

การเปรียบเทียบตัวแบบ GM(1,1), DGM(1,1), FGM(1,1) และ FDGM(1,1) สำหรับการทำนายจำนวนผู้ป่วยโรคมะเร็งเต้านม

ศวิตา ทองขุนวงศ์^{1*} และ ภัคพล สวัสดิ์กมล²

¹โรงพยาบาลมหาวิทยาลัยเทคโนโลยีสุรนารี มหาวิทยาลัยเทคโนโลยีสุรนารี

²สาขาวิชาคณิตศาสตร์และภูมิสารสนเทศ สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี

^{1,2}111 ถนนมหาวิทยาลัย ตำบลสุรนารี อำเภอเมือง จังหวัดนครราชสีมา 30000

รับบทความ 6 พฤษภาคม 2568 แก้ไขบทความ 2 ตุลาคม 2568 ตอรับบทความ 20 ตุลาคม 2568

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบทำนายข้อมูลอนุกรมเวลาของผู้ป่วยโรคมะเร็งเต้านม จำนวน 4 ตัวแบบ ได้แก่ ตัวแบบระบบเกรย์ GM(1,1) ตัวแบบเกรย์แยกส่วน DGM(1,1) ตัวแบบ FGM(1,1) และตัวแบบ FDGM(1,1) โดยตัวแบบ FGM(1,1) และ FDGM(1,1) ได้รับการปรับปรุงความแม่นยำในการพยากรณ์ด้วยอนุกรมฟูริเยร์ ใช้ข้อมูลรายปีจากสถาบันมะเร็งแห่งชาติในการวิเคราะห์ ประสิทธิภาพของตัวแบบประเมินด้วยค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสัมบูรณ์ (MAPE) ผลการศึกษาพบว่า ตัวแบบ FDGM(1,1) ให้ประสิทธิภาพสูงที่สุด โดยมีค่า MAPE เท่ากับ 15.25% ขณะที่ตัวแบบ GM(1,1), DGM(1,1) และ FGM(1,1) ให้ค่า MAPE เท่ากับ 17.78%, 17.41% และ 15.49% ตามลำดับ ทั้งนี้ เกณฑ์ของสำนักงานสถิติแห่งชาติระบุว่า หากค่า MAPE อยู่ระหว่าง 10%–20% ถือว่าตัวแบบมีความแม่นยำในระดับค่อนข้างดี ดังนั้น สำหรับการศึกษาครั้งนี้จึงสรุปได้ว่า ตัวแบบ FDGM(1,1) มีความเหมาะสมที่สุดในการใช้ทำนายจำนวนผู้ป่วยโรคมะเร็งเต้านม

คำสำคัญ : ตัวแบบเกรย์; อนุกรมฟูริเยร์; ตัวแบบ FDGM(1,1)

Comparison of GM(1,1), DGM(1,1), FGM(1,1), and FDGM(1,1) Models for Forecasting Breast Cancer Cases

Sawita Thongkunwong^{1*} and Phakkhaphon Sawatkamon²

¹Suranaree University of Technology Hospital, Suranaree University of Technology

²School of Mathematical Sciences and Geoinformatics, Institute of Science, Suranaree University of Technology

^{1,2}111 University Avenue, Suranaree Subdistrict, Muang District, Nakhon Ratchasima, 30000

Received 6 May 2025; Revised 2 October 2025; Accepted 20 October 2025

Abstract

The objective of this study is to compare the forecasting performance of four time series models for predicting the number of breast cancer patients. The models considered include the Grey Model GM(1,1), the Discrete Grey Model DGM(1,1), the Fourier Grey Model FGM(1,1), and the Fourier Discrete Grey Model FDGM(1,1). The FGM(1,1) and FDGM(1,1) models were enhanced using Fourier series to improve forecasting accuracy. Annual data from the National Cancer Institute were used for the analysis. Model performance was evaluated using the Mean Absolute Percentage Error (MAPE). The results show that the FDGM(1,1) model achieved the highest accuracy, with a MAPE value of 15.25%, whereas the GM(1,1), DGM(1,1), and FGM(1,1) models yielded MAPE values of 17.78%, 17.41%, and 15.49%, respectively. According to the National Statistical Office's criteria, a MAPE between 10% and 20% indicates a reasonably accurate forecasting model. Therefore, it can be concluded that, for this study, the FDGM(1,1) model is the most appropriate choice for forecasting the number of breast cancer patients.

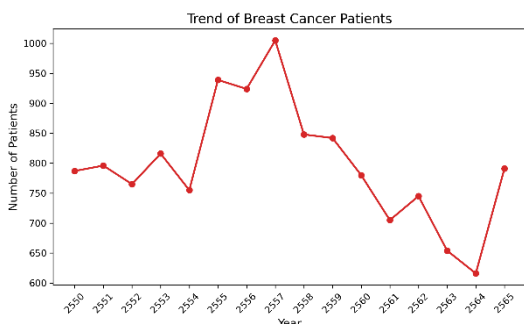
Keywords : Grey models; Fourier series; FDGM(1,1) model

* Corresponding Author. Tel.: +666 3795 4249, E-mail Address: sawiphak2424@gmail.com

1. บทนำ

ข้อมูลจากสถาบันมะเร็งแห่งชาติ พ.ศ. 2564 ระบุว่า มะเร็งที่พบมากที่สุดในเพศหญิงเป็น 3 อันดับแรก ได้แก่ มะเร็งเต้านม มะเร็งปากมดลูกและมะเร็งลำไส้ โดยมะเร็งเต้านมเป็นชนิดที่มีอุบัติการณ์มากที่สุดในผู้หญิงไทย [1] ทั้งนี้ มะเร็งเต้านมสามารถเกิดขึ้นในเพศชายได้ในอัตราส่วนประมาณ 1 ใน 100 ของผู้ป่วยมะเร็งเต้านมทั้งหมด [2] ในระดับสากล องค์การอนามัยโลก (WHO) รายงานว่ามีผู้ป่วยมะเร็งเต้านมรายใหม่ทั่วโลกราว 2.3 ล้านคน และมีอัตราการเสียชีวิตราวประมาณ 685,000 คนต่อปี [3] ปัจจัยเสี่ยงที่สัมพันธ์กับการเกิดโรคมะเร็งเต้านม ได้แก่ อายุที่มากขึ้น พันธุกรรม ภาวะอ้วน มีประจำเดือนครั้งแรกก่อนอายุ 12 ปี มีประวัติเคยเป็นมะเร็งเต้านม มีบุตรคนแรกมากกว่าอายุ 30 ปี การได้รับฮอร์โมนเอสโตรเจนเป็นเวลานาน [4] อย่างไรก็ตาม โรคมะเร็งเต้านมหากได้รับการวินิจฉัยและรักษาในระยะแรกเริ่ม จะสามารถรักษาให้มีโอกาสรอดชีวิตสูงขึ้นและช่วยลดอัตราการเสียชีวิตลงได้ [2]

การคาดการณ์จำนวนผู้ป่วยโรคมะเร็งเต้านมในอนาคตเป็นเครื่องมือสำคัญที่ช่วยให้หน่วยงานสาธารณสุขสามารถกำหนดแนวทางป้องกัน วางแผนการรักษา และบริหารจัดการทรัพยากรทางการแพทย์ได้อย่างมีประสิทธิภาพ เพื่อรองรับการรักษาผู้ป่วยที่เกิดขึ้นในอนาคต



รูปที่ 1 แสดงจำนวนผู้ป่วยโรคมะเร็งเต้านมจากสถาบันมะเร็งแห่งชาติช่วงปี พ.ศ. 2550-2565

ข้อมูลแสดงให้เห็นว่า จำนวนผู้ป่วยมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่องในช่วงปี 2550-2557 โดยมีจุดสูงสุดที่ 1,005 รายในปี 2557 หลังจากนั้นจำนวนผู้ป่วยมีความผันผวนและลดลงอย่างค่อยเป็นค่อยไป โดยเฉพาะในช่วงปี 2563-2564 ที่มีจำนวนผู้ป่วยลดลงอย่างเห็นได้ชัด ซึ่งสอดคล้องกับช่วงการระบาดของโรคโควิด-19 ที่ส่งผลให้ประชาชนจำนวนมากไม่สะดวกในการเข้ารับการตรวจคัดกรองและการวินิจฉัย ส่งผลให้ตัวเลขผู้ป่วยรายใหม่ที่ถูกรายงานลดลงชั่วคราว ลักษณะของข้อมูลที่มีความผันผวนในบางช่วงและไม่ได้เป็นเส้นแนวโน้มที่ราบเรียบ แสดงให้เห็นว่า ข้อมูลนี้มีความไม่แน่นอนและไม่เป็นเส้นตรงชัดเจน จึงเหมาะสมกับการใช้แบบจำลองระบบเกรย์ (Grey Models) ซึ่งสามารถประมวลผลข้อมูลจำนวนน้อยและข้อมูลที่มีความไม่แน่นอนได้อย่างยืดหยุ่น ตรงกันข้ามกับแบบจำลองเชิงสถิติแบบพื้นฐาน เช่น Moving Average หรือ ARIMA ที่มีข้อสมมติว่าข้อมูลควรมีความนิ่ง (stationary) และมีจำนวนตัวอย่างมากพอ ซึ่งอาจไม่เหมาะกับชุดข้อมูลที่มีความผันผวนและขนาดเล็กเช่นนี้

ทฤษฎีระบบเกรย์ถูกสร้างขึ้นโดยนักคณิตศาสตร์ชาวจีน J. Deng ในปี ค.ศ. 1982 มีจุดเด่นในการนำไปใช้กับข้อมูลที่มีขนาดเล็กหรือไม่สมบูรณ์ จึงเหมาะสมอย่างยิ่งกับปัญหาที่มีข้อมูลจำกัด ทฤษฎีระบบเกรย์ถูกประยุกต์ใช้ในหลายสาขา เช่น การแพทย์ วิศวกรรม และการเกษตร เป็นต้น [5] ตัวอย่างงานวิจัยที่ประยุกต์ใช้ทฤษฎีระบบเกรย์ เช่น การนำตัวแบบทฤษฎีระบบเกรย์ไปทำนายจำนวนผู้เสียชีวิตจากโรคปอดบวมและโรคเลือดออกในสมอง ผลการศึกษาพบว่าตัวแบบทำนายจำนวนผู้เสียชีวิตจากโรคปอดบวมและโรคเลือดออกในสมองให้ค่า MAPE เท่ากับ 2.83% และ 1.20% ตามลำดับ [6] การนำตัวแบบ GM(1,1) ไปพยากรณ์ค่าฝุ่น PM10 ในบริเวณภาคเหนือตอนบน ผลการพยากรณ์พบว่าในช่วงปี 2557 – 2560 ฝุ่น PM10

มีแนวโน้มที่จะมีปริมาณเพิ่มขึ้น [7] อีกทั้งยังมีการเปรียบเทียบประสิทธิภาพระหว่างแบบจำลอง GM(1,1) และแบบจำลองเกรย์แยกส่วน DGM(1,1) ในการพยากรณ์วิกฤตการณ์ทางเศรษฐกิจของประเทศไทย ซึ่งพบว่า DGM(1,1) ให้ผลแม่นยำมากกว่า [8] นอกจากนี้ยังมีการปรับปรุงความแม่นยำของตัวแบบด้วยการใช้ออนุกรมฟูริเยร์ โดยพบว่า ตัวแบบ FDGM(1,1) ซึ่งเป็นการประยุกต์ออนุกรมฟูริเยร์กับ DGM(1,1) ให้ค่า MAPE ต่ำกว่าตัวแบบ GM(1,1) [9] เป็นต้น ตัวแบบทฤษฎีระบบเกรย์นั้น เป็นตัวแบบที่ได้รับความนิยมเนื่องจากมีวิธีการคำนวณพารามิเตอร์ที่ไม่ซับซ้อน และสามารถใช้ได้กับจำนวนข้อมูลที่มีขนาดเล็กหรือข้อมูลตั้งแต่ 4 ข้อมูลขึ้นไป [10]

ดังนั้น งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาตัวแบบการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งเต้านมโดยประยุกต์ใช้ทฤษฎีระบบเกรย์ ผ่านการสร้างแบบจำลอง GM(1,1) และ DGM(1,1) พร้อมทั้งปรับปรุงความแม่นยำของการพยากรณ์ด้วยออนุกรมฟูริเยร์ เพื่อสร้างแบบจำลอง FGM(1,1) และ FDGM(1,1) จากนั้นทำการประเมินและเปรียบเทียบประสิทธิภาพของแบบจำลองทั้งสี่ประเภท โดยใช้ค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error: MAPE) เป็นเกณฑ์ในการประเมินผล

2. ระเบียบวิธีวิจัย

2.1 ขั้นตอนการวิจัย

ขั้นตอนที่ 1 ทำการเก็บข้อมูลจำนวนผู้ป่วยโรคมะเร็งเต้านม จากสถาบันมะเร็งแห่งชาติ [11] ซึ่งเป็นข้อมูลรายปีตั้งแต่ปี พ.ศ. 2550 ถึง พ.ศ. 2565 รวมทั้งสิ้น 16 ปี โดยเป็นข้อมูลผู้ป่วยรายใหม่ของแต่ละปีที่เข้ารับการรักษาที่สถานมะเร็งแห่งชาติ

ขั้นตอนที่ 2 แบ่งข้อมูลออกเป็น 2 ส่วนเพื่อใช้ในการตรวจสอบความแม่นยำของตัวแบบ (Validation) คือ ข้อมูลชุดฝึกหัดและข้อมูลชุดทดสอบในอัตราส่วน 8 : 2 นั่นคือ ข้อมูลชุดฝึกหัดจำนวน 13 ปี และ ข้อมูลชุดทดสอบจำนวน 3 ปี

ขั้นตอนที่ 3 สร้างตัวแบบ GM(1,1) และ DGM(1,1) ภายใต้การใช้ภาษา Python ผ่านไลบรารี numpy

ขั้นตอนที่ 4 ปรับปรุงค่าพยากรณ์ของตัวแบบด้วยการประยุกต์ใช้ออนุกรมฟูริเยร์ (Fourier Series) ซึ่งจะได้ตัวแบบ FGM(1,1) และ FDGM(1,1)

ขั้นตอนที่ 5 เปรียบเทียบประสิทธิภาพตัวแบบทั้งหมดโดยใช้ค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสัมบูรณ์ (MAPE) กับข้อมูลชุดทดสอบ

2.2 ตัวแบบระบบเกรย์ GM(1,1)

ตัวแบบ GM(1,1) เป็นแบบจำลองที่ได้รับการพัฒนาจาก J.Deng ในปี 1982 คำว่า “เกรย์” หมายถึงระบบข้อมูลที่ไม่สมบูรณ์หรือไม่แน่นอน ซึ่งขั้นตอนการสร้างตัวแบบ GM(1,1) มีดังนี้ [12]

1) สร้างลำดับของข้อมูลตั้งต้น (Original Data) หรือ ข้อมูลจริงจำนวน n ข้อมูล

$$X^{(0)} = [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)]$$

2) สร้างลำดับข้อมูลสะสมของข้อมูลตั้งต้น

$$X^{(1)} = [x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)]$$

$$\text{โดยที่ } x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$$

3) ประมาณค่าพารามิเตอร์ a, b จากสมการเชิงอนุพันธ์ของตัวแบบ GM(1,1)

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b \quad (1)$$

เมื่อ

$x^{(1)}(t)$ = ข้อมูลที่ผ่านการสะสมครั้งที่ 1 (1-AGO)

ณ เวลา t

a = ค่าสัมประสิทธิ์การหน่วง (Development Coefficient)

b = ค่าอินพุตเกรย์ (Grey input)

ผลเฉลยของสมการที่ (1) คือ

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (2)$$

$$\text{เมื่อ } z^{(1)}(k+1) = \frac{x^{(1)}(k) + x^{(1)}(k+1)}{2}$$

โดยที่

$$k = 1, 2, \dots, n-1$$

การประมาณค่าพารามิเตอร์จะใช้วิธีกำลังสองน้อยสุด ซึ่งคำนวณได้จาก

$$\begin{bmatrix} a \\ b \end{bmatrix}_{2 \times 1} = (B^T B)^{-1} B^T Y \quad (3)$$

$$\text{เมื่อ } B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}_{(n-1) \times 2}$$

$$\text{และ } Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}_{(n-1) \times 1}$$

4) คำนวณค่าทำนายสะสมของตัวแบบ GM(1,1)

จากสมการ

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a} \quad (4)$$

โดยที่

$$\hat{x}^{(1)}(1) = x^{(0)}(1) \text{ และ } k = 1, 2, \dots, n-1$$

เมื่อต้องการค่าทำนาย ณ เวลา $k+1$ จะคำนวณโดย

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \quad (5)$$

ลักษณะสำคัญประการหนึ่งของแบบจำลอง GM(1,1) คือความเหมาะสมในการพยากรณ์ด้วยข้อมูลที่มีจำนวนจำกัด โดยสามารถใช้ข้อมูลเพียง 4-5 จุดได้ [13] แบบจำลองนี้ช่วยลดความไม่แน่นอนโดยการแปลงข้อมูลดั้งเดิมให้เป็นลำดับสะสม ซึ่งช่วยเผยให้เห็นรูปแบบที่ซ่อนอยู่ในข้อมูล และตั้งอยู่บนสมมติฐานของการเติบโตแบบเอ็กซ์โปเนนเชียล จึงเหมาะสำหรับข้อมูลที่มีแนวโน้มชัดเจน อย่างไรก็ตาม ความแม่นยำของแบบจำลองจะลดลงอย่างมากหากข้อมูลไม่เป็นไปตามสมมติฐานการเติบโตแบบเอ็กซ์โปเนนเชียลดังกล่าว [14]

2.2 ตัวแบบเกรย์แยกส่วน DGM(1,1)

DGM(1,1) หรือแบบจำลองเกรย์ไม่ต่อเนื่อง (Discrete Grey Model (1,1)) เป็นแบบจำลองที่ได้รับการพัฒนาต่อยอดจาก GM(1,1) แบบดั้งเดิม โดยมีวัตถุประสงค์เพื่อเพิ่มความแม่นยำในการพยากรณ์ ด้วยการแทนที่สมการเชิงอนุพันธ์ใน GM(1,1) ด้วยสมการเชิงผลต่าง ซึ่งเหมาะสมกับข้อมูลที่เป็นลำดับเวลาแบบไม่ต่อเนื่องมากกว่าซึ่งมีงานวิจัยที่ระบุว่า DGM(1,1) สามารถให้ผลการพยากรณ์ที่แม่นยำกว่า GM(1,1) โดยเฉพาะในกรณีที่ข้อมูลมีการเปลี่ยนแปลงไม่สม่ำเสมอ [8] โดยที่ตัวแบบ DGM(1,1) มีสมการดังต่อไปนี้

$$x^{(1)}(k+1) = \beta_1 x^{(1)}(k) + \beta_2 \quad (6)$$

เมื่อ β_1, β_2 คือ พารามิเตอร์ของตัวแบบการประมาณค่าพารามิเตอร์จะกระทำโดยใช้วิธีกำลังสองน้อยสุดซึ่งสามารถคำนวณได้จาก

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}_{2 \times 1} = (M^T M)^{-1} M^T P \quad (7)$$

$$\text{เมื่อ } M = \begin{bmatrix} x^{(1)}(1) & 1 \\ x^{(1)}(2) & 1 \\ \vdots & \vdots \\ x^{(1)}(n-1) & 1 \end{bmatrix}_{(n-1) \times 2}$$

$$\text{และ } P = \begin{bmatrix} x^{(1)}(2) \\ x^{(1)}(3) \\ \vdots \\ x^{(1)}(n) \end{bmatrix}_{2 \times (n-1)}$$

การหาค่าทำนายสะสมของ DGM(1,1) ณ เวลา $k+1$ สามารถคำนวณจากสมการ

$$\hat{x}^{(1)}(k+1) = \beta_1^k x^{(0)}(1) + \frac{1-\beta_1^k}{1-\beta_1} \times \beta_2 \quad (8)$$

โดยที่ $\hat{x}^{(1)}(1) = x^{(0)}(1)$ และ $k=1, 2, \dots, n-1$ เมื่อต้องการหาค่าทำนาย ณ เวลา $k+1$ สามารถหาได้จากสมการ

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \quad (9)$$

แบบจำลอง DGM(1,1) มีข้อดีหลายประการ โดยช่วยลดข้อผิดพลาดในการพยากรณ์ที่มักพบในแบบจำลอง GM(1,1) โดยเฉพาะในกรณีที่มีข้อมูลไม่เป็นไปตามสมมติฐานการเติบโตแบบเอ็กซ์โปเนนเชียลแบบจำลองนี้ยังมีความเรียบง่ายในการคำนวณ เนื่องจากใช้สมการเชิงผลต่างแทนการอินทิเกรต นอกจากนี้

DGM(1,1) ยังเหมาะสมสำหรับชุดข้อมูลขนาดเล็ก โดยปกติมีจำนวนข้อมูลตั้งแต่ 4 ถึง 10 จุดเท่านั้น [15]

2.3 การปรับปรุงค่าคลาดเคลื่อนด้วยอนุกรมฟูริเยร์

การปรับปรุงค่าคลาดเคลื่อนด้วยอนุกรมฟูริเยร์ของตัวแบบ GM(1,1) และ DGM(1,1) นั้น [9] จะมีขั้นตอนดังต่อไปนี้

1) คำนวณค่าคลาดเคลื่อน ณ เวลาที่ k

$$\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \quad (10)$$

เมื่อ $k=2, 3, \dots, n$ จากนั้นนำมาสร้างเป็นเมทริกซ์ของค่าคลาดเคลื่อนดังนี้

$$\varepsilon = [\varepsilon(2) \quad \varepsilon(3) \quad \dots \quad \varepsilon(n)]_{1 \times (n-1)}^T$$

2) การประมาณค่าคลาดเคลื่อนที่ k ด้วยอนุกรมฟูริเยร์ สามารถประมาณได้โดย

$$\begin{aligned} \hat{\varepsilon}(k) = & \frac{1}{2} a_0 + \sum_{i=1}^Z \left[a_i \cos\left(\frac{2\pi i}{n-1}(k)\right) \right] \\ & + \sum_{i=1}^Z \left[b_i \sin\left(\frac{2\pi i}{n-1}(k)\right) \right] \end{aligned} \quad (11)$$

$$\text{เมื่อ } Z = \left\lfloor \frac{n-1}{2} \right\rfloor - 1$$

ซึ่ง เมทริกซ์ค่าคลาดเคลื่อนสามารถเขียนอยู่ในรูปของสมการ

$$\varepsilon = Q \times C \quad (12)$$

เมื่อ

$$Q = \begin{bmatrix} \frac{1}{2} & \cos\left(\frac{2\pi z}{n-1} \cdot 2\right) & \dots & \sin\left(\frac{2\pi z}{n-1} \cdot 2\right) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} & \cos\left(\frac{2\pi z}{n-1} \cdot n\right) & \dots & \sin\left(\frac{2\pi z}{n-1} \cdot n\right) \end{bmatrix}_{(n-1) \times (2Z+1)}$$

และ $C = [a_0 \ a_1 \ b_1 \ \cdots \ a_z \ b_z]^T_{1 \times (2Z+1)}$

3) ประมาณค่าพารามิเตอร์ของเมทริกซ์ C ใช้วิธีกำลังสองน้อยสุด ซึ่งคำนวณจาก

$$C = (Q^T Q)^{-1} Q^T \varepsilon \quad (13)$$

จากสมการที่ (13) เมื่อได้ค่าพารามิเตอร์ก็นำไปแทนในสมการที่ (11) เพื่อคำนวณค่าคลาดเคลื่อน ณ เวลาที่ k

4) ปรับค่าทำนายของตัวแบบ GM(1,1) และ DGM(1,1) ดังนี้

$$\hat{f}_k = \begin{cases} \hat{f}_1 = x^{(0)}(1) \\ \hat{f}_k = \hat{x}^{(0)}(k) + \hat{\varepsilon}(k); k = 2, 3, \dots, n \end{cases} \quad (14)$$

หลังจากปรับค่าคลาดเคลื่อนก็จะทำให้ได้ตัวแบบ FGM(1,1) และ FDGM(1,1) การแก้ไขค่าคลาดเคลื่อนในแบบจำลอง GM(1,1) และ DGM(1,1) ด้วยอนุกรมฟูริเยร์ช่วยเพิ่มความแม่นยำในการพยากรณ์ โดยเฉพาะอย่างยิ่งในกรณีที่มีข้อมูลแสดงลักษณะตามฤดูกาล หรือมีรูปแบบคลื่นซ้ำ ๆ ซึ่งแบบจำลองเกรย์ทั่วไปไม่สามารถจับลักษณะเหล่านี้ได้ เทคนิคนี้ช่วยให้สามารถตรวจจับรูปแบบฤดูกาลที่ซ่อนอยู่ ลดสัญญาณรบกวน เพิ่มความแม่นยำของการพยากรณ์ และยังสามารถประยุกต์ใช้ได้ทั้งกับ GM(1,1) และ DGM(1,1) อย่างไรก็ตาม การใช้งานอนุกรมฟูริเยร์อย่างมีประสิทธิภาพจำเป็นต้องมีข้อมูลที่เพียงพอเพื่อจับรูปแบบความถี่ต่าง ๆ ของแบบจำลองได้อย่างแม่นยำ

2.4 เกณฑ์ประเมินประสิทธิภาพของตัวแบบ

ในงานวิจัยนี้ใช้ค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Percentage Error, MAPE) เพื่อทดสอบความแม่นยำของแบบจำลอง โดยที่ตัวแบบที่ดีจะมีค่า MAPE ที่ต่ำ ซึ่งสามารถคำนวณจากสมการ [16]

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i} \times 100 \quad (15)$$

เมื่อ

y_t แทน อนุกรมเวลา ณ เวลา t

$\hat{y}_t$ แทน ค่าพยากรณ์ของอนุกรมเวลา ณ เวลา t

n แทน จำนวนข้อมูลที่ศึกษา

และได้มีเกณฑ์ที่ใช้ในการกำหนดความแม่นยำสำหรับการพยากรณ์ดังนี้ [17]

ตารางที่ 1 เกณฑ์การประเมินความแม่นยำของตัวแบบ

MAPE	ผลการประเมินความแม่นยำ
MAPE < 10%	แม่นยำสูง
10% < MAPE < 20%	ค่อนข้างแม่นยำ
20% < MAPE < 50%	พอใช้
MAPE > 50%	ไม่มีความแม่นยำ

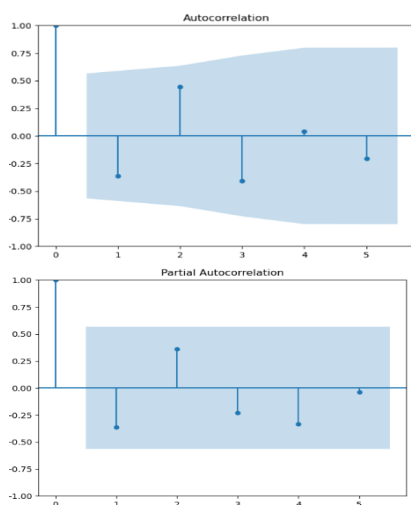
3. ผลการศึกษาและอภิปรายผล

จากการเก็บรวบรวมจำนวนผู้ป่วยโรคมะเร็งเต้านมเป็นข้อมูลรายปี ตั้งแต่ปี พ.ศ. 2550 ถึง พ.ศ. 2565 จากสถาบันมะเร็งแห่งชาติ รวมทั้งสิ้น 16 ปี ข้อมูลดังกล่าวถูกแบ่งออกเป็น 2 ชุด ได้แก่ ชุดฝึกหัด (Training set) และ ชุดทดสอบ (Testing set) ในอัตราส่วน 80 : 20 ซึ่งเป็นแนวทางที่เหมาะสมและได้รับการยอมรับกันอย่างแพร่หลายในการสร้างแบบจำลอง [18] โดยชุดฝึกหัดประกอบด้วยข้อมูล 13 ปี และชุดทดสอบจำนวน 3 ปี

ตารางที่ 2 แสดงจำนวนผู้ป่วยโรคมะเร็งเต้านมที่ใช้เป็นข้อมูลสำหรับการฝึกแบบจำลอง

ปี	จำนวนผู้ป่วย	ปี	จำนวนผู้ป่วย
2550	787	2551	796
2552	765	2553	816
2554	755	2555	939
2556	924	2557	1005
2558	848	2559	842
2560	780	2561	705
2562	745		

วิเคราะห์ความนิ่งของข้อมูลอนุกรมเวลาจำนวนผู้ป่วยมะเร็งเต้านม โดยใช้การทดสอบ Augmented Dickey–Fuller (ADF) เพื่อประเมินความเหมาะสมของการนำแบบจำลอง ARIMA มาใช้ในการเปรียบเทียบประสิทธิภาพกับแบบจำลองเกรย์ ผลการทดสอบเบื้องต้นพบว่าข้อมูลจริงมีค่า p-value เท่ากับ 0.09 ซึ่งมากกว่า 0.05 บ่งชี้ว่าข้อมูลยังไม่เป็นแบบนิ่ง (non-stationary) จึงได้ทำการหาความต่างอันดับที่ 1 เพื่อลดแนวโน้มของข้อมูล ผลการทดสอบภายหลังพบว่า p-value เท่ากับ 0.0003 ซึ่งน้อยกว่า 0.05 แสดงว่าข้อมูลหลังการแปลงมีลักษณะเป็นแบบนิ่ง (stationary) จากนั้นจึงได้วิเคราะห์กราฟ ACF และ PACF เพื่อช่วยในการระบุโครงสร้างของโมเดล ARIMA ดังแสดงในรูปที่ 2



รูปที่ 2 การวิเคราะห์กราฟ ACF และ PACF

จากการวิเคราะห์ข้อมูลหลังการหาผลต่างอันดับที่ 1 ด้วยกราฟฟังก์ชันสหสัมพันธ์ (ACF) และฟังก์ชันสหสัมพันธ์ย่อย (PACF) พบว่าข้อมูลมีลักษณะของความสัมพันธ์เชิงเวลาทั้งส่วนของกระบวนการ AR และ MA มีบทบาทร่วมกันอย่างชัดเจน โดยสังเกตได้จากค่าสหสัมพันธ์ที่โดดเด่นอย่างมีนัยสำคัญในหลายช่วงเวลา และไม่มีจุดที่กราฟตัดลงอย่างชัดเจนหลัง lag 1 หรือ 2 ดังนั้นจึงเลือกใช้แบบจำลอง ARIMA(2,1,2)

ซึ่งสามารถสะท้อนโครงสร้างของข้อมูลที่มีทั้งความสัมพันธ์เชิงเวลาและความผันผวนสะสมได้ดี เพื่อใช้เป็นแบบจำลองเชิงสถิติสำหรับเปรียบเทียบประสิทธิภาพกับแบบจำลองเกรย์ ผลการทำนายแสดงดังตารางที่ 3

ตารางที่ 3 แสดงผลการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งเต้านมของตัวแบบ ARIMA(2,1,2) และจำนวนผู้ป่วยจริง

ปี	จำนวนผู้ป่วยจริง	ตัวแบบ ARIMA(2,1,2)
2563	654	699
2564	616	771
2565	791	703
MAPE (%)		15.92 %

ขั้นตอนต่อไปนำชุดข้อมูลฝึกหัดมาสร้างแบบจำลอง GM(1,1) และ DGM(1,1) เพื่อใช้ในการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งเต้านม โดยดำเนินการด้วยภาษา Python และไลบรารี numpy ในการประมวลผล แบบจำลองแต่ละแบบถูกนำไปใช้พยากรณ์ข้อมูลในช่วงปี พ.ศ. 2563 ถึง พ.ศ. 2565 และเปรียบเทียบผลการพยากรณ์กับข้อมูลจริงในช่วงเวลาเดียวกัน เพื่อเป็นการประเมินความแม่นยำและประสิทธิภาพของแบบจำลอง ผลการพยากรณ์แสดงในตารางที่ 4

ตารางที่ 4 แสดงผลการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งเต้านมของตัวแบบ GM(1,1) ตัวแบบ DGM(1,1) และ จำนวนผู้ป่วยจริง

ปี	จำนวนผู้ป่วยจริง	ตัวแบบ GM(1,1)	ตัวแบบ DGM(1,1)
2563	654	803	801
2564	616	800	797
2565	791	796	793
MAPE (%)		17.78 %	17.41 %

จากการเปรียบเทียบผลการพยากรณ์ของแบบจำลอง GM(1,1) และ DGM(1,1) กับข้อมูลจริงในตารางที่ 4 พบว่าทั้งสองแบบจำลองมีค่าคลาดเคลื่อนที่ค่อนข้างสูงในช่วงปี พ.ศ. 2563 และปี พ.ศ. 2564 ดังนั้น จึงดำเนินการนำค่าคลาดเคลื่อนของแบบจำลองเดิมทั้งสองมาปรับด้วยอนุกรมฟูริเยร์ (Fourier Series) เพื่อเป็นการเพิ่มความแม่นยำในการพยากรณ์ ส่งผลให้ได้แบบจำลองใหม่ คือ FGM(1,1) และ FDGM(1,1)

ผลการพยากรณ์ของแบบจำลองหลังปรับค่าคลาดเคลื่อน พร้อมเปรียบเทียบกับข้อมูลจริง แสดงไว้ในตารางที่ 5 ซึ่งใช้เพื่อประเมินประสิทธิภาพของการปรับปรุงแบบจำลองด้วยอนุกรมฟูริเยร์

ตารางที่ 5 แสดงผลการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งเต้านมของตัวแบบ FGM(1,1) ตัวแบบ FDGM(1,1) และ จำนวนผู้ป่วยจริง

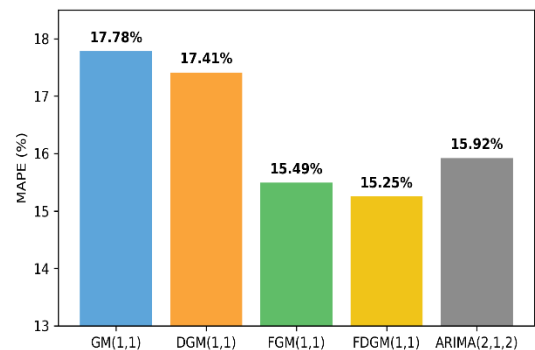
ปี	จำนวนผู้ป่วยจริง	ตัวแบบ FGM(1,1)	ตัวแบบ FDGM(1,1)
2563	654	725	721
2564	616	776	773
2565	791	716	712
MAPE (%)		15.49 %	15.25 %

จากตารางที่ 5 พบว่าแบบจำลอง FGM(1,1) และ FDGM(1,1) สามารถพยากรณ์จำนวนผู้ป่วยโรคมะเร็งเต้านมได้ใกล้เคียงกับข้อมูลจริงโดยรวม อย่างไรก็ตาม ในปี พ.ศ. 2563 และ 2564 ผลการพยากรณ์มีค่าสูงกว่าจำนวนผู้ป่วยจริง ในขณะที่ปี พ.ศ. 2565 ผลการพยากรณ์กลับมีค่าต่ำกว่าจำนวนผู้ป่วยจริง ซึ่งความคลาดเคลื่อนดังกล่าวอาจมีสาเหตุมาจากสถานการณ์การแพร่ระบาดของโรค COVID-19 ที่ส่งผลให้จำนวนผู้ป่วยไม่ได้เข้ารับการตรวจวินิจฉัยหรือรักษาอย่างครบถ้วน เนื่องจากมาตรการควบคุมการแพร่ระบาดของโรคในช่วงเวลาดังกล่าว แต่เมื่อเข้าสู่ปี พ.ศ. 2565 สถานการณ์เริ่มกลับเข้าสู่ภาวะปกติ ทำให้ประชาชนสามารถเข้าถึงบริการทางการแพทย์ได้มากขึ้น

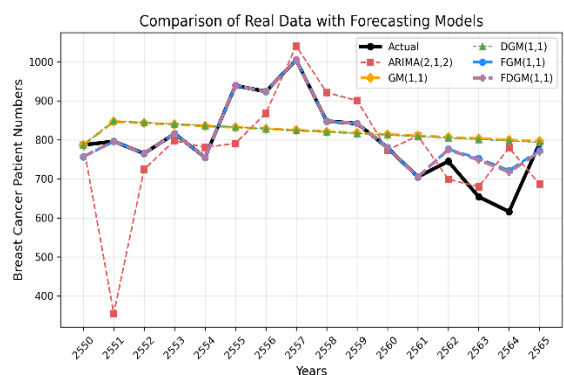
ส่งผลให้จำนวนผู้ป่วยที่ได้รับการตรวจวินิจฉัยเพิ่มขึ้นตามจริง ซึ่งลักษณะการเปลี่ยนแปลงดังกล่าวเป็นความผันผวนที่แบบจำลองไม่สามารถสะท้อนได้อย่างแม่นยำ

การเปรียบเทียบประสิทธิภาพของแบบจำลองทั้งสี่กระทำโดยใช้ค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสัมบูรณ์ (MAPE) แสดงในรูปที่ 3

ผลการเปรียบเทียบในรูปที่ 3 แสดงให้เห็นว่าแบบจำลองทั้งสี่แบบมีค่า MAPE อยู่ในช่วง 10% ถึง 20% ซึ่งจัดอยู่ในระดับความแม่นยำค่อนข้างสูงตามเกณฑ์ที่ระบุในตารางที่ 1 โดยแบบจำลอง FDGM(1,1) มีค่า MAPE ต่ำที่สุด สะท้อนถึงประสิทธิภาพในการพยากรณ์ที่เหนือกว่าแบบจำลองอื่น



รูปที่ 3 แสดงค่า MAPE ของตัวแบบพยากรณ์



รูปที่ 4 กราฟแสดงการเปรียบเทียบข้อมูลจริงและตัวแบบพยากรณ์

ผลการศึกษานี้สอดคล้องกับงานวิจัยของ Wang et al. [19] ที่ใช้แบบจำลอง GM(1,1) ในการพยากรณ์

จำนวนผู้ป่วยโรคตับอักเสบบีในประเทศจีน และพบว่าแบบจำลอง GM(1,1) เหมาะกับข้อมูลขนาดเล็กและมีลักษณะข้อมูลที่ไม่ต่อเนื่องหรือไม่สมบูรณ์ อย่างไรก็ตามแบบจำลองดังกล่าวมีข้อจำกัดในการตอบสนองต่อความผันผวนของข้อมูล โดยเฉพาะในช่วงที่มีเหตุการณ์ที่ไม่ปกติ เช่น การระบาดของโรค ขณะเดียวกัน งานของ Zhao et al. [20] ได้เปรียบเทียบประสิทธิภาพของแบบจำลอง GM(1,1), ARIMA และ LSTM ในการพยากรณ์จำนวนผู้ป่วยวัณโรคในจีน พบว่าแบบจำลอง ARIMA และ LSTM ให้ผลลัพธ์ที่แม่นยำกว่าเมื่อข้อมูลมีลักษณะต่อเนื่องและมีปริมาณข้อมูลที่เพียงพอ ในขณะที่ GM(1,1) มีข้อได้เปรียบในกรณีข้อมูลมีจำนวนน้อยและไม่สมบูรณ์ ซึ่งสอดคล้องกับลักษณะของข้อมูลในงานวิจัยนี้ ดังนั้นงานวิจัยนี้จึงเลือกใช้ GM(1,1) และ DGM(1,1) พร้อมทั้งปรับค่าคลาดเคลื่อนด้วยอนุกรมฟูรีเยร์เพื่อเพิ่มความแม่นยำในการพยากรณ์กับลักษณะของข้อมูลที่มีจำกัด

4. สรุป

สำหรับงานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบที่ใช้ทำนายจำนวนผู้ป่วยโรคมะเร็งเต้านม โดยใช้ข้อมูลจากสถาบันมะเร็งแห่งชาติจำนวนทั้งสิ้น 16 ปี ตั้งแต่ พ.ศ. 2550 ถึง พ.ศ. 2565 ผลการศึกษาแสดงให้เห็นว่าตัวแบบทั้งสี่ตัวแบบ ได้แก่ GM(1,1), DGM(1,1), FGM(1,1) และ FDGM(1,1) ให้ค่า MAPE อยู่ในเกณฑ์ที่ค่อนข้างแม่นยำเมื่อเทียบกับเกณฑ์การประเมินในตารางที่ 1 จากการนำเทคนิคการปรับค่าคลาดเคลื่อนด้วยฟูรีเยร์มาประยุกต์ใช้ร่วมกับตัวแบบ ผลปรากฏว่าค่าคลาดเคลื่อนมีค่าต่ำลง โดยที่ตัวแบบ FDGM(1,1) มีค่า MAPE เท่ากับ 15.25% เป็นค่าที่ต่ำที่สุดเมื่อเปรียบเทียบกับตัวแบบ GM(1,1), DGM(1,1) และ FGM(1,1) จึงเป็นข้อสรุปว่าตัวแบบ FDGM(1,1) เป็นตัวแบบที่เหมาะสมที่สุดในการใช้พยากรณ์จำนวนผู้ป่วยโรคมะเร็งเต้านม

ทั้งนี้ ตัวแบบ FDGM(1,1) ให้ผลการพยากรณ์ที่แม่นยำกว่าแบบจำลอง GM(1,1), DGM(1,1) และ FGM(1,1) เนื่องจากตัวแบบนี้ผสมข้อดีของแบบจำลองเกรย์เข้ากับแนวคิดการแยกส่วนข้อมูล ซึ่งเหมาะสมกับข้อมูลอนุกรมเวลาที่มีลักษณะไม่ต่อเนื่อง อีกทั้งยังใช้เทคนิคการปรับค่าคลาดเคลื่อนด้วยอนุกรมฟูรีเยร์ ทำให้สามารถรองรับทั้งลักษณะข้อมูลแบบไม่ต่อเนื่องและความผันผวนที่ไม่เป็นเส้นตรงได้ดีกว่าแบบจำลองพื้นฐาน ส่งผลให้ค่าความคลาดเคลื่อนของการพยากรณ์ลดลงมากที่สุด

ข้อจำกัดของตัวแบบที่ควรพิจารณา ได้แก่ ตัวแบบ GM(1,1) ซึ่งอาศัยสมการเชิงอนุพันธ์ต่อเนื่อง มักให้ความแม่นยำน้อยลงเมื่อข้อมูลมีความผันผวนสูงหรือมีลักษณะไม่ต่อเนื่อง ขณะที่ตัวแบบ DGM(1,1) แม้จะเหมาะกับข้อมูลแบบ discrete มากกว่า แต่ก็ยังมีความไวต่อความไม่สม่ำเสมอของข้อมูลเช่นกัน นอกจากนี้ทั้งสองแบบจำลองเป็นแบบจำลองตัวแปรเดียว จึงไม่สามารถสะท้อนปัจจัยอื่นที่อาจส่งผลต่อจำนวนผู้ป่วยได้อย่างครบถ้วน ทางผู้วิจัยเสนอแนะว่า หากสามารถรวบรวมข้อมูลปัจจัยที่เกี่ยวข้องกับโรคมะเร็งเต้านมเพิ่มเติมได้ ควรนำมาประยุกต์ใช้ร่วมกับแบบจำลองทำนาย เช่น แบบจำลอง GM(1,N) ซึ่งสามารถพิจารณาตัวแปรต้นหลายตัวแปรร่วมกับตัวแปรตามได้ นอกจากนี้ ควรพิจารณาทดลองใช้เทคนิคการปรับค่าคลาดเคลื่อนรูปแบบอื่น ๆ เพื่อเพิ่มความแม่นยำของแบบจำลองให้ดียิ่งขึ้น

5. เอกสารอ้างอิง

- [1] Nakornthon Hospital, "Top 3 most common cancers in women," *Nakornthon Hospital*, Bangkok, Thailand, 2024. [Online].
<https://www.nakornthon.com/article/detail>

- 1/3 โรคมาเร็ยยอดฮิตพบบ่อยสุดในผู้หญิง.
[Accessed: Jan. 20, 2025].
- [2] National Cancer Institute (Thailand), *Fight cancer guide: Roo Soo Cancer*, Bangkok, Thailand. [Online].
https://www.nci.go.th/th/File_download/fight_cancer/รู้สู้มะเร็ง.pdf. [Accessed: Jan. 20, 2025].
- [3] Hfocus, "Current issues in Thailand's healthcare system," *Hfocus*, Bangkok, Thailand, Mar. 27, 2022. [Online].
<https://www.hfocus.org/content/2022/03/24791> [Accessed: Jan. 20, 2025].
- [4] R. Chotsangsakorn, S. Thanasilp, and N. Pudtong, "Predicting factors of breast cancer screening behaviors among menopausal women," *Kuakarun Journal of Nursing*, vol. 26, no. 2, pp. 107–120, Jul.-Dec. 2019.
- [5] P. Khruasom, C. Puripat, P. Nuchpho, and P. Kongsangson, "Grey system model for forecasting dengue outbreak: Case study of Bangkok metropolis," *Disease Control Journal*, vol. 49, no. 2, pp. 353–363, 2023.
- [6] B. Yunan and V. Jayathavaj, "Forecasting the Number of Deaths from Pneumonia and Intracerebral Hemorrhage," *Regional Health Promotion Center 9 Journal*, vol. 18, no. 3, pp. 946–956, Jul. 2024.
- [7] P. Amphanthong and P. Busababodhin, "Forecasting PM10 in the upper northern area of Thailand with Grey System Theory," *Burapha Science Journal*, vol. 20, no. 1, pp. 15–24, 2015.
- [8] N. Doungdee and M. Chiangpradit, "Forecasting crisis peaks of economics in Thailand using Grey model," *Burapha Science Journal*, vol. 23, no. 2, pp. 1111–1122, May–Aug. 2018.
- [9] V.-T. Phan, Z. Malara, and N. T. Nguyen, "Using Fourier series to improve the Discrete Grey Model (1,1)," *Advances in Computational Collective Intelligence*, pp. 99–109, 2020.
- [10] A. Nimcharoen, P. Busababodhin, and M. Chiangpradit, "Residual modification of forecasting with Fourier series," *Burapha Science Journal*, vol. 23, no. 3, pp. 1396–1408, Sep.-Dec. 2018.
- [11] National Cancer Institute of Thailand, "Cancer incidence statistics in Thailand," [Online].
https://www.nci.go.th/th/cancer_record/cancer_rec1.html. [Accessed: Feb. 17, 2025].
- [12] J. Deng, *Introduction to Grey System Theory*. The Journal of Grey System, vol. 1, pp. 1–24, 1989.
- [13] Z. Ceylan, "Short-term prediction of COVID-19 spread using grey rolling model optimized by particle swarm optimization," *Applied Soft Computing*, vol. 109, p. 107592, Sept. 2021.
- [14] S. Liu and Y. Lin, *Grey Systems: Theory and Applications*. New York: Springer, 2010.

- [15] N.-M. Xie and S. F. Liu, "Discrete grey forecasting model and its optimization," *Applied Mathematical Modelling*, vol. 33, no. 2, pp. 1173–1186, 2009.
- [16] J. Rattanahon and V. Jayathavaj, "Predicting the number of neonatal deaths in Thailand using grey system theory," *Journal of Medicine and Public Health, Ubon Ratchathani University*, vol. 7, no. 2, pp. 154–163, May.-Aug. 2024.
- [17] National Statistical Office, Analytical report on unemployment 2023 Bangkok, Thailand: National Statistical Office, 2023. [Online]. https://www.nso.go.th/public/e-book/Analytical-Reports/Report_Unemployed_2566/17/. [Accessed: Mar. 3, 2025].
- [18] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. New York: Springer, 2013.
- [19] Y.-W. Wang, Z.-Z. Shen, and Y. Jiang, "Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China," *PLoS ONE*, vol. 13, no. 9, pp. e0201987, 2018.
- [20] D. Zhao, H. Zhang, Q. Cao, Z. Wang, S. He, M. Zhou, and R. Zhang, "The research of ARIMA, GM(1,1), and LSTM models for prediction of TB cases in China," *PLoS ONE*, vol. 17, no. 2, pp. e0262734, 2022.