$-18,670.66$	$-18,525.75$

within a 13-year timeframe from January 2009 to December 2021. These data were collected from the Stock Exchange of Thailand (SET), the Thai Bond Market Association (Thai BMA), and Bloomberg as follows.

1. **SET**: daily (logarithmic) return of the Stock Exchange of Thailand Index.
2. **MAI**: daily (logarithmic) return of the Market for Alternative Investment Index.
3. **ZeroShort**: daily variation in the average yield of 3M/6M/1Y Thai government zero-coupon bonds.
4. **ZeroLong**: daily variation in the average yield of 2Y/5Y/10Y Thai government zero-coupon bonds.
5. **CorpBond**: daily change in spread between BBB-rated corporate bonds and 2Y/5Y/10Y Thai government zero-coupon yields.
6. **THB**: daily (logarithmic) return of the Thai baht against the US dollar (THB/USD), equivalent to the inverse of USD/THB.
7. **EquityFlow**: daily net flow (in USD million) of foreign investment into and out of Thai equity securities (SET).
8. **BondFlow**: daily net flow (in THB million) from non-resident investors into or out of Thai bonds (Thai BMA).
9. **EMAsiaEquity**: daily (logarithmic) return of Bloomberg MSCI Emerging Markets Asia Index (M2MS).
10. **SP500**: daily (logarithmic) return of the S&P 500 Index.
11. **EMBond**: daily (logarithmic) return of the Bloomberg Emerging Markets USD Aggregate Total Return Index (EMUSTRUU).
12. **USBond**: daily (logarithmic) return of the Bloomberg US Universal Total Return Index (LC07TRUU).
13. **EMAsiaFX**: daily (logarithmic) return of the Bloomberg JPMorgan Asia Dollar Index (ADXY).

14. **USD**: daily (logarithmic) return of the Bloomberg Dollar Spot Index (BBDXY).

Descriptive statistics of financial series are presented in Table 4. To assess model performance on the new dataset, we also apply R-, C-, D-, Gaussian, and SL-vine copulas as benchmarks. As summarized in Table 5, among the various copula models applied to the capital-market dataset, the SL-vine copula demonstrates strong overall performance, ranking just behind the R-vine copula in terms of model quality. This occurs because the R-vine employs a greedy algorithm based on Kendall's Tau to select its tree structure at each level, allowing for far greater structural flexibility than the SL-vine. However, this very flexibility can be a drawback: although the first few trees of an R-vine may outperform those of an SL-vine, there is no guarantee that the subsequent levels will maintain the optimal structure. For example, as demonstrated with the abalone dataset, later R-vine trees do not always achieve the best possible configuration.

Table 4: Descriptive Statistics of Financial Series

Series	**Mean**	**SD**	**Median**	**Min**	**Max**	**Skew**	**Kurtosis**
SET	0.0367	1.059	0.012	-11.43	7.65	-0.923	14.91
MAI	0.0368	1.102	0.044	-8.00	8.05	-1.070	11.51
ZeroShort	-0.0004	0.013	0.000	-0.25	0.09	-5.679	99.13
ZeroLong	-0.0003	0.027	0.000	-0.21	0.10	0.652	13.96
CorpBond	-0.0001	0.011	0.000	-0.12	0.10	-0.008	19.58
THB	0.0015	0.286	0.010	-1.56	1.37	-0.068	4.87
EquityFlow	-7.2057	66.876	-2.42	-543.16	826.99	0.584	16.92
BondFlow	1,082.91	4,571.04	154.96	-26,973.48	37,136.74	1.553	10.40
EMAsiaEquity	0.0384	1.107	0.070	-5.87	5.62	-0.341	6.24
SP500	0.0483	1.125	0.050	-12.77	8.97	-0.692	16.76
EMBond	0.0289	0.252	0.030	-3.71	2.02	-2.538	37.42
USBond	0.0164	0.198	0.020	-2.19	1.25	-0.735	10.59
EMAsiaFX	0.0005	0.234	0.010	-1.64	1.32	-0.219	5.74
USD	0.0031	0.397	0.000	-2.51	2.26	0.038	5.29

Table 5: Comparison of copula models for capital-market dataset based on Log-Likelihood, AIC, and BIC.

Copula model	Log-likelihood	AIC	BIC
R-vine copula	$8,038.15$	$-15,790.31$	$-14,913.96$
SL-vine copula	$8,012.37$	$-15,716.73$	$-14,772.97$
C-vine copula	$7,969.66$	$-15,653.32$	$-14,776.98$
D-vine copula	$7,900.42$	$-15,486.85$	$-14,524.71$
Gaussian vine copula	$6,615.55$	$-13,049.10$	$-12,491.43$

5 Conclusions and Outlook

In conclusion, this study introduces the SL-vine copula model as a novel framework for multivariate dependence modeling, addressing the limitations of traditional copula models in handling high-dimensional data and mixed dependency structures. By imitating the structure of canonical (C-) and drawable (D-) vines, the SL-vine captures both hierarchical and sequential dependencies, making it well-suited for real-world datasets with complex relationships. The model's formal definition and compliance with regular vine (R-vine) tree sequence requirements demonstrate its mathematical validity and its joint density function. The application of the SL-vine to the abalone dataset illustrates its superior performance in modeling intricate relationships between variables such as weight and size. Compared to Gaussian and other conventional vine copula models, the SL-vine copula model performs better than all other models. While applying the SL-vine model to the capital-market dataset, it still delivers superior performance compared to the other models, being second only to the R-vine model. Overall, the SL-vine model offers an innovative alternative for multivariate analysis, providing new avenues for exploring complex data structures in both theoretical and practical contexts.

Although the structure of an SL-vine is constrained to a single form once T_1 is specified, it can be made more flexible, as demonstrated at the end of Section 3. An interesting question is whether it is possible to define a more flexible structure than the SL-vine without specifying T_1. In Figure 4, it is evident that in each possible tree structure, no node is connected to more than three edges. In other words, the vine structure has a degree of at most three. Furthermore, this structure maintains a sequential arrangement and features an intermediary node that connects to others, which are characteristic properties of D-vine and C-vine, respectively. Therefore, as a direction for future research, our aim is to investigate vine structures with a degree constraint of at most three.

Acknowledgment. The author appreciates the referees for their thorough review of the manuscript and their valuable feedback.

References

[1] H. Joe, Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters, Lecture Notes-Monograph Series 28 (1996) 120–141.

[2] T. Bedford, R.M. Cooke, Vines: A new graphical model for dependent random variables, Ann. Stat. 30 (2002) 1031–1068.

[3] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Stat. Univ. Paris 8 (1959) 229–231.

[4] D. Kurowicka, R. Cooke, Uncertainty Analysis with High Dimensional Dependence Modelling, Wiley Ser. Probab. Stat., John Wiley & Sons, Ltd, Chichester, 2006. `https://doi.org/10.1002/0470863072`.

[5] I.H. Haff, K. Aas, A. Frigessi, On the simplified pair-copula construction - simply useful or too simplistic?, J. Multivar. Anal. 101 (2010) 1296–1310.

[6] J. Stöber, H. Joe, C. Czado, Simplified pair copula constructions - limitations and extensions, J. Multivar. Anal. 119 (2013) 101–118.

[7] M.S. Kurz, F. Spanhel, Testing the simplifying assumption in high-dimensional vine copulas, Electron. J. Stat. 16 (2) (2022) 5226–5276. https://doi.org/10.1214/22-EJS2051.

[8] A. Derumigny, J.D.F., About tests of the "simplifying" assumption for conditional copulas, Depend. Model. 5 (1) (2017) 154–197.

[9] R.B. Millar, Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB, Wiley Ser. Probab. Stat., John Wiley & Sons, Ltd, Chichester, 2011. https://doi.org/10.1002/9780470094846.

[10] J. Dißmann, E.C. Brechmann, C. Czado, D. Kurowicka, Selecting and estimating regular vine copulae and application to financial returns, Comput. Stat. Data Anal. 59 (2013) 52–69. https://doi.org/10.1016/j.csda.2012.08.010.

[11] T. Nagler, T. Vatter, rvinecopulib: High performance algorithms for vine copula modeling, R package version 0.6.3.1 (2023). https://CRAN.R-project.org/package=rvinecopulib.

[12] P. Nacaskul, K. Kalakan, P. Khlaisamniang, P. Veerabulyarith, I. Sukcharoenchaikul, Multi-Paradigm Analysis of Thai Capital Market Linkages: Bivariate/Vine Copulas, Granger Causality, Network Centrality, and Graph Neural Network/Graph Embedding Approaches, (2023). https://paperswithcode.com/paper/multi-paradigm-analysis-of-thai-capital.

Portfolio selection by GJR-GARCH model and Gaussian Mixture model

Kornprapa Chotsirikunnawat[1,*] and Dawud Thongtha[1,†]

[1]Department of Mathematics, Faculty of Science,
King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand

Abstract

In financial markets, asset prices constantly change and volatilize due to various factors such as economic conditions, politics, and significant global events. Volatility is one of the key factors that influences investor's decisions. Using models that are able to estimate such volatility is important in constructing investment portfolios. This research aims to construct portfolio based on the mean-variance theory by utilizing the Gaussian Mixture model (GMM) and the GJR-GARCH model with constant conditional correlations. Both models are used to estimate covariance matrices which are used for constructing portfolios. Daily asset price data from 2014 to 2024 were used in this study. Additionally, the portfolio allocation was adjusted every 10 days. The performance of the two portfolios is compared based on cumulative amount of money, standard deviation of return rate, and Sharpe ratio. The results of the study indicated that the portfolio created using the GJR-GARCH model achieved a cumulative amount of growth of 4.211739 times the initial capital, which is higher than the GMM. However, the 10-day return standard deviation was 0.041234 which is slightly higher than that from GMM. Additionally, the GJR-GARCH model exhibited a superior Sharpe ratio.

Keywords: Gaussian Mixture model, GJR-GARCH-CCC, Portfolio optimization
2020 MSC: 62M10, 62P05, 91G10

1 Introduction

Financial markets are inherently complex and uncertain. A key characteristic of the stock market is the fluctuation in prices and volatility, which is influenced by a range of factors such as changes in economic conditions, political instability, and global events. According to Eugene Fama (1965) [1], stock prices movement have been changed responding to new information, suggesting that markets are efficient in processing current data. To manage investment risks and enhance returns, Markowitz (1952) [2] introduced the concept of portfolio optimization. This framework is based on diversification, aiming to minimize total portfolio risk by allocating assets across different investments. This foundation was later extended by Elton and Gruber (1997) [3], who emphasized constructing portfolios that maximize expected return for a given

*Speaker (kornprapa.chot@mail.kmutt.ac.th)
†Corresponding author (dawud.tho@kmutt.ac.th)

level of risk. However, the static assumptions of MPT, particularly concerning constant volatility and correlation, have been challenged by empirical evidence from real-world markets, leading to the adoption of more flexible, data-driven models.

One approach involves the use of Gaussian Mixture Models (GMM) in financial modeling. GMM offer a probabilistic framework that represents return distributions as a combination of multiple normal distributions, each capturing different market conditions. This enables the model to better describe the non-normality and heteroskedasticity often observed in financial time series. GMM have been applied to estimate conditional expectations and covariances for portfolio optimization, effectively capturing the presence of heavy tails and skewness in asset returns. For instance, Xu and Gencay (2003) [4] showed that GMM forecasts can outperform single-regime models in capturing nonlinearities in asset returns. Similarly, Wang and Zhou (2012) [5] used GMM to improve asset allocation under non-Gaussian distributions, demonstrating the model's potential in enhancing portfolio performance.

Another widely adopted method in financial econometrics is the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced by Bollerslev (1986) [6] as an extension of Engle's (1982) [7] ARCH model. GARCH models have become a cornerstone in modeling time-varying volatility, which is crucial for risk management and forecasting. However, standard GARCH models assume symmetric responses to shocks, whereas empirical studies reveal that negative news often has a greater impact on volatility than positive news—a phenomenon known as the leverage effect. To address this, Glosten, Jagannathan, and Runkle (1993) [8] proposed the GJR-GARCH model, which introduces an asymmetric component that increases volatility more after negative returns. When dealing with portfolios of multiple assets, Engle (2002) [9] proposed models such as the Constant Conditional Correlation (CCC) and the Dynamic Conditional Correlation (DCC) GARCH models to estimate time-varying covariances. The CCC-GJR-GARCH model combines asymmetry in volatility with a constant correlation framework, and has been applied by researchers like Arouri et al. (2011) [10] and Conrad et al. (2014) [11] to better capture the joint dynamics of asset returns under asymmetric market conditions.

This paper is organized as follows. Section 2 outlines the theoretical background and preliminary concepts, including the portfolio optimization, estimation methods for both GMM and GJR-GARCH. Section 3 presents the empirical results from a portfolio optimization perspective, including the performance comparison between the two models. Finally, Section 4 offers a discussion of the findings and their implications for financial decision-making under volatile conditions.

2 Preliminaries

In this section, various models used in this research are described. The portfolio selection is estimated under an optimization problem. An expected return and covariance matrix are two main factors used in the portfolio construction. Two main methods, the Gaussian mixture model (GMM) and the GJR-GARCH model, for estimating these factors are also explained.

2.1 Portfolio Optimization Model

The concept portfolio selection was first introduced by Markowitz in 1952 [2]. His idea based on optimization problem for investigating optimal weights for minimizing risk. His model is known

as the mean-variance (MV) optimization. His concept was extended by many researchers such as the Modern Portfolio Theory (MPT) by Sharpe (1964) [12], the Multi-Objective Optimization Problem (MOPT) framework [13], and the OGARCH model developed by Luo et al. (2015) [14]. For example, Engle (1982) introduced a mathematical model for finding optimal weights with maximum return under a fixed level of risk through the use of the ARCH model [7]. In this work, by following the optimization problem in Saunders [15], the portfolio weights are determined by simultaneously maximizing expected return and minimizing risk.

Consider a portfolio consisting of n assets, let $r_{i,t}$ denote the return of asset i at time t for $i = 1, 2, \ldots, n$ and $t \geq 0$ and R_t be the return vector at time t. That is,

$$R_t = (r_{1,t}, r_{2,t}, \ldots, r_{n,t})$$

The optimization problem for portfolio selection given in Saunders [15] is as follows:

$$\min_{\omega_t} \left\{ \lambda\, \omega_t^\top \operatorname{Var}_{t-1}(R_t)\, \omega_t - \omega_t^\top \mathbb{E}_{t-1}(R_t) \right\} \tag{2.1}$$

subject to the constraint:

$$\sum_{i=1}^{k} \omega_{i,t} = 1$$

where $\omega_{i,t}$ represents the proportion of wealth allocated to asset i at time t, $w_t = (w_{1,t}, w_{2,t}, \ldots, w_{n,t})$ is the weight vector, w_t^T is the transpose of w_t, $\mathbb{E}_{t-1}(R_t)$ represents the conditional expected return vector, and $\operatorname{Var}_{t-1}(R_t)$ denotes the conditional covariance matrix of asset returns when the information is available up to time $t-1$. The parameter λ reflects the investor's risk aversion. In this study, a constant risk aversion coefficient of $\lambda = 11.39$ is used throughout the analysis, consistent with the value employed in Saunder [15].

2.2 Gaussian Mixture Model

The Gaussian Mixture Model (GMM) is a regime switching model introduced by Hamilton in 1989 [16] The GMM assumes independent and identically distributed (i.i.d.) of state process (S_t). Although the i.i.d. assumption may not be fully held in financial time series due to features like autocorrelation and volatility clustering, GMM remains valuable for modeling the cross-sectional distribution of returns within a rolling window framework. The first step involves classifying return observations into a number of regimes. In this work, the GMM is considered in two regimes. These regimes are considered based on the sign of the daily log-return. The state 0 is positive return regime (log-return > 0) and state 1 is negative or zero return regime (log-return ≤ 0).

After separating the data into two regimes, the expected return and the covariance matrix can be computed by using the transition probabilities between two regimes. The probabilities p and q are defined as $p = P(S_t = 1)$ and $q = P(S_t = 0)$, where S_t denotes the regime at time t. Note that, in this case, $p + q = 1$. In order to construct an optimal portfolio from optimization as in equation (2.1), we need the conditional expected return vector and conditional variance matrix. In 2013 [15], Saunders provided formulas for computing the expected return μ and covariance Σ of n assets. Their formulas for one step period are as follow:

$$\mu = (1 - p)\mu_0 + p\mu_1 \tag{2.2}$$

and

$$\Sigma = (1-p)\Sigma_0 + p\Sigma_1 + (1-p)p(\mu_1 - \mu_0)(\mu_1 - \mu_0)^\top \tag{2.3}$$

where μ_k and Σ_k are the $(n \times 1)$ vector of expected values and the $(n \times n)$ covariance matrix at state k, for $k = 0, 1$, respectively. The conditional covariance elements are computed as:

$$\Sigma_k^{(i,j)} = C_k^{(i,j)} \sigma_k^i \sigma_k^j$$

where C_k is the correlation matrix at state k, $C_k^{(i,j)}$ is the $(i,j)^{th}$ entry of C_k, σ_k^i and σ_k^j are the conditional standard deviation of asset i and j at state k, respectively. In this work, the regime expected return μ and the covariance matrix Σ are adjusted for each window and then used as inputs to the optimization problem in equation 2.1, allowing the portfolio to dynamically adapt based on market regimes identified via GMM.

2.3 GJR-GARCH Model

The GJR-GARCH model was introduced by Glosten, Jagannathan, and Runkle (1993) [8] as an extension of the standard GARCH framework to capture asymmetric effects in conditional volatility. The GJR-GARCH model incorporates this asymmetry by including an indicator function that distinguishes between positive and negative past innovations.
Let r_t denote the return of a financial asset at time t. The return process can be written as:

$$r_t = \mu + \varepsilon_t,$$

where μ is the constant mean of the return and ε_t is the shock term at time t.
The shock term is modeled as:

$$\varepsilon_t = \sigma_t z_t, \quad z_t \sim \mathcal{N}(0,1),$$

where σ_t^2 is the conditional variance and z_t is an i.i.d. standard normal variable.
The conditional variance in the GJR-GARCH(1,1) model is defined as:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \gamma I_{t-1}\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{2.4}$$

where $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $\gamma \geq 0$ are estimated parameters. The indicator function I_{t-1} captures the asymmetry and is defined as:

$$I_{t-1} = \begin{cases} 1, & \text{if } \varepsilon_{t-1} < 0 \\ 0, & \text{otherwise} \end{cases}$$

The term $\gamma I_{t-1}\varepsilon_{t-1}^2$ allows the model to account for the increased sensitivity of volatility to negative shocks. When $\gamma > 0$, the model implies that negative returns increase future volatility more than positive returns, consistent with empirical observations in financial markets. For multidimensional GJR-GARCH model, the CCC-GJR-GARCH model was proposed by Bollerslev (1990) [17]. The model is as follows:

$$H_t = D_t R D_t \tag{2.5}$$

where H_t is conditional covariance matrix, D_t is diagonal matrix of conditional standard deviations, and R is constant correlation matrix.
According to the model given in equation (2.5), the step for estimating the conditional covariance matrix H_t are as follows:

1. **Estimating conditional volatility of each asset via univariate GJR-GARCH(1,1):** The conditional variances $h_{i,t}$ of asset i, $i = 1, 2, \ldots, n$ can be estimated by fitting model:

$$h_{i,t} = \omega_i + (\alpha_i + \gamma_i I_{i,t-1})\epsilon_{i,t-1}^2 + \beta_i h_{i,t-1} \tag{2.6}$$

where the indicator function $I_{i,t-1} = \begin{cases} 1 & \text{if } r_{i,t-1} < \mu_i \\ 0 & \text{otherwise} \end{cases}$

2. **Forming the diagonal matrix $\mathbf{D_t}$**: The diagonal entries of D_t is the conditional standard deviation obtain from step 1. Thus,

$$D_t = \begin{pmatrix} \sqrt{h_{1,t}} & 0 & \cdots & 0 \\ 0 & \sqrt{h_{2,t}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{h_{n,t}} \end{pmatrix}$$

3. **Estimating the constant correlation matrix R**: This correlation matrix can be computed by using historical data.

$$R = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n} & \rho_{2n} & \cdots & 1 \end{pmatrix}$$

4. **Construct the conditional covariance matrix $\mathbf{H}_t$** as in equation(2.5)

$$H_t = D_t R D_t$$

This decomposition separates the time-varying volatilities (captured by D_t) and the constant cross-asset correlations (captured by R), simplifying the multivariate modeling and ensuring positive semi-definiteness of H_t. For the empirical implementation, we use a rolling window approach to estimate the time-varying conditional volatility for each asset. Specifically, historical stock prices are first converted into log returns and segmented into non-overlapping windows. For each window, a univariate GJR-GARCH(1,1) model is fitted independently to each asset's return to extract conditional variances. These are then used to construct D_t, while the constant correlation matrix R is estimated from the daily returns in the same window. The resulting covariance matrix H_t is then used in a rolling mean-variance optimization framework. Within each window, the expected returns are estimated as the mean of log returns, and the portfolio weights ω_t are determined by solving the optimization problem in equation (2.1).

The optimized portfolio is then rebalanced at each time step using updated estimation of the expected return and the covariance matrix from GMM and GJR-GARCH models, allowing for dynamic allocation based on recent market conditions.

3 Main Results

This research investigates the performance of two optimal portfolios. The construction of the portfolios is based on the daily closing prices of four stocks: Apple Inc. (AAPL), Nike Inc.

(NKE), Alphabet Inc. (GOOGL), and Chevron Corp. (CVX). These stocks were chosen to represent different sectors of the U.S. equity market: technology, consumer discretionary, communication services, and energy, respectively. Such sectoral diversification helps mitigate unsystematic risk and provides a more realistic and robust testing ground for portfolio optimization models. The dataset spans from January 2014 to January 2024. By using expected return and covariance matrix for constructing the portfolios, the first portfolio is constructed by using the Gaussian Mixture model (GMM) while the second portfolio is constructed by using GJR-GARCH model. In order to dynamically adjust the portfolio weights over time responding to market conditions, the portfolios are rebalanced with rolling window size 10 days (about two weeks as Saunders [15]). In this study, the historical data of 200 days—approximately 80% of a year—are used to compute the expected return and covariance matrix for constructing portfolios. This window length is selected to ensure that there are enough data to estimate model parameters, without depending on outdated information that may no longer represent current market behavior. Consequently, the first rolling window starts at day 201 of the dataset (17th October 2014). The performance of the portfolios are measured by cumulative amount of money, Sharpe ratio and VaR@95% .

3.1 Expected Returns

This section presents the expected returns computed using two different methods: the GMM, and the historical average approach. Both methods rely on daily log-returns. In the GMM method, MSCI is used as an indicator variable to determine the regime (state) of the market at each point in time. The expected return from the GMM is computed by following equation (2.2), which incorporates regime-dependent statistics and the estimated transition probabilities between regimes.

The first rolling window numerical daily expected return computing by GMM as in equation (2.2) is presented. The estimated parameters of the initial window are:

$$p = 0.482412, \quad \mu_0 = \begin{bmatrix} 0.003135 \\ 0.005168 \\ 0.005140 \\ 0.001319 \end{bmatrix}, \quad \mu_1 = \begin{bmatrix} -0.001041 \\ -0.004354 \\ -0.005767 \\ -0.002579 \end{bmatrix}$$

where the $(i, 1)$ entries is the values of asset i, for $i = 1, 2, \ldots, 4$, In this situation, the asset 1, 2, 3, and 4 are AAPL, NKE, GOOGL, CVX, respectively. Substituting the values into the equation (2.2) gives the expected return

$$\mu = (1-p)\mu_0 + p\mu_1 = \begin{bmatrix} 0.001120 \\ 0.000574 \\ -0.000122 \\ -0.000561 \end{bmatrix}.$$

This calculation demonstrates how the expected return is estimated based on the GMM. The transition probabilities p, the means μ_0 and μ_1 as well as the expected return for each asset are computed in a similar manner in each rolling window.

The expected returns obtained from the GMM and the historical average approach are plotted in Figure 1. The historical expected return is calculated as the simple average of log-returns over historical data of 200 days, with the window moving forward every 10 days. This

window length and step size were selected to align with the data structure used in the GMM framework. Both approaches exhibit similar overall patterns and are slightly different in values. Notably, both methods capture significant declines in expected returns during the 2022–2023 period. These declines coincide with several macroeconomic events that adversely affected the stock market.

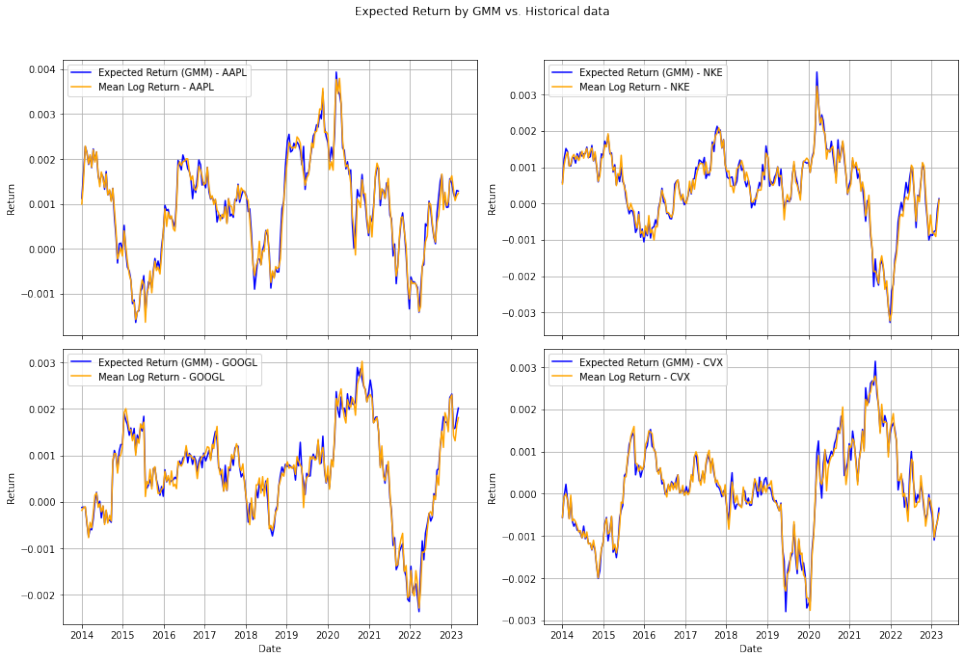


Figure 1: Comparison of expected return: (Left) expected by GMM vs. (Right) expected by historical data

3.2 Volatilities and Covariance Matrices

This section presents the volatility estimation obtained from two approaches: the GMM and the GJR-GARCH(1,1) models. To better understand the steps of these models, the following example demonstrates how conditional volatility is computed based on the model parameters and historical data.

First, we now demonstrate the computation of the conditional covariance matrix based on the GMM. The conditional covariance matrix for a 10-day rolling window is computed by combining the covariance matrices of both regimes, weighted by the transition probabilities, as given by equation (2.3). The GMM parameters estimated from the first rolling window using 200 days historical data are:

$$p = 0.482412, \quad q = 0.517588, \quad \mu_0 = \begin{bmatrix} 0.003135 \\ 0.005168 \\ 0.005140 \\ 0.001319 \end{bmatrix}, \quad \mu_1 = \begin{bmatrix} -0.001041 \\ -0.004354 \\ -0.005767 \\ -0.002579 \end{bmatrix}$$

$$\Sigma_0 = \begin{bmatrix} 0.000172 & 0.000037 & 0.000015 & 0.000001 \\ 0.000037 & 0.000206 & 0.000043 & 0.000025 \\ 0.000015 & 0.000043 & 0.000152 & -0.000003 \\ 0.000001 & 0.000025 & -0.000003 & 0.000071 \end{bmatrix},$$

$$\Sigma_1 = \begin{bmatrix} 0.000196 & 0.000027 & 0.000042 & 0.000030 \\ 0.000027 & 0.000137 & 0.000088 & 0.000046 \\ 0.000042 & 0.000088 & 0.000169 & 0.000057 \\ 0.000030 & 0.000046 & 0.000057 & 0.000098 \end{bmatrix}.$$

The covariance matrix is computed by using equation (2.3) as

$$\Sigma = \begin{bmatrix} 0.000188 & 0.000042 & 0.000040 & 0.000019 \\ 0.000042 & 0.000195 & 0.000091 & 0.000045 \\ 0.000040 & 0.000091 & 0.000190 & 0.000037 \\ 0.000019 & 0.000045 & 0.000037 & 0.000088 \end{bmatrix}.$$

Next, numerical results of asset volatilities from the first window via GJR-GARCH(1,1) model are presented. The conditional variance at time t of the first window data fitting by the model in equation (2.6) is

$$h_t^{AAPL} = 0.000071 + (0.042321 + 1.124651 \cdot I_{AAPL,t-1})\epsilon_{AAPL,t-1}^2 + 0.305045 \cdot h_{t-1}^{AAPL}$$

With this model, the estimated conditional standard deviation at the end of window becomes

$$\sqrt{h}^{\text{AAPL}} = 0.019252$$

From the fitted model and historical data, we got the conditional volatilities for the four assets (AAPL, NKE, GOOGL, CVX) are:

$$\sqrt{h}^{\text{AAPL}} = 0.019252, \quad \sqrt{h}^{\text{NKE}} = 0.010780, \quad \sqrt{h}^{\text{GOOGL}} = 0.018936, \quad \sqrt{h}^{\text{CVX}} = 0.013552$$

By estimating the covariance matrix via CCC-GJR-GARCH(1,1) model the conditional covariance matrix is given by

$$H = DRD$$

where R is correlation matrix and diagonal matrix $D = \text{diag}(\sqrt{h}^{AAPL}, \sqrt{h}^{NKE}, \sqrt{h}^{GOOGL}, \sqrt{h}^{CVX})$
For the initial window the diagonal volatility matrix is:

$$D = \begin{bmatrix} 0.019252 & 0 & 0 & 0 \\ 0 & 0.010780 & 0 & 0 \\ 0 & 0 & 0.018936 & 0 \\ 0 & 0 & 0 & 0.013552 \end{bmatrix}$$

and the constant correlation matrix R computed from 200 days historical data is estimated as:

$$R = \begin{bmatrix} 1.000000 & 0.217047 & 0.205454 & 0.148036 \\ 0.217047 & 1.000000 & 0.472166 & 0.342432 \\ 0.205454 & 0.472166 & 1.000000 & 0.284618 \\ 0.148036 & 0.342432 & 0.284618 & 1.000000 \end{bmatrix}.$$

Thus, conditional covariance matrix is as follows:

$$H = DRD = \begin{bmatrix} 0.000370 & 0.000045 & 0.000075 & 0.000039 \\ 0.000045 & 0.000116 & 0.000096 & 0.000050 \\ 0.000075 & 0.000096 & 0.000359 & 0.000073 \\ 0.000039 & 0.000050 & 0.000073 & 0.000184 \end{bmatrix}.$$

This results in a symmetric matrix capturing the time-varying variances and covariances among the assets. The conditional covariance matrices from the GMM and the GJR-GARCH models are used directly in the portfolio optimization process to balance return and risk dynamically at each time step. These illustrates the procedure for a single rolling window. The same process is repeated throughout the entire dataset. At each step, new GMM and GJR-GARCH models are fitted to updated return data, and a new conditional covariance matrices are computed.

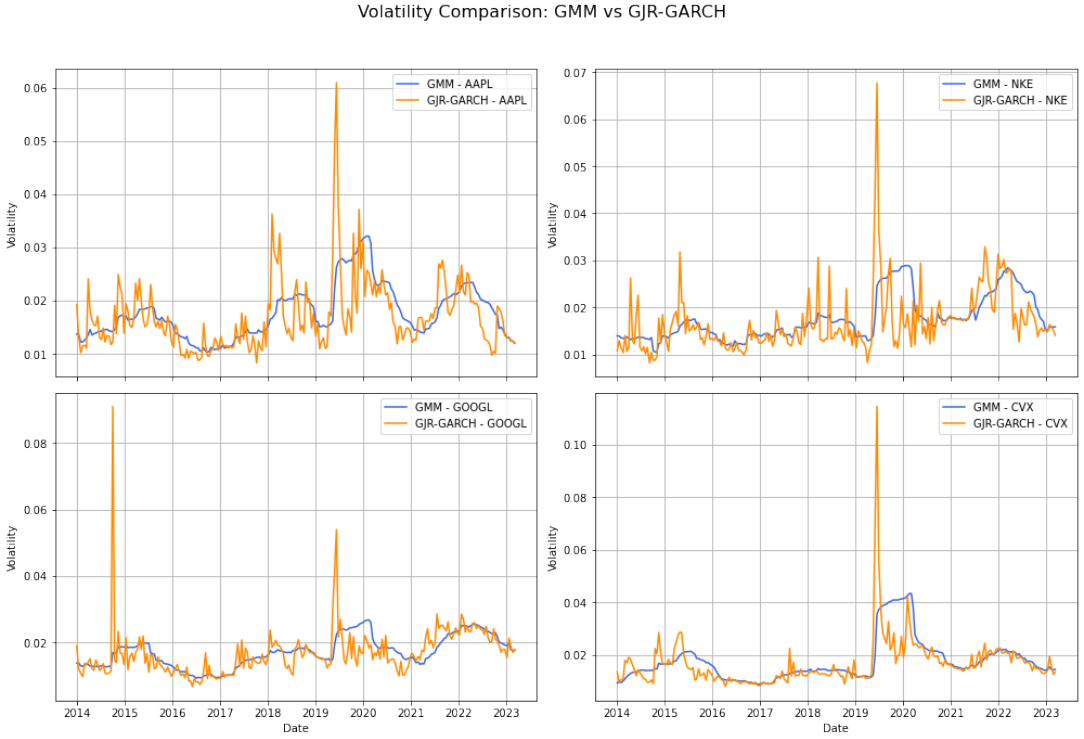


Figure 2: Comparison of volatilty: (Left) volatility by GMM vs. (Right) volatility by GJR-GARCH

The volatility estimation for each asset obtained from both models are shown in Figures 2. These figures compare the volatility behavior over time for each asset between the GMM and GJR-GARCH models. In comparison, Figure 2 presents the volatility dynamics from the GJR-GARCH(1,1,1) model, which accounts for past returns and volatility shocks. The GJR-GARCH model captures similar volatility pattern to GMM volatility, but it responds to the asymmetric shocks, which is different from the regime-based approach of GMM.

3.3 Results from Portfolio Optimization

In this section, we perform a rolling window portfolio optimization using the optimization problem given in equation(2.1). For each 10-day window, the expected returns and conditional covariance matrix are estimated, and portfolio weights are optimized according to each window parameters.

By using command "minimize" from scipy.optimize in python, the weights of each 10-days window are given. Figure 3 and Figure 4 show the portfolio weights obtained from the GMM and GJR-GARCH models, respectively. The results indicate that the GMM model tends to allocate most of the weight to a single asset in each period. On the other hand, the GJR-GARCH model produces more diversified portfolio weights across all assets, which reflects a more diversified portfolio and consequently reduces in risk exposure.

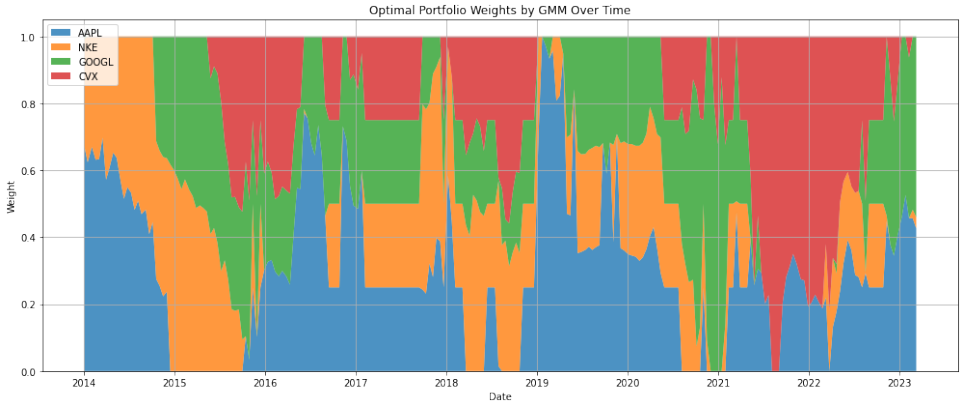


Figure 3: Rolling window portfolio weights using GMM model

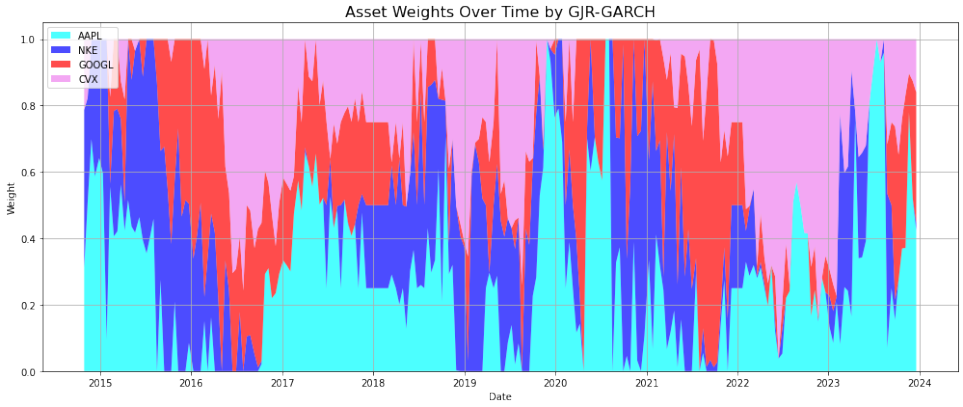


Figure 4: Rolling window portfolio weights using GJR model

To evaluate the performance of each modeling approach in a practical investment context, we implement a portfolio in various window framework. Using the estimated parameters from both models, portfolio weights are re-optimized within a rolling window, and the results of these two methods are estimated over time. Figure 5 compares the cumulative amount of money of the two portfolios. The experimental analysis, conducted from 2014 to 2024, start with an

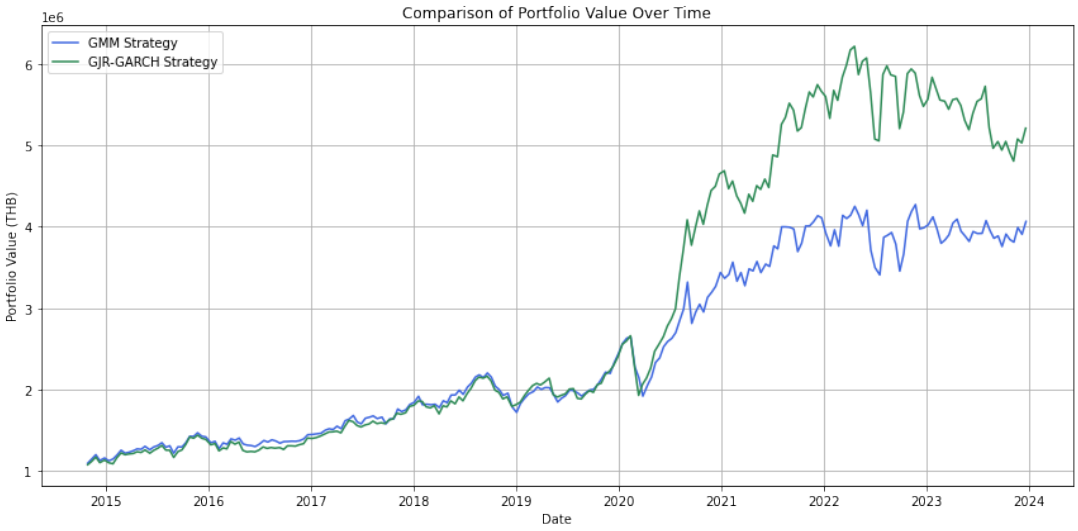


Figure 5: Comparison of portfolios value

initial investment of 1 million dollars. From the graph, the performance of both models are observed to be relatively similar in the first five years. However, in 2020, the portfolio under the GJR-GARCH model begin to outperform the GMM model. From 2020 onward, GJR-GARCH consistently delivered superior results, and by the end of the sample, the GJR-GARCH portfolio had increased to slightly above five million dollars, while the GMM portfolio reached approximately four million dollars, indicating the superior performance of the GJR-GARCH approach.

Table 1: Descriptive statistics and out-of-sample performance of minimum variance and maximum return portfolio with short sale for the econometric models under study from January 2014 to January 2024.

Performance measure	GMM	CCC-GJR-GARCH
Cumulative Amount of Growth (times the initial capital)	3.065854	4.211739
Standard Deviation	0.041234	0.041880
Sharpe Ratio	0.145374	0.169543
VaR@95%	-0.059059	-0.060270

From Table 1, the GJR-GARCH model delivers superior performance compared to the GMM approach in terms of cumulative growth. The excess of GJR-GARCH portfolio growth is approximately 4.21 times the initial capital, whereas that of the GMM portfolio reaches around 3.07. This suggests that the GJR-GARCH model is more effective in capturing profitable opportunities during the sample period. In terms of risk, both models exhibit similar levels of risk, with standard deviations of approximately 0.041. With the risk free rate of 0.000954 per 10 days, the GJR-GARCH model achieves a higher Sharpe ratio (0.1695) comparing to the GMM model (0.1454), indicating better risk-adjusted returns. Interestingly, the Value at Risk (VaR)

at the 95% confidence level for the GMM model at -0.0591 is slightly lower (i.e., less negative) compared to -0.0603 for GJR-GARCH.

Overall, while both models present comparable volatility, the GJR-GARCH model outperforms in terms of cumulative return and Sharpe ratio, reinforcing its strength as a more profitable and efficient approach to portfolio management under the examined period.

4 Conclusion and Discussion

This research focuses on portfolio allocation by employing econometric models and compare the performance of two prominent approaches: the GMM and the GJR-GARCH models. The study aims to constructing optimal investment portfolios, analyzing their returns, risks, and risk-adjusted performance over a ten-year period from January 2014 to December 2023. Key measures such as cumulative excess return, standard deviation, Sharpe ratio, and value-at-risk (VaR) at 0.95 confidence levels were utilized to assess and compare the portfolio results. The results revealed that both portfolios grew over time, but the GJR-GARCH portfolio began to consistently outperform the GMM portfolio from 2020 onward. This turning point can be attributed to the distinct mathematical properties of the GJR-GARCH model that captures asymmetric volatility—where negative shocks (e.g., market crashes or bad news) have a greater effect on future volatility than positive shocks of the same magnitude. During periods of market turbulence such as the COVID-19 pandemic in early 2020, this asymmetry becomes highly relevant.

Mathematically, GJR-GARCH includes an indicator function that increases the conditional variance when lagged negative returns are observed. This sensitivity allows the model to more accurately update the volatility estimation during stress periods. Consequently, the covariance matrix obtained from CCC-GJR-GARCH responds faster to market conditions, enabling the optimizer to allocate capital more efficiently, especially in volatile environments. On the other hand, the GMM framework captures regime switching between latent states with distinct return distributions, but it may react slower to abrupt shifts due to its reliance on state transition probabilities and smoothing over historical data. Thus, while GMM is useful for modeling structural changes over time, its slower reaction might have led to suboptimal portfolio weights during periods of rapid volatility change.

This difference in model responsiveness is reflected in the post-2020 trajectory: the GJR-GARCH portfolio captured the volatility spike and adapted allocations accordingly, resulting in better performance both in terms of cumulative value and risk-adjusted return as reflected in the Sharpe ratio and overall capital growth. This research contributes to investors for suggesting an outperform method used in portfolio allocation.

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and for their constructive suggestions, which have helped to improve the quality of this paper. This work was supported by the Development and Promotion of Science and Technology Talents Project (DPST), Thai government scholarship.

References

[1] Fama, E. F. (1965). *The behavior of stock-market prices.* Journal of Business, 38(1), 34-105.

[2] Markowitz, H. (1952). *Portfolio Selection*. Journal of Finance, 7(1), 77-91.

[3] Elton, E. J., & Gruber, M. J. (1997). *Modern Portfolio Theory and Investment Analysis*. John Wiley & Sons, Inc.

[4] Xu, D., & Gencay, R. (2003). *Gaussian Mixture Models in Financial Time Series*. Computational Statistics & Data Analysis, 43(1), 1-20.

[5] Wang, X., & Zhou, H. (2012). *Improving Asset Allocation with Gaussian Mixture Models: Application to Non-Gaussian Distributions*. Journal of Financial Econometrics, 10(3), 345-374.

[6] Bollerslev, T. (1986). *Generalized Autoregressive Conditional Heteroskedasticity*. Journal of Econometrics, 31(3), 307-327.

[7] Engle, R. F. (1982). *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation*. Econometrica, 50(4), 987-1007.

[8] Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). *On the Relationship Between the Expected Return and the Volatility of the Nominal Excess Return on Stocks*. Journal of Finance, 48(5), 1779-1801.

[9] R. Engle, *Dynamic Conditional Correlation – A Simple Class of Multivariate GARCH Models*, Journal of Business & Economic Statistics, 20(3), 2002, pp. 339-350.

[10] Arouri, M. E. H., Benkraiem, R., & Nguyen, D. K. (2011). *Dynamic Correlations and Volatility Spillovers Between International Stock Markets: Evidence from the CCC-GARCH and DCC-GARCH Models*. Journal of International Financial Markets, Institutions & Money, 21(4), 664-688.

[11] Conrad, J., Wahal, S., & Zhang, M. (2014). *Volatility and Correlation Modeling in the CCC-GJR-GARCH Framework*. Journal of Financial Economics, 114(2), 318-338.

[12] W. F. Sharpe, *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*, Journal of Finance, 19(3), 1964, pp. 425-442.

[13] X. Li, J. Lee, and Z. Zhan, *A multi-objective optimization model for portfolio selection*, European Journal of Operational Research, 123(1), 2000, pp. 1-14.

[14] C. Luo, L. Seco, L.-L. B. Wu, *Portfolio optimization in hedge funds by OGARCH and Markov switching model*, Journal of Management Science, 57(Part A), 2015, pp. 34-39.

[15] D. Saunders, L. Seco, C. Vogt, R. Zagst (2013). *A fund of hedge funds under regime switching*, J. Altern. Invest. 15 (4) 8–23.

[16] J.D. Hamilton, *A new approach to the economic analysis of nonstationary time series and the business cycle*, Econometrica 57 (1989) 357–384.

[17] T. Bollerslev, *Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model*, Rev. Econ. Stat. 72 (1990) 498–505.

[18] Y. Wang and T. Aste. (2023). *Dynamic portfolio optimization with inverse covariance clustering*, Expert Systems with Applications, vol. 213, p. 118739

การจำแนกระยะการเป็นโรคไตเรื้อรังโดยตัวแบบการถดถอยลอจิสติกพหุ

จริยาวัฒน์ ลี้พันธ์[1,*] และ วิภาวรรณ เล้าอรุณ[1,†]

[1]ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร 73000

บทคัดย่อ

งานวิจัยนี้มุ่งศึกษาปัจจัยเสี่ยงที่สำคัญต่อการจำแนกระยะการป่วยเป็นโรคไตเรื้อรัง ซึ่งแบ่งออกเป็น 3 ระยะ คือ Non CKD (1), Mild stage (2) และ Middle to end stage (3) ผู้วิจัยใช้การถดถอยลอจิสติกพหุ (Multinomial logistic regression) ภายใต้ตัวแบบลอจิตเทียบกลุ่มฐานในการวิเคราะห์ข้อมูล โดยกำหนดกลุ่ม Non CKD เป็นกลุ่มฐาน โดยตัวแบบประกอบด้วยอิทธิพลหลักจากตัวแปรอธิบาย 12 ตัว ได้แก่ อายุ เพศ ดัชนีมวลกาย การสูบบุหรี่ ประวัติครอบครัวเป็นโรคไต ค่าความดันขณะหัวใจบีบตัว ค่าความดันขณะหัวใจคลายตัว ระดับน้ำตาลในเลือด ระดับเซรั่มครีเอตินิน ปริมาณโปรตีนในปัสสาวะ ระดับความเข้มข้นเลือด และไขมันในเลือด โดยใช้ข้อมูลจากผู้ป่วย 1,659 คน ที่เผยแพร่ในเว็บไซต์ Kaggle ในการวิเคราะห์ผู้วิจัยได้แบ่งข้อมูลเป็น 2 ชุด คือ ชุดข้อมูลฝึกฝน สัดส่วน 80% ใช้ในการวิเคราะห์ตัวแบบ และชุดข้อมูลทดสอบ สัดส่วน 20% ผลการวิจัยพบว่า เพศ อายุ และระดับเซรั่มครีเอตินิน เป็นปัจจัยที่ส่งผลต่อการจำแนกระยะการเป็นโรคไตเรื้อรัง ทั้งนี้ผู้ที่มีอายุเพิ่มขึ้น และมีระดับเซรั่มครีเอตินินสูง และเป็นเพศหญิง จะมีความเสี่ยงต่อการเป็นโรคไตเรื้อรังในระยะที่สูงขึ้น ผลการศึกษาพบว่า ตัวแบบการถดถอยลอจิสติกพหุที่ประกอบด้วยปัจจัย อายุ เพศ และระดับเซรั่มครีเอตินิน ประสิทธิภาพของตัวแบบในการจำแนกระยะการป่วยเป็นโรคไตเรื้อรัง พิจารณาจากค่าความแม่น (Accuracy) ความเที่ยง (Precision) ความไว (Sensitivity) และคะแนน F1 (F1-Score) ที่คำนวณจากชุดข้อมูลทดสอบ โดยมีค่าความแม่น ร้อยละ 96.08 ความเที่ยง ร้อยละ 95.94 ความไว ร้อยละ 96.08 และคะแนน F1 ร้อยละ 95.95 ตามลำดับ

คำสำคัญ: โรคไตเรื้อรัง, การถดถอยลอจิสติกพหุ, ตัวแบบลอจิตเทียบกลุ่มฐาน

2020 MSC: 62P10

*ผู้นำเสนอ: (Leepun_j@su.ac.th)

†ผู้ประพันธ์บรรณกิจ (laoarun_w@su.ac.th)

1 บทนำ

ปัจจุบันโรคไตเรื้อรัง (Chronic kidney disease: CKD) เป็นปัญหาสาธารณสุขระดับโลกและกำลังมีแนวโน้มเพิ่มมากขึ้นในประเทศไทยซึ่งเป็นปัจจัยเสี่ยงสำคัญของการเกิดโรคเบาหวาน โรคหัวใจ โรคความดันโลหิตสูง เป็นต้น และมีการดำเนินโรคไปสู่โรคไตวายเรื้อรังระยะสุดท้าย (End-stage renal disease: ESRD) ที่ต้องได้รับการรักษาด้วยการบำบัดทดแทนไตหรือการปลูกถ่ายไต ประชากรส่วนใหญ่ไม่รู้ตัวว่าตนเองมีโรคไตเรื้อรังอยู่ อีกทั้งไม่ได้รับการวินิจฉัยจนกว่าจะถึงระยะลุกลามของโรค ด้วยเหตุผลที่ว่าตัวเลือกการรักษามีจำกัดและการพยากรณ์โรคไม่ดีเท่าที่ควร ส่งผลให้ผู้ป่วยโรคไตเรื้อรังระยะสุดท้ายมีความต้องการในการปลูกถ่ายไตและฟอกไตที่มีค่าใช้จ่ายค่อนข้างสูงและมักไม่สามารถเข้าถึงบริการการรักษาได้ ผลกระทบทางสังคมดังกล่าว จึงมีความสำคัญนำไปสู่การสูญเสียรายได้ คุณภาพชีวิตลดลง และเป็นการเพิ่มภาระของผู้ดูแล การจัดการกับภาระของโรคไตเรื้อรังจึงต้องเพิ่มความตระหนักให้แก่ประชาชนในการตรวจหาและการรักษาแต่เนิ่น ๆ [1]

ได้มีงานวิจัยเกี่ยวกับการวิจัยที่เกี่ยวข้องกับโรคไตเรื้อรังเป็นจำนวนมาก ได้แก่ ปี 2560 ณิชกานต์ วงษ์ประกอบ และ ลัฆวี ปิยะบัณฑิตกุล [2] ได้ทำการศึกษาปัจจัยทำนายการเกิดโรคไตเรื้อรังในผู้ป่วยเบาหวานและความดันโลหิตสูง และวิเคราะห์ความสัมพันธ์ระหว่างปัจจัยด้านบุคคล การเจ็บป่วยและพฤติกรรมทางสุขภาพกับการเกิดโรคไตเรื้อรังในผู้ป่วยเบาหวานและความดันโลหิตสูง กลุ่มตัวอย่าง คือ ผู้ป่วยที่เป็นทั้งโรคเบาหวานและความดันโลหิตสูงและ ได้รับการวินิจฉัยว่าเป็นโรคไตเรื้อรัง จำนวน 160 คน เปรียบเทียบกับกลุ่มควบคุมซึ่งเป็นผู้ป่วยโรคเบาหวานและโรคความดันโลหิตสูง มาไม่น้อยกว่า 6 เดือน แต่ไม่เป็นโรคไตเรื้อรัง จำนวน 320 คน ศึกษาข้อมูลตั้งแต่เดือนพฤศจิกายน พ.ศ. 2559 ถึงเดือน เมษายน พ.ศ. 2560 ใช้แบบสัมภาษณ์ ที่ประกอบด้วยข้อมูลทั่วไป ประวัติการเจ็บป่วย พฤติกรรม ทางสุขภาพ และผลตรวจทางห้องปฏิบัติการจากเวชระเบียน ผลการวิเคราะห์ปัจจัยทำนายการเกิดโรคไตเรื้อรังโดยใช้ การวิเคราะห์การถดถอยลอจิสติก (Logistic regression analysis) โดยกำหนดตัวแปรตอบสนองเป็น 2 กลุ่ม คือ เป็นและไม่เป็นโรคไตเรื้อรัง อธิบายผลตัวแปรอธิบายแต่ละตัวที่เกี่ยวพันกับตัวแปรตอบสนองอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 ผ่านค่า Adjusted odds ratio และช่วงความเชื่อมั่น 95% ที่ไม่คลุม 1.00 เมื่อควบคุมอิทธิพลของตัวแปรรบกวน พบว่า ปัจจัยทำนายการเกิดโรคไตเรื้อรังในผู้ป่วยเบาหวาน และความดันโลหิตสูง อย่างมีนัยสำคัญทางสถิติ ได้แก่ การควบคุมระดับน้ำตาลสะสม (HbA1C) การมีโรคประจำตัวร่วม การควบคุมระดับความดันโลหิต การรับประทานอาหารที่มีไขมันสูง Diastolic blood pressure สูง Triglyceride สูง อายุระหว่าง 70-79 ปี ระดับ LDL-Cholesterol สูง และเพศหญิง จากผลการศึกษาทำให้ทราบถึงปัจจัยทำนายการเกิดโรคไตเรื้อรังในผู้ป่วยเบาหวานและ ความดันโลหิตสูง และสามารถนำไปใช้ในการคัดกรองโรคไตเรื้อรังในผู้ป่วยเบาหวาน และความดันโลหิตสูง ในชุมชนได้ตั้งแต่ระยะเริ่มต้น

ในปี 2566 จรูญ บุญธกานนท์ [3] ศึกษาปัจจัยที่มีความสัมพันธ์กับการลดลงของอัตราการกรองของไตอย่างรวดเร็วในผู้ป่วยโรคไตเรื้อรังระยะ 3-4 ที่เข้ารับการรักษาในคลินิกโรคไตเรื้อรัง โรงพยาบาลเกษตรสมบูรณ์ จังหวัดชัยภูมิ เก็บรวบรวมข้อมูลตัวแปรที่ต้องการศึกษา โดยการดึงข้อมูลจากฐานข้อมูลอิเล็กทรอนิกส์

(โปรแกรมสำเร็จรูป HOSxP) จำนวน 1,184 ราย ได้ผู้เข้าร่วมวิจัยที่ผ่านเกณฑ์คัดเข้าและเกณฑ์คัดออก จำนวน 434 ราย ผลการศึกษาพบว่า ตัวอย่างส่วนใหญ่เป็นเพศหญิง (ร้อยละ 63.4) อายุเฉลี่ย 71 ปี มีโรคประจำตัวเป็นโรคเบาหวานหรือโรคความดันโลหิตสูง (ร้อยละ 94.47) จากการติดตามในระยะเวลา 1 ปี พบการเปลี่ยนแปลงระยะของโรคไตคงที่ ร้อยละ 71.89 มีอัตราการกรองของไตลดลงอย่างรวดเร็ว eGFR ≥ 4 mL/min/1.73 m^2/year ร้อยละ 68.43 ผลการวิเคราะห์ความสัมพันธ์โดยพิจารณาจากอัตราส่วนออดส์ (Odds ratio: OR) จากการวิเคราะห์การถดถอยลอจิสติกพหุ (Multinomial logistic regression) พบว่าระดับน้ำตาลในเลือดหลังอดอาหาร (Fasting blood suga: FBS) สัมพันธ์กับการลดลงของอัตราการกรองของไตอย่างรวดเร็วอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 (OR = 2.95, 95% CI ของ OR เท่ากับ (1.23, 7.07), p-value = 0.015) นอกจากนี้ยังพบว่ามีปัจจัยอื่นที่อาจจะส่งผลต่อการลดลงของอัตราการกรองของไต อย่างไม่มีนัยสำคัญทางสถิติ ได้แก่ เพศ อายุ ดัชนีมวลกาย โรคเบาหวาน โรคความดันโลหิตสูง ค่าความดันขณะหัวใจบีบตัว (SBP) ระดับน้ำตาลสะสมในเลือด (HbA1c), ระดับ Cholesterol, ระดับไตรกลีเซอร์ไรด์ (Triglyceride) ระดับ HDL, LDL ปริมาณโปรตีนในปัสสาวะ (Urine proteinuria) และยารักษาโรคเรื้อรังบางชนิด

นอกจากนี้ ในปัจจุบันยังพบงานวิจัยที่ได้นำเอาเทคนิคการเรียนรู้ของเครื่อง (Machine learning: ML) มาใช้ในการคัดกรองโรคต่าง ๆ รวมถึงโรคไตเรื้อรัง ได้แก่ งานวิจัยของ Elias dritsas and maria trigka ในปี 2022 [4] ศึกษาโดยใช้ข้อมูลจากผู้ป่วย 400 ราย ที่เผยแพร่ในเว็บไซต์ kaggle ประกอบด้วยตัวแปรคุณลักษณะ (Feature) 13 ตัว และตัวแปรตอบสนอง 1 ตัว (จำแนกได้ 2 กลุ่ม คือ เป็น/ไม่เป็นโรคไตเรื้อรัง) โดยมีวัตถุประสงค์เพื่อประยุกต์ใช้ได้เทคนิคการเรียนรู้ของเครื่องจำนวน 15 ตัวแบบ ในการสร้างเครื่องมือที่มีประสิทธิภาพในการทำนายการเกิดโรคไตเรื้อรัง (เป็น/ไม่เป็นโรคไตเรื้อรัง) โดยมีการจัดการปัญหาข้อมูลไม่สมดุลระหว่างกลุ่ม หลังจากนั้นจะมีการจัดอันดับและวิเคราะห์คุณลักษณะ (Features) และสุดท้ายจะมีการฝึกฝนและประเมินประสิทธิภาพของตัวแบบจากตัวชี้วัดที่หลากหลาย ผลการศึกษาพบว่า Rotation forest (RotF) มีประสิทธิภาพในการทำนายมากที่สุด โดยมีค่า Area under the curve (AUC) ที่ 100%, Precision, Recall, F-Measure และ Accuracy เท่ากับ 99.2%

ปี 2023 M.M. El Sherbiny, E. Abdelhalim และ H.El. D. Mostafa [5] ได้ทำการศึกษาโดยมุ่งเน้นไปที่การใช้เทคนิคการเรียนรู้ของเครื่องในการพยากรณ์และตรวจจับโรคไตเรื้อรัง ในแต่ละระยะ (5 ระยะ) โดยแหล่งข้อมูลที่ใช้ในงานวิจัยนี้มาจากคลังข้อมูล Machine learning มหาวิทยาลัยแคลิฟอร์เนีย เออร์ไวน์ (University of California Irvine ML Repository) ซึ่งรวบรวมมาจากโรงพยาบาลแห่งหนึ่งในประเทศอินเดียเป็นระยะเวลาสองเดือน ชุดข้อมูลประกอบด้วยตัวอย่างจำนวน 400 รายการ ที่มีตัวแปรคุณลักษณะทั้งหมด 25 ตัวแปร ซึ่งประกอบด้วยตัวแปรคุณลักษณะเชิงปริมาณ 14 ตัวแปร ตัวแปรคุณลักษณะแบบจำแนกประเภท (Categorical) 10 ตัวแปร และอีก 1 ตัวแปรเป็นตัวแปรเป้าหมาย (Target variable) ในการจำแนกประเภท ซึ่งระบุว่าเป็นผู้ป่วยโรคไตเรื้อรัง (CKD) หรือไม่เป็นโรคไตเรื้อรัง (Non CKD) ชุดข้อมูลนี้ประกอบด้วยข้อมูลของบุคคลที่เป็นโรคไตเรื้อรังจำนวน 250 ราย และบุคคลที่ไม่เป็นโรคไตเรื้อรังจำนวน 150 ราย จำนวนตัวอย่างที่สมบูรณ์ (ไม่มีค่าสูญหาย) ในชุดข้อมูลนี้มีทั้งหมด 158

รายการ ตัวแบบที่นำเสนอประกอบด้วยการประยุกต์ใช้ตัวแบบการจำแนกประเภทที่ใช้การเรียนรู้ของเครื่องทั้งหมด 7 ตัวแบบ ได้แก่ การถดถอยลอจิสติก (Logistic regression: LR), ซัพพอร์ตเวกเตอร์แมชชีน (Support vector machine: SVM), ต้นไม้ตัดสินใจ (Decision tree: DT), แรนดอมฟอเรสต์ (Random forest: RF), นาอีฟเบส์ (Naïve bayes: NB), เอ็กซ์ตรีมกราเดียนต์บูสติ้ง (Extreme gradient boosting: XGBoost) และอะแดพทีฟบูสติ้ง (Adaptive boosting: AdaBoost) มีการทดลองหลายรูปแบบในงานวิจัยนี้ ซึ่งรวมถึงเทคนิคการทดแทนค่าข้อมูลสูญหาย (Data imputation) และวิธีการคัดเลือกคุณลักษณะ (Feature selection) ประสิทธิภาพของตัวแบบต่าง ๆ ถูกประเมินโดยใช้ตัวชี้วัด 4 ตัวได้แก่ ความแม่น (Accuracy) ความเที่ยง (Precision) คะแนน F-1 (F1-Score) และค่าการเรียกคืน (Recall) ผลการศึกษาพบว่า XGBoost และ Random Forest ให้ผลลัพธ์ที่ดีกว่าเทคนิคอื่น ๆ

ในปัจจุบันถึงแม้จะมีเทคนิคการเรียนรู้ของเครื่องหลายเทคนิคเข้ามามีบทบาทสำคัญในงานวิเคราะห์จำแนกประเภทสำหรับงานวิจัยทางการแพทย์ ผลลัพธ์ที่ได้จากการวิเคราะห์อาจทราบถึงปัจจัยที่ส่งผลต่อการจำแนกประเภท แต่มักจะไม่พบสารสนเทศในแง่ความสัมพันธ์เชิงสาเหตุมากนัก แต่สารสนเทศดังกล่าวสามารถพบได้จากการวิเคราะห์การถดถอยลอจิสติก ซึ่งสามารถตีความจากค่าประมาณสัมประสิทธิ์ (Coefficients) หรืออาจตีความผ่านค่าอัตราส่วนออดส์ (Odds ratio) เช่น บอกได้ว่าตัวแปรอธิบายแต่ละตัวแปรมีผลกระทบเชิงบวกหรือเชิงลบ ต่อความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ นอกจากนี้ยังมีกระบวนการทดสอบความเหมาะสมของตัวแบบเพื่อให้มั่นใจว่าสามารถนำตัวแบบที่วิเคราะห์ได้ไปใช้ในการทำนายกลุ่มหรือจำแนกประเภทต่อไปได้

ดังนั้น ในงานวิจัยนี้จึงสนใจการนำตัวแบบลอจิตเทียบกลุ่มฐาน (Baseline category logit model) มาวิเคราะห์ปัจจัยเสี่ยงที่สำคัญต่อการป่วยเป็นโรคไตเรื้อรังของผู้ป่วย โดยกำหนดตัวแปรตอบสนอง คือ ระยะในการเป็นโรคไตเรื้อรัง แบ่งระยะในการเป็นโรคไตเรื้อรัง ออกเป็น 3 ระยะ คือ ไม่เป็นโรคไตเรื้อรัง (Non CKD) (1), เป็นโรคไตเรื้อรังระยะเริ่มต้น (Mild stage) (2) และ เป็นโรคไตเรื้อรังระยะปานกลางถึงระยะสุดท้าย (Middle to end stage) (3) โดยใช้ข้อมูลของผู้ป่วยจำนวน 1,659 คน จากเว็บไซต์ https://www.kaggle.com /datasets/rabieelkharoua/chronic-kidney-disease-dataset-analysis?resource=download จากฐานข้อมูล Kaggle และตัวแปรอธิบายที่เป็นปัจจัยที่คาดว่าจะเกี่ยวพันกับการเป็นโรคไตเรื้อรัง จำนวน 12 ตัวแปร รายละเอียดจะกล่าวถึงในหัวข้อ 3 นอกจากนี้ยังสามารถนำตัวแบบที่ได้จากการวิเคราะห์ไปใช้ประโยชน์ในการคำนวณความเสี่ยงต่อการเป็นโรคไตเรื้อรังในแต่ละระยะได้ จะช่วยให้แพทย์วางแผนการรักษาได้อย่างเหมาะสม โดยในระยะต้นจะเป็นการควบคุมปัจจัยเสี่ยง ในระยะปลายจะเป็นการเตรียมตัวสำหรับการล้างไตหรือปลูกถ่ายไต ทั้งนี้โรคไตเรื้อรังเป็นโรคที่รักษาได้แต่ไม่หายขาด ดังนั้นการวินิจฉัยที่รวดเร็วและแม่นยำจะช่วยลดค่าใช้จ่ายในการรักษาระยะสุดท้ายและยังช่วยลดภาระของระบบสาธารณสุขได้

2 ความรู้พื้นฐาน

2.1 โรคไตเรื้อรัง (Chronic kidney disease: CKD)

โรคไตเรื้อรัง หรือเรียกสั้น ๆ ว่า CKD หมายถึง ภาวะที่มีการทำลายของเนื้อไตซึ่งอาศัยหลักฐานที่บ่งบอกทางอ้อม (Maker) หรือจากการตรวจชิ้นเนื้อไตโดยตรงหรือการประเมินอัตราการกรองของไต (Glomerular filtration rate: GFR) ที่น้อยกว่า 60 ml/min/1.73 m^2 มาเป็นเวลานานกว่า 3 เดือน โดย Maker ของ Kidney damage คือ การตรวจปัสสาวะพบ Proteinuria มี Urine sediment หรือความผิดปกติจากการตรวจทางรังสีวิทยา (Imaging study) ของไต โดยค่าปกติของ GFR ในวัยผู้ใหญ่มีค่าประมาณ 120 - 130 ml/min/1.73 m^2 เมื่ออายุมากกว่า 30 ปีขึ้นไป ระดับ GFR จะลดลงตามอายุในอัตรา 1 ml/min/1.73 m^2 ต่อปี ดังนั้นระดับ GFR ที่ลดลงต่ำกว่า 60 ml/min/1.73 m^2 จึงแสดงถึงระดับการทำงานของไตลดลงแล้วประมาณครึ่งหนึ่งของภาวะปกติและเป็นระดับที่เริ่มเกิดภาวะแทรกซ้อนต่าง ๆ ของ CKD ได้มากขึ้น [1]

ในปัจจุบันมีการตื่นตัวในการวินิจฉัยโรคไตเรื้อรังตั้งแต่เนิ่น ๆ เพื่อให้สามารถเริ่มใช้มาตรการชะลอ การเสื่อมหน้าที่ไตให้เร็วขึ้น จึงได้แบ่งโรคไตเรื้อรังออกเป็น 5 ระยะ ตามระดับการเปลี่ยนแปลงของอัตรา การกรองของไต (eGFR) แสดงดังตารางที่ 1 โดยค่าดังกล่าวสามารถคำนวณโดยสูตรคำนวณที่ขึ้นกับ เพศ อายุ และระดับเซรั่มครีเอตินิน [6]

ตารางที่ 1: การจำแนกระยะของโรคไตเรื้อรังในปัจจุบัน โดยพิจารณาจากอัตราการกรองของไต

ระยะของไตเรื้อรัง	อัตราการกรองของไตใน 1 นาที eGFR (ml/min/1.73 m^2)	คำนิยาม
ระยะที่ 1	มากกว่าหรือเท่ากับ 90	ภาวะไตทำงานปกติ อัตราการกรองปกติ
ระยะที่ 2	60 - 89	ภาวะไตผิดปกติ มีการรั่วของโปรตีนในปัสสาวะ แต่อัตราการกรองลดลงเล็กน้อย
ระยะที่ 3a	45 - 59	อัตราการกรองลดลงเล็กน้อยถึงปานกลาง
ระยะที่ 3b	30 - 44	อัตราการกรองลดลงปานกลางถึงรุนแรง
ระยะที่ 4	15 - 29	อัตราการกรองลดลงมากหรืออย่างรุนแรง
ระยะที่ 5	น้อยกว่า 15	ไตระยะสุดท้ายหรือระยะไตวาย

ที่มา: ประเสริฐ ธนกิจจารุ, สกานต์ บุนนาค, วรางคณา พิชัยวงศ์. โรคไตเรื้อรัง (Chronic Kidney Disease: CKD)

จากระยะของโรคไตเรื้อรังข้างต้น ผู้วิจัยเลือกที่จะศึกษาระยะของโรคไตเรื้อรังเพื่อเป็นการคัดกรองเบื้องต้น ดังนี้ Recode CKD stage ตามระดับ eGFR โดย Stage 2 กำหนดว่าเป็น Mild stage of CKD, Stage 3 - 5 กำหนดว่าเป็น Middle to end stage of CKD และในระดับ eGFR ที่มากกว่าหรือเท่ากับ 90 กำหนดว่าเป็น Non CKD ดังตารางที่ 2

ตารางที่ 2: Recode CKD stage ตามระดับ eGFR

eGFR (ml/min/1.73 m^2)	Description	Stage
มากกว่าหรือเท่ากับ 90	ภาวะปกติ	Non CKD
60 - 89	ภาวะไตผิดปกติ มีการรั่วของโปรตีนในปัสสาวะ แต่อัตราการกรองยังคงปกติถึงลดลงเล็กน้อย	Mild stage
น้อยกว่าหรือเท่ากับ 59	อัตราการกรองลดลงปานกลาง มากหรืออย่างรุนแรง, ไตระยะสุดท้ายหรือระยะไตวาย	Middle to end stage

2.2 การวิเคราะห์การถดถอยลอจิสติกพหุ (Multinomial logistic regression analysis)

การวิเคราะห์การถดถอยลอจิสติกพหุ เป็นวิธีการทางสถิติ ที่ใช้วิเคราะห์ความเกี่ยวพันระหว่างตัวแปรอธิบายกับตัวแปรตอบสนองที่มีมากกว่า 2 กลุ่มและมาตรวัดเป็นแบบมาตรนามบัญญัติ [7] โดยอาศัยตัวแบบ Multinomial logit

ในการศึกษาครั้งนี้ผู้วิจัยใช้ตัวแบบลอจิตเทียบกลุ่มฐาน (Baseline category logit) เป็นตัวแบบที่ใช้กำหนดความสัมพันธ์ระหว่างตัวแปรอธิบายและตัวแปรตอบสนอง สมมติให้ตัวแปรตอบสนอง (Y) จำแนกได้ 3 กลุ่ม และกำหนดกลุ่มที่ 1 เป็นกลุ่มฐาน และตัวแปรอธิบาย (X) p ตัว ตัวแบบลอจิตเทียบกลุ่มฐานกำหนดได้ดังตัวแบบ (1)

$$log\left(\frac{\pi_j}{\pi_1}\right) = \alpha_j + \beta_{j_1}X_1 + \beta_{j_2}X_2 + \ldots + \beta_{j_p}X_p,\ j = 2,3 \qquad (1)$$

โดยที่ $\pi_j = P(Y = j),\ j = 2,\ 3$ แทน ความน่าจะเป็นที่หน่วยสังเกตจะมีค่าตอบสนองอยู่ในกลุ่มที่ j

π_1 แทน ความน่าจะเป็นที่หน่วยสังเกตจะมีค่าตอบสนองอยู่ในกลุ่มที่ 1 (กลุ่มฐาน)

α_j, β_{j_i} แทน ค่าคงที่การถดถอย และสัมประสิทธิ์การถดถอยลอจิสติกสำหรับตัวแปรอธิบาย X_i สำหรับกลุ่มของตัวแปรตอบสนองกลุ่มที่ j ตามลำดับ เมื่อ $j = 2,3$ และ i=1, 2..., p

$X_1, X_2, \ldots, X_p$ แทน ตัวแปรอธิบายหรือปัจจัยที่คาดว่าจะส่งผลต่อการทำนายกลุ่มของตัวแปรตอบสนอง กรณีที่ตัวแปรตอบสนองจำแนกได้ 3 กลุ่ม จากตัวแบบ (1) จะได้ตัวแบบ $log\left(\frac{\pi_2}{\pi_1}\right)$ และ $log\left(\frac{\pi_3}{\pi_1}\right)$ ซึ่งแสดงถึง Log-odds ที่ค่าตอบสนองอยู่ในกลุ่มที่ j [8] หมายถึง log ของอัตราส่วนความน่าจะเป็นที่หน่วยสังเกตใด ๆ จะมีค่าตอบสนองอยู่ในกลุ่มที่ 2, 3 ต่อความน่าจะเป็นที่หน่วยสังเกตใด ๆ จะมีค่าตอบสนองอยู่ในกลุ่มฐาน ตามลำดับ

นอกจากนี้ผู้วิจัยยังมีการตีความหมายของค่าประมาณพารามิเตอร์สำหรับตัวแปรอธิบายแต่ละตัวในตัวแบบการถดถอยลอจิสติกพหุ จากค่าอัตราส่วนออดส์ (Odds ratio: OR) โดยที่ $OR = exp\ (B_{j_i})$ เมื่อ B_{j_i} แทนค่าประมาณของ β_{j_i}

จากตัวแบบ (1) สามารถทดสอบว่าตัวแปรอธิบาย (โดยรวม) มีเกี่ยวพันกับตัวแปรตอบสนองหรือไม่ ภายใต้ระดับนัยสำคัญ α (ในที่นี้กำหนด เท่ากับ 0.05) โดยมีสมมติฐานดังนี้

$H_0: \beta_{j_i} = \ldots = \beta_{j_p} = 0$; j = 2, 3, i = 1, 2, ..., p

$H_1: \beta_{j_i}$ อย่างน้อย 1 ตัว ที่ไม่เท่ากับ 0

ตัวสถิติทดสอบ คือ Likelihood ratio test (LRT) = (-2 log L (Null model)) - (-2 log L (Fitted model))

เกณฑ์การตัดสินใจปฏิเสธสมมติฐานว่าง จะปฏิเสธสมมติฐาน H_0 เมื่อ LRT > $\chi^2_{\alpha,df}$ หรือ p-value < α

หากตัวแปรอธิบายอย่างน้อย 1 ตัวเกี่ยวพันกันกับตัวแปรตอบสนอง จะมีการทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอยลอจิสติกแต่ละตัว ภายใต้ระดับนัยสำคัญ 0.05 โดยมีสมมติฐานดังนี้

$H_0: \beta_{j_i} = 0$; j = 2, 3, i = 1, 2, ..., p

$H_1: \beta_{j_i} \neq 0$

ตัวสถิติทดสอบ คือ $Z = \frac{B_{j_i}}{S.E.(B_{j_i})}$ มีการแจกแจงปรกติมาตรฐาน ภายใต้สมมติฐานว่างจริง

เกณฑ์การตัดสินใจปฏิเสธสมมติฐานว่าง จะปฏิเสธสมมติฐาน H_0 เมื่อ $Z > Z_{\frac{\alpha}{2}}$ หรือ $Z < -Z_{\frac{\alpha}{2}}$ หรือ p-value < α หรืออาจใช้ตัวสถิติวัลด์ (Wald statistic) กำหนดโดย $Z^2 = Wald = \left(\frac{B_{j_i}}{SE(B_{j_i})}\right)^2$ มีการแจกแจงไคสแควร์ที่องศาเสรีเท่ากับ 1 ภายใต้สมมติฐานว่างจริง

เกณฑ์ในการตัดสินใจปฏิเสธสมมติฐานว่าง จะปฏิเสธสมมติฐาน H_0 เมื่อ $Wald > \chi^2_{\alpha;1}$

เมื่อพบว่าตัวแปรมีนัยสำคัญ จะตรวจสอบความเหมาะสมของตัวแบบ โดยใช้ตัวสถิติทดสอบอัตราส่วนภาวะน่าจะเป็น G^2 หรือ ผลต่าง Deviance คือ [-2 log L (Fitted model)] - [-2 log L (Saturated model)] ที่มีการแจกแจงแบบไคสแควร์ด้วยองศาเสรี เท่ากับ ผลต่างของจำนวนพารามิเตอร์จาก 2 ตัวแบบ เมื่อ log L แทนค่าของ Log-likelihood [9] โดยหากค่าผลต่างของ Deviance มีค่าน้อย หรือ p-value สูงกว่าระดับนัยสำคัญที่กำหนด กล่าวได้ว่าตัวแบบ Fitted model นั้นมีความเหมาะสมหรือตัวแบบสามารถอธิบายข้อมูลได้ดี สามารถนำไปใช้ในการพยากรณ์ความน่าจะเป็นที่หน่วยตัวอย่างจะตกอยู่ในกลุ่มของตัวแปรตอบสนองกลุ่มใด ๆ ได้ จนนำไปสู่การพยากรณ์กลุ่มได้ต่อไป

ในกรณีที่มีตัวแบบที่เหมาะสมมากกว่า 1 ตัวแบบ ผู้วิจัยจะทำการคัดเลือกตัวแบบที่เหมาะสมที่สุด จากค่าสถิติ Akaike's Information criterion (AIC) โดยตัวแบบที่ให้ค่า AIC น้อยที่สุด ถือว่าเป็นตัวแบบที่เหมาะสมที่สุด

2.3 การประเมินประสิทธิภาพของตัวแบบ

ในการประเมินประสิทธิภาพของตัวแบบในการจำแนกระยะของโรคไตเรื้อรัง จะพิจารณาจากตัวชี้วัดต่าง ๆ ได้แก่ ความแม่น (Accuracy) ความเที่ยง (Precision) และความไว (Sensitivity) ในการจำแนก ที่คำนวณจากการนำตัวแบบที่วิเคราะห์ได้จากชุดข้อมูลฝึกฝนไปใช้กับชุดข้อมูลทดสอบ ทั้งนี้ตัวชี้วัดต่าง ๆ มีที่มาจากการคำนวณค่าจาก Confusion matrix ในงานวิจัยนี้พิจารณาปัญหาการจำแนกเป็น 3 คลาส จึงคำนวณตัววัดประสิทธิภาพและคำนวณ

ค่า TN, TP, FN, FP ของแต่ละคลาสต่างหาก เพื่อง่ายต่อการคำนวณจึงได้แยกตามสีที่แตกต่างกัน โครงสร้างอย่างง่ายของ Confusion Matrix ที่มี 3 คลาส [10] แสดงดังตารางที่ 3

ตารางที่ 3: ตาราง Confusion matrix กรณีตัวแปรตอบสนองจำแนกเป็น 3 คลาส

Confusion Matrix		Prediction		
		Class 1	Class 2	Class 3
Actual	Class 1	A = TP	B = FN	C = FN
	Class 2	D = FP	E = TN	F =TN
	Class 3	G = FP	H = TN	I = TN

เมื่อพิจารณา Class 1 กำหนดให้ แทน True positives แทน True negatives แทน Misclassified cases

รายละเอียดจากตาราง Confusion matrix ตามหลักการจำแนกแบบทวิภาค อธิบายได้ดังนี้

True positive (TP) คือ จำนวนตัวอย่างของสิ่งที่ทำนาย ตรงกับสิ่งที่เกิดขึ้นจริง ในกรณีทำนายว่าจริง และสิ่งที่เกิดขึ้นคือจริง

True negative (TN) คือ จำนวนตัวอย่างของสิ่งที่ทำนายตรงกับสิ่งที่เกิดขึ้น ในกรณีทำนายว่าไม่จริง และสิ่งที่เกิดขึ้นคือไม่จริง

False positive (FP) คือ จำนวนตัวอย่างของสิ่งที่ทำนายไม่ตรงกับสิ่งที่เกิดขึ้น คือ ทำนายว่าจริง แต่สิ่งที่เกิดขึ้นคือไม่จริง

False negative (FN) คือ จำนวนตัวอย่างของสิ่งที่ทำนายไม่ตรงกับที่เกิดขึ้นจริง คือ ทำนายว่าไม่จริง แต่สิ่งที่เกิดขึ้นคือจริง

ตัวชี้วัดประสิทธิภาพของตัวแบบในการจำแนกที่ใช้ในงานวิจัยนี้ พิจารณาจาก

1. **ความแม่น (Accuracy)** ของการจำแนก คำนวณจาก

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

หรือ

$$Accuracy = \frac{(A+E+I)}{(A+E+I+D+G+B+C+F+H)}$$

โดยทั่วไปแล้วความแม่นในการจำแนกประเภทยังไม่เพียงพอที่จะตัดสินใจว่าตัวแบบนั้นมีประสิทธิภาพหรือดีพอที่จะใช้ในการทำนายได้อย่างถูกต้อง ปัญหาของความแม่น คือ ไม่สามารถแยกระหว่างการจำแนกประเภทที่หลากหลายได้ จึงคำนวณตัววัดประสิทธิภาพตัวอื่น ๆ เช่น ความเที่ยง (Precision), ความไว (Sensitivity หรืออาจเรียกว่า Recall), ความจำเพาะ (Specificity), คะแนน F1 (F1-score)

2. **ความเที่ยง (Precision)** คือ อัตราส่วนระหว่างผลลัพธ์การทำนายและผลลัพธ์จริงเป็นจริง (TP) ต่อจำนวนตัวอย่างที่ผลลัพธ์ทำนายเป็นจริงทั้งหมด กำหนดโดย

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}}$$

หรือ $$Precision = \frac{\mathrm{A}}{\mathrm{A+D+G}}$$

3. **ความไว (Sensitivity หรือ Recall)** คือ ความสามารถของตัวแบบในการจำแนกกลุ่มนั้น ๆ ได้ถูกต้อง กำหนดโดย

$$Sensitivity\ (\text{หรือ}\ Recall) = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$$

หรือ $$Sensitivity\ (\text{หรือ}\ Recall) = \frac{\mathrm{A}}{\mathrm{A+B+C}}$$

ในงานวิจัยทางการแพทย์ เช่น การจำแนกโรค ค่า Sensitivity สูงจะสำคัญมาก เพราะหมายถึงโอกาสตรวจพบโรคถูกต้อง

4. **คะแนน F1 (F1-score)** คือ ค่าที่ใช้วัดความสมดุลระหว่างความเที่ยง และความไว โดยเฉพาะในกรณีที่มีความไม่สมดุลของข้อมูล (Class imbalance) คำนวณจากค่าเฉลี่ยฮาร์โมนิกระหว่างค่า Precision และ Recall กำหนดโดย

$$F1 - Score = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall}\right)$$

3 ระเบียบวิธีวิจัย

3.1 ข้อมูลที่ใช้ในการศึกษา

ในการศึกษางานวิจัยครั้งนี้ศึกษาข้อมูลผู้ป่วยโรคไตเรื้อรัง จำนวน 1,659 คน จากชุดข้อมูลที่เผยแพร่บนเว็บไซต์ https://www.kaggle.com/datasets/rabieelkharoua/chronic-kidney-disease-dataset-analysisresource =download โดยประกอบด้วยอิทธิพลหลักจากตัวแปรอธิบาย 12 ตัว ได้แก่ อายุ เพศ ดัชนีมวลกาย การสูบบุหรี่ ประวัติครอบครัวเป็นโรคไต ค่าความดันขณะหัวใจบีบตัว ค่าความดันขณะหัวใจ คลายตัว ระดับน้ำตาลในเลือด ระดับเซรั่มครีเอตินิน ปริมาณโปรตีนในปัสสาวะ ระดับความเข้มข้นเลือด และระดับไขมันในเลือด จำแนกระยะการป่วยเป็นโรคไตเรื้อรัง เป็น 3 ระยะ คือ ไม่เป็นโรคไตเรื้อรัง (Non CKD) (1), เป็นโรคไตเรื้อรังระยะเริ่มต้น (Mild stage) (2) และเป็นโรคไตเรื้อรังระยะปานกลางถึงระยะสุดท้าย (Middle to end stage) (3)

3.2 วิธีดำเนินงานวิจัย

1. การทำความเข้าใจข้อมูล (Data understanding)

จากการศึกษาเอกสารที่เกี่ยวข้องและปัจจัยเสี่ยงของโรคไตเรื้อรัง ประกอบด้วยตัวแปรอธิบาย 12 ตัวแปร ได้แก่ อายุ เพศ ดัชนีมวลกาย การสูบบุหรี่ ประวัติครอบครัวเป็นโรคไต ค่าความดันขณะหัวใจบีบตัว ค่าความดันขณะหัวใจคลายตัว ระดับน้ำตาลในเลือด ระดับเซรั่มครีเอตินิน ปริมาณโปรตีนในปัสสาวะ ระดับความ เข้มข้นเลือด และระดับไขมันในเลือด และตัวแปรตอบสนอง คือ ระยะของโรคไตเรื้อรัง แสดงรายละเอียดตัวแปร ดังตารางที่ 4

ตารางที่ 4: รายละเอียดของตัวแปรที่ใช้ในงานวิจัย

ตัวแปร	ชื่อตัวแปร	คุณลักษณะ	ค่า/หน่วยตัวแปร
Y	Stage	ระยะของโรคไตเรื้อรัง	1 = ไม่เป็นโรคไตเรื้อรัง (Non CKD) 2 = เป็นโรคไตเรื้อรังระยะเริ่มต้น (Mild stage) 3 = เป็นโรคไตเรื้อรังระยะปานกลางถึงระยะสุดท้าย (Middle to end stage)
X_1	Age	อายุ	ปี (Year)
X_2	Gender	เพศ	0 = เพศชาย (Male) (กลุ่มฐาน); 1 = เพศหญิง (Female)
X_3	BMI	ดัชนีมวลกาย	กิโลกรัม / เมตร^2 (kg/m^2)
X_4	Smoking	การสูบบุหรี่	0 = ไม่สูบบุหรี่ (กลุ่มฐาน); 1 = สูบบุหรี่
X_5	FamilyhistoryKidneydisease	ประวัติครอบครัวเป็นโรคไต	0 = ไม่มี (กลุ่มฐาน); 1 = มี
X_6	SystolicBP	ค่าความดันขณะหัวใจบีบตัว	มิลลิเมตรปรอท (mmHg.)
X_7	DiastolicBP	ค่าความดันขณะหัวใจคลายตัว	มิลลิเมตรปรอท (mmHg.)
X_8	HbA1c	ระดับน้ำตาลสะสมในเลือด	มิลลิกรัมเปอร์เซ็นต์ (mg%)
X_9	SerumCreatinine	ระดับเซรั่มครีเอตินิน	มิลลิกรัมต่อเดซิลิตร (mg/dL)
X_{10}	ProteinInUrine	ปริมาณโปรตีนในปัสสาวะ	มิลลิกรัมต่อกรัม (mg/g)
X_{11}	HemoglobinLevels	ระดับความเข้มข้นเลือด	มิลลิกรัมต่อเดซิลิตร (mg/dL)
X_{12}	CholesterolLDL	ระดับไขมันในเลือด	มิลลิกรัมต่อเดซิลิตร (mg/dL)

2. การจัดเตรียมข้อมูล (Data preparation)

โดยการทำความสะอาดข้อมูล ได้แก่ การจัดการกับข้อมูลบางส่วนที่มีความซ้ำซ้อน การตรวจสอบค่าผิดปรกติ การประมาณค่าสูญหาย (ถ้ามี) รวมถึงการวิเคราะห์ข้อมูลเบื้องต้นสำหรับทุกตัวแปรด้วยค่าสถิติ หรือแผนภาพที่เหมาะสมกับมาตรวัดของตัวแปร

3. การแบ่งข้อมูล (Data splitting)

ข้อมูลจะถูกแบ่งออกเป็น 2 ชุดคือ ชุดข้อมูลฝึกฝน (Training set) จำนวน 1,327 คน คิดเป็นร้อยละ 80 สำหรับสร้างตัวแบบและชุดข้อมูลทดสอบ (Test set) จำนวน 332 คน คิดเป็นร้อยละ 20 สำหรับทดสอบตัวแบบและใช้คำนวณตัวชี้วัดประสิทธิภาพของตัวแบบ

4. จากชุดข้อมูลฝึกฝน ทำการสร้างตัวแบบการจำแนกระยะของโรคไตเรื้อรัง (Modeling) โดยตัวแบบลอจิสติกพหุ แบบเทียบกลุ่มฐาน โดยกำหนดให้ Stage = 1 (ไม่เป็นโรคไตเรื้อรัง (Non CKD)) เป็นกลุ่มฐาน มีรายละเอียดดังนี้

4.1 เขียนตัวแบบแสดงความสัมพันธ์ระหว่างตัวแปรอธิบายทั้งหมดและตัวแปรตอบสนอง

4.2 ทดสอบว่าตัวแปรอธิบาย (โดยรวม) มีความเกี่ยวพันกับตัวแปรตอบสนองหรือไม่ภายใต้ระดับนัยสำคัญ 0.05

4.3 หากผลการทดสอบในข้อ 4.2 พบว่ามีตัวแปรอธิบายอย่างน้อย 1 ตัวเกี่ยวพันกับตัวแปรตอบสนอง จะมีการทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอยลอจิสติกแต่ละตัวต่อไป

4.4 ทดสอบความเหมาะสมของตัวแบบเต็มรูปที่ประกอบด้วยตัวแปรอธิบายทั้งหมด

4.5 นำตัวแปรอธิบายเฉพาะที่ผลการทดสอบในขั้น 4.3 มีความเกี่ยวพันกับตัวแปรตอบสนองอย่างมีนัยสำคัญ มาสร้างตัวแบบอีกครั้งหนึ่ง

4.6 ดำเนินการตามขั้นตอน 4.2 และ 4.3

4.7 ทดสอบความเหมาะสมของตัวแบบลดรูป ที่ประกอบด้วยตัวแปรอธิบายที่มีนัยสำคัญเท่านั้น

4.8 เลือกตัวแบบที่เหมาะสมที่สุดภายใต้เกณฑ์ค่า AIC ต่ำที่สุด

5. นำตัวแบบที่เหมาะสมที่สุดจากขั้นตอน 4.8 ไปทดสอบกับชุดข้อมูลทดสอบ และคำนวณค่าประสิทธิภาพของตัวแบบในการจำแนกระยะการเป็นโรคไตเรื้อรัง

6. สรุปผลและอภิปรายผล

3.3 เครื่องมือที่ใช้ในการวิจัย

1) โปรแกรม Microsoft excel เพื่อใช้ในการคัดกรองข้อมูล

2) โปรแกรม R studio เพื่อใช้ในการสร้างตัวแบบ

4 ผลการวิจัย

4.1 ผลการวิเคราะห์สถิติเชิงพรรณนา

จากชุดข้อมูลของผู้ป่วยจำนวน 1,659 คน พบว่า ร้อยละ 84.27 เป็นผู้ป่วยโรคไตเรื้อรังระยะปานกลางถึงระยะสุดท้าย และผู้ที่ป่วยเป็นโรคไตระยะปานกลาง และไม่เป็นโรคไตเรื้อรังมีสัดส่วนใกล้เคียงกัน คือ ร้อยละ 8.74 และ 6.99 ตามลำดับ ผลการวิเคราะห์สถิติเชิงพรรณนาของตัวแปรอธิบายเชิงคุณภาพและตัวแปรอธิบายเชิงปริมาณจำแนกตามระยะโรคไตเรื้อรังพร้อมแสดงค่าเกณฑ์มาตรฐานของแต่ละตัวแปร แสดงดังตารางที่ 5 และ ตารางที่ 6 ตามลำดับ โดยพบว่า เพศหญิงมีสัดส่วนการป่วยเป็นโรคไตเรื้อรังระยะปานกลางถึงระยะสุดท้ายสูงกว่าเพศชาย (ร้อยละ 87.25) ในกลุ่มผู้สูบบุหรี่และไม่สูบบุหรี่สัดส่วนการป่วยเป็นโรคไตเรื้อรังในแต่ละระยะใกล้เคียงกัน และผู้ที่มีประวัติคนในครอบครัวเคยป่วยเป็นโนคไตเรื้อรัง จะมีสัดส่วนการป่วยเป็นโรคไตเรื้อรังระยะปานกลางถึงระยะสุดท้ายสูงกว่าผู้ที่ไม่มีประวัติคนในครอบครัวเคยป่วยด้วยโรคนี้ (ร้อยละ 87.61) และจะพบว่าอายุเฉลี่ยของผู้ป่วยเป็นโรคไตเรื้อรังจะสูงกว่าผู้ที่ไม่เป็นโรคไตเรื้อรัง และระดับเซรั่มครีเอตินินซึ่งเป็นค่าที่ใช้วัดอัตราการกรองของไตโดยเฉลี่ย พบว่ามีค่าเฉลี่ยที่สูงกว่าค่ามาตรฐาน (0.5-1.2 mg/dL โดยรวม และ 0.74-1.35 mg/dL สำหรับผู้ชาย และ 0.59-1.04 mg/dL สำหรับผู้หญิง) และสูงขึ้นตามระยะของโรค นอกจากนี้ในส่วนของค่าเฉลี่ยของดัชนีมวลกาย ค่าความดันขณะหัวใจบีบตัว ค่าความดันขณะหัวใจคลายตัว และปริมาณโปรตีนในปัสสาวะ จะพบว่ามีค่าสูงในกลุ่มผู้ป่วยโรคไตระยะปานกลาง แต่สูงกว่าอีก 2 กลุ่มไม่มากนัก โดยผู้ป่วยทุกกลุ่มจะมีค่าเฉลี่ยของดัชนีมวลกาย และปริมาณโปรตีนในปัสสาวะ สูงกว่าค่ามาตรฐาน สำหรับค่าเฉลี่ยของระดับน้ำตาลในเลือด ในกลุ่มผู้ไม่เป็นโรคไตเรื้อรังสูงกว่าอีก 2 กลุ่มไม่มากนัก ทั้งนี้ยังพบว่าในผู้ป่วยทุกกลุ่มจะมีค่าเฉลี่ยของระดับน้ำตาลในเลือดสูงกว่าเกณฑ์มาตรฐาน

ตารางที่ 5: จำนวนและร้อยละของผู้ป่วยโรคไตเรื้อรังจำแนกตามระยะและตัวแปรอธิบายเชิงคุณภาพ

ปัจจัย	ระยะโรคไตเรื้อรัง							
	Non CKD		Mild stage		Middle to end stage		รวม	
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
รวม	**116**	**6.99**	**145**	**8.74**	**1398**	**84.27**	**1659**	**100.00**
Gender								
0 = Male	78	9.70	74	9.20	652	81.10	804	100.00
1 = Female	38	4.44	71	8.31	746	87.25	855	100.00
Smoking								
0 = No	83	7.08	101	8.61	989	84.31	1173	100.00
1 = Yes	33	6.79	44	9.05	409	84.16	486	100.00
FamilyHistoryKidneyDisease								
0 = No	104	7.30	128	8.98	1193	83.72	1425	100.00
1 = Yes	12	5.13	17	7.26	205	87.61	234	100.00

ตารางที่ 6: ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของปัจจัยเสี่ยงเชิงปริมาณของผู้ป่วยโรคไตเรื้อรังระยะต่าง ๆ

ปัจจัย	ระยะโรคไตเรื้อรัง				เกณฑ์มาตรฐาน
	Non CKD	Mild stage	Middle to end stage	รวม	
	mean ± S.D.	mean ± S.D.	mean ± S.D.	mean ± S.D.	
Age	46.28 ± 18.40	56.74 ± 21.32	54.88 ±20.50	54.44 ± 20.55	
BMI	26.68 ± 6.75	28.13 ± 7.29	27.65 ± 7.33	27.62 ± 7.29	18.5 – 22.9 kg/m^2
SystolicBP	133.81 ± 26.52	135.34 ± 24.26	134.34 ± 25.87	134.39 ± 25.77	120 -140 mmHg
DiastolicBP	90.49 ± 18.85	90.37 ± 17.70	89.11 ± 17.19	89.31 ± 17.35	80 - 90 mmHg
HbA1c	7.06 ± 1.67	6.89 ± 1.79	6.99 ± 1.73	6.98 ± 1.73	< 6.0 mg%
SerumCreatinine	0.71 ± 0.15	1.00 ± 0.22	3.10 ± 1.12	2.75 ± 1.32	0.5 – 1.2 mg/dL
ProteinInUrine	2.50 ± 1.41	2.53 ± 1.51	2.49 ± 1.44	2.49 ± 1.45	< 1.5 mg/g
HemoglobinLev	13.72 ± 2.31	14.00 ± 2.35	13.94± 2.34	13.93 ± 2.34	12 – 17.4 g/dL
CholesterolLDL	122.62 ± 41.82	126.49 ± 44.63	125.10 ± 45.53	125.04 ± 42.65	<100 – 129 mg/dL

4.2 ผลการวิเคราะห์ตัวแบบ

ในการวิเคราะห์ส่วนนี้ผู้วิจัยใช้ข้อมูลจากชุดข้อมูลฝึกฝนในการสร้างตัวแบบ โดยทำการสร้างตัวแบบใน 2 ตัวแบบย่อย คือ ตัวแบบเต็มรูป (Full model) ประกอบด้วยตัวแปรอธิบาย 12 ตัวแปร และ ตัวแบบลดรูป (Reduced model) ที่ประกอบด้วยตัวแปรอธิบายที่มีความเกี่ยวพันอย่างมีนัยสำคัญทางสถิติกับความน่าจะเป็นของการเกิดโรคไตเรื้อรังระยะต่าง ๆ แสดงผลการวิเคราะห์ดังตารางที่ 7 และผลการทดสอบความเหมาะสมของแต่ละตัวแบบโดยค่าผลต่างของ Deviance ดังตารางที่ 8 ผลการทดสอบพบว่าตัวแบบที่เหมาะสมที่จะใช้ในการจำแนกระยะโรคไตเรื้อรัง คือ ตัวแบบลดรูปที่ประกอบด้วยตัวแปรอธิบาย 3 ตัวแปร โดยค่าประมาณสัมประสิทธิ์และการทดสอบความมีนัยสำคัญของแต่ละตัวแปร และค่าอัตราส่วนออดส์ แสดงดังตารางที่ 9

ตารางที่ 7: ผลการทดสอบความมีนัยสำคัญของตัวแปรอธิบายโดยรวม โดยสถิติอัตราส่วนภาวะน่าจะเป็น

ตัวแบบ	Residual deviance	d.f.	log L	LRT	d.f.	p-value	จำนวนตัวแปรอธิบายที่มีนัยสำคัญ
Null model	1396.719	2652	-698.3596				
Full model	189.8797	2628	-94.9398	1206.8	24	< 2.2X10^{-16}	3
Reduced model	198.3166	2646	-99.1583	1198.4	6	< 2.2X10^{-16}	3

ตารางที่ 8: ผลการทดสอบความเหมาะสมของตัวแบบ และค่า AIC

ตัวแบบ	Diff. deviance	d.f.	p-value	AIC
Full model	189.8797	2628	1.0000	241.8797
Reduced model	198.3166	2646	1.0000	214.3166

ตารางที่ 9: ค่าประมาณสัมประสิทธิ์ของตัวแบบถดถอยลอจิสติกพหุ (ตัวแบบลดรูป) ผลการทดสอบความมีนัยสำคัญ และค่าประมาณอัตราส่วนออดส์

ตัวแปร	B	S.E.(B)	z value	p-value	OR = exp(B)
Intercept (2)	-12.0029	1.6391	-7.323	$2.43\text{x}10^{-13}$ ***	$6.1267\text{x}10^{-6}$
Intercept (3)	-26.9659	2.5196	-10.702	$<2.00\text{x}10^{-16}$ ***	$1.9448\text{x}10^{-12}$
Age (2)	0.0727	0.0125	5.796	$6.80\text{x}10^{-9}$ ***	1.0754
Age (3)	0.1325	0.0171	7.755	$8.82\text{x}10^{-15}$ ***	1.1416
Gender1 (2)	2.0576	0.4712	5.322	$1.03\text{x}10^{-7}$ ***	12.2754
Gender1 (3)	4.5359	0.6416	7.070	$1.55\text{x}10^{-12}$ ***	93.3063
SerumCreatinine (2)	8.3722	1.1953	7.004	$2.48\text{x}10^{-12}$ ***	4325.256
SerumCreatinine (3)	17.2799	1.5898	10.869	$<2.00\text{x}10^{-16}$ ***	31955260

*** ผลการทดสอบ พบว่า ค่าพารามิเตอร์แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติ ที่ระดับนัยสำคัญ 0.001

หมายเหตุ $Gender1 = \begin{cases} 1; \text{สำหรับเพศหญิง} \\ 0; \text{สำหรับเพศชาย} \end{cases}$

โดยปัจจัยที่เกี่ยวพันกับความน่าจะเป็นในการเกิดโรคไตเรื้อรังระยะต่าง ๆ อย่างมีนัยสำคัญทางสถิติ ที่ระดับนัยสำคัญ 0.05 ได้แก่ อายุ เพศ และระดับเซรั่มครีเอตินิน

จากผลลัพธ์ในตารางที่ 9 จะได้ตัวแบบถดถอยลอจิสติกพหุที่ประมาณได้ (Fitted model) 2 ตัวแบบย่อยดังนี้

$$log\left(\frac{\hat{\pi}_2}{\hat{\pi}_1}\right) = -12.0029 + 0.0727 * Age + 2.0576 * Gender1 + 8.3722 * SerumCreatinine \quad (2)$$

$$log\left(\frac{\hat{\pi}_3}{\hat{\pi}_1}\right) = -26.9659 + 0.1325 * Age + 4.5359 * Gender1 + 17.2799 * SerumCreatinine \quad (3)$$

จากตัวแบบ (2) สามารถอธิบายความหมายอิทธิพลของแต่ละปัจจัยต่อการเป็นโรคไตเรื้อรังในระยะเริ่มต้นเทียบกับการไม่เป็นโรคไตเรื้อรัง (เมื่อควบคุมให้อิทธิพลของตัวแปรอื่น ๆ คงที่) ได้ดังนี้

หากผู้ป่วยมีอายุต่างกัน 1 ปี ผู้ที่มีอายุมากกว่าจะมีความเสี่ยงในการเป็นโรคไตเรื้อรังในระยะเริ่มต้นสูงกว่าประมาณ ร้อยละ 7.54 เมื่อเทียบกับการไม่เป็นโรคไตเรื้อรัง และผู้หญิงจะมีความเสี่ยงในการเป็นโรคไตเรื้อรังในระยะเริ่มต้นสูงกว่าผู้ชายคิดเป็น 12.2754 เท่า เมื่อเทียบกับการไม่เป็นโรคไตเรื้อรัง นอกจากนี้ยังพบว่า ผู้ป่วยที่มีระดับ

เซรั่มครีเอตินินสูงขึ้น 1 mg/dL จะมีความเสี่ยงในการเป็นโรคไตเรื้อรังในระยะเริ่มต้นสูงขึ้นคิดเป็น 4325.256 เท่า เมื่อเทียบกับการไม่เป็นโรคไตเรื้อรัง

และจากตัวแบบ (3) สามารถอธิบายความหมายอิทธิพลของแต่ละปัจจัยต่อการเป็นโรคไตเรื้อรังในระยะปานกลางถึงระยะสุดท้าย เทียบกับการไม่เป็นโรคไตเรื้อรัง (เมื่อควบคุมให้อิทธิพลของตัวแปรอื่น ๆ คงที่) ได้ดังนี้

หากผู้ป่วยมีอายุต่างกัน 1 ปี ผู้ที่มีอายุมากกว่าจะมีความเสี่ยงในการเป็นโรคไตเรื้อรังในระยะปานกลางถึงระยะสุดท้ายสูงกว่าประมาณ ร้อยละ 14.16 เมื่อเทียบกับการไม่เป็นโรคไตเรื้อรัง และผู้หญิงจะมีความเสี่ยงในการเป็นโรคไตเรื้อรังในระยะดังกล่าวสูงกว่าผู้ชายคิดเป็น 93.3063 เท่า เมื่อเทียบกับการไม่เป็นโรคไตเรื้อรัง นอกจากนี้ยังพบว่า ผู้ป่วยที่มีระดับเซรั่มครีเอตินินสูงขึ้น 1 mg/dL จะมีความเสี่ยงในการเป็นโรคไตเรื้อรังในระยะปานกลางถึงระยะสุดท้ายสูงขึ้นคิดเป็น 31955260 เท่า เมื่อเทียบกับการไม่เป็นโรคไตเรื้อรัง

4.3 ผลการประเมินประสิทธิภาพของตัวแบบ

ผู้วิจัยได้นำตัวแบบการถดถอยลอจิสติกพหุที่ประกอบด้วยปัจจัย อายุ เพศ และระดับเซรั่มครีเอตินิน ที่วิเคราะห์ได้จากชุดข้อมูลฝึกฝน ไปใช้ในการจำแนกระยะของการเป็นโรคไตเรื้อรังกับชุดข้อมูลทดสอบ ได้ผลการทำนายหรือการจำแนก (Prediction) เทียบกับค่าจริง (Actual) แสดงในตาราง Confusion matrix ดังตารางที่ 10 และได้ค่าความแม่น (Accuracy) และค่าของตัวชี้วัดประสิทธิภาพของตัวแบบแสดงดังตารางที่ 11

ตารางที่ 10: Confusion matrix การจำแนกระยะโรคไตเรื้อรัง โดยตัวแบบลดรูป คำนวณจากชุดข้อมูลทดสอบ

Testing		Prediction		
		Stage 1	Stage 2	Stage 3
Actual	Stage 1	17	6	0
	Stage 2	2	32	5
	Stage 3	0	0	270

ตารางที่ 11: ค่าของตัวชี้วัดประสิทธิภาพของตัวแบบลดรูป (ตัวแปรอธิบาย 3 ตัว) คำนวณจากชุดข้อมูลทดสอบ

	ระยะโรคไตเรื้อรัง			ค่าเฉลี่ยถ่วงน้ำหนัก
	Non CKD	Mild stage	Middle to end stage	
Accuracy				**0.9608**
Precision	0.8947	0.8421	0.9818	**0.9594**
Sensitivity	0.7391	0.8205	1.0000	**0.9608**
F1-score	0.8095	0.8312	0.9908	**0.9595**

จากตารางที่ 11 จะได้ว่า ตัวแบบการถดถอยลอจิสติกพหุที่ประกอบด้วยปัจจัย อายุ เพศ และระดับเซรั่มครีเอตินิน สามารถนำไปใช้จำแนกระยะการป่วยเป็นโรคไตได้ถูกต้องร้อยละ 96.08 มีความเที่ยงในการทำนายระยะการป่วยเป็นโรคไตแต่ละระยะ โดยเฉลี่ยร้อยละ 95.94 และมีความไวในการทำนายการเป็นโรคไตเรื้อรังระยะปานกลางถึงระยะสุดท้ายได้มากที่สุด คิดเป็นร้อยละ 100.00 และเมื่อพิจารณาความสามารถโดยรวมระหว่างความเที่ยงและความไวจากคะแนน F1 พบว่า ตัวแบบนี้มีความสามารถจำแนกระยะการเป็นโรคไตระยะปานกลางถึงระยะสุดท้ายได้สูงสุด

5 สรุปผลและอภิปรายผล

ในการศึกษาครั้งนี้ ผู้วิจัยมุ่งเน้นที่จะวิเคราะห์หาปัจจัยที่ส่งผลต่อการจำแนกระยะการเป็นโรคไตเรื้อรัง 3 ระยะ ด้วยการวิเคราะห์การถดถอยลอจิสติกพหุ โดยใช้ข้อมูลที่เผยแพร่บนเว็บไซต์ https://www.kaggle.com/datasets/rabieelkharoua/chronic-kidney-disease-dataset-analysisresource =download เพื่อค้นหาตัวแบบที่แสดงความเกี่ยวพันระหว่างความน่าจะเป็นของการเป็นโรคไตเรื้อรังแต่ละระยะกับปัจจัยต่าง ๆ โดยศึกษาเฉพาะอิทธิพลหลัก (Main effect) รวมถึงการอธิบายอิทธิพลของแต่ละปัจจัยผ่านค่าอัตราส่วนออดส์ ทั้งนี้ผู้วิจัยยังได้นำเสนอผลการตรวจสอบความเหมาะสมของตัวแบบก่อนที่จะนำไปใช้ในการทำนายระยะของการเป็นโรคไตเรื้อรัง ผลการวิจัยพบว่าตัวแบบที่เหมาะสมกับข้อมูลชุดนี้ คือ ตัวแบบที่ประกอบด้วยปัจจัยหลัก 3 ตัวที่เกี่ยวพันกับความน่าจะเป็นของการเป็นโรคไตเรื้อรังแต่ละระยะอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 จากตัวแบบอธิบายได้ว่า อายุที่เพิ่มขึ้น ระดับเซรั่มครีเอตินินที่สูง รวมถึงเพศหญิง เป็นปัจจัยเสี่ยงต่อการเป็นโรคไตเรื้อรังในระยะที่สูงขึ้น ซึ่งถือเป็นปัจจัยที่เกี่ยวข้องกับอัตราการกรองของไตโดยเฉลี่ย (eGFR) ทั้งนี้ในงานวิจัยนี้ ไม่ได้มีการกำหนดเกณฑ์การคัดเข้าของผู้ป่วย ได้แก่ การกำหนดเฉพาะโรคประจำตัวของผู้ป่วย หรือระยะของโรคไตเรื้อรังที่สนใจศึกษา จึงทำให้ปัจจัยที่ได้จากงานวิจัยนี้อาจมีความแตกต่างจากงานวิจัยของ ณิชกานต์ วงษ์ประกอบ และ ลัฆวี ปิยะบัณฑิตกุล (2560) ที่ทำการศึกษาปัจจัยที่ส่งผลต่อการจำแนกโรคไตเรื้อรัง (เป็นโรคไตเรื้อรัง / ไม่เป็นโรคไตเรื้อรัง) เฉพาะในกลุ่มผู้ป่วยเบาหวาน และความดันโลหิตสูง โดยปัจจัยที่ส่งผลต่อการจำแนกโรคไตเรื้อรัง คือ การควบคุมระดับน้ำตาลสะสม (HbA1C) การมีโรคประจำตัวร่วม การควบคุมระดับความดันโลหิต การรับประทานอาหารที่มีไขมันสูง Diastolic blood pressure สูง Triglyceride สูง อายุระหว่าง 70-79 ปี ระดับ LDL-Cholesterol สูง และเพศหญิง รวมถึงงานวิจัยของจรูญ บุญธกานนท์ (2566) ที่ทำการศึกษาปัจจัยที่มีความสัมพันธ์กับการลดลงของอัตราการกรองของไตอย่างรวดเร็วในผู้ป่วยโรคไตเรื้อรังระยะ 3-4 พบว่ามีเพียงระดับน้ำตาลในเลือดหลังอดอาหาร (Fasting blood sugar: FBS) ที่สัมพันธ์กับการลดลงของอัตราการกรองของไตอย่างรวดเร็ว

ในส่วนของประสิทธิภาพของตัวแบบในการทำนายระยะของโรคไตจากตัวชี้วัด 4 ตัว คือ ค่าความแม่น ความเที่ยง ความไว และคะแนน F1 พบว่า จากงานวิจัยนี้ ค่าเฉลี่ยของตัวชี้วัดทุกตัวมีค่าสูงเกินกว่าร้อยละ 95.00 ซึ่งถือว่ามีประสิทธิภาพดีมากเมื่อเทียบกับความไม่ซับซ้อนของตัวแบบ ตัวแบบที่ใช้ในงานของ Elias Dritsas และ Maria

Trigka (2022) รวมถึง M.M. El Sherbiny และคณะ (2023) ที่ให้ความแม่นยำสูงกว่าร้อยละ 99 โดยเฉพาะเทคนิค Rotation forest และ XGBoost แต่แลกมากับความซับซ้อนและการตีความที่ยากขึ้น

นอกจากนี้จากผลการวิเคราะห์จะเห็นว่าสัดส่วนของข้อมูลในโรคไตเรื้อรังแต่ละระยะ มีความแตกต่างกันค่อนข้างมาก ในกระบวนการวิเคราะห์ครั้งนี้ไม่ได้มีการแก้ปัญหาข้อมูลไม่สมดุล (Imbalanced dataset) ก่อนการสร้างตัวแบบ ผลการวิเคราะห์ได้ตัวแบบที่ประกอบด้วย 3 ปัจจัยหลัก ซึ่งถือเป็นตัวแบบที่มีความซับซ้อนค่อนข้างน้อย ง่ายต่อการตีความ และให้ความสามารถในการจำแนกค่อนข้างสูง โดยผลการวิจัยนี้อาจเป็นประโยชน์ต่อประชาชนทั่วไปให้ตระหนักต่อความเสี่ยงการเป็นโรคไตเรื้อรัง และดูแลตนเองให้ห่างจากการเป็นโรคไตเรื้อรัง

สำหรับงานวิจัยในอนาคต ผู้ที่สนใจอาจพิจารณาวิธีการแก้ปัญหาความสัมพันธ์ระหว่างตัวแปรอธิบาย และพัฒนาตัวแบบเทคนิคการเรียนรู้ของเครื่อง ตัวแบบลอจิต เพื่อใช้การวิเคราะห์จำแนกระยะการเป็นโรคไตเรื้อรัง

กิตติกรรมประกาศ ผู้วิจัยขอขอบคุณภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร ที่สนับสนุนการทำการวิจัยในครั้งนี้ รวมถึงผู้ทรงคุณวุฒิทุกท่านที่ได้ให้ข้อคิดเห็นและข้อเสนอแนะต่าง ๆ เพื่อปรับปรุงบทความวิจัยนี้ให้สมบูรณ์

เอกสารอ้างอิง

[1] KDIGO (Kidney Disease: Improving Global Outcome). (2012).

[2] ณิชกานต์ วงษ์ประกอบ และ ลัฆวี ปิยะบัณฑิตกุล. ปัจจัยทำนายการเกิดโรคไตเรื้อรังในผู้ป่วยเบาหวานและความดันโลหิตสูง. *The Journal of Baromarajonani College of Nusing, Nakhonratchasima*. 2560; 23(2): กรกฎาคม - ธันวาคม.

[3] จรูญ บุญธกานนท์. ปัจจัยที่มีความสัมพันธ์กับการลดลงของอัตราการกรองของไตอย่างรวดเร็วในผู้ป่วยโรคไตเรื้อรังระยะ 3-4 โรงพยาบาลเกษตรสมบูรณ์ จังหวัดชัยภูมิ. *วารสารศูนย์อนามัยที่ 9 : วารสารส่งเสริมสุขภาพ และอนามัยสิ่งแวดล้อม*. 2566: 17(1): มกราคม - เมษายน.

[4] Elias Dritsas and Maria Trigka (2022). Machine Learning Techniques for Chronic Kidney Disease Risk Prediction. Department of Computer Engineering and Informatics, University of Patras, 26504 Patras, Greece. Big Data Cogn. Comput. 2022, 6(3), 98

[5] M.M. El Sherbiny, E. Abdelhalim, H.El. D. Mostafa, Severity prediction and multi classification of chronic kidney disease based on machine learning techniques, *Journal of Survey in Fisheries Sciences* 10 (3S) (2023) 4246-4256. Special Issue 3.

[6] ประเสริฐ ธนกิจจารุ, สกานต์ บุนนาค, วรางคณา พิชัยวงศ์. โรคไตเรื้อรัง (Chronic Kidney Disease: CKD). ใน: ชุษณะ มะกรสาร, บรรณาธิการ. *Thailand Medical Service Profile 2011–2014*. กรุงเทพฯ: กรมการแพทย์ กระทรวงสาธารณสุข; 2557. หน้า 8–1 – 8-55.

[7] วีรานันท์ พงศาภักดี, *การวิเคราะห์ข้อมูลเชิงกลุ่ม: ทฤษฎีและการประยุกต์ (กับ GLIM และ SPSS/FW)*, โรงพิมพ์มหาวิทยาลัยศิลปากร, นครปฐม, 2544.

[8] Multi category Logit Models, 2017. https://galton.uchicago.edu/~yibi/teaching/stat226/2017/lectures/C06.pdf. (accessed March 30, 2025).

[9] ผึ้งพร ลาภส่งผล, ณัฐภาภรณ์ รอดรัตษะ, วีรานันท์ พงศาภักดี และทิติยา จิตติหรรษา, การใช้ตัวแบบโลจิตสะสมในการศึกษาข้อมูลการชะลอการเจริญของเส้นใยรา *Ascophaera apis*. *วารสารวิทยาศาสตร์มหาวิทยาลัยนเรศวร*. 2551; 5(2): 176-189.

[10] Ali M., Son D., Kang S., & Nam S. An Accurate CT Saturation Classification Using a Deep Learning Approach Based on Unsupervised Feature Extraction and Supervised Fine-Tuning Strategy. *MDPI*. 2017; 10(11): 1830. https://doi.org/10.3390/en10111830.

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

A Comparative Study of Odds Ratio Estimators Under a Fixed Effect Model in Meta-Analysis Studies

Phimpatta Ariyawutthichai[1,*], Chukiat Viwatwongkasem[1] and Jutatip Sillabutra[1,†]

[1]Department of Biostatistics, Faculty of Public Health,
Mahidol University, Bangkok 10400, Thailand

Abstract

This research aimed to evaluate the performance of the pooled odds-ratio (θ) estimators: the profile likelihood estimator (PL), the inverse-variance weighted estimator (INV), Peto's estimator (PT), and Mantel-Haenszel estimator (MH), based on a fixed effect model via a simulation plan in meta-analysis studies. The results revealed that under no treatment effect $(\theta = 1)$, INV and PT perform the best with the smallest MSE, regardless of study sizes $(k = 4, 8, 16, 32, 64, 128)$, the sample size of each study in the treatment arm (n_i^T) and the control arm (n_i^C) at 10,50,100,200,500,1000. All four estimators (INV, PT, MH, PL) perform the best with the smallest MSE, (1) when study sizes are small $(k = 4, 8)$ in combination with moderate to large sample sizes $(n_i^T, n_i^C \geq 200)$, (2) when study sizes are moderate $(k = 16, 32)$ in combination with slightly moderate to large sample sizes $(n_i^T, n_i^C \geq 100)$, (3) when study sizes are large $(k = 64, 128)$ in combination with strongly small to large sample sizes $(n_i^T, n_i^C \geq 50)$. Under existence of treatment effect$(\theta = 2)$, INV and PT perform the best with the smallest MSE in most situations with small to heavily moderate study sizes $(k = 4, 8, 16, 32)$ and $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$. All four estimators (INV, PT, MH, PL) perform the best with the smallest MSE, (1) when study sizes are heavily moderate $(k = 32)$ in combination with large sample sizes $(n_i^T, n_i^C \geq 500)$, (2) when study sizes are large $(k = 64)$ in combination with strongly moderate to large sample sizes $(n_i^T, n_i^C \geq 200)$, (3) when study sizes are very large $(k = 128)$ in combination with slightly moderate to large sample sizes $(n_i^T, n_i^C \geq 100)$.

Keywords: Meta-Analysis, Profile likelihood, Inverse-Variance Weighted estimator, Peto's estimator, Mantel-Haenszel estimator

2020 MSC: 60A10,60E05

*Speaker (phimpatta@gmail.com)

†Corresponding author (jutatip.sil@mahidol.ac.th)

1 Introduction

Meta-Analysis is a statistical technique to combine all previous effect results of multiple research studies under a similar research topic to produce a new summarizing effect result as a weighted mean of estimates for determining the effect size of particular interventions or exposure factors affecting on a defined outcome. There are two classical methods, either the fixed effect model or the random effect model, to combine the previously different effects into a new single conclusion effect. The basic distinction between the fixed effects and the random effects is that under homogeneity of effect measures overall k studies $(\theta_1 = \theta_2 = \cdots = \theta_k = \theta)$, there is a single true effect size (θ) assumed, or in other words all population effect sizes from multiple k studies are assumed equal and the fixed effect model is suitable enough to analyze the data. Conversely, under heterogeneity or the presence of variation among study outcomes, there are multiple true effect sizes assumed, or in other words the effect sizes have a specific distribution, and the random effect model is more appropriate for analysis. Before choosing either the fixed effects model or the random effects model to compute a summary effect size, the Cochran's Q test statistic is used to test the null hypothesis of homogeneity of effect sizes overall studies.

Based on binary outcomes, the effect size measure of interest in this study compared between treatment and control groups is considerably referred to a common odds ratio (θ). In a single trial, we have known that the odds ratio estimate can be computed from a 2-by-2 table. Then, a meta-analysis is an extensively progressive ground of pooling the odds ratio estimators into a summarizing odds ratio measure in the form of a weighted mean. The odds ratio as a popular effect size can be used in all observational study designs, including the cohort study, the case-control study, and the cross-sectional study. The odds ratio provides a range value of $[0, \infty)$; if its value is greater than 1, the risk of exposed group or treatment group is larger than the risk of non-exposed group or control group. In contrast, if the value of odds ratio is less than 1, the exposure or treatment group is referred to a protective factor.

Under a common odds ratio (θ) across k studies or under homogeneity of the odds ratios $(\theta_1 = \theta_2 = \cdots = \theta_k = \theta)$ overall k studies, there are several weights used to find the summary odds ratio estimators under the fixed effect model of interest. Firstly, the inverse-variance weighted method (INV) was first described by Cochran (1937) [1]; each weight of the INV is inversely proportional to each variance of study estimate directly. Secondly, the Mantel-Haenszel (MH) method was first arisen by Mantel and Haenszel (1959) [2]; each weight of the MH is proportional to sample size (total number of people) in treatment and control groups. Next, the Peto's method (PT) was first described by Peto, et al. (1977) [3]; each weight of the PT depends on the difference between the observed (O) and the expected (E) number of events from both treatment and control groups. Lastly, Böhning, et al. (2008) [4] proposed the profile likelihood (PL) estimator of interest by elimination of the nuisance parameter and the result of PL provides the iterative processes.

Since there are different summary estimators to combine odds ratios available under the simple and wildly used fixed effects model in meta-analysis, the gap of this study is based on the desirable guideline to indicate which a particular estimator is most appropriate to use in different situations. As it has been discovered in the past, the Peto method (PT) is suitable

for rare events and small sample sizes and it also works well for studies with zero-count cells; however, Greenland and Salvan [5] in 1990 stated that PT can yield large biases and large standard errors in the imbalanced sample sizes of two treatment groups ($n_i^T \neq n_i^C$). The Mantel-Haenszel (MH) and the inverse-variance weighted (INV) methods are simple and most wildly used in general situations. Fleiss (1981) [6] compared the performance between the MH and the INV and it has been found that the MH is better if there are many studies to combine and the sample size within study is small. The INV is more suitable for large sample sizes but small study sizes. Up to now, we do not yet know the performance of the profile likelihood method (PL), especially in comparisons among other summary odds ratio estimators.

2 Methodology

A comparison of the performance of odds ratio estimators under the fixed effect model in meta-analysis focuses on 1) Examples of meta-analysis used, 2) All candidates of estimation, 3) Simulation plan and process.

2.1 Examples of meta-analysis used

All four examples were selected to describe the number of events (x_i^T, x_i^C) out of the sample sizes (n_i^T, n_i^C) of the treatment and control groups in each study i $(i = 1,...,k)$, leading to the odds ratio estimators $\hat{O}R_i = x_i^T \times (n_i^C - x_i^C)/x_i^C \times (n_i^T - x_i^T)$ as the treatment effects in each study i, also the number of studies (k).

Sun et al. [7] in 2018 conducted a meta-analysis of all 6 studies to study the effect of diet interventions on iron deficient children with anemia. The results of meta-analysis revealed that diet interventions are effective in improving iron deficiency with iron deficiency anemia. From the k results of odds ratios in Table 1, the pooling odds ratio as a single value of the summary odds ratio over all k studies is computed by means of the inverse-variance weighted method (INV), Peto's method (PT), Mantel-Haenszel method (MH), and the profile likelihood method (PL) as $\hat{\theta}_{INV} = \hat{O}R_{INV} =$ Nan, $\hat{\theta}_{PT} = \hat{O}R_{PT} =$ 5.03, $\hat{\theta}_{MH} = \hat{O}R_{MH} =$ 6.82, and $\hat{\theta}_{PL} = \hat{O}R_{PL} =$ 7.00, respectively.

Yang et al. (2018) [8] conducted a meta-analysis of all 11 studies to determine the association between the occupational noise exposure and ECG in meta-analysis. The relationship shows that working in a noisy environment affects the abnormalities of the ECG, leading to the coronary heart disease and hypertension with odds ratio value greater than 1 in each study i. Moreover, from the k results of odds ratios in Table 2, the summary odds ratios are computed as $\hat{\theta}_{INV} = \hat{O}R_{INV} =$ 0.7766, $\hat{\theta}_{PT} = \hat{O}R_{PT} =$ 2.0687, $\hat{\theta}_{MH} = \hat{O}R_{MH} =$ 2.1976, and $\hat{\theta}_{PL} = \hat{O}R_{PL} =$ 2.2054.

Table 1: Sun et al.'s data

Study (i)	Treatment		Control		$\hat{OR}_i$
	Event (x_i^T)	Total (n_i^T)	Event (x_i^C)	Total (n_i^C)	
Lan H 2012	50	52	46	51	2.72
Liu M 2010	71	75	50	75	8.87
Qi GJ 2011	70	70	65	70	Inf.
Tian J 2010	58	60	54	60	3.22
Wang SH 2008	47	50	31	49	9.09
Zhu M 2011	31	32	27	32	5.74

Table 2: Yang et al.'s data

Study (i)	Treatment		Control		$\hat{OR}_i$
	Event (x_i^T)	Total (n_i^T)	Event (x_i^C)	Total (n_i^C)	
Chen Zhengqi 2003	102	574	35	293	1.5930
Chen Zlran 2007	61	463	16	286	2.2606
Fu Xuying 2011	25	76	6	45	3.1863
Lin Jinmo 2005	42	267	19	252	2.2891
Liu Yimin 2007	33	130	11	134	3.8041
Lou Shuyan 2006	273	843	26	183	2.8921
Su Guangxiao 2016	32	85	10	90	4.8302
Wang Hulwen 2004	33	124	13	113	2.7895
Xiao Qihua 2008	159	953	93	953	1.8518
Yang Shule 2002	171	1023	42	486	2.1217
Zu Aihua 2012	37	233	16	233	2.5603

Cao, L. (2018) [9] conducted a meta-analysis of all 8 studies to determine the risk of thyroid cancer in patients with and without inflammatory bowel disease (IBD). The study found that patients with inflammatory bowel disease have a higher risk of developing disease of the gland's Cao cancer. From the k results of the odds ratios in Table 3, the summary odds ratio is computed as $\hat{\theta}_{INV} = \hat{OR}_{INV} = 1.7465$, $\hat{\theta}_{PT} = \hat{OR}_{PT} = 1.8525$, $\hat{\theta}_{MH} = \hat{OR}_{MH} = 1.7494$, and $\hat{\theta}_{PI} = \hat{OR}_{PI} = 1.7316$.

Table 3: Cao's data

Study (i)	IBD		Control		$\hat{OR}_i$
	Event (x_i^T)	Total (n_i^T)	Events (x_i^C)	Total (n_i^C)	
Jess 2004	1	374	18	37400	5.5678
Jung 2017	38	15291	2548	1529100	1.4926
Jussila 2013	34	21964	1859	2196400	1.8302
So 2017	1	1603	188	160300	0.5316
Sonu 2013	13	2686	9	1638	0.8803
This study 2018	11	1392	3	1392	3.6879
Wadhwa 2016	145	289935	80	315145	1.9706
Yano 2013	1	770	55	77000	1.8192

Ma et al. (2017) [10] conducted a study on the relationship between Traditional Chinese Medicine (TCM) constitution and dyslipidemia in a total of 11 studies of meta-analysis. The results in Table 4 showed that phlegm-dampness constitution was a risk factor of dyslipidemia as $\hat{\theta}_{INV} = \hat{OR}_{INV} = 2.4266$, $\hat{\theta}_{PT} = \hat{OR}_{PT} = 2.5254$, $\hat{\theta}_{MH} = \hat{OR}_{MH} = 2.4971$ and $\hat{\theta}_{PI} = \hat{OR}_{PI} = 2.4624$.

Table 4: Data of Ma et al.

Study (i)	Treatment		Control		$\hat{OR}_i$
	Event (x_i^T)	Total (n_i^T)	Events (x_i^C)	Total (n_i^C)	
Cal et al. 2016	148	343	27	143	3.2608
Huang. 2015	165	638	52	470	2.8041
Shi et al. 2016	68	511	243	3383	1.9835
Wang et al. 2013	9	200	6	305	2.3482
Wang et al. 2014	349	1145	314	1956	2.2927
Wang et al. 2016	23	87	246	1869	2.3710
Wang 2012	79	342	29	279	2.5895
Xie et al. 2013	32	258	4	231	8.0354
Xu 2016	13	40	5	40	3.3704
Zhang et al. 2015	53	238	24	199	2.0890
Zhou et al. 2014	21	107	2	106	12.6977

2.2 All Candidates of Estimation

Under the fixed effect model, all effect sizes are assumed equal, or there is no heterogeneity between the studies, or containing only a single true effect size over all the studies. The differences of the studies in observed effects are due to sampling error. In this paper, let parameter of interest be a common odds ratio ($\theta = OR$) that is assumed to be a

constant over all k studies $(i=1,...,k)$ under the fixed-effects model. There are three popular approaches to estimate this common odds ratio over all k studies as the treatment effect with binary outcomes: the inverse-variance weighted method, Peto's method, and Mantel-Haenszel method. In addition, the profile likelihood method of interest for estimating the summary odds ratio over all studies from the work of Böhning et al. (2008) [4] is also used to compare the performance with three those approaches.

2.2.1 Inverse Variance Weighted Method

Since the odds ratio estimators $\hat{\theta}_i = \hat{OR}_i$ from study i often have a right skew distribution, the natural logarithm of the odds ratio estimator, $\log \hat{OR}_i$, is preferable to use under the approximate normal distribution and better to find the weighted mean and variance. Then the logarithm of inverse-variance weighted estimator for a true constant, $\log OR$, over all k studies under a fixed-effect model $(\log OR_1 = ... = \log OR_k = \log OR)$ is in the summary estimator of the form:

$$\hat{\theta}_{IVW} = \frac{\sum_{i=1}^{k} \widehat{w_i}(log\widehat{OR})}{\sum_{i=1}^{k} \widehat{w_i}}$$

where the natural logarithm of the odds ratio estimator, $\log \hat{OR}_i$, for each study i is denoted by

$$\log \hat{OR}_i = \log\left(\frac{\hat{p}_i^T}{1-\hat{p}_i^T}\right) - \log\left(\frac{\hat{p}_i^C}{1-\hat{p}_i^C}\right),$$

where the $\hat{p}_i^T = x_i^T / n_i^T$ and $\hat{p}_i^C = x_i^C / n_i^C$ are the estimated risk probabilities in the treatment group and the control group, and the weight estimates $\hat{w}_i$ are defined by the inverse variance estimates

$$\hat{w}_i = \frac{1}{\hat{\text{var}}[\log \hat{OR}_i]} = \left(\frac{1}{x_i^T} + \frac{1}{n_i^T - x_i^T} + \frac{1}{x_i^C} + \frac{1}{n_i^C - x_i^C}\right)^{-1}$$

Additionally, the variance of the summary estimator, $\log \hat{OR}_{INV}$, is given by

$$\widehat{var}\left(log\widehat{OR}_{INV}\right) = \frac{1}{\sum_{i=1}^{k} \widehat{w}_i}$$

An approximate 95% confidence interval for a true value $\log OR$ is given by

$$log\widehat{OR}_{INV} \pm 1.96\sqrt{\widehat{var}(log\widehat{OR}_{INV})}$$

Thus, the summary estimates for a true odds ratio constant $\theta = OR$ over all k studies by the means of the inverse-variance weighted method can be obtained as

$$\hat{\theta}_{INV} = \widehat{OR}_{INV} = \exp\left(log\widehat{OR}_{INV}\right)$$

An approximate 95% confidence interval for a true value $\theta = OR$ is given by

$$exp\left(log\widehat{OR}_{INV} \pm 1.96\sqrt{\widehat{var}(log\widehat{OR}_{INV})} \right)$$

2.2.2 Peto's Method

Peto's method for pooling odds ratios is a modification of the Mantel-Haenszel method plus the inverse variance weighted method for finding the pooled variance. To calculate the pooled odds ratio by Peto's method, we consider a 2x2 table for each study i in a clinical trial study.

Table 5 A 2x2 table for each study i to calculate a Peto's odds ratio estimate

Study i	Event outcome	Non-event outcome	Total
Treatment	$O_i = x_i^T$	$n_i^T - x_i^T$	n_i^T
Control	x_i^C	$n_i^C - x_i^C$	n_i^C
Total	$d_i = x_i^T + x_i^C$	$(n_i^T - x_i^T) + (n_i^C - x_i^C)$	$n_i = n_i^T + n_i^C$

Defining $O_i = x_i^T$ is the observed number of events in the treatment arm,

$d_i = x_i^T + x_i^C$ is the number of events in both treatment and control groups,

$n_i^T,\ n_i^C$ are the number of patients in the treatment and control groups,

$n_i = n_i^T + n_i^C$ is a total number of patients in the ith study,

$E_i = (n_i^T/n_i)d_i$ is the expected number of events in the treatment group.

Thus, the Peto's odds ratio estimate over all k studies can be obtained as

$$\hat{\theta}_{PETO} = exp\left(\frac{\sum_{i=1}^{k}(O_i - E_i)}{\sum_{i=1}^{k} V_i} \right)$$

where V_i is the weighting factor or the variance of the difference between the observed O_i and the expected values E_i as $O_i - E_i$, under assumption of no difference effects between the treatment and the control groups.

$$V_i = E_i\left(\frac{n_i - n_i^T}{n_i}\right)\left(\frac{n_i - d_i}{n_i - 1}\right)$$

For k studies, the variance estimates of the natural log of the pooled estimate $\hat{\theta}_{PETO}$ is provided by

$$\widehat{var}\left(\log \hat{\theta}_{PETO}\right) = \left(\sum_{i=1}^{k} V_i\right)$$

The $100(1-\alpha)\%$ confidence interval is given by

$$exp\left(\frac{\sum_{i=1}^{k}(O_i - E_i) \pm Z_{\alpha/2}\sqrt{\left(\sum_{i=1}^{k} V_i\right)}}{\sum_{i=1}^{k} V_i}\right)$$

where $Z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of a standard normal distribution.

2.2.3 Mantel-Haenszel Method

Mantel-Haenszel proposed a pooled estimator for combining odds ratios from several stratified case-control studies to have a single common odds ratio estimate, and this method can also be used in a meta-analysis. The pooled odds ratio estimator as a Mantel-Haenszel method was introduced by Böhning et al. (2008) [4] as

$$\hat{\theta}_{MH} = \frac{\sum_{i=1}^{k} \frac{\left(n_i^C - x_i^C\right)x_i^T}{n_i^T + n_i^C}}{\sum_{i=1}^{k} \frac{\left(n_i^T - x_i^T\right)x_i^C}{n_i^T + n_i^C}} = \frac{\sum_{i=1}^{k} a_i d_i / n_i}{\sum_{i=1}^{k} b_i c_i / n_i}$$

where $a_i = x_i^T, b_i = n_i^T - x_i^T$, $c_i = x_i^C$, $d_i = n_i^C - x_i^C$, $n_i = n_i^T + n_i^C$. A variance estimate for the estimated summary odds ratio in the log form was derived by Robin et al. (1999) as

$$\widehat{var}(\widehat{log\,\theta}_{MH}) = \frac{\sum_{i=1}^{k} P_i R_i}{2\left(\sum_{i=1}^{k} R_i\right)^2} + \frac{\sum_{i=1}^{k}(P_i S_i + Q_i R_i)}{2\left(\sum_{i=1}^{k} R_i\right)\left(\sum_{i=1}^{k} S_i\right)} + \frac{\sum_{i=1}^{k} Q_i S_i}{2\left(\sum_{i=1}^{k} S_i\right)^2}$$

where $P_i = (a_i + d_i)/n_i, Q_i = (b_i + c_i)/n_i$, $R_i = a_i d_i / n_i$, $S_i = b_i c_i / n_i$. Thus,the $100(1-\alpha)\%$ confidence interval for the summary odds ratio is given by

$$exp\left(\widehat{log\,\theta}_{MH} \pm Z_{\alpha/2}\sqrt{\widehat{var}\,(\widehat{log\,\theta}_{MH})}\right)$$

where $Z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of a standard normal distribution.

2.2.4 Profile Likelihood Method

In fixed effects models, the profile likelihood is particularly useful when the focus is on a subset of the parameters. Here, the parameters are divided into two sets, the interest parameters, and the nuisance parameters. Let's assume a parameter vector ϑ was split into these two classes:

$$\vartheta = (\alpha, \beta)$$

where α is of a vector parameter of interest and β is a vector of the nuisance parameters. Generally, the profile likelihood technique utilizes the identical feature as the maximum likelihood model; however, unlike the maximum likelihood model, the profile likelihood estimation deals with the elimination of the nuisance parameters to find the parameters of interest of the profile likelihood. Suppose the ordinary likelihood function referring to the sampling model of observed data $y = (y_1, \dots y_k)$ is written as

$$L(\vartheta \; ; \; y) \; = \; L(\alpha, \beta \; ; \; y) \; = \; \prod_{i=1}^{k} f \; (y_i \; ; \; \alpha, \beta)$$

Generally, the maximum likelihood estimators $\hat{\vartheta} = (\hat{\alpha}, \hat{\beta})$ of parameters $\vartheta = (\alpha, \beta)$ are obtained by maximizing the likelihood function with respect to α and β:

$$Max_{(\alpha,\beta)} L(\hat{\alpha}, \hat{\beta} \; ; \; y)$$

But the profile likelihood function $L_p(\alpha \; ; \; y)$ of the parameter of interest α is obtained by maximizing over the nuisance parameters $\hat{\beta}_\alpha$ as a function of α such that $L(\beta_\alpha | \alpha) \geq L(\beta | \alpha)$ for all β. That is, after substituting $\hat{\beta}_\alpha$, the profile likelihood function $L_p(\alpha \; ; \; y)$ now depends only on the parameter of interest α:

$$L_p(\alpha \; ; \; y) = Max_{(\beta)} L(\alpha, \hat{\beta}_\alpha \; ; \; y)$$

Finally, the profile maximum likelihood estimator $\hat{\alpha}$ of parameter of interest α is achieved by

$$Max_{(\alpha)} L_p(\hat{\alpha} \; ; \; y)$$

Next, observing the work of Böhning et al. (2008) [4] in the means of the profile likelihood estimation to find a summary odds ratio from combining all k-studies odds ratios. Nonetheless, some terms which are not available in the derivation of Böhning et al. (2008) can be extended their derivation in the elaborate details. After complete derivation, let the odds estimate in control arm $\hat{q}_i^C = \hat{p}_i^C / (1 - \hat{p}_i^C)$ be the nuisance parameters with $\hat{p}_i^C = x_i^C / n_i^C$; while the odds ratio estimate of interest for each study i is denoted by $\hat{\theta}_i = \hat{q}_i^T / \hat{q}_i^C$ where $\hat{q}_i^T = \hat{p}_i^T / (1 - \hat{p}_i^T)$ and $\hat{p}_i^T = x_i^T / n_i^T$. The nuisance parameters $q_i^C[\theta_i]_1$ as a function of θ_i are found as

$$q_i^C[\theta_i]_1 = -\frac{2(x_i^T + x_i^C)}{x_i^T + x_i^C - n_i^C + (x_i^T + x_i^C - n_i^T)\theta_i - \sqrt{r_i(\theta_i)}}$$

where $r_i(\theta_i) = -4(x_i^T + x_i^C)(x_i^T + x_i^C - n_i^T - n_i^C)\theta_i + (x_i^T + x_i^C - n_i^C + (x_i^T + x_i^C - n_i^T)\theta_i)^2$

Under homogeneity situations of $\theta_1 = \theta_2 = \dots = \theta_k = \theta$, the associated fixed-point iteration for a common odds ratio θ is denoted by

$$\theta = \frac{\sum_{i=1}^{k} \theta_i^T}{\sum_{i=1}^{k} \frac{n_i^T \left(\theta \times (q_i^C)'[\theta]\right) + (q_i^C[\theta])}{1 + \theta \times q_i^C[\theta]} + \frac{n_i^C (q_i^C)'[\theta]}{1 + q_i^C[\theta]} \frac{(x_i^T + x_i^C)(q_i^C)'[\theta]}{q_i^C[\theta]}} := \Phi_{OR}(\theta)$$

In other formula, the pooled odds ratio by the profile likelihood method is obtained as

$$\theta = \frac{\sum_{i=1}^{k} \frac{\left(n_i^C - x_i^C\right)x_i^T}{\sqrt{r_i(\theta)}}}{\sum_{i=1}^{k} t_i\,(\theta)}$$

where

$$t_i(\theta) = \frac{n_i^T\left(\theta(q_i^C)'[\theta]+(q_i^C[\theta])\right)}{1+\theta\times q_i^C[\theta]} + \frac{n_i^C(q_i^C)'[\theta]}{1+q_i^C[\theta]} - \frac{(x_i^T+x_i^C)(q_i^C)'[\theta]}{q_i^C[\theta]} + \frac{(n_i^C-x_i^C)(x_i^T)-x_i^T\sqrt{r_i[\theta]})}{\theta\sqrt{r_i[\theta]}}$$

$$r_i(\theta) = -4(x_i^T + x_i^C)(x_i^T + x_i^C - n_i^T - n_i^C)\theta + (x_i^T + x_i^C - n_i^C + (x_i^T + x_i^C - n_i^T)\theta)^2$$

$$q_i^C[\theta]_1 = -\frac{2(x_i^T+x_i^C)}{x_i^T+x_i^C-n_i^C+(x_i^T+x_i^C-n_i^T)\theta-\sqrt{r_i(\theta)}}$$

$$(q_i^C)'[\theta]_1 = \frac{2(x_i^T + x_i^C)\left((x_i^T + x_i^C - n_i^T) - \frac{(r_i)'[\theta]}{2\sqrt{r_i(\theta)}}\right)}{\left(x_i^T + x_i^C - n_i^C + (x_i^T + x_i^C - n_i^T)\theta - \sqrt{r_i(\theta)}\right)^2}$$

$$(r_i)'[\theta] = -4(x_i^T + x_i^C)(x_i^T + x_i^C - n_i^T - n_i^C) + 2(x_i^T + x_i^C - n_i^C + (x_i^T + x_i^C - n_i^T)\theta)(x_i^T + x_i^C - n_i^T)$$

Under $\theta = 1$ with the result of first iterative step, the profile likelihood odds ratio estimator is identical to the Mantel-Haenszel estimator as

$$\hat{\theta}_{MH} = \frac{\sum_{i=1}^{k} \frac{\left(n_i^C - x_i^C\right)x_i^T}{n_i^T + n_i^C}}{\sum_{i=1}^{k} \frac{\left(n_i^T - x_i^T\right)x_i^C}{n_i^T + n_i^C}}$$

2.3 Simulation plan and process

A summary of parameters from four examples used for creating simulation plan is presented in Table 6 where k is the number of studies being combined, n_i^T is the total number persons in the treatment group, n_i^C is the total number persons in the control group, $\hat{p}_i^T$ is the observed proportion of an event in the treatment arm, $\hat{p}_i^C$ is the observed proportion of an event in the control arm, and $\hat{\theta}$ is the pooled odds ratio estimate over all k studies.

Table 6: Summary of parameters used in simulation plans

Example	1	2	3	4
Study size (k)	6	11	8	11
Frequency (n_i^T)	32-75	76-1023	374-289935	20-1145
Frequency (n_i^C)	32-75	45-923	1392-2196400	20-3383
Proportion risk ($\hat{p}_i^T$)	0.94-1	0.13-0.37	0.001-0.008	0.05-0.43
Proportion risk ($\hat{p}_i^C$)	0.63-0.93	0.06-0.14	0-0.005	0.02-0.19
Method	Fixed	Fixed	Fixed	Fixed
OR's Pooled ($\hat{\theta}$)	5.03	2.27	1.75	2.50

Parameter Setting:

- The true common odds ratio OR (θ) is predetermined and fixed over all studies $(i = 1,...,k)$ at the values of 1, 2, 5, 10.
- The number of studies (k) was chosen at the values of 4, 8, 16, 32, 64, 128.
- The sample sizes in the treatment arm (n_i^T) and the control arm (n_i^C) for each study i were generated using Poisson distributions with parameter mean n_T, n_C where n_T, n_C were equal in balanced cases and were unequal in unbalanced cases. The mean parameters were specified as 10, 50, 100, 200, 500, 1000. Note that $n_i^T \sim Poi(n_T)$ and $n_i^C \sim Poi(n_C)$.
- Let the risk parameter in control arm be a uniform distribution as $p_i^C \sim U(0,0.8)$ and the associated risk parameter in treatment arm can be rewritten as $p_i^T = \theta p_i^C / (1 - p_i^C + \theta p_i^C)$.

Statistical calculation:

- The number of events in the treatment and the control arms were generated using binomial distributions $x_i^T \sim B(n_i^T, p_i^T)$ and $x_i^C \sim B(n_i^C, p_i^C)$ for each study i .
- Computational statistics at each study i are as $\hat{p}_i^C = x_i^C / n_i^C$, $\hat{p}_i^T = x_i^T / n_i^T$, $\hat{q}_i^C = \hat{p}_i^C / (1 - \hat{p}_i^C)$, $\hat{q}_i^T = \hat{p}_i^T / (1 - \hat{p}_i^T)$ and $\hat{\theta}_i = \hat{q}_i^T / \hat{q}_i^C$.
- Four statistical pooling methods used to estimate the treatment effect (odds ratio) in this research are Inverse-Variance Weighted estimator (INV), Peto's estimator (PT), Mantel-Haenszel estimator (MH), and Profile maximum likelihood estimator (PL).
- All simulations were performed with the statistical software R for Windows, Version 3.6.3 with 10,000 replications.

- Compare the efficient performance of the pooled odds ratio estimators with 1) Mean (or Biasedness), 2) Variance (or SE: standard error), and 3) MSE (or RMSE: squared root of mean square error) under 10,000 replications per each combination scenario of simulation study. It yields:

  Mean: $\bar{\theta}_m = \sum_{j=1}^{10,000} \frac{\hat{\theta}_{(m)_j}}{10,000}$

  Bias: $Bias_m = \sum_{j=1}^{10,000} \frac{\hat{\theta}_{(m)_j}}{10,000} - \theta$

  Variance: $VAR_m = \sum_{j=1}^{10,000} \frac{\left(\hat{\theta}_{(m)_j} - \bar{\theta}_m\right)^2}{10,000}$

  MSE: $MSE_m = \sum_{j=1}^{10,000} \frac{\left(\hat{\theta}_{(m)_j} - \theta\right)^2}{10,000}$

  where m is estimation method.

3 Results

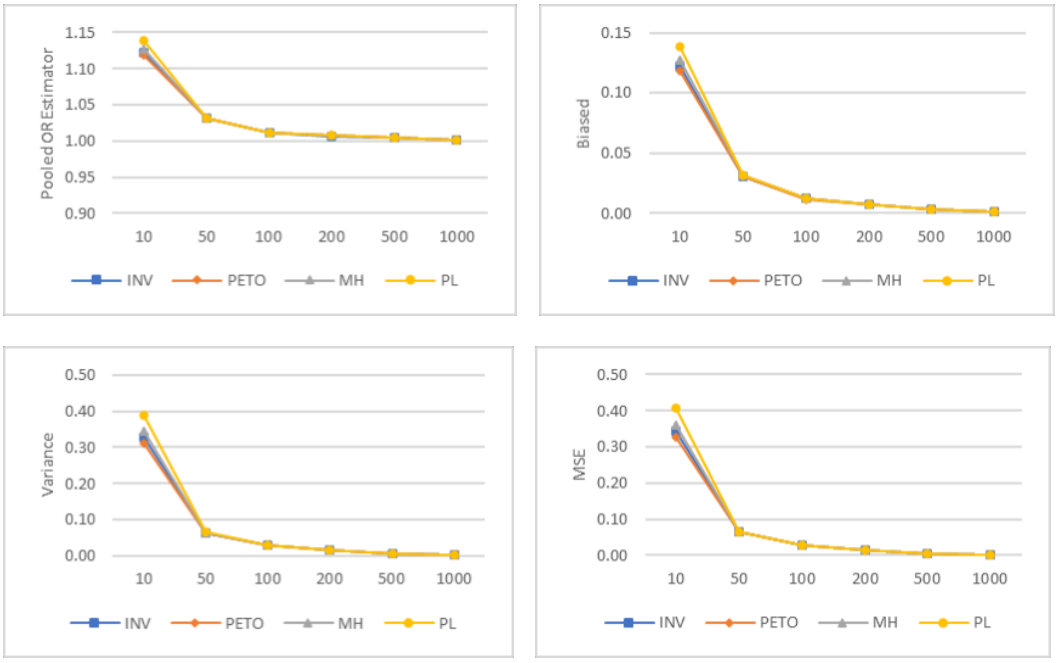
3.1 Balanced case $(n_i^T = n_i^C = n)$

Before describing the detailed results, we would like to summarize overall results under the balanced cases as follows:

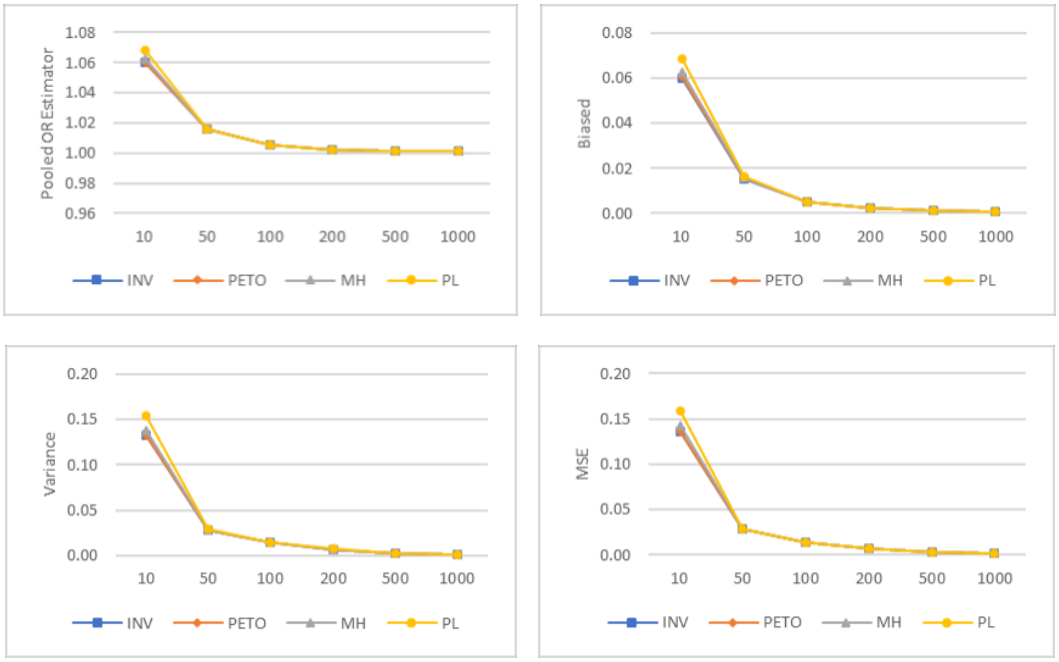
1. Generally, under none of treatment or exposure effect in terms of odds ratio equal to 1 $(\theta = 1)$, INV and PT perform the best with the smallest MSE in every situation, regardless of study sizes $(k = 4, 8, 16, 32, 64, 128)$, and regardless of sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.
2. Under none of treatment or exposure effect $(\theta = 1)$, all four estimators (INV, PT, MH, PL) perform the best with the smallest MSE, (1) when study sizes are small $(k = 4, 8)$ in combination with moderate to large sample sizes $(n_i^T, n_i^C \geq 200)$, (2) when study sizes are moderate $(k = 16, 32)$ in combination with slightly moderate to large sample sizes $(n_i^T, n_i^C \geq 100)$, (3) when study sizes are large $(k = 64, 128)$ in combination with strongly small to large sample sizes $(n_i^T, n_i^C \geq 50)$.
3. Under existence of treatment effect $(\theta = 2)$, INV and PT still perform the best with the smallest MSE in most situations with small to heavily moderate study sizes $(k = 4, 8, 16, 32)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.
4. Under existence of treatment effect $(\theta = 2)$, all four estimators (INV, PT, MH, PL) perform the best with the smallest MSE, (1) when study sizes are heavily moderate $(k = 32)$ in combination with large sample sizes $(n_i^T, n_i^C \geq 500)$, (2) when study sizes are large $(k = 64)$ in combination with strongly moderate to large sample sizes

$(n_i^T, n_i^C \geq 200)$, (3) when study sizes are very large $(k = 128)$ in combination with slightly moderate to large sample sizes $(n_i^T, n_i^C \geq 100)$.

5. Generally, under large treatment effect $(\theta = 5)$, all three estimators (PL, MH, INV) performed the best with the smallest MSE, followed by PT. In more details, if study sizes are small to moderate $(k \leq 32)$ in combination with heavily small to large sample sizes $(n_i^T, n_i^C \geq 50)$, the INV performs the best; meanwhile, if study sizes are moderate to large $(k \geq 16)$ at almost every sample sizes $(n_i^T, n_i^C = 50, 100, 200, 500, 1000)$, the MH is best practically. Additionally, PL is to be preferred, (1) when study size is large $(k \geq 64)$ in combination with large sample size $(n_i^T, n_i^C \geq 500)$, and (2) when study size is very large $(k \geq 128)$ at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$. Moreover, PT provided the worst performance because it produced seriously large bias and large variance in these large effects $(\theta = 5)$.
6. Generally, under very large treatment effect $(\theta = 10)$, all three estimators (PL, MH, INV) performed the best with the smallest MSE, followed by PT. In more details, if study sizes are small to moderate $(k \leq 32)$ in combination with heavily small to large sample sizes $(n_i^T, n_i^C \geq 50)$, the INV performs the best; meanwhile, if study sizes are moderate to large $(k \geq 16)$ at almost every sample sizes $(n_i^T, n_i^C = 50, 100, 200, 500, 1000)$, the MH is best practically. Additionally, PL is to be preferred when study size is large $(k \geq 64)$ at every sample size. Moreover, PT provided the worst performance because it produced seriously large bias and large variance in these large effects $(\theta = 10)$.
7. PL, in addition to the above general cases, under no effect size $(\theta = 1)$, performs well with smallest MSE, (1) when study sizes are small $(k = 4,8)$ in combination with moderate to large sample sizes $(n_i^T,\ n_i^C \geq 200)$, (2) when study sizes are moderate $(k = 16,32)$ in combination with slightly moderate to large sample sizes $(n_i^T,\ n_i^C \geq 100)$, (3) when study sizes are large $(k = 64,128)$ in combination with heavily small to large sample sizes $(n_i^T,\ n_i^C \geq 50)$.
8. PL, under mild effect size $(\theta = 2)$, performs best, (1) when study sizes are small to moderate $(k \leq 32)$ in combination with large sample sizes $(n_i^T, n_i^C \geq 500)$, (2) when study sizes are large $(k = 64, 128)$ in combination with small and slightly moderate to large sample sizes $(n_i^T, n_i^C = 10\ and \geq\ 200)$.
9. PL, under strong effect size $(\theta \geq 5)$, performs best practically with smallest MSE, (1) when study sizes are small to moderate $(k \leq 32)$ in combination with small and large sample sizes $(n_i^T, n_i^C = 10, 1000)$, (2) when study sizes are large $(k = 64, 128)$ and at every sample sizes $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.
10. INV, under no treatment effect $(\theta = 1)$, performs the best with the smallest MSE in every situations, regardless of study sizes $(k = 4, 8, 16, 32, 64, 128)$, and regardless of sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.
11. INV, under existence of treatment effect $(\theta = 2)$, still performs the best with the smallest MSE in most situations with small to heavily moderate study sizes $(k = 4, 8, 16, 32)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.

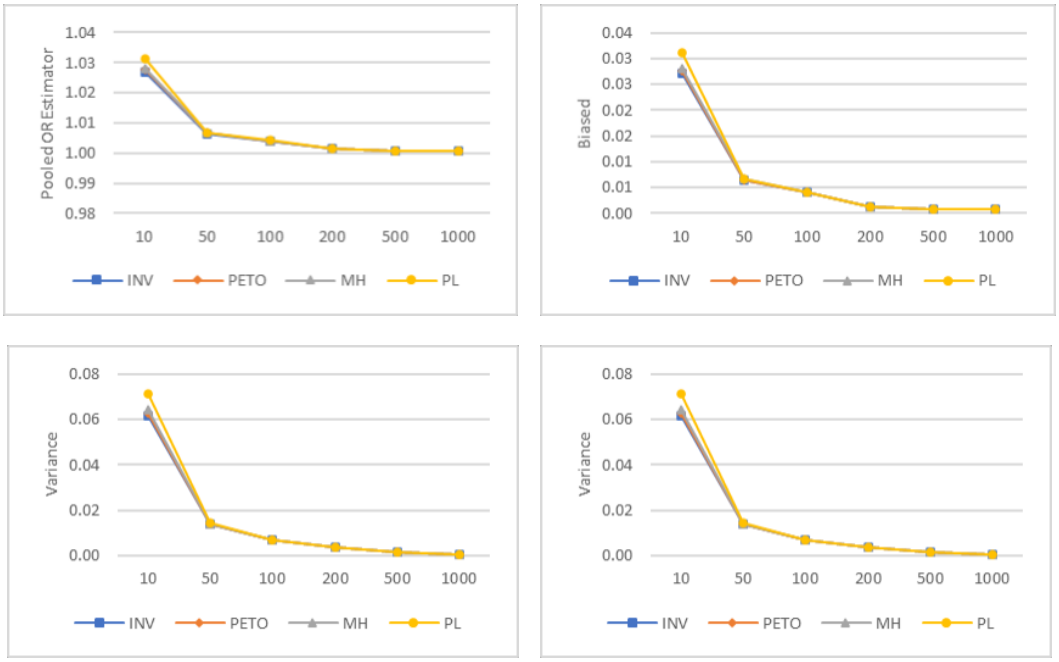


(a) θ= 1, k = 4

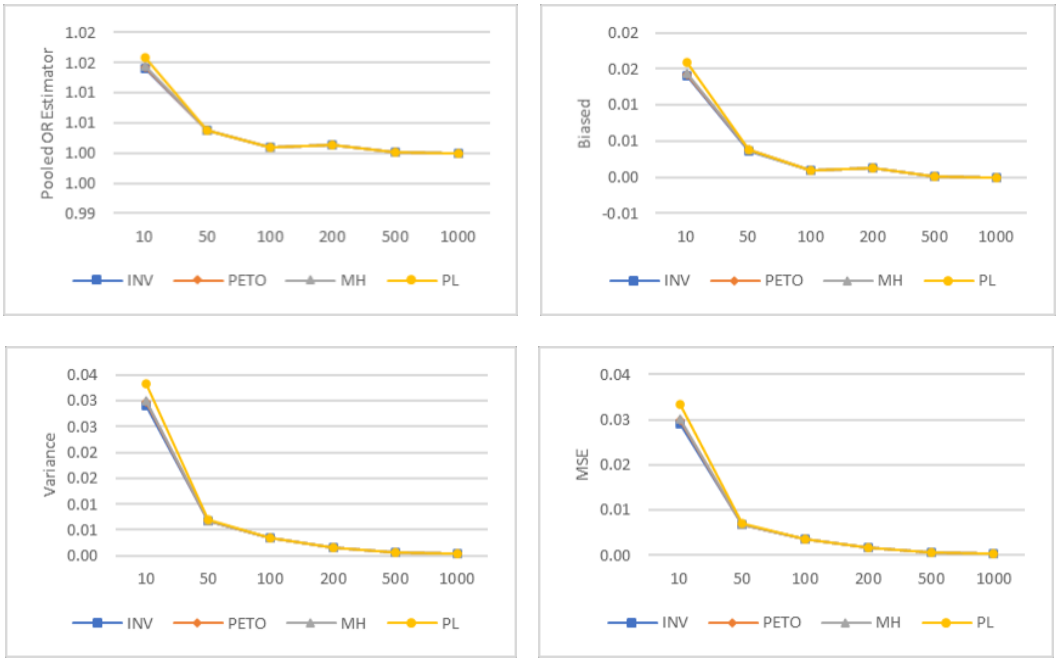


(b) θ= 1, k = 8

Figure 1: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 1$ with balance case scenarios.

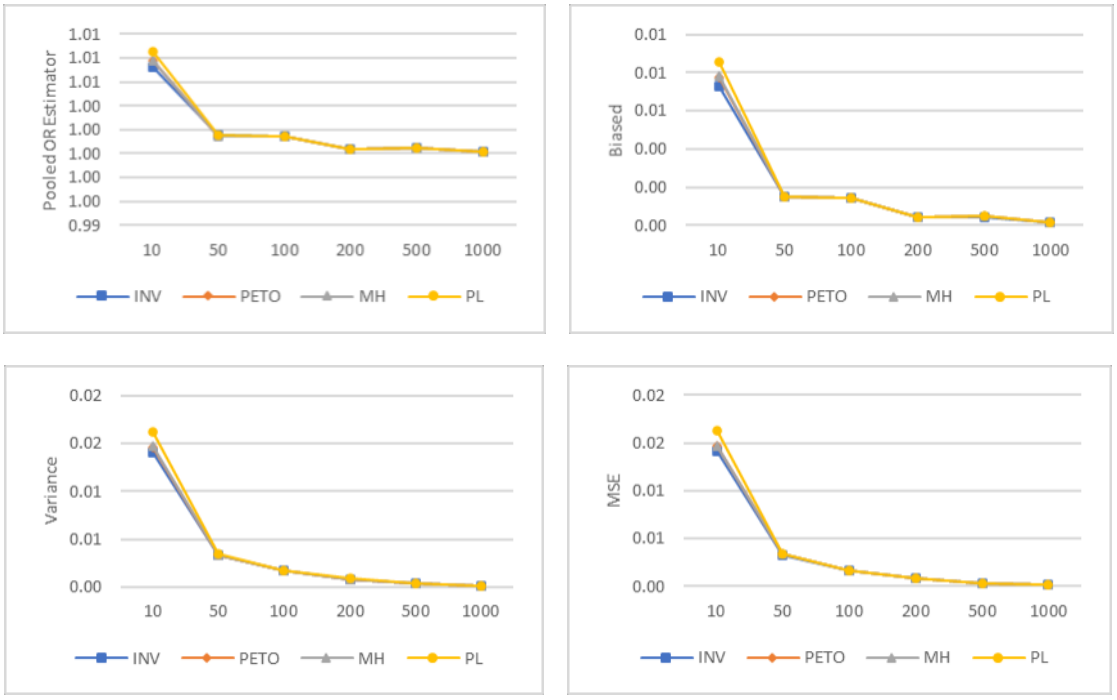


(c) θ= 1, k = 16

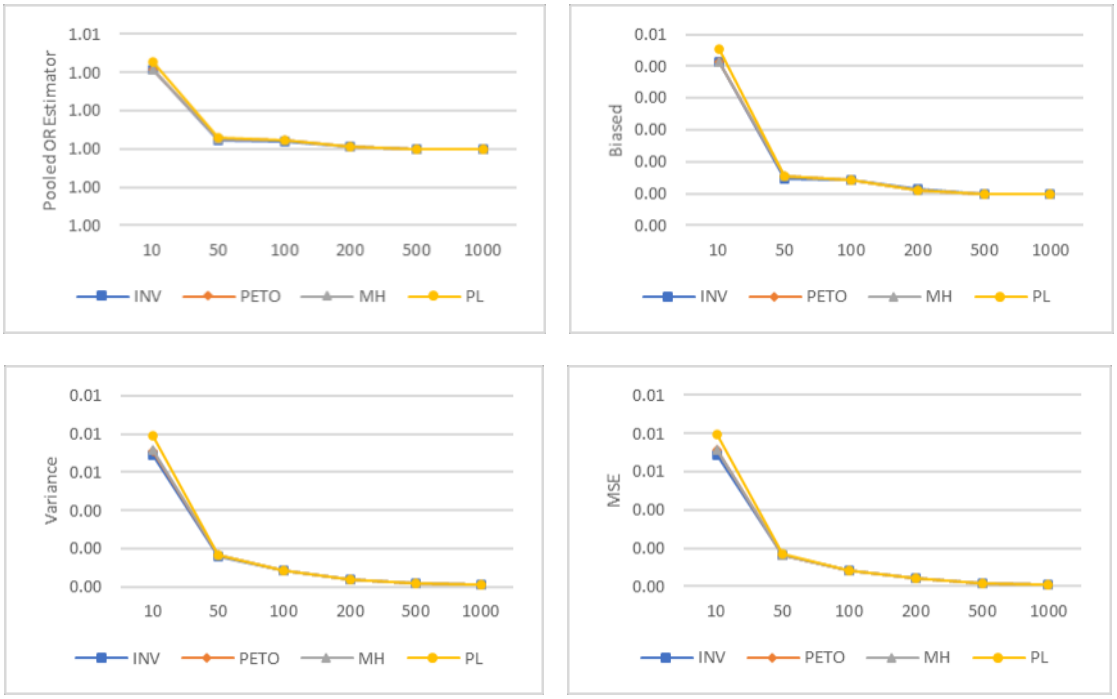


(d) θ= 1, k = 32

Figure 1: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 1$ with balance case scenarios.(Continued)

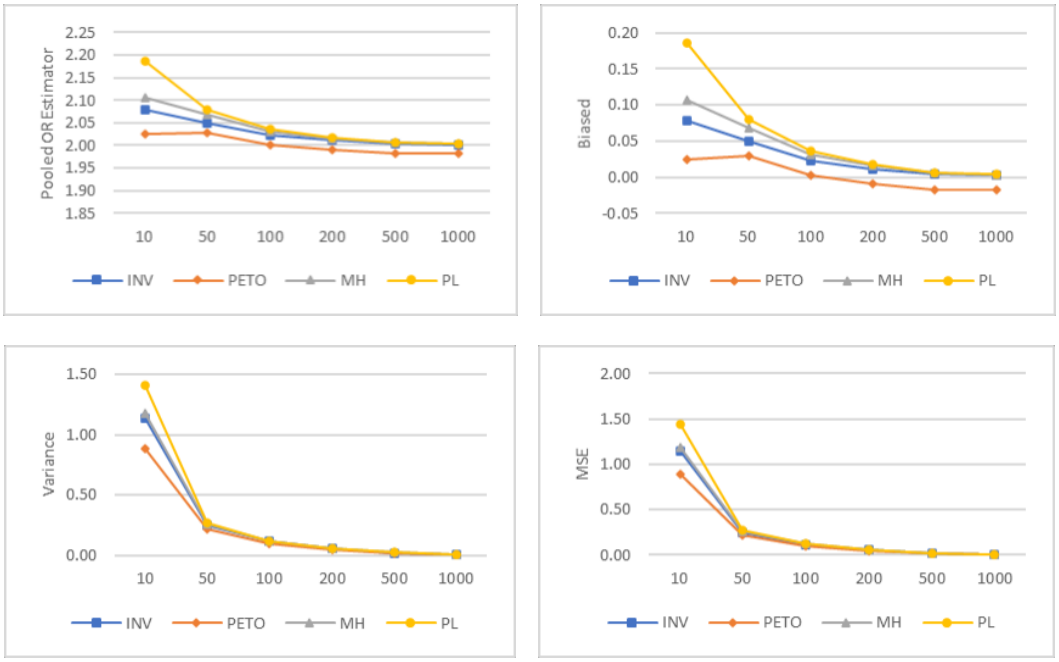


(e) θ= 1, k = 64

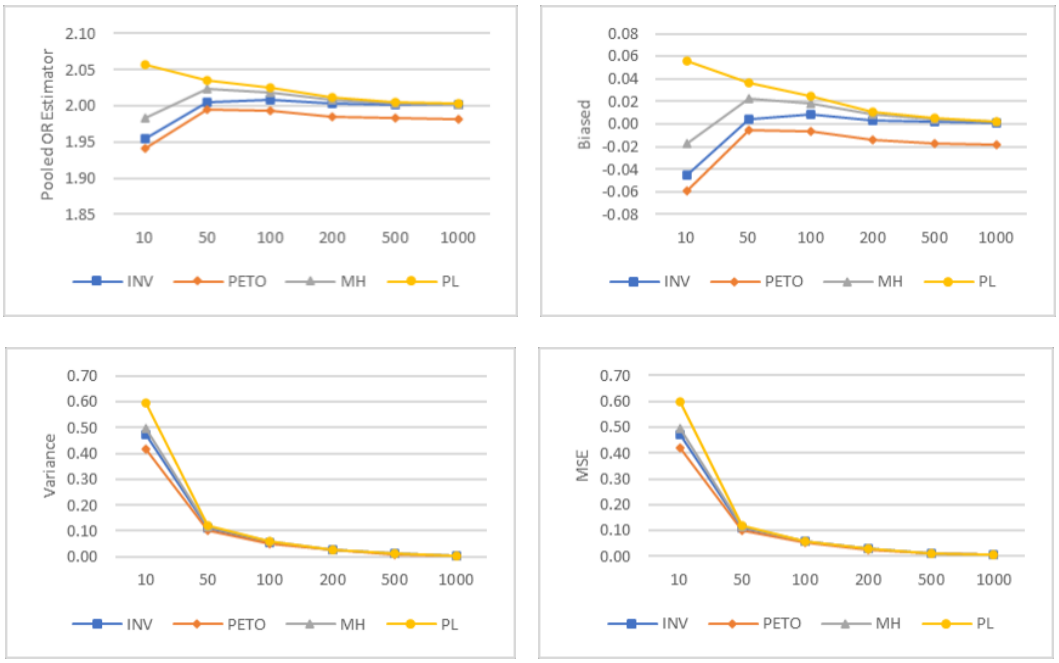


(f) θ= 1, k = 128

Figure 1: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 1$ with balance case scenarios.(Continued)

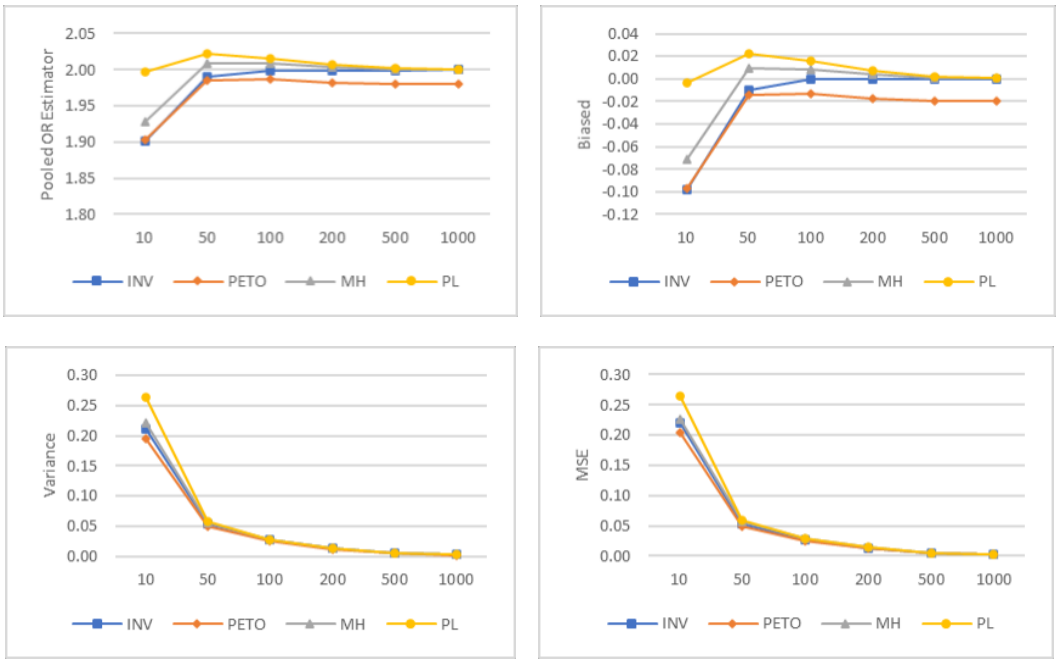


(a) $\theta = 2$, $k = 4$

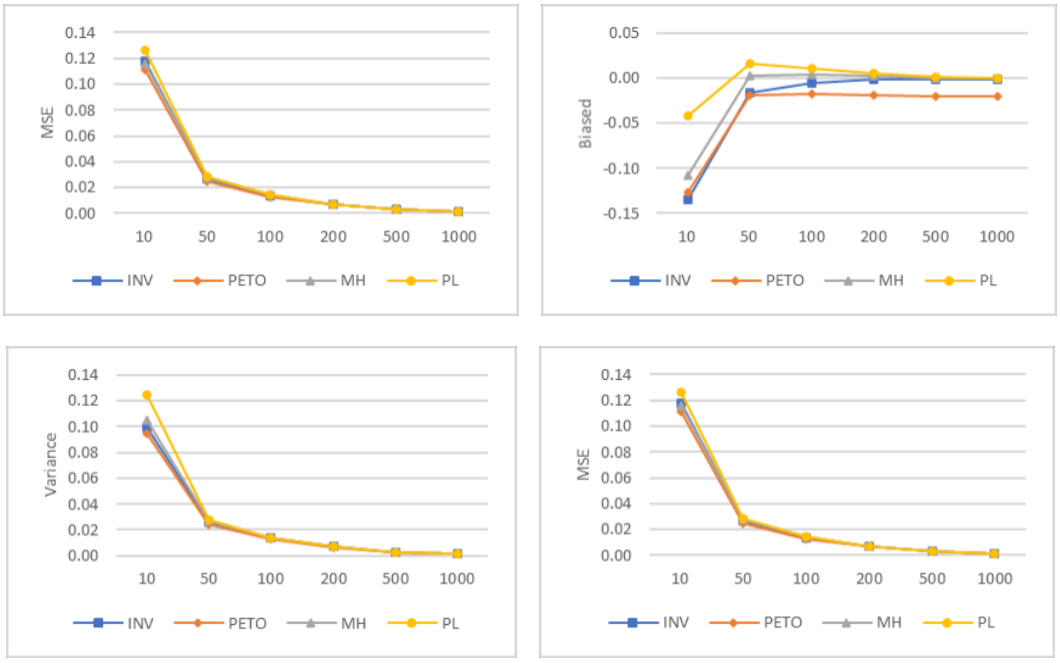


(b) $\theta = 2$, $k = 8$

Figure 2: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 2$ with balance case scenarios.(Continued)

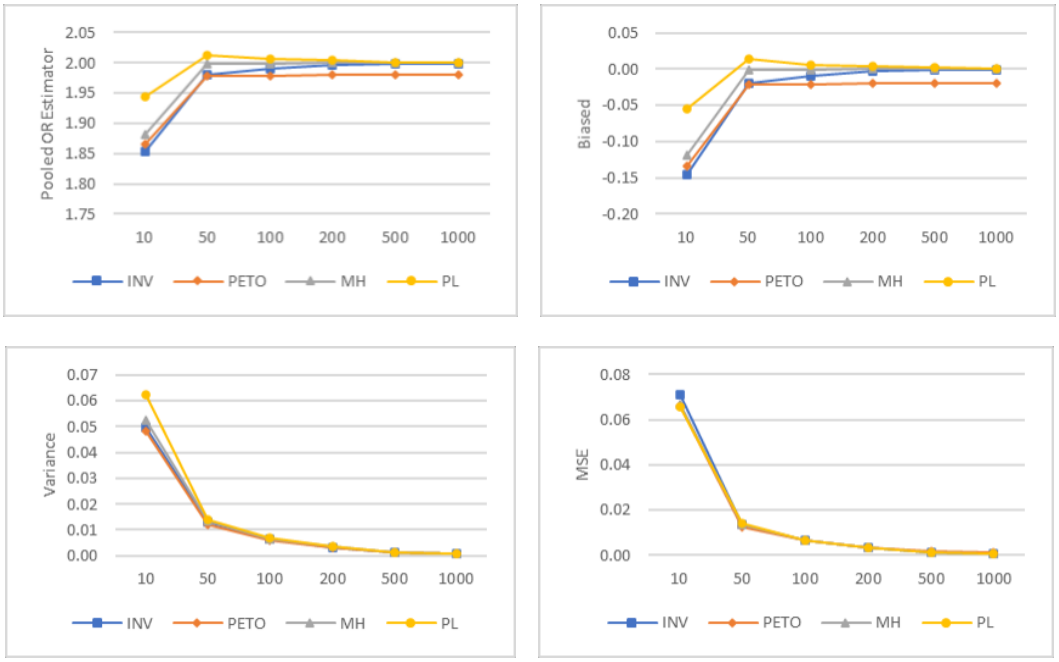


(c) θ= 2, k = 16

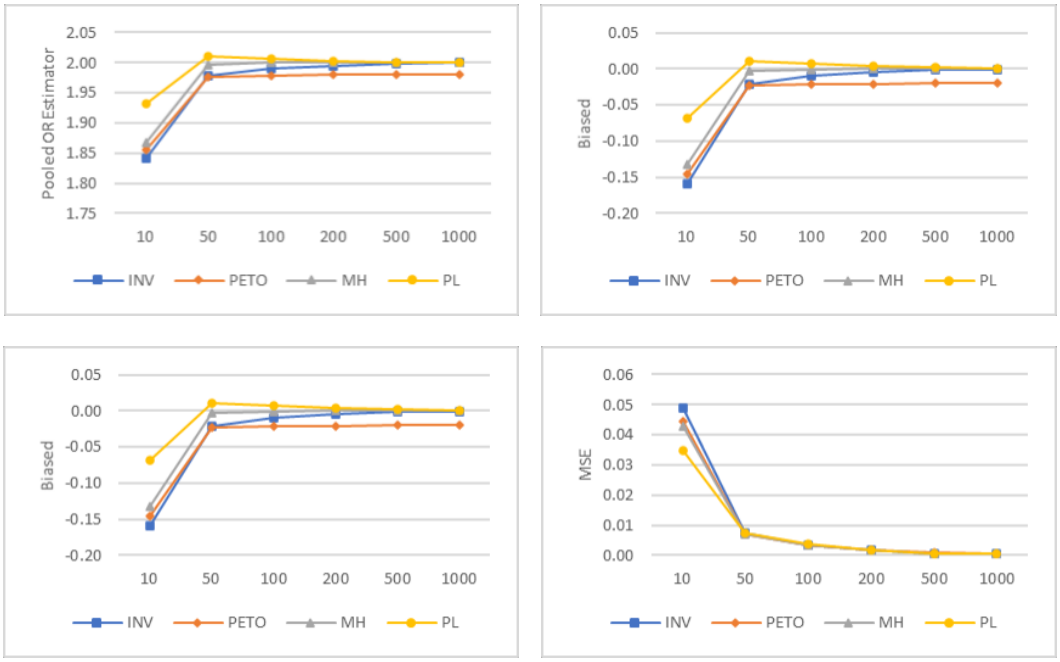


(d) θ= 2, k = 32

Figure 2: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 2$ with balance case scenarios.(Continued)

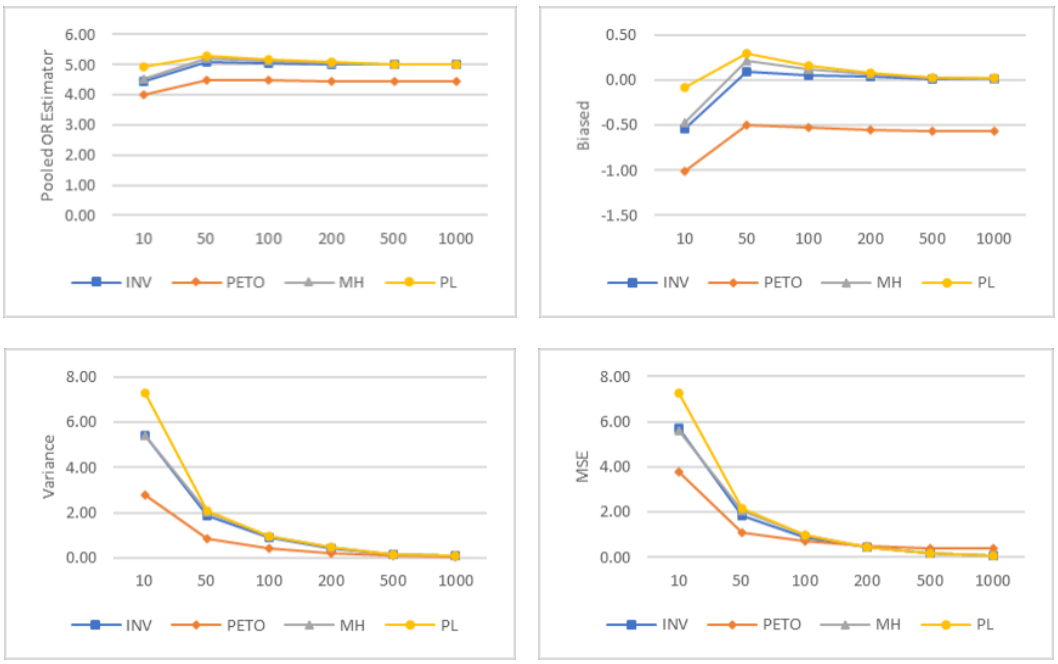


(e) θ= 2, k = 64

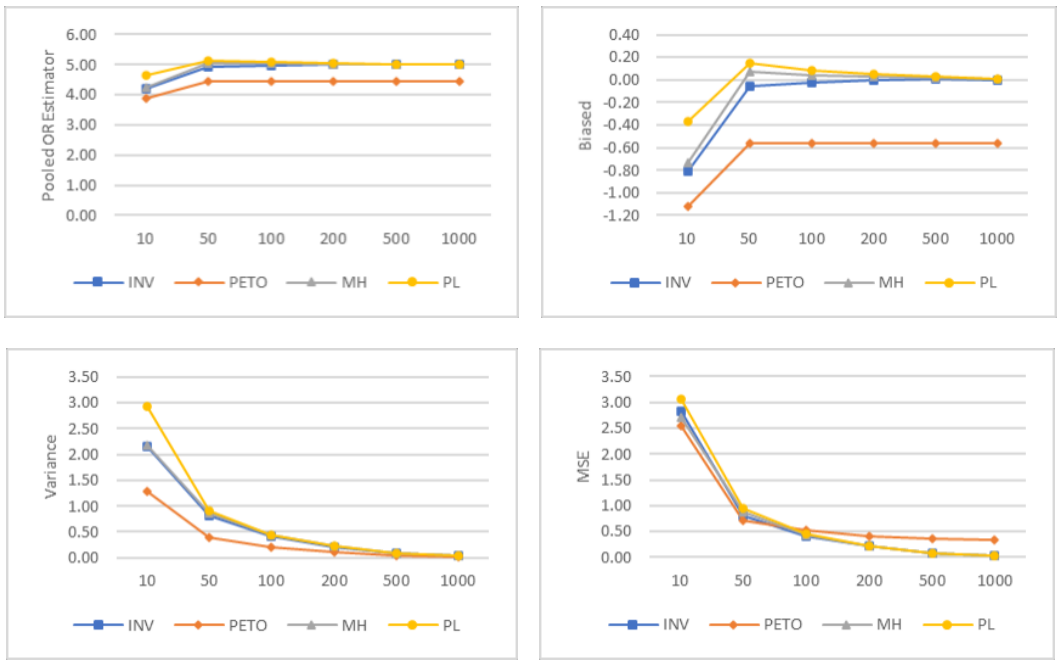


(f) θ= 2, k = 128

Figure 2: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 2$ with balance case scenarios.(Continued)

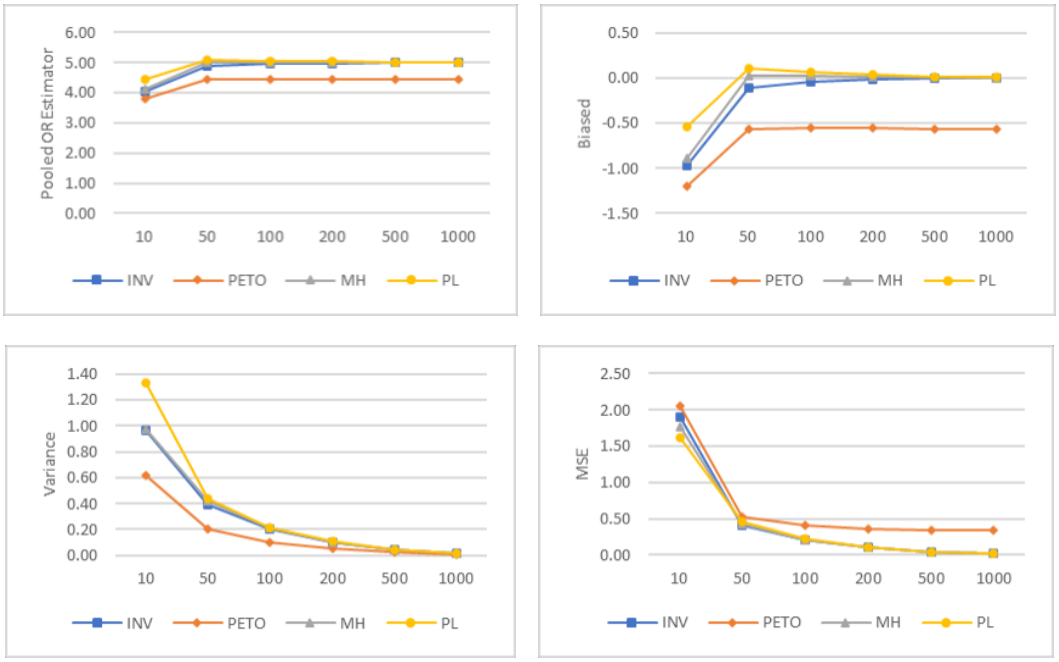


(a) θ= 5, k = 4

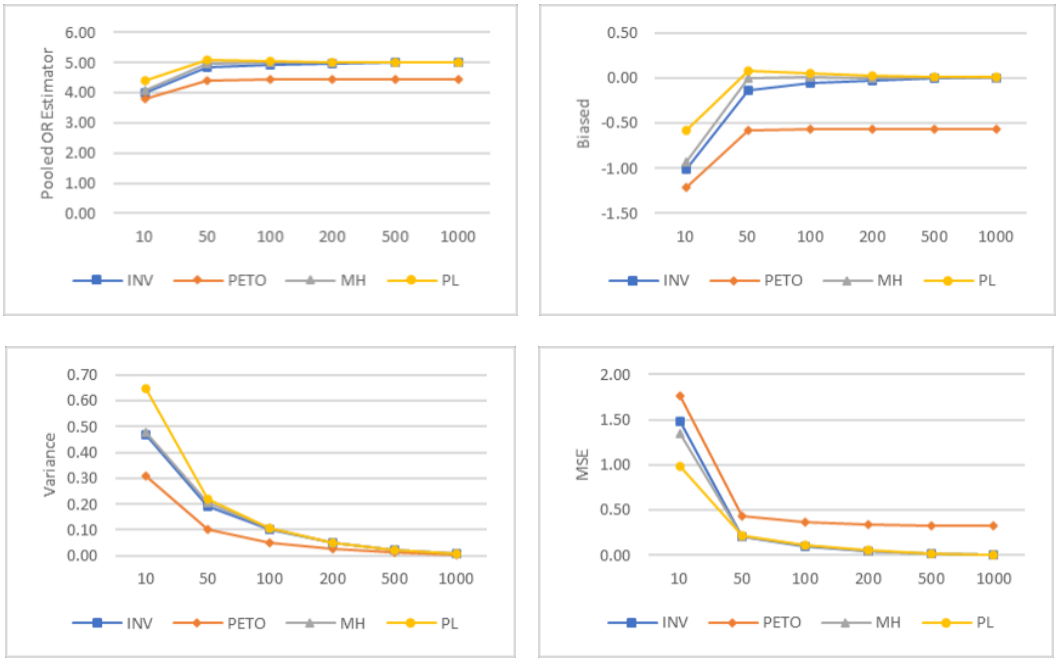


(b) θ= 5, k = 8

Figure 3: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 5$ with balance case scenarios.(Continued)

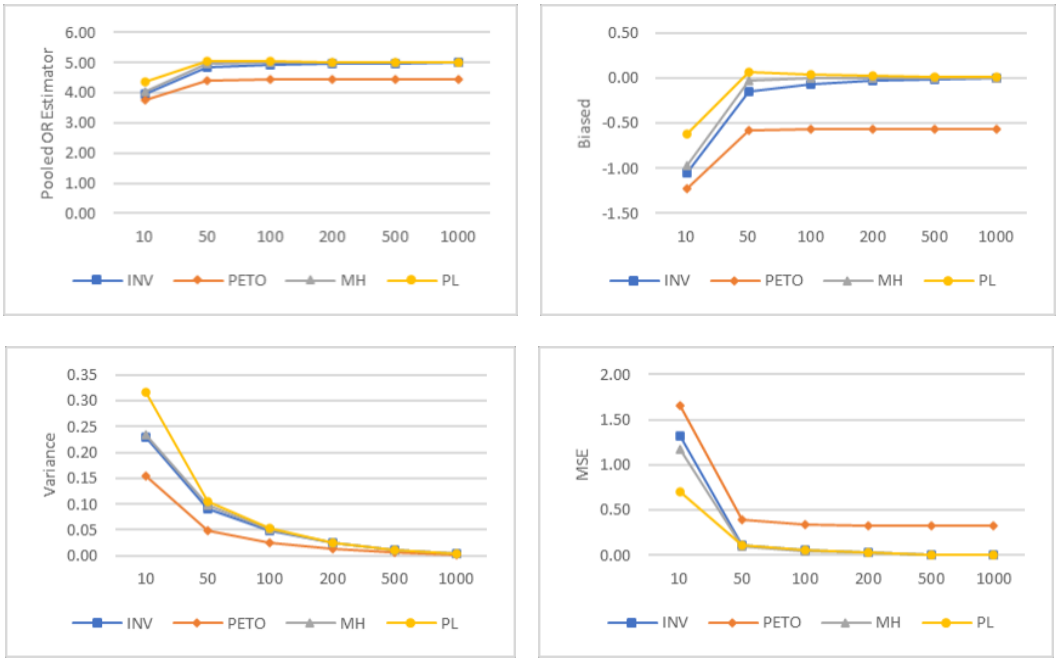


(c) θ= 5, k = 16

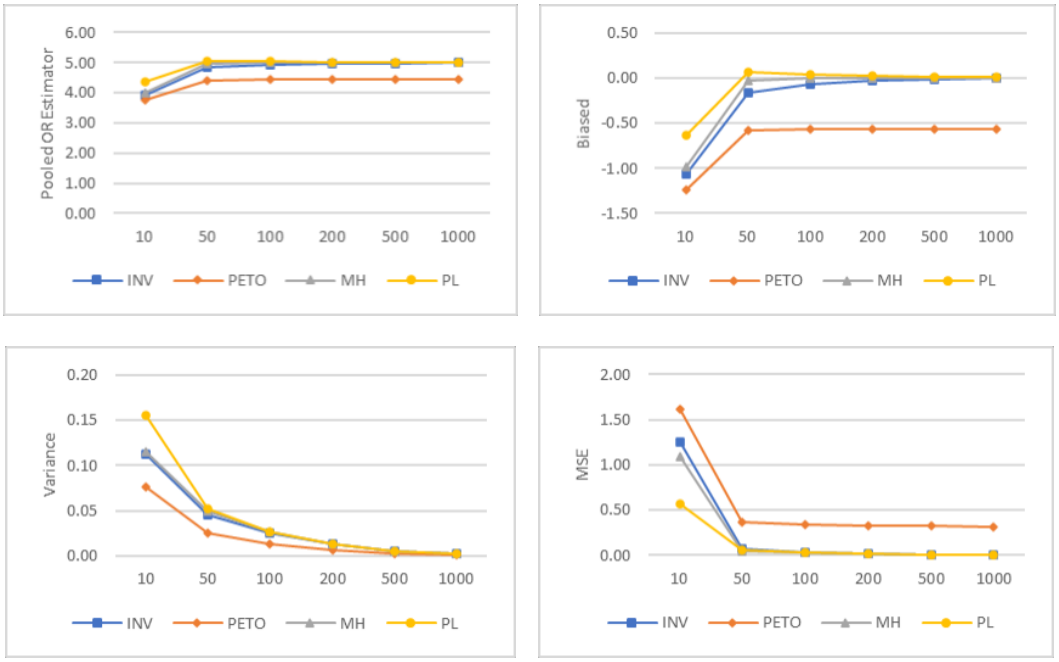


(d) θ= 5, k = 32

Figure 3: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 5$ with balance case scenarios.(Continued)

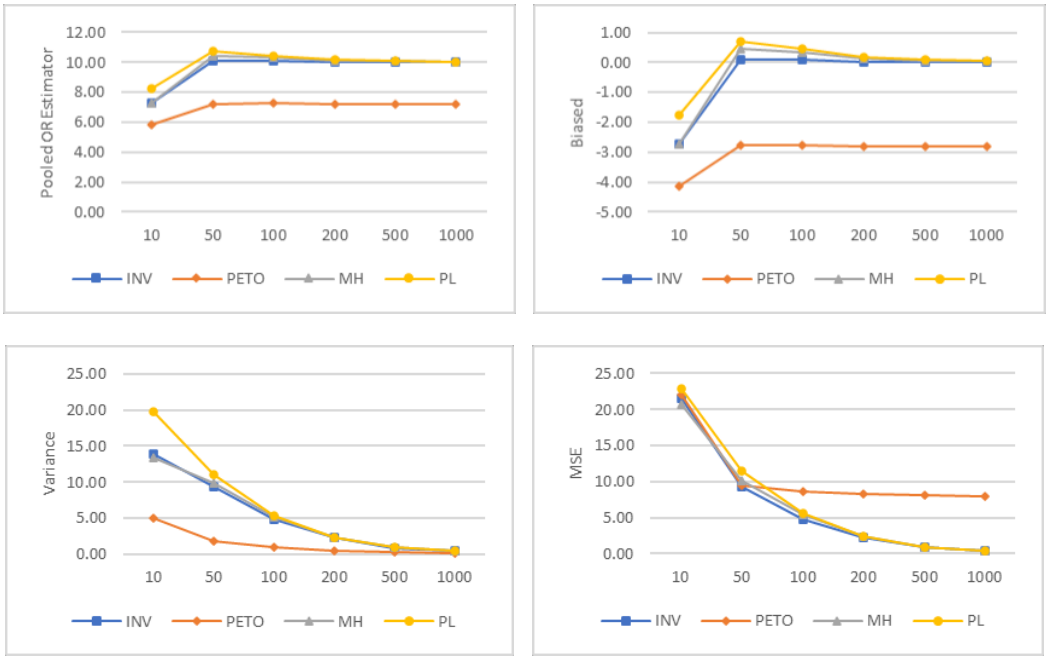


(e) θ= 5, k = 64

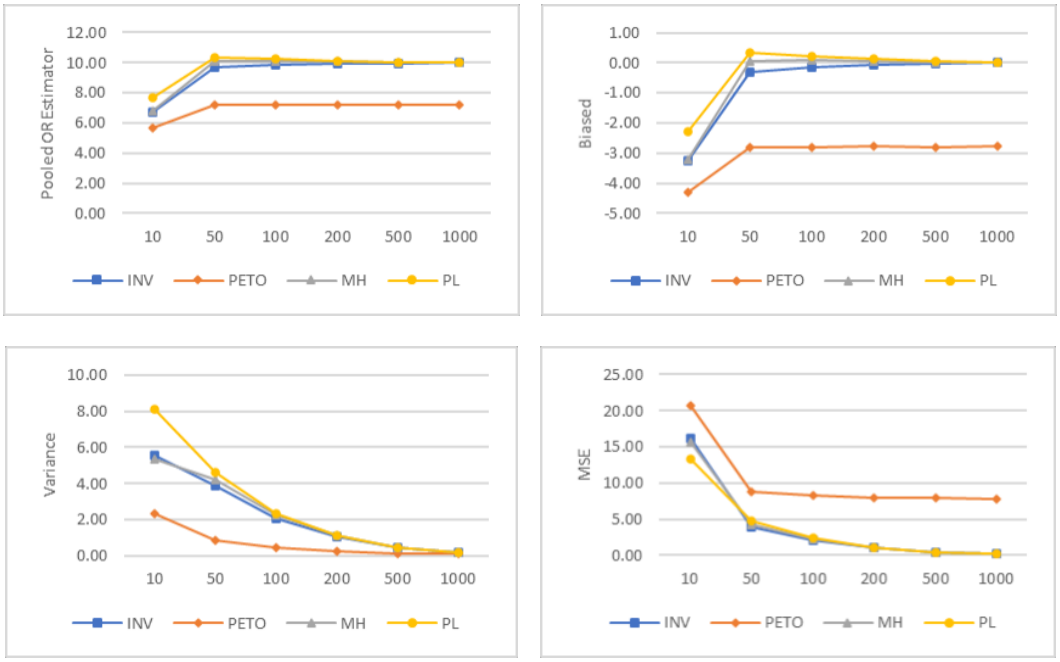


(f) θ= 5, k = 128

Figure 3: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 5$ with balance case scenarios.(Continued)

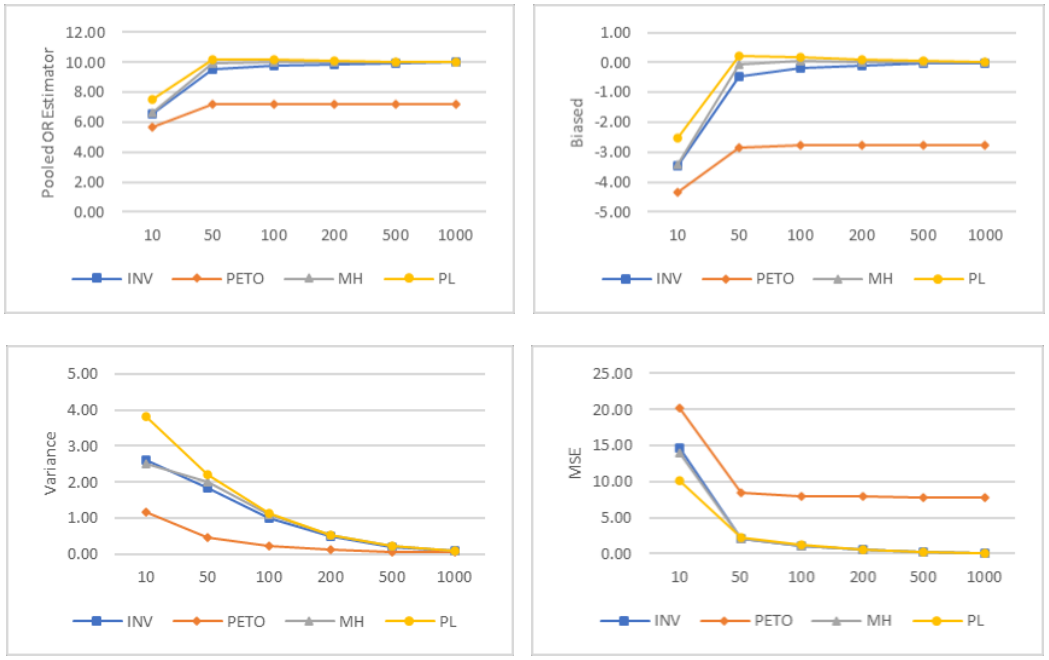


(a) θ= 10, k = 4

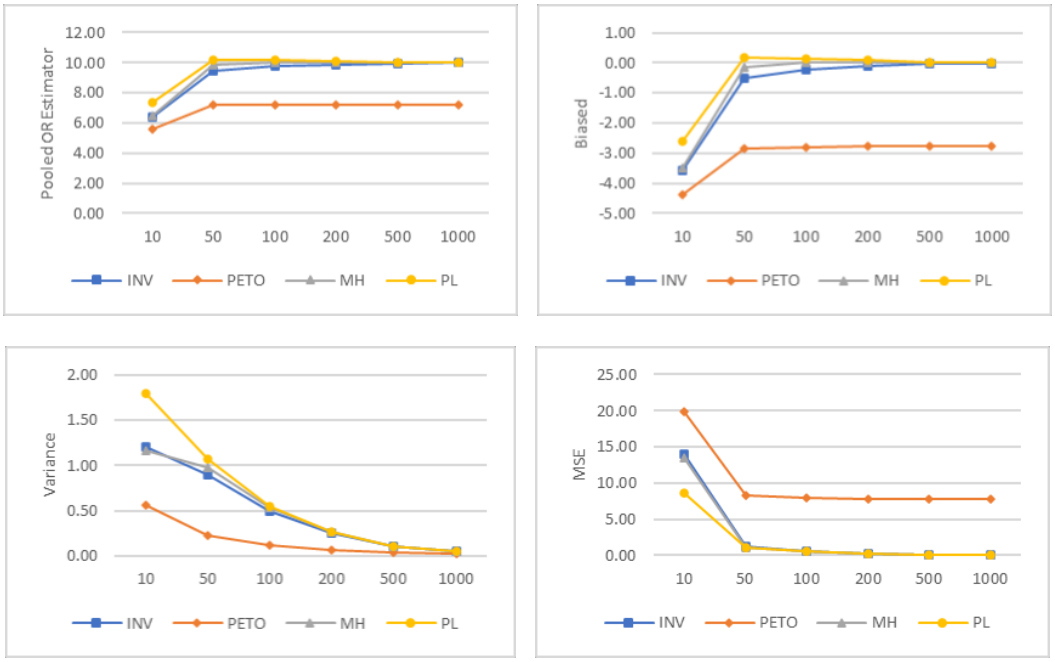


(b) θ= 10, k = 8

Figure 4: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 10$ with balance case scenarios.(Continued)

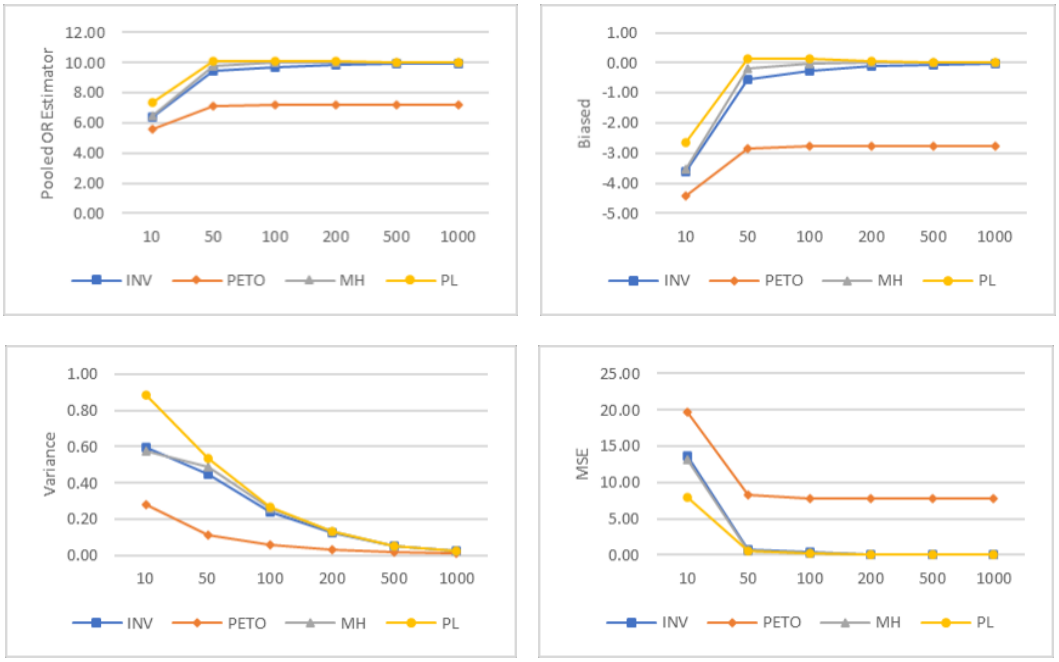


(c) θ= 10, k = 16

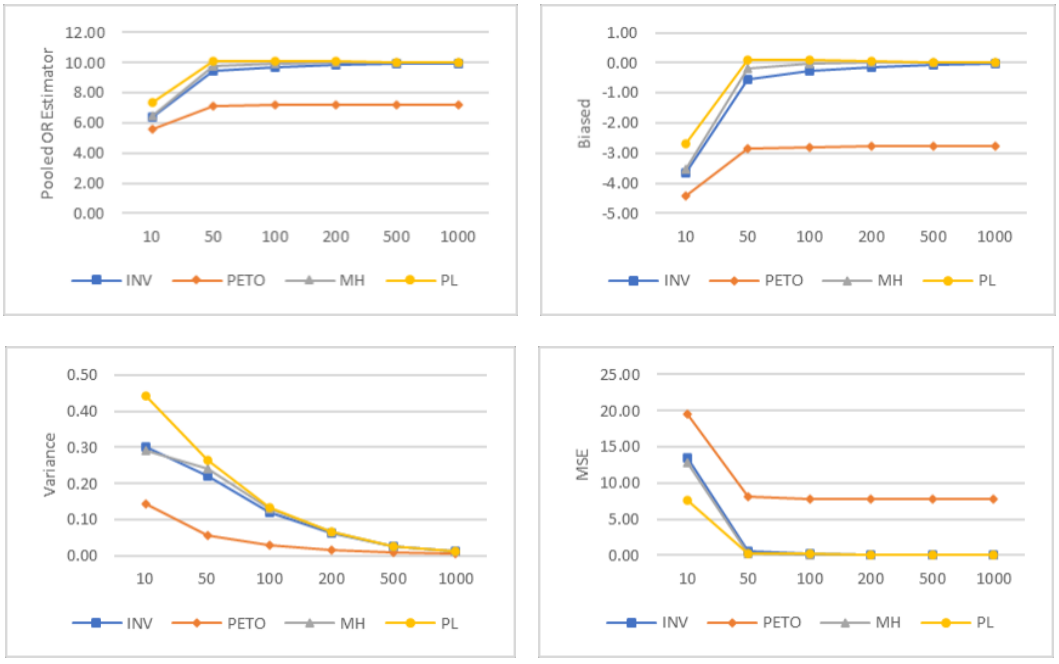


(d) θ= 10, k = 32

Figure 4: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 10$ with balance case scenarios.(Continued)



(e) θ= 10, k = 64



(f) θ= 10, k = 128

Figure 4: Comparison of the pooled odds ratio estimator, biased, variance and MSE of the four point estimators base on $\theta = 10$ with balance case scenarios.(Continued)

12. INV, under very large treatment effect $(\theta \geq 5)$, performs the best with the smallest MSE if study sizes are small to moderate $(k \leq 32)$ in combination with heavily small to large sample sizes $(n_i^T, n_i^C \geq 50)$, the INV performs the best.
13. MH, under no treatment effect $(\theta = 1)$, performs acceptably with smaller MSE, (1) when study sizes are small $(k = 4, 8)$ in combination with moderate to large sample sizes $(n_i^T, n_i^C \geq 200)$, (2) when study sizes are moderate $(k = 16, 32)$ in combination with slightly moderate to large sample sizes $(n_i^T, n_i^C \geq 100)$, (3) when study sizes are large $(k = 64, 128)$ in combination with strongly small to large sample sizes $(n_i^T, n_i^C \geq 50)$.
14. MH, under existence of treatment effect $(\theta = 2)$, performs better with smaller MSE, (1) when study sizes are small $(k = 4, 8)$ in combination with very large sample sizes $(n_i^T, n_i^C = 1000)$, (2) when study sizes are moderate $(k = 16, 32)$ in combination with large sample sizes $(n_i^T, n_i^C \geq 500)$, (3) when study sizes are large $(k = 64)$ in combination with strongly moderate to large sample sizes $\left(n_i^T, n_i^C \geq 200\right)$, (4) when study sizes are very large $(k = 128)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.
15. MH, under large treatment effect $(\theta \geq 5)$, performs better with smaller MSE, (1) when study sizes are small $(k = 4, 8)$ in combination with small and moderate to large sample sizes $(n_i^T, n_i^C = 10 \; and \; \geq 100)$, (2) when study sizes are moderate to large $(k \geq 16)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.
16. PT, under no treatment effect $(\theta = 1)$, performs the best with the smallest MSE in every situations, regardless of study sizes $(k = 4, 8, 16, 32, 64, 128)$, and regardless of sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.
17. PT, under existence of treatment effect $(\theta = 2)$, works well in most situations with (1) small to heavily moderate study sizes $(k = 4, 8, 16, 32)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$, (2) when study sizes are large $(k = 64)$ in combination with strongly moderate to large sample sizes $(n_i^T, n_i^C \geq 200)$, (3) when study sizes are very large $(k = 128)$ in combination with slightly moderate to large sample sizes $(n_i^T, n_i^C \geq 100)$.
18. PT, under large treatment effect $(\theta \geq 5)$, surprisingly can perform well when study size is small $(k \leq 8)$ in combination with small sample sizes $\left(n_i^T, n_i^C = 10, 50\right)$. However, PT is not recommended to use since it provides the worst performance and produces large bias and large standard error.

4 Discussion

Referring to the gap of study, we wish to have the guideline to indicate which a particular estimator is most appropriate to use in different situations; especially, we don't yet know about the performance of the profile likelihood method (PL).

PL, for none and mild effect sizes $(\theta = 1, 2)$, can be preferred when study sizes are large $(k \geq 64)$ in combination with moderate to large sample size $(n_i^T, n_i^C \geq 100)$. For large effect sizes$(\theta \geq 5)$, PL performs best practically with smallest MSE when study sizes are large $(k \geq 64)$ and at every sample sizes $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.

INV, under no treatment effect $(\theta = 1)$, performs the best with the smallest MSE in every situation, regardless of study sizes $(k = 4, 8, 16, 32, 64, 128)$, and regardless of sample

size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$. For existence of treatment effect $(\theta = 2)$, INV still performs the best with the smallest MSE in most situations with small to heavily moderate study sizes $(k = 4, 8, 16, 32)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$. Also, under very large treatment effect $(\theta \geq 5)$, INV performs the best with the smallest MSE if study sizes are small to moderate $(k \leq 32)$ in combination with heavily small to large sample sizes $(n_i^T, n_i^C \geq 50)$, the INV performs the best.

MH, under no treatment effect $(\theta = 1)$, performs acceptably, (1) when study sizes are small $(k = 4, 8)$ in combination with moderate to large sample sizes $(n_i^T, n_i^C \geq 200)$, (2) when study sizes are moderate $(k = 16, 32)$ in combination with slightly moderate to large sample sizes $(n_i^T, n_i^C \geq 100)$, (3) when study sizes are large $(k = 64, 128)$ in combination with strongly small to large sample sizes $(n_i^T, n_i^C \geq 50)$. For existence of treatment effect $(\theta = 2)$, MH, performs better with smaller MSE, (1) when study sizes are small $(k = 4, 8)$ in combination with very large sample sizes $(n_i^T, n_i^C = 1000)$, (2) when study sizes are moderate $(k = 16, 32)$ in combination with large sample sizes $(n_i^T, n_i^C \geq 500)$, (3) when study sizes are large $(k = 64)$ in combination with strongly moderate to large sample sizes $\left(n_i^T, n_i^C \geq 200\right)$, (4) when study sizes are very large $(k = 128)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$. Under large treatment effect $(\theta \geq 5)$, MH performs acceptably, (1) when study sizes are small $(k = 4, 8)$ in combination with small and moderate to large sample sizes $(n_i^T, n_i^C = 10 \; and \; \geq 100)$, (2) when study sizes are moderate to large $(k \geq 16)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$.

PT, under no treatment effect $(\theta = 1)$, performs the best with the smallest MSE in every situation, regardless of study sizes $(k = 4, 8, 16, 32, 64, 128)$, and regardless of sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$. Under existence of treatment effect $(\theta = 2)$, PT, works well in most situations with (1) small to heavily moderate study sizes $(k = 4, 8, 16, 32)$ and at every sample size $(n_i^T, n_i^C = 10, 50, 100, 200, 500, 1000)$, (2) when study sizes are large $(k = 64)$ in combination with strongly moderate to large sample sizes $(n_i^T, n_i^C \geq 200)$, (3) when study sizes are very large $(k = 128)$ in combination with slightly moderate to large sample sizes $(n_i^T, n_i^C \geq 100)$. For large treatment effect $(\theta \geq 5)$, surprisingly PT can perform well when study size is small $(k \leq 8)$ in combination with small sample sizes $\left(n_i^T, n_i^C = 10, 50\right)$. However, PT is not recommended to use since it provides the worst performance and produces large bias and large standard error.

Limitation

The limitation of our method is that it is limited to odd ratios. An extension to other effect measures, such as risk ratios, risk difference, and mean difference, might be interesting to pursue in future work; especially, PL is an estimator which we don't know much about it. However, for the case of sparse data, the difference between odds and risks becomes negligible, and the odds ratio and the risk ratio are quite similar in interpretation as effect sizes for practical purposes.

Recommendation

To provide recommendations regarding the optimal approach in more variable scenarios, future work could focus on comparing in simulations the fixed effects model with the random-effects method, various approaches to Bayesian approach, many kinds of bootstrap approach, rare events and sparse data situations.

The discussion interprets the findings, connects them to existing research, and highlights their significance for the field. It should also acknowledge the study's limitations and propose directions for future research, emphasizing the broader impact of the results.

References

[1] Cochran, W.G.,Problem arising in the analysis of a series of similar experiments. J.Roy. Statist. Soc. 4: (Supplement) (1937) 102-18.

[2] N. Mantel, W. Haenszel, Statistical aspects of the analysis of data from retrospective studies of disease. J. Nat. Cancer Inst. 22 (1959) 719-48.

[3] R. Peto , M.C. Pike, P. Armitage, N.E. Breslow, D.R. Cox, S.V. Howard, N. Mantel, K. McPherson, J. Peto, P.G. Smith, Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II: Analysis and examples. Br. J. Cancer 35 (1977) 1-39.

[4] D. Böhning , R. Kuhnert, S. Rattanasiri, Meta-Analysis of Binary data using profile likelihood. New York: Chapman & Hall/CRC Press. (2008).

[5] S. Greenland, A. Salvan, Bias in the one-step method for pooling study results. Stat. Med. 9 (1990) 247-52.

[6] J.L. Fleiss, Statistical Methods for Rates and Proportions. 2nd ed. New York: Wiley.(1981)

[7] J. Sun, L. Zhang, J. Cui, S. Li, H. Lu, Y. Zhang, J. Sun, Z. Baloch, Effect of dietary intervention treatment on children with iron deficiency anemia in China: a meta-analysis. Published online 2018 May 10. doi: 10.1186/s12944-018-0749-x. (2018).

[8] Y. Yang, E. Zhang, J. Zhang, S. Chen, G. Yu, X. Liu, H. Shao,Relationship between occupational noise exposure and the risk factors of cardiovascular disease in China. Medicine. (2018) 1-7.

[9] L. Cao, Assessment of thyroid cancer risk in more than 334,000 patients with inflammatory bowel disease: a case-control study and a meta-analysis. World Journal of Surgical Oncology, (2018) 1-9.

[10] Y. -I. Ma, H. Yao, W. -j. Yang, X. -x. Ren, L. Teng, & M. -c. Yang, Correlation between tradition Chinese medicine constitution and dyslipidemia: a systematic review and meta-analysis. Evidence-Based Complementary and Alternative Medicine, (2017) 1-7.

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

The Properties of Function for Searching the Bound of Nieuwland Constant in Rupert's Problem

Patid Srikirinth[1,*] and Wacharin Wichiramala[1]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University Phayathai Rd., Pathumwan, Bangkok 10330 Thailand

Abstract

A subset of $\mathbb{R}^3$ is defined as a solid if it is convex, compact, and has a non-empty interior. We say that a solid $\mathcal{S}$ has the Rupert property if a congruent copy $\mathcal{S}$ can pass directly through $\mathcal{S}$ (allow only translation). The Nieuwland constant of a convex solid $\mathcal{S}$, denoted by $\nu(\mathcal{S})$, is roughly defined as the largest possible enlargement of a congruent copy of $\mathcal{S}$ compared with $\mathcal{S}$ that can pass directly through $\mathcal{S}$. We can observe that $\nu(\mathcal{S}) > 1$ if and only if $\mathcal{S}$ has the Rupert property. We have two main objectives for this work. Firstly, we aim to find an upper bound of $\nu(\mathcal{S})$. We found that $\nu(\mathcal{S})$ is at most the ratio of diameter of $\mathcal{S}$ to its width. However, we can do a better bound if $\mathcal{S}$ is a polyhedron. Secondly, we develop theorems about properties of the function related to Nieuwland constant that can be implemented in the Grid-search algorithm for finding the bound of $\nu(\mathcal{S})$ in the future.

Keywords: Rupert problem, Grid search, Polyhedral
2020 MSC: 52A40, 52B55

1 Introduction

Like most researchers on this topic, the introduction begins with a narrative about Prince Rupert of the Rhine, who lived in the 17th century. In his retired life, he dedicated himself to scientific inquiry and notably won a wager by betting on the fact

"*It is possible that a hole can be cut in one of two equal cubes large enough to permit the other cube to pass through*"

This fact mostly surprises people who haven't heard it before, since it is hard to imagine the picture of how it is true. Figure 1 shows the unit square (a projection of the unit cube) fits inside a regular hexagon with side $\sqrt{2/3}$ (another projection of the unit cube). In other words, Figure 1 shows that we can align a unit cube into the position that its projection is a regular hexagon with side $\sqrt{2/3}$, then we cut a unit square shape hole into it without tearing it apart. Obviously, the other unit cube can directly pass this hole by using only translation. We name this property as "the Rupert property". Since we know that a sphere does not have this property (which is easy to verify), it is natural to ask further about the Rupert property of other solids such as Tetrahedron, Octahedron, etc.

*Speaker and corresponding author (6572011023@student.chula.ac.th), supported by Chulalongkorn University Graduate School, to celebrate the 72nd Birthday Anniversary of His Majesty the King.

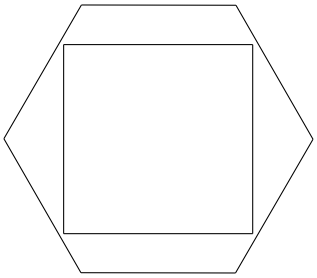


Figure 1: The unit square fits inside the regular hexagon with side length $\sqrt{2/3}$.

We say $\mathcal{S}_1$ *may pass directly through* $\mathcal{S}_2$, where $\mathcal{S}_1$ and $\mathcal{S}_2$ are solids, if there exists a shadow P of $\mathcal{S}_1$ and a shadow Q of $\mathcal{S}_2$ such that P fits in Q° (the interior of Q). A solid $\mathcal{S}$ *has the Rupert property* if $\mathcal{S}$ can pass directly through $\mathcal{S}$. The concept of the Rupert property originates from the fact that a solid $\mathcal{S}_1$ can traverse along a straight line through a solid $\mathcal{S}_2$ (allowing only translation) if and only if there exists a shadow P of $\mathcal{S}_1$ that can fit in the interior of the shadow Q of $\mathcal{S}_2$. Various studies [1–3] might describe the Rupert property in alternative ways, but the main ideas remain consistent. However, the exact definition of the Rupert property is still under debate, since "a solid could be passed through a suitable hole in another copy of it" may include traversing along a path in $\mathbb{R}^3$ or rotation. Maehara and Tokushige [4] have shown that a Tetrahedron can pass through a hole where its area is less than the area of any shadow of a Tetrahedron by allowing rotation while passing through the hole. Tetrahedron is impossible to pass through this hole if we allow only traversing along a straight line. So, this paper challenges our definition of the Rupert property.

However, another natural question to ask is "How large can the other cube that still passes through a hole of the unit cube ?". Alignment, as in Figure 1, shows the second cube can be enlarged into side length $\sqrt{6} - \sqrt{2} \approx 1.03527$. A century after the death of Rupert, Pieter Nieuwland gives a new way to cut a hole in the unit cube that increases the side length of the second cube into $3\sqrt{2}/4 \approx 1.06066$. This constant turns out to be the optimal solution later, which was proved in [5]. This concept leads to the definition of the Nieuwland constant. The Nieuwland constant of a solid $\mathcal{S}$, denoted by $\nu(\mathcal{S})$, is the largest possible scaling of a congruent copy of $\mathcal{S}$ that can pass directly through $\mathcal{S}$.

For the rest of this introduction, we provide a brief overview of this topic. In 1968, [6] showed that Tetrahedron and Octahedron have the Rupert property. In 2017, Jerrard, Wetzel, and Yuan [2] published a paper proving that all Platonic solids have the Rupert property. After this paper had been published until recently, many papers about this topic came out. There are two approaches to answering this problem, analytic traditional way and numerical way, which focus on finding the Nieuwland constant by algorithms. In 2018, Chai, Yuan, and Zamfirescu [7] provided the Rupert property of 8 of 13 Archimedean solids using an analytic approach. They also state the following conjecture.

Conjecture 1.1. ***(Chai, Yuan, Zamfirescu) [7]*** *Every convex polyhedron has the Rupert property.*

In the same year, the master thesis of Tonpho [8], an advisee of Wichiramala, used a solely numerical method to search for a lower bound of the Nieuwland constant. They got 7 of 13

Archimedean solids, 2 distinct from the previous one. Most results are a lower bound of the Nieuwland constant for each solid, not exactly the Nieuwland constant. There are 3 Archimedean solids left: Snub cube, Rhomicosidodecahedron, and Snub dodecahedron. Huber, Shultz, and Wetzel [1] showed that $n-$dimensional cube also has the Rupert property in 2018. More recently, Steininger and Sergey [9, 10] provided probabilistic and deterministic algorithms that approach Nieuwland constant in 2022. This paper conjectures on the contrary of Conjecture 1.1, since 100 million tries of Rhombicosidodecahedron, one of the Archimedean solids, found no solution.

Conjecture 1.2. ***(Steininger and Sergey) [9]*** *The Rhombicosidodecahedron does not have the Rupert property.*

Two years later, Fredriksson [11] provided a nonlinear optimization method, which showed the Rupert property of some solids other than Platonic and Archimedean solids. However, the Rupert property of 3 Archimedean solids: Snub cube, Rhomicosidodecahedron, and Snub dodecahedron, are still unknown until now.

Observe that many studies on this topic have focused on searching for a lower bound of the Nieuwland constant of a given solid. In this work, we aim to take another approach, which is searching for an upper bound of the Nieuwland constant of a given solid. This work provides two ways to obtain an upper bound without implementation. The first upper bound (Theorem 3.8) is easy to calculate. The second upper bound (Proposition 3.15) can converge to the Nieuwland constant of solid, but it may use a lot of time to compute. Implementation of these two boundaries should support how close the lower bounds of the previous works are to the Nieuwland constants.

2 Preliminaries

For the convenience of the readers, we state elementary definitions of this work here including definitions of *the Rupert property* and *the Nieuwland constant* again. People, who are familiar with these words, may skip these definitions.

A subset of $\mathbb{R}^3$ is called a *solid*, if it is convex compact with non-empty interior.

A hyperplane of $\mathbb{R}^n$ is called *supporting hyperplane* of a set $S \subseteq \mathbb{R}^n$, if S is entirely contained in one of the two closed half-spaces bounded by the hyperplane and there exists at least one $s \in \partial S$ (boundary of S) on the hyperplane. In particular, H supports the solid $\mathcal{S}$ at s, referring to H as a tangent hyperplane of $\mathcal{S}$ at s.

Diameter of a compact set S in $\mathbb{R}^n$, denoted by $\text{diam}(S)$, is the longest Euclidean distance between elements in S.

Width of a compact set S in $\mathbb{R}^n$, denoted by $\text{width}(S)$, is the shortest Euclidean distance between two distinct parallel supporting hyperplanes of S.

Width in the direction θ of a compact set S in $\mathbb{R}^n$, denoted by $w_\theta(S)$, is the Euclidean distance between two distinct parallel supporting lines of S, at the angle $\theta + \pi/2$ around the $X-$axis. Note that $\text{width}(S) = \min_{\theta \in [0,\pi)} w_\theta(S)$.

Area of a compact set S in $\mathbb{R}^2$, denoted by $\text{area}(S)$, Note that area is defined as a function by the measure axioms of region size.

Length of perimeter of a compact set S in $\mathbb{R}^2$, denoted by $\text{perim}(S)$, is defined by

$$\text{perim}(S) := \sup\left\{\sum_{i=1}^{n} d(\rho(t_{i-1}), \rho(t_i)) : 0 = t_0 \leq \cdots \leq t_n = 1 \text{ and } n \in \mathbb{N}\right\},$$

where the path $\rho : [0,1] \to \mathbb{R}^2$ is the path representing ∂S.
We note that $\text{area}, \text{diam}, \text{width}, \text{w}_\theta$ and perim, are all well-defined functions.

$R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation in $\mathbb{R}^2$ around the origin by an angle θ.

$R_{\beta,\phi} : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear function that rotates around the $X-$axis by an angle β and then rotates around the $Y-$axis by an angle $-\phi$.

$\pi : \mathbb{R}^3 \to \mathbb{R}^2$ be the standard projection that is $(x, y, z) \mapsto (x, y)$.
In particular, for a solid $\mathcal{S}$, then the image $R_\theta \circ \pi \circ R_{\beta,\phi}(\mathcal{S})$ is

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix} \mathcal{S}.$$

A compact set in $\mathbb{R}^2$ is called *shadow* of a solid $\mathcal{S}$ if it is a projection of $\mathcal{S}$ into $\mathbb{R}^2$. In particular, a shadow P of $\mathcal{S}$ can be written in the form $P = R_\theta \circ \pi \circ R_{\beta,\phi}(\mathcal{S}) + (x, y)$ for some $\theta, \beta, \phi, x, y$. The set $\text{Shad}(\mathcal{S})$ is defined as the set of all shadows of $\mathcal{S}$.

Given compact subsets of $\mathbb{R}^2$ P and Q. We say P *fits in* Q, if there exist θ, x, y such that $R_\theta(P) + (x, y)$.

Given solids $\mathcal{S}_1$ and $\mathcal{S}_2$. We say $\mathcal{S}_1$ *(can) pass directly through* $\mathcal{S}_2$, if there exists a shadow P of $\mathcal{S}_1$ and a shadow Q of $\mathcal{S}_2$ such that P fits in Q° (interior of Q).

A solid $\mathcal{S}$ *has the Rupert property* if $\mathcal{S}$ can pass directly through $\mathcal{S}$.

The *Nieuwland constant* of a solid $\mathcal{S}$ is defined as

$$\nu(\mathcal{S}) := \sup\{\alpha > 0 : \alpha\mathcal{S} \text{ can pass directly through } \mathcal{S}\}.$$

Equivalently,

$$\nu(\mathcal{S}) = \sup_{\substack{P,Q\in\text{Shad}(\mathcal{S}),\\ \theta\in[0,2\pi),(x,y)\in\mathbb{R}^2}} \{\alpha > 0 : \alpha R_\theta(P) + (x, y) \subseteq Q^\circ\}.$$

Remark 2.1. Let $\mathcal{S}$ be a solid. Then $\nu(\mathcal{S}) > 1$ if and only if $\mathcal{S}$ has the Rupert property. Moreover, $\nu(\mathcal{S})\mathcal{S}$ cannot pass directly through $\mathcal{S}$.

Remark 2.1 is straightforward to understand, but it demonstrates the connection between the Rupert property and the computable value such as the Nieuwland constant. This is the reason why we focus on studying Nieuwland constants.

2.1 Fitting two compact sets in $\mathbb{R}^2$

As in the elementary definition at the introduction of Section 2, Rupert property depends on existing non-empty interior compact subsets P and Q of $\mathbb{R}^2$, such that P fits in Q°. So, we gather some necessary conditions for P fits in Q° in Lemma 2.2 and Proposition 2.3.

Lemma 2.2. *Let P and Q be compact subsets of $\mathbb{R}^2$ such that $P \subseteq Q^\circ$ and $Q^\circ \neq \emptyset$. Let θ be an angle. Then*

$$w_\theta(P) < w_\theta(Q).$$

Proof. This lemma is obvious, so we skip the proof. □

Proposition 2.3. *Let P, Q be compact subsets of $\mathbb{R}^2$ such that P fits in Q°, and $P^\circ, Q^\circ \neq \emptyset$. The following are necessary:*

(i) $\operatorname{area}(P) < \operatorname{area}(Q)$;

(ii) $\operatorname{diam}(P) < \operatorname{diam}(Q)$;

(iii) $\operatorname{width}(P) < \operatorname{width}(Q)$;

(iv) If P is convex, $\operatorname{perim}(P) < \operatorname{perim}(Q)$.

Proof. Since the area, diameter, width, and boundary length are invariant under rotation and translation, we assume $P \subseteq Q^\circ$ instead of P fits in Q°. The statement (i) is obvious because the area is a measure. The statements (ii) and (iii) are follows easily from Lemma 2.2. For (iv), the case $\operatorname{perim}(P) \leq \operatorname{perim}(Q)$, which $P \subseteq Q$, of statement (iv) is proved by lemma 2.1 in [13]. Note that convexity of P is used in this step. It also has a proof without words in [9] for the case P, Q are polygons. Assume $P \subseteq Q^\circ$, without loss of generality $(0,0) \in P$. We can find a small enough $\epsilon > 0$ such that $(1+\epsilon)P \subseteq Q$. Note that $\operatorname{perim}((1+\epsilon)P) = (1+\epsilon)\operatorname{perim}(P)$. This implies $\operatorname{perim}(P) < \operatorname{perim}((1+\epsilon)P)$. Since $(1+\epsilon)P \subseteq Q$, $\operatorname{perim}((1+\epsilon)P) \leq \operatorname{perim}(Q)$. Hence, $\operatorname{perim}(P) < \operatorname{perim}(Q)$. □

Note that if P and Q are polygons, all of the conditions in 2.3 are quite easy to check. So, checking these conditions should be done before trying to verify that P fits in Q°.

Lemma 2.4 is stated here without proof since this fact has appeared from time to time as mentioned in [12].

Lemma 2.4. *Let S be a polygon. There is a vertex v of S and an edge e of S such that* $\operatorname{width}(S)$, *measuring between v and the line containing e, exists.*

2.2 Objective function ($\mathcal{A}_\mathcal{S}$)

As we focus on the Nieuwland constant of polyhedrons, we define a function in which the Nieuwland constant is its optimal(supremum) value.

Definition 2.5. Given two convex compact sets $P, Q \subseteq \mathbb{R}^2$ which contain the origin, an angle $\theta \in [0, 2\pi)$, and a coordinate $(x, y) \in \mathbb{R}^2$. If $(x, y) \in Q$, we can scale (shrink or enlarge) P by a constant $k > 0$ to satisfy the condition

$$kR_\theta(P) + (x, y) \subseteq Q.$$

The largest possible constant k satisfies this condition, called α. So, we define $\mathcal{A}$ by

$$\mathcal{A}(P, Q, \theta, x, y) := \begin{cases} \alpha & \text{if } (x, y) \in Q \\ 0 & \text{otherwise.} \end{cases}$$

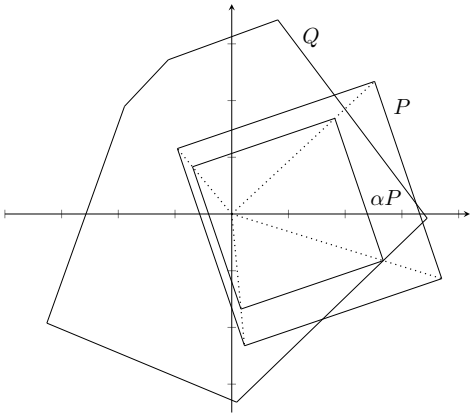


Figure 2: P, Q, and $\mathcal{A}(P, Q, \theta, x, y)P$ as αP.

This function represents "The size of P that is maximal inside Q" as shown in Figure 2.

Definition 2.6. For a solid $\mathcal{S}$, we define $\mathcal{A}_\mathcal{S} : \mathbb{R}^7 \to \mathbb{R}$ by

$$\mathcal{A}_\mathcal{S}(\phi_1, \beta_1, \phi_2, \beta_2, \theta, x, y) := \mathcal{A}(P_\mathcal{S}, Q_\mathcal{S}, \theta, x, y),$$

where $P_\mathcal{S} = \pi \circ R_{\beta_1,\phi_1}(\mathcal{S})$, and $Q_\mathcal{S} = \pi \circ R_{\beta_2,\phi_2}(\mathcal{S})$. Note that $P_\mathcal{S}, Q_\mathcal{S}$ are shadows of $\mathcal{S}$. For convenience, we usually write P, Q instead of $P_\mathcal{S}, Q_\mathcal{S}$.

We can separate parameters of $\mathcal{A}_\mathcal{S}$ into three types.

1. $(\phi_1, \beta_1, \phi_2, \beta_2)$: these represent the rotation of $\mathcal{S}$ in $\mathbb{R}^3$ to get the shadow $P_\mathcal{S}, Q_\mathcal{S}$.
2. (θ): this represents the rotation of the shadow $P_\mathcal{S}$ in $\mathbb{R}^2$.
3. (x, y): this represents the translation of the shadow $P_\mathcal{S}$.

The value of $\mathcal{A}_\mathcal{S}$ represents the largest scale of a copy of $\mathcal{S}$ (depending on ϕ_1, β_1) that can pass through another copy (depending on ϕ_2, β_2) by using this alignment (depending on θ, x, y).

Proposition 2.7. *A solid $\mathcal{S}$ has the Rupert property if and only if there exist $\phi_1, \beta_1, \phi_2, \beta_2, \theta$, x, y that $\mathcal{A}_\mathcal{S}(\phi_1, \beta_1, \phi_2, \beta_2, \theta, x, y) > 1$, Moreover, $\nu(\mathcal{S})$ is the supremum of $\mathcal{A}_\mathcal{S}$.*

Proposition 2.7 is clear by the definitions of the Rupert property, $\mathcal{A}_\mathcal{S}$, and the Nieuwland constant. For more details, if $\mathcal{A}_\mathcal{S}$ gives an output greater than 1, then $\mathcal{S}$ and its copy can adjust their position to pass through itself directly.

Through the end of this subsection, we assume our solid $\mathcal{S}$ is a convex polyhedron containing the origin. So its shadows are convex polygons. As in the Definition of 2.6, it needs to check that the point (x, y) is contained in the convex polygon. We verify this property by using Lemma 2.8.

Lemma 2.8. *Let $(q_1, q_2, \ldots, q_n)$ be the sequence of vertices of convex polygon Q ordered in counter-clockwise direction $(q_{n+1} := q_1)$, and a point $p \in \mathbb{R}^2$.*

$$p \in Q^\circ \quad \text{if and only if} \quad \overrightarrow{q_i q_{i+1}} \otimes \overrightarrow{q_i p} > 0 \text{ for all } i \in \{1, 2, \ldots, n\},$$

where $(a_1, a_2) \otimes (b_1, b_2) := a_1 b_2 - a_2 b_1$.

This lemma is a well-known fact in computational geometry; people who may not be familiar may take a look at Theorem 2.2 in the textbook Computational Geometry: An Introduction by F.P. Preparata and M.I. Shamos [14]. In other words, this lemma is telling that a point p is contained inside a polygon Q, if and only if p is on the left-hand side of the ray $\overrightarrow{q_i q_{i+1}}$ for all $i \in \{1, 2, \ldots, n\}$. To calculate $\mathcal{A}(P, Q, \theta, x, y)$, we use Lemma 2.8 and algorithm constructed in the appendix of [8].

3 Main Results

We divide our main results into two parts following our objective as we mentioned in the abstract. The first part suggests an easy way to calculate an upper bound of $\nu(\mathcal{S})$ where $\mathcal{S}$ is a convex polyhedron. However, the upper bound, provided from the first part, is not approaching $\nu(\mathcal{S})$. So, we developed our understanding of the continuity of $\mathcal{A}_\mathcal{S}$ that we intend to use in finding an upper bound approaching $\nu(\mathcal{S})$ in the second part.

3.1 An upper bound of the Nieuwland constant

The results of this part mainly originate from Proposition 2.3. Then, we try to find an upper bound of $\nu(\mathcal{S})$ in some special cases. Before anything else, let us introduce another definition that we use quite often in this part.

Definition 3.1. Let $\mathcal{S}$ be a solid, and $f : C \to \mathbb{R}$ where C is the collection of compact subsets of $\mathbb{R}^2$. We denote;

$$f^*(\mathcal{S}) = \sup\{f(S) : S \in \mathrm{Shad}(\mathcal{S})\},$$
$$f_*(\mathcal{S}) = \inf\{f(S) : S \in \mathrm{Shad}(\mathcal{S})\}.$$

In this work, we treat area and perim as well-defined functions from the collection of compact subsets of $\mathbb{R}^2$ to $\mathbb{R}$. The Definition 3.1 mainly intends to be used f as diameter, width, perimeter, or area of compact subsets of $\mathbb{R}^2$. For example, a given solid $\mathcal{S}$, diam_* means the smallest diameter of shadows of $\mathcal{S}$. It is clear that $\mathrm{diam}^*(\mathcal{S}) = \mathrm{diam}(\mathcal{S})$, and $\mathrm{width}_*(\mathcal{S}) = \mathrm{width}(\mathcal{S})$.

Theorem 3.2. *Let $\mathcal{S}$ be a solid. Then*

$$\nu(\mathcal{S}) \leq \min\left\{\frac{\mathrm{diam}(\mathcal{S})}{\mathrm{diam}_*(\mathcal{S})}, \frac{\mathrm{width}^*(\mathcal{S})}{\mathrm{width}(\mathcal{S})}, \frac{\mathrm{perim}^*(\mathcal{S})}{\mathrm{perim}_*(\mathcal{S})}, \sqrt{\frac{\mathrm{area}^*(\mathcal{S})}{\mathrm{area}_*(\mathcal{S})}}\right\}. \tag{3.1}$$

Proof. Denote $\nu := \nu(\mathcal{S})$. By definition of ν, the set $(\nu - \epsilon)\mathcal{S}$ can pass directly through $\mathcal{S}$, for all $\epsilon \in (0, \nu)$. So, we can find shadows P, Q of $\mathcal{S}$ such that $(\nu - \epsilon)P \subseteq Q^\circ$. Let $f \in \{\mathrm{diam}, \mathrm{width}, \mathrm{perim}, \sqrt{\mathrm{area}}\}$. By Proposition 2.3, we get that $f((\nu - \epsilon)P) < f(Q)$. Hence,

$$\nu - \epsilon < \frac{f(Q)}{f(P)} \leq \frac{f^*(\mathcal{S})}{f(P)} \leq \frac{f^*(\mathcal{S})}{f_*(\mathcal{S})}.$$

Since ϵ is arbitrary in $(0, \nu)$, we get $\nu(\mathcal{S}) \leq \min\left\{\frac{\mathrm{diam}(\mathcal{S})}{\mathrm{diam}_*(\mathcal{S})}, \frac{\mathrm{width}^*(\mathcal{S})}{\mathrm{width}(\mathcal{S})}, \frac{\mathrm{perim}^*(\mathcal{S})}{\mathrm{perim}_*(\mathcal{S})}, \sqrt{\frac{\mathrm{area}^*(\mathcal{S})}{\mathrm{area}_*(\mathcal{S})}}\right\}$. □

Corollary 3.3. *Let $\mathcal{S}$ be a solid. Then $\nu(\mathcal{S}) \leq \operatorname{diam}(\mathcal{S})/\operatorname{width}(\mathcal{S})$.*

This corollary is straightforward from Theorem 3.2, $\nu(\mathcal{S}) \leq \operatorname{diam}(\mathcal{S})/\operatorname{diam}_*(\mathcal{S})$, with the fact that $\operatorname{diam}_*(\mathcal{S}) \geq \operatorname{width}(\mathcal{S})$.

However, we are interested in more special cases such as $\mathcal{S}$ is a polyhedral (or right prism). We can tell more about its $\operatorname{diam}_*(\mathcal{S})$ and $\operatorname{width}^*(\mathcal{S})$.

Readers, who want a challenge, may consider the following fact: Let $\mathcal{S}$ be a right prism (or cylinder) in which the base is a polygon (circle) with diameter k. If the height h of $\mathcal{S}$ is k or higher, then

$$\operatorname{diam}_*(\mathcal{S}) = \operatorname{width}^*(\mathcal{S}) = k. \tag{3.2}$$

In the case, $\mathcal{S}$ is polyhedral, we can consider its vertices, edges, and faces. Through the end of this subsection, we denote sets of vertices, edges, and faces of the mentioned polyhedral by V, E, and F. Let Λ be the set of pairs of parallel distinct support planes of $\mathcal{S}$. Define

$$V/E := \{(M, N) \in \Lambda : \text{planes } M, N \text{ support } \mathcal{S} \text{ at some } u \in V, e \in E \text{ respectively}\}.$$

We define $V/V, V/F, E/E, E/F$, and F/F in the same manner. For convenience, we also write the set of Euclidean distances $d(X) = \{d(x, y) : (x, y) \in X\}$.

Remark 3.4. Observe that

$$F/F \subseteq E/F \begin{matrix} \subsetneqq \\ \subsetneqq \end{matrix} V/F \begin{matrix} \subsetneqq \\ \subsetneqq \end{matrix} V/E \subseteq V/V.$$

In particular, V/F and E/E are both in between E/F and V/E. Moreover, $V/V = \Lambda$. It implies $\operatorname{diam}(\mathcal{S}) = \sup d(V/V)$ and $\operatorname{width}(\mathcal{S}) = \inf d(V/V)$.

Lemma 3.5. *Let $l_1, l_2 \subseteq \mathbb{R}^2$ be distinct parallel lines, and $p_1 \in l_1, p_2 \in l_2$. Let $\phi < \frac{\pi}{2}$ be the angle between $\overline{p_1p_2}$ and l_1. For a given angle θ such that $0 < \theta \leq \phi$. We get l_1' from rotating l_1 around p_1 toward $\overline{p_1p_2}$ by angle θ. Also, l_2' in the same manner (rotating in the same direction). Then $d(l_1, l_2) \geq d(l_1', l_2')$.*

Proof. Proof of this lemma follows some elementary algebra steps, so we omit it. However, a geometric example of this lemma is shown in Figure 3. □

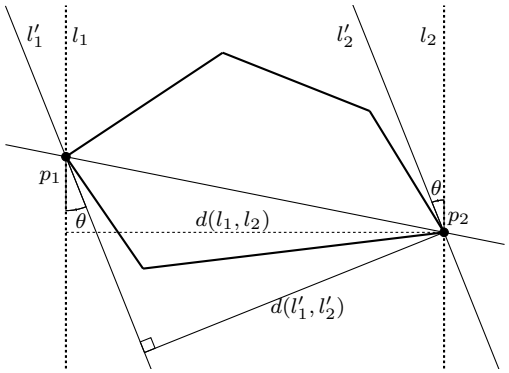


Figure 3: A geometric example for Lemma 3.5.

Lemma 3.6. *Let $\mathcal{S}$ be a polyhedron. Then*

$$\text{width}(\mathcal{S}) = \inf d(V/V) = \inf d(V/E) = \inf d(V/F \cup E/E).$$

The proof of this lemma is omitted since this lemma follows from the fact about the width of the convex polyhedrons as appeared in [12]. Readers, who might be interested, may try to prove by considering shadows of $\mathcal{S}$ and using Lemma 3.5.

Proposition 3.7. *Let $\mathcal{S}$ be a convex polygon.*

$$\begin{aligned} \text{diam}_*(\mathcal{S}) &\geq \inf d\left((E/E)\backslash(V/F)\right), \\ \text{width}^*(\mathcal{S}) &\leq \sup d(V/E). \end{aligned}$$

Proof. Firstly, we show $\text{diam}_*(\mathcal{S}) \geq \inf d\left((E/E)\backslash(V/F)\right)$. Claim that we can choose $(M, N) \in \Lambda$ such that $d(M, N) = \text{diam}_*(\mathcal{S})$. Since any shadows of $\mathcal{S}$ can be obtained from rotate $\mathcal{S}$ around $X-$axis by an angle β then rotate around $Y-$axis by an angle ϕ, shadows of $\mathcal{S}$ can be represented by $(\beta, \phi) \in [0, 2\pi]^2$. Compactness of $[0, 2\pi]^2$ implies that we can choose a shadow S of $\mathcal{S}$ such that $\text{diam}(S) = \text{diam}_*(\mathcal{S})$. We assume $\mathcal{S}$ aligned in the $\mathbb{R}^3$ in the way that its projection is S. Pick $m, n \in \mathcal{S}$ such that $\text{diam}(S) = d(\pi(m), \pi(n))$. We can find the parallel lines l_m, l_n support S at $\pi(m), \pi(n)$ and $d(l_m, l_n) = \text{diam}(S)$. The lines l_m, l_n form $(M, N) \in \Lambda$ such that projection of M, N are l_m, l_n respectively. This choice of M, N guarantees that M, N support $\mathcal{S}$ at m, n respectively, and perpendicular to $XY-$plane. So $d(M, N) = \text{diam}_*(\mathcal{S})$. Since $d(\pi(m), \pi(n)) = \text{diam}(S)$, we get that $\pi(m), \pi(n)$ are vertices of S. It implies that m, n can not be contained in the interior of any faces of $\mathcal{S}$. This also implies that (M, N) cannot be contained in V/F. So, $d(M, N) \geq \inf d\left((V/V)\backslash(V/F)\right)$. Lemma 3.6 implies that for any $(P, Q) \in V/V$, we can find $(P^*, Q^*) \in V/F \cup E/E$ such that $d(P^*, Q^*) \leq d(P, Q)$. So, $\text{diam}_*(\mathcal{S}) = d(M, N) \geq \inf d\left((V/V)\backslash(V/F)\right) = \inf d\left((E/E)\backslash(V/F)\right)$.

Secondly, we show $\text{width}^*(\mathcal{S}) \leq \sup d(V/E)$. By a similar argument as in the first step, we obtain a shadow S of $\mathcal{S}$ such that $\text{width}(S) = \text{width}^*(\mathcal{S})$. Lemma 2.4 guarantees that we can pick points $p, q \in \mathcal{S}$ such that $d(\pi(p), \pi(q)) = \text{width}(S)$, $\pi(p)$ is a vertex of S, and $\pi(q)$ is in an edge of S. We construct lines $l_p, l_q \subseteq \mathbb{R}^2$ and $(P, Q) \in \Lambda$ in the same manner as the first step so $d(P, Q) = d(l_p, l_q) = \text{width}^*(\mathcal{S})$. Since $\pi(q)$ is in an edge of S, Q has to support $\mathcal{S}$ at an edge or a face of $\mathcal{S}$. We get $(P, Q) \in V/E$. Hence, $\text{width}^*(\mathcal{S}) = d(P, Q) \leq \sup d(V/E)$. □

Theorem 3.8. *Let $\mathcal{S}$ be a polyhedral. Then*

$$\nu(\mathcal{S}) \leq \min\left\{\frac{\text{diam}(\mathcal{S})}{\inf d\left((E/E)\backslash(V/F)\right)}, \frac{\sup d(V/E)}{\text{width}(\mathcal{S})}\right\}.$$

Moreover, if S is a right prism (or cylinder) in which the base is a polygon (circle) with diameter k and the height h of $\mathcal{S}$ is k or higher. We can replace $\inf d\left((E/E)\backslash(V/F)\right)$ *and* $\sup d(V/E)$ *by k instead.*

This theorem is a result of Theorem 3.2, with Proposition 3.7 and the Equation 3.2, so we skip the proof. However, to find $\inf d\left((E/E)\backslash(V/F)\right)$ and $\sup d(V/E)$ of a given solid, we claim that some modifications on *Rotating Caliper Algorithm* in [12] can be used. Observe that

$$\begin{aligned} \inf d\left((E/E)\backslash(V/F)\right) \geq \text{width}(\mathcal{S}) &\Rightarrow \frac{\text{diam}(\mathcal{S})}{\inf d\left((E/E)\backslash(V/F)\right)} \leq \frac{\text{diam}(\mathcal{S})}{\text{width}(\mathcal{S})}, \\ \sup d(V/E) \leq \text{diam}(\mathcal{S}) &\Rightarrow \frac{\sup d(V/E)}{\text{width}(\mathcal{S})} \leq \frac{\text{diam}(\mathcal{S})}{\text{width}(\mathcal{S})}. \end{aligned}$$

So, compared with Corollary 3.3, this theorem is given a better upper bound.

3.2 Analyzing of $\mathcal{A}_{\mathcal{S}}$

Our objective for this subsection is to study the function $\mathcal{A}_{\mathcal{S}}$ for a given solid $\mathcal{S}$ that we have defined in the section 2.2.

Given a convex compact set with non-empty interior $P \subseteq \mathbb{R}^n$ containing the origin, we define

$$\begin{aligned} b_P &:= \text{minimum distance between the origin and boundary of P,} \\ R_P &:= \text{the radius of the inscribed circles of } P, \\ &\quad \text{(largest circle that can be contained in } P\text{).} \end{aligned}$$

Since this subsection will frequently mention a constant b_P, we note some facts about b_P that we can see easily here:

1. $b_P \leq R_P$.
2. We can translate P until $b_P = R_P$.
3. If P is a shadow of $\mathcal{S}$, then $R_P \geq R_{\mathcal{S}}$, where $R_{\mathcal{S}}$ is the radius of the inscribed spheres of $\mathcal{S}$.

In Proposition 3.9, we try to analyze $\mathcal{A}_{\mathcal{S}}$ when fixing all other variables except translation (x, y).

Proposition 3.9. *Let $P, Q \subseteq \mathbb{R}^2$ be convex compact sets with non-empty interior containing the origin. Fix $\theta \in [0, 2\pi)$, and $q_0, q_1 \in Q$. Denote $\vec{\delta} = q_1 - q_0$, $\delta = \|\vec{\delta}\|$, $\alpha_0 = \mathcal{A}(P, Q, \theta, q_0)$, and $\alpha_1 = \mathcal{A}(P, Q, \theta, q_1)$. If $\alpha_0 b_P > \delta$, then*

$$\alpha_0 - \frac{\delta}{b_P} \leq \alpha_1 \leq \alpha_0 + \frac{\delta}{b_P}.$$

Proof. Let $p' \in \left(\alpha_0 - \frac{\delta}{b_P}\right) R_\theta(P) + \vec{\delta}$. So there exists $p \in R_\theta(P)$ such that

$$p' = \left(\alpha_0 - \frac{\delta}{b_P}\right) p + \vec{\delta} = \left(\frac{\alpha_0 b_P - \delta}{\alpha_0 b_P}\right) \alpha_0 p + \left(\frac{\delta}{\alpha_0 b_P}\right) (\alpha_0 b_P)\, \widehat{\delta}.$$

Since $0 < \frac{\delta}{\alpha_0 b_P} < 1$, p' is in between $\alpha_0 p$ and $(\alpha_0 b_P)\, \widehat{\delta}$. Note that $\alpha_0 p$ and $(\alpha_0 b_P)\, \widehat{\delta}$ are contained in $\alpha_0 R_\theta(P)$. Since $\alpha_0 R_\theta(P)$ is convex, $p' \in \alpha_0 R_\theta(P)$. Now, we have

$$\left(\alpha_0 - \frac{\delta}{b_P}\right) R_\theta(P) + q_1 \subseteq \alpha_0 R_\theta(P) + q_0 \subseteq Q.$$

Hence $\alpha_0 - \frac{\delta}{b_P} \leq \alpha_1$, by the definition of α_1.

For the other, we show similarly. Let $p'' \in \left(\frac{b_P}{\alpha_0 b_P + \delta}\right) \left(\alpha_0 R_\theta(P) - \vec{\delta}\right)$. So there exists $p \in R_\theta(P)$ such that

$$p'' = \left(\frac{b_P}{\alpha_0 b_P + \delta}\right) \left(\alpha_0 p - \vec{\delta}\right) = \left(\frac{\alpha_0 b_P}{\alpha_0 b_P + \delta}\right) p + \left(\frac{\delta}{\alpha_0 b_P + \delta}\right) (-b_P) \widehat{\delta}.$$

Since $0 < \frac{\delta}{\alpha_0 b_P + \delta} < 1$, p'' is in between p and $(-b_P)\,\widehat{\delta}$ contained in $R_\theta(P)$ by a similar argument as in the above. Now, we get that

$$\left(\frac{b_P}{\alpha_0 b_P + \delta}\right)\left(\alpha_0 R_\theta(P) - \vec{\delta}\right) \subseteq R_\theta(P) \Rightarrow \alpha_0 R_\theta(P) + q_0 \subseteq \left(\alpha_0 + \frac{\delta}{b_P}\right) R_\theta(P) + q_1.$$

Suppose to the contrary, $\alpha_1 > \alpha_0 + \frac{\delta}{b_P}$. This implies

$$\alpha_0 R_\theta(P) + q_0 \subseteq \left(\alpha_0 + \frac{\delta}{b_P}\right) R_\theta(P) + q_1 \subseteq (\alpha_1 R_\theta(P) + q_1)^\circ \subseteq Q^\circ.$$

So, $\alpha_0 R_\theta(P) + p_0 \subset Q$ which contradicts the definition of α_0. □

Now, we will consider the rotation variables $(\phi_1, \beta_1, \phi_2, \beta_2, \theta)$. We consider them by using the following consideration: " How far the shadows (of the solids $\mathcal{S}$) will be changed if we change these variables ?". To represent this idea, we define the following. Let S be a subset of $\mathbb{R}^n$, and $R > 0$. We define the notation

$$[S]_R := \bigcup_{x \in S} B_R(x),$$

where $B_R(x)$ is the open ball radius R at the center x.

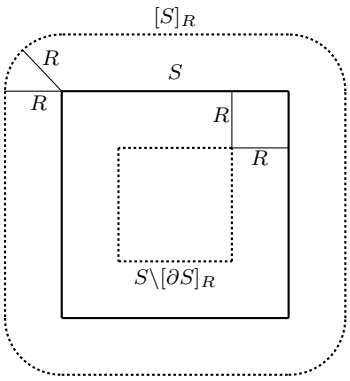


Figure 4: Illustration of $[S]_R$ and $S\backslash[\partial S]_R$ where S is a unit square.

Lemma 3.10. *Let P be a convex compact subset of $\mathbb{R}^n$ with non-empty interior. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous function such that $f(P)$ is also convex, and $R = \max\{d(f(x), x) : x \in P\}$. Then*

$$P\backslash[\partial P]_R \subseteq f(P) \subseteq \overline{[P]_R},$$

where $\overline{[P]_R}$ is the closure of $[P]_R$.

Proof. The part $f(P) \subseteq \overline{[P]_R}$ is easy to see. Since $d(f(x), x) \leq R$ for any $x \in P$, $f(x) \in \overline{B_R(x)} \subseteq \overline{[P]_R}$. So, we will show the other part. Assume $x \in P\backslash[\partial P]_R$, and suppose to the contradiction $x \notin f(P)$. Since $f(P)$ is a compact convex set, we can find a hyperplane H that separates x and $f(P)$. Now consider H with P. We can separate R^n by H into two open half-planes called R_1, R_2. Assume $x \in R_1$, notice that the compactness of P provides $\partial P \cap R_1 \neq \emptyset$.

Hence, we can find a point $y \in \partial P \subseteq R_1$ such that the vector $\overrightarrow{y-x}$ is perpendicular to H. Since $y \in \partial P$, by our assumption of this part, $d(x,y) \geq R$. Observe that the construction of H makes x on the opposite side with $f(P)$, so $f(P) \subseteq R_2$. This implies $y \in R_1$ is on the opposite side with $f(y) \in f(P) \subseteq R_2$. We also get that $d(H,y) \geq d(x,y) \geq R$, since the vector $\overrightarrow{y-x}$ is perpendicular to H. Hence,

$$\begin{aligned} d\left(f(y),y\right) &= d\left(f(y),h\right) + d(h,y) \text{ ; for some } h \in H \\ &\geq d\left(f(y),H\right) + d(H,y) \\ &\geq d\left(f(y),H\right) + R. \end{aligned}$$

Note that $f(y) \notin H$, so $d\left(f(y),y\right) \geq d\left(f(y),H\right) + R > R$. This contradicts the definition of R in the statement, so we are done. □

Proposition 3.11. *Suppose P and Q are convex compact subsets of $\mathbb{R}^2$ with non-empty interior containing the origin, an angle θ, and $(x,y) \in Q$. Given a continuous function $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that $f(P)$ is also convex. Denote $R = \max\{d\left(f(x),x\right) : x \in P\}$, $\alpha_0 = \mathcal{A}(P,Q,\theta,x,y)$, and $\alpha_1 = \mathcal{A}(f(P),Q,\theta,x,y)$. If $b_P > R$, then*

$$\left(\frac{b_P}{b_P+R}\right)\alpha_0 \leq \alpha_1 \leq \left(\frac{b_P}{b_P-R}\right)\alpha_0.$$

Proof. Observe that, Lemma 3.10 implies $P\backslash[\partial P]_R \subseteq f(P) \subseteq \overline{[P]_R}$.

We will show $\left(\frac{b_P}{b_P+R}\right)\alpha_0 \leq \alpha_1$. Assume there exists a constant $k_1 > 0$ such that $k_1\overline{[P]_R} \subseteq \alpha_0 P$. By our observation, we get that $k_1 f(P) \subseteq k_1\overline{[P]_R}$. Hence,

$$k_1 R_\theta\left(f(P)\right) + (x,y) \subseteq k_1 R_\theta\left(\overline{[P]_R}\right) + (x,y) \subseteq \alpha_0 R_\theta\left(P\right) + (x,y) \subseteq Q.$$

By the definition of α_1, $k_1 \leq \alpha_1$. It has left to show only, $\left(\frac{b_P}{b_P+R}\right)$ can be our constant k_1, $\left(\frac{b_P}{b_P+R}\right)\alpha_0\overline{[P]_R} \subseteq \alpha_0 P$. Let $p' \in \overline{[P]_R}$. If $p' \in P$, our proof is done easily by convexity of P and $0 < \frac{b_P}{b_P+R} < 1$. So, we assume $p' \in \overline{[P]_R}\backslash P$. We can write p' in the form $p' = p + (nR)\hat{p}$ for some $p \in \partial P$, and $0 < n \leq 1$. We note that $\left(\frac{b_P}{b_P+R}\right)p' = \left(\frac{b_P}{b_P+R}\right)(\|p\| + nR)\hat{p}$. By the fact $b_P \leq \|p\|$, we get that

$$0 < \left(\frac{b_P}{b_P+R}\right)(\|p\| + nR) < \left(\frac{\|p\|}{\|p\|+R}\right)(\|p\| + nR) < \|p\|.$$

This means $\left(\frac{b_P}{b_P+R}\right)p'$ is in between the origin and p. Convexity of P implies $\left(\frac{b_P}{b_P+R}\right)p' \in P$, so $\left(\frac{b_P}{b_P+R}\right)\alpha_0\overline{[P]_R} \subseteq \alpha_0 P$.

Next, we will show $\alpha_1 \leq \left(\frac{b_P}{b_P-R}\right)\alpha_0$. Assume there exists a constant $k_2 > 0$ such that $(k_2 + \epsilon)R_\theta\left(P\backslash[\partial P]_R\right) + (x,y) \nsubseteq Q$ for any $\epsilon > 0$. Again, our observation implies $(k_2+\epsilon)R_\theta\left(f(P)\right) + (x,y) \nsubseteq Q$. Hence, the definition of α_1 implies $\alpha_1 \leq k_2$. As similar as before, it has left to show only $\left(\frac{b_P}{b_P-R}\right)\alpha_0$ can be chosen as our constant k_2. By definition of α_0, we can choose a point $p \in \partial P$ such that $\alpha_0 R_\theta(p) + (x,y) \in \partial Q$ and $(x,y) \in Q^\circ$. Since $\|p\| \geq b_P > R$, we get that $p' := (\|p\| - R)\hat{p} \in P\backslash[\partial P]_R$. We note that $0 < \frac{\|p\|}{\|p\|-R} \leq \frac{b_P}{b_P-R}$, so $\left(\frac{\|p\|}{\|p\|-R}\right)\alpha_0 p' = \alpha_0 p$ is in

between the origin and $\left(\frac{b_P}{b_P-R}\right)\alpha_0 p'$. Hence, $\alpha_0 R_\theta(p)+(x,y) \in \partial Q$ is in between $(x,y) \in Q^\circ$ and $\left(\frac{b_P}{b_P-R}\right)\alpha_0 R_\theta(p')+(x,y)$. Convexity of Q implies that, for any $\epsilon > 0$, $\left(\left(\frac{b_P}{b_P-R}\right)\alpha_0+\epsilon\right) R_\theta(p')+(x,y) \notin Q$. Since $p' \in P\backslash[\partial P]_R$, our proof is done. □

Proposition 3.12. *Suppose P and Q are convex compact subsets of $\mathbb{R}^2$ with non-empty interior containing the origin, an angle θ, and $(x,y) \in Q$. Given a continuous function $g : \mathbb{R}^2 \to \mathbb{R}^2$ such that $g(Q)$ is also convex. Denote $R = \max\{d(g(x),x) : x \in Q\}$, $\alpha_0 = \mathcal{A}(P,Q,\theta,x,y)$, and $\alpha_1 = \mathcal{A}(P,g(Q),\theta,x,y)$. If $b_P > R$, then*

$$\left(\frac{b_P-R}{b_P}\right)\alpha_0 \le \alpha_1 \le \left(\frac{b_P+R}{b_P}\right)\alpha_0.$$

Proof. Since a proof of this proposition is similar to the above proposition, we provide only the sketch of the proof. For $\left(\frac{b_P-R}{b_P}\right)\alpha_0 \le \alpha_1$, we find a constant $k_1 > 0$ such that $k_1\alpha_0 P \subseteq (\alpha_0 P)\setminus[\partial(\alpha_0 P)]_R$. Hence,

$$k_1\alpha_0 R_\theta(P)+(x,y) \subseteq (\alpha_0 R_\theta(P))\setminus[\partial(\alpha_0 R_\theta(P))]_R+(x,y) \subseteq Q\backslash[\partial Q]_R \subseteq g(Q).$$

It means $k_1\alpha_0 \le \alpha_1$. Then we show, $k_1 = \frac{b_P-R}{b_P}$ met the condition. For $\alpha_1 \le \left(\frac{b_P+R}{b_P}\right)\alpha_0$, we find a constant $k_2 > 0$ such that $(k_2+\epsilon)\alpha_0 R_\theta(P)+(x,y) \nsubseteq \overline{[Q]_R}$ for any $\epsilon > 0$. This implies,

$$(k_2+\epsilon)\alpha_0 R_\theta(P)+(x,y) \nsubseteq g(Q).$$

It means $k_2\alpha_0 \ge \alpha_1$. Then we show, $k_2 = \frac{b_P+R}{b_P}$ met the condition. □

Remark 3.13. Suppose we are considering $\mathcal{A}_\mathcal{S}$ of some given solid $\mathcal{S}$. We notice that the constant b_P in Proposition 3.9, 3.11, and 3.12 is depending on the shadow P, but we can do an adjustment. As mentioned at the beginning of section 3.2, before calculating α_0 and α_1 in the propositions, we can translate P until $b_P = R_P$. So, we can change the constant b_P in these propositions into R_P by following the same proof. Moreover, we know that $R_P \ge R_\mathcal{S}$, which means changing b_P into $R_\mathcal{S}$ still satisfies all inequalities in these propositions.

Theorem 3.14. *Let $\mathcal{S}$ be a solid containing the origin. The function $\mathcal{A}_\mathcal{S}$ is continuous.*

Proof. In this proof, we define $P_\mathcal{S}$ and $Q_\mathcal{S}$ as follow;

$$P_\mathcal{S}(x_1,x_2,x_3,x_4,x_5,x_6,x_7) := R_{x_5} \circ \pi \circ R_{x_2,x_1}(\mathcal{S})+(x_6,x_7),$$
$$Q_\mathcal{S}(x_1,x_2,x_3,x_4,x_5,x_6,x_7) := \pi \circ R_{x_4,x_3}(\mathcal{S}).$$

We fix $X = (\phi_1,\beta_1,\phi_2,\beta_2,\theta,x,y)$, and assume $\Delta = (\delta_1,\delta_2\ldots,\delta_7)$. Now, we will consider $|\mathcal{A}_\mathcal{S}(X+\Delta)-\mathcal{A}_\mathcal{S}(X)|$ which equivalence with considering

$$\left|\mathcal{A}\left(P_\mathcal{S}(X+\Delta),Q_\mathcal{S}(X+\Delta),0,0,0\right)-\mathcal{A}\left(P_\mathcal{S}(X),Q_\mathcal{S}(X),0,0,0\right)\right|.$$

Since we are going to consider the case that $\|\Delta\|$ is small, we assume that it is small enough so that $(x_6+\delta_6,x_7+\delta_7) \in P_\mathcal{S}(X)$. Let $\omega \in (-\pi,\pi]$ be an angle. We denote $r_P^0(\omega)$ and $r_P^1(\omega)$ to be the length of the vector emanating from (x_6,x_7) at angle ω extending to the boundary of $P_\mathcal{S}(X)$ and $P_\mathcal{S}(X+\Delta)$ respectively. Note that $r_P^0(\omega)$ and $r_P^1(\omega)$ are well defined

since $(x_6+\delta_6, x_7+\delta_7) \in P_{\mathcal{S}}(X)$. Also, $r_Q^0(\omega)$ and $r_Q^1(\omega)$ are denoted as the length of the ray emanating from the origin at angle ω extending to the boundary of $Q_{\mathcal{S}}(X), Q_{\mathcal{S}}(X+\Delta)$ respectively. Then, we define functions f^* and g^* mapping a polar coordinate (r,ω) by;

$$f^*(r,\omega) := \left(r\left(\frac{r_P^1(\omega)}{r_P^0(\omega)}\right),\omega\right), \text{ and } g^*(r,\omega) := \left(r\left(\frac{r_Q^1(\omega)}{r_Q^0(\omega)}\right),\omega\right).$$

Then, we define functions $f, g : \mathbb{R}^2 \to \mathbb{R}^2$ by;

$$\begin{aligned} f(a,b) &:= \text{carte} \circ f^* \circ \text{polar}\left((a,b)-(x_6,x_7)\right)+(x_6,x_7), \text{ and} \\ g(a,b) &:= \text{carte} \circ g^* \circ \text{polar}\left((a,b)\right), \end{aligned}$$

where polar : $\mathbb{R}^2 \to \mathbb{R}\times(-\pi,\pi]$ is a function that changes Cartesian coordinates into polar coordinates, and carte : $\mathbb{R}\times(-\pi,\pi] \to \mathbb{R}^2$ is a function that changes polar coordinates into Cartesian coordinates. In particular, the function carte is defined by $\text{carte}(r,\omega) = (r\cos\omega, r\sin\omega)$, and the function polar is defined by

$$\text{polar}(x,y) = \begin{cases} (0,0) & \text{,if } (x,y)=(0,0) \\ (\sqrt{x^2+y^2}, \text{atan2}(y,x)) & \text{,otherwise.} \end{cases}$$

Note that the function atan2 may be known as 2−argument arctangent, and polar is designed to guarantee that f and g are well-defined. Notice that, the functions f, g are continuous and send $P_{\mathcal{S}}(X), Q_{\mathcal{S}}(X)$ into $P_{\mathcal{S}}(X+\Delta), Q_{\mathcal{S}}(X+\Delta)$ respectively, so $f(P_{\mathcal{S}}(X))$ and $g(Q_{\mathcal{S}}(X))$ are convex. Then, we define $R_f = \max\{d(f(p),p) : p \in P_{\mathcal{S}}(X)\}$ and $R_g = \max\{d(g(q),q) : q \in Q_{\mathcal{S}}(X)\}$. We claim without proof that we can find $\delta > 0$ in the way that if $\|\Delta\| \le \delta$, then R_f and R_g are as small as we wish and also $(x_6+\delta_6, x_7+\delta_7) \in P_{\mathcal{S}}(X)$. So, we assume both $R_f, R_g < R_{\mathcal{S}}$. For convenience, we denote $\mathcal{A}(P_{\mathcal{S}}(X), Q_{\mathcal{S}}(X), 0, 0, 0)$, $\mathcal{A}(P_{\mathcal{S}}(X+\Delta), Q_{\mathcal{S}}(X), 0, 0, 0)$, and $\mathcal{A}(P_{\mathcal{S}}(X+\Delta), Q_{\mathcal{S}}(X+\Delta), 0, 0, 0)$ as α_0, α_0', and α_1 respectively. Now, we apply Proposition 3.11 and Remark 3.13, we get that

$$\left(\frac{R_{\mathcal{S}}}{R_{\mathcal{S}}+R_f}\right)\alpha_0 \le \alpha_0' \le \left(\frac{R_{\mathcal{S}}}{R_{\mathcal{S}}-R_f}\right)\alpha_0.$$

Then, we apply Proposition 3.12 and Remark 3.13, we get that

$$\left(\frac{R_{\mathcal{S}}-R_g}{R_{\mathcal{S}}}\right)\left(\frac{R_{\mathcal{S}}}{R_{\mathcal{S}}+R_f}\right)\alpha_0 \le \alpha_1 \le \left(\frac{R_{\mathcal{S}}+R_g}{R_{\mathcal{S}}}\right)\left(\frac{R_{\mathcal{S}}}{R_{\mathcal{S}}-R_f}\right)\alpha_0.$$

Consider the above inequality to simplify $|\alpha_1-\alpha_0|$, we get that

$$|\mathcal{A}_{\mathcal{S}}(X+\Delta)-\mathcal{A}_{\mathcal{S}}(X)| = |\alpha_1-\alpha_0| \le \left(\frac{R_f+R_g}{R_{\mathcal{S}}-R_f}\right)\alpha_0. \tag{3.3}$$

By our claim (R_f, R_g are as small as we desire), $|\mathcal{A}_{\mathcal{S}}(X+\Delta)-\mathcal{A}_{\mathcal{S}}(X)|$ converges to 0. So, the proof is done here. □

Theorem 3.14 confirms our presumption about continuity of $\mathcal{A}_{\mathcal{S}}$. It provides the inequality 3.3, which we want to use in searching for the bound of $\nu(\mathcal{S})$, but it needs simplification to apply it. Observe that this inequality needs the condition, $\delta \ge \|\Delta\|$ is small enough to guarantee that R_f and R_g are less than the constant $R_{\mathcal{S}}$ of the solid $\mathcal{S}$ of interest.

Proposition 3.15. *Let $\mathcal{S}$ be a solid containing the origin. There exists a finite subset D of the domain of $\mathcal{A}_\mathcal{S}$ and a constant $k^* > 0$ such that*

$$\nu(\mathcal{S}) \leq \max\{(1+k^*)\mathcal{A}_\mathcal{S}(X) : X \in D\}.$$

Proof. We use the same setup as in the proof of Theorem 3.14. Note that R_f and R_g are depending on 7 variables $(\phi_1, \beta_1, \phi_2, \beta_2, \theta, x, y)$. By our claim in the proof of Theorem 3.14, we can find a small constant $\delta^* \geq \|\Delta\| > 0$ and a constant $R > 0$ that it makes both $R_f, R_g \leq R < R_\mathcal{S}$ for any X in the domain of $\mathcal{A}_\mathcal{S}$. We can bound the inequality 3.3, by a constant $k^* \geq \frac{2R}{R_\mathcal{S}-R} > 0$, into

$$|\mathcal{A}_\mathcal{S}(X+\Delta) - \mathcal{A}_\mathcal{S}(X)| \leq k^*\mathcal{A}_\mathcal{S}(X) \Rightarrow \mathcal{A}_\mathcal{S}(X+\Delta) \leq (1+k^*)\mathcal{A}_\mathcal{S}(X). \tag{3.4}$$

Then, we construct a finite subset D of the domain of $\mathcal{A}_\mathcal{S}$. Observe that any projection of $\mathcal{S}$ is a subset of the set $[-\operatorname{diam}(\mathcal{S}), \operatorname{diam}(\mathcal{S})]^2$. So, $C := [0, 2\pi]^5 \times [-\operatorname{diam}(\mathcal{S}), \operatorname{diam}(\mathcal{S})]^2$ covers the set $\{X' : \mathcal{A}_\mathcal{S}(X') > 0\}$. Since C is compact we can choose a set of finite points D such that $\{X' : \mathcal{A}_\mathcal{S}(X') > 0\} \subseteq \bigcup_{X\in D} B_{\delta^*}(X)$. Use the inequality 3.4, we get that $\nu(\mathcal{S}) \leq \max\{(1+k^*)\mathcal{A}_\mathcal{S}(X) : X \in D\}$. □

Remark 3.16. We note that smaller δ^* also gives us smaller k^*. But it also increases the minimum number of elements that we have to use in the construction of the subset D. The above corollary provides that $\max\{(1+k^*)\mathcal{A}_\mathcal{S}(X) : X \in D\}$ is an upper bound of $\nu(\mathcal{S})$. On the other hands, $\max\{\mathcal{A}_\mathcal{S}(X) : X \in D\}$ is obviously a lower bound of $\nu(\mathcal{S})$.

4 Discussion

Before the discussion, let us conclude what we got from the section 3. We have gotten two main theorems; Theorem 3.8 and Theorem 3.14. For convenience of reading, we recall the significant parts of them here.
Theorem 3.8; for any convex polyhedron $\mathcal{S}$,

$$\nu(\mathcal{S}) \leq \min\left\{\frac{\operatorname{diam}(\mathcal{S})}{\inf d\left((E/E)\backslash(V/F)\right)}, \frac{\sup d(V/E)}{\operatorname{width}(\mathcal{S})}\right\}. \tag{4.1}$$

Theorem 3.14; for any solid $\mathcal{S}$ containing the origin. The function $\mathcal{A}_\mathcal{S}$ is continuous. Proposition 3.15, after Theorem 3.14, provides that for designed finite subset D of the domain of $\mathcal{A}_\mathcal{S}$ and the small constant k^* in Proposition,

$$\nu(\mathcal{S}) \leq \max\{(1+k^*)\mathcal{A}_\mathcal{S}(X) : X \in D\}. \tag{4.2}$$

Our work provides two upper bounds of a given convex polyhedron $\mathcal{S}$, which are shown above. The big difference between these two is how hard it can get. Obviously, the first one is much easier than the second one. It can be calculated by using some algorithms and getting the result in a short time. But, as it should be, the disadvantage of the first one is fixed, which means it can not converge to $\nu(\mathcal{S})$. On the other hand, the second one can find a smaller k^*, but it will enlarge the subset D (which means the time of calculation is also increased) to improve the upper bound.

As mentioned in the Section 1, most of the works in this field focus on searching for the lower bound of the Nieuwland constant of a given polyhedron. So, this work will help to confirm how close the lower bound is to the Nieuwland constant of a given polyhedron.

Grid-search algorithm operates by constructing a grid on variables and calculating the value of our objective function ($\mathcal{A}_\mathcal{S}$) at each node on the grid. As mentioned in Section 1, we aim to develop an effective way to use the Grid-search algorithm for finding an upper bound of $\nu(\mathcal{S})$ for given polyhedrons $\mathcal{S}$. More specifically, Grid-search algorithm itself can provide us with a lower bound of $\mathcal{A}_\mathcal{S}$, by running through 7 variables $(\phi_1, \beta_1, \phi_2, \beta_2, \theta, x, y)$. But to find an upper bound, we do some theories on the function $\mathcal{A}_\mathcal{S}$ in Section 3.2. We have to investigate the difference of $\mathcal{A}_\mathcal{S}$ when we change a small amount of the initial variables. This is equivalent to considering the continuity of $\mathcal{A}_\mathcal{S}$. So, this work also proves a presumption that $\mathcal{A}_\mathcal{S}$ is continuous. This fact seems true, but it is not easy to prove it rigorously.

In the future, we aim to calculate our upper bounds using Grid-search algorithm on well-known polyhedrons, e.g., Archimedean solids, especially the Rhombicosidodecahedron and Snub cube. An obstacle to apply Corollary 3.15 is choosing constant δ^* in the proof of Corollary 3.15. So, we want a formula (or algorithm) to help us select a constant δ^* and a related constant k^*. In the way we used them in the proof, we are positive that there should be an easy way to select them. This may be a future research direction for this topic.

Acknowledgment. The first author is supported by scholarships from Chulalongkorn University Graduate School, to celebrate the 72nd Birthday Anniversary of His Majesty the King. A special thanks to Pongbunthit Tonpho, who gave some advice at the beginning of this work. He also gave useful comments and suggestions on this manuscript. The authors wish to express their sincere gratitude for the insightful comments provided by the referee. These suggestions have been enhancing the quality of this work.

References

[1] G. Huber, K.P. Shultz, J.E. Wetzel, The n -cube is Rupert, Amer. Math. Monthly. 125 (6) (2018) 505–512.

[2] R.P. Jerrard, J.E. Wetzel, L. Yuan, Platonic passages, Math. Mag. 90 (2) (2017) 87–98.

[3] E. Scott, Two sufficient conditions for a polyhedron to be (locally) Rupert, Cornell University Library. <`https://arxiv.org/abs/2208.12912`>, 2022 (accessed 18.02.25).

[4] H. Maehara, N. Tokushige, Regular simplices passing through holes, Geom. Dedicata 145 (1) (2010) 19–32.

[5] A. Bezdek, Z. Guan, M. Hujter, A. Joós, Cubes and Boxes Have Rupert's Passages in Every Nontrivial Direction, Amer. Math. Monthly. 128 (6) (2021) 534–542.

[6] C.J. Scriba, Das Problem des Prinzen Ruprecht von der Pfalz. Praxis der Mathematik 10 (1968) 241–246.

[7] Y. Chai, Y. Liping, Z. Tudor, Rupert property of Archimedean solids, Amer. Math. Monthly. 125 (6) (2018) 497–504.

[8] P. Tonpho, Covering of objects related to Rupert property, (Master thesis), Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn Uni-

versity, (2018) <https://digital.car.chula.ac.th/cgi/viewcontent.cgi?article=3464&context=chulaetd>.

[9] J. Steininger, S. Yurkevich, extended abstract for: solving Rupert's problem algorithmically, ACM Commun. Comput. Algebra. 56 (2) (2022) 32–35.

[10] J. Steininger, S. Yurkevich, An algorithmic approach to Rupert's problem, Math. Comp. 92 (342) (2023) 1905—1929.

[11] A. Fredriksson, Optimizing for the Rupert property, Amer. Math. Monthly. 131 (3) (2024) 255–261.

[12] M.E. Houle, G.T. Toussaint, Computating the width of a set, Proceedings of the first annual symposium on Computational geometry - SCG'85, ACM Press, Baltimore, Maryland, United States (1985), pp. 1–7.

[13] N.H.B. Sahoo, Extremal Values of Pi, Amer. Math. Monthly. 129 (10) (2022) 933–951.

[14] F.P. Preparata, M.I. Shamos, Computational Geometry: An Introduction, Springer New York, New York, 1985.

รูปเรขาคณิตของผลลัพธ์ของผลคูณเชิงเวกเตอร์

พัฒนพรรณ เหลืองปฐมธนา[1,*] และ เป็นหญิง โรจนกุล[1,†]

[1]ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ 50200

บทคัดย่อ

ผลคูณเชิงเวกเตอร์เป็นแนวคิดพื้นฐานทางคณิตศาสตร์ที่มีความสำคัญทั้งในแง่ทฤษฎีและการประยุกต์ใช้ งานวิจัยนี้ศึกษาเกี่ยวกับผลลัพธ์ของผลคูณเชิงเวกเตอร์ในระบบพิกัดฉากสามมิติ โดยกำหนดให้ $\vec{\mathbb{T}}(\vec{x},\vec{y})$ เป็นรูปเชิงเวกเตอร์ของรูปสี่เหลี่ยมจัตุรัสที่ระบุด้านประกอบได้จากเวกเตอร์ $\vec{x}$ และ $\vec{y}$ ที่มีความยาวเท่ากันและตั้งฉากกัน และ $\vec{u}$ เป็นเวกเตอร์ใด ๆ ที่ไม่ใช่เวกเตอร์ศูนย์ โดยพิจารณาผลลัพธ์ของ $\vec{u}\times\vec{\mathbb{T}}(\vec{x},\vec{y})$ ภายใต้เงื่อนไขที่กำหนด ผลการศึกษาพบว่าผลลัพธ์ของผลคูณเชิงเวกเตอร์สามารถเขียนอยู่ในรูปเรขาคณิตที่มีลักษณะเฉพาะตัว ซึ่งผลลัพธ์ที่ได้จะเป็นรูปสี่เหลี่ยมผืนผ้า รูปสี่เหลี่ยมจัตุรัส รูปสี่เหลี่ยมขนมเปียกปูน และเส้นตรง ขึ้นอยู่กับมุมระหว่างเวกเตอร์ $\vec{u}$ กับเวกเตอร์ใน $\vec{\mathbb{T}}(\vec{x},\vec{y})$

คำสำคัญ: เวกเตอร์, ผลคูณเชิงเวกเตอร์, รูปเรขาคณิต

2020 MSC: 47L65, 97G50, 97G70

1 บทนำ

เวกเตอร์เป็นปริมาณที่มีทั้งขนาดและทิศทาง ซึ่งเป็นรากฐานที่สำคัญต่อศาสตร์หลายแขนง ทั้งด้านคณิตศาสตร์ วิทยาศาสตร์ และวิศวกรรมศาสตร์ เวกเตอร์เป็นเครื่องมือพื้นฐานในการคำนวณพีชคณิตเชิงเส้นและเรขาคณิตวิเคราะห์ [7, 9] เป็นพื้นฐานของแคลคูลัสเวกเตอร์ที่ใช้ในการคำนวณอัตราการเปลี่ยนแปลงของฟังก์ชันหลายตัวแปร และสามารถใช้คำนวณการเคลื่อนที่ในหลายมิติ ซึ่งเวกเตอร์ในทางคณิตศาสตร์มีการดำเนินการที่สำคัญ คือ การบวกเวกเตอร์ ผลคูณเชิงสเกลาร์ และผลคูณเชิงเวกเตอร์ อีกทั้งเวกเตอร์ยังมีบทบาทสำคัญในการเขียนสมการแสดงการอธิบายและวิเคราะห์ปรากฏการณ์ทางฟิสิกส์ และช่วยให้เห็นภาพพจน์ของความคิดทางด้านฟิสิกส์ได้ง่ายและกระจ่างขึ้น เช่น การเคลื่อนที่ของวัตถุ แรงและสมดุลกลศาสตร์ สนามแม่เหล็กและสนามไฟฟ้า และกลศาสตร์ ควอนตัมและสัมพัทธภาพ [1, 8] นอกจากนี้ยังเป็นองค์ความรู้สำคัญในการสร้างภาพเคลื่อนไหวเชิงคอมพิวเตอร์กราฟิก [10] ตลอดจนสามารถนำมาประยุกต์ใช้จริงในชีวิตประจำวันด้านการเดินทาง จีพีเอส กีฬา และดนตรี ซึ่งถือเป็นเหตุผลว่าทำไมการเรียนการสอนในเรื่องเวกเตอร์ในระดับชั้นมัธยมศึกษาจึงสำคัญมาก

*ผู้นำเสนอ (pattanapanzii@gmail.com)
†ผู้ประพันธ์บรรณกิจ (penying.rochanakul@cmu.ac.th)

หนึ่งในเนื้อหาที่สำคัญคือ ผลคูณเชิงเวกเตอร์ อันเป็นการคูณระหว่างเวกเตอร์สองเวกเตอร์ที่ได้ผลลัพธ์เป็นเวกเตอร์ ซึ่งในบทความนี้จะกล่าวถึงเฉพาะเวกเตอร์ในระบบพิกัดฉากสามมิติ โดยเป็นการดำเนินการทางคณิตศาสตร์เพื่อให้ผลลัพธ์ใหม่ได้เป็นเวกเตอร์มีที่ทิศทางตั้งฉากกับเวกเตอร์ทั้งสองตัวที่เป็นเวกเตอร์ตั้งต้นเสมอ

งานวิจัยนี้ได้จัดทำขึ้นเพื่อการศึกษาค้นคว้าผลลัพธ์ที่ได้จากผลคูณเชิงเวกเตอร์ โดยพิจารณาว่าผลคูณของเวกเตอร์ใด ๆ กับเวกเตอร์สองเวกเตอร์ที่มีความยาวเท่ากันและตั้งฉากกันสามารถได้ผลลัพธ์เป็นรูปเรขาคณิตที่แตกต่างกันอย่างไรบ้าง หากกำหนดเงื่อนไขเวกเตอร์เริ่มต้น

2 ความรู้พื้นฐาน

ในหัวข้อนี้จะกล่าวถึงความรู้พื้นฐาน และเอกลักษณ์ของเวกเตอร์

บทนิยาม 2.1 ([2]). *เวกเตอร์ (vector)* คือ ปริมาณที่มีขนาดและทิศทาง ซึ่งให้สัญลักษณ์ $\overrightarrow{AB}$ แทนเวกเตอร์ที่มีจุด*เริ่มต้น (initial point)* ที่จุด A และ *จุดสิ้นสุด (terminal point)* ที่จุด B *ขนาด* ของ $\overrightarrow{AB}$ คือ ความยาวของส่วนของเส้นตรง AB แทนด้วยสัญลักษณ์ $|\overrightarrow{AB}|$

บทนิยาม 2.2 ([2]). *เวกเตอร์ตำแหน่ง (position vector)* คือเวกเตอร์ที่ระบุตำแหน่งของจุดใด ๆ ก็ตามเมื่อเทียบกับจุดอ้างอิง

ในปริภูมิสามมิติ เมื่อจุดอ้างอิงเป็นจุดกำเนิด $O = (0, 0, 0)$ และ $P = (x, y, z)$ เวกเตอร์ตำแหน่ง $\vec{v}$ ของจุด P สามารถแสดงเป็น $\vec{v} = x\vec{i} + y\vec{j} + z\vec{k}$

ตัวอย่าง 2.3. เวกเตอร์ที่มีจุดเริ่มต้นที่จุด $P_1(x_1, y_1, z_1)$ และจุดสิ้นสุดที่จุด $P_2(x_2, y_2, z_2)$ ซึ่งเขียนแทนด้วย $\overrightarrow{P_1P_2}$ หมายถึง เวกเตอร์ $\begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix}$

นอกจากนี้ $\begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix}$ ยังสามารถแทนเวกเตอร์ที่มีจุดเริ่มต้นที่จุดกำเนิดและจุดสิ้นสุดที่จุด $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$ ด้วย

บทนิยาม 2.4 ([3, 6]). ให้ $\vec{u} = a\vec{i} + b\vec{j} + c\vec{k}$ และ $\vec{v} = d\vec{i} + e\vec{j} + f\vec{k}$ เป็นเวกเตอร์ในระบบพิกัดฉากสามมิติ

1. *ผลคูณเชิงสเกลาร์ (scalar product)* ของ $\vec{u}$ และ $\vec{v}$ เขียนแทนด้วย $\vec{u} \cdot \vec{v}$ กำหนดโดย

$$\vec{u} \cdot \vec{v} = ad + be + cf$$

2. *ผลคูณเชิงเวกเตอร์ (cross product)* ของ $\vec{u}$ และ $\vec{v}$ เขียนแสดงแทนด้วย $\vec{u} \times \vec{v}$ กำหนดโดย

$$\vec{u} \times \vec{v} = (bf - ce)\vec{i} - (af - cd)\vec{j} + (ae - bd)\vec{k}$$

หมายเหตุ 2.5. นิยมเขียนผลลัพธ์ของผลคูณเชิงเวกเตอร์ในรูปดีเทอร์มิแนนต์ดังนี้

$$\begin{aligned}\vec{u}\times\vec{v} &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a & b & c \\ c & e & f \end{vmatrix} \\ &= (bf-ce)\vec{i}-(af-cd)\vec{j}+(ae-bd)\vec{k}\end{aligned}$$

บทนิยาม 2.6 ([6]). ให้ $\vec{u}$ และ $\vec{v}$ ไม่ใช่เวกเตอร์ศูนย์
$\vec{u}$ *ขนานกับ* $\vec{v}$ เขียนแทนด้วย $\vec{u}//\vec{v}$ ก็ต่อเมื่อ มีจำนวนจริง k ที่ $k\neq 0$ ซึ่ง $\vec{u}=k\vec{v}$
ถ้า $k>0$ แล้ว $\vec{u}$ และ $\vec{v}$ มีทิศทางเดียวกัน
ถ้า $k<0$ แล้ว $\vec{u}$ และ $\vec{v}$ มีทิศทางตรงข้ามกัน

ทฤษฎีบท 2.7 ([6]). *กำหนด* $\vec{u},\vec{v}$ *และ* $\vec{w}$ *เป็นเวกเตอร์ใด ๆ ในระบบพิกัดฉากสามมิติ และ* a *เป็นจำนวนจริงใด ๆ จะได้ว่า*

1. $a(\vec{u}+\vec{v})=a\vec{u}+a\vec{v}$
2. $\vec{u}\times(\vec{v}+\vec{w})=(\vec{u}\times\vec{v})+(\vec{u}\times\vec{w})$
3. $(a\vec{u})\times\vec{v}=a(\vec{u}\times\vec{v})$
4. $\vec{u}\perp\vec{v}$ *ก็ต่อเมื่อ* $\vec{u}\cdot\vec{v}=0$
5. $\vec{u}\cdot(\vec{v}+\vec{w})=\vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w}$ *และ* $(\vec{u}+\vec{v})\cdot\vec{w}=(\vec{u}\cdot\vec{w})+(\vec{v}\cdot\vec{w})$
6. $\vec{u}\cdot\vec{v}=|\vec{u}||\vec{v}|\cos\theta$ *เมื่อ* θ *เป็นขนาดของมุมระหว่าง* $\vec{u}$ *และ* $\vec{v}$ *โดยที่* $0^\circ\leq\theta\leq 180^\circ$
7. $|\vec{u}\times\vec{v}|=|\vec{u}|\ |\vec{v}|\sin\theta$ *เมื่อ* θ *เป็นขนาดของมุมระหว่าง* $\vec{u}$ *และ* $\vec{v}$ *โดยที่* $0^\circ\leq\theta\leq 180^\circ$
8. $\vec{u}\times(\vec{v}\times\vec{w})=(\vec{u}\cdot\vec{w})\vec{v}-(\vec{u}\cdot\vec{v})\vec{w}$
9. $\vec{u}\cdot(\vec{v}\times\vec{w})=\vec{v}\cdot(\vec{w}\times\vec{u})=\vec{w}\cdot(\vec{u}\times\vec{v})$

บทนิยาม 2.8 ([4, 5]). ให้ L *เส้นตรงในปริภูมิสามมิติ* ที่ผ่านจุด $P_0(x_0,y_0,z_0)$ และขนานกับเวกเตอร์ $\vec{v}=v_1\vec{i}+v_2\vec{j}+v_3\vec{k}$ จะได้ L คือเซตของจุด $P(x,y,z)$ ทั้งหมดที่ทำให้เวกเตอร์ $\overrightarrow{P_0P}$ ขนานกับ $\vec{v}$ นั่นคือ $\overrightarrow{P_0P}=t\vec{v}$ เมื่อ $t\neq 0$ เป็นจำนวนจริง หรือ

$$(x-x_0)\vec{i}+(y-y_0)\vec{j}+(z-z_0)\vec{k}=t(v_1\vec{i}+v_2\vec{j}+v_3\vec{k}) \tag{2.1}$$

ถ้าให้ $\vec{r}$ เป็นเวกเตอร์ตำแหน่งของจุด P และ $\vec{r_0}$ เป็นเวกเตอร์ตำแหน่งของจุด P_0 จากสมการ (2.1) จะได้ว่า *สมการเชิงเวกเตอร์ของเส้นตรงที่ผ่านจุด* $P_0(x_0,y_0,z_0)$ และขนานกับ $\vec{v}$ คือ

$$\vec{r}=\vec{r_0}+t\vec{v}, \qquad -\infty<t<\infty \tag{2.2}$$

บทนิยาม 2.9 ([11]). กำหนดให้ V และ W เป็นปริภูมิเวกเตอร์บนฟิลด์ของจำนวนจริง $\mathbb{R}$
เรียกฟังก์ชัน $T : V \to W$ ว่า *การแปลงเชิงเส้น (linear transformations)* จาก V ไปยัง W
เมื่อ T สอดคล้องกับสมบัติ 2 ข้อ คือ สำหรับเวกเตอร์ u และ v ใด ๆ ใน V และ จำนวนจริง c ใด ๆ ใน $\mathbb{R}$

1. $T(u+v) = T(u) + T(v)$
2. $T(cu) = cT(u)$

ทฤษฎีบท 2.10. *ในระบบพิกัดฉากสามมิติ ผลคูณเชิงเวกเตอร์เป็นการแปลงเชิงเส้น*

พิสูจน์. ให้ $T : \mathbb{R}^3 \to \mathbb{R}^3$ กำหนดโดย $T(\vec{x}) = \vec{a} \times \vec{x}$ สำหรับทุกเวกเตอร์ $\vec{x} \in \mathbb{R}$
ให้ c เป็นจำนวนจริงใด ๆ และให้ $\vec{u}, \vec{v}$ และ $\vec{a}$ เป็นเวกเตอร์ใด ๆ ในระบบพิกัดฉากสามมิติ จะได้ว่า

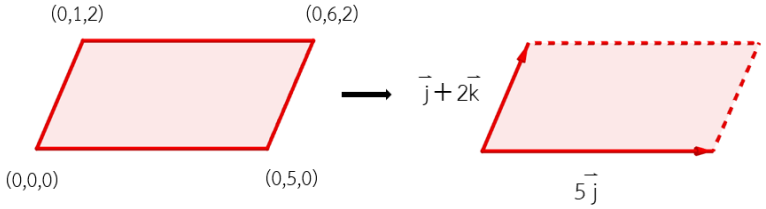
1. $T(\vec{u}+\vec{v}) = \vec{a} \times (\vec{u}+\vec{v}) = (\vec{a} \times \vec{u}) + (\vec{a} \times \vec{v}) = T(\vec{u}) + T(\vec{v})$
2. $T(c\vec{u}) = \vec{a} \times (c\vec{u}) = c(\vec{a} \times \vec{u}) = cT(\vec{u})$

นั่นคือผลคูณเชิงเวกเตอร์เป็นการแปลงเชิงเส้น □

ในหัวข้อถัดไป เราจะกำหนดสัญลักษณ์เพื่อความสะดวกในการวิเคราะห์ผลลัพธ์ของบทความ

3 กำหนดสัญลักษณ์

รูปสี่เหลี่ยมด้านขนานใด ๆ ในปริภูมิสามมิติ เมื่อกำหนดให้พิกัดของจุดยอดหนึ่งเป็นจุดกำเนิดจะสามารถอ้างถึงรูปสี่เหลี่ยมด้านขนานนั้นได้ด้วยเวกเตอร์ตำแหน่งของจุดยอดซึ่งเป็นจุดสิ้นสุดของด้าน 2 ด้านที่เป็นด้านประกอบของจุดดังกล่าวดังรูป 1



รูปที่ 1: รูปสี่เหลี่ยมด้านขนานที่มีจุดยอดหนึ่งเป็นจุดกำเนิดและระบุด้านประกอบได้จากเวกเตอร์ตำแหน่ง $\vec{j} + 2\vec{k}$ และ $5\vec{j}$

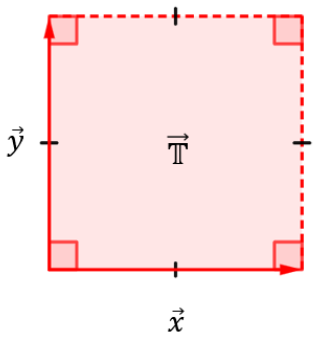
เราจึงนิยามรูปสี่เหลี่ยมด้านขนานในปริภูมิสามมิติในรูปของเวกเตอร์ตำแหน่งดังนี้

บทนิยาม 3.1. ให้ $\vec{x}$ และ $\vec{y}$ เป็นเวกเตอร์ กำหนดสัญลักษณ์ $\vec{\mathbb{T}}(\vec{x}, \vec{y})$ แทนรูปสี่เหลี่ยมด้านขนานที่ระบุด้านประกอบได้จากเวกเตอร์ตำแหน่ง $\vec{x}$ และ $\vec{y}$ นั่นคือ $\vec{\mathbb{T}}(\vec{x}, \vec{y}) = \{t\vec{x} + s\vec{y} \mid t, s \in [0,1]\}$

หมายเหตุ 3.2. เมื่อ $\vec{x}$ และ $\vec{y}$ ตั้งฉากกันและมีขนาดเท่ากันจะได้ว่า $\vec{\mathbb{T}}(\vec{x}, \vec{y})$ เป็นรูปสี่เหลี่ยมจัตุรัส
เมื่อ $\vec{x}$ หรือ $\vec{y}$ เวกเตอร์ใดเวกเตอร์หนึ่งเป็นเวกเตอร์ศูนย์ จะได้ว่า $\vec{\mathbb{T}}(\vec{x}, \vec{y})$ เป็นเส้นตรง
เมื่อทั้ง $\vec{x}$ และ $\vec{y}$ เป็นเวกเตอร์ศูนย์ จะได้ว่า $\vec{\mathbb{T}}(\vec{x}, \vec{y})$ เป็นจุด

บทนิยาม 3.3. ให้ $\vec{u} \in \mathbb{R}^3$ และ $W \subseteq \mathbb{R}^3$ กำหนด $\vec{u} \times W = \{\vec{u} \times \vec{v} \mid \vec{v} \in W\}$ เรียก $\vec{u}$ ว่า *เวกเตอร์ตั้งต้น*

ในบทความนี้จะศึกษาลักษณะเฉพาะของผลคูณเชิงเวกเตอร์กับรูปสี่เหลี่ยมจัตุรัสและเงื่อนไขที่เกี่ยวข้อง ซึ่งเราจะพิจารณา $\vec{\mathbb{T}}(\vec{x},\vec{y})$ ในกรณีที่ $\vec{x}$ และ $\vec{y}$ มีความยาวเท่ากันและตั้งฉากกัน แต่ไม่เป็นเวกเตอร์ศูนย์ ดังแสดงในรูปที่ 2 นั่นคือรูปสี่เหลี่ยมจัตุรัสที่ระบุด้านประกอบได้จาก เวกเตอร์ตำแหน่ง $\vec{x}$ และ $\vec{y}$ โดยเรามีสมมติฐานว่าผลลัพธ์จะเป็นรูปสี่เหลี่ยมจัตุรัส รูปสี่เหลี่ยมผืนผ้า รูปสี่เหลี่ยมด้านขนาน ฯลฯ ขึ้นกับความสัมพันธ์ระหว่างเวกเตอร์ตั้งต้น $\vec{u}$ และเวกเตอร์ $\vec{x}$ และ $\vec{y}$ ที่ใช้ระบุด้านประกอบของรูปสี่เหลี่ยมจัตุรัส



รูปที่ 2: $\vec{\mathbb{T}}(\vec{x},\vec{y})$ แทนรูปสี่เหลี่ยมจัตุรัสที่ระบุด้านประกอบได้จากเวกเตอร์ $\vec{x}$ และ $\vec{y}$

4 ผลการศึกษา

จากทฤษฎีบท 2.10 ผลคูณเชิงเวกเตอร์เป็นการแปลงเชิงเส้น เราจึงได้ตามมาว่าผลคูณเชิงเวกเตอร์ ของเวกเตอร์ใด ๆ $\vec{u}$ กับรูปสี่เหลี่ยมจัตุรัส $\vec{\mathbb{T}}(\vec{x},\vec{y})$ จะยังคงเป็นรูปสี่เหลี่ยม ดังทฤษฎีบทต่อไปนี้

ทฤษฎีบท 4.1. *ให้ $\vec{x},\vec{y}$ และ $\vec{u}$ เป็นเวกเตอร์ในระบบพิกัดฉากสามมิติ จะได้ว่า*

$$\vec{u}\times\vec{\mathbb{T}}(\vec{x},\vec{y})=\vec{\mathbb{T}}(\vec{u}\times\vec{x},\vec{u}\times\vec{y})$$

ในทฤษฎีบทถัดจากนี้จะศึกษาเงื่อนไขที่สอดคล้องกับลักษณะเฉพาะของรูปสี่เหลี่ยมที่เป็นผลลัพธ์

ทฤษฎีบท 4.2. *ให้ $\vec{u}$ เป็นเวกเตอร์ในระบบพิกัดฉากสามมิติใด ๆ ที่ไม่ใช่เวกเตอร์ศูนย์ ให้ $\vec{x}$ และ $\vec{y}$ เป็นเวกเตอร์ในระบบพิกัดฉากสามมิติที่ ตั้งฉากกันและมีขนาดเท่ากัน โดยที่ $\vec{u}$ ไม่ขนานกับ $\vec{x}$ และ $\vec{y}$ และให้ $\vec{\mathbb{V}}=\vec{u}\times\vec{\mathbb{T}}(\vec{x},\vec{y})$ จะได้ว่า*

1. *$\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมผืนผ้า ก็ต่อเมื่อ $\theta_{ux}=90^\circ$ หรือ $\theta_{uy}=90^\circ$*
   โดยจะเป็นรูปสี่เหลี่ยมจัตุรัส เมื่อ $\theta_{ux}=90^\circ$ และ $\theta_{uy}=90^\circ$
2. *$\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมขนมเปียกปูน ก็ต่อเมื่อ $\theta_{ux}=\theta_{uy}$ หรือ $\theta_{ux}=180^\circ-\theta_{uy}$*
3. *$\vec{\mathbb{V}}$ เป็นเส้นตรง ก็ต่อเมื่อ $\vec{u}$ และ $\vec{\mathbb{T}}(\vec{x},\vec{y})$ อยู่บนระนาบเดียวกัน กล่าวคือ*
   $\theta_{ux}+\theta_{uy}=90^\circ$ หรือ $|\theta_{ux}-\theta_{uy}|=90^\circ$

โดยที่ θ_{ux} แทนมุมระหว่าง $\vec{u}$ กับ $\vec{x}$ และ θ_{uy} แทนมุมระหว่าง $\vec{u}$ กับ $\vec{y}$

พิสูจน์. จาก $\vec{\mathbb{V}} = \vec{u} \times \vec{\mathbb{T}}(\vec{x}, \vec{y})$ เป็นรูปสี่เหลี่ยมที่ระบุด้านประกอบได้จากเวกเตอร์ $\vec{u} \times \vec{x}$ และ $\vec{u} \times \vec{y}$

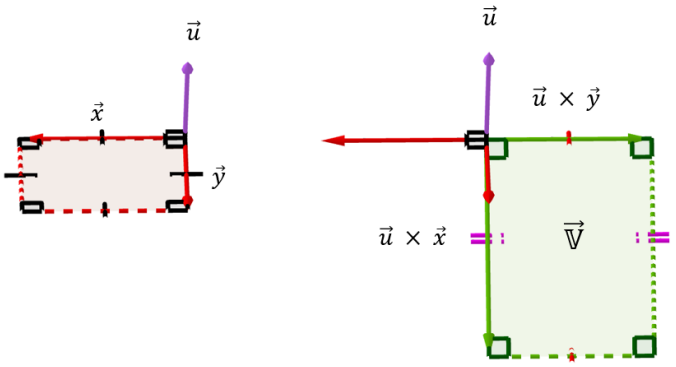
1. สมมติ $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมผืนผ้า จะได้ว่า $\vec{u} \times \vec{x} \perp \vec{u} \times \vec{y}$ กล่าวคือ $(\vec{u} \times \vec{x}) \cdot (\vec{u} \times \vec{y}) = 0$
   จะแสดงว่า $\theta_{ux} = 90^\circ$ หรือ $\theta_{uy} = 90^\circ$

พิจารณา
$$\begin{aligned}
0 = (\vec{u} \times \vec{x}) \cdot (\vec{u} \times \vec{y}) &= \Big[(\vec{u} \times \vec{y}) \times \vec{u}\Big] \cdot \vec{x} && \text{(จากทฤษฎีบท 2.7 ข้อที่ 9)} \\
&= \Big[(\vec{u} \cdot \vec{u})\vec{y} - (\vec{y} \cdot \vec{u})\vec{u}\Big] \cdot \vec{x} && \text{(จากทฤษฎีบท 2.7 ข้อที่ 8)} \\
&= \Big[(\vec{u} \cdot \vec{u})\vec{y}\Big] \cdot \vec{x} - \Big[(\vec{y} \cdot \vec{u})\vec{u}\Big] \cdot \vec{x} && \text{(จากทฤษฎีบท 2.7 ข้อที่ 5)} \\
&= \Big[(|\vec{u}||\vec{u}| \cos 0^\circ)\vec{y}\Big] \cdot \vec{x} - \Big[(|\vec{y}||\vec{u}| \cos \theta_{yu})\vec{u}\Big] \cdot \vec{x} && \text{(จากทฤษฎีบท 2.7 ข้อที่ 6)} \\
&= (|\vec{u}|^2\vec{y}) \cdot \vec{x} - \Big[(|\vec{y}||\vec{u}| \cos \theta_{yu})\vec{u}\Big] \cdot \vec{x} \\
&= |\vec{u}|^2|\vec{y}||\vec{x}| \cos 90^\circ - |\vec{y}||\vec{u}| \cos \theta_{yu} \, |\vec{u}||\vec{x}| \cos \theta_{xu} && \text{(จากทฤษฎีบท 2.7 ข้อที่ 6)} \\
&= 0 - |\vec{y}||\vec{u}| \cos \theta_{yu} \, |\vec{u}||\vec{x}| \cos \theta_{xu} && \because \cos 90^\circ = 0 \\
&= -|\vec{y}||\vec{u}| \cos \theta_{yu} \, |\vec{u}||\vec{x}| \cos \theta_{xu}
\end{aligned}$$

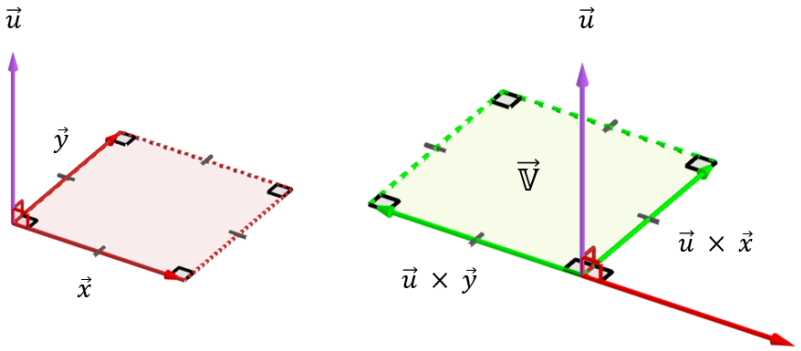
สมมติ $\theta_{ux} = 90^\circ$ หรือ $\theta_{uy} = 90^\circ$ จะแสดงว่า $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมผืนผ้า
โดยไม่เสียนัยทั่วไป สมมติ $\theta_{ux} = 90^\circ$ นั่นคือ $\vec{u} \perp \vec{x}$
พิจารณา
$$\begin{aligned}
(\vec{u} \times \vec{x}) \cdot (\vec{u} \times \vec{y}) &= \vec{y} \cdot \Big[(\vec{u} \times \vec{x}) \times \vec{u}\Big] && \text{(จากทฤษฎีบท 2.7 ข้อที่ 9)} \\
&= \vec{y} \cdot \Big[(\vec{u} \cdot \vec{u})\vec{x} - (\vec{u} \cdot \vec{x})\vec{u}\Big] && \text{(จากทฤษฎีบท 2.7 ข้อที่ 8)} \\
&= (\vec{u} \cdot \vec{u})(\vec{x} \cdot \vec{y}) - (\vec{u} \cdot \vec{x})(\vec{u} \cdot \vec{y}) && \text{(จากทฤษฎีบท 2.7 ข้อที่ 5)} \\
&= (\vec{u} \cdot \vec{u})(0) - (\vec{u} \cdot \vec{x})(\vec{u} \cdot \vec{y}) && \because \vec{x} \perp \vec{y} \\
&= -(\vec{u} \cdot \vec{x})(\vec{u} \cdot \vec{y}) && \because \vec{u} \perp \vec{x} \\
&= 0
\end{aligned}$$

ดังนั้น $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมผืนผ้า ดังแสดงในรูปที่ 3
กรณี $\theta_{ux} = \theta_{uy} = 90^\circ$ จะได้เพิ่มเติมว่า $|\vec{u} \times \vec{x}| = |\vec{u}||\vec{x}| \sin \theta_{ux} = |\vec{u}||\vec{y}| \sin \theta_{uy} = |\vec{u} \times \vec{y}|$
นั่นคือ $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมจัตุรัส ดังรูปที่ 4

2. $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมขนมเปียกปูน ก็ต่อเมื่อ $\theta_{ux} = \theta_{uy}$ หรือ $\theta_{ux} = 180^\circ - \theta_{uy}$
   สมมติว่า $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมขนมเปียกปูน นั่นคือ $|\vec{u} \times \vec{x}| = |\vec{u} \times \vec{y}|$
   จาก $|\vec{u} \times \vec{x}| = |\vec{u}||\vec{x}| \sin \theta_{ux}$ และ $|\vec{u} \times \vec{y}| = |\vec{u}||\vec{y}| \sin \theta_{uy}$ (จากทฤษฎีบท 2.7 ข้อที่ 7)
   จะได้ว่า $\sin \theta_{ux} = \dfrac{|\vec{u} \times \vec{x}|}{|\vec{u}||\vec{x}|} = \dfrac{|\vec{u} \times \vec{y}|}{|\vec{u}||\vec{y}|} = \sin \theta_{uy}$
   ดังนั้น $\theta_{ux} = \theta_{uy}$ หรือ $\theta_{ux} = 180^\circ - \theta_{uy}$
   ในทางกลับกัน สมมติ $\theta_{ux} = \theta_{uy}$ หรือ $\theta_{ux} = 180^\circ - \theta_{uy}$

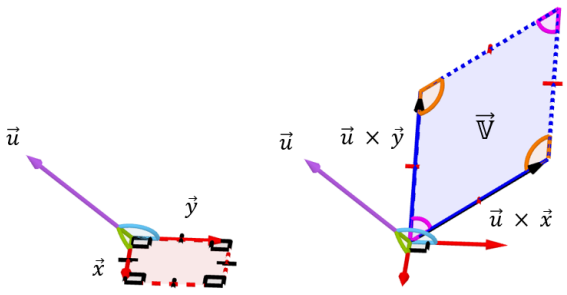


รูปที่ 3: รูปสี่เหลี่ยมผืนผ้า $\vec{\mathbb{V}}$ เมื่อ $\vec{u} \perp \vec{x}$



รูปที่ 4: รูปสี่เหลี่ยมจัตุรัส $\vec{\mathbb{V}}$ เมื่อ $\vec{u} \perp \vec{x}$ และ เมื่อ $\vec{u} \perp \vec{y}$

จะได้ว่า $\sin\theta_{ux} = \sin\theta_{uy}$
ดังนั้น $|\vec{u} \times \vec{x}| = |\vec{u}||\vec{x}|\sin\theta_{ux} = |\vec{u}||\vec{y}|\sin\theta_{uy} = |\vec{u} \times \vec{y}|$
นั่นคือ $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมขนมเปียกปูน ดังแสดงในรูปที่ 5



รูปที่ 5: รูปสี่เหลี่ยมขนมเปียกปูน $\vec{\mathbb{V}}$ เมื่อ $\theta_{ux} = \theta_{uy}$ หรือ $\theta_{ux} = 180^\circ - \theta_{uy}$

3. $\vec{\mathbb{V}}$ เป็นเส้นตรง ก็ต่อเมื่อ $\vec{u}$ และ $\vec{\mathbb{T}}(\vec{x}, \vec{y})$ อยู่บนระนาบเดียวกัน
สังเกตว่า $\vec{u} \times \vec{x}$ กับ $\vec{u} \times \vec{y}$ จะเป็นเส้นตรงก็ต่อเมื่อ $(\vec{u} \times \vec{x}) \times (\vec{u} \times \vec{y}) = \vec{0}$

สมมติ $\vec{\mathbb{V}}$ เป็นเส้นตรง นั่นคือ $\vec{u}\times\vec{x}$ และ $\vec{u}\times\vec{y}$ มีทิศทางเดียวกัน หรือ มีทิศทางตรงข้ามกัน
ซึ่งจะได้ตามมาว่า $(\vec{u}\times\vec{x})\times(\vec{u}\times\vec{y})=\vec{0}$

พิจารณา $\vec{0}=(\vec{u}\times\vec{x})\times(\vec{u}\times\vec{y})=[(\vec{u}\times\vec{x})\cdot\vec{y}\,]\vec{u}-[(\vec{u}\times\vec{x})\cdot\vec{u}\,]\vec{y}$ (ทฤษฎีบท 2.7 ข้อที่ 8)
$=[(\vec{u}\times\vec{x})\cdot\vec{y}\,]\vec{u}-(0)\vec{y}$ $\because(\vec{u}\times\vec{x})\perp\vec{u}$
$=[(\vec{x}\times\vec{y})\cdot\vec{u}\,]\vec{u}$

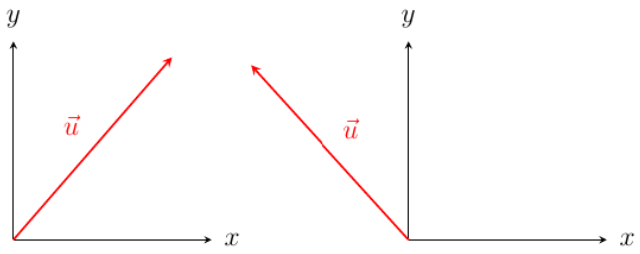
จาก $\vec{u}\neq 0$ จะได้ $(\vec{x}\times\vec{y})\cdot\vec{u}=0$
ดังนั้น $(\vec{x}\times\vec{y})\perp\vec{u}$
นั่นคือ $\vec{u}$ และ $\vec{\mathbb{T}}(\vec{x},\vec{y})$ อยู่บนระนาบเดียวกัน

ขั้นต่อไปสมมติว่า $\vec{x},\vec{y}$ และ $\vec{u}$ อยู่บนระนาบเดียวกัน จะได้ว่า $(\vec{x}\times\vec{y})\perp\vec{u}$ และ $(\vec{u}\times\vec{x})\perp\vec{u}$
ดังนั้น $(\vec{x}\times\vec{y})\cdot\vec{u}=0$ และ $(\vec{u}\times\vec{x})\cdot\vec{u}=0$ พิจารณาได้จากสองกรณีต่อไปนี้

พิจารณา $(\vec{u}\times\vec{x})\times(\vec{u}\times\vec{y})=[(\vec{u}\times\vec{x})\cdot\vec{y}]\,\vec{u}-[(\vec{u}\times\vec{x})\cdot\vec{u}]\,\vec{y}$ (ทฤษฎีบท 2.7 ข้อที่ 8)
$=[(\vec{x}\times\vec{y})\cdot\vec{u}]\,\vec{u}-[(\vec{u}\times\vec{x})\cdot\vec{u}]\,\vec{y}$ (ทฤษฎีบท 2.7 ข้อที่ 9)
$=(0)\vec{u}-(0)\vec{y}$ $\because\vec{x}\perp\vec{y}$ และ $(\vec{u}\times\vec{x})\perp\vec{u}$
$=\vec{0}$

จะได้ว่า $\vec{u}\times\vec{x}$ และ $\vec{u}\times\vec{y}$ มีทิศทางเดียวกัน หรือ มีทิศทางตรงข้ามกัน
กล่าวคือ $\vec{\mathbb{V}}$ เป็นเส้นตรง ดังแสดงในรูปที่ 7
นั่นคือ $\vec{\mathbb{V}}$ เป็นเส้นตรง ก็ต่อเมื่อ $\vec{x},\vec{y}$ และ $\vec{u}$ อยู่บนระนาบเดียวกัน กล่าวคือ $\vec{u}$ และ $\vec{\mathbb{T}}(\vec{x},\vec{y})$ อยู่ระนาบเดียวกัน

หมายเหตุ 4.3. เมื่อ $\vec{x},\vec{y}$ และ $\vec{u}$ อยู่บนระนาบเดียวกัน จะเป็นได้สองกรณีดังนี้
กรณีที่ 1 $\vec{u}$ อยู่ระหว่าง $\vec{x}$ กับ $\vec{y}$ จะได้ $\theta_{ux}+\theta_{uy}=90^\circ$
กรณีที่ 2 $\vec{u}$ ไม่ได้อยู่ระหว่าง $\vec{x}$ กับ $\vec{y}$ จะได้ $|\theta_{ux}-\theta_{uy}|=90^\circ$
กรณีทั้งสองสามารถแสดงได้ดังรูปที่ 6 □

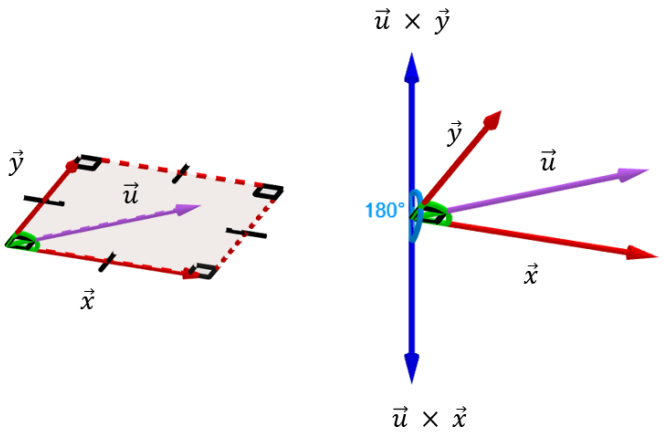


กรณีที่ 1 กรณีที่ 2

รูปที่ 6: $\vec{x},\vec{y}$ และ $\vec{u}$ อยู่บนระนาบเดียวกัน

5 สรุปผลและอภิปรายผล

จากการศึกษาผลลัพธ์ของผลคูณเชิงเวกเตอร์จากการพิสูจน์ โดยกำหนดเงื่อนไข ให้ $\vec{\mathbb{T}}(\vec{x},\vec{y})$ แทนรูปสี่เหลี่ยมจัตุรัสที่ระบุด้านประกอบได้จากเวกเตอร์ $\vec{x}$ และ $\vec{y}$ ที่มีความยาวเท่ากันและตั้งฉากกันแต่ไม่เท่ากับศูนย์ และ $\vec{u}$ เป็นเวกเตอร์



รูปที่ 7: เส้นตรง $\vec{\mathbb{V}}$ เมื่อ $\vec{x}, \vec{y}$ และ $\vec{u}$ อยู่บนระนาบเดียวกัน

ใด ๆ ที่ไม่ใช่เวกเตอร์ศูนย์และไม่ขนานกับ $\vec{x}$ และ $\vec{y}$ ให้ $\vec{\mathbb{V}} = \vec{u} \times \vec{\mathbb{T}}(\vec{x}, \vec{y})$ จะได้ผลลัพธ์ของผลคูณเชิงเวกเตอร์ทั้งหมด 3 แบบ ตามเงื่อนไขดังนี้

1. $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมผืนผ้า ก็ต่อเมื่อ $\theta_{ux} = 90°$ หรือ $\theta_{uy} = 90°$ โดยจะเป็นรูปสี่เหลี่ยมจัตุรัส เมื่อ $\theta_{ux} = 90°$ และ $\theta_{uy} = 90°$
2. $\vec{\mathbb{V}}$ เป็นรูปสี่เหลี่ยมขนมเปียกปูน ก็ต่อเมื่อ $\theta_{ux} = \theta_{uy}$ หรือ $\theta_{ux} = 180° - \theta_{uy}$
3. $\vec{\mathbb{V}}$ เป็นเส้นตรง ก็ต่อเมื่อ $\vec{u}$ และ $\vec{\mathbb{T}}(\vec{x}, \vec{y})$ อยู่บนระนาบเดียวกัน กล่าวคือ $\theta_{ux} + \theta_{uy} = 90°$ หรือ $|\theta_{ux} - \theta_{uy}| = 90°$

นอกเหนือจากเงื่อนไขทั้ง 3 ข้างต้น ผลลัพธ์ของผลคูณเชิงเวกเตอร์ของรูปสี่เหลี่ยมจัตุรัสจะเป็นรูปสี่เหลี่ยมด้านขนาน และเราสามารถขยายผลลัพธ์นี้ไปยัง ผลคูณเชิงเวกเตอร์ของทรงลูกบาศก์ด้วยหลักการเดียวกัน โดยใช้เวกเตอร์สามเวกเตอร์ที่ตั้งฉากกันและมีขนาดเท่ากันในการระบุด้านประกอบสาม ด้านที่ติดกับจุดกำเนิด และแยกพิจารณาทีละคู่ เพื่อพิจารณาผลลัพธ์ของพื้นผิวทีละด้าน

กิตติกรรมประกาศ ผู้วิจัยขอขอบคุณผู้ทรงคุณวุฒิทุกท่านที่ได้ให้ข้อคิดเห็นและข้อเสนอแนะต่าง ๆ เพื่อปรับปรุงบทความวิจัยนี้

เอกสารอ้างอิง

[1] ก่องกัญจน์ ภัทรากาญจน์, ฟิสิกส์ทั่วไป 1, มหาวิทยาลัยรามคำแหง, กรุงเทพมหานคร, 2531.

[2] กวิยา เนาวประทีป, เทคนิคการเรียนคณิตศาสตร์ เวกเตอร์, สำนักพิมพ์ฟิสิกส์เซ็นเตอร์, กรุงเทพมหานคร, 2555.

[3] จักรินทร์ วรรณโพธิ์กลาง, สุดยอดคำนวณและเทคนิคลัด คณิตศาสตร์ม.5 ค 014, สำนักพิมพ์พัฒนาศึกษา, กรุงเทพมหานคร, 2543

[4] ดำรงค์ ทิพย์โยธา, ณัฏฐนาถ ไตรภพ, และสุรชัย สมบัติบริบูรณ์, แคลคูลัส 2, สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร, 2558.

[5] ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่, แคลคูลัสสำหรับวิศวกรรมศาสตร์. <https://www.math.science.cmu.ac.th/thawinan/206161/2559-1/Intro_to_vector.pdf>, 2559 (สืบค้น 16 ก.พ. 68).

[6] สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, หนังสือเรียนรายวิชาเพิ่มเติม คณิตศาสตร์ เล่ม 1 ชั้นมัธยมศึกษาปีที่ 5 ตามผลการเรียนรู้ กลุ่มสาระการเรียนรู้คณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551, โรงพิมพ์ สกสค.ลาดพร้าว, 2563.

[7] สาขาการสอนคณิตศาสตร์ คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง, เวกเตอร์ในระบบสามมิติ. <http://old-book.ru.ac.th/e-book/m/ME504(54)/ME504-8.pdf>, (สืบค้น 16 ก.พ. 68).

[8] สมพงษ์ ใจดี, ฟิสิกส์ทั่วไป มหาวิทยาลัย 1, สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร, 2548.

[9] วไลลักษณ์ คีรีเพ็ชร, การศึกษาผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์เรื่องเวกเตอร์ในปริภูมิสามมิติ ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 และชั้นมัธยมศึกษาปีที่ 5, (ปริญญานิพนธ์การศึกษามหาบัณฑิต), มหาวิทยาลัยทักษิณ, 2546.

[10] C. Zhu, Q. Peng, J. X. Chen, and H. Zhang, ,Using geometric algebra for 3D linear transformations, Computing in Science Engineering. 8 (3) 68–75.

[11] Machine Learning Glossary and Terms, Linear Transformation. <https://deepai.org/machine-learning-glossary-and-terms/linear-transformation>, (สืบค้น 16 ม.ค. 68).

Generalized Final Topologies and Hereditary Classes

Paphatphong Painee[1,*] and Pongdate Montagantirud[1]

[1]Department of Mathematics, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

For a family of functions from generalized topological spaces X_α for all $\alpha \in \Lambda$ to a nonempty set X, we introduce the concept of a generalized final topology on X, which is more general than the generalized quotient topology, and survey some examples. Additionally, we study the relation between the generalized topologies induced by hereditary classes and the genralized final topology via particular hereditary classes.

Keywords: Generalized topology, Generalized final topology, Hereditary classes
2020 MSC: 54A05, 54B99

1 Introduction

In 1997, Császár[2] introduced the idea of open sets within a topological space using monotonic mappings, referred to as γ-open sets. He noted that for any monotonic map γ, the empty set is γ-open, and the union of any γ-open sets remains γ-open. Based on this observation, he defined a generalized topology (GT) as a family of subsets of a given nonempty set that includes the empty set and is closed under arbitrary unions of its elements. In 1990, Janković and Hamlett[7] defined a hereditary class $\mathcal{H}$ on a topological space X as a collection of subsets of X where every subset of a member of $\mathcal{H}$ also belongs to $\mathcal{H}$. Moreover, they extended the concept of the closure in a topological space using the special hereditary class called an ideal. After that, Császár[4] adopted this concept and applied on generalized topologies. Moreover, quotient spaces play an important role in topological constructions. In 2011, L.E. De Arruda Saraiva[1] constructed a generalized quotient topology $\mathfrak{g}_\pi$ on a non-empty set Y induced by a surjective map $\pi : (X, \mathfrak{g}_X) \to Y$ and proved the characterisation of the generalized quotient topology that $\mathfrak{g}_\pi$ is the largest generalized topology for which π is generalized continuous. Consequently, P. Montagantirud and S. Phonrakkhet[8] studied properties of generalized quotient topologies when they have hereditary classes on the domain and the codomain of a generalized quotient map. The preceding characterization of the generalized quotient topology, when looked at from a more general point of view (i.e. instead of a single function f from GT Z to GT X we consider family of functions f_α from generalized topological spaces $(X_\alpha, \mathfrak{g}_\alpha)$ to a nonempty set X) leads us to the concept of the generalized final topology, which is defined to be the largest generalized topology on X that make f_α generalized continuous for all α, and determine

*Speaker and corresponding author (3p.painee@gmail.com)

some properties. We continue to survey some examples of generalized final topological spaces in addition to the generalized quotient space. For examples, the generalized disjoint union topology is the generalized final topologies with respect to the generalized inclusion maps. In 2018, P. Montagantirud and W. Thaikua [9], studied generalized topologies via hereditary classes. n more detail, they focused on the following questions: What happens if the generalized topology of the domain of a generalized quotient map is extended by some hereditary class, or if it can even be expressed as an extension of some generalized topology via a hereditary class? The same questions was also posed for generalized quotient topologies. They answered this question in two situations, each involving a hereditary class on both the domain and codomain. In our work, we adopt some approaches from [9], but focus on the case where a hereditary class is given on each domain and consider the generalized final topology on the codomain which is more general than generalized quotient topology. More precisely, let $((X_\alpha, \mathfrak{g}_\alpha, \mathcal{H}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces with hereditary classes to a nonempty set X. Let $\mathfrak{g}_X$ be a generalized final topology on X induced by the family $(f_\alpha)_{\alpha\in\Lambda}$. If we consider the following data $((X_\alpha, (\mathfrak{g}_\alpha)^*_{\mathcal{H}_\alpha}, \mathcal{H}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ where $(\mathfrak{g}_\alpha)^*_\alpha$ is a generalized topology on X_α via the hereditary class $\mathcal{H}_\alpha$, we can construct the new generalized final topology G_X on X that relate to each $(\mathfrak{g}_\alpha)^*_{\mathcal{H}_\alpha}$. By virtue of the fact that $f_\alpha(\mathcal{H}_\alpha)$ is again a hereditary class on X for each $\alpha \in \Lambda$, this leads us to study the relation among the generalized final topology $\mathfrak{g}_X$, the generalized final topology G_X and another $(\mathfrak{g}_X)^*_{\mathcal{K}}$ where $\mathcal{K}$ is a hereditary class on X that relates to the hereditary class $f_\alpha(\mathcal{H}_\alpha)$.

2 Preliminaries

The concept of generalized topologies and generalized topological spaces are defined by [3] as follows.

Definition 2.1. A subset $\mathfrak{g}_X \subseteq \mathcal{P}(X)$ is called a generalized topology (GT) on a non-empty set X if

1. $\varnothing \in \mathfrak{g}_X$,
2. an arbitrary union of elements in $\mathfrak{g}_X$ is in $\mathfrak{g}_X$.

The pair $(X, \mathfrak{g}_X)$ is called a **generalized topological space** (GTS). An element in $\mathfrak{g}_X$ is called a $\mathfrak{g}_X$**-open**(or **genralized open set** or **open set** if no confusion can arise). A set $F \subseteq X$ is said to be $\mathfrak{g}_X$**-closed** (or **genralized closed set** or **closed set** if no confusion can arise) if $X - F$ is $\mathfrak{g}_X$-open.

The **closure** of a set $A \subseteq X$ in $(X, \mathfrak{g}_x)$, denoted by $\operatorname{cl} A$, can be defined by

$$\operatorname{cl} A = \{x \in X : M \cap A \neq \varnothing \text{ for each } M \in \mathfrak{g}_X \text{ containing } x\}.$$

Next we address the important concepts of functions between GTS.

Definition 2.2. A function $f : (X, \mathfrak{g}_X) \to (Y, \mathfrak{g}_Y)$ between GTS's is said to be

(i) $(\mathfrak{g}_X, \mathfrak{g}_Y)$**-continuous** (or **generalized continuous**) if $f^{-1}(V) \in \mathfrak{g}_X$ whenever $V \in \mathfrak{g}_Y$.

(ii) $(\mathfrak{g}_X, \mathfrak{g}_Y)$**-open** (or **generalized open**) if $f(U) \in \mathfrak{g}_Y$ whenever $U \in \mathfrak{g}_X$.

(iii) $(\mathfrak{g}_X, \mathfrak{g}_Y)$**-homeomorphism** (or **generalized homeomorphism**) if f is bijective, generalized continuous and generalized open.

Remark 2.3. It is straightforward to check that a composition of generalized continuous (resp. generalized open, generalized homeomorphism) functions is also a generalized continuous (resp. generalized open, generalized homeomorphism) function.

We now introduce some notations. Let $f : X \to Y$ be a function. For a collection μ (resp. ν) of subsets of X (resp. Y), we write a $f(\mu) := \{f(M) : M \in \mu\}$ (resp. $f^{-1}(\nu) := \{f^{-1}(B) : B \in \nu\}$).

Proposition 2.4 ([8]). *Let (X, μ) be a GTS and $f : X \to Y$ a function. Then $f(\mu)$ is a GT on Y.*

Remark 2.5. This result may be false in topological spaces.

Definition 2.6. Let $\pi : (X, \mathfrak{g}_X) \to Y$ be a surjective function. The **generalized quotient topology** $\mathfrak{g}_\pi$ is defined to be a collection $\{O \subseteq Y : \pi^{-1}(O) \in \mathfrak{g}_X\}$, and the map π is called the **generalized quotient map**.

Proposition 2.7 ([1]). *The generalized quotient topology $\mathfrak{g}_\pi$ is the largest GT $\mathfrak{g}_Y$ on Y making π $(\mathfrak{g}_X, \mathfrak{g}_Y)$-continuous.*

Hamlett and Janković gave the idea of a hereditary class and an ideal in [7].

Definition 2.8. A **hereditary class** on X is defined to be a collection $\mathcal{H}$ of subsets of X such that

$$A \subseteq B \subseteq X \text{ and } B \in \mathcal{H} \text{ imply } A \in \mathcal{H}.$$

We shall denote $(X, \mathfrak{g}_X, \mathcal{H})$ a generalized topological space $(X, \mathfrak{g}_X)$ together with a hereditary class $\mathcal{H}$.

Definition 2.9. A hereditary class $\mathcal{I}$ on X is said to be an **ideal** if each $A, B \in \mathcal{I}$ implies $A \cup B \in \mathcal{I}$.

Definition 2.10. Let $(X, \mathfrak{g}_X, \mathcal{H})$ be a GTS with a hereditary class. For each $A \subseteq X$, we define

$$A^*_{\mathcal{H}} = \{x \in X : M \cap A \notin \mathcal{H} \text{ for any } M \in \mathfrak{g}_X \text{ containing } x\} \text{ and } c^*_{\mathfrak{g}_X,\mathcal{H}} A = A \cup A^*_{\mathcal{H}}.$$

If there is no confusion about $\mathfrak{g}_X$ and $\mathcal{H}$, then we may write $A^*_{\mathcal{H}}$ and $c^*_{\mathfrak{g}_X,\mathcal{H}}$ as A^* and c^*, respectively.

Remark 2.11. Observe that if $\mathcal{H} = \{\varnothing\}$, $\operatorname{cl} A$ and $A^*_{\{\varnothing\}}$ coincide.

According to [4], if we have a hereditary class on a generalized topology $(X, \mathfrak{g}_X)$, then we will have a new generalized topology on X.

Definition 2.12. Let $(X, \mathfrak{g}_X, \mathcal{H})$ be a GTS with a hereditary class. We define an extension of the generalized topology $\mathfrak{g}_X$ to be

$$(\mathfrak{g}_X)^*_{\mathcal{H}} = \{M \subseteq X : c^*(X \setminus M) = X \setminus M\},$$

and we call $(\mathfrak{g}_X)^*_{\mathcal{H}}$ the generalized topology via the hereditary class.

Proposition 2.13 ([4]). *Let $(X, \mathfrak{g}_X, \mathcal{H})$ be a GTS. Then $\mathfrak{g}_X \subseteq (\mathfrak{g}_X)^*_{\mathcal{H}}$.*

Remark 2.14. Every $(\mathfrak{g}_X, \mathfrak{g}_Y)$-continuous function is also a $((\mathfrak{g}_X)^*_{\mathcal{H}}, \mathfrak{g}_Y)$-continuous function. This follows from Proposition 2.13.

Let $(X, \mathfrak{g}_X)$ be a GTS. A collection $\mathcal{B}$ of subsets of X is called a **base for** $\mathfrak{g}_X$ if every $\mathfrak{g}_X$-open is a union of elements of $\mathcal{B}$. The following theorem tells us about the elements that belong to a generalized topology via a hereditary class.

Theorem 2.15 ([8]). *Let $(X, \mathfrak{g}_X, \mathcal{H})$ be a GTS. Then the collection $\{M - H : M \in \mathfrak{g}_X$ and $H \in \mathcal{H}\}$ is a base for $(\mathfrak{g}_X)^*_{\mathcal{H}}$.*

Definition 2.16. Let X be a nonempty set, $\mathcal{H}$ a hereditary class on X, and $A \subseteq X$. The relative hereditary class $\mathcal{H}_A$ on A is

$$\mathcal{H}_A = \{H \cap A : H \in \mathcal{H}\}.$$

Theorem 2.17 ([9]). *Let $(X, \mu, \mathcal{H})$ be a GTS with hereditary class and $A \subseteq X$. For the relative hereditary class $\mathcal{H}_A$ on A, we have $(\mu_A)^* = (\mu^*)_A$.*

The following proposition allows us to talk about the hereditary class on the codomain.

Proposition 2.18 ([8]). *Let $\mathcal{H}$ be a hereditary class on X. Then for any map $\pi : X \to Y, \pi(\mathcal{H}) := \{\pi(H) : H \in \mathcal{H}\}$ is a hereditary class on Y.*

Theorem 2.19 ([9]). *Let $(X, \mu, \mathcal{H})$ and (Y, ν) be GTS. If f is (μ, ν)-continuous bijection from X onto Y, then for the hereditary class*

$$f(\mathcal{H}) = \{f(H) : H \in \mathcal{H}\}),$$

the function f is $(\mu^*_{\mathcal{H}}, \nu^*_{f(\mathcal{H})})$-continuous.

Theorem 2.20 ([9]). *Let (X, μ) and (Y, ν) be generalized topological spaces and $f : X \to Y$. If $X \in \mu$, then there is a hereditary class $\mathcal{H}_f = \{A \subseteq X - f^{-1}(V) : V \in \nu\}$ on X such that f is $(\mu^*_{\mathcal{H}_f}, \nu)$-continuous.*

If a continuous function f behaves well enough, it will affect the continuity of the generalized topology via a hereditary class.

Theorem 2.21 ([9]). *Let (X, μ) and $(Y, \nu, \mathcal{H}_Y)$ be GTS. If f is a (μ, ν)-continuous injection from X into Y, then for the hereditary class on X defined by*

$$\mathcal{H}_X = \{f^{-1}(A) | A \in \mathcal{H}_Y\},$$

the function f is $(\mu^*_{\mathcal{H}_X}, \nu^*_{\mathcal{H}_Y})$-continuous.

3 Main Results

3.1 Generalized final topology

Similar to continuous functions in topological spaces, holomorphic functions in complex analysis, differentiable functions in real analysis, and linear transformations in linear algebra which all play fundamental roles, in the generalized topological spaces, the kind of function that we are

interested in is a generalized continuous function. Let $f : Z \to X$ be a function from a generalized topology Z to a nonempty set X. Of course, we could not tell anyting about the generalized continuity of f. However we can ask whether there is a generalized topology on X that makes f to be a generalized continuous function. The simplest answer is a trivial generalized topology. The next question: is there the largest generalized topology? In the situation of a generalized quotient map, the answer is "yes" by Thereom 2.7. From this, we obtain the characterization of the generalized quotient topology. Let $((X_\alpha, \mathfrak{g}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces $(X_\alpha, \mathfrak{g}_\alpha)$ to a nonempty set X. We use the same idea as in the characterization of the generalized quotient map to give the generalized topology on X.

Definition 3.1. Let $((X_\alpha, \mathfrak{g}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces $(X_\alpha, \mathfrak{g}_\alpha)$ to a nonempty set X. The **generalized final topology** (GFT) $\mathfrak{g}_X \subseteq \mathcal{P}(X)$ on X induced by the family $(f_\alpha)_{\alpha\in\Lambda}$ is the largest generalized topology on X such that f_α is continuous for all $\alpha \in \Lambda$. More precisely, if $\mathfrak{g}'$ is a GT on X that make f_α is $(\mathfrak{g}_\alpha, \mathfrak{g}')$-continuous for all $\alpha \in \Lambda$ then $\mathfrak{g}' \subseteq \mathfrak{g}_X$.

Remark 3.2. The generalized quotient space is a generalized final topology with respect to a single generalized quotient map.

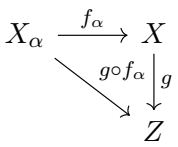
We shall show that a GFT always exists.

Proposition 3.3. *Given $((X_\alpha, \mathfrak{g}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ as before, the GFT $\mathfrak{g}_X$ on X exists and is characterised by the following condition: If $U \subseteq X$, then $U \in \mathfrak{g}_X$ if and only if $f_\alpha^{-1}(U) \in \mathfrak{g}_\alpha$ for all $\alpha \in \Lambda$.*

Proof. Define $\mathfrak{g}_X = \{U \subseteq X |\ f_\alpha^{-1}(U) \in \mathfrak{g}_\alpha \text{ for all } \alpha \in \Lambda\}$. Fisrt, we will show that $\mathfrak{g}_X$ is a generalized topology. Clearly, $\varnothing \in \mathfrak{g}_X$. The formula $f_\alpha^{-1}\left(\bigcup_{i\in\Omega} U_i\right) = \bigcup_{i\in\Omega} f_\alpha^{-1}(U_i)$ for all α shows that the union of arbitrary elements in $\mathfrak{g}_X$ is in $\mathfrak{g}_X$. Thus $\mathfrak{g}_X$ does define a generalized topology on X. By construction we have f_α is $(\mathfrak{g}_\alpha, \mathfrak{g}_X)$-continuous for all α. We now show that this GT is the GFT. Assume $\mathfrak{g}'$ is another GT on X such that f_α is $(\mathfrak{g}_\alpha, \mathfrak{g}')$-continuous for all $\alpha \in \Lambda$. Let $V \in \mathfrak{g}'$. Then $f_\alpha^{-1}(V) \in \mathfrak{g}_\alpha$ for all α by assumption, and so $V \in \mathfrak{g}_X$. Therefor $\mathfrak{g}' \subseteq \mathfrak{g}_X$ as desired. □

The generalized final topology has its own universal property. This amounts to saying precisely which functions out of the GFT are generalized continuous.

Theorem 3.4. *Let $((X_\alpha, \mathfrak{g}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces $(X_\alpha, \mathfrak{g}_\alpha)$ to a nonempty set X. If X is endowed with the GFT $\mathfrak{g}_X$, then a function $g : (X, \mathfrak{g}_X) \to (Z, \mathfrak{g}_Z)$ is $(\mathfrak{g}_X, \mathfrak{g}_Z)$-continuous if and only if a composition $g \circ f_\alpha$ is $(\mathfrak{g}_\alpha, \mathfrak{g}_Z)$-continuous for every $\alpha \in \Lambda$.*



Proof. ($\Rightarrow$) This follows immediately from Remark 2.3
($\Leftarrow$) Assume that $g \circ f_\alpha$ is $(\mathfrak{g}_\alpha, \mathfrak{g}_Z)$-continuous for every $\alpha \in \Lambda$. Let $\mathfrak{g}' = \{g^{-1}(U) \subseteq X : U \in \mathfrak{g}_Z\}$. Observe that $\mathfrak{g}'$ is a GT on X. For each $\alpha \in \Lambda$ and each $U \in \mathfrak{g}_Z$ we have $f_\alpha^{-1}(g^{-1}(U)) = (g \circ f_\alpha)^{-1}(U) \in \mathfrak{g}_\alpha$ since $g \circ f_\alpha$ is $(\mathfrak{g}_\alpha, \mathfrak{g}_Z)$-continuous. Thus f_α is $(\mathfrak{g}_\alpha, \mathfrak{g}')$-continuous for each $\alpha \in \Lambda$. By definition of GFT, $\mathfrak{g}' \subseteq \mathfrak{g}_X$, and so g is $(\mathfrak{g}_X, \mathfrak{g}_Z)$-continuous as desired. □

The converse of Theorem 3.4 holds as well.

Theorem 3.5. *Let $((X_\alpha, \mathfrak{g}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces $(X_\alpha, \mathfrak{g}_\alpha)$ to a nonempty set X. If a GT $\mathfrak{g}'$ on X satisfies the property that a function $g : (X, \mathfrak{g}') \to (Z, \mathfrak{g}_Z)$ is $(\mathfrak{g}', \mathfrak{g}_Z)$-continuous if and only if a composition $g \circ f_\alpha$ is $(\mathfrak{g}_\alpha, \mathfrak{g}_Z)$-continuous for every $\alpha \in \Lambda$, then $\mathfrak{g}'$ is the GFT on X.*

Proof. Firstly, we will show that for each $\alpha \in \Lambda$, f_α is $(\mathfrak{g}_\alpha, \mathfrak{g}')$-continuous. It is easy to see that the identity map $\mathrm{Id}_X : X \to X$ is $(\mathfrak{g}', \mathfrak{g}')$-continuous. Thus f_α is $(\mathfrak{g}_\alpha, \mathfrak{g}')$-continuous for all $\alpha \in \Lambda$ since $\mathrm{Id}_X \circ f_\alpha = f_\alpha$ for all $\alpha \in \Lambda$. Next, let $\mathfrak{g}''$ be another GT on X such that $f_\alpha : X_\alpha \to X$ is $(\mathfrak{g}_\alpha, \mathfrak{g}'')$-continuous for all $\alpha \in \Lambda$. We will show that $\mathfrak{g}'' \subseteq \mathfrak{g}'$. Let $\mathrm{Id}_X : (X, \mathfrak{g}') \to (X, \mathfrak{g}'')$ be the identity function. Then $\mathrm{Id}_X \circ f_\alpha = f_\alpha : X_\alpha \to X$ is $(\mathfrak{g}_\alpha, \mathfrak{g}'')$-continuous for each $\alpha \in \Lambda$. Thus Id_X is $(\mathfrak{g}', \mathfrak{g}'')$-continuous, and hence $\mathfrak{g}'' \subseteq \mathfrak{g}'$ as desired. Therefore $\mathfrak{g}'$ is indeed a GFT on X. □

Observe that for each $\alpha \in \Lambda$ we have $f_\alpha(\mathfrak{g}_\alpha)$ is a GT on X by Proposition 2.4, and so we can talk about the relation between $f_\alpha(\mathfrak{g}_\alpha)$ for all $\alpha \in \Lambda$ and GFT $\mathfrak{g}_X$.

Theorem 3.6. *Let $((X_\alpha, \mathfrak{g}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces $(X_\alpha, \mathfrak{g}_\alpha)$ to a nonempty set X. If f_α is bijective for all $\alpha \in \Lambda$, then $\bigcap_{\alpha\in\Lambda} f_\alpha(\mathfrak{g}_\alpha)$ and the GFT $\mathfrak{g}_X$ are the same.*

Proof. Note that $\bigcap_{\alpha\in\Lambda} f_\alpha(\mathfrak{g}_\alpha)$ is indeed a GT on X. Let U be an open subset in $\bigcap_{\alpha\in\Lambda} f_\alpha(\mathfrak{g}_\alpha)$. It follows that for each $\alpha \in \Lambda$, there is $V_\alpha \in \mathfrak{g}_\alpha$ such that $U = f_\alpha(V_\alpha)$. Since f is injective, $f_\alpha^{-1}(U) \in \mathfrak{g}_\alpha$ for all $\alpha \in \Lambda$. Hence $\bigcap_{\alpha\in\Lambda} f_\alpha(\mathfrak{g}_\alpha) \subseteq \mathfrak{g}_X$. Next, let $W \in \mathfrak{g}_X$. Then for each $\alpha \in \Lambda$ we have $f_\alpha^{-1}(W) \in \mathfrak{g}_\alpha$, and so $W = f_\alpha(f_\alpha^{-1}(W)) \in f_\alpha(\mathfrak{g}_\alpha)$ since f_α is surjective. This implies that $W \in \bigcap_{\alpha\in\Lambda} f_\alpha(\mathfrak{g}_\alpha)$. From this we see that $\mathfrak{g}_X = \bigcap_{\alpha\in\Lambda} f_\alpha(\mathfrak{g}_\alpha)$. □

Here are important examples of generalized final topologies.

Definition 3.7. Let $(X_\alpha, \mathfrak{g}_\alpha)_{\alpha\in\Lambda}$ be a family of generalized topological spaces. The **disjoint union** is the set $\coprod_{\alpha\in\Lambda} X_\alpha$ consisting of all ordered pairs (x, α) where $\alpha \in \Lambda$ and $x \in X_\alpha$. For each $\alpha \in \Lambda$, there is a canonical injection $\iota_\alpha : X_\alpha \to \coprod_{\alpha\in\Lambda} X_\alpha$ given by $x \mapsto (x, \alpha)$. We define the **generalized disjoint union topology (GDUT)** on $\coprod_{\alpha\in\Lambda} X_\alpha$ to be the largest generalized topology such that ι_α is continuous for each $\alpha \in \Lambda$. Sometimes we will write (x, α) by x_α.

Remark 3.8. Indeed, ι_α is an injective map for all $\alpha \in \Lambda$.

The following proposition tells us about elements in GDUT.

Proposition 3.9. *A subset $U \subseteq \coprod_{\alpha\in\Lambda} X_\alpha$ belongs to the GDUT $\mathfrak{g}_X$ if and only if it is of the form $U = \bigcup_{\alpha\in\Lambda}(U_\alpha, \alpha)$ where each $U_\alpha \subseteq X_\alpha$ is open and (U_α, α) is a set consisting of (u, α) for all $\alpha \in U_\alpha$.*

Proof. Let $U \subseteq \coprod_{\alpha\in\Lambda} X_\alpha$. Then $\iota_\alpha^{-1}(U) \in \mathfrak{g}_\alpha$ for all $\alpha \in \Lambda$ by definition. Thus $U = \bigcup_{\alpha\in\Lambda}(\iota_\alpha^{-1}(U), \alpha)$. Conversely, we write $U = \bigcup_{\alpha\in\Lambda}(U_\alpha, \alpha)$ where $U_\alpha \in \mathfrak{g}_\alpha$ for all $\alpha \in \Lambda$. It is easy to check that $\iota_\alpha^{-1}(U) = U_\alpha$ for all $\alpha \in \Lambda$. Thus U is open in $\coprod_{\alpha\in\Lambda} X_\alpha$ as desired. □

In algebraic topology, the concept of an adjunction space plays an important role. Not surprisingly, we can introduce the concept of generalized adjunction space as follows.

Definition 3.10. Let X, Y be GTS, A a subspace of Y, and let $f : A \to X$ be a generalized continuous function. Let $\sim$ be an equivalence relation on the disjoint union $X \sqcup Y$ generated by $a \sim f(a)$ for all $a \in A$, and define the **generalized adjunction space** by

$$X \cup_f Y = (X \sqcup Y)/\sim .$$

Let $\tilde{\pi} : X \sqcup Y \to X \cup_f Y$ be a canaonical projection i.e., $x \mapsto [x]$. We have the following two composition: $X \xrightarrow{\iota_1} X \sqcup Y \xrightarrow{\tilde{\pi}} X \cup_f Y$ and $Y \xrightarrow{\iota_2} X \sqcup Y \xrightarrow{\tilde{\pi}} X \cup_f Y$ where ι_i is the canonical injection. The generalized topology on $X \cup_f Y$ is defined to be the generalized final topology with respect to $\tilde{\pi} \circ \iota_1$ and $\tilde{\pi} \circ \iota_2$. We call this GT the generalized adjunction topology (GAT).

Proposition 3.11. *The GAT on $X \cup_f Y$ and the GQT induced by $\tilde{\pi} : X \sqcup Y \to X \cup_f Y$ are coincide.*

Proof. Let $\mathfrak{g}_X$ and $\mathfrak{g}_Y$ be GT's on X and Y, respectively. By Proposition 3.3, the GAT on $X \cup_f Y$ is $\mathfrak{g}_1 = \{U \subseteq X \cup_f Y : (\tilde{\pi} \circ \iota_1)^{-1}(U) \in \mathfrak{g}_X \text{ and } (\tilde{\pi} \circ \iota_2)^{-1}(U) \in \mathfrak{g}_Y\}$. Similarly, the GQT on $X \cup_f Y$ induced by $\tilde{\pi}$ is $\mathfrak{g}_2 = \{U \subseteq X \cup_f Y : \tilde{\pi}^{-1}(U) \in \mathfrak{g}_{X \sqcup Y}\}$ where $\mathfrak{g}_{X \sqcup Y}$ is a GDUT on $X \sqcup Y$. By Definition 3.7, it is easy to verify that $\mathfrak{g}_1 = \mathfrak{g}_2$. □

3.2 Hereditary Classes

In this section, we consider the case that we have a hereditary class $\mathcal{H}_\alpha$ on X_α for all $\alpha \in \Lambda$. Let $((X_\alpha, \mathfrak{g}_\alpha, \mathcal{H}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha \in \Lambda}$ be a family of functions f_α from generalized topological spaces with hereditary classes to a nonempty set X. Theorem 2.18 guarantees that, $f_\alpha(\mathcal{H}_\alpha)$ is a hereditary class on X for all $\alpha \in \Lambda$. Consequently, we have a lot of information about hereditary classes on X. The following lemma is an one way to deal with information of hereditary classes on X.

Lemma 3.12. *If $(\mathcal{H}_\alpha)_{\alpha \in \Lambda}$ is a collection of hereditary classes on X, then $\bigcup_{\alpha \in \Lambda} \mathcal{H}_\alpha$ and $\bigcap_{\alpha \in \Lambda} \mathcal{H}_\alpha$ are hereditary classes on X.*

Proof. Clearly, $\varnothing \in \bigcup_{\alpha \in \Lambda} \mathcal{H}_\alpha$ and $\varnothing \in \bigcap_{\alpha \in \Lambda} \mathcal{H}_\alpha$. Let $A \subseteq B \subseteq X$ be such that $B \in \bigcup_{\alpha \in \Lambda} \mathcal{H}_\alpha$. Then there is $\alpha_0 \in \Lambda$ such that $B \in \mathcal{H}_{\alpha_0}$, and so $A \in \mathcal{H}_{\alpha_0}$. Hence $A \in \bigcup_{\alpha \in \Lambda} \mathcal{H}_\alpha$. This implies that $\bigcup_{\alpha \in \Lambda} \mathcal{H}_\alpha$ is a hereditary class on X. Similarly, $\bigcap_{\alpha \in \Lambda} \mathcal{H}_\alpha$ is a hereditary class on X. □

Corollary 3.13. *If $((X_\alpha, \mathfrak{g}_\alpha, \mathcal{H}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha \in \Lambda}$ is a family of functions f_α from generalized topological spaces with hereditary classes to a nonempty set X, then $\bigcup_{\alpha \in \Lambda} f_\alpha(\mathcal{H}_\alpha)$ and $\bigcap_{\alpha \in \Lambda} f_\alpha(\mathcal{H}_\alpha)$ are also hereditary classes on X.*

Proof. This follows immediately from Proposition 2.18 and Lemma 3.12. □

Remark 3.14. In the case of a collection of ideals $(\mathcal{I}_\alpha)_{\alpha \in \Lambda}$ on X, we can show that $\bigcap_{\alpha \in \Lambda} \mathcal{I}_\alpha$ is again an ideal on X (Proposition 3.15). But $\bigcup_{\alpha \in \Lambda} \mathcal{I}_\alpha$ is not necessarily an ideal on X (Example 3.16).

Proposition 3.15. *Let $(\mathcal{I}_\alpha)_{\alpha \in \Lambda}$ be a collection of ideals on X. Then $\bigcap_{\alpha \in \Lambda} \mathcal{I}_\alpha$ is again an ideal on X.*

Proof. Let $A, B \in \bigcap_{\alpha \in \Lambda} \mathcal{I}_\alpha$. Then $A, B \in \mathcal{I}_\beta$ for each $\beta \in \Lambda$ and so $A \cup B \in \mathcal{I}_\beta$ for all $\beta \in \Lambda$ since $\mathcal{I}_\beta$ is an ideal for all $\beta \in \Lambda$. Thus $A \cup B \in \bigcap_{\alpha \in \Lambda} \mathcal{I}_\alpha$. Therefore $\bigcap_{\alpha \in \Lambda} \mathcal{I}_\alpha$ is an ideal on X as desired. □

Example 3.16. Let $X = \{a, b, c, d\}$, $\mathcal{I}_1 = \{\varnothing, \{a\}, \{a, b\}\}$ and $\mathcal{I}_2 = \{\varnothing, \{c\}\}$. Clearly, $\mathcal{I}_1$ and $\mathcal{I}_2$ are ideals on X. But $\mathcal{I}_1 \cup \mathcal{I}_2$ is not an ideal on X since $\{a\} \cup \{c\} = \{a, c\} \notin \mathcal{I}_1 \cup \mathcal{I}_2$.

Under the following condition we have arbitrary union of ideals on X is again an ideal on X.

Proposition 3.17. *Let $(\mathcal{I}_\alpha)$ be a collection of ideals on X. If for all $\beta, \gamma \in \Lambda$, $\mathcal{I}_\beta \subseteq \mathcal{I}_\gamma$ or $\mathcal{I}_\gamma \subseteq \mathcal{I}_\beta$, then $\bigcup_{\alpha\in\Lambda} \mathcal{I}_\alpha$ is an ideal on X.*

Proof. Let $A, B \in \bigcup_{\alpha\in\Lambda}$. Then there are $\beta, \gamma \in \Lambda$ such that $A \in \mathcal{I}_\beta$ and $B \in \mathcal{I}_\gamma$. WLOG, we assume $\mathcal{I}_\beta \subseteq \mathcal{I}_\gamma$. It follows that $A \cup B \in \mathcal{I}_\gamma \subseteq \bigcup_{\alpha\in\Lambda} \mathcal{I}_\alpha$ since $\mathcal{I}_\gamma$ is an ideal on X. Thus $\bigcup_{\alpha\in\Lambda} \mathcal{I}_\alpha$ is an ideal on X. □

Proposition 3.18. *Under the hypotheses of Proposition 2.18 with a hereditary class $\mathcal{H}$ replaced by an ideal $\mathcal{I}$. Then $\pi(\mathcal{I})$ is an ideal on Y.*

Proof. It suffices to show that $\pi(H_1) \cup \pi(H_2) \in \pi(\mathcal{I})$ for all $H_1, H_2 \in \mathcal{I}$. Let $H_1, H_2 \in \mathcal{I}$. Then $H_1 \cup H_2 \in \mathcal{I}$ since $\mathcal{I}$ is an ideal on X. Thus $\pi(H_1) \cup \pi(H_2) = \pi(H_1 \cup H_2) \in \pi(\mathcal{I})$ as desired. □

We can now state an analogue of Corollary 3.13.

Corollary 3.19. *Let $((X_\alpha, \mathfrak{g}_\alpha, H_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces with ideal to a nonempty set X. If for all $\beta, \gamma \in \Lambda$, $f_\beta(\mathcal{H}_\beta) \subseteq f_\gamma(\mathcal{H}_\gamma)$ or $f_\beta(\mathcal{H}_\beta) \supseteq f_\gamma(\mathcal{H}_\gamma)$ then $\bigcup_{\alpha\in\Lambda} f_\alpha(\mathcal{H}_\alpha)$ and $\bigcap_{\alpha\in\Lambda} f_\alpha(\mathcal{H}_\alpha)$ are also ideals on X.*

Recall that when we studied a generalized topology via a hereditary class, we know its concrete base elements of its by Theorem 2.15. We offer some useful lemmas for dealing with the generalized topology via the hereditary class.

Lemma 3.20. *Let $(X, \mathfrak{g}_X)$ be a GTS. If $\mathcal{B}$ is a base for $\mathfrak{g}_X$, then $\mathfrak{g}_X$ is the smallest GT on X containing $\mathcal{B}$.*

Proof. Let $\mathfrak{g}'$ be a GT on X such that $\mathcal{B} \subseteq \mathfrak{g}'$. Let $U \in \mathfrak{g}_X$. Then $U = \bigcup_{\alpha\in\Lambda} B_\alpha$ where $B_\alpha \in \mathcal{B}$ for all $\alpha \in \Lambda$. By assumption, $B_\alpha \in \mathfrak{g}'$ for each $\alpha \in \Lambda$. Since $\mathfrak{g}'$ is a GT on X, $U = \bigcup_{\alpha\in\Lambda} B_\alpha \in \mathfrak{g}'$. Hence $\mathfrak{g} \subseteq \mathfrak{g}'$. □

Here is an application of this lemma.

Proposition 3.21. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be hereditary classes on X. If $\mathcal{H}_1 \subseteq \mathcal{H}_2$, then $(\mathfrak{g}_X)^*_{\mathcal{H}_1} \subseteq (\mathfrak{g}_X)^*_{\mathcal{H}_2}$.*

Proof. Suppose that $\mathcal{H}_1 \subseteq \mathcal{H}_2$. By Theorem 2.15, $\mathcal{B}_1 := \{M - H_1 : M \in \mathfrak{g}_X, H_1 \in \mathcal{H}_1\}$ and $\mathcal{B}_2 := \{M - H_2 : M \in \mathfrak{g}_X, H_2 \in \mathcal{H}_2\}$ are bases for $(\mathfrak{g}_X)^*_{\mathcal{H}_1}$ and $(\mathfrak{g}_X)^*_{\mathcal{H}_2}$, respectively. By the assumption, we have $\mathcal{B}_1 \subseteq \mathcal{B}_2$. It follows that $(\mathfrak{g}_X)^*_{\mathcal{H}_2} \supseteq \mathcal{B}_1$. By Lemma 3.20, $(\mathfrak{g}_X)^*_{\mathcal{H}_1} \subseteq (\mathfrak{g}_X)^*_{\mathcal{H}_2}$ as desired. □

Remark 3.22. The converse is not true in general; see Example 2.14 in [9].

Corollary 3.23. *Let $((X_\alpha, \mathfrak{g}_\alpha, \mathcal{H}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces with hereditary classes to a nonempty set X. If $\mathfrak{g}_X$ is the GFT on X induced by the family of functions $(f_\alpha)_{\alpha\in\Lambda}$, then $(\mathfrak{g}_X)^*_{\bigcap_{\alpha\in\Lambda} f_\alpha(\mathcal{H}_\alpha)} \subseteq (\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)} \subseteq (\mathfrak{g}_X)^*_{\bigcup_{\alpha\in\Lambda} f_\alpha(\mathcal{H}_\alpha)}$ for all $\alpha \in \Lambda$.*

Proof. This follows immediately from Corollary 3.13 and Proposition 3.21. □

Next, if we consider $((X_\alpha, (\mathfrak{g}_\alpha)^*_{\mathcal{H}_\alpha}, \mathcal{H}_\alpha)) \xrightarrow{f_\alpha} X$, it is natural to ask about the final topology G_X on X. By Proposition 3.3, we can write

$$G_X = \{O \subseteq X : f_\alpha^{-1}(O) \in (\mathfrak{g}_\alpha)^*_{\mathcal{H}_\alpha} \text{ for all } \alpha \in \Lambda\}.$$

From now on, we let G_X be as defined above and $\mathfrak{g}_X$ be the generalized final topology on X induced by the family of functions $(f_\alpha)_{\alpha\in\Lambda}$.

Remark 3.24. In fact, we can view G_X as the largest GT on X that make f_α $((\mathfrak{g}_\alpha)^*_{\mathcal{H}_\alpha}, G_X)$-continuous for all $\alpha \in \Lambda$.

Proposition 3.25. *Let $((X_\alpha, \mathfrak{g}_\alpha, \mathcal{H}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces with hereditary classes to a nonempty set X. Then $\mathfrak{g}_X \subseteq G_X$.*

Proof. Let $O \in \mathfrak{g}_X$. Then $f_\alpha^{-1}(O) \in \mathfrak{g}_\alpha$ for all $\alpha \in \Lambda$. By Proposition 2.13, we have $f_\alpha^{-1}(O) \in (\mathfrak{g}_\alpha)^*_{\mathcal{H}_\alpha}$ for all $\alpha \in \Lambda$, and so $O \in G_X$ by the construction of G_X. Hence $\mathfrak{g}_X \subseteq G_X$ as desired. □

The following example shows that the inclusion is strict. Before we see an example, we introduced some useful terminology from [8]. Let $\pi : (X, \mu) \to Y$ be a generalized quotient map and ν a generalized quotient topology on Y. We will say that π is a (μ, ν)**-generalized quotient map** and ν is **induced by a** (μ, ν)**-generalized quotient map.**

Example 3.26. Let $X = \{a, b, c, d, e\}, \mathfrak{g}_X = \{\varnothing, \{a, b\}, \{a, b, c, d, e\}\}$ and $\mathcal{H} = \{\varnothing, \{e\}\}$. It follows that the set $\mathcal{B} := \{M - H : M \in \mathfrak{g}_X \text{ and } H \in \mathcal{H}\} = \{\varnothing, \{a, b\}, \{a, b, c, d\}, \{a, b, c, d, e\}\}$ is a GT on X. Thus $\mathcal{B} = (\mathfrak{g}_X)^*_{\mathcal{H}}$ by Lemma 3.20. Next, let $\pi : X \to Y := \{x, y, z, w\}$ be defined by $\pi(a) = x, \pi(b) = y, \pi(c) = z = \pi(d)$, and $\pi(e) = w$. Recall that every GQT is a GFT with respect to generalized quotient map. Now, we have the GQT on Y, $\mathfrak{g}_\pi = \{\varnothing, \{x, y\}, \{x, y, z, w\}\}$. Next, we consider G_π induced by a $((\mathfrak{g}_X)^*_{\mathcal{H}}, G_\pi)$-generalized quotient map. It is easy to verify that $G_\pi = \{\varnothing, \{x, y\}, \{x, y, z\}, \{x, y, z, w\}\}$. Thus $\mathfrak{g}_\pi \subsetneq G_\pi$.

In general, a bijective generalized continuous function is not necessarily a generalized homeomorphism. However, an injective generalized quotient map is always a generalized open map, and hence a generalized homeomorphism. This fact follows immediately from Proposition 3.3. Nevertheless, in the case of a generalized final topology, a bijective map need not be a generalized open map.

Example 3.27. Let $(X_1, \mathfrak{g}_1) = (\{a, b\}, \{\varnothing, \{a\}, \{a, b\}\}), (X_2, \mathfrak{g}_2) = (\{d, e\}, \{\varnothing, \{d\}\})$ be generalized topological spaces and $X = \{x, y\}$ a set. Defind $f_1 : X_1 \to X$ and $f_2 : X_2 \to X$ by $f_1(a) = x = f_2(d), f_1(b) = y = f_2(e)$. Clearly, both f_1 and f_2 are bijective function. By construction in Proposition 3.3, we can check that the generalized final topology $\mathfrak{g}_X$ on X is indeed a set $\{\varnothing, \{x\}\}$. But f_1 is not a generalized open map since $f_1(X_1) = X \notin \mathfrak{g}_X$.

The following theorem tells us that, given an arbitrary family of hereditary classes, under a nice condition on functions, we can describe the relation between G_X and the generalized final topology via the hereditary class on the codomain.

Theorem 3.28. *Let $((X_\alpha, \mathfrak{g}_\alpha, \mathcal{H}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces with hereditary classes to a nonempty set X. If f_α is bijective for all $\alpha \in \Lambda$ and $\bigcup_{\alpha\in\Lambda} f_\alpha(\mathfrak{g}_\alpha) \subseteq \mathfrak{g}_X$, then $G_X = \bigcap_{\alpha\in\Lambda} (\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)}$.*

Proof. To show $G_X \supseteq \bigcap_{\alpha\in\Lambda}(\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)}$, it is enough to show that f_α is $((\mathfrak{g}_\alpha)^*_{\mathcal{H}}, \bigcap_{\alpha\in\Lambda}(\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)})$-continuous for all $\alpha \in \Lambda$. Note that f_α is $(\mathfrak{g}_\alpha, \mathfrak{g}_X)$-continuous for all $\alpha \in \Lambda$. Since f_α is bijective for all $\alpha \in \Lambda$, f_α is $((\mathfrak{g}_\alpha)^*_{\mathcal{H}_\alpha}, (\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)})$-continuous for all $\alpha \in \Lambda$ by Theorem 2.19. This implies that f_α is $((\mathfrak{g}_\alpha)^*_{\mathcal{H}}, \bigcap_{\alpha\in\Lambda}(\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)})$-continuous for all $\alpha \in \Lambda$. It remains to show that $G_X \subseteq \bigcap_{\alpha\in\Lambda}(\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)}$. Let $G \in G_X$ and $\alpha \in \Lambda$. Then $f_\alpha^{-1}(G) \in (\mathfrak{g}_\alpha)^*_{\mathcal{H}_\alpha}$, i.e.

$$f_\alpha^{-1}(G) = \bigcup_{\beta\in\Omega}(M_\beta - H_\beta)$$

where $M_\beta \in \mathfrak{g}_\alpha$ and $H_\beta \in \mathcal{H}_\alpha$ for all $\beta \in \Omega$. Clearly, $f_\alpha(H_\beta) \in f_\alpha(\mathcal{H}_\alpha)$ for all $\beta \in \Omega$. Since $\bigcup_{\alpha\in\Lambda} f_\alpha(\mathfrak{g}_\alpha) \subseteq \mathfrak{g}_X$, $f_\alpha(M_\beta) \in \mathfrak{g}_X$ for all $\beta \in \Omega$. By surjectivity of f_α,

$$G = f_\alpha(f_\alpha^{-1}(G)) = f_\alpha(\bigcup_{\beta\in\Omega}(M_\beta - H_\beta)) = \bigcup_{\beta\in\Omega}(f_\alpha(M_\beta) - f_\alpha(H_\beta)).$$

It follows that $G_X \subseteq (\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)}$ for all $\alpha \in \Lambda$. Thus $G_X \subseteq \bigcap_{\alpha\in\Lambda}(\mathfrak{g}_X)^*_{f_\alpha(\mathcal{H}_\alpha)}$. □

Corollary 3.29. *Let the assumptions as in Theorem 3.28 be satisfied. Then*

$$(\mathfrak{g}_X)^*_{\bigcap_{\alpha\in\Lambda} f_\alpha(\mathcal{H}_\alpha)} \subseteq G_X \subseteq (\mathfrak{g}_X)^*_{\bigcup_{\alpha\in\Lambda} f_\alpha(\mathcal{H}_\alpha)}.$$

Proof. This follows immediately from Theorem 3.28 and Corollary 3.23. □

Remark 3.30. Example 3.31 and 3.32 show that we could not drop the assumption that f_α is bijective for all $\alpha \in \Lambda$ in Theorem 3.28.

Example 3.31. Let $X_1 = \{a, b, c, d\}, X_2 = \{b, c, d\}, \mathfrak{g}_1 = \{\varnothing, \{a\}, \{a, b, d\}\}, \mathfrak{g}_2 = \{\varnothing, \{c\}\}, \mathcal{H}_1 = \{\varnothing, \{a\}, \{b\}, \{a, b\}\}$, and $\mathcal{H}_2 = \{\varnothing, \{d\}\}$. Let $X = X_1 \sqcup X_2$, equipped with the canonical injection maps ι_j where $j \in \{1, 2\}$. Then

$$\mathfrak{g}_X = \{\varnothing, \{a_1\}, \{a_1, b_1, d_1\}, \{c_2\}, \{a_1, c_2\}, \{a_1, b_1, d_1, c_2\}\}$$

by Proposition 3.9. Clearly, $\iota_2(\mathcal{H}_2) = \{\varnothing, \{d_2\}\}$. Thus

$$(\mathfrak{g}_X)^*_{\iota_2(\mathcal{H}_2)} = \{\varnothing, \{a_1\}, \{a_1, b_1, d_1\}, \{c_2\}, \{a_1, c_2\}, \{a_1, b_1, d_1, c_2\}\}.$$

It is not hard to see that $(\mathfrak{g}_1)^*_{\mathcal{H}_1} = \{\varnothing, \{a\}, \{a, b, d\}, \{b, d\}, \{a, d\}, \{d\}\}$
and $(\mathfrak{g}_2)^*_{\mathcal{H}_2} = \{\varnothing, \{c\}\}$. It follows that

$$\begin{aligned} G_X = \{&\varnothing, \{a_1\}, \{a_1, b_1, d_1\}, \{b_1, d_1\}, \{a_1, d_1\}, \{d_1\}, \{c_2\}, \{a_1, c_2\}, \{a_1, b_1, d_1, c_2\}, \\ &\{b_1, d_1, c_2\}, \{a_1, d_1, c_2\}, \{d_1, c_2\}\} \end{aligned}$$

by Proposition 3.9. Therefore $\bigcap_{j\in\{1,2\}}(\mathfrak{g}_X)^*_{\iota_j(\mathcal{H}_j)} \subseteq (\mathfrak{g}_X)^*_{\iota_2(\mathcal{H}_2)} \subsetneq G_X$, and so $\bigcap_{j\in\{1,2\}}(\mathfrak{g}_j)^*_{\mathcal{H}_j} \neq G_X$.

Example 3.32. Let $X = \{a, b, c, d, e\}, \mathfrak{g}_X = \{\varnothing, \{a, b\}, X\}$, and $\mathcal{H} = \{\varnothing, \{a, b, c\}, \{c\}\}$ a hereditary class on X. Define $\pi : X \to Y := \{x, y, z, w\}$ by $\pi(a) = x, \pi(b) = y, \pi(c) = z = \pi(d)$ and $\pi(e) = w$. Then the GQT induced by π is $\mathfrak{g}_\pi = \{\varnothing, \{x, y\}, Y\}$. Clearly, $\pi(\mathcal{H}) = \{\varnothing, \{z\}, \{x, y, z\}\}$. Observe that the bases for GT $(\mathfrak{g}_X)^*_{\mathcal{H}}$ and $(\mathfrak{g}_\pi)^*_{\pi(\mathcal{H})}$ are
$\mathcal{B} = \{\varnothing, \{a, b\}, \{d, e\}, \{a, b, d, e\}, X\}$ and $\mathcal{B}' = \{\varnothing, \{w\}, \{x, y\}, \{x, y, w\}\}$, respectively. Thus $\mathcal{B} = (\mathfrak{g}_X)^*_{\mathcal{H}}$ and $\mathcal{B}' = (\mathfrak{g}_\pi)^*_{\pi(\mathcal{H})}$ by Lemma 3.20. It is not hard to check that $G_X = \{\varnothing, \{x, y\}, Y\}$. Therefore $G_X \neq (\mathfrak{g}_\pi)^*_{\pi(\mathcal{H})}$.

4 Discussion

There are several directions in which this work can be extended. First, the construction of underlying set of the generalized disjoint union topology and the generalized quotient topology is a special case of a more general, categorical construction, **colimit**. For more details about the concept of categorical constructions, we refer the reader to [5] and [6]. In Section 3.2, we studied the case in which we have a collection of GTS's equipped with hereditary classes. We investigated the relationships between these GTS's via hereditary classes and explored how particular hereditary classes induce a GFT on the codomain. Next, let $((X_\alpha, \mathfrak{g}_\alpha) \xrightarrow{f_\alpha} X)_{\alpha\in\Lambda}$ be a family of functions f_α from generalized topological spaces $(X_\alpha, \mathfrak{g}_\alpha)$ to a nonempty set X. By Remark 3.25 and Example 3.26, this approach raises several questions:

1. Can we find a condition that makes $\mathfrak{g}_X \subseteq G_X$ to be $\mathfrak{g}_X = G_X$?
2. Can we find a hereditary class $\mathcal{H}$ on X such that $(\mathfrak{g}_X)^*_{\mathcal{H}} = G_X$?

Let us now move on to discuss the following situation. Suppose we have a collection of GTS's and GFT together with hereditary class on X. We seek to understand what can be said about appropriate hereditary classes on each X_α. Furthermore, we aim to explore how the GTS's relate to one another through these hereditary classes, and how they relate to the GFT via the hereditary class on X. Let $\mathcal{H}$ be a hereditary class on X. Unfortunately, $f_\alpha^{-1}(\mathcal{H})$ may not a hereditary class on X_α. For example, see [[8], page 12]. However, we have some result about the condition of a map that makes $f_\alpha^{-1}(\mathcal{H})$ be a hereditary class on X_α.

Theorem 4.1. *Let $f : (Z, \mu) \to (Y, \nu, \mathcal{K})$ be an injective map. Then $f^{-1}(\mathcal{K}) = \{f^{-1}(K) : K \in \mathcal{K}\}$ is a hereditary class on Z.*

Proof. Let $\mathcal{K}$ be a hereditary class on Y , $A \subseteq B \subseteq X$ and $B \in f^{-1}(\mathcal{K})$. Clearly, $f(A) \subseteq f(B)$. Since f is injective, $f(B) = f(f^{-1}(K)) \subseteq K$ for some $K \in \mathcal{K}$. It follows that $f(B) \in \mathcal{K}$ because $\mathcal{K}$ is a hereditary class, and hence $f(A) \in \mathcal{K}$. Finally, $A = f^{-1}(f(A)) \in f^{-1}(\mathcal{K})$ since f is injective. Therefore $f^{-1}(\mathcal{K})$ is a hereditary class on Z as desired. □

Based on the preceding theorem, we can demonstrate that GDUT serves as an example such that, whenever there is a hereditary class $\mathcal{H}$ on the codomain, $f_\alpha^{-1}(\mathcal{H})$ froms a hereditary class on X_α.

Corollary 4.2. *Let $((X_\alpha, \mathfrak{g}_\alpha) \xrightarrow{\iota_\alpha} (X := \coprod_{\alpha\in\Lambda} X_\alpha, \mathfrak{g}_X, \mathcal{H}))_{\alpha\in\Lambda}$ where $\mathfrak{g}_X$ is the GDUT on X. Then $\iota_\alpha^{-1}(\mathcal{H})$ is a hereditary class on X for all $\alpha \in \Lambda$.*

Proof. This follows from Remark 3.8 and Theorem 4.1. □

In our future work, we aim to further investigate the setting in which a hereditary class is given on the codomain, according to two main directions:

1. Find other conditions on each map f_α such that $f_\alpha^{-1}(\mathcal{H})$ is a hereditary class on X_α
2. construct hereditary classes on each X_α that are related to the hereditary class $\mathcal{H}$ on X.

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments. The first author acknowledges support from a Development and Promotion of Science and Technology Talents Project (DPST) scholarship

References

[1] de Arruda Saraiva, L. E.: Generalized quotient topologies, *Acta Math. Hungar.* **132**, 168–173 (2011).

[2] Császár, Á.: Generalized open sets, *Acta Math. Hungar.* **75**, 65–87 (1997).

[3] Császár, Á.: Generalized topology, generalized continuity, *Acta Math. Hungar.* **96**, 351–357 (2002).

[4] Császár, Á.: Modifications of generalized topologies via hereditary classes, *Acta Math. Hungar.* **120**, 275–279 (2008).

[5] Bradley, T. D., Bryson, T., Terilla, J.: Topology a categorical approach, *The MIT Press*, (2020).

[6] Brown, R: Topology and groupoids, *BookSurge Publishing*, (2006).

[7] D. Janković , and T. R. Hamlett. 1990. "New Topologies from Old via Ideals." The American Mathematical Monthly 97 (4) (1990) 295–310.

[8] Montagantirud, P., Phonrakkhet, S.: Generalized quotient topologies and hereditary classes, *Acta Math. Hungar.* **161**, 1–15 (2020).

[9] Montagantirud, P., Thaikua, W.: Continuity on generalised topological spaces via hereditary classes, *Bull. Aust. Math. Soc.* **97**, 320–330 (2018).

ความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลม

เวธน์วศิล พิรุณกาญจน์[1,*] อัคเรศ เตี๋ยวสกุล[1,*] กรวิกา ก้องกุล[2,†] อลงกรณ์ แซ่ตั้ง[2] และ เกษราภรณ์ แซ่ตั้ง[2]

[1]สาขาวิชาคณิตศาสตร์ คณะศึกษาศาสตร์ มหาวิทยาลัยทักษิณ 90000

[2]สาขาวิชาคณิตศาสตร์และการจัดการข้อมูล คณะวิทยาศาสตร์และนวัตกรรมดิจิทัล มหาวิทยาลัยทักษิณ 93210

บทคัดย่อ

ในการศึกษางานวิจัยของ Paul Goron (2023) ที่ได้แสดงความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลม โดยแสดงการพิสูจน์ความสัมพันธ์ระหว่างความยาวด้านของรูปสามเหลี่ยม ABC ในระนาบ ที่มีเส้นตรง AC เป็นเส้นสัมผัสวงกลมที่มี r เป็นความยาวรัศมี โดยที่จุด B และ C อยู่บนวงกลมดังกล่าว เมื่อกำหนดความยาวด้านของรูปสามเหลี่ยม $AB = c, AC = b$ และ $BC = a$ ทำให้ได้ความสัมพันธ์

$$r^2 = \frac{a^4b^2}{2(a^2b^2 + a^2c^2 + b^2c^2) - a^4 - b^4 - c^4}$$

สำหรับงานวิจัยนี้มีวัตถุประสงค์เพื่อหาความสัมพันธ์ระหว่างความยาวเส้นมัธยฐานของรูปสามเหลี่ยมและความยาวรัศมีของรูปวงกลมในรูปของความยาวด้านของรูปสามเหลี่ยม พร้อมแสดงการพิสูจน์ความสัมพันธ์ที่ได้โดยใช้เรขาคณิตวิเคราะห์และอัตราส่วนตรีโกณมิติ

ผลการศึกษาพบว่าความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลมที่กำหนดเงื่อนไขข้างต้นเป็นไปตามความสัมพันธ์ของความยาวรัศมีของรูปวงกลมกับความยาวด้านและมุมของรูปสามเหลี่ยมดังนี้

$$r^2 = \frac{a^2}{4\sin^2 C}$$

และเมื่อพิกัดจุด $B(p, q)$ สามารถพิจารณาจากเงื่อนไขของมุม C ที่ทำให้จุด B เปลี่ยนจตุภาค โดยที่ $p = r\cos 2C$ และ $q = a\cos C$ ความสัมพันธ์ของความยาวเส้นมัธยฐานของรูปสามเหลี่ยม AH ในรูปของความยาวด้านของรูปสามเหลี่ยมและความยาวรัศมีของรูปวงกลมเป็นดังนี้

$$|AH| = \frac{1}{2}\sqrt{2r^2 + 4b^2 - 2rp - 4bq}$$

คำสำคัญ: ความสัมพันธ์เชิงระยะทาง

2020 MSC: 51K05, 51M25

*ผู้นำเสนอ (arm611518021@gmail.com)

†ผู้ประพันธ์บรรณกิจ (konvika@tsu.ac.th)

1 บทนำ

เรขาคณิตเป็นพิมพ์เขียวของการสร้างสรรค์และก่อกำเนิดรูปแบบทั้งหมด ทั้งยังเกี่ยวข้องกับจำนวนในปริภูมิระดับพื้นฐานที่แสดงให้เห็นเลขคณิตที่เกี่ยวข้องกับการวัดเชิงเรขาคณิต รวมทั้ง จำนวนในปริภูมิยังแสดงให้เห็นเรขาคณิตเชิงสัดส่วนซึ่งสะท้อนความหมายและแนวคิด สัดส่วนทางเรขาคณิตทำหน้าที่เป็นเครื่องมือออกแบบโดยเฉพาะอย่างยิ่งสำหรับการออกแบบลวดลาย บทบาทสำคัญของเรขาคณิตในลวดลาย เช่น ในการศึกษาลวดลายอิสลาม จะถูกกล่าวถึงโดยสัมพันธ์กับสัดส่วนซึ่งทำหน้าที่เป็นโครงสร้างพื้นฐานของกระบวนการออกแบบ สัดส่วนจะเชื่อมโยงกับเรขาคณิต ซึ่งสามารถมองเห็นวิธีการที่พิสูจน์ที่แสดงว่าสวยงามตามหลักการออกแบบ นอกจากนี้ สัดส่วนทางเรขาคณิตควบคุมลำดับของลวดลาย ทำหน้าที่เป็นตัวกลางระหว่างขั้วของลำดับและความหลากหลาย [1] การศึกษาความสัมพันธ์ระหว่างรูปเรขาคณิตสองมิติและสามมิติจึงเป็นไปอย่างกว้างขวาง ทั้งในระดับพื้นฐานที่ศึกษาในระดับโรงเรียน จนถึงระดับที่เป็นปัญหาคณิตศาสตร์โอลิมปิก อย่างไรก็ตามได้มีงานวิจัยที่นำเสนอการพิสูจน์ความสัมพันธ์เชิงระยะทางของรูปเรขาคณิตในระนาบทำให้ได้สูตรการคำนวณหรือได้แนวทางการสร้างทางเรขาคณิตที่หลากหลาย ในปี พ.ศ. 2561 ภิญโญ มนูศิลป์และยุพร ริมชลการ [2] ได้ศึกษาสมบัติของพื้นที่วงกลมเก้าจุดของรูปสามเหลี่ยมของอาร์คิมีดีส ผลจากการศึกษาพบว่า พื้นที่วงกลมเก้าจุดของรูปสามเหลี่ยมของอาร์คิมีดีสเท่ากับ 1/4 เท่าของพื้นที่วงกลมล้อมรอบรูปสามเหลี่ยมของอาร์คิมีดีสและพื้นที่รูปสามเหลี่ยมของออยเลอร์เท่ากับ 1/4 เท่าของพื้นที่รูปสามเหลี่ยมของอาร์คิมีดีส จากบทความวิจัยทำให้ได้แนวคิดในการสร้างวงกลมและรูปสามเหลี่ยมที่มีสมบัติพิเศษรวมทั้ง ได้เห็นความสัมพันธ์เชิงระยะทางของรูปเรขาคณิตที่นำมาใช้ในการพิสูจน์ นอกจากนี้ ยังมีงานวิจัยในปี ค.ศ. 2024 ของ Parolin and Santos [3] ที่ศึกษาสมบัติบางประการของวงกลมสามวงที่สัมพันธ์กับรูปสามเหลี่ยมโดย วงกลมแต่ละวงอยู่ภายในรูปสามเหลี่ยม สัมผัสกับด้านสองด้านของรูปสามเหลี่ยม และสัมผัสด้านนอกของวงกลมด้านที่สามซึ่งเป็นเส้นผ่านศูนย์กลาง จากการศึกษาได้พบความสัมพันธ์ที่เกี่ยวข้องกับรัศมีของวงกลมอพอลโลเนียสภายในและภายนอกของวงกลมทั้งสามวงนั้น อย่างไรก็ตาม จากการศึกษางานวิจัยของ Paul Goron ที่เผยแพร่ในปี ค.ศ. 2023 [4] ที่ได้แสดงความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลม โดยแสดงการพิสูจน์ความสัมพันธ์ระหว่างความยาวด้านของรูปสามเหลี่ยม ABC ในระนาบ ที่มีเส้นตรง AC เป็นเส้นสัมผัสวงกลมที่มี r เป็นความยาวรัศมี โดยที่จุด B และ C อยู่บนวงกลมดังกล่าว เมื่อกำหนดความยาวด้านของรูปสามเหลี่ยม $AB = c,\ AC = b$ และ $BC = a$ ทำให้ได้ความสัมพันธ์

$$r^2 = \frac{a^4b^2}{2(a^2b^2 + a^2c^2 + b^2c^2) - a^4 - b^4 - c^4}$$

สำหรับงานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) หาความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลมที่กำหนดเงื่อนไขเช่นเดียวกับงานวิจัยของ Paul Goron ในรูปของความยาวรัศมีของรูปวงกลมกับความยาวด้านและมุมของรูปสามเหลี่ยม 2) หาความสัมพันธ์ระหว่างความยาวเส้นมัธยฐานของรูปสามเหลี่ยมในรูปของความยาวรัศมีของรูปวงกลมในรูปและความยาวด้านของรูปสามเหลี่ยม พร้อมแสดงการพิสูจน์ความสัมพันธ์ที่ได้โดยใช้เรขาคณิต

วิเคราะห์และอัตราส่วนตรีโกณมิติ สำหรับประโยชน์ของผลจากการศึกษาครั้งนี้นอกจากจะได้สูตรความสัมพันธ์ r^2 กับความยาวด้านและมุมของรูปสามเหลี่ยมแล้ว ยังสามารถขยายแนวคิดสู่การหาความยาวเส้นมัธยฐานของรูปสามเหลี่ยมซึ่งอ้างอิงจากวงกลมที่สร้างขึ้น

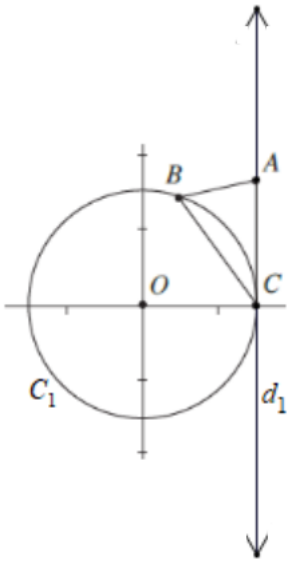
2 ความรู้พื้นฐาน

ในการศึกษาความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลมจำเป็นต้องมีความรู้พื้นฐานเกี่ยวกับเรขาคณิตวิเคราะห์ โดยใช้แนวคิดการพิสูจน์ความสัมพันธ์ดังกล่าวบนระนาบพิกัดฉากซึ่งกำหนดตำแหน่งของจุดต่าง ๆ บนระนาบโดยใช้คู่อันดับของจำนวนจริง ซึ่งเป็นเครื่องมือสำคัญที่ใช้พิสูจน์ทฤษฎีบทในงานวิจัยของ Paul Goron ซึ่งเป็นทฤษฎีหลักหลักที่เป็นฐานในการศึกษาวิจัยในครั้งนี้

ทฤษฎีบท 2.1 [4] *ให้ ABC เป็นรูปสามเหลี่ยมบนระนาบที่มีเส้นตรง AC เป็นเส้นสัมผัสวงกลมที่มี O เป็นจุดศูนย์กลาง โดยที่จุด B และ C อยู่บนวงกลมดังกล่าว เมื่อ $a,\ b,\ c,\ r \in \mathbb{R}^+$ ซึ่ง $AB = c,\ AC = b, BC = a$ และ r เป็นความยาวรัศมีของวงกลม จะได้*

$$r^2 = \frac{a^4b^2}{2(a^2b^2 + a^2c^2 + b^2c^2) - a^4 - b^4 - c^4}$$

พิสูจน์. ให้ $d,\ f,\ p,\ q \in \mathbb{R}$ โดยที่ $d \neq 0,\ f \neq 0$ และให้ C_1 เป็นวงกลมที่มี O เป็นจุดศูนย์กลาง ซึ่งจุด B และ C อยู่บนวงกลม C_1 โดยการหมุนและการเลื่อนขนานซึ่งสมมิติ (isometry) กำหนดพิกัดจุดเป็น $O(0,0),\ B(p,q)$ และ $C(d,0)$ ดังรูปที่ 1



รูปที่ 1: รูปสามเหลี่ยม ABC และวงกลม C_1 ที่มี O เป็นจุดศูนย์กลาง

ให้ d_1 เป็นเส้นสัมผัสวงกลม C_1 ที่จุด C จะได้สมการเส้นตรง d_1 คือ $x = d$ โดยที่จุด A เป็นจุดบนเส้นตรง d_1 ทำให้ได้พิกัดจุด $A(d, f)$ นั่นคือ

$$a^2 = (d-p)^2 + q^2 \tag{2.1}$$

$$b^2 = f^2 \tag{2.2}$$

$$c^2 = (d-p)^2 + (f-q)^2 \tag{2.3}$$

$$r^2 = d^2 = p^2 + q^2 \tag{2.4}$$

ดังนั้น $$c^2 = a^2 + b^2 - 2fq \tag{2.5}$$

ทำให้ได้

$$q = \frac{c^2 - a^2 - b^2}{-2f}$$

นั่นคือ

$$q^2 = \frac{(c^2-a^2-b^2)^2}{4b^2} \tag{2.6}$$

และจะได้ว่า $|q|$ เป็นความสูงของจุด B จาก $\overline{OC}$ ในรูปสามเหลี่ยม OBC

จากรูปที่ 1 ให้ A_1 เป็นพื้นที่ของรูปสามเหลี่ยมหน้าจั่ว OBC ซึ่งมี $\overline{BC}$ เป็นฐาน ดังนั้น

$$A_1 = \frac{a\sqrt{r^2 - \frac{1}{4}a^2}}{2}$$

ให้ A_2 เป็นพื้นที่ของรูปสามเหลี่ยมหน้าจั่ว OBC ซึ่งมี $\overline{OC}$ เป็นฐานและสูง $|q|$ ดังนั้น

$$A_2 = \frac{r|q|}{2}$$

ดังนั้น $\frac{r|q|}{2} = \frac{a\sqrt{r^2-\frac{1}{4}a^2}}{2}$ ทำให้ได้ $|q| = \frac{a\sqrt{4r^2-a^2}}{2r}$

นั่นคือ $$q^2 = a^2 - \frac{a^4}{4r^2} \tag{2.7}$$

สมการที่ (2.6) และ (2.7) จะได้ $\frac{(c^2-a^2-b^2)^2}{4b^2} = a^2 - \frac{a^4}{4r^2}$

เพราะฉะนั้น

$$r^2 = \frac{a^4b^2}{2(a^2b^2 + a^2c^2 + b^2c^2) - a^4 - b^4 - c^4}$$

□

เนื่องจากการหมุนและการเลื่อนขนานสมมิติ (isometry) นั่นคือ การเปลี่ยนพิกัดของรูปสามเหลี่ยมที่เกิดจากการหมุนหรือการเลื่อนขนานในระนาบยังคงสภาพความยาวด้านและขนาดของมุมคงเดิม

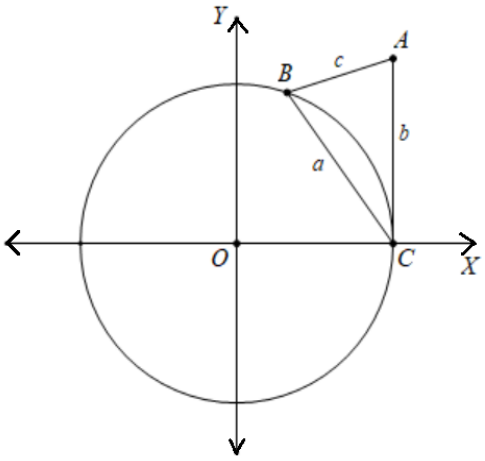
3 ผลการศึกษา

จากการสร้างรูปสามเหลี่ยมและรูปวงกลมตามเงื่อนไขของทฤษฎีบท 2.1 นำไปสู่การพิสูจน์ทฤษฎีบทที่แสดงความสัมพันธ์ของความยาวรัศมีของรูปวงกลมกับความยาวด้านและมุมของรูปสามเหลี่ยมดังต่อไปนี้

ทฤษฎีบท 3.1 *ให้ ABC เป็นรูปสามเหลี่ยมใด ๆ บนระนาบ โดยกำหนดด้านตรงข้ามมุม A, B และ C ยาว a, b และ c ตามลำดับ และมีด้าน AC สัมผัสวงกลมที่มี O เป็นจุดศูนย์กลางที่จุด C โดยที่ทั้งจุด B และ C อยู่บนวงกลมดังกล่าว เมื่อกำหนด r เป็นความยาวรัศมีของวงกลม จะได้*

$$r^2 = \frac{a^2}{4\sin^2 C}$$

พิสูจน์. เนื่องจากการหมุนและการเลื่อนขนานสมมิติ (isometry) สำหรับรูปสามเหลี่ยมและวงกลม โดยไม่เสียนัย สร้างวงกลมที่มี $O(0,0)$ เป็นจุดศูนย์กลาง และ ABC เป็นรูปสามเหลี่ยมใด ๆ ที่มีด้าน AC สัมผัสวงกลมที่มี O เป็นจุดศูนย์กลางที่จุด C ซึ่งเป็นจุดบนแกน X ในจตุภาคที่ 1 โดยที่ทั้งจุด B และ C อยู่บนวงกลมดังกล่าว และสร้างจุดยอด A ที่อยู่นอกวงกลมในจตุภาคที่ 1 ดังรูปที่ 2



รูปที่ 2: การสร้างรูปสามเหลี่ยม ABC และวงกลมที่มี O เป็นจุดศูนย์กลางเพื่อการพิสูจน์

จากรูปสามเหลี่ยม ABC จะได้ $AB = c,\ AC = b, BC = a$ และจากกฎของโคไซน์ จะได้ว่า

$$c^2 = a^2 + b^2 - 2ab\cos C \tag{3.1}$$

จากสมการ (2.5) และสมการ (3.1) จะได้

$$-2ab\cos C = -2fq$$

$$a\cos C = q \tag{3.2}$$

$$q^2 = a^2 \cos^2 C \tag{3.3}$$

จากสมการ (2.7) และ (3.3) จะได้ว่า

$$a^2 - \frac{a^4}{4r^2} = a^2 \cos^2 C$$

$$1 - \frac{a^2}{4r^2} = \cos^2 C$$

$$-\frac{a^2}{4r^2} = \cos^2 C - 1$$

เพราะฉะนั้น

$$r^2 = -\frac{a^2}{4(\cos^2 C - 1)}$$

ดังนั้น

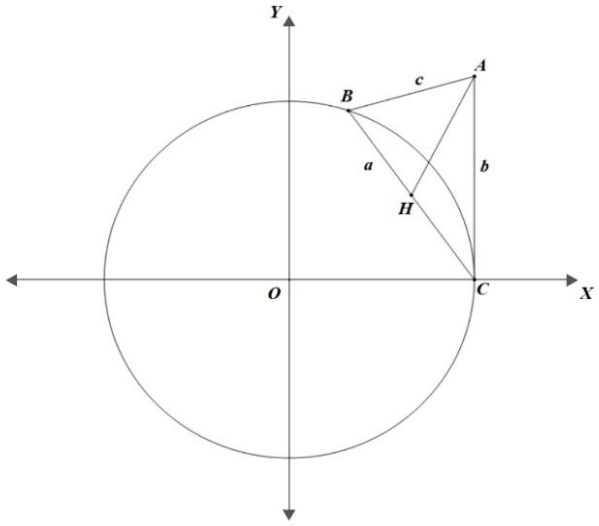
$$r^2 = \frac{a^2}{4 \sin^2 C}$$

□

ทฤษฎีบท 3.2 *ให้ ABC เป็นรูปสามเหลี่ยมใด ๆ บนระนาบ ที่มีด้าน AC สัมผัสวงกลมที่มี O เป็นจุดศูนย์กลางที่จุด C โดยที่ทั้งจุด B และ C อยู่บนวงกลมดังกล่าว เมื่อกำหนดความยาวด้านของรูปสามเหลี่ยม $AC = b$, $BC = a$, r เป็นความยาวรัศมีของวงกลม, $\overline{AH}$ เป็นเส้นมัธยฐานของรูปสามเหลี่ยม ABC และ $|AH|$ เป็นความยาวเส้นมัธยฐาน จะได้*

$$|AH| = \frac{1}{2}\sqrt{2r^2 + 4b^2 - 2rp - 4bq} \text{ เมื่อ } p = r\cos 2C \text{ และ } q = a\cos C$$

พิสูจน์. ลากส่วนของเส้นตรงจากจุด A ไปแบ่งครึ่งด้าน BC ที่จุด H จะได้ว่า $\overline{AH}$ เป็นเส้นมัธยฐานเส้นหนึ่งของรูปสามเหลี่ยม ABC ดังภาพ



รูปที่ 3: การสร้าง $\overline{AH}$ เป็นเส้นมัธยฐานเส้นหนึ่งของรูปสามเหลี่ยม ABC

จากการกำหนดให้ $d,\ f,\ p,\ q \in \mathbb{R}$ โดยที่ $d \neq 0,\ f \neq 0$ และกำหนดพิกัดจุดเป็น $O(0,0)$, $B(p,q)$ และ $C(d,0)$ ซึ่งได้พิกัดจุด $A(d,f)$ และจากพิกัดจุด $B(p,q)$ และ $C(d,0)$ ทำให้ได้พิกัดของจุดกึ่งกลางด้าน H คือ $\left(\frac{p+d}{2},\frac{q}{2}\right)$

จากสมการ (2.4) และจากทฤษฎีบท 3.1 ซึ่ง $r^2 = \frac{a^2}{4\sin^2 C}$ จะได้

$$\frac{a^2}{4\sin^2 C} = p^2 + q^2 \tag{3.4}$$

และจากสมการ (3.2) $$q = a\cos C$$

ทำให้ได้

$$\frac{a^2}{4\sin^2 C} = p^2 + (a\cos C)^2$$

$$p^2 = \frac{a^2}{4\sin^2 C} - a^2\cos^2 C = \frac{a^2}{4\sin^2 C}(1 - 4\sin^2 C\cos^2 C) = r^2(1-\sin^2 2C)$$

$$p^2 = r^2\cos^2 2C$$

นั่นคือ $$p = r\cos 2C$$

ซึ่งจะให้ค่าเป็นจำนวนบวกหรือลบขึ้นอยู่กับค่าของ $\cos 2C$ ดังนั้น พิกัดจุด $B(p,q)$ สามารถพิจารณาจากเงื่อนไขของมุม C ที่ทำให้จุด B เปลี่ยนจตุภาค โดยที่ $p = r\cos 2C$ และ $q = a\cos C$ และจะได้ความยาวเส้นมัธยฐานโดยหาระยะทางระหว่างจุด A และ H นั่นคือ

$$|AH| = \sqrt{\left(d - \frac{p+d}{2}\right)^2 + \left(f - \frac{q}{2}\right)^2}$$

$$= \sqrt{\left(\frac{d-p}{2}\right)^2 + \left(\frac{2f-q}{2}\right)^2}$$

$$= \frac{1}{2}\sqrt{(d-p)^2 + (2f-q)^2}$$

$$= \frac{1}{2}\sqrt{2d^2 + 4f^2 - 2dp - 4fq}$$

จากสมการที่ (2.2) และ (2.4) จะได้

$$|AH| = \frac{1}{2}\sqrt{2r^2 + 4b^2 - 2rp - 4bq} \tag{3.5}$$

ทั้งนี้ ความยาวเส้นมัธยฐานที่ได้จากสมการ (3.5) จะสอดคล้องสำหรับความยาวด้านของรูปสามเหลี่ยม a และ b คู่หนึ่ง ๆ รวมทั้งค่าของ p และ q นั้น ถูกกำหนดชัดเจนเพียงค่าเดียวโดยมุม C

□

4 สรุปผลและอภิปรายผล

จากการศึกษาความสัมพันธ์เชิงระยะทางระหว่างรูปสามเหลี่ยมและรูปวงกลม พบว่า ความสัมพันธ์ระหว่างความยาวด้านของรูปสามเหลี่ยม ABC ใด ๆ บนระนาบ โดยกำหนดด้านตรงข้ามมุม A, B และ C ยาว a, b และ c ตามลำดับ และมีด้าน AC สัมผัสวงกลมที่มี O เป็นจุดศูนย์กลางที่จุด C โดยที่ทั้งจุด B และ C อยู่บนวงกลมดังกล่าว เมื่อกำหนด r เป็นความยาวรัศมีของวงกลม $\overline{AH}$ เป็นเส้นมัธยฐานของรูปสามเหลี่ยม ABC และ $|AH|$ เป็นความยาวเส้นมัธยฐาน ทำให้ได้ความสัมพันธ์ต่อไปนี้

1) $r^2 = \frac{a^2}{4\sin^2 C}$

2) $|AH| = \frac{1}{2}\sqrt{2r^2 + 4b^2 - 2rp - 4bq}$ เมื่อ $p = r\cos 2C$ และ $q = a\cos C$

จากผลของการศึกษาความสัมพันธ์ที่ได้สำหรับค่าของ r^2 ในข้อ 1) พบว่า ถ้าเป็นวงกลมเดียวกันหรือวงกลมที่มีความยาวรัศมีเท่ากันแล้ว ตำแหน่งของจุดสัมผัส C และตำแหน่งของจุด B ที่เปลี่ยนไป รวมทั้ง ความยาวด้านทั้งสามของรูปสามเหลี่ยมและขนาดของมุม C ที่เปลี่ยนไปไม่ส่งผลต่อค่าที่ได้จากความสัมพันธ์ เมื่อการสร้างรูปสามเหลี่ยมและรูปวงกลมตามเงื่อนไขของทฤษฎีบท นอกจากนี้ ความสัมพันธ์ของความยาวเส้นมัธยฐานของรูปสามเหลี่ยมในข้อ 2) พบว่า ความยาวเส้นมัธยฐานที่ได้จะสอดคล้องสำหรับความยาวด้านของรูปสามเหลี่ยม a และ

b คู่หนึ่ง ๆ รวมทั้ง พิกัดจุด $B(p,q)$ ที่เปลี่ยนไปในแต่ละจตุภาค ซึ่งถูกกำหนดชัดเจนเพียงค่าเดียวจากเงื่อนไขของมุม C

กิตติกรรมประกาศ ผู้แต่งขอขอบคุณผู้ทรงคุณวุฒิทุกท่านที่ได้ให้ข้อคิดเห็นและข้อเสนอแนะต่าง ๆ เพื่อปรับปรุงบทความวิจัยนี้

เอกสารอ้างอิง

[1] L. M. Dabbour, Geometric proportions: The underlying structure of design process for Islamic geometric patterns, Frontiers of Architectural Research. 1 (2012) 380--391.

[2] ภิญโญ มนูศิลป์, ยุพร ริมชลการ, สมบัติของพื้นที่วงกลมเก้าจุดของรูปสามเหลี่ยมของอาร์คิมีดีส, Journal of Science and Technology MSU. 37 (1) (2018) 136--150.

[3] R. de S. Parolin, A. D. Santos, Trigonometric solutions that relate geometrical quantities of the triangle and the inscribed circle, Ci. e Nat. 46 (2024) 1--16.

[4] P. Goron, A metric relation on the triangle and the circle, The Mathematical Gazette. 107 (568) (2023) 146--147.

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

Counting Umbilical Points on Generalised Fermat Surfaces

Phemmatad Tansoontorn* and Nithi Rungtanapirom[1,†]

[1]Department of Mathematics and Computer Science, Faculty of Science,
Chulalongkorn University, Bangkok 10330, Thailand

Abstract

Umbilical points are points on a surface where the two principal curvatures are equal. In this work, we use both differential geometry and algebraic geometry to compute the number of real and complex umbilical points of generalised Fermat surfaces of the form $ax^n + by^n + cz^n = 1$ where $n \geq 3$. We found that there are either four real umbilical points if n is odd, and either 0, 2, 4, or 14 real umbilical points if n even. Furthermore, there are at most $n^3 - 4n^2 + 7n$ complex umbilical points on a generalised Fermat surface of degree n.

Keywords: Enumerative geometry, generalised Fermat surfaces, Umbilical points

2020 MSC: 14N10, 53A05

1 Introduction

Curvature is a fundamental concept in differential geometry. For surfaces in the three-dimensional space $\mathbb{R}^3$, their geometry can be stated in terms of the principal curvatures. These curvatures are defined as the eigenvalues of the shape operator and are precisely the minimum and maximum curvatures of the surface among all the directions.

In this work, we are particularly interested in *umbilical points*. These are the points where the principal curvatures are equal. For example, every point on a sphere is umbilical [1], therefore one can regard the umbilical points as the points which are locally spherical.

The study of enumerative geometry of umbilical points can be traced back to the works of George Salmon who, in 1862, proved that a general surface of degree d in $\mathbb{R}^3$ has $10d^3 - 28d^2 + 22d$ complex umbilical points [2]. It is also known that an ellipsoid has four real umbilical points. More recently, in [3], P. Breiding, K. Ranestad and M. Weinstein studied umbilical points, and critical curvature points in general, using algebraic geometry. They defined the *curvature variety* $\mathcal{C}(f)$ associated to a hypersurface $f = 0$ in $\mathbb{R}^d$ as real roots of some polynomial equations, before using it to define complex umbilical points, which we will describe below. In addition, they also used the polynomial equations of the curvature variety to determine all the real and complex umbilical points on a quadric $ax^2 + by^2 + cz^2 = 1$ in $\mathbb{R}^3$, and found that there are 12 complex umbilical points, four of which are real in the case of ellipsoids.

*Speaker (prame_tansuntorn@hotmail.com)
†Corresponding author (nithi.r@chula.ac.th)

This work extends the works of [3] discussed in the previous paragraph to the cases of higher degrees. More precisely, we consider surfaces of the form

$$ax^n + by^n + cz^n = 1$$

where $a, b, c \in \mathbb{R}$ and $n \geq 3$. We call them *generalised Fermat surfaces*, alluding to the fact that, when $a = b = c = 1$, the projectivisation of this surface is known as Fermat varieties. Using the same method as in [3], we derive a polynomial equation for umbilical points on generalised Fermat surfaces. Then, by finding all the possible zeroes of the polynomial, we can determine all the real and complex umbilical points on the surfaces.

Umbilical points are interesting objects of study in both differential geometry and applied geometry. They are points on a surface for which a local representation by means of lines of curvature is complicated, which has important ramifications in applications such as computer aided geometric design, shape analysis and surface interrogation [4]. On the theoretical side, several works have been done regarding the geometry of umbilical points on a surface. This includes [5], where local geometry of umbilical points on a surface in $\mathbb{R}^3$ is considered. An algorithm for computing the critical curvature points, which are generalisation of umbilical points to higher dimensions, is provided in [6]. In [7], it is shown that the set of umbilical points on an immersed surface in a three-dimensional space form can have positive measure.

2 Preliminaries

2.1 Differential Geometry of Surfaces

We shall restate some basic results in differential geometry here. Our notation and terminology follow that of [8].

Definition 2.1. A *(regular) surface* is a subset $S \subseteq \mathbb{R}^3$ such that each point $p \in S$ has a neighbourhood U of p, open in S, such that U is diffeomorphic to an open subset $V \subseteq \mathbb{R}^2$, i.e. it is a two-dimensional embedded submanifold of $\mathbb{R}^3$.

One can construct a surface given an equation of the form $f(x, y, z) = 0$ thanks to the following theorem.

Theorem 2.2 (Regular Level Set Theorem)**.** *Let $U \subseteq \mathbb{R}^3$ be an open set and let $f\colon U \to \mathbb{R}$ be a smooth function. We say that $\lambda \in \mathbb{R}$ is a* regular value *of f if $\lambda \in f(U)$ and the Jacobian matrix*

$$df_p = \begin{pmatrix} \frac{\partial f}{\partial x}(p) & \frac{\partial f}{\partial y}(p) & \frac{\partial f}{\partial z}(p) \end{pmatrix} \neq (0, 0, 0)$$

is non-zero for all $p \in f^{-1}(\lambda)$. If this is the case, then $f^{-1}(\lambda)$ is a regular surface.

For a proof, see [8, Theorem 9.9].

Proposition 2.3. *Suppose $a, b, c \in \mathbb{R}$, not all zero, and suppose $n \geq 2$ is an integer. The equation*

$$ax^n + by^n + cz^n = 1$$

defines a regular surface when n is odd, or, when n is even and at least one of a, b, c is positive.

Proof. Set $f(x,y,z) = ax^n + by^n + cz^n - 1$. By the regular level set theorem, it suffices to show that 0 is a regular value of f. The Jacobian matrix of f is

$$df_p = \begin{pmatrix} nax^{n-1} & nby^{n-1} & ncz^{n-1} \end{pmatrix},$$

which is zero only at $(x,y,z) = (0,0,0)$. But $f(0,0,0) = -1 \neq 0$, hence $(0,0,0)$ does not lie on the set $f^{-1}(0)$, and so the Jacobian is non-zero on $f^{-1}(0)$.

The only obstruction left to $f^{-1}(\lambda)$ being a regular surface is that $f^{-1}(0)$ might be empty. If n is odd, then this does not happen since we can take one of the non-zero term, say cz^n, and rearrange the equation into

$$z = \left(\frac{1 - ax^n - by^n}{c}\right)^{1/n},$$

which is always solvable over the reals.

For even n, we need at least one of a, b, c to be positive, otherwise $ax^n + by^n + cz^n < 1$. So we may suppose $c > 0$. Rearranging the equation gives

$$z = \pm\left(\frac{1 - ax^n - by^n}{c}\right)^{1/n}.$$

The region where $1 - ax^n - by^n \geq 0$ is always non-empty, at least in a neighbourhood of $(x,y) = (0,0)$. This proves the claim. □

We now turn to the curvature of surfaces in $\mathbb{R}^3$. If S is a surface and $p \in S$, let T_pS be the tangent space at p. The inclusion $S \hookrightarrow \mathbb{R}^3$ gives rise to an injection $T_pS \hookrightarrow T_p\mathbb{R}^3$, and so we may use the standard frame $(\partial/\partial x, \partial/\partial y, \partial/\partial z)$ for vector fields along S.

Theorem 2.4. *The directional derivative $D_{X_p}Y$ at p of a smooth vector field $Y = \sum b^i \partial/\partial x^i$ on $\mathbb{R}^3$ in the direction $X_p = \sum_i a^i \left.\frac{\partial}{\partial x^i}\right|_p \in T_p\mathbb{R}^3$ is defined by*

$$D_{X_p}Y \coloneqq \sum_{i=1}^{3} (X_p b^i) \left.\frac{\partial}{\partial x^i}\right|_p = \sum_{i=1}^{3} \left(\sum_{j=1}^{3} a^j \left.\frac{\partial b^i}{\partial x^j}\right|_p\right) \left.\frac{\partial}{\partial x^i}\right|_p.$$

This is the local description of the Levi-Civita connection $\nabla_X Y =: D_X Y$, and allows us to define the shape operator at a point.

Definition 2.5. Let p be a point on a surface S in $\mathbb{R}^3$ and let N be a smooth unit normal vector field on S. The *shape operator* or the *Weingarten map* is a map $L_p \colon T_pS \to T_pS$ defined by

$$L_p(X_p) = -D_{X_p}N.$$

The vector $L_p(X_p)$ is orthogonal to N_p, which is a normal vector at p [8, p.30]. Therefore the range of this map is indeed T_pS.

The shape operator is self-adjoint [8, p.31]. Hence all its eigenvalues are real. The eigenvalues are called the *principal curvatures* and the eigenvectors are the *principal directions*. This leads us to the next definition.

Definition 2.6. An *umbilical point* on a surface S is a point $p \in S$ at which both principal curvatures are equal.

Note that there are two possible choices for a smooth unit normal vector field on an orientable surface. Namely, if N is a smooth unit normal vector field, then $-N$ is also a smooth unit normal vector field. If we compute the principal curvatures using the two normal vector fields N and $-N$, then the resulting principal curvatures will have different signs. However, umbilical points are independent of the unit normal vector field chosen.

We now derive an equation for umbilical points for a surface $S = f^{-1}(0)$ where 0 is a regular value of f. The gradient $N = \nabla f = \sum_{i=1}^{3} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^i}$ is a nowhere-vanishing normal vector field along the regular surface. It is smooth because f is smooth. Since it is nowhere-vanishing, the normal vector field $\tilde{N} = N/\|\nabla f\|$ is smooth and of unit length.

Now, let $p \in S$ be a point on S and $X_p = \sum_i a^i \partial/\partial x^i\big|_p$ be a tangent vector at p. The directional derivative of $\tilde{N}$ in the direction of X_p is

$$D_{X_p}\tilde{N} = D_{X_p}\left(\sum_{i=1}^{3} \frac{1}{\|\nabla f\|}\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^i}\right) = \sum_{j=1}^{3}\left(\sum_{i=1}^{3} a^i \frac{\partial}{\partial x^i}\left(\frac{1}{\|\nabla f\|}\frac{\partial f}{\partial x^j}\right)\Big|_p\right)\frac{\partial}{\partial x^j}\Big|_p.$$

The j-th coordinate of $D_{X_p}\tilde{N}$ is equal to

$$\sum_{i=1}^{3} a^i \frac{\partial}{\partial x^i}\left(\frac{1}{\|\nabla f\|}\frac{\partial f}{\partial x^j}\right)\Big|_p = \sum_{i=1}^{3} a^i \frac{1}{\|\nabla f(p)\|}\frac{\partial^2 f}{\partial x^i \partial x^j}(p) + a^i \frac{\partial}{\partial x^i}\frac{1}{\|\nabla f\|}(p)\frac{\partial f}{\partial x^j}(p).$$

Hence we may write a matrix representation for $D_{X_p}\tilde{N}$ as

$$\begin{aligned} D_{X_p}\tilde{N} = \frac{1}{\|\nabla f(p)\|}&\begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(p) & \frac{\partial^2 f}{\partial y \partial x}(p) & \frac{\partial^2 f}{\partial z \partial x}(p) \\ \frac{\partial^2 f}{\partial x \partial y}(p) & \frac{\partial^2 f}{\partial y^2}(p) & \frac{\partial^2 f}{\partial z \partial y}(p) \\ \frac{\partial^2 f}{\partial x \partial z}(p) & \frac{\partial^2 f}{\partial y \partial z}(p) & \frac{\partial^2 f}{\partial z^2}(p) \end{bmatrix}\begin{bmatrix} a^1 \\ a^2 \\ a^3 \end{bmatrix} \\ &+ \begin{bmatrix} \frac{\partial f}{\partial x}(p)\frac{\partial}{\partial x}\frac{1}{\|\nabla f\|}(p) & \frac{\partial f}{\partial x}(p)\frac{\partial}{\partial y}\frac{1}{\|\nabla f\|}(p) & \frac{\partial f}{\partial x}(p)\frac{\partial}{\partial z}\frac{1}{\|\nabla f\|}(p) \\ \frac{\partial f}{\partial y}(p)\frac{\partial}{\partial x}\frac{1}{\|\nabla f\|}(p) & \frac{\partial f}{\partial y}(p)\frac{\partial}{\partial y}\frac{1}{\|\nabla f\|}(p) & \frac{\partial f}{\partial y}(p)\frac{\partial}{\partial z}\frac{1}{\|\nabla f\|}(p) \\ \frac{\partial f}{\partial z}(p)\frac{\partial}{\partial x}\frac{1}{\|\nabla f\|}(p) & \frac{\partial f}{\partial z}(p)\frac{\partial}{\partial y}\frac{1}{\|\nabla f\|}(p) & \frac{\partial f}{\partial z}(p)\frac{\partial}{\partial z}\frac{1}{\|\nabla f\|}(p) \end{bmatrix}\begin{bmatrix} a^1 \\ a^2 \\ a^3 \end{bmatrix}. \end{aligned}$$

We claim that the term in the second line in the above equation is a normal vector to S. To see this, let $B_p = \begin{bmatrix} b^1 & b^2 & b^3 \end{bmatrix}^T$ be any tangent vector at p, so that $B_p^T \nabla f = 0$. For simplicity, we employ the notation f_x for the partial derivative of f by x, etc. Then we can calculate the dot product between them as follows:

$$\begin{aligned} &\begin{bmatrix} b^1 & b^2 & b^3 \end{bmatrix}\begin{bmatrix} f_x\|\nabla f\|_x^{-1} & f_x\|\nabla f\|_y^{-1} & f_x\|\nabla f\|_z^{-1} \\ f_y\|\nabla f\|_x^{-1} & f_y\|\nabla f\|_y^{-1} & f_y\|\nabla f\|_z^{-1} \\ f_z\|\nabla f\|_x^{-1} & f_z\|\nabla f\|_y^{-1} & f_z\|\nabla f\|_z^{-1} \end{bmatrix}\begin{bmatrix} a^1 \\ a^2 \\ a^3 \end{bmatrix} \\ &\quad = \begin{aligned}[t] &a^1 b^1 f_x\|\nabla f\|_x^{-1} + a^2 b^1 f_x\|\nabla f\|_y^{-1} + a^3 b^1 f_x\|\nabla f\|_z^{-1} \\ &+a^1 b^2 f_y\|\nabla f\|_x^{-1} + a^2 b^2 f_y\|\nabla f\|_y^{-1} + a^3 b^2 f_y\|\nabla f\|_z^{-1} \\ &+a^1 b^3 f_z\|\nabla f\|_x^{-1} + a^2 b^3 f_z\|\nabla f\|_y^{-1} + a^3 b^3 f_z\|\nabla f\|_z^{-1}. \end{aligned} \end{aligned}$$

Since $B_p^T \nabla f = 0$, the sum along each column is zero. Therefore the dot product is zero, and so this proves the claim. Thus we may write the directional derivative of $\tilde{N}$ at X_p as

$$D_{X_p}\tilde{N} = \frac{1}{\|\nabla f(p)\|} H f_p X_p + \mu \nabla f, \tag{2.1}$$

where

$$Hf_p = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(p) & \frac{\partial^2 f}{\partial y \partial x}(p) & \frac{\partial^2 f}{\partial z \partial x}(p) \\ \frac{\partial^2 f}{\partial x \partial y}(p) & \frac{\partial^2 f}{\partial y^2}(p) & \frac{\partial^2 f}{\partial z \partial y}(p) \\ \frac{\partial^2 f}{\partial x \partial z}(p) & \frac{\partial^2 f}{\partial y \partial z}(p) & \frac{\partial^2 f}{\partial z^2}(p) \end{bmatrix}$$

is the Hessian matrix of f at p, and μ is a real constant. If X_p is a principal direction at p with principal curvature λ, then $L(X_p) = \lambda X_p$ and so

$$\begin{aligned} -L(X_p) = D_{X_p}\tilde{N} &= \frac{1}{\|\nabla f(p)\|} Hf_p X_p + \mu \nabla f \\ -\lambda X_p &= \frac{1}{\|\nabla f(p)\|} Hf_p X_p + \mu \nabla f \\ 0 &= Hf_p X_p + \lambda \|\nabla f(p)\| X_p + \mu \|\nabla f(p)\| \nabla f. \end{aligned}$$

We have just proved one implication of the following lemma.

Lemma 2.7. *Let $X_p = \begin{bmatrix} x^1 & x^2 & x^3 \end{bmatrix}^T$ be a tangent vector at p of a regular surface $S = f^{-1}(0)$. Then X_p is a principal direction at p if and only if it satisfies the equation*

$$Hf_p \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} + y_1 \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} + y_2 \nabla f = 0.$$

If so, then the principal curvature at p along X_p is equal to $y_1/\|\nabla f\|$.

Proof. The condition is necessary as shown above. On the other hand, if X_p satisfies $Hf_p X_p + y_1 X_p + y_2 \nabla f = 0$, then by taking the directional derivative of $\tilde{N}$ along the direction X_p as in (2.1), we have

$$D_{X_p}\tilde{N} = \frac{1}{\|\nabla f\|} Hf_p X_p + \mu \nabla f.$$

Substituting $Hf_p X_p = -y_1 X_p - y_2 \nabla f$ gives

$$\begin{aligned} D_{X_p}\tilde{N} &= -\frac{y_1}{\|\nabla f\|} X_p - \frac{y_2}{\|\nabla f\|} \nabla f + \mu \nabla f \\ D_{X_p}\tilde{N} + \frac{y_1}{\|\nabla f\|} X_p &= \left(\mu - \frac{y_2}{\|\nabla f\|} \right) \nabla f. \end{aligned}$$

Since the shape operator $L_p(X_p) = -D_{X_p}\tilde{N}$ takes values in T_pS, it follows that the left hand side is a vector in T_pS, whereas the right hand side is orthogonal to T_pS. This is only possible if $\mu - \frac{y_2}{\|\nabla f\|} = 0$. Hence $D_{X_p}\tilde{N} = -\frac{y_1}{\|\nabla f\|} X_p$, and so X_p is an eigenvector of L_p with eigenvalue $\frac{y_1}{\|\nabla f\|}$. □

The set of equations describing a unit tangent vector at a point (x, y, z) on a surface given by $S = \{f = 0\}$ are

$$f(x, y, z) = 0 \tag{2.2}$$

$$u_1 \frac{\partial f}{\partial x} + u_2 \frac{\partial f}{\partial y} + u_3 \frac{\partial f}{\partial z} = 0 \tag{2.3}$$

$$u_1^2 + u_2^2 + u_3^2 - 1 = 0. \tag{2.4}$$

The second and the third equation establish that $u = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T$ is a unit-length tangent vector in T_pS. To get principal curvatures, we impose the equation

$$Hf_p \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + y_1 \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + y_2 \nabla f = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{2.5}$$

It says that u is an eigenvector of the shape operator with eigenvalue $y_1/\|\nabla f\|$, with some normal component given by y_2, by Lemma 2.7.

2.2 Complex Umbilical Points

We use the following definition for complex umbilical points, following [3]. Write $u = (u_1, u_2, u_3)$ and $y = (y_1, y_2)$. The equation for principal curvatures in Lemma 2.7 involves the norm of the gradient

$$\|\nabla f\| = \sqrt{\sum_{i=1}^{3} \left(\frac{\partial f}{\partial x^i}\right)^2},$$

which is not a polynomial in general. To get an algebraic description of principal curvatures entirely in terms of polynomials, we define an auxiliary equation

$$\lambda^2 \|\nabla f\|^2 - 1 = \lambda^2 \left(\sum_{i=1}^{3} \left(\frac{\partial f}{\partial x^i}\right)^2 \right) - 1 = 0. \tag{2.6}$$

For a surface defined by a polynomial f, this is indeed a polynomial equation. Then, with the notation of Lemma 2.7, a principal curvature at the point (x, y, z) of the surface $S = \{f = 0\}$ is then equal to $y_1/\|\nabla f\| = \pm\lambda y_1$. The ambiguity of the sign does not concern us here, because principal curvatures are defined up to sign.

Definition 2.8. The *curvature variety* $\mathcal{C}(f)$ of a surface $S = \{f = 0\}$ is a real algebraic variety such that

$$\mathcal{C}(f) \coloneqq \{(x, \lambda, u, y) \in S \times \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^2 : (x, \lambda, u, y) \text{ satisfies the equations (2.2)-(2.6)}\}.$$

Denote by π the projection $\pi \colon \mathcal{C}(f) \to S$ to the first coordinate. An umbilical point $x \in S$ is by definition a point such that there exist $(x, \lambda, u, y), (x, \lambda', u', y') \in \pi^{-1}(x)$ with $u' \neq \pm u$ such that $y_1 = y_1'$.

Definition 2.9. For a surface $S = \{f = 0\}$, we define $U(f)$ to be the set of its umbilical points.

We let $\mathcal{P}_{3,n}$ denote the vector space of real polynomials in three variables of degree at most n. Then we consider the space of smooth hypersurfaces and their umbilical points:

$$U_{3,n} \coloneqq \bigcup_{f \in \mathcal{P}_{3,n} : \mathrm{Sing}(f) = \varnothing} \{f\} \times U(f).$$

Definition 2.10. A *complex umbilical point* x of a hypersurface $X = \{f = 0\}$ is the point such that $(f, x) \in \overline{U_{3,n}}$, where $\overline{U_{3,n}}$ is the complex Zariski closure of $U_{3,n}$.

For more information on complex Zariski closure, see [9].

3 Main Results

3.1 Polynomial for Umbilical Points

We now turn towards studying umbilical points on generalised Fermat surfaces.

Definition 3.1. A *generalised Fermat surface* is a surface given by the equation

$$ax^n + by^n + cz^n = 1,$$

where a, b, c and n satisfy the conditions of Proposition 2.3, so that it is a regular surface.

For such surfaces, the equations (2.2)-(2.6) take the form

$$ax^n + by^n + cz^n = 0 \tag{3.1}$$

$$u_1 ax^{n-1} + u_2 by^{n-1} + u_3 cz^{n-1} = 0 \tag{3.2}$$

$$u_1^2 + u_2^2 + u_3^2 - 1 = 0 \tag{3.3}$$

$$n(n-1)\begin{bmatrix} ax^{n-2} & 0 & 0 \\ 0 & by^{n-2} & 0 \\ 0 & 0 & cz^{n-2} \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + y_1\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + y_2\begin{bmatrix} nax^{n-1} \\ nby^{n-1} \\ ncz^{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.4}$$

$$n^2\lambda^2(a^2x^{2(n-1)} + b^2y^{2(n-1)} + c^2z^{2(n-1)}) - 1 = 0. \tag{3.5}$$

In this section, we will derive a polynomial in the variable y_1 such that its roots correspond to principal curvatures at the point (x, y, z). To ease notation, we use cyclic notation with the convention that ax, by and cz occur together in the cyclic sum.

In what follows, let S be a generalised Fermat surface given by the equation $ax^n+by^n+cz^n = 1$.

Lemma 3.2. *For each point $(x, y, z) \in S$, define the polynomial*

$$m(y_1) = \mu_0 y_1^2 + \mu_1 y_1 + \mu_2, \tag{3.6}$$

where

$$\mu_0 = a^2x^{2(n-1)} + b^2y^{2(n-1)} + c^2z^{2(n-1)}, \tag{3.7}$$

$$\mu_1 = n(n-1)\sum_{\text{cyc}} a^2x^{2(n-1)}(by^{n-2} + cz^{n-2}) \tag{3.8}$$

$$\mu_2 = n^2(n-1)^2\sum_{\text{cyc}} a^2x^{2(n-1)}by^{n-2}cz^{n-2}. \tag{3.9}$$

Then y_1 is a root of this polynomial if and only if $y_1/\|\nabla f(x)\|$ is a principal curvature of S at (x, y, z).

Proof. First suppose $y_2 \neq 0$ in (3.4). Then we have

$$\begin{bmatrix} n(n-1)ax^{n-2} + y_1 & 0 & 0 \\ 0 & n(n-1)by^{n-2} + y_1 & 0 \\ 0 & 0 & n(n-1)cz^{n-2} + y_1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = -y_2\begin{bmatrix} nax^{n-1} \\ nby^{n-1} \\ ncz^{n-1} \end{bmatrix}. \tag{3.10}$$

Multiply each row of (3.10) by the other terms in the diagonal of the left matrix, so that the matrix in the left hand side is a scalar multiple of the identity matrix, we obtain

$$\prod_{\text{cyc}}(n(n-1)ax^{n-2}+y_1)\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}=-y_2\begin{bmatrix}nax^{n-1}(n(n-1)by^{n-2}+y_1)(n(n-1)cz^{n-2}+y_1)\\nby^{n-1}(n(n-1)ax^{n-2}+y_1)(n(n-1)cz^{n-2}+y_1)\\ncz^{n-1}(n(n-1)ax^{n-2}+y_1)(n(n-1)by^{n-2}+y_1)\end{bmatrix}. \tag{3.11}$$

Left-multiply the above equation by $\begin{bmatrix}ax^{n-1} & by^{n-1} & cz^{n-1}\end{bmatrix}$, and use the fact that $u_1ax^{n-1}+u_2by^{n-1}+u_3cz^{n-1}=0$ to make the left hand side become zero. Since $y_2\neq 0$, we must have

$$0=\begin{bmatrix}ax^{n-1} & by^{n-1} & cz^{n-1}\end{bmatrix}\begin{bmatrix}nax^{n-1}(n(n-1)by^{n-2}+y_1)(n(n-1)cz^{n-2}+y_1)\\nby^{n-1}(n(n-1)ax^{n-2}+y_1)(n(n-1)cz^{n-2}+y_1)\\ncz^{n-1}(n(n-1)ax^{n-2}+y_1)(n(n-1)by^{n-2}+y_1)\end{bmatrix}. \tag{3.12}$$

After multiplying out, the right hand side is the polynomial (3.6).

We now consider the case when $y_2=0$. The equation (3.10) implies that

$$(n(n-1)ax^{n-2}+y_1)u_1=(n(n-1)by^{n-2}+y_1)u_2=(n(n-1)cz^{n-2}+y_1)u_3=0.$$

Since $u=(u_1,u_2,u_3)^T\neq 0$, at least one of the terms

$$(n(n-1)ax^{n-2}+y_1),(n(n-1)by^{n-2}+y_1),(n(n-1)cz^{n-2}+y_1)$$

is zero. There are two possible cases: either at least two of the terms are zero, or only one term is. In the first case, we see that the polynomial in (3.12) must be zero, and hence $m(y_1)=0$. For the second case, we may suppose that only $n(n-1)ax^{n-2}+y_1=0$. This implies $u_2=u_3=0$, and from (3.2) we must have $u_1ax^{n-1}=0$. But since $u_1\neq 0$, it follows that $ax^{n-1}=0$. Therefore the polynomial in (3.12) is also zero in this case, and so $m(y_1)=0$.

From the above proof, we see that if λy_1 is a principal curvature of S at (x,y,z), then y_1 satisfies the equations (3.1) – (3.5) and so $m(y_1)=0$. Since there are two principal curvatures to every point on a surface in $\mathbb{R}^3$ (counting multiplicity), and $\deg m=2$, it follows that all the roots of m correspond to principal curvatures of S at (x,y,z). □

Remark 3.3. It is easy to determine the umbilical points of a generalised Fermat surface $ax^n+by^n+cz^n=1$ if some of the constants a,b,c are zero. If two of them are zero, then the surface is a plane, and it is known that every point is umbilical because the curvature is zero.

If only one of them is zero, say $c=0$, then $m(y_1)$ reduces to

$$m(y_1)=(a^2x^{2(n-1)}+b^2y^{2(n-1)})y_1^2+n(n-1)\left(a^2x^{2(n-1)}by^{n-2}+b^2y^{2(n-1)}ax^{n-2}\right)y_1.$$

One root of $m(y_1)$ is 0, and so one of the principal curvatures is $y_1=0$. We obtain an umbilical point if the other root is also 0, i.e.

$$\begin{aligned}0&=n(n-1)\left(a^2x^{2(n-1)}by^{n-2}+b^2y^{2(n-1)}ax^{n-2}\right)\\0&=ax^{n-2}by^{n-2}(ax^n+by^n).\end{aligned}$$

By the surface equation, $ax^n+by^n=1$, and so this is zero if and only if $x^{n-2}y^{n-2}=0$. In particular, a cylinder $x^2+y^2=1$ has no umbilical points, whereas higher-degree cylinders

$$x^{2n}+y^{2n}=1,\quad (n\geq 2)$$

have umbilical points along the lines $(0,\pm 1,z)$ and $(\pm 1,0,z)$.

From now on we suppose that $abc \neq 0$ for a generalised Fermat surface $ax^n + by^n + cz^n = 1$.

Corollary 3.4. *Umbilical points of a generalised Fermat surface are precisely the points such that the determinant of $m(y_1)$ is zero, i.e. they are the zeroes of the equation*

$$\Delta := \mu_1^2 - 4\mu_0\mu_2 = 0. \tag{3.13}$$

Proof. This follows from Theorem 3.2 and the fact that umbilical points are points where the two principal curvatures are equal. □

3.2 Counting Real Umbilical Points

We now determine the real roots of $\Delta = 0$, which will gives us all the real umbilical points. Since counting umbilical points on quadrics has already been done in [3], we now suppose that $n \geq 3$. A direct computation gives

$$\begin{aligned}\Delta = n^2(n-1)^2\Bigg(&\sum_{\text{cyc}} a^4x^{4(n-1)}(by^{n-2} - cz^{n-2})^2 \\ &- \sum_{\text{cyc}} 2a^2x^{2(n-1)}b^2y^{2(n-1)}(by^{n-2} - cz^{n-2})(cz^{n-2} - ax^{n-2})\Bigg).\end{aligned} \tag{3.14}$$

This is done with the aid of Macaulay2. (The code is available upon request.) It can be further factored over the complex numbers:

$$\begin{aligned}\Delta = n^2(n-1)^2 &\left(ax^{n-1}\sqrt{by^{n-2} - cz^{n-2}} + by^{n-1}\sqrt{cz^{n-2} - ax^{n-2}} + cz^{n-1}\sqrt{ax^{n-2} - by^{n-2}}\right) \\ &\left(ax^{n-1}\sqrt{by^{n-2} - cz^{n-2}} - by^{n-1}\sqrt{cz^{n-2} - ax^{n-2}} + cz^{n-1}\sqrt{ax^{n-2} - by^{n-2}}\right) \\ &\left(ax^{n-1}\sqrt{by^{n-2} - cz^{n-2}} + by^{n-1}\sqrt{cz^{n-2} - ax^{n-2}} - cz^{n-1}\sqrt{ax^{n-2} - by^{n-2}}\right) \\ &\left(ax^{n-1}\sqrt{by^{n-2} - cz^{n-2}} - by^{n-1}\sqrt{cz^{n-2} - ax^{n-2}} - cz^{n-1}\sqrt{ax^{n-2} - by^{n-2}}\right).\end{aligned} \tag{3.15}$$

If (x, y, z) is an umbilical point, then $\Delta = 0$ and thus one of the factors in (3.15) is zero. We claim that, for umbilical points, two of the factors vanish at the same time.

Firstly, notice that

$$(ax^{n-2} - by^{n-2}) + (by^{n-2} - cz^{n-2}) + (cz^{n-2} - ax^{n-2}) = 0.$$

Since the sum is zero, either all of the expressions

$$ax^{n-2} - by^{n-2}, \quad by^{n-2} - cz^{n-2}, \quad cz^{n-2} - ax^{n-2}$$

are zero, or one of them differs in sign from the other two. In what follows, we assume that one of the terms, say $cz^{n-2} - ax^{n-2}$, is non-positive and the other term is non-negative. Hence $\sqrt{cz^{n-2} - ax^{n-2}}$ is either zero or purely imaginary. Now suppose that

$$ax^{n-1}\sqrt{by^{n-2} - cz^{n-2}} + by^{n-1}\sqrt{cz^{n-2} - ax^{n-2}} + cz^{n-1}\sqrt{ax^{n-2} - by^{n-2}} = 0.$$

Then, by comparing the imaginary part,

$$\begin{aligned}&\mathrm{Im}(ax^{n-1}\sqrt{by^{n-2}-cz^{n-2}}+by^{n-1}\sqrt{cz^{n-2}-ax^{n-2}}+cz^{n-1}\sqrt{ax^{n-2}-by^{n-2}})\\&=\mathrm{Im}(by^{n-1}\sqrt{cz^{n-2}-ax^{n-2}})=0.\end{aligned}$$

This implies that $by^{n-1}\sqrt{cz^{n-2}-ax^{n-2}}=0$, and so

$$ax^{n-1}\sqrt{by^{n-2}-cz^{n-2}}+cz^{n-1}\sqrt{ax^{n-2}-by^{n-2}}=0.$$

In other words, (x,y,z) is a common zero of

$$\left(ax^{n-1}\sqrt{by^{n-2}-cz^{n-2}}+by^{n-1}\sqrt{cz^{n-2}-ax^{n-2}}+cz^{n-1}\sqrt{ax^{n-2}-by^{n-2}}\right)$$

and

$$\left(ax^{n-1}\sqrt{by^{n-2}-cz^{n-2}}-by^{n-1}\sqrt{cz^{n-2}-ax^{n-2}}+cz^{n-1}\sqrt{ax^{n-2}-by^{n-2}}\right).$$

The other cases can be treated similarly. This shows that a zero of $\mu_1^2-4\mu_0\mu_2$ is a common zero of two of the factors in the right hand side of (3.15), which proves the claim.

Hence, to find umbilical points, we determine all pairwise common zeros of the four factors in (3.15). For example, if we consider the zeroes of

$$\begin{cases}ax^{n-1}\sqrt{by^{n-2}-cz^{n-2}}+by^{n-1}\sqrt{cz^{n-2}-ax^{n-2}}+cz^{n-1}\sqrt{ax^{n-2}-by^{n-2}}=0,\\ax^{n-1}\sqrt{by^{n-2}-cz^{n-2}}-by^{n-1}\sqrt{cz^{n-2}-ax^{n-2}}+cz^{n-1}\sqrt{ax^{n-2}-by^{n-2}}=0,\end{cases}\tag{3.16}$$

we obtain

$$by^{n-1}\sqrt{cz^{n-2}-ax^{n-2}}=0.\tag{3.17}$$

After substitution, the equation $\Delta=0$ becomes

$$0=(ax^{n-1}\sqrt{by^{n-2}-cz^{n-2}})^2-(cz^{n-1}\sqrt{ax^{n-2}-by^{n-2}})^2,$$

which is equivalent to

$$(ax^{n-1})^2(by^{n-2}-cz^{n-2})=(cz^{n-1})^2(ax^{n-2}-by^{n-2}).\tag{3.18}$$

Conversely, every solution (x,y,z) of (3.17) and (3.18) is a common zero of two of the factors of (3.15), and so is an umbilical point. Consequently, finding $(x,y,z)\in S$ satisfying (3.16) is equivalent to finding $(x,y,z)\in S$ satisfying both (3.17) and (3.18). From (3.17), we divide this into two cases depending on whether $y=0$ or $cz^{n-2}-ax^{n-2}=0$.

- *(Case I)* If $y=0$, then $(ax^{n-1})^2(-cz^{n-2})=(cz^{n-1})^2(ax^{n-2})$. By rearranging the equation, we get

  $$acx^{n-2}z^{n-2}(ax^n+cz^n)=0.$$

  But $ax^n+cz^n=1$ by the surface equation, hence either $x=0$ or $z=0$.

  First consider the case when n is odd. If $x=0$, then $cz^n-1=0$ and so the umbilical point is $(0,0,\sqrt[n]{1/c})$, and if $z=0$ then $ax^n-1=0$ and so the umbilical point is $(\sqrt[n]{1/a},0,0)$. Hence we get two umbilical points $(0,0,\sqrt[n]{1/c}),(\sqrt[n]{1/a},0,0)$.

  Of course, if n is even, then taking the root is subjected to $a>0$ or $c>0$ and we get two roots $(\pm\sqrt[n]{1/a},0,0)$ or $(0,0,\pm\sqrt[n]{1/c})$, respectively.

- *(Case II)* If $cz^{n-2} - ax^{n-2} = 0$, then from the first equation of (**??**),

$$(ax^{n-1})^2(by^{n-2} - ax^{n-2}) = (ax^{n-2}z)^2(ax^{n-2} - by^{n-2}).$$

This is equivalent to

$$(ax^{n-2})^2(by^{n-2} - ax^{n-2})(x^2 + z^2) = 0.$$

If $x = 0$, then also $z = 0$. Hence we get the umbilical point $(0, \sqrt[n]{1/b}, 0)$, with the same condition to the above that $b > 0$ when n is even.

If $x^2 + z^2 = 0$, then this occurs only when $(x, y, z) = (0, \sqrt[n]{1/b}, 0)$. Thus we are left with the case when $by^{n-2} - ax^{n-2} = 0$.

Hence $ax^{n-2} = by^{n-2} = cz^{n-2}$. In the case when n is odd, by solving this equation on the generalised Fermat surface, we get

$$\begin{aligned} ax^n + by^n + cz^n - 1 &= 0 \\ ax^n + b\left(\frac{\sqrt[n-2]{a}}{\sqrt[n-2]{b}}x\right)^n + c\left(\frac{\sqrt[n-2]{a}}{\sqrt[n-2]{c}}x\right)^n &= 1 \\ \sqrt[n-2]{a^n}x^n\left(\frac{a}{\sqrt[n-2]{a^n}} + \frac{b}{\sqrt[n-2]{b^n}} + \frac{c}{\sqrt[n-2]{c^n}}\right) &= 1, \end{aligned}$$

which implies that

$$x = \frac{1}{\sqrt[n-2]{a}}\left(\frac{a}{\sqrt[n-2]{a^n}} + \frac{b}{\sqrt[n-2]{b^n}} + \frac{c}{\sqrt[n-2]{c^n}}\right)^{-1/n}. \tag{3.19}$$

Note that dividing the both sides by

$$\frac{a}{\sqrt[n-2]{a^n}} + \frac{b}{\sqrt[n-2]{b^n}} + \frac{c}{\sqrt[n-2]{c^n}}$$

is valid because each of its summand is positive, hence this constant is not zero. Therefore the umbilical point in this case is given by

$$(x, y, z) = \left(\frac{a}{\sqrt[n-2]{a^n}} + \frac{b}{\sqrt[n-2]{b^n}} + \frac{c}{\sqrt[n-2]{c^n}}\right)^{-1/n}\left(\frac{1}{\sqrt[n-2]{a}}, \frac{1}{\sqrt[n-2]{b}}, \frac{1}{\sqrt[n-2]{c}}\right).$$

Lastly, if n is even, then real umbilical points such that $ax^{n-2} = by^{n-2} = cz^{n-2}$ only occur if $a, b, c > 0$. Solving for x, as in (3.19), shows that there are two possible values of x. Similarly, there are two possible values of y and of z, independent of the signs of each other variables. In total, we get eight umbilical points:

$$(x, y, z) = \left(\frac{a}{\sqrt[n-2]{a^n}} + \frac{b}{\sqrt[n-2]{b^n}} + \frac{c}{\sqrt[n-2]{c^n}}\right)^{-1/n}\left(\frac{\pm 1}{\sqrt[n-2]{a}}, \frac{\pm 1}{\sqrt[n-2]{b}}, \frac{\pm 1}{\sqrt[n-2]{c}}\right).$$

The other pairwise common factors also give the same set of umbilical points on the generalised Fermat surface. We have just proved the following results.

Theorem 3.5. *Let $n \geq 3$ be an odd number. The number of real umbilical points on the generalised Fermat surface $ax^n + by^n + cz^n = 1$ with $abc \neq 0$ is four. They are given by*

$$\left(\frac{1}{\sqrt[n]{a}}, 0, 0\right), \left(0, \frac{1}{\sqrt[n]{b}}, 0\right), \left(0, 0, \frac{1}{\sqrt[n]{c}}\right), \left(\frac{a}{\sqrt[n-2]{a^n}} + \frac{b}{\sqrt[n-2]{b^n}} + \frac{c}{\sqrt[n-2]{c^n}}\right)^{-1/n} \left(\frac{1}{\sqrt[n-2]{a}}, \frac{1}{\sqrt[n-2]{b}}, \frac{1}{\sqrt[n-2]{c}}\right).$$

Theorem 3.6. *Let $n \geq 4$ be an even number. The number of real umbilical points on the generalised Fermat surface $ax^n + by^n + cz^n = 1$ with $abc \neq 0$ is either 2, 4 or 14, depending on the following conditions:*

- *It has 2 umbilical points if only one of a, b, c is positive.*
- *It has 4 umbilical points if only two of a, b, c are positive.*
- *It has 14 umbilical points if all of a, b, c are positive.*

3.3 Counting Complex Umbilical Points

From the definition of complex umbilical points, we seek polynomial equations satisfied by real umbilical points. From (3.16), we obtain the polynomial equations

$$\begin{aligned} b^2 y^{2(n-1)}(cz^{n-2} - ax^{n-2}) &= 0 \\ a^2 x^{2(n-1)}(by^{n-2} - cz^{n-2}) - c^2 z^{2(n-1)}(ax^{n-2} - by^{n-2}) &= 0 \end{aligned} \tag{3.20}$$

which must also be satisfied by the complex umbilical points. Other pairs of factors of (3.15) give a similar polynomial equation.

The equation (3.20) and its counterparts for other factors can be solved in a same fashion as in the above section. In Case I, we notice that umbilical points are n-th roots of a real number. Recall that taking the n-th roots gives n solutions over the complex numbers, and so we get n count of points. For example, the real umbilical point $(\sqrt[n]{1/a}, 0, 0)$ corresponds to n complex umbilical points, for each n-th root of $1/a$.

In Case II, we proved that complex umbilical points must satisfy

$$\begin{cases} ax^{n-2} = by^{n-2} = cz^{n-2}, \\ ax^n + by^n + cz^n - 1 = 0. \end{cases} \tag{3.21}$$

We solved for x by taking the n-th root of a complex number as in (3.19), but with an assumption that

$$\frac{a}{\sqrt[n-2]{a^n}} + \frac{b}{\sqrt[n-2]{b^n}} + \frac{c}{\sqrt[n-2]{c^n}} \neq 0,$$

otherwise there are no complex umbilical points. Under this assumption, there are n possible values for x. The values of y and z can be found by taking the $(n-2)$-th root of ax^{n-2}. Therefore there are at most $n(n-2)^2$ complex umbilical points in this case.

This can be summarised into the following result.

Theorem 3.7. *Let $n \geq 3$. The number of complex umbilical points on the generalised Fermat surface $ax^n + by^n + cz^n = 1$ with $abc \neq 0$ is at most $3n + n(n-2)^2 = n^3 - 4n^2 + 7n$.*

4 Discussion

Our proof of Theorem 3.6 involves equations with the terms x^{n-2}, y^{n-2}, z^{n-2}, which are identically equal to 1 if $n = 2$, but otherwise could be zero if $n \geq 3$. As a consequence, the argument is not readily applicable to the case $n = 2$, leading to the discrepancy between Theorem 3.6 and the fact that there are only four real umbilical points on an ellipsoid.

As previously mentioned in the introduction, Salmon proved that there are $10n^3 - 28n^2 + 22n$ complex umbilical points on a general surface of degree n in $\mathbb{R}^3$. For generalised Fermat surfaces, there are at most $n^3 - 4n^2 + 7n$ complex umbilical points. The difference is expected, since a general surface of degree n has non-diagonal terms which are missing from the equation of generalised Fermat surfaces.

Acknowledgment. The authors are grateful to the referees for their careful reading of the manuscript and their useful comments. The work presented here is a continuation of the author's previous work, resulting in the first author's bachelor's thesis. We would like to wholeheartedly thank Assistant Professor Dr. Rapeepan Kraikaew, co-advisor of the thesis, for her advice and support during its preparation at Khon Kaen University. The first author would also like to extend his gratitude for the support provided by Development and Promotion of Science and Technology Talents Project (DPST) scholarship during his studies.

References

[1] W. Kühnel, Differential Geometry: Curves–Surfaces–Manifolds, third ed., American Mathematical Society, 2015.

[2] G. Salmon, A Treatise on the Analytic Geometry of Three Dimensions, Hodges, Smith, and Company, 1865.

[3] P. Breiding, K. Ranestad, M. Weinstein, Critical Curvature of Algebraic Surfaces in Three-Space, arXiv:2206.09130 [math.AG].

[4] M. Malinen, Local Reparametrization by Approximating Lines of Curvature, Math. Comput. Simul. 201 (2022) 275-290.

[5] G. Gangopadhyay: Geometry of Umbilics on Surfaces in the Three-space, Master Thesis, Higher School of Economics, 2018.

[6] E. Horobeţ: The critical curvature degree of an algebraic variety, J. Symbolic Computation 121 (2024).

[7] G. Catino, A. Roncoroni, L. Vezzoni, On the Umbilic Set of Immersed Surfaces in Three-dimensional Space Forms, Bull. Sci. Math. 165 (2020) 102917.

[8] L. Tu, Differential Geometry: Connections, Curvature, and Characteristic Classes. Graduate Texts in Mathematics. Springer, 2017.

[9] H. Whitney, Elementary Structure of Real Algebraic Varieties, Ann. Math. 66 (3)(1957).

Conference Proceedings of the 29th Annual Meeting in Mathematics (AMM 2025)
Department of Mathematics, Faculty of Science, Srinakharinwirot University
21–23 May 2025

Geometric Data Analysis for Classifying Patterns of Points Derived from Partial Circles of Gaze Positions

Asama Jampeepan[1,*] and Supanut Chaidee[2]

[1]Program in Applied Mathematics, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

[2]Advanced Research Center for Computational Simulation (ARCCoS), Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

Abstract

In this study, we focus on the geometric data analysis of the point patterns derived from a partial circle based on the modeling of gaze positions. The problem is addressed to analyze and classify the shape of partial circles using roundness measurement and the shape of a convex polygon formed by data points derived from the partial circle. Geometric models are introduced to analyze roundness as a function, and the necessary conditions of roundness with respect to the shape of a partial circle are provided. In addition, the analysis of shape is done using ideally generated data using the convex hull framework. From the study, we can observe the tendency to classify the shape of data based on the roundness measurement.

Keywords: Gaze positions, Convex hull, Roundness measurement

2020 MSC: 68U05, 52A38, 52C35

1 Introduction

Finding the exact location of objects in an image is a fundamental challenge that researchers have studied for a long time. Another closely related problem is identifying which objects are present in an image. For example, applications such as facial recognition [1] and medical imaging [2] rely on the detection and classification of objects accurately.

One key method for identifying objects in images is point pattern recognition, which focuses on analyzing the arrangement of points in an image. Classifying these point patterns helps us to understand the shapes and structures of objects. Currently, there are many algorithms for analyzing point patterns (see [3] and [4]). An important problem in object identification is iris detection, where the goal is to determine the position and size of the iris in an eye image. One effective method is the circle fitting method, which finds the best-fit circle that matches a set of points. This method helps determine the location and radius the iris more accurately. One of the most commonly used techniques is the least squares method [5].

Although recognition methods for dealing with the data are widely presented in this era, attaching those problems to the known pattern would be more effective. In the case of iris

*Speaker and corresponding author (asama_ja@cmu.ac.th)

detection, we can assume that the set of points on the iris boundary forms part of a circle, analyzing the geometric properties of these point patterns makes the detection process more robust and efficient. One of effective tools for describing the pattern of the data is the use of suitable index to measure convex shape, such as the roundness index. This index serves as a numerical value that indicates roundness, helping to specify the pattern of data derived from the partial circle. Once the pattern of data sets is determined, an appropriate algorithm can be employed to fit a circle accurately.

The main purpose of this study is to analyze and classify the shape of an incomplete circle using roundness measurement. The modeling assumptions are based on the gaze positions as compiled in [6]. We construct simple geometrical models of the roundness of the data shape using the convexity of the data shape, including the convex hull. Then, we mathematically analyze the model to understand the roundness as a function of shape parameters to clarify the necessary condition of the roundness function. Finally, we classify the shape of a convex polygon formed by data points derived from a partial circle by observing the pattern of data of roundness value.

The paper is structured as follows: Section 2 provides the basic idea of geometrical shapes and states the modeling assumptions. We define the roundness functions based on assumptions and analyze them mathematically in Section 3. The experiments from ideally generated data to reflect the shape of data with classification will be shown in Section 4. Finally, we conclude the paper with a brief summary and discuss future work.

2 Preliminaries

In this study, we mainly focus on the analysis of geometrical shapes, including fundamental aspects of convex polygons and computations involving polygons, convex sets, and convex hulls. Additionally, we explore the geometric properties of roundness for measuring the shape of convex polygons, along with related problem formulation.

2.1 Convex polygon

In discrete and computational geometry, the fundamental object is a polygon, which represents a region on a plane enclosed by a finite number of straight line segments forming a simple closed curve. A subset S of the plane is considered convex if, for any two points x and y within S, the line segment $\overline{xy}$ connecting them lies entirely within the set S. More formally, S is convex if for every pair of points $x, y \in S$ and any $\alpha \in [0, 1]$, the combination $\alpha x + (1 - \alpha)y$ also lies inside S.

The convex polygon can be considered as a polygon whose all interior angles do not exceed 180°. This means that the vertices of the polygon point outward and no side is directed inward.

For a given set S, the convex hull of S, denoted by $\mathcal{CH}(S)$, is the smallest convex set that encompasses S. It can be understood as the intersection of all convex sets that cover S. In other words, for every convex set C_α that contains S, the convex hull is defined as

$$\mathcal{CH}(S) = \bigcap_{\alpha \in \Lambda} C_\alpha.$$

For a finite set of points $S = \{p_1, p_2, \ldots, p_n\}$ in the two-dimensional Euclidean space, the convex hull of S is the smallest convex polygon that contains all the points in S. In the special case where all the points in S lie on the boundary of the convex hull, the convex hull is a convex polygon whose vertices are exactly the points in S [7]. There are several algorithms for computing

the convex hull of S, such as Graham's Scan and the Divide-and-Conquer method [8]. The lower bound for the complexity of computing the convex hull has been proven to be $O(n \log n)$, where n is the number of points in S.

The centroid of a convex polygon is the point (x_c, y_c) where its medians intersect. To obtain its coordinates, we sum the x-coordinates and divide by the number of vertices for the x-coordinate, and perform the same process for y-coordinates to determine the y-coordinate. The formulas are:

$$(x_c, y_c) = \left(\frac{x_1 + x_2 + \cdots + x_n}{n}, \frac{y_1 + y_2 + \cdots + y_n}{n} \right).$$

When a polygon is given such that the vertices $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ of a simple polygon are listed counterclockwise, the area of the polygon can be computed by the Surveyor's formula [9]:

$$A = \frac{1}{2} \left(\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \cdots + \begin{vmatrix} x_{n-1} & x_n \\ y_{n-1} & y_n \end{vmatrix} + \begin{vmatrix} x_n & x_1 \\ y_n & y_1 \end{vmatrix} \right), \tag{2.1}$$

where $|\cdot|$ denotes the determinant operation.

2.2 Roundness Measurement

Measurements of roundness are used to assess how closely a feature or object resembles a perfect circle. This geometric property is used to measure how far a shape deviates from a real circle.

Definition 2.1. [10] (Diameter of P) Let P be a convex hull of finite points. The diameter of P, denoted by $diam(P)$, is the largest distance between any pair of $x, y \in P$. In other words, it is the length of the longest diagonal of P (e.g., straight line segment joining two vertices), that is,

$$diam(P) = \max_{x,y \in P} d(x, y). \tag{2.2}$$

We define the roundness of an object by the formula based on [11] as follows:

Definition 2.2. (Roundness) Let $\mathcal{O}$ be a bounded set without hole. The roundness of $\mathcal{O}$ is defined as the ratio of the area of $\mathcal{O}$ to the area of a circle with the same perimeter. It is calculated using the formula:

$$RF = \frac{4A}{\pi F_{\max}^2}, \tag{2.3}$$

where $F_{\max}$ is the longest distance between two points along the boundary of $\mathcal{O}$, and A is the area of $\mathcal{O}$. For the ideal shape, which is a circle with radius r, $F_{\max}$ is $2r$ and the roundness is equal to 1.

In the case that the considered object is a convex polygon, the largest distance in the polygon can be regarded as $diam(P)$. i.e., Equation (2.3) can be written as

$$RF = \frac{4Area}{\pi \left(diam(P)\right)^2}. \tag{2.4}$$

The roundness of a convex polygon is a measure of how closely its shape resembles that of a circle. A higher roundness value, approaching 1, indicates a shape more similar to a circle. Therefore, RF is bounded by $0 \leq RF \leq 1$.

Example 2.3. Given a set of points $S = \{(1, 2.8), (0, 2), (0, 1), (0.4, 0.4), (2, 0), (3, 1), (2, 2)\}$ with its convex hull. We first compute the diameter of P using 2.2 to obtain $diam(P) = \sqrt{10}$. By Equation 2.4, the roundness value is given by $RF = \frac{4 \cdot 5.2}{\pi(\sqrt{10})^2} \approx 0.662$.

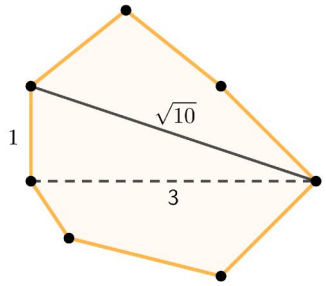


Figure 1: An example of convex hull of a set of points S.

2.3 Modeling Assumptions

This study aims to analyze and categorize data patterns obtained from partial circles, focusing on their roundness. The problem originally focuses on the analysis of gaze position. Therefore, to simplify a gaze position model, we begin by assuming that the model consists of an upper eyelid, a lower eyelid, and an iris boundary, as shown in Figure 2.

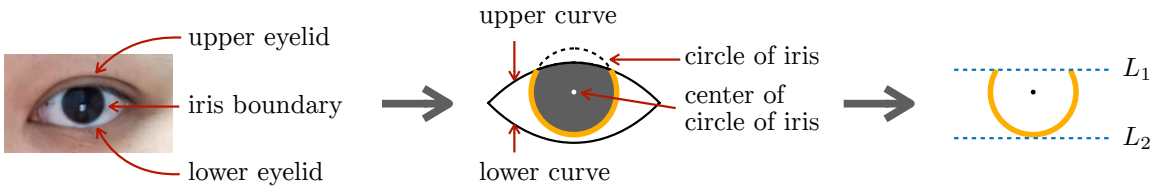


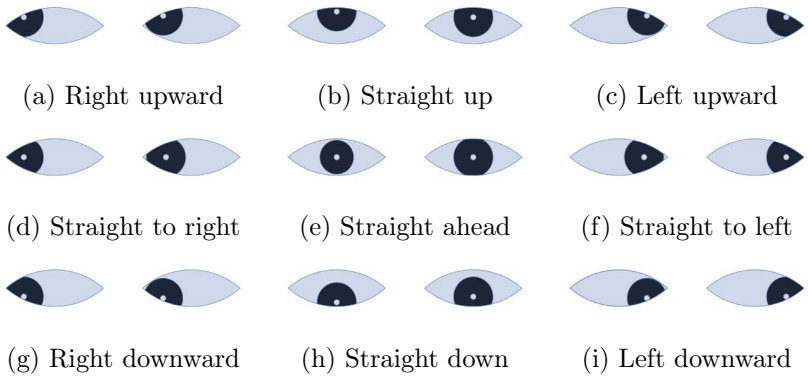
Figure 2: Modeling procedure in this study; (left) an eye photo (center) a geometric model of the eye photo (right) the modeling of the circle of iris based on circle and lines L_1 and L_2. The illuminated boundary shows the considered partial circular arc.

Specifically, we are focusing on classifying ideally generated data along the iris boundary of open eyes. To achieve this, we assume that there are 9 positions of gaze based on [6], as illustrated in Figure 3:

We assume that the point cloud of iris boundary is given. Then the boundary approximated as a circle is assumed to be cut into a maximum of two chords. For each chord, which is a part of the line L_i, we assume that the chord cuts a circle so that the angle at the center measured to the intersections between chord L_i and the circle is θ_i, as shown in Figure 4.

For a circle cut by two chords, if a chord does not intersect with another chord or the intersecting point is on the circle, we can consider the following subcases, as shown in Figure 4a, with angles θ_1 and θ_2. In this case, when it is cut by one chord, it can be separated into two cases, as shown in Figure 4b. In this case, we suppose that the considered angle θ_1 is θ, and another angle is $\theta_2 = 0$.

We first assume that the data points are ideally on the part of the circle boundary, i.e., they are the circular arcs of the circle. In this case, the area is considered as a segment of a circle.

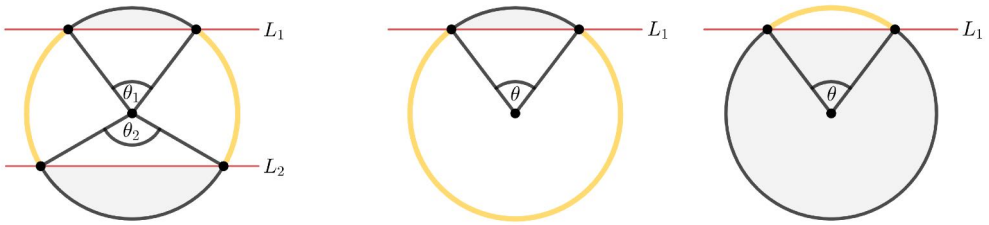


(a) Right upward (b) Straight up (c) Left upward

(d) Straight to right (e) Straight ahead (f) Straight to left

(g) Right downward (h) Straight down (i) Left downward

Figure 3: Gaze positions are analyzed based on the study of [6].



(a) (b)

Figure 4: (a) A circle is cut by two chords in different positions. (b) A circle is cut by one chord; (left) a minor segment of a circle and (right) a major segment of a circle. The highlight part is set as the possible locations of a point cloud to compute the roundness.

Without loss of generality, if the circle is not cut by the chord, the arc is the full circumference. The partial arcs of Figure 4(a) are divided into two cases: there is only one partial arc if $\theta_1 = 0$ or $\theta_2 = 0$, and there are two partial arcs if $\theta_1, \theta_2 \neq 0$. In Figure 4(b), if the circle is cut by a minor segment, the arc will be more than half of the circular arc, whereas if the circle is cut by a major segment, the arc will be less than half of the circular arc.

Now, we proceed with the assumption that each element can be represented mathematically using the equations presented in Table 1 based on gaze positions shown in Figure 3.

The symbol $\otimes$ in Table 1 indicates that there is no such information on the intersecting line, that is, the line is tangent to the circle with angle $\theta = 0$. For each position of Figure 3, the left and right positions indicate the subcases shown by the left and right figures, respectively.

In this study, our objective is to analyze the roundness represented by Definition 2.2 analytically. After considering these behaviors, we observe the shape of the data illustrated by its convex hull, as presented in Section 4. We expect to classify the roundness using the ideally generated data.

3 Geometrical Properties of Roundness

Based on the assumption shown in Table 1 and the chord cuts a circle as described in Section 2.3, we construct a roundness function using Definition 2.2. We assume that the data are all

Table 1: The mathematical equations correspond to the positions of gaze

Gaze position in Figure 3		Angles (θ_1, θ_2)	Lines (L_1, L_2)
(a)	left	$(117.08, 0)$	$(-1.42x + 0.94y = 1.02, \otimes)$
	right	$(159.64, 0)$	$(-0.7x + 1.84y = 1.54, \otimes)$
(b)	left	$(116.54, 0)$	$(y = 0.7, \otimes)$
	right	$(159.64, 0)$	$(y = 0.65, \otimes)$
(c)	left	$(159.64, 0)$	$(0.7x + 1.84y = 1.54, \otimes)$
	right	$(117.08, 0)$	$(1.42x + 0.94y = 1.02, \otimes)$
(d)	left	$(69.64, 0)$	$(x = -1.18, \otimes)$
	right	$(126.87, 126.87)$	$(y = 0.5x + 1, y = -0.5x - 1)$
(e)	left	$(0, 0)$	$(\otimes, \otimes)$
	right	$(82.82, 82.82)$	$(y = 0.75, y = -0.75)$
(f)	left	$(126.87, 126.87)$	$(y = -0.5x + 1, y = 0.5x - 1)$
	right	$(69.64, 0)$	$(x = 1.18, \otimes)$
(g)	left	$(159.64, 0)$	$(0.7x + 1.84y = -1.54, \otimes)$
	right	$(117.08, 0)$	$(1.42x + 0.94y = -1.02, \otimes)$
(h)	left	$(116.54, 0)$	$(y = -0.7, \otimes)$
	right	$(159.64, 0)$	$(y = -0.65, \otimes)$
(i)	left	$(159.64, 0)$	$(-0.7x + 1.84y = -1.54, \otimes)$
	right	$(117.08, 0)$	$(-1.42x + 0.94y = -1.02, \otimes)$

points on circular arcs of the unit circle satisfying one of the following conditions: two nearly opposite arcs or one arc taking the major arc of the unit circle. Furthermore, we assume that the circle can be cut by at most two chords, as shown in Figure 4. Since the area of a segment is calculated by

$$A_1 = \frac{1}{2}\left(\theta_1 - \sin\theta_1\right) r^2, \tag{3.1}$$

where θ_1 is the angle of the sector with respect to the considered segment in radians, and given that the circle can be cut by at most two non-intersecting chords, the area of the segments should be

$$A_1 + A_2 = \frac{1}{2}\left(\left(\theta_1 - \sin\theta_1\right) + \left(\theta_2 - \sin\theta_2\right)\right) r^2. \tag{3.2}$$

In this model, we assume that the circle has a radius of 1. As a result, the diameter of the circle is 2, which is also the diameter of the full circle. We can compute the roundness using the fact that the roundness of a full circle is given by $RF_1(\theta_1, \theta_2) = 1$, where $\theta_1, \theta_2 = 0$. If the circle is cut by segments A_i, for $i = 1, 2$, the roundness can be computed as

$$RF_1\left(\theta_1, \theta_2\right) = 1 - \frac{1}{2\pi}\left[\left(\theta_1 - \sin\theta_1\right) + \left(\theta_2 - \sin\theta_2\right)\right]. \tag{3.3}$$

where $\theta_1, \theta_2 \in [0, \pi]$. It is clear that RF_1 with respect to θ_1 and θ_2 is well-defined. From Equation (3.3), we obtain the following lemma:

Lemma 3.1. *Let* $RF_1(\theta_1, \theta_2) = 1 - \frac{1}{2\pi}((\theta_1 - \sin\theta_1) + (\theta_2 - \sin\theta_2))$ *be a function defined on* $[0, \pi] \times [0, \pi]$. *Suppose* R *is a roundness value within the interval* $[0, 1]$. *Then, there exists at least one pair of points* $(\theta_1^*, \theta_2^*) \in [0, \pi] \times [0, \pi]$ *such that* $RF_1(\theta_1^*, \theta_2^*) = R$.

Proof. Obviously, the function $RF_1(\theta_1, \theta_2)$ is differentiable everywhere on $[0, \pi] \times [0, \pi]$. We evaluate that $RF_1(\theta_1, \theta_2)$ takes on all values in the interval $[0, 1]$ by firstly fixed θ_2 to be an arbitrary number in $[0, \pi]$. We remark that $RF_1^*(\theta_1) := RF_1(\theta_1, \theta_2)$ is decreasing function; i.e. $\frac{\partial RF_1^*}{\partial \theta_1} < 0$ except the end of the interval. Hence,

$$RF_1(\pi, \theta_2) \leq RF_1(\theta_1, \theta_2) \leq RF_1(0, \theta_2).$$

Similarly, we interchange θ_2 to θ_1; i.e. for any arbitrary $\theta_1 \in [0, \pi]$, we have

$$RF_1(\theta_1, \pi) \leq RF_1(\theta_1, \theta_2) \leq RF_1(\theta_1, 0).$$

By the intermediate value theorem, there exists (θ_1^*, θ_2^*) such that $RF_1(\theta_1^*, \theta_2^*) = R$. □

Lemma 3.1 implies that for any roundness R, we can find the pair of (θ_1, θ_2) satisfying the shape by analyzing a graph of $RF_1(\theta_1, \theta_2)$ with the space of angle (θ_1, θ_2). Remark that the roundness R is represented as a plane parallel to the plane-$\theta_1\theta_2$ called the plane of roundness. The trace from the cut of $RF_1(\theta_1, \theta_2)$ and R is shown as a plane curve on the plane R. The plot of $RF_1(\theta_1, \theta_2)$ with the plane R is shown in Figure 6.

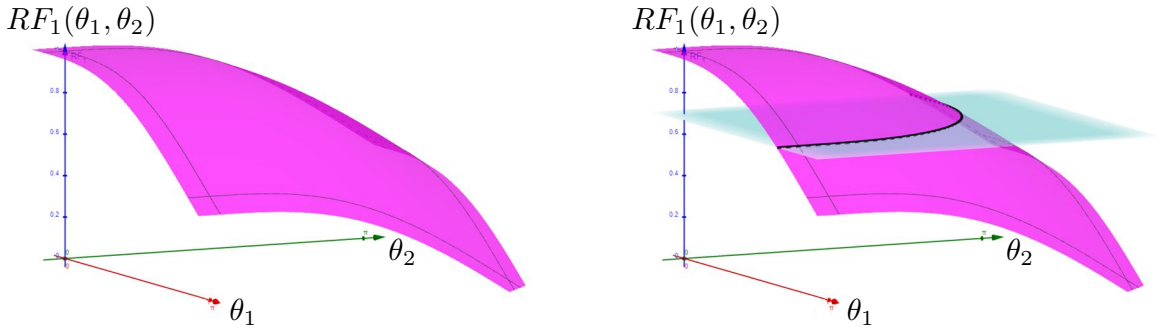


Figure 6: (left) the relation between θ_1 and θ_2 satisfying the roundness value, (right) the plane curve of the intersection between the plane of roundness and the function $RF_1(\theta_1, \theta_2)$.

For example, as shown in Figure 6 (right), it is evident that if we choose $RF_1(\theta_1, \theta_2) = 0.5$, there exists a set of values for the parameters (θ_1, θ_2) along the curve where the condition $RF_1(\theta_1, \theta_2) = 0.5$ is satisfied. The curve exhibits points or regions where the $RF_1(\theta_1, \theta_2)$ value corresponds precisely to 0.5, showcasing the relationship and dependencies between the variables $RF_1(\theta_1, \theta_2)$, θ_1, and θ_2 along the black curve.

The remaining case occurs when the points are on the minor circular arc, as shown in Figure 7.

For a segment of a circle in Figure 7, the roundness function can be expressed as the ratio of the area of the segment to the area of the circle from which the segment is derived. In this situation, the diameter of the considered segment depends on the length of the chord. That is,

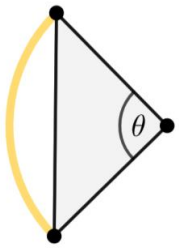


Figure 7: The area of a segment of circle.

the chord of length $2\sin\left(\frac{\theta}{2}\right)$ will be a diameter when the length is greater than the radius of the unit circle. Since the area of a minor segment of a circle is given by

$$A_{\text{minor}} = \frac{1}{2}r^2(\theta - \sin\theta), \tag{3.4}$$

where θ is the central angle in radians, we define the roundness of the minor segment using Equation 2.3 as

$$RF_2(\theta) = \frac{4A_{\text{minor}}}{\pi M^2}, \tag{3.5}$$

where $M = \max\left\{1, 2\sin\left(\frac{\theta}{2}\right)\right\}$ and $\theta \in (0, \pi)$. Assuming the radius $r = 1$, the roundness function simplifies to

$$RF_2(\theta) = \frac{2\left(\theta - \sin\theta\right)}{\pi M^2}. \tag{3.6}$$

For the expression $M = \max\{1, 2\sin\left(\frac{\theta}{2}\right)\}$ where $\theta \in (0, \pi)$, we remark that the smallest angle θ which satisfies $2\sin\left(\frac{\theta}{2}\right) > 1$ is $\theta = \frac{\pi}{3}$. Therefore,

$$RF_2(\theta) = \begin{cases} \frac{2(\theta-\sin\theta)}{\pi} & \text{if } \theta \in \left(0, \frac{\pi}{3}\right), \\ \frac{2(\theta-\sin\theta)}{\pi\left(2\sin\left(\frac{\theta}{2}\right)\right)^2} & \text{if } \theta \in \left[\frac{\pi}{3}, \pi\right). \end{cases} \tag{3.7}$$

If $\theta \in \left(0, \frac{\pi}{3}\right)$, let $M = 1$. Then the roundness function $RF_2(\theta) = \frac{2(\theta-\sin\theta)}{\pi}$ is bounded by $0 < RF_2(\theta) < \frac{2}{\pi}\left(\frac{\pi}{3} - \frac{\sqrt{3}}{2}\right)$. For $\theta \in \left[\frac{\pi}{3}, \pi\right)$, let $M = 2\sin\left(\frac{\theta}{2}\right)$, and define the function $RF_2(\theta) = \frac{2(\theta-\sin\theta)}{\pi\left(2\sin\left(\frac{\theta}{2}\right)\right)^2}$. In this case, the function is bounded by $\frac{2}{\pi}\left(\frac{\pi}{3} - \frac{\sqrt{3}}{2}\right) \leq RF_2(\theta) < 0.5$. The graph shown in Figure 8.

We can conclude the results with the following corollary:

Corollary 3.2. *For a circular arc of the unit circle whose central angle is smaller than $\frac{\pi}{3}$, the roundness $RF_2(\theta)$ of a segment cut by a chord is bounded by $0 < RF_2(\theta) < \frac{2}{\pi}\left(\frac{\pi}{3} - \frac{\sqrt{3}}{2}\right) \approx 0.1153$. In the case that the central angle is in the interval $\theta \in \left[\frac{\pi}{3}, \pi\right)$, the roundness is bounded by $0.1153 \approx \frac{2}{\pi}\left(\frac{\pi}{3} - \frac{\sqrt{3}}{2}\right) \leq RF_2(\theta) \leq 0.5$.*

Corollary 3.2 states the necessary conditions for a shape which have a roundness smaller than 0.5. By considering $RF_1(\theta_1, \theta_2)$ such that $\theta_2 = 0$ and $\theta_1 \in [0, \pi]$, the roundness of the shape is larger than 0.5 with Corollary 3.2.

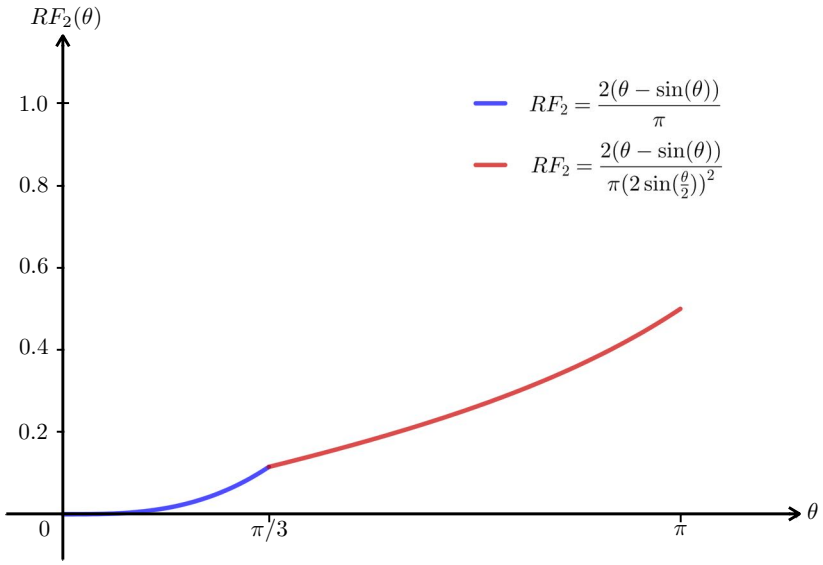


Figure 8: The plot of the roundness function $RF_2(\theta) = \frac{2(\theta - \sin\theta)}{\pi}$, where $\theta \in (0, \pi)$.

4 Experimental Results

In this part, we will consider the shape of a convex hull when the points are close to a circle. We conducted experiments to observe the roundness behavior based on the orientation of points based on Figure 3.

4.1 Data

If the set of points S on the plane is given, in the experiment, we first construct the convex hull $\mathcal{CH}(S)$ of the points. After that, the area of the convex hull is computed using the formula (2.1), and the diagonal of the convex hull $\mathcal{CH}(S)$ is computed using (2.2). Therefore, the roundness of the convex hull $\mathcal{CH}(S)$ is obtained. We suppose that the numerical issue in the computing area is solved using robust geometric computation.

From the application viewpoint, a photo of the iris boundary is extracted as pixels instead of points on the plane. Therefore, we compute the reasonable number of points needed to simulate the framework by assuming the points are on the unit circle. First, we calculate the number of ideally generated data points n. This involves determining the quantity of data by counting the pixels the circumference passes through. To facilitate this calculation, we consider a circle with a radius of 1 cm and a diameter of 2 cm. Employing UnitConverter.net [12] for the conversion of centimeters to pixels, we found that 2 centimeters is approximately equivalent to 76×76 pixels.

To ascertain the number of pixels constituting the circumference, we utilize the formula $n = 4N - 4$, equivalent to the count of pixels in Minecraft [13], where N represents the size of pixels. In the case $N = 76$, the number of random points is $n = (4 \cdot 76) - 4 = 300$. It is important to note that we introduce a consideration for potential error by assuming that the detection of the circumference has an allowance of ± 1 pixel, approximately ± 0.04 in the measurement on the 2D plane. Together with those points on the plane equivalent to the pixels, we assume that the points on the plane are uniformly distributed in the annulus of the unit circle; that is, we generate n points in the region bounded by the circle: $c_1 : x^2 + y^2 = (1 + \epsilon)^2$

and $c_2 : x^2 + y^2 = (1 - \epsilon)^2$.

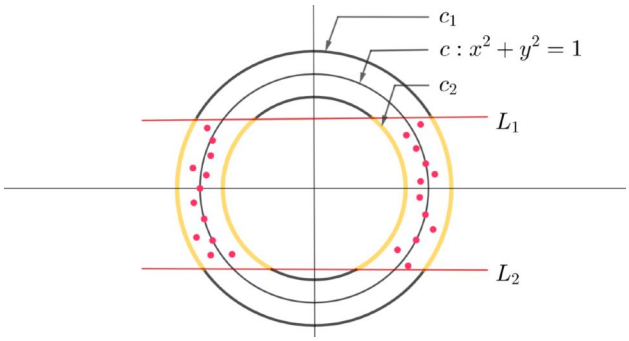


Figure 9: The red data points represent the data of interest, located in the region bounded by circles c_1 and c_2 and lines L_1 and L_2.

In our study, we chose $\epsilon = 0.04$ to random 300 points in the annulus, where the outer circle is defined by $c_1 : x^2 + y^2 = 1.04^2$, and the inner is defined by $c_2 : x^2 + y^2 = 0.96^2$. This generates the points corresponding to the case (e) of Figure 3.

For other cases appearing in Figure 3, we generate a similar situation to the previous case; however, the region is limited by the equation of lines shown in Table 2, and the number of randomized points is also limited, as shown in the following table. Therefore, there are 18 cases to generate the experiments.

Table 2: Approximated number of points on the annulus

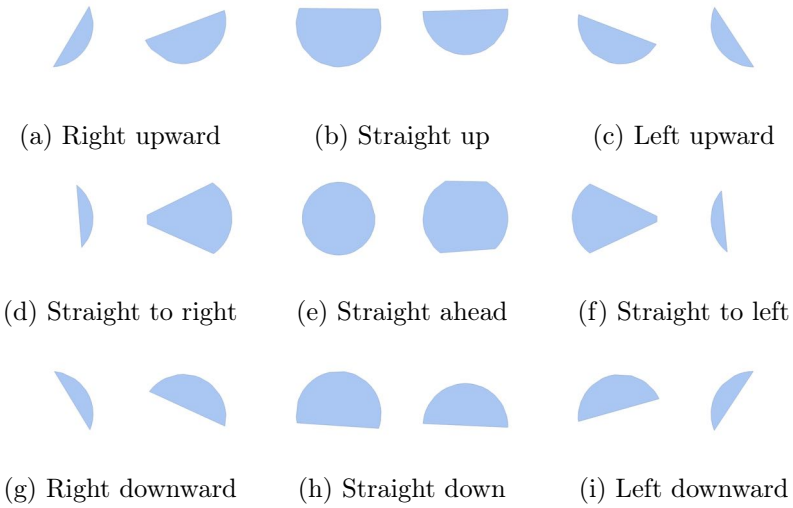
Case	No. of points	
	left	right
(a)	98	168
(b)	203	134
(c)	168	98
(d)	59	89
(e)	300	163
(f)	89	59
(g)	168	98
(h)	203	134
(i)	168	98

4.2 Experiments

In the experiments, we consider each case by randomizing 1,000 sets of points. We compute the convex hull, its area, diameter, and roundness. Hence, 18,000 values of roundness are recorded and plotted as a histogram. We used Wolfram Mathematica 14.0 and employed the ConvexHullMesh command [14] to compute the convex hull of the provided data from the previous subsection. Examples of the convex hull of each case are shown in Figure 10 for 9 positions of gaze.

4.3 Results

From the experiments, the roundness of each data set is computed as a roundness of a convex polygon. 18,000 values of roundness are plotted in the histogram as shown in Figure 11. We



(a) Right upward (b) Straight up (c) Left upward

(d) Straight to right (e) Straight ahead (f) Straight to left

(g) Right downward (h) Straight down (i) Left downward

Figure 10: The convex hull of a set of points based on positions of gaze.

found that the roundness value for each case is bounded by $\min(RF) \leq RF \leq \max(RF)$, as illustrated in Table 3.

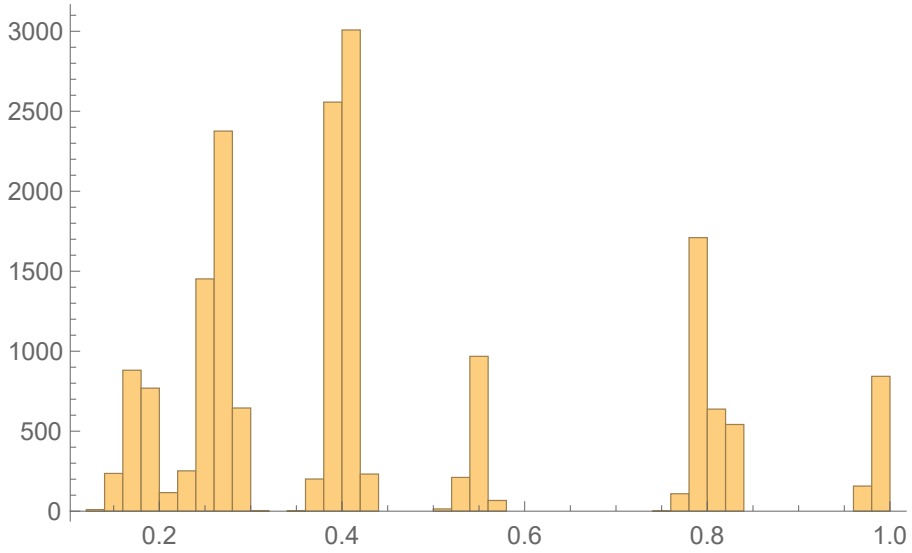


Figure 11: The histogram of roundness values of all 9 positions of gaze

From Table 3, it indicates that when data is distributed in a circular pattern, the resulting convex shape tends to be nearly circular, approaching a roundness value of 1. Specifically, data points generated on a circular arc cut by a minor segment exhibit a roundness value greater than 0.5. Conversely, data generated on a circular arc cut by major segments tends to yield a roundness value of less than 0.5. These observations align with the geometric properties of roundness discussed in Section 3.

For $\theta \in [0, \pi]$, the roundness of the convex hull, denoted as $RF_2(\theta)$, varies within the interval $[0.127985, 0.438148]$, which is consistent with Corollary 3.2, where the roundness is bounded by

$0.1153 \approx \frac{2}{\pi}\left(\frac{\pi}{3} - \frac{\sqrt{3}}{2}\right) \leq RF_2(\theta) \leq 0.5$.

When considering angle pairs $(\theta_1, \theta_2) \in (0, \pi) \times (0, \pi)$, based on Lemma 3.1, we observed that in the case where the circle is not cut at all, that is, when $\theta_1 = \theta_2 = 0$, as shown in case (e) left, the roundness of the convex hull is close to 1. When $\theta_1 \in [0, \pi]$ and $\theta_2 = 0$, the roundness is in the range 0.759905 to 0.809748, as shown in case (b) left and case (h) left. For angle pairs $(\theta_1, \theta_2) \in (0, \pi) \times (0, \pi)$, the roundness of the convex hull can take any value in the range from 0 to 1, depending on θ_1 and θ_2. However, it is interesting to investigate the roundness of the data with respect to the cases (d) right, (e) right, and (f) left. The experiments suggest refining the analysis of function RF, satisfying the shape of the cases (d) right, (e) right, and (f) left to obtain the smaller bounds for better classification.

Finally, it can be conclude that measuring roundness using the convex hull can result in slight differences compared to roundness calculated from a full circle or from circular segments defined by chords. These differences occur because the convex hull only approximates the boundary of the data and may not fully capture the original circular shape, particularly when the data is unevenly distributed or taken from partial arcs.

Moreover, from Figure 11, we can see that the roundness values can be roughly grouped based on the gaps observed in the graph, with noticeable clusters in the ranges of 0.127985 to 0.301037, 0.352913 to 0.438148, 0.508521 to 0.571732, 0.759905 to 0.83354, and 0.973095 to 0.98861. Therefore, based on Lemma 3.1 and Corollary 3.2, it is possible to classify the data into different groups according to roundness: the values of the first two groups correspond to the data where the circle is cut with $\theta \in [\pi/3, \pi]$, while those in the remaining three groups correspond to data where the circle is either not cut or is cut with $\theta_1, \theta_2 \in [0, \pi]$. These observations are consistent with the results proven in Section 3.

5 Discussion

This study modeled the shape of data on the edge of the iris boundary by proposing and examining two roundness functions, RF_1 and RF_2. We provided the necessary conditions for the data shape to illustrate the roundness in some cases. Experimental results demonstrate the application of the roundness function to classify patterns derived from partial circle data obtained from various models of gaze positions. The results suggest that using values facilitates the classification of convex polygons formed by data points, with the need for further refinement in certain groups based on the number of large chords present. Overall, the research provides a mathematical framework for assessing and classifying incomplete circle shapes.

For future work, we plan to validate this framework by applying it to real data obtained from the edge of the iris boundary. This involves acquiring a dataset of actual iris images and assessing how well different algorithms perform under real-world conditions, considering factors like lighting variations and occlusions. The empirical validation aims to refine the framework, identify limitations, and provide informed recommendations for selecting the most suitable circle-fitting algorithm for practical applications in iris recognition.

Acknowledgment. The first author acknowledges the support of the DPST of IPST, Ministry of Education, Thailand. We would like to thank my friends for the discussion. Our sincere thanks go to the reviewers for their thoughtful feedback and for reviewing our work.

Table 3: The minimum and maximum roundness values for each convex shape based on Figure 10

Case	Convex shape	Roundness value	
		min (RF)	max (RF)
(e) left		0.973095	0.98861
(e) right		0.799416	0.83354
(b) right (h) right		0.364352	0.438148
(b) left (h) left		0.759905	0.809748
(d) right (f) left		0.508521	0.571732
(d) left (f) right		0.127985	0.217332
(a) right (c) left (g) right (i) left		0.352913	0.435572
(a) left (c) right (g) left (i) right		0.207294	0.301037

6 Appendix

This section shows an example (case (e), left) of how we construct and obtain data in the annular region using Wolfram Mathematica. For the remaining cases, we construct the annulus based on the gaze position. For the circle of the iris, we define the radius as $r = 1$, and the arcs of the upper and lower curves are obtained from circles with fixed centers: the upper curve is centered at $(0, -2)$ and the lower curve at $(0, 2)$, both with a radius of $r = 2\sqrt{2}$. For each case, the number of random points is based on Table 2. The minimum and maximum roundness values may vary due to the random distribution of points in the annular region for each case.

```
Annulus1 = RegionDifference[Disk[{0, 0}, 1.04], Disk[{0, 0}, 0.96]];
R1 = Region[Annulus1];
iterations1 = 1000;
resultTable1 = {};

Do[
  Points1 = RandomPoint[Annulus1, 300];

  (* Construct a convex hull and compute the roundness of a convex polygon *)
  CH1 = ConvexHullMesh[Points1];
  CH1Vertices = MeshCoordinates[CH1];
  MinBounding1 = BoundingRegion[CH1Vertices, "MinDisk"];
  Center1 = RegionCentroid[MinBounding1];
  Circle1 = Graphics[{{LightBlue, MinBounding1}, Point[CH1Vertices]}];
  Distances1 = DistanceMatrix[CH1Vertices];
  {index11, index12} = Position[Distances1, Max[Flatten[Distances1]]][[1]];
  point11 = CH1Vertices[[index11]];
  point12 = CH1Vertices[[index12]];

  (* Roundness Measure *)
  Distance1 = EuclideanDistance[point11, point12];
  Roundness1[i] = 4 * Area[CH1]/(Pi * Distance1^2);
  Print[Roundness1[i]];
  AppendTo[resultTable1, {i, Roundness1[i]}],
  {i, 1, iterations1}
]

resultTable1;
Result1 = resultTable1[[All, 2]];
H1 = Histogram[Result1, {0.02}, PlotRange -> {{0, 1}, Automatic},
     ChartStyle -> Directive[Red, Opacity[0.5]]];
MinResult1 = Min[Result1];
```

Listing 1: Mathematica code for constructing data points and computing roundness of convex hulls in an annular region for the case (e) left.

References

[1] M. A. Khan, H. F. B. M. Zaki, A. B. M. Ibrahim, S. A. Hoq, and A. A. bin Shafie, *A systematic review on facial detection and recognition: Limitations and opportunities*, Asian Journal of Electrical and Electronic Engineering, 4(2), 61–76, (2024).

[2] K. Yan, Y. Tang, Y. Peng, V. Sandfort, M. Bagheri, Z. Lu, and R. M. Summers, *MULAN: multitask universal lesion analysis network for joint lesion detection, tagging, and segmentation*, In Medical Image Computing and Computer Assisted Intervention – MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22, Springer International Publishing, 194–202, (2019).

[3] B. N. Boots and A. Getis, *Point pattern analysis*, Reprint edition, edited by Grant Ian Thrall, WVU Research Repository, (2020).

[4] J. Illian, J. Penttinen, H. Stoyan, and D. Stoyan, *Statistical analysis and modelling of spatial point patterns*, John Wiley & Sons, (2008).

[5] N. Chernov and C. Lesort, *Least squares fitting of circles*, Journal of Mathematical Imaging and Vision, 23, 239–252, (2005).

[6] H. W. Lim, D. E. Lee, J. W. Lee, M. H. Kang, and M. Seong, *Clinical measurement of the angle of ocular movements in the nine cardinal positions of gaze*, Ophthalmology, 121(4), 870–876, (2014).

[7] M. Smid, *Computing the convex hull of a planar point set*, (2003).

[8] S. L. Devadoss and J. O'Rourke, *Discrete and computational geometry*, Princeton University Press, (2011).

[9] B. Braden, *The surveyor's area formula*, The College Mathematics Journal, 17(4), 326–337, (1986).

[10] S. W. Bae, M. Korman, and Y. Okamoto, *The geodesic diameter of polygonal domains*, Discrete & Computational Geometry, 50, 306–329, (2013).

[11] D. W. Hetzner, *Comparing binary image analysis measurements–euclidean geometry, centroids and corners*, Microscopy Today, 16(4), 10–17, (2008).

[12] *Unit converter*, <`https://www.unitconverters.net/`>, (accessed June 5, 2024).

[13] J. G. Donat, *Pixel Circle / Oval Generator (Minecraft)*, Donat Studios, <`https://donatstudios.com/PixelCircleGenerator`>, 2012 (accessed June 5, 2024).

[14] Wolfram Research, *ConvexHullMesh*, Wolfram Language & System Documentation Center, <`https://reference.wolfram.com/language/ref/ConvexHullMesh.html`>, (accessed July 28, 2024).

การสร้างแบบจำลองแผนผังสำหรับตลาดนัด มศว

ปริยาภัทร นิ่มเจริญ[1,*] ณัฐณิชา ใจฝัน[1] ธัญสินี มะนะโส[1] ธันยพร นาคหรั่ง[1] และ อนัญญา เมฆเกิดชู[1,†]

[1]ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ 10110

บทคัดย่อ

ตลาดนัด มศว เป็นแหล่งค้าขายยอดนิยมที่มีผู้ใช้บริการหนาแน่น การศึกษานี้จึงมีเป้าหมายในการพัฒนาแบบจำลองแผนผังตลาดนัดที่ใช้พื้นที่ได้อย่างมีประสิทธิภาพสูงสุด โดยสามารถเพิ่มจำนวนร้านค้าได้อย่างน้อยร้อยละ 25 โดยไม่กระทบต่อความสะดวกในการสัญจร แนวคิดการออกแบบอิงหลักเรขาคณิต โดยใช้เส้นโค้งรูปพาราโบลาและวงรีในการกำหนดรูปทรงหน้าร้านใหม่ เพื่อเพิ่มพื้นที่หน้าร้านโดยไม่รบกวนเส้นทางเดิน พร้อมทั้งปรับความลึกของร้านให้เหมาะสมผ่านการวิเคราะห์เชิงคณิตศาสตร์ โดยการคำนวณอนุพันธ์อันดับที่สามของอัตราส่วนพื้นที่ที่ถูกตัดออกต่อความกว้างหน้าร้าน เพื่อประเมินความคุ้มค่าของพื้นที่

การศึกษานี้นำเสนอแผนผังใหม่สองแนวทาง ได้แก่ (1) การปรับปรุงจากผังปัจจุบัน พบว่าสามารถเพิ่มจำนวนร้านค้าได้ 66 ร้าน คิดเป็นร้อยละ 25.38 จากจำนวนร้านค้าปัจจุบัน และ (2) การออกแบบจากผลสำรวจพื้นที่จริง พบว่าสามารถเพิ่มจำนวนร้านค้าได้ 82 ร้าน คิดเป็นร้อยละ 31.54 จากจำนวนร้านค้าปัจจุบัน ผลการศึกษาแสดงให้เห็นว่าแผนผังที่พัฒนาแล้วสามารถเพิ่มจำนวนร้านค้าได้ตามเป้าหมาย

คำสำคัญ : แผนผังตลาด, พาราโบลา, วงรี, การใช้พื้นที่อย่างมีประสิทธิภาพ, อนุพันธ์อันดับที่สาม

2020 MSC: 26A06

*ผู้นำเสนอ (pariyapat.nim@g.swu.ac.th)

†ผู้ประพันธ์บรรณกิจ (ananya.make@g.swu.ac.th)

1 บทนำ

ตลาดนัดในเขตเมืองเป็นองค์ประกอบสำคัญของวิถีชีวิตผู้คน ไม่เพียงทำหน้าที่เป็นสถานที่ซื้อขายสินค้าเท่านั้น แต่ยังเป็นพื้นที่ทางสังคมที่สร้างโอกาสทางเศรษฐกิจให้กับผู้ประกอบการรายย่อย โดยเฉพาะตลาดที่ตั้งอยู่ในย่านเศรษฐกิจหรือสถาบันการศึกษา ซึ่งมักมีความหนาแน่นของผู้ใช้บริการสูง เช่นเดียวกับ [1] ตลาดนัดมหาวิทยาลัยศรีนครินทรวิโรฒ หรือ “ตลาดนัด มศว” ซึ่งตั้งอยู่บริเวณย่านอโศก ได้รับความนิยมจากประชาชนทั่วไปและชาวต่างชาติมาอย่างต่อเนื่องกว่า 20 ปี ด้วยความหลากหลายของสินค้าและทำเลที่ตั้งอยู่ใจกลางเมือง แม้ตลาดนัด มศว จะมีความนิยมอย่างต่อเนื่อง แต่จากการสำรวจพื้นที่จริงพบว่าการจัดสรรพื้นที่ภายในตลาดยังไม่มีประสิทธิภาพเท่าที่ควร บางร้านมีพื้นที่ภายในว่างมากเกินไป ส่งผลให้จำนวนร้านค้าที่สามารถให้เช่าได้ลดลง อีกทั้งพื้นที่หน้าร้านส่วนใหญ่ไม่มีพื้นที่ยืนรอสำหรับลูกค้า ทำให้ลูกค้าต้องยืนรอบนทางเดินส่วนกลาง ส่งผลให้เกิดปัญหาความแออัดขัดขวางการสัญจรของผู้อื่นในตลาด โดยเฉพาะในช่วงเวลาที่มีผู้มาใช้บริการหนาแน่น

จากการสืบค้นพบว่า มี [2] การออกแบบทางสัญจรสาธารณะที่กล่าวว่า การออกแบบทางสวนกัน หรือทางคู่ 2 คน เท่ากับ 48 นิ้ว หรือ 121.9 เซนติเมตร และ การจัดพื้นที่ในตลาดโดยใช้หลักการออกแบบสถาปัตยกรรมและผังเมือง เช่น [3] การกำหนดระยะทางเดินเท้าที่เหมาะสมบนถนนสายรองควรมีความกว้างไม่น้อยกว่า 1.5 เมตร ซึ่งเป็นข้อมูลสำคัญในการวางผังตลาดในด้านการใช้งาน อย่างไรก็ตาม ยังขาดการประยุกต์ใช้ความรู้ทางคณิตศาสตร์ในเชิงลึกเพื่อคำนวณและวิเคราะห์พื้นที่แบบจำลอง การศึกษานี้จึงนำแนวคิดจาก [4] รูปพาราโบลา [5] รูปวงรี และ [6] แคลคูลัสมาบูรณาการ เพื่อออกแบบแผนผังร้านค้าใหม่ภายในตลาดนัด มศว ให้ใช้พื้นที่ได้อย่างคุ้มค่า รองรับจำนวนร้านค้าได้มากขึ้น และช่วยลดปัญหาความแออัดบริเวณหน้าร้าน

โดยวัตถุประสงค์หลักของการศึกษานี้คือ เพื่อสร้างแบบจำลองแผนผังตลาดนัด มศว ที่เพิ่มจำนวนร้านค้าอย่างน้อยร้อยละ 25 ภายใต้ข้อจำกัดของพื้นที่เดิม พร้อมออกแบบพื้นที่หน้าร้านให้รองรับการยืนรอของลูกค้าโดยไม่กีดขวางทางเดิน การศึกษาครั้งนี้ไม่เพียงมุ่งแก้ปัญหาเชิงพื้นที่ของตลาดนัด มศว เท่านั้น แต่ยังเป็นต้นแบบของการประยุกต์ใช้คณิตศาสตร์กับปัญหาจริงในสังคม ซึ่งสามารถต่อยอดและปรับใช้กับพื้นที่ตลาดอื่นในอนาคตได้อย่างมีประสิทธิภาพ

2 ความรู้พื้นฐาน

2.1 นิยามศัพท์เฉพาะ

2.1.1. แผนผังตลาด หมายถึง การย่อส่วนของพื้นที่ตลาดเพื่อแสดงตำแหน่งที่ตั้งของร้านค้าแต่ละร้านค้าที่อยู่ภายในตลาด

2.1.2. แบบจำลองแผนผังตลาด หมายถึง แบบจำลองสองมิติและแบบจำลองสามมิติของแผนผังตลาดที่สร้างขึ้นใหม่

2.1.3. การใช้พื้นที่อย่างมีประสิทธิภาพ หมายถึง การออกแบบพื้นที่ของตลาดให้รองรับร้านค้าได้เพิ่มขึ้นอย่างน้อยร้อยละ 25 แม้พื้นที่จะมีขนาดเท่าเดิม โดยคำนึงถึงความสะดวกในการเดินของผู้มาใช้บริการ

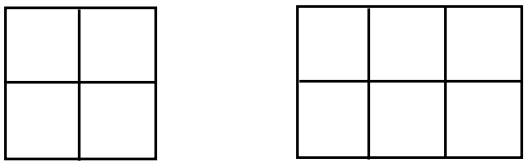
2.1.4. การวางพื้นที่ร้านค้าแบบมีประสิทธิภาพ หมายถึง การจัดวางผังร้านที่สามารถเพิ่มจำนวนร้านค้าให้เช่ามากขึ้นโดยใช้พื้นที่ของตลาดเท่าเดิม รวมถึงสามารถลดปัญหาการยืนกีดขวางทางเดินส่วนรวม

2.1.5. ความคุ้มค่าของพื้นที่ที่ถูกตัด หมายถึง ผลที่ได้จากการลดพื้นที่ภายในร้านลง แต่ทำให้ได้รับบริเวณหน้าร้านสำหรับรองรับลูกค้าได้มากขึ้น

2.1.6. ความสะดวกในการเดิน หมายถึง ผู้คนสามารถเดินได้อย่างคล่องตัว ไม่ต้องเบียดเสียดหรือหลบหลีกผู้อื่นมากนัก

2.1.7. รูปร่างของพื้นที่ร้าน หมายถึง ขอบเขตพื้นที่ใช้สอยของแต่ละร้านค้า

2.1.8. บล็อกร้านค้า หมายถึง ร้านค้าที่ไม่จัดวางเป็นแถวต่อกันในแนวยาว และเป็นร้านค้าที่ตั้งอยู่ติดกันจำนวน 4 ร้าน ขนาด 2 ร้านค้า × 2 ร้านค้า และร้านค้าที่ตั้งอยู่ติดกันจำนวน 6 ร้าน ขนาด 2 ร้านค้า × 3 ร้านค้า ดังรูป (1) และ (2) ตามลำดับ



บล็อกร้านค้า (1) บล็อกร้านค้า (2)

รูปที่ 1: บล็อกร้านค้า

2.2 ตลาดนัด มศว

ในปัจจุบัน [1] ตลาดนัดมหาวิทยาลัยศรีนครินทรวิโรฒ ได้มีการเปิดให้บริการทุกวันอังคารและวันพฤหัสบดี ตั้งแต่เวลา 06:30 น. ถึง 14:00 น. บริเวณด้านหลังอาคารบริการ ศ.ม.ล.ปิ่น มาลากุล ระหว่างอาคารบริการ ศ.ม.ล. ปิ่น มาลากุลกับอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี และหลังอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี รวมถึงมีพื้นที่ในการขายของแต่ละร้านค้าขนาด 2.5 x 2.5 เมตร โดยตลาดยังมีผู้มาใช้บริการและเป็นที่นิยมของผู้คนอย่างต่อเนื่อง

2.3 แคลคูลัส (Calculus)

2.3.1 อนุพันธ์ (Derivative)

1) ([7]) อนุพันธ์อันดับที่สาม (Third Derivative)

ผลลัพธ์ที่ได้จากการหาอนุพันธ์ของ $f''(x)$ เราเรียกว่า อนุพันธ์อันดับที่สาม ของ f ใช้สัญลักษณ์ $f'''(x), \frac{d^3y}{dx^3}, \frac{d^3}{dx^3}f(x)$ และ y'''

2) การประยุกต์ใช้อนุพันธ์อันดับที่สาม ในกราฟ

- ช่วยตรวจสอบลักษณะของกราฟในพื้นที่ที่อนุพันธ์อันดับที่หนึ่งและสองที่ไม่สามารถอธิบายได้ชัดเจน เช่น การเปลี่ยนแปลงลักษณะการโค้งในระดับสูง จุดที่ฟังก์ชันมีอัตราการเพิ่มสูงที่สุด
- ใช้ในการคาดการณ์รูปทรงของกราฟในบริเวณที่มีการเปลี่ยนแปลงซับซ้อน

2.3.2 ความยาวส่วนโค้ง

([6]) บทนิยาม 2.3.1 ถ้า $y = f(x)$ เป็นเส้นโค้งปรับเรียบบนช่วง $[a, b]$ แล้วความยาวส่วนโค้ง L ของเส้นโค้งบนช่วง $[a, b]$ คือ

$$L = \int_a^b \sqrt{1 + [f'(x)]^2}\,dx$$

ผลลัพธ์นี้ให้ทั้งคำจำกัดความและสูตรสำหรับการคำนวณหาความยาวเส้นโค้ง และบทนิยามดังกล่าวสามารถเขียนให้อยู่ในรูปได้เป็น

$$L = \int_a^b \sqrt{1 + [f'(x)]^2}\,dx = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx$$

นอกเหนือจากนั้น ส่วนโค้งในรูปของ $x = g(y)$ ที่ g' ต่อเนื่องบน $[c, d]$ ความยาวส่วนโค้ง L จาก $y = c$ ถึง $y = d$ สามารถเขียนให้อยู่ในรูป

$$L = \int_c^d \sqrt{1 + [g'(y)]^2}\,dy = \int_c^d \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\,dy$$

2.4 ([4]) พาราโบลา (Parabola)

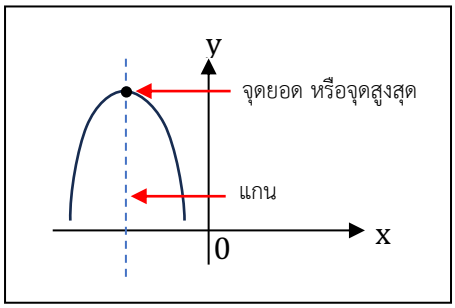
1) สมการของพาราโบลา

บทนิยาม 2.3.2 พาราโบลา (Parabola) เป็นชื่อเรียกกราฟของสมการ $y = ax^2 + bx + c$ เมื่อ x เป็นตัวแปรที่เป็นจำนวนจริง a, b และ c เป็นค่าคงตัว โดยที่ $a \neq 0$ และเรียกสมการดังกล่าวว่า สมการของพาราโบลา (Parabola Equation)

2) ลักษณะของกราฟพาราโบลาคว่ำ

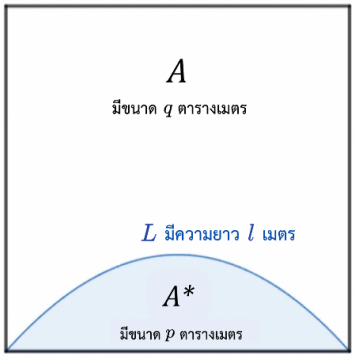
จากสมการพาราโบลา $y = ax^2 + bx + c$ เมื่อ $a \neq 0$ จะมีลักษณะกราฟเป็นรูปคว่ำ ขึ้นอยู่กับค่าของ a ซึ่งพิจารณาได้ ดังนี้

ถ้า $a < 0$ จะได้พาราโบลาคว่ำ



รูปที่ 2: ลักษณะของกราฟพาราโบลา

พิจารณาส่วนโค้งพาราโบลา สมการพาราโบลาทั่วไปที่เลือกใช้ คือ $y = ax^2 + k$ โดยที่ $y > 0$ และ $a < 0$ และให้ k เป็นความลึกของจุดยอด



รูปที่ 3: ตัวแปรที่ใช้

- กำหนดให้ A^* คือ บริเวณพื้นที่ภายใต้เส้นโค้ง (พื้นที่ที่ตัดออก) มีขนาดเป็น p ตารางเมตร
- กำหนดให้ A คือ บริเวณพื้นที่ใช้สอยภายในร้านมีขนาด เป็น q ตารางเมตร
- กำหนดให้ L คือ เส้นโค้งหน้าร้าน มีความยาวเป็น l เมตร
- กำหนดระยะห่างระหว่างเสาของร้านค้าหนึ่งกับร้านค้าตรงข้าม เป็นทางสัญจรในตลาด มีระยะห่าง 1.5 เมตร
- กำหนดให้ค่า $k \geq 0.457$ เมตร เพื่อให้พื้นที่หน้าร้านเพียงพอสำหรับลูกค้าหนึ่งคนสามารถยืนซื้อของหรือรอรับสินค้าได้สะดวก โดยที่ไม่กระทบทางสัญจรในตลาด โดยค่านี้มาจากการพิจารณาสัดส่วนพื้นที่ที่บุคคลหนึ่งต้องการในขณะยืนใช้งานหน้าร้าน

โดยกำหนดให้บริเวณหน้าร้านอยู่บนแกน X จุดกำเนิด $(0,0)$ เป็นจุดกึ่งกลางของหน้าร้านและเพื่อให้ง่ายต่อการคำนวณ พาราโบลาจะต้องผ่านจุด $(-1.25,0)$ และ $(1.25,0)$ เนื่องจากความยาวหน้าร้าน 2.5 เมตร จากพาราโบลาผ่านจุด $(1.25,0)$ จะได้ว่า $0 = a(1.25)^2 + k$ นั่นคือ $a = \frac{-k}{(1.25)^2}$

การหาพื้นที่ภายใต้เส้นโค้ง

พิจารณา $\int_{-1.25}^{1.25}(ax^2+k)\,dx = \int_{-1.25}^{1.25}\left(\frac{-k}{(1.25)^2}x^2+k\right)dx$

$$= \left[\frac{-kx^3}{3(1.25)^2}+kx\right]_{x=-1.25}^{x=1.25}$$

$$= \left[\frac{-k(1.25)^3}{3(1.25)^2}+k(1.25)\right]-\left[\frac{-k(-1.25)^3}{3(1.25)^2}+k(-1.25)\right]$$

$$= \frac{5k}{3}$$

จะได้พื้นที่ของ A^* มีขนาดคือ $p=\frac{5k}{3}$ ตารางเมตร

1. **การหาความยาวเส้นโค้ง**

จะได้ $l=\int_{-1.25}^{1.25}\sqrt{1+\left(\frac{-2kx}{(1.25)^2}\right)^2}\,dx$

พิจารณา $\int\sqrt{1+(u)^2}\,du$

ให้ $u=\tan\theta$ จะได้ $du=\sec^2\theta\,d\theta$ และ $\sqrt{1+u^2}=sec\,\theta$

จะได้ $\int\sqrt{1+(u)^2}\,du = \int\sec^3\theta\,d\theta$

$$= \frac{1}{2}\sec\theta\tan\theta+\frac{1}{2}\ln|\sec\theta+\tan\theta|+C$$

$$= \frac{1}{2}u\sqrt{1+u^2}+\frac{1}{2}\ln\left|\sqrt{1+u^2}+u\right|+C$$

ให้ $u=\frac{-2kx}{(1.25)^2}$ จะได้ $du=\frac{-2k}{(1.25)^2}dx$

เมื่อ $x=-1.25$ จะได้ $u=\frac{2k}{1.25}$

และเมื่อ $x=1.25$ จะได้ $u=\frac{-2k}{1.25}$

ดังนั้นจาก $l=\int_{-1.25}^{1.25}\sqrt{1+\left(\frac{-2kx}{(1.25)^2}\right)^2}\,dx$

$$= \frac{(1.25)^2}{-2k}\int_{\frac{2k}{1.25}}^{\frac{-2k}{1.25}}\sqrt{1+(u)^2}\,du$$

$$= \frac{(1.25)^2}{2k}\int_{\frac{-2k}{1.25}}^{\frac{2k}{1.25}}\sqrt{1+(u)^2}\,du$$

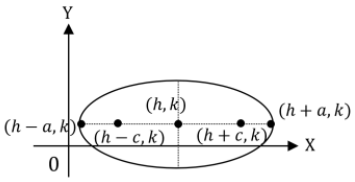
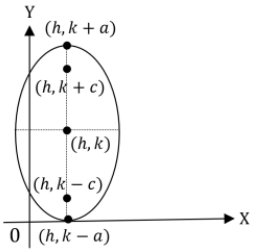
$$= \frac{(1.25)^2}{2k}\left[\frac{1}{2}u\sqrt{1+u^2}+\frac{1}{2}\ln\left|\sqrt{1+u^2}+u\right|\right]_{u=\frac{-2k}{1.25}}^{u=\frac{2k}{1.25}}$$

$$= \frac{1}{4}\sqrt{25+64k^2}+\frac{25}{32k}\ln\left|\frac{\sqrt{25+64k^2}+8k}{5}\right|$$

จะได้ L มีความยาวคือ $l=\frac{1}{4}\sqrt{25+64k^2}+\frac{25}{32k}\ln\left|\frac{\sqrt{25+64k^2}+8k}{5}\right|$ เมตร

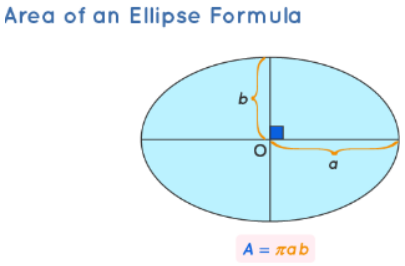
2.5 ([5]) วงรี (Ellipse)

สมการวงรีที่มีจุดศูนย์กลางอยู่ที่จุด (h,k) มีข้อสรุป ดังนี้

ลักษณะของวงรี	แกนเอกขนานแกน X	แกนเอกขนานแกน Y
		
สมการรูปแบบมาตรฐาน	$\frac{(x-h)^2}{a^2}+\frac{(y-k)^2}{b^2}=1$ โดยที่ $a>b>0$	$\frac{(x-h)^2}{b^2}+\frac{(y-k)^2}{a^2}=1$ โดยที่ $a>b>0$
จุดยอด	$(h-a,k)$ และ $(h+a,k)$	$(h,k-a)$ และ $(h,k+a)$
โฟกัส	$(h-c,k)$ และ $(h+c,k)$ โดยที่ $c^2=a^2-b^2$	$(h,k-a)$ และ $(h,k+a)$ โดยที่ $c^2=a^2-b^2$
ความยาวแกนเอก	$2a$ หน่วย	$2a$ หน่วย
ความยาวแกนโท	$2b$ หน่วย	$2b$ หน่วย

ตารางที่ 1: ข้อสรุปสมการวงรี

พื้นที่ของรูปวงรีที่มีแกนเอกยาว $2a$ หน่วย และแกนโทยาว $2b$ หน่วย จะมีสูตรในการคำนวณคือ πab เมื่อ a คือความยาวครึ่งแกนเอกของวงรี และ b คือความยาวครึ่งแกนโทของวงรี



([8]) รูปที่ 4: พื้นที่ของรูปวงรี

พิจารณาที่ส่วนโค้งวงรี สมการวงรีทั่วไปที่เลือกใช้ คือ $1=\frac{x^2}{a^2}+\frac{y^2}{b^2}$ โดยที่ $y>0$ และ $a=1.25$ และ ให้ k เป็นความสูงของจุดยอด จะได้ว่า $b=k$

- กำหนดให้ A^* คือบริเวณพื้นที่ภายใต้เส้นโค้ง (พื้นที่ที่ตัดออก) มีขนาดเป็น p ตารางเมตร
- กำหนดให้ A คือ บริเวณพื้นที่ใช้สอยภายในร้านมีขนาด เป็น q ตารางเมตร
- กำหนดให้ L คือ เส้นโค้งครึ่งวงรี มีความยาวเป็น l เมตร
- กำหนดระยะห่างระหว่างเสาของร้านค้าหนึ่งกับร้านค้าตรงข้าม เป็นทางสัญจรในตลาด มีระยะห่าง 1.5 เมตร

- กำหนดให้ค่า $k \geq 0.457$ เมตร เพื่อให้พื้นที่หน้าร้านเพียงพอสำหรับลูกค้าหนึ่งคนสามารถยืนซื้อสินค้าหรือรอรับสินค้าได้สะดวก โดยไม่กระทบทางสัญจรในตลาด โดยค่านี้มาจากการพิจารณาสัดส่วนพื้นที่ที่บุคคลหนึ่งต้องการในขณะยืนใช้งานหน้าร้าน

โดยวงรีจะต้องผ่านจุด $(-1.25,0)$ และ $(1.25,0)$ เนื่องจากความยาวหน้าร้าน 2.5 เมตร

1. การหาพื้นที่ภายใต้เส้นโค้ง

จากสูตรการหาพื้นที่ใต้วงรีจะได้ $p = \pi ab$

ในที่นี้ พิจารณาวงรีเฉพาะที่ $y > 0$ นั่นคือครึ่งวงรี จะได้ $p = \frac{\pi ab}{2}$

ดังนั้น พื้นที่ของ A^* มีขนาดคือ $p = \frac{1.25\pi k}{2}$ ตารางเมตร

2. การหาความยาวเส้นโค้ง

จาก $1 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$

พิจารณารูปวงรีที่ต้องการ จะได้ $1 = \frac{x^2}{1.25^2} + \frac{y^2}{k^2}$ นั่นคือ $a = 1.25$ และ $b = k$

ให้ $x = a\cos t$ และ $y = bsint$ โดยที่ $0 \leq t \leq 2\pi$

จะได้สูตรความยาวของส่วนโค้งวงรีทั้งหมดคือ $4a\int_0^{\frac{\pi}{2}} \sqrt{1 - m^2\cos^2 t}\, dt$

โดยที่ $m = \frac{\sqrt{a^2-b^2}}{a} = \frac{\sqrt{1.25^2-k^2}}{1.25}$

นั่นคือ $m^2 = \frac{1.25^2-k^2}{(1.25)^2}$

ดังนั้น ความยาวของส่วนโค้งวงรีทั้งหมดคือ $5\int_0^{\frac{\pi}{2}} \sqrt{1 - \frac{1.25^2-k^2}{(1.25)^2}\cos^2 t}\, dt$

จะได้ เส้นโค้ง L มีความยาวคือ $l = \frac{5}{2}\int_0^{\frac{\pi}{2}} \sqrt{1 - \frac{1.25^2-k^2}{(1.25)^2}\cos^2 t}\, dt$ เมตร

2.6 ([9]) การทำ Polynomial regression model fitting ด้วยโปรแกรม Microsoft Excel

การประยุกต์ใช้โปรแกรม Microsoft Excel สำหรับคำนวณการวิเคราะห์การถดถอยและสหสัมพันธ์ เพื่อพิจารณาความสัมพันธ์ของข้อมูล โดยทั่วไปอาจนำมาใช้ในการวิเคราะห์ข้อมูล 2 ประเภท ดังนี้

1) ความสัมพันธ์เชิงฟังก์ชันแบบกราฟเส้นตรง (linear function) ซึ่งอยู่ในรูปสมการ $y = ax + b$ เมื่อ y คือ ตัวแปรตาม x คือ ตัวแปรอิสระ a คือ ความชันของเส้นตรง และ b คือ จุดตัดแกน Y
2) ความสัมพันธ์เชิงฟังก์ชันแบบกราฟไม่เป็นเส้นตรง (non-linear function) เส้นกราฟอาจมีลักษณะเป็นรูปพาราโบลา หรือเอกซ์โพเนนเชียล หรือไฮเพอร์โบลา

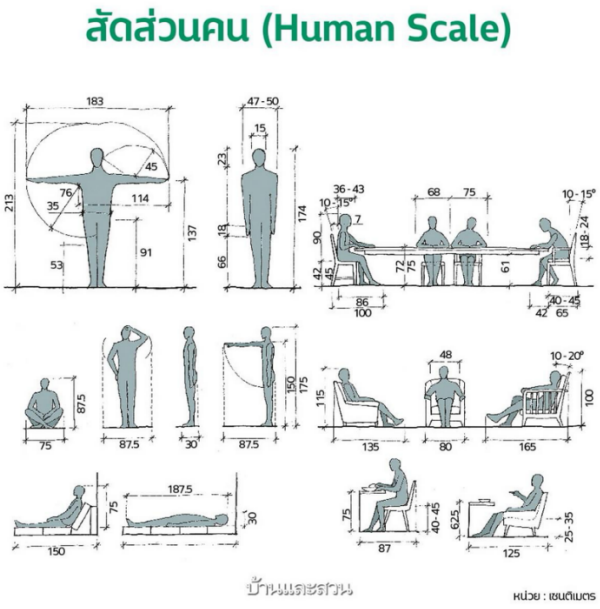
ค่าสัมประสิทธิ์สหสัมพันธ์เป็นค่าที่แสดงว่าตัวแปร x มีอิทธิพลต่อตัวแปร y มากน้อยเพียงใด โดยที่ค่าสัมประสิทธิ์สหสัมพันธ์แสดงด้วย r^2 โดยที่ $0 \leq r^2 \leq 1$ สรุปได้ดังนี้

- ถ้า r มีค่าใกล้ 1 แสดงว่า ตัวแปรสองตัวนั้นมีความสัมพันธ์กันมาก และมีทิศทางเดียวกัน กล่าวคือ ถ้า x มีค่ามาก y จะมีค่ามากด้วย
- ถ้า r มีค่าใกล้ -1 แสดงว่า ตัวแปรสองตัวนั้นมีความสัมพันธ์กันมาก แต่มีทิศทางตรงข้ามกัน กล่าวคือ ถ้า x มีค่ามาก y จะมีค่าน้อย หรือ x มีค่าน้อย y จะมีค่ามาก
- ถ้า r จะเข้าใกล้ 0 แสดงว่า x และ y มีความสัมพันธ์กันน้อย

2.7 สัดส่วนคน (Human scale)

[2] ในการออกแบบพื้นที่ใช้สอยทั่วไปภายในร้านค้า บริเวณที่ใช้สำหรับการปรุงอาหาร ควรกำหนดระยะทางสัญจรให้มีความกว้างอย่างน้อย 121.9 เซนติเมตร และความกว้างของทางเดินบริเวณหน้าร้านซึ่งลูกค้าใช้สำหรับหยุดยืนเลือกซื้อสินค้า ควรมีความกว้างไม่น้อยกว่า 45.7 เซนติเมตร หรือ 0.457 เมตร

สำหรับการออกแบบผังตลาดนั้น จำเป็นต้องคำนึงถึงความสะดวกและการใช้งานได้จริงในบริบทของผู้ใช้งานทุกกลุ่ม การออกแบบจึงต้องเริ่มต้นจากการพิจารณา [10] สัดส่วนและพฤติกรรมของคน โดยเฉพาะระยะที่เกี่ยวข้องกับท่าทางพื้นฐาน เช่น การเดิน การนั่ง และการยืน เพื่อใช้ข้อมูลเหล่านี้เป็นพื้นฐานในการออกแบบที่ตอบสนองต่อความต้องการและพฤติกรรมของผู้ใช้งานได้อย่างมีประสิทธิภาพ



รูปที่ 5: สัดส่วนคน

3 ผลการศึกษา

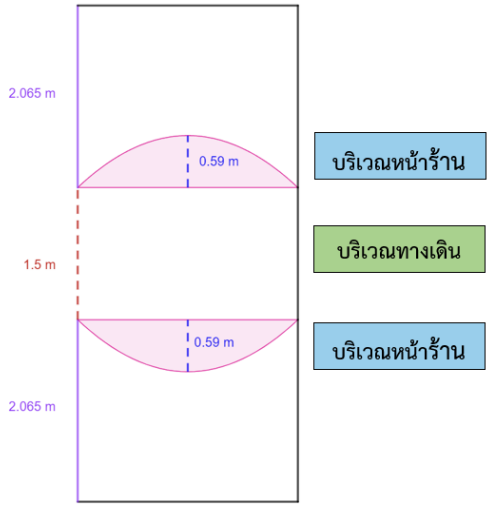
3.1 การหาค่า k ที่เหมาะสม สำหรับส่วนโค้งของพาราโบลา

แบ่งออกเป็น 2 กรณี

3.1.1 การหาค่า k ที่เหมาะสม กรณีที่ 1 กำหนดค่า k จากการอ้างอิงความกว้างของพื้นที่ที่แคบที่สุด แต่ต้องมีค่ามากกว่า 5 เมตร ดังนั้นเพื่อให้พื้นที่เพียงพอสำหรับวางร้านค้าสองร้านค้าภายในบริเวณตลาดโดยอิงจากผังตลาดปัจจุบัน เราจะเลือกจากบริเวณที่มีความกว้าง 5.63 เมตร และกำหนดระยะห่างระหว่างเสาร้านค้าหนึ่งกับเสาของร้านค้าฝั่งตรงข้าม เป็น 1.5 เมตร

จากการสังเกต พบว่าพื้นที่มีความกว้าง 5.63 เมตร หากวางร้านค้าสองร้านหันหน้าเข้าหากัน จะเหลือระยะห่างระหว่างเสาของร้านค้าเพียง 0.63 เมตร ซึ่งไม่สอดคล้องกับเงื่อนไขที่กำหนดให้ทางสัญจรต้องมีความกว้างอย่างน้อย 1.5 เมตร ดังนั้น ทางคณะผู้จัดทำจึงพิจารณาปรับพื้นที่โดย กำหนดความกว้างขั้นต่ำของพื้นที่ร้านค้า เพื่อให้เพียงพอต่อการใช้งานพื้นฐานของร้านค้าแต่ละร้าน และลดความลึกของร้านค้า เพื่อเพิ่มพื้นที่ทางเดินและคงไว้ซึ่งความสมดุลในการบริหารจัดการพื้นที่ของร้านค้า โดยการปรับขนาดดังกล่าวจะช่วยให้พื้นที่สามารถรองรับทั้งร้านค้าและความสะดวกในการสัญจรของลูกค้าได้อย่างเหมาะสม

โดยเราจะกำหนดความกว้างขั้นต่ำของพื้นที่ร้านค้า โดยพิจารณาความกว้างของพื้นที่ลบกับทางเดิน จะได้ $5.63 - 1.5 = 4.13$ จากนั้นนำ 4.13 หารด้วย 2 เพื่อได้ความลึกของร้าน นั่นคือ 2.065 เมตร จะหาค่า k จากการกำหนดความกว้างของโต๊ะทั่วไป ซึ่งมีความกว้างคือ 0.6 เมตร โดยจะเริ่มนับจากจุดยอดของพาราโบลา และจากสัดส่วนคนเพื่อหาความกว้างของคนยืนและสามารถยื่นมือไปได้หนึ่งช่วงแขน นั่นคือ 0.875 เมตร เพื่อความสะดวกต่อการค้าขาย จะได้ค่า $k = 2.065 - (0.6 + 0.875) = 0.59$



รูปที่ 6: ร้านค้า 2 ร้านค้าที่หันหน้าร้านเข้าหากัน

จะได้ $p = \frac{5(0.59)}{3} \approx 0.983$ ตารางเมตร

และ $l = \frac{1}{4}\sqrt{25+64(0.59)^2} + \frac{25}{32(0.59)} \ln\left|\frac{\sqrt{25+64(0.59)^2}+8(0.59)}{5}\right| \approx 2.833$ เมตร

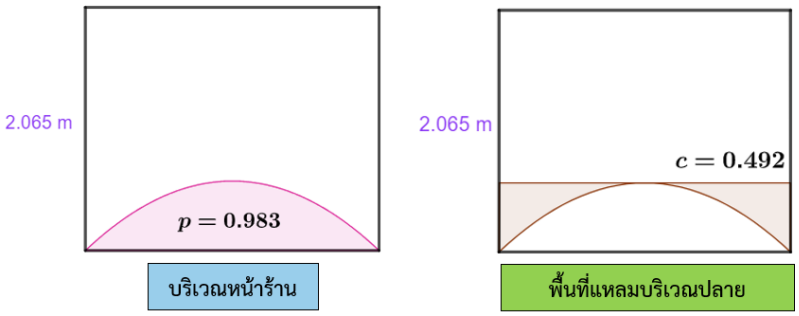
พื้นที่แหลมบริเวณปลายทั้งสองด้านหาได้โดยการอินทิเกรตระหว่าง $y = k$ และ

$y = \frac{-k}{(1.25)^2}x^2 + k$ โดย $-1.25 \le x \le 1.25$

นั่นคือ $\int_{-1.25}^{1.25}\left[(0.59) - \frac{-0.59}{(1.25)^2}x^2 - 0.59\right]dx \approx 0.492$ ตารางเมตร

และอัตราส่วนระหว่าง p และ l คือ $\frac{0.983}{2.833} \approx 0.347$

ได้ดังรูป

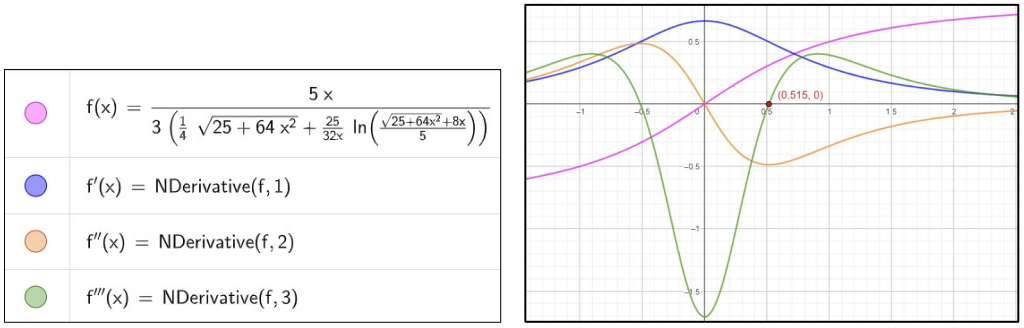


รูปที่ 7: ร้านค้าในกรณีที่ 1 ของส่วนโค้งพาราโบลา

3.1.2 การหาค่า k ที่เหมาะสม กรณีที่ 2 กำหนดค่า k โดยใช้การหาอนุพันธ์อันดับที่สามของอัตราส่วนระหว่าง p และ l

จะได้อัตราส่วนระหว่าง p และ l คือ $\frac{5k}{3\left(\frac{1}{4}\sqrt{25+64k^2}+\frac{25}{32k}\ln\left|\frac{\sqrt{25+64k^2}+8k}{5}\right|\right)}$

เพื่อให้ง่ายต่อการหาอนุพันธ์เราจะหาอนุพันธ์อันดับที่สาม ด้วยการพิจารณาลักษณะกราฟผ่านโปรแกรม GeoGebra ดังรูป



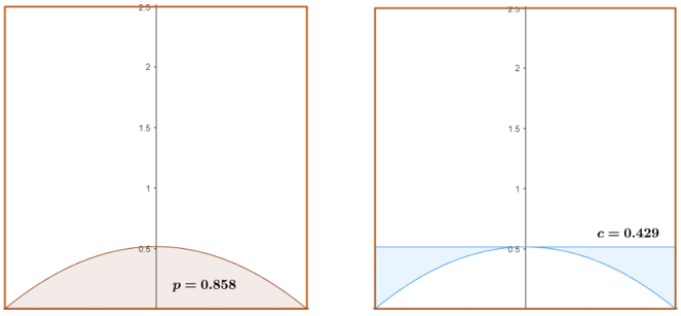
รูปที่ 8: กราฟอนุพันธ์อันดับต่าง ๆ ของ f

ดังนั้นเราจะได้จุดตัดแกน X ของอนุพันธ์อันดับที่สาม ของ $\frac{5k}{3\left(\frac{1}{4}\sqrt{25+64k^2}+\frac{25}{32k}\ln\left|\frac{\sqrt{25+64k^2}+8k}{5}\right|\right)}$ อยู่ที่ $(0.515,0)$ เมื่อ $k>0$

นั่นหมายความว่า เมื่อ $k=0.515$ จะทำให้เกิดความคุ้มค่าของพื้นที่ที่ถูกตัดไปเมื่อเทียบกับความยาวหน้าร้านที่ได้รับมามากที่สุด

จะได้ $p=\frac{5(0.515)}{3}\approx 0.858$ ตารางเมตร และ $l=\frac{1}{4}\sqrt{25+64(0.515)^2}+\frac{25}{32(0.515)}\ln\left|\frac{\sqrt{25+64(0.515)^2}+8(0.515)}{5}\right|\approx 2.759$ เมตร

พื้นที่แหลมบริเวณปลายทั้งสองด้านหาได้โดยการอินทิเกรตระหว่าง $y=k$ และ $y=\frac{-k}{(1.25)^2}x^2+k$ โดย $-1.25\le x\le 1.25$
นั่นคือ $\int_{-1.25}^{1.25}\left[(0.515)-\frac{-0.515}{(1.25)^2}x^2-0.515\right]dx\approx 0.429$ ตารางเมตร
และอัตราส่วนระหว่าง p และ l คือ $\frac{0.858}{2.759}=0.311$
ได้ดังรูป



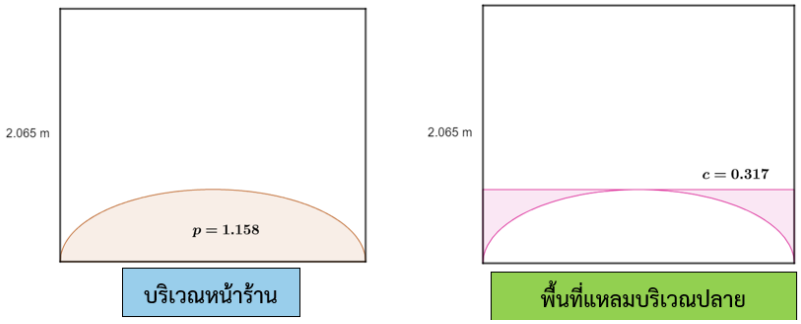
รูปที่ 9: ร้านค้าในกรณีที่ 2 ของส่วนโค้งพาราโบลา

3.2 การหาค่า k ที่เหมาะสม สำหรับส่วนโค้งของวงรี

แบ่งออกเป็น 2 กรณีเช่นเดียวกับส่วนโค้งของพาราโบลา แต่ในการหาค่า k ที่เหมาะสมสำหรับส่วนโค้งของวงรี กรณีที่ 2 มีความแตกต่างจากพาราโบลา เนื่องจากสมการความยาวส่วนโค้งแสดงอยู่ในรูป $l=\frac{5}{2}\int_0^{\frac{\pi}{2}}\sqrt{1-\frac{1.25^2-k^2}{(1.25)^2}\cos^2 t}\,dt$ ซึ่งไม่สามารถหาค่าอนุพันธ์อันดับที่สาม ได้โดยตรง ในการแก้ปัญหานี้เราจึงเลือกใช้การประมาณค่า l ด้วยสมการพหุนามแทน จากนั้นจึงคำนวณหาอนุพันธ์อันดับที่สาม ของอัตราส่วนระหว่าง p และ l

3.2.1 การหาค่า k ที่เหมาะสม กรณีที่ 1 กำหนดค่า k จากการอ้างอิงความกว้างของพื้นที่ที่แคบที่สุด แต่ต้องมีค่ามากกว่า 5 เมตร ดังนั้นเพื่อให้พื้นที่เพียงพอสำหรับวางร้านค้าสองร้านค้าภายในบริเวณตลาด โดยอิงจากผังตลาดปัจจุบัน เราจะเลือกจากบริเวณที่มีความกว้าง 5.63 เมตร และกำหนดระยะห่างระหว่างเสาร้านค้าหนึ่งกับเสาของร้านค้าฝั่งตรงข้าม เป็น 1.5 เมตร

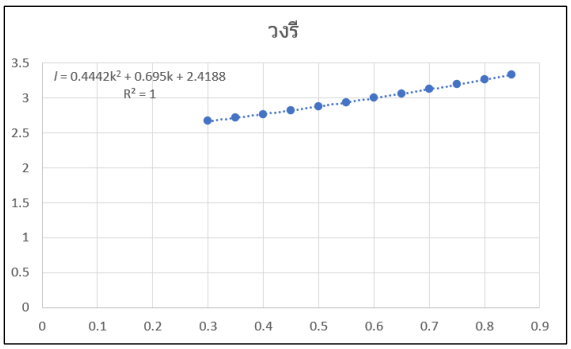
จากกรณีเดียวกับพาราโบลาจะได้ $k = 0.59$ และ $p \approx 1.158$ ตารางเมตร และ $l \approx 2.984$ เมตร และพื้นที่แหลมบริเวณปลายทั้งสองด้านหาได้โดยการอินทิเกรต คือ 0.317 ตารางเมตร และอัตราส่วนระหว่าง p และ l คือ 0.388 ได้ดังรูป



รูปที่ 10: ร้านค้าในกรณีที่ 1 ของส่วนโค้งวงรี

3.2.2 การหาค่า k ที่เหมาะสม กรณีที่ 2 กำหนดค่า k โดยใช้การหาอนุพันธ์อันดับที่สาม ของอัตราส่วนระหว่าง p และ l เพื่อหาความคุ้มค่าของพื้นที่ที่ถูกตัดไปเมื่อเทียบกับความยาวหน้าร้านที่เพิ่มขึ้น

เนื่องจากค่า l สำหรับกรณีวงรีอยู่ในรูปของอินทิกรัลซึ่งยากต่อการคำนวณ เราจึงจะประมาณค่า l ด้วยพหุนามจากโปรแกรม Microsoft Excel โดย $l = \frac{5}{2}\int_0^{\frac{\pi}{2}} \sqrt{1 - \frac{1.25^2 - k^2}{(1.25)^2}\cos^2 t}\, dt$ จะได้ว่า $l = 0.4442k^2 + 0.695k + 2.4188$ ดังรูป



รูปที่ 11: การประมาณค่า l ด้วยพหุนาม

จะได้ อัตราส่วนระหว่าง p และ l คือ $\frac{0.625\pi k}{0.4442k^2+0.695k+2.4188}$ ดังนั้นเราจะหาอนุพันธ์อันดับที่สาม ของ $\frac{0.625\pi k}{0.4442k^2+0.695k+2.4188}$ โดยที่ $\frac{0.625\pi k}{0.4442k^2+0.695k+2.4188}=0$ จะได้จุดวิกฤตของอนุพันธ์อันดับที่สอง อยู่ที่ $(0.357,0)$ เมื่อ $k>0$ ดังรูป



รูปที่ 12: กราฟอนุพันธ์อันดับต่าง ๆ ของอัตราส่วนระหว่าง p และ l ของวงรี

นั่นหมายความว่า เมื่อ $k=0.357$ จะทำให้เกิดความคุ้มค่าของพื้นที่ที่ถูกตัดไปเมื่อเทียบกับความยาวหน้าร้านที่ได้รับมามากที่สุด

จาก $k=0.357$

จะได้ $p=\frac{1.25\pi(0.357)}{2}\approx 0.701$ ตารางเมตร

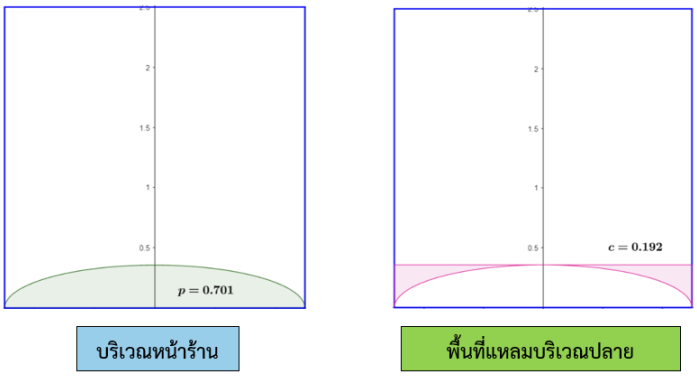
และ $l=\frac{5}{2}\int_0^{\frac{\pi}{2}}\sqrt{1-\frac{1.25^2-(0.357)^2}{(1.25)^2}\cos^2 t}\,dt\approx 2.723$ เมตร

และพื้นที่แหลมบริเวณปลายทั้งสองด้านหาได้โดยการอินทิเกรตระหว่าง $y=k$ และ $y=\sqrt{k^2-\frac{(kx)^2}{(1.25)^2}}$ โดย $-1.25\le x\le 1.25$

นั่นคือ $\int_{-1.25}^{1.25}\left[(0.357)-\sqrt{(0.357)^2-\frac{((0.357)x)^2}{(1.25)^2}}\right]dx\approx 0.192$ ตารางเมตร

และอัตราส่วนระหว่าง p และ l คือ $\frac{0.701}{2.723}\approx 0.257$

ได้ดังรูป



รูปที่ 13: ร้านค้าในกรณีที่ 2 ของส่วนโค้งวงรี

3.3 การออกแบบพื้นที่หน้าร้านค้าโดยใช้เส้นโค้งและรูปเรขาคณิตสองมิติ

ทางคณะผู้จัดทำได้พิจารณาเลือกส่วนโค้งพาราโบลาและส่วนโค้งวงรีมาใช้ในการออกแบบพื้นที่หน้าร้าน โดยกำหนดให้เส้นโค้งแต่ละประเภทมีคุณสมบัติ ดังนี้

1. เส้นโค้งทั้งหมดอยู่บนแกน X โดยจุดกึ่งกลางของเส้นโค้งอยู่บนแกน Y และมีแกนสมมาตรเป็นแกน Y
2. ส่วนโค้งพาราโบลา เป็นรูปพาราโบลาคว่ำ
3. ส่วนโค้งวงรี เป็นรูปครึ่งวงรีแนวนอน

และมีสมการทั่วไป ดังนี้

1. ส่วนโค้งพาราโบลา : $y = ax^2 + k$; โดยที่ $(0, k)$ เป็นจุดยอดของพาราโบลา และ $a < 0$ และ $k > 0$
2. ส่วนโค้งวงรี : $y = \sqrt{b^2 - \left(\frac{bx}{a}\right)^2}$; โดยที่ $(0, k)$ เป็นค่าสูงสุดของวงรีและ $k > 0$

การพิจารณาหาค่า k ที่เหมาะสมในการศึกษานี้ คณะผู้จัดทำได้เลือกศึกษา 2 กรณี

กรณีที่ 1 : กำหนดค่า k จากการอ้างอิงความกว้างของพื้นที่ที่แคบที่สุด แต่ต้องมีค่ามากกว่า 5 เมตร ดังนั้นเพื่อให้พื้นที่เพียงพอสำหรับวางร้านค้าสองร้านค้าภายในบริเวณตลาดโดยอิงจากผังตลาดปัจจุบัน เราจะเลือกจากบริเวณที่มีความกว้าง 5.63 เมตร และกำหนดระยะห่างระหว่างเสาร้านค้าหนึ่งกับเสาของร้านค้าฝั่งตรงข้าม เป็น 1.5 เมตร นั่นคือ $k = 0.59$

กรณีที่ 2 : กำหนดค่า k โดยใช้การหาอนุพันธ์อันดับที่สาม ของอัตราส่วนระหว่าง p และ l เพื่อหาความคุ้มค่าของพื้นที่ที่ถูกตัดไปเมื่อเทียบกับความยาวหน้าร้านที่เพิ่มขึ้น นั่นคือ $k = 0.515$ ในกรณีของรูปพาราโบลา และ $k = 0.357$ ในกรณีของรูปครึ่งวงรี

3.4 การสร้างแบบจำลองแผนผังสำหรับตลาดนัด มศว

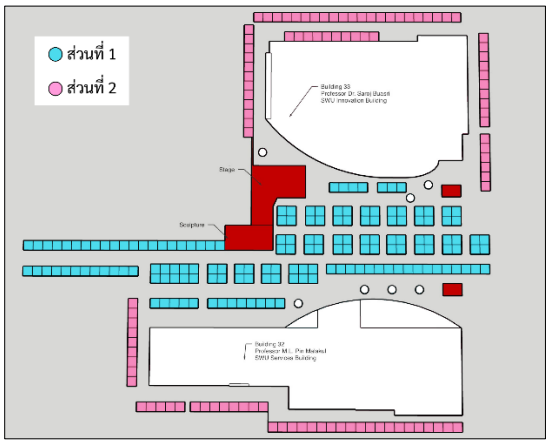
คณะผู้จัดทำได้พิจารณาการจัดวางแผนผังโดยเลือกกรณีของร้านค้าที่มีความคุ้มค่าที่สุดในแต่ละพื้นที่ ดังนี้

1. บริเวณระหว่างอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี เลือกกรณีที่ 2 ของรูปพาราโบลา ซึ่งพิจารณาความคุ้มค่าของพื้นที่ที่ถูกตัดไปเมื่อเทียบกับความยาวหน้าร้านที่เพิ่มขึ้น และเนื่องจากค่า k ในกรณีที่ 2 ของรูปครึ่งวงรี มีค่าเพียง 0.357 เมตร ซึ่งส่งผลให้พื้นที่ใช้สอยมีขนาดเล็กเกินไป ไม่เพียงพอสำหรับให้ลูกค้ายืนเลือกซื้อสินค้าได้อย่างสะดวก
2. บริเวณหลังอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และหลังอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี เลือกกรณีที่ 1 ของทั้งรูปพาราโบลาและครึ่งวงรี เนื่องจากบริเวณนี้มีความกว้างของพื้นที่ที่จำกัดซึ่งแคบกว่า

พื้นที่บริเวณอื่น ดังนั้นการเลือกใช้กรณีนี้จะช่วยให้การจัดวางร้านค้า มีจำนวนร้านเพิ่มมากขึ้น โดยยังคงคำนึงถึงการรักษาพื้นที่สัญจรและพื้นที่ภายในร้านค้าในบริเวณดังกล่าว

ในการสร้างแบบจำลองแผนผังตลาดนัด มศว คณะผู้จัดทำขอเลือกใช้รูปพาราโบลาเพียงรูปเดียว เนื่องจากง่ายต่อการออกแบบ และพื้นที่หน้าร้านที่ได้จากรูปครึ่งวงรีในกรณีที่ 2 อาจไม่สามารถรองรับการยืนเลือกซื้อสินค้าของลูกค้าได้อย่างเหมาะสม ส่งผลให้เกิดความแออัด และขาดความสะดวกในการเลือกซื้อสินค้า

การจัดรูปแบบแผนผังตลาดเพื่อเพิ่มจำนวนร้านค้าและยังคงรักษาความสะดวกระหว่างพื้นที่สัญจรและทางร้านค้า โดยคณะผู้จัดทำได้พิจารณาออกแบบแผนผังมาทั้งหมด 2 กรณี โดยมีวัตถุประสงค์คือเพิ่มจำนวนร้านค้าอย่างน้อยร้อยละ 25 จากจำนวนร้านค้าปัจจุบันทั้งหมด โดยมีข้อกำหนด คือ ความกว้างของทางสัญจรอย่างน้อย 1.5 เมตร และความลึกของร้านค้ามากกว่า 0.457 เมตร ($k > 0.457$) โดยกำหนดพื้นที่ 2 ส่วน ดังนี้

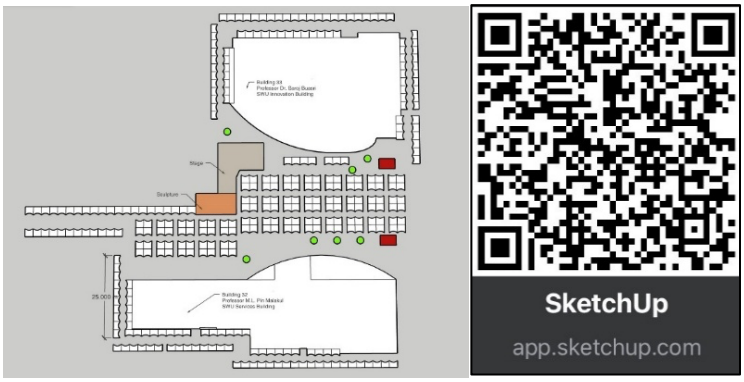


รูปที่ 14: การกำหนดพื้นที่ตลาด

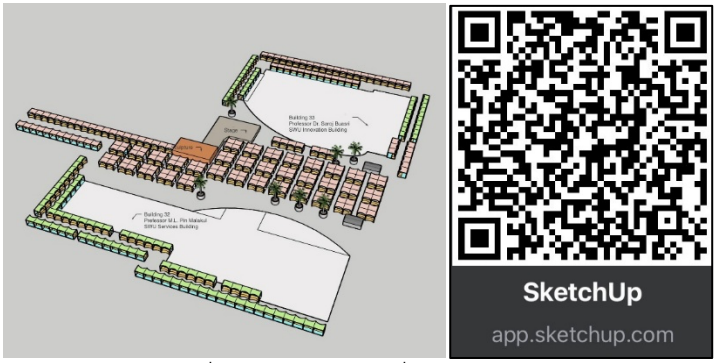
1. บริเวณระหว่างอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี บริเวณนี้ถูกแบ่งเป็น 2 กรณี คือ อิงจากผังปัจจุบันแล้วปรับรูปแบบโดยการตัดส่วนโค้งออกเพื่อเพิ่มพื้นที่ และกรณีที่ 2 คือ เพิ่มจำนวนร้านค้าในแต่ละบล็อกร้านค้า จาก 4 ร้าน เป็น 6 ร้าน เนื่องจากการสำรวจตลาด พบว่าผู้คนส่วนใหญ่มักเดินในแนวยาวของตลาดและไม่ค่อยเดินออกระหว่างบล็อกร้านค้าในแต่ละจุด ทางคณะผู้จัดทำจึงออกแบบให้เพิ่มจำนวนร้านค้าต่อบล็อกเพื่อใช้พื้นที่ให้มีประสิทธิภาพสูงที่สุด
2. บริเวณหลังอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และหลังอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี สำหรับพื้นที่นี้ บริเวณที่มีความกว้างมากกว่า 5 เมตร ทางคณะผู้จัดทำเลือกใช้แนวทางการเพิ่มแถวร้านค้าเพื่อให้มีร้านค้าเพิ่ม โดยรักษาความกว้างของทางเดินให้ได้อย่างน้อย 1.5 เมตร โดยข้อนี้จะอยู่ในแผนผังทั้ง 2 กรณี ของข้อ 1.

3.4.1 แผนผังกรณีที่ 1

บริเวณระหว่างอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี ออกแบบโดยอิงจากผังปัจจุบันแล้วปรับรูปแบบโดยการตัดส่วนโค้งออกเพื่อเพิ่มพื้นที่ จะได้จำนวนร้านค้าทั้งหมด 326 ร้าน



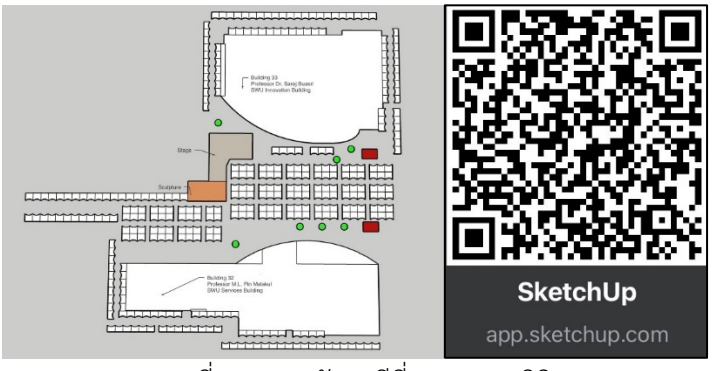
รูปที่ 15: แผนผังกรณีที่ 1 แบบ 2 มิติ



รูปที่ 16: แผนผังกรณีที่ 1 แบบ 3 มิติ

3.4.2 แผนผังกรณีที่ 2

บริเวณระหว่างอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี ออกแบบโดยเพิ่มจำนวนร้านค้าในแต่ละบล็อกร้านค้า จาก 4 ร้าน เป็น 6 ร้าน จะได้จำนวนร้านค้าทั้งหมด 342 ร้าน



รูปที่ 17: แผนผังกรณีที่ 2 แบบ 2 มิติ



รูปที่ 18: แผนผังกรณีที่ 2 แบบ 3 มิติ

4 สรุปผลและอภิปรายผล

4.1 สรุปผล

จากการศึกษา วิเคราะห์ และสร้างแบบจำลองแผนผังสำหรับตลาดนัด มศว ทำให้ได้ว่า การทำให้หน้าร้านเป็นส่วนโค้งครึ่งวงรีเหมาะสำหรับการเพิ่มพื้นที่รับลูกค้า แต่ต้องแลกกับการลดพื้นที่ใช้งานภายในร้านค้าบางส่วนไป และไม่สามารถใช้ได้ในทุกพื้นที่ของตลาดนัด มศว กล่าวคือ ใช้ได้เพียงบริเวณหลังอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และหลังอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี เท่านั้น และรูปพาราโบลาอาจใช้พื้นที่อย่างคุ้มค่ามากกว่าในภาพรวม แต่มีข้อเสียในบริเวณปลายแคบที่ไม่สามารถใช้งานได้ ส่วนการเลือกใช้งานรูปใดรูปหนึ่งอาจขึ้นอยู่กับวัตถุประสงค์การใช้งานและข้อจำกัดในพื้นที่ ที่อาจต้องพิจารณาจากปัจจัยอื่น ๆ เพิ่มเติม เช่น จำนวนลูกค้า สินค้าที่วางขาย และลักษณะพื้นที่ของร้านค้า

ในการศึกษานี้คณะผู้จัดทำได้เลือกศึกษากรณีที่ครอบคลุมทั้งความต้องการด้านการใช้งานพื้นที่ตลาดและร้านค้าโดยเลือกศึกษา 2 กรณี ดังนี้

กรณีที่ 1 : กำหนดค่า k จากการอ้างอิงความกว้างของพื้นที่ที่แคบที่สุด แต่ต้องมีค่ามากกว่า 5 เมตร ดังนั้นเพื่อให้พื้นที่เพียงพอสำหรับวางร้านค้าสองร้านค้าภายในบริเวณตลาดโดยอิงจากผังตลาดปัจจุบัน เราจะเลือกจากบริเวณที่มีความกว้าง 5.63 เมตร และกำหนดระยะห่างระหว่างเสาร้านค้าหนึ่งกับเสาของร้านค้าฝั่งตรงข้าม เป็น 1.5 เมตร นั่นคือ $k = 0.59$

กรณีที่ 2 : กำหนดค่า k โดยใช้การหาอนุพันธ์อันดับที่สาม ของอัตราส่วนระหว่าง p และ l เพื่อหาความคุ้มค่าของพื้นที่ที่ถูกตัดไปเมื่อเทียบกับความยาวหน้าร้านที่เพิ่มขึ้น นั่นคือ $k = 0.515$ ในกรณีของรูปพาราโบลา

สำหรับการศึกษานี้ คณะผู้จัดทำได้ทำการจัดรูปแบบแผนผังตลาดเพื่อเพิ่มจำนวนร้านค้าและยังคงรักษาความสะดวกระหว่างพื้นที่สัญจรและทางร้านค้า จึงออกแบบแผนผังทั้งหมด 2 กรณี ที่มีวัตถุประสงค์คือเพิ่มจำนวน

ร้านค้าอย่างน้อยร้อยละ 25 จากจำนวนร้านค้าปัจจุบันทั้งหมด โดยมีข้อกำหนด คือ ความกว้างของทางสัญจรอย่างน้อย 1.5 เมตร และความลึกของร้านค้ามากกว่า 0.457 เมตร ($k > 0.457$) โดยกำหนดพื้นที่ 2 ส่วน ดังนี้

1. บริเวณระหว่างอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี บริเวณนี้ถูกแบ่งเป็น 2 กรณี คือ อิงจากผังปัจจุบันแล้วปรับรูปแบบโดยการตัดส่วนโค้งออกเพื่อเพิ่มพื้นที่ และกรณีที่ 2 คือ เพิ่มจำนวนร้านค้าในแต่ละบล็อคร้านค้า จาก 4 ร้าน เป็น 6 ร้าน เนื่องจากการสำรวจตลาด พบว่าผู้คนส่วนใหญ่มักเดินในแนวยาวของตลาดและไม่ค่อยเดินออกระหว่างบล็อคร้านค้าในแต่ละจุด ทางคณะผู้จัดทำจึงออกแบบให้เพิ่มจำนวนร้านค้าต่อบล็อกเพื่อใช้พื้นที่ให้มีประสิทธิภาพสูงที่สุด
2. บริเวณหลังอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล และหลังอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี สำหรับพื้นที่นี้บริเวณที่มีความกว้างมากกว่า 5 เมตร ทางคณะผู้จัดทำเลือกใช้แนวทางการเพิ่มแถวร้านค้า เพื่อให้มีร้านค้าเพิ่ม โดยรักษาความกว้างของทางเดินให้ได้อย่างน้อย 1.5 เมตร โดยข้อนี้จะอยู่ในแผนผังทั้ง 2 กรณี ของข้อ 1.

จากข้างต้นทำให้ได้ข้อมูล ดังนี้

ตารางที่ 2: สรุปจำนวนร้านค้าที่ได้ในแต่ละกรณี

	หลังอาคารบริการ ศ.ม.ล. ปิ่น มาลากุล (พื้นที่ส่วนที่ 2)	หลังอาคารนวัตกรรม ศ.ดร. สาโรช บัวศรี (พื้นที่ส่วนที่ 2)	บริเวณระหว่างอาคาร (พื้นที่ส่วนที่ 1)	รวม	จำนวนร้านที่เพิ่มขึ้นจากผังปัจจุบัน (ร้าน)
จำนวนร้านผังปัจจุบัน (ร้าน)	42	60	158	260	-
จำนวนร้านผังใหม่กรณีที่ 1 (ร้าน)	75	82	169	326	66
จำนวนร้านผังใหม่กรณีที่ 2 (ร้าน)	75	82	185	342	82

4.2 อภิปรายผล

จากการพิจารณาแผนผังใหม่ทั้ง 2 กรณี สรุปได้ว่าแผนผังใหม่ในกรณีที่ 1 มีจำนวนร้านค้าเพิ่มขึ้น 66 ร้าน คิดเป็นร้อยละ 25.38 จากจำนวนร้านค้าปัจจุบันทั้งหมด และแผนผังใหม่ในกรณีที่ 2 มีจำนวนร้านค้าเพิ่มขึ้น 82 ร้าน คิดเป็นร้อยละ 31.54 จากจำนวนร้านค้าปัจจุบันทั้งหมด โดยแผนผังใหม่ทั้ง 2 กรณีมีจำนวนร้านค้าเพิ่มขึ้นมากกว่าร้อยละ 25 จากจำนวนร้านค้าปัจจุบัน และมีขนาดความกว้างของทางสัญจรตรงตามวัตถุประสงค์ที่กำหนดไว้ข้างต้น

ข้อเสนอแนะ

ในกรณีที่ต้องการศึกษาข้อมูลเชิงลึกใด ๆ ที่เกี่ยวข้องกับตลาดนัด มศว ควรจะติดต่อดำเนินการขอข้อมูลก่อนล่วงหน้า เนื่องจากงานเอกสารต้องดำเนินการตามขั้นตอนของมหาวิทยาลัยซึ่งต้องใช้เวลาในการประสานงานอย่างเป็นระบบ

กิตติกรรมประกาศ

การจัดทำการศึกษานี้ สำเร็จลุล่วงไปได้ด้วยความกรุณาและความช่วยเหลืออย่างสูง ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.ฐิตารีย์ รุ่งรัตน์เกษม อาจารย์ที่ให้คำปรึกษา คำแนะนำ ชี้ให้เห็นถึงปัญหา และแนะแนวทางแก้ไขปัญหา อ.นพดล วิชิตสงคราม อาจารย์ที่ปรึกษาในการคำนวณคณิตศาสตร์ และชี้แนะแนวทาง ผู้ช่วยศาสตราจารย์ ดร.ขวัญ เพียซ้าย และผู้ช่วยศาสตราจารย์ ดร.สุกัญญา หะยีสาและ ที่ช่วยตรวจสอบจนเสร็จสมบูรณ์

ตลอดจนขอขอบคุณผู้ที่เกี่ยวข้องทุกท่านที่ไม่ได้กล่าวนามไว้ ณ ที่นี้ ที่ได้ให้กำลังใจและมีส่วนช่วยเหลือในการศึกษานี้ให้สำเร็จลุล่วงไปได้ด้วยดี

เอกสารอ้างอิง

[1] ชวิตรา ตันติมาลา, อัตลักษณ์ที่ลื่นไหลของตลาดนัดนานาชาติ มหาวิทยาลัศรีนครินทรวิโรฒ, วารสารสถาบันวัฒนธรรมและศิลปะ, 17(2), 159--170. <https://so02.tci-thaijo.org/index.php/jica/article/view/99323>, 2559 (สืบค้น 4 พ.ค. 68)

[2] ชูใจ ประสาทเสรี, มนุษย์มิติ, สำนักพิมพ์ สถาบันเทคโนโลยีพระจอมเกล้า พระนครเหนือ, กรุงเทพมหานคร, 2526.

[3] ณัฏฐาศักดิ์ ทิษฏิญาณิณ, ดลินา อมรเหมานนท์, การศึกษาวิเคราะห์ระบบการสัญจรทางเท้าและสภาพแวดล้อมทางกายภาพของเส้นทางในเมืองย่านตาขาว จังหวัดตรัง, วารสารวิชาการคณะสถาปัตยกรรม สจล., 36(1) (2566) 61--62.

[4] รังสรรค์ ทองสุกนอก, คณิตศาสตร์มัธยมศึกษาตอนต้น (ม.3) พาราโบลา, 2559. https://www.facebook.com/GTRmath/photos/a.1600300720184914/1773414052873579/?type=3.

[5] เตโชวัฒน์ เนื่องจำนงค์, นวพล คงปลื้ม, ชลดา บัวทอง, โครงงานคณิตศาสตร์ เรื่อง การศึกษาการหาปริมาตรทรงกรวยหน้าตัดรูปวงรีที่เกิดจากการตัดกรวยด้วยระนาบที่ทำมุมองศากับฐาน, 2566. https://anyflip.com/kkcpd/qjrs/basic.

[6] H. Anton, I. Bivens, S. Davis, Calculus, 8th ed., John Wiley & Sons, Inc., 2005.

[7] OpenDurian, อนุพันธ์อันดับสูง (high order derivative), (สืบค้น 28 ต.ค. 67). https://www.opendurian.com/learn/high_order_derivative/.

[8] Cuemath, Area Of Ellipse, 2023. https://www.cuemath.com/geometry/area-of-an-ellipse/.

[9] ฉลอง สีแก้วสิ่ว, การทำ Polynomial regression model fitting โดยใช้โปรแกรม Microsoft Excel, (สืบค้น 28 ต.ค. 67). https://www.geocities.ws/chalong_sri/fitpoly.htm.

[10] บ้านและสวน, แบบแปลนห้อง และการวางเฟอร์นิเจอร์ ทำเองได้, 2567. https://www.baanlaesuan.com/276020/ideas/house-ideas/room-plan-sml.

บรรณาธิการ

บรรณาธิการ	รองศาสตราจารย์ ดร.วรานุช แขมมณี
	อาจารย์กาญจนา พานิชการ
บรรณาธิการฝ่ายจัดทำ	ผู้ช่วยศาสตราจารย์ ดร.ธันวา ธีระกาญจน์
	ผู้ช่วยศาสตราจารย์ ดร.วิศรุต โพธิ์อ้น
ผู้ออกแบบปก	ผู้ช่วยศาสตราจารย์ ดร.สุชาดา พงษ์ประเสริฐ

ผู้ทรงคุณวุฒิประเมินบทความวิจัย

ศาสตราจารย์ ดร.วัชรินทร์ วิชิรมาลา	จุฬาลงกรณ์มหาวิทยาลัย
รองศาสตราจารย์ ดร.จริยา อุ่ยยะเสถียร	จุฬาลงกรณ์มหาวิทยาลัย
รองศาสตราจารย์ ดร.สุจินต์ คมฤทัย	จุฬาลงกรณ์มหาวิทยาลัย
ผู้ช่วยศาสตราจารย์ ดร.นิธิ รุ่งธนาภิรมย์	จุฬาลงกรณ์มหาวิทยาลัย
ผู้ช่วยศาสตราจารย์ ดร.จันฒิมา มากมูล	มหาวิทยาลัยเกษตรศาสตร์
รองศาสตราจารย์ ดร.กิตติกร นาคประสิทธิ์	มหาวิทยาลัยขอนแก่น
รองศาสตราจารย์ ดร.นิมิต นิมานะ	มหาวิทยาลัยขอนแก่น
รองศาสตราจารย์ ดร.วิจารณ์ สดศิริ	มหาวิทยาลัยขอนแก่น
ผู้ช่วยศาสตราจารย์ ดร.นวรัตน์ เอกก้านตรง	มหาวิทยาลัยขอนแก่น
ดร.สร้างสรรค์ วรัคคกุล	มหาวิทยาลัยขอนแก่น
รองศาสตราจารย์ ดร.ดวงหทัย กาศวิบูลย์	มหาวิทยาลัยเชียงใหม่
ผู้ช่วยศาสตราจารย์ ดร.เจนสมุทร แสงพันธ์	มหาวิทยาลัยเชียงใหม่
ผู้ช่วยศาสตราจารย์ ดร.ศุภณัฐ ชัยดี	มหาวิทยาลัยเชียงใหม่
ดร.ปิ่นแก้ว ศิริวงศ์	มหาวิทยาลัยทักษิณ
รองศาสตราจารย์ ดร.ทรงพล ศรีวงค์ษา	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
รองศาสตราจารย์ ดร.ปริญญา เสงี่ยมสุนทร	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
รองศาสตราจารย์ ดร.วิราวรรณ ชินวิริยสิทธิ์	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
รองศาสตราจารย์ ดร.เศรษฐภัทร ชินวิริยสิทธิ์	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

ผู้ช่วยศาสตราจารย์ ดร.ดาวุด ทองทา	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
ผู้ช่วยศาสตราจารย์ ดร.ธีระพล สลีวงศ์	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
ดร.ปรินทร์ ไชยปัญญา	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
รองศาสตราจารย์ ดร.ศุภวัชร์ อัศวสัมฤทธิ์	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
รองศาสตราจารย์ ดร.สุรัตนา สังข์หนุน	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
รองศาสตราจารย์ ดร.อรไท พลเสน	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ผู้ช่วยศาสตราจารย์ ดร.จารุณี สุนทรานนท์	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ผู้ช่วยศาสตราจารย์ ดร.ชนนิกานต์ รอดมรณ์	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ผู้ช่วยศาสตราจารย์ ดร.ธนวัฒน์ วิเชียรไพศาล	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ดร.ดอนนี่ พัสสาหรี่	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ผู้ช่วยศาสตราจารย์ ดร.เพ็ญภัค ศิริมาก	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน
ผู้ช่วยศาสตราจารย์ ดร.เจษฎา ตัณฑนุช	มหาวิทยาลัยเทคโนโลยีสุรนารี
ผู้ช่วยศาสตราจารย์ ดร.ธิดารัตน์ อารีรักษ์	มหาวิทยาลัยเทคโนโลยีสุรนารี
ผู้ช่วยศาสตราจารย์ ดร.วรรณวรัตม์ อันล้ำเลิศ	มหาวิทยาลัยธรรมศาสตร์
รองศาสตราจารย์ ดร.นวลักษณ์ ทองจับ	มหาวิทยาลัยธรรมศาสตร์
รองศาสตราจารย์ ดร.ภทรวรรณ แสงนวกิจ	มหาวิทยาลัยธรรมศาสตร์
รองศาสตราจารย์ ดร.มณฑิรา ดวงสาพล	มหาวิทยาลัยธรรมศาสตร์
รองศาสตราจารย์ ดร.วรเวทย์ ลีลาอภิรดี	มหาวิทยาลัยธรรมศาสตร์
รองศาสตราจารย์ ดร.วันหยก อติเศรษฐพงศ์	มหาวิทยาลัยธรรมศาสตร์
รองศาสตราจารย์ศิริจันทร์ เวสารัชศาต	มหาวิทยาลัยธรรมศาสตร์
ผู้ช่วยศาสตราจารย์กานต์ณัฐ ณ บางช้าง	มหาวิทยาลัยธรรมศาสตร์
ผู้ช่วยศาสตราจารย์ ดร.บวร คูหิรัญ	มหาวิทยาลัยธรรมศาสตร์
ดร.รวินทร์ ยังน้อย	มหาวิทยาลัยธรรมศาสตร์
รองศาสตราจารย์ ดร.กิจติ รอดเทศ	มหาวิทยาลัยนเรศวร
ผู้ช่วยศาสตราจารย์ ดร.กัลยา บุญหล้า	มหาวิทยาลัยนเรศวร
ผู้ช่วยศาสตราจารย์ ดร.พรดนัย กาศเกษม	มหาวิทยาลัยนเรศวร
รองศาสตราจารย์ ดร.อรวิชญ์ กุมพล	มหาวิทยาลัยมหาสารคาม
รองศาสตราจารย์ ดร.พรรณนภา ช่างเพ็ชร	มหาวิทยาลัยมหิดล
ผู้ช่วยศาสตราจารย์ ดร.ชัยวัฒน์ มณีสว่าง	มหาวิทยาลัยมหิดล
ผู้ช่วยศาสตราจารย์ ดร.มัณฑนา ชุดทอง	มหาวิทยาลัยมหิดล
ผู้ช่วยศาสตราจารย์ ดร.รู้ธ จ. สกุลคู	มหาวิทยาลัยมหิดล

ดร.ระวี สุวรรณเดโชไชย	มหาวิทยาลัยมหิดล
ผู้ช่วยศาสตราจารย์ ดร.จีรวรรณ พัชรประกิติ	มหาวิทยาลัยแม่โจ้
ผู้ช่วยศาสตราจารย์ ดร.มัลลิกา ราชกิจ	มหาวิทยาลัยแม่โจ้
ผู้ช่วยศาสตราจารย์ ดร.สิตา ชากฤษณ์	มหาวิทยาลัยแม่โจ้
ผู้ช่วยศาสตราจารย์เยาวลักษณ์ คงธรรม	มหาวิทยาลัยแม่โจ้
รองศาสตราจารย์ ดร.ปิยะนุช ศิริวัฒน์	มหาวิทยาลัยแม่ฟ้าหลวง
ผู้ช่วยศาสตราจารย์ ดร.สุภาภรณ์ สดวกดี	มหาวิทยาลัยราชภัฏจันทรเกษม
ผู้ช่วยศาสตราจารย์ ดร.ศุภโชค อิสริยปาลกุล	มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา
ผู้ช่วยศาสตราจารย์ ดร.สุริยัน เขตบรรจง	มหาวิทยาลัยรามคำแหง
รองศาสตราจารย์ ดร.วรานุช แขมมณี	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.ขวัญ เพียซ้าย	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.จิตตินาถ รัตนมุง	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.ชิรา ลำดวนหอม	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.ญานิน กองทิพย์	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.ณหทัย ฤกษ์ฤทัยรัตน์	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.ธันวา ธีระกาญจน์	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.นิสรา สิรสุนทร	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.รัตน์ชัยนันท์ ธรรมสุจริต	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.เรืองวรินท์ สราญรักษ์สกุล	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.วิศรุต โพธิ์อ้น	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.ศญาพัฒน์ สุขใส	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.สุกัญญา หะยีสาและ	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.สุกัญญา อินทรภักดิ์	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.อิทธิเทพ นวาระสุจิตร	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.เอนก จันทรจรูญ	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ศศิวิมล สุขพัฒน์	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.เขมจิรา เที่ยงอยู่	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.จิดาภา ธาดาจรัสศิริ	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.ชนากานต์ จรเสถียร	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.ธีรศักดิ์ ฉลาดการณ์	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.นพดล วิชิตสงคราม	มหาวิทยาลัยศรีนครินทรวิโรฒ

ดร.รุ่งฟ้า จันท์จารุภรณ์	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.เรืองศักดิ์ ตระกูลพุทธิรักษ์	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.เสริมศรี ไทยแท้	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.อังคนา เกาะแก้ว	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.อัจฉริยะ วัธนวิสูตร	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.อุทุมพร มาโต	มหาวิทยาลัยศรีนครินทรวิโรฒ
อาจารย์กาญจนา พานิชการ	มหาวิทยาลัยศรีนครินทรวิโรฒ
ดร.สุภาพ เกิดแสง	มหาวิทยาลัยศิลปากร
ผู้ช่วยศาสตราจารย์ ดร.ภานุพงศ์ วิจิตรคุณากร	มหาวิทยาลัยสงขลานครินทร์
ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐ์ ดิษเจริญ	มหาวิทยาลัยอุบลราชธานี
ผู้ช่วยศาสตราจารย์ ดร.ธารชุดา พันธ์นิกุล	มหาวิทยาลัยอุบลราชธานี
ผู้ช่วยศาสตราจารย์ ดร.สมบัติ สินธุเชาวน์	มหาวิทยาลัยอุบลราชธานี
ดร.ณัฐนันท์ วิจิตรอักษร	สถาบันวิจัยเพื่อการพัฒนาประเทศไทย

คณะกรรมการดำเนินงาน

ผศ.ดร.ณหทัย ฤกษ์ฤทัยรัตน์	ประธานกรรมการ
ผศ.ดร.ธันวา ธีระกาญจน์	รองประธานกรรมการ
ผศ.ดร.จิตตินาถ รัตนมุง	เลขานุการและประธานฝ่ายการเงิน
รศ.ดร.วรานุช แขมมณี	ประธานฝ่ายวิชาการ
ผศ.ดร.พิศุทธวรรณ ศรีภิรมย์ สิรินิลกุล	ประธานฝ่ายสถานที่และอาหาร
ผศ.ดร.วิศรุต โพธิ์อ้น	ประธานฝ่ายการนำเสนอและเอกสารฯ
ผศ.ดร.สุชาดา พงษ์ประเสริฐ	ประธานฝ่ายประชาสัมพันธ์
ผศ.ดร.ฐิตารีย์ รุ่งรัตน์เกษม	ประธานฝ่ายพิธีการ
อ.ดร.อัจฉริยะ วัธนวิสูตร	ประธานฝ่ายลงทะเบียน
อ.ดร.เสริมศรี ไทยแท้	ประธานฝ่ายพัสดุ
อ.กาญจนา พานิชการ	ประธานฝ่ายรับรอง

ผศ.ดร.ขวัญ เพียซ้าย	ผศ.ดร.ชิรา ลำดวนหอม
ผศ.ดร.ญานิน กองทิพย์	ผศ.ดร.นิสรา สิรสุนทร
ผศ.ดร.เรืองวรินท์ อินทรวงษ์ สราญรักษ์สกุล	ผศ.ดร.ศญาพัฒน์ สุขใส
ผศ.ดร.สายัณห์ โสธะโร	ผศ.ดร.สุกัญญา หะยีสาและ
ผศ.ดร.สุกัญญา อินทรภักดิ์	ผศ.ดร.อิทธิเทพ นวาระสุจิตร
ผศ.ดร.เอนก จันทรจรูญ	ผศ.ปัญญวัฒน์ หาอาษา
อ.ดร.เขมจิรา เฒี่ยงอยู่	อ.ดร.จิดาภา ธาดาจรัสศิริ
อ.ดร.ชนากานต์ จรเสถียร	อ.ดร.ธีรศักดิ์ ฉลาดการณ์
อ.ดร.นพดล วิชิตสงคราม	อ.ดร.รุ่งฟ้า จันท์จารุภรณ์
อ.ดร.อังคนา เกาะแก้ว	อ.ดร.อุทุมพร มาโต
นายเกียรติศักดิ์ ภูตะมี	นางสาวฉันทนา บุญโล่ห์
นายณัฐดนัย คำนวณ	นางสาวลัดดา แซ่อึ้ง
นายศิรภพ ประสิทธิ์	นายสมนึก ปุณยกร
นางสาวสุพรรณี เหม็นภัย	นายอิทธิพล โพธิ์ทองคำ

ผู้สนับสนุนการประชุมวิชาการทางคณิตศาสตร์ ครั้งที่ 29 ประจำปี 2568



บริษัท พรูเด็นเชียล ประกันชีวิต (ประเทศไทย) จำกัด (มหาชน)



บริษัท ปตท. จำกัด (มหาชน)



บริษัท ยิบอินซอย จำกัด



บริษัท โอเพ่นเทค จำกัด



บริษัท แลคตาซอย จำกัด