

ตัวแบบสถิติเพื่อพยากรณ์คุณภาพอากาศ: กรณีศึกษาค่าฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน (PM2.5) ในจังหวัดเชียงใหม่

Statistical Model for Air Quality Forecasting: A Case Study of Dust Particles No Larger Than 2.5 Microns (PM2.5) in Chiang Mai Province

กฤติชัย โรจนนะ¹ และ รัชนีวรรณ วงศ์พระจันทร์^{2*}

^{1,2}สาขาวิชาสถิติและการจัดการสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้

Kittichai Rochana¹ and Ratchaneewan Wongprachan^{2*}

^{1,2}Division of Statistics and Information Management, Faculty of Science, Maejo University

ratchaneewan@mju.ac.th

Received 20 April 2025

Revised 30 May 2025

Accepted 11 June 2025

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบจำลองเชิงสถิติสำหรับคาดการณ์ระดับ PM2.5 ในจังหวัดเชียงใหม่ โดยเปรียบเทียบประสิทธิภาพเทคนิคการพยากรณ์ 3 วิธี ได้แก่ แบบจำลองบ็อกซ์-เจนกินส์ แบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ และแบบจำลองการถดถอยเชิงเส้น ข้อมูลที่ใช้วิเคราะห์เป็นข้อมูลค่าเฉลี่ยรายเดือนของระดับ PM2.5 ที่จัดเก็บโดยกรมควบคุมมลพิษ ระหว่างปี พ.ศ. 2556 - 2567 รวมทั้งสิ้น 144 เดือน โดยข้อมูลถูกแบ่งเป็น 2 ชุด ได้แก่ ชุดฝึกฝน จำนวน 120 เดือน (พ.ศ. 2556 - 2565) เพื่อใช้ในการพัฒนาแบบจำลอง และชุดทดสอบ จำนวน 24 เดือน (พ.ศ. 2566 - 2567) เพื่อประเมินประสิทธิภาพของแบบจำลอง ผลการศึกษาพบว่า แบบจำลองด้วยวิธีของวินเทอร์สามารถทำนายปริมาณฝุ่นละออง PM2.5 ได้แม่นยำกว่าวิธีอื่น โดยมีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด แบบจำลองที่ดีที่สุดนี้ถูกนำไปใช้ในการพยากรณ์ค่า PM2.5 รายเดือนสำหรับปี พ.ศ. 2568 ซึ่งพบว่าช่วงเดือนมีนาคมถึงเมษายนมีแนวโน้มที่ระดับ PM2.5 จะสูงเกินเกณฑ์ที่กรมควบคุมมลพิษกำหนดไว้ที่ 37.5 มคก./ลบ.ม. เน้นย้ำถึงความจำเป็นในการบริหารจัดการคุณภาพอากาศเชิงรุกในช่วงเวลาดังกล่าว

คำสำคัญ: บ็อกซ์-เจนกินส์, การปรับให้เรียบด้วยวิธีวินเทอร์, การถดถอยเชิงเส้น

Abstract

This research aims to develop a statistical model for forecasting PM2.5 levels in Chiang Mai Province. Three forecasting techniques were compared: the Box-Jenkins model, Winter's exponential smoothing model, and the linear regression model. The analysis utilized monthly average PM2.5 data from the Pollution Control Department covering 144 observations from 2013 to 2024. The results show that the Winter's model performed best in predicting PM2.5 levels, with the lowest average absolute percentage error. This model was used for forecasting monthly PM2.5 levels for 2025, which shows a trend of exceeding the PM2.5 threshold set by the Pollution Control Department (37.5 µg/m³) from March to April. The study emphasizes the need for proactive air quality management during this period.

2024. The dataset was divided into two subsets: a training set of 120 months (2013–2022) for model development and a testing set of 24 months (2023–2024) for model performance evaluation. Results showed that Winter’s exponential smoothing model yielded the most accurate predictions, with the smallest mean absolute percentage error (MAPE) among the three methods. The best-performing model was then used to forecast monthly PM2.5 levels for the year 2025. The forecast suggested that PM2.5 levels from February to April are likely to exceed the standard threshold of 37.5 $\mu\text{g}/\text{m}^3$ set by the Pollution Control Department, underscoring the importance of proactive air quality management during this period.

Keywords: Box-Jenkins Model, Winter’s Exponential Smoothing Model, Linear Regression Model

1. บทนำ

ปัญหาหมอกควันและคุณภาพอากาศในภาคเหนือตอนบนของประเทศไทย ยังคงมีค่าฝุ่นละอองในอากาศเกินมาตรฐานที่กำหนด และมีแนวโน้มทวีความรุนแรงเพิ่มขึ้นอย่างต่อเนื่อง เนื่องจากภูมิประเทศที่ถูกล้อมรอบด้วยภูเขา ซึ่งจำกัดการไหลเวียนของอากาศ เกิดฝุ่นละอองสะสมในชั้นบรรยากาศ ทำให้อากาศไม่หมุนเวียน หรือที่เรียกว่า ปรากฏการณ์ฝาชีครอบต่ำ จึงพบหมอกควันหนาแน่นเกินจากค่ามาตรฐานที่กำหนด ฝุ่นละออง PM2.5 เป็นฝุ่นอนุภาคขนาดเล็กที่มีเส้นผ่านศูนย์กลางไม่เกิน 2.5 ไมครอน สามารถแทรกซึมเข้าสู่ระบบทางเดินหายใจส่วนลึกและเข้าสู่กระแสเลือดได้ ส่งผลต่อสุขภาพในระยะยาว ก่อให้เกิดโรคต่าง ๆ เช่น โรคมะเร็งปอด โรคหลอดเลือดสมอง โรคหลอดเลือดหัวใจ โรคปอดอุดกั้นเรื้อรัง โรคหอบหืด และโรคภูมิแพ้ เป็นต้น รวมถึงภาวะผิดปกติทางผิวหนังและดวงตา ทั้งนี้ข้อมูลจากกระทรวงสาธารณสุขในปี พ.ศ. 2566 พบว่ามลพิษทางอากาศส่งผลให้ประชาชนเจ็บป่วยจากโรคที่เกี่ยวข้องกว่า 10.5 ล้านราย [1, 2]

นอกจากลักษณะภูมิประเทศที่เอื้อต่อการสะสมของฝุ่นละอองแล้ว ต้นเหตุสำคัญของการเกิดฝุ่นละออง PM2.5 ในภาคเหนือมีความเชื่อมโยงกับการเผาไหม้ในรูปแบบต่าง ๆ เช่น การเผาในที่โล่ง การเผาเพื่อเตรียมพื้นที่เพาะปลูก หรือทำการเกษตร การเผาเพื่อหาของป่า ทำไร่เลื่อนลอย หรือเตรียมพื้นที่เลี้ยงสัตว์ การเผาในประเพณีเพื่อบ้าน การเผาไหม้ขยะที่ไม่เหมาะสมหรือไม่ได้ควบคุมอย่างเหมาะสม ปัจจัยเพิ่มเติมยังรวมถึงการใช้พลังงานจากการเผาไหม้ที่ไม่มีประสิทธิภาพ เช่น การใช้เชื้อเพลิงที่ไม่ได้มาตรฐาน กระบวนการเผาที่ไม่สะอาดในการผลิตไฟฟ้า หรือการใช้พลังงานอื่น ๆ ตลอดจนกิจกรรมด้านการขนส่ง การจราจรทางถนน และการขนส่งทางอากาศ เป็นต้น [3] ผลกระทบจากมลภาวะทางอากาศไม่ได้จำกัดเพียงด้านสุขภาพเท่านั้น แต่ยังส่งผลในวงกว้างต่อระบบเศรษฐกิจของจังหวัดเชียงใหม่และพื้นที่ใกล้เคียงโดยเฉพาะในอุตสาหกรรมการท่องเที่ยว ที่เผชิญกับภาวะซบเซาจากการหดตัวของนักท่องเที่ยว อันเป็นผลสืบเนื่องจากความกังวลเกี่ยวกับปัญหาคุณภาพอากาศ ส่งผลต่อรายได้ที่ลดลงของร้านอาหาร โรงแรม และสถานที่ท่องเที่ยวต่าง ๆ นอกจากนี้การแพร่กระจายของฝุ่นละอองในอากาศยังทำให้ทัศนวิสัยแย่งนำไปสู่ความไม่สะดวกในการคมนาคมทั้งภาคพื้นดินและทางอากาศ

ปัญหา PM2.5 ยังคงอยู่ในระดับที่น่าเป็นห่วงโดยเฉพาะในพื้นที่จังหวัดเชียงใหม่ จากรายงานค่าดัชนีคุณภาพอากาศ เมื่อวันที่ 15 มีนาคม พ.ศ. 2567 พบว่ามีระดับ PM2.5 อยู่ในช่วงระหว่าง 74.0 – 181.3 มคก./ลบ.ม. ซึ่งส่งผลให้จังหวัดเชียงใหม่ถูกจัดให้อยู่ในอันดับหนึ่งของเมืองที่มีมลพิษทางอากาศรุนแรงที่สุดในโลกในวันดังกล่าว [4] ปัญหามลพิษจากฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ยังคงเป็นประเด็นด้านสิ่งแวดล้อมที่สำคัญ เนื่องจากส่งผลกระทบต่อ

โดยตรงต่อสุขภาพของประชาชน เพื่อยกระดับการควบคุมคุณภาพอากาศ กรมควบคุมมลพิษจึงได้ปรับปรุงเกณฑ์มาตรฐานใหม่และประกาศบังคับใช้ในวันที่ 1 มิถุนายน พ.ศ. 2566 โดยลดระดับความเข้มข้นเฉลี่ยราย 24 ชั่วโมงของ PM2.5 จากเดิมไม่เกิน 50 ลงมาเป็น 37.5 มคก./ลบ.ม. [5]

การจัดการปัญหามลพิษทางอากาศจำเป็นต้องดำเนินการอย่างทันท่วงที โดยต้องอาศัยการบูรณาการความร่วมมือจากทุกภาคส่วน เพื่อควบคุมแหล่งกำเนิดมลพิษและคุ้มครองสุขภาพของประชาชน ดังนั้นการเฝ้าระวังและประเมินระดับ PM2.5 ควบคู่ไปกับการวางมาตรการเพื่อลดการสะสมของมลภาวะในชั้นบรรยากาศ จึงถือเป็นประเด็นความสำคัญอย่างยิ่ง ผู้วิจัยจึงได้ตระหนักถึงความสำคัญและผลที่เกิดจากปัญหานี้ จึงดำเนินการศึกษาพัฒนาแบบจำลองทางสถิติเพื่อพยากรณ์ปริมาณฝุ่น PM2.5 ในจังหวัดเชียงใหม่

งานวิจัยที่ผ่านมาได้มีความพยายามในการพยากรณ์ค่าฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ด้วยวิธีการที่หลากหลาย โดย วัฒนา ชยธวัช [6] ได้ใช้วิธีแยกส่วนประกอบแบบดั้งเดิมกับข้อมูลรายเดือนในช่วงปี พ.ศ. 2563 – 2565 รวมระยะเวลา 36 เดือน ผลการพยากรณ์ให้ค่า ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error, MAPE) เท่ากับ 28.72 ซึ่งแม้ว่าวิธีดังกล่าวจะสามารถแสดงองค์ประกอบของแนวโน้มและฤดูกาลได้อย่างคร่าว ๆ แต่ข้อจำกัดด้านความแม่นยำของแบบจำลองสะท้อนให้เห็นว่าวิธีดังกล่าวอาจไม่สามารถจับรูปแบบที่ซับซ้อนของข้อมูลมลพิษทางอากาศได้อย่างเพียงพอ ขณะเดียวกัน ปรัชญา สิงหรวงศ์ [7] ได้ประยุกต์ใช้แบบจำลองหน่วยความจำระยะยาว (LSTM) สำหรับพยากรณ์ค่า PM2.5 รายชั่วโมงล่วงหน้า โดยใช้ข้อมูลระหว่างปี พ.ศ. 2561 - 2562 พบว่าแบบจำลอง LSTM ให้ค่า MAPE ที่สถานีศูนย์ราชการจังหวัดเชียงใหม่ เท่ากับ 20.42 อย่างไรก็ตาม เมื่อใช้แบบจำลองเดียวกันกับข้อมูลจากสถานีโรงเรียนยุพราชวิทยาลัย กลับให้ค่า MAPE สูงถึง 46.40 ซึ่งแสดงถึงความไม่เสถียรของแบบจำลองเมื่อนำไปใช้กับข้อมูลจากพื้นที่ต่างกัน อันอาจเนื่องมาจากความแตกต่างของคุณภาพข้อมูล ขอบเขตเวลา และการจัดการกับข้อมูลที่ขาดหาย

เพื่อแก้ไขข้อจำกัดที่พบในงานวิจัยข้างต้น งานวิจัยนี้จึงมุ่งพัฒนาแบบจำลองที่แตกต่างกับที่เคยมีการศึกษา โดยใช้ตัวแบบสถิติ 3 เทคนิค ได้แก่ แบบจำลองบ็อกซ์-เจนกินส์ แบบจำลองการปรับให้เรียบด้วยวิธีของวินเทอร์ และแบบจำลองการถดถอยเชิงเส้น โดยใช้ข้อมูลค่าเฉลี่ยรายเดือนย้อนหลังเป็นระยะเวลา 10 ปี ซึ่งเป็นช่วงเวลาที่ข้อมูลต่อเนื่องมากขึ้น ช่วยให้สามารถตรวจสอบและจำลองรูปแบบการเปลี่ยนแปลงของค่าฝุ่น PM2.5 ได้ชัดเจนยิ่งขึ้น จากนั้นจะทำการเปรียบเทียบประสิทธิภาพของแต่ละเทคนิค เพื่อเลือกตัวแบบจำลองที่ดีที่สุดมาทำนายค่าเฉลี่ยรายเดือนของฝุ่น PM2.5 ล่วงหน้า 12 เดือน ผลการศึกษาที่ได้จะช่วยให้หน่วยงานที่เกี่ยวข้องสามารถคาดการณ์ระดับ PM2.5 ในจังหวัดเชียงใหม่ล่วงหน้าได้ ซึ่งจะช่วยในการเฝ้าระวังและประเมินสถานการณ์มลพิษทางอากาศได้ดียิ่งขึ้น การติดตามแนวโน้มฝุ่น PM2.5 ล่วงหน้า ช่วยเสริมสร้างความพร้อมในการจัดการและป้องกันปัญหามลพิษทางอากาศอย่างมีระบบ

2. ขอบเขตงานวิจัย

2.1 ขอบเขตด้านเนื้อหา ในการศึกษาครั้งนี้มุ่งเน้นพัฒนาแบบจำลองมลพิษทางอากาศเพื่อพยากรณ์ปริมาณฝุ่นละออง PM2.5 โดยใช้ตัวแบบทางสถิติ 3 เทคนิค ได้แก่ แบบจำลองบ็อกซ์-เจนกินส์ แบบจำลองการปรับให้เรียบด้วยวิธีของวินเทอร์ และแบบจำลองการถดถอยเชิงเส้น

2.2 ขอบเขตด้านพื้นที่ ในการศึกษาครั้งนี้มุ่งเน้นพัฒนาแบบจำลองปริมาณฝุ่นละออง PM2.5 ในเขตจังหวัดเชียงใหม่ โดยใช้ข้อมูลค่า PM2.5 เฉลี่ย 24 ชั่วโมง ที่จัดเก็บบันทึกรายเดือนในฐานะข้อมูลของกรมควบคุมมลพิษ ณ จุดตรวจวัด ในเขต ต.ศรีภูมิ อ.เมือง จ.เชียงใหม่

2.3 ขอบเขตด้านเวลา ทำการพัฒนาตัวแบบโดยใช้ข้อมูลฝุ่นละออง PM2.5 ในช่วง ปี พ.ศ. 2556 ถึง 2567

3. ทฤษฎีที่เกี่ยวข้อง

3.1 องค์ประกอบของข้อมูลอนุกรมเวลา [6, 8]

แนวโน้ม (Trend) เป็นการเคลื่อนไหวของข้อมูลที่เกิดขึ้นอย่างต่อเนื่องในช่วงเวลาหนึ่ง ซึ่งมักจะแสดงทิศทางที่ชัดเจนในระยะยาว โดยแนวโน้มสามารถมีในทิศทางเพิ่มขึ้น ลดลง หรือเปลี่ยนแปลงไปมา โดยไม่จำเป็นต้องเป็นเส้นตรงเสมอไป

ฤดูกาล (Seasonal) เป็นการเคลื่อนไหวของข้อมูลที่เกิดขึ้นเป็นระยะๆ ในช่วงเวลาที่กำหนดและสามารถคาดการณ์ล่วงหน้าได้ เช่น การเคลื่อนไหวในช่วงรายเดือน ไตรมาส หรือปี มักเกิดจากปัจจัยภายนอก เช่น สภาพอากาศ หรือนโยบาย

วัฏจักร (Cyclic) เป็นรูปแบบการขึ้นลงของข้อมูลที่ไม่มีช่วงเวลาที่แน่นอน มักเกี่ยวข้องกับปัจจัยทางเศรษฐกิจหรือวัฏจักรทางธุรกิจ เช่น ภาวะเศรษฐกิจขาขึ้นหรือขาลง โดยทั่วไปแล้ว วัฏจักรจะมีระยะเวลายาวนานกว่ารูปแบบตามฤดูกาล และขนาดของวัฏจักรมักมีความผันแปรมากกว่าขนาดของรูปแบบตามฤดูกาล ระยะเวลาของวัฏจักรมักยาวนานกว่า 2 ปี และไม่สามารถคาดการณ์ได้อย่างแม่นยำเหมือนกับรูปแบบฤดูกาล

ความผิดปกติ (Irregular Fluctuation) เป็นองค์ประกอบสุ่มที่ผันผวนจากเหตุการณ์ที่ไม่สามารถคาดเดาได้ ก่อนที่เหตุการณ์จะเกิดขึ้น เช่น ภัยจากธรรมชาติ โรคระบาด เป็นต้น

3.2 เทคนิคการพยากรณ์

3.2.1 แบบจำลองด้วยวิธีบ็อกซ์-เจนกินส์ (Box-Jenkins Model) เป็นเทคนิคการพยากรณ์ข้อมูลอนุกรมเวลาที่ต้องอาศัยข้อมูลที่มีลักษณะคงที่ (Stationary) โดยสามารถตรวจสอบลักษณะของข้อมูลได้จากกราฟฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation Function; ACF) และกราฟฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (Partial Autocorrelation Function; PACF) หากพบว่าข้อมูลมีลักษณะไม่คงที่ จำเป็นที่จะต้องทำให้คงที่โดยการแปลงข้อมูลโดยการหาผลต่างของอนุกรมเวลาในกรณีที่พบว่าแนวโน้ม การหาผลต่างฤดูกาลของอนุกรมเวลาในกรณีที่พบว่าฤดูกาล เมื่อดำเนินการแปลงข้อมูลอนุกรมเวลาให้คงที่เรียบร้อยแล้ว จะพิจารณากราฟ ACF และ PACF อีกครั้ง เพื่อใช้เป็นแนวทางในการกำหนดโครงสร้างของแบบจำลองที่สอดคล้องและเหมาะสมที่สุดต่อไป

ในการศึกษานี้ได้พิจารณาองค์ประกอบของฤดูกาลร่วมด้วย เนื่องจากข้อมูลค่าฝุ่น PM2.5 มีลักษณะความผันแปรตามฤดูกาล จึงเลือกใช้แบบจำลอง SARIMA(p, d, q)(P, D, Q)_s (Seasonal Autoregressive Integrated Moving Average) ซึ่งประกอบด้วยองค์ประกอบหลัก ได้แก่ ค่าคงที่ (α) รวมถึง ตัวดำเนินการสหสัมพันธ์ในตัวเองทั้งแบบไม่มีฤดูกาลและแบบมีฤดูกาล ซึ่งมีอันดับเท่ากับ p และ P ตามลำดับ ตลอดจนตัวดำเนินการค่าเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาลและแบบมีฤดูกาล ซึ่งมีอันดับเท่ากับ q และ Q ตามลำดับ ตัวแบบ SARIMA(p, d, q)(P, D, Q)_s สามารถแสดงในรูปสมการได้ดังนี้ [9, 10]

$$\nu_p(B)N_p(B^s)(1-B)^d(1-B^s)^DY_t = \alpha + \kappa_q(B)K_Q(B^s)\xi_t \quad (1)$$

เมื่อ	Y_t	คือ ข้อมูลอนุกรม ณ เวลา t
	ξ_t	คือ ความคลาดเคลื่อนสุ่ม ณ เวลา t ภายใต้ข้อสมมติ $\xi_t \sim Nid(0, \sigma^2)$
	B	คือ ตัวดำเนินการย้อนกลับ โดยที่ $B^k Y_t = Y_{t-k}$
	d, D	คือ อันดับผลต่างและผลต่างฤดูกาลของอนุกรมเวลา
	s	คือ จำนวนฤดูกาลต่อปี

$$\begin{aligned}\alpha &= \mu \nu_p(B) N_p(B^s) \quad \text{เมื่อ } \mu \text{ เป็นเฉลี่ยของอนุกรมเวลาที่มีลักษณะคงที่} \\ \nu_p(B) &= 1 - \nu_1 B - \nu_2 B^2 - \dots - \nu_p B^p \\ N_p(B^s) &= 1 - N_1 B^s - N_2 B^{2s} - \dots - N_p B^{Ps} \\ \kappa_q(B) &= 1 - \kappa_1 B - \kappa_2 B^2 - \dots - \kappa_q B^q \\ K_Q(B^s) &= 1 - K_1 B^s - K_2 B^{2s} - \dots - K_Q B^{Qs}\end{aligned}$$

3.2.2 แบบจำลองการปรับให้เรียบด้วยวิธีของวินเทอร์ (Winters' Exponential Smoothing) เป็นเทคนิคที่เหมาะสมกับอนุกรมเวลาที่มีประกอบด้วยแนวโน้มที่เป็นเส้นตรงร่วมกับความผันแปรตามฤดูกาล โดยสามารถจำแนกได้ 2 รูปแบบ คือ รูปแบบบวกและรูปแบบคูณ ในการศึกษาครั้งนี้จะทำการพยากรณ์ด้วยรูปแบบคูณ เนื่องจากข้อมูลค่าฝุ่น PM2.5 มีลักษณะของความผันแปรตามฤดูกาลที่ไม่แน่นอน และเปลี่ยนแปลงไปตามช่วงเวลา พยากรณ์ที่ใช้จึงต้องให้น้ำหนักกับข้อมูลในอดีตอย่างไม่เท่ากัน โดยอาศัยค่าปรับให้เรียบ (Smoothing Constants) ทั้งหมด 3 ค่า ได้แก่ ค่าคงที่สำหรับระดับ α สำหรับความชัน γ และสำหรับฤดูกาล δ ($0 \leq \alpha, \gamma, \delta \leq 1$) ตัวแบบจำลองด้วยเทคนิคการปรับให้เรียบด้วยวิธีของวินเทอร์แบบคูณ แสดงได้ดังสมการที่ (2) [11, 12]

$$Y_t = (\beta_0 + \beta_1 t) M_t \xi_t \quad (2)$$

เมื่อ Y_t คือ ข้อมูลอนุกรม ณ เวลา t
 β_0, β_1 คือ ค่าพารามิเตอร์ที่บ่งบอกระดับของอนุกรมเวลา และความชันของแนวโน้ม
 M_t คือ ความผันแปรตามฤดูกาล ณ เวลา t เมื่อ s คือ จำนวนฤดูกาลต่อปี
 ξ_t คือ ความคลาดเคลื่อนสุ่ม ณ เวลา t ภายใต้ข้อสมมติ $Y_t = \theta \exp(\hat{y}_t)$
 และมีตัวแบบการพยากรณ์ สำหรับช่วงเวลาในอนาคต r คาบ ณ เวลา t แสดงได้ดังสมการที่ (3)

$$\hat{Y}_{t+r} = (v_t + u_t r) \hat{M}_t \quad (3)$$

โดยที่ $\hat{Y}_{t+r}$ คือ ค่าพยากรณ์ ณ เวลา $t + r$ ซึ่ง r หมายถึงจำนวนคาบเวลาที่พยากรณ์ล่วงหน้า

$$\begin{aligned}v_t &= \alpha \left(\frac{Y_t}{\hat{M}_{t-s}} \right) + (1 - \alpha)(v_{t-1} + u_{t-1}) \\ u_t &= \gamma (v_t - v_{t-1}) + (1 - \gamma)u_{t-1} \\ \hat{M}_t &= \delta \left(\frac{Y_t}{v_t} \right) + (1 - \delta)\hat{M}_{t-s}\end{aligned}$$

3.2.3 แบบจำลองการถดถอยเชิงเส้น (Linear Regression Model) เทคนิคทางสถิติที่ใช้เพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรต้น (Independent Variables) กับตัวแปรตาม (Dependent Variable) โดยมีวัตถุประสงค์เพื่อพยากรณ์หรือคาดการณ์ค่าของตัวแปรตาม จากค่าของตัวแปรต้นที่กำหนด ในการสร้างแบบจำลองคาดการณ์ระดับฝุ่นละออง PM2.5 ภายใต้ชุดข้อมูลอนุกรมเวลาขนาด n ซึ่งแสดงลักษณะของแนวโน้มและฤดูกาล ดังนั้นตัวแบบการถดถอยจะประกอบไปด้วยส่วนของแนวโน้มเชิงเส้นซึ่งพิจารณาจากตัวแปรเวลา และส่วนของฤดูกาล

ซึ่งตัวแปรฤดูกาลจะถูกกำหนดให้อยู่ในรูปแบบของตัวแปรหุ่นจำนวน $s-1$ ตัว เมื่อ s คือจำนวนฤดูกาลต่อปี โดยจะกำหนดให้ตัวแปรหุ่นมีค่าเท่ากับ 0 เมื่ออนุกรมเวลาไม่ได้เกิดในฤดูกาลที่กำหนด และมีค่าเท่ากับ 1 เมื่ออนุกรมเวลาเกิดในฤดูกาลที่กำหนด จะได้ตัวแบบจำลองดังสมการที่ (4) [8]

$$Y_t = \alpha + \beta t + \sum_{i=1}^{s-1} \delta_i M_{it} + \xi_t \quad (4)$$

เมื่อ Y_t คือ ข้อมูลอนุกรมเวลา ณ เวลา t
 α คือ ระยะตัดแกนตั้ง
 β คือ สัมประสิทธิ์ถดถอยของแนวโน้ม
 δ_i คือ สัมประสิทธิ์ถดถอยของฤดูกาลที่ i ; $i = 1, 2, \dots, s-1$
 $M_{it} = \begin{cases} 1 & \text{เมื่อข้อมูลอนุกรมเวลา ณ เวลา } t \text{ เกิดในฤดูกาลที่ } i \\ 0 & \text{เมื่อข้อมูลอนุกรมเวลา ณ เวลา } t \text{ ไม่ได้เกิดในฤดูกาลที่ } i \end{cases}$
 ξ_t คือ ความคลาดเคลื่อนสุ่ม ณ เวลา t ภายใต้อสมมติ $\xi_t \sim \text{Nid}(0, \sigma^2)$

ตัวแบบการพยากรณ์ ณ เวลา t แสดงได้ดังสมการที่ (5)

$$\hat{Y}_t = a + bt + \sum_{i=1}^{s-1} c_i M_{it} \quad (5)$$

3.3 การประเมินประสิทธิภาพของแบบจำลอง

ข้อมูลอนุกรมเวลาชุดที่ 2 ซึ่งประกอบไปด้วยข้อมูล PM2.5 จำนวน 24 เดือน ครอบคลุมช่วงปี พ.ศ. 2566 – 2567 จะถูกนำมาประเมินประสิทธิภาพของแบบจำลอง โดยใช้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error, MAPE) เป็นเกณฑ์ในการประเมินความแม่นยำของแบบจำลอง โดยสามารถแปลผลระดับความแม่นยำได้ดังตารางที่ 1 [13]

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{e_t}{Y_t} \times 100 \right|}{n} \quad (6)$$

โดยที่ Y_t คือ ข้อมูลอนุกรม ณ เวลา t
 e_t คือ ค่าความคลาดเคลื่อน ณ เวลา t ; $e_t = Y_t - \hat{Y}_t$
 n คือ จำนวนข้อมูลอนุกรมเวลา

ภายหลังการประเมินความแม่นยำของแต่ละวิธี ตัวแบบจำลองต่าง ๆ จะถูกตรวจสอบข้อสมมติที่สำคัญโดยใช้การทดสอบทางสถิติต่าง ๆ ได้แก่ การทดสอบของ Ljung-Box Q เพื่อตรวจสอบการกระจายตัวที่ไม่เป็นอิสระกันของความคลาดเคลื่อน การทดสอบ t สำหรับความมีนัยสำคัญของพารามิเตอร์ การทดสอบ Kolmogorov-Smirnov เพื่อตรวจสอบการแจกแจงปกติของค่าความคลาดเคลื่อน และการทดสอบ Levene เพื่อประเมินความแปรปรวนคงที่

ตารางที่ 1 เกณฑ์ความแม่นยำในการพยากรณ์ด้วยวิธี MAPE [14]

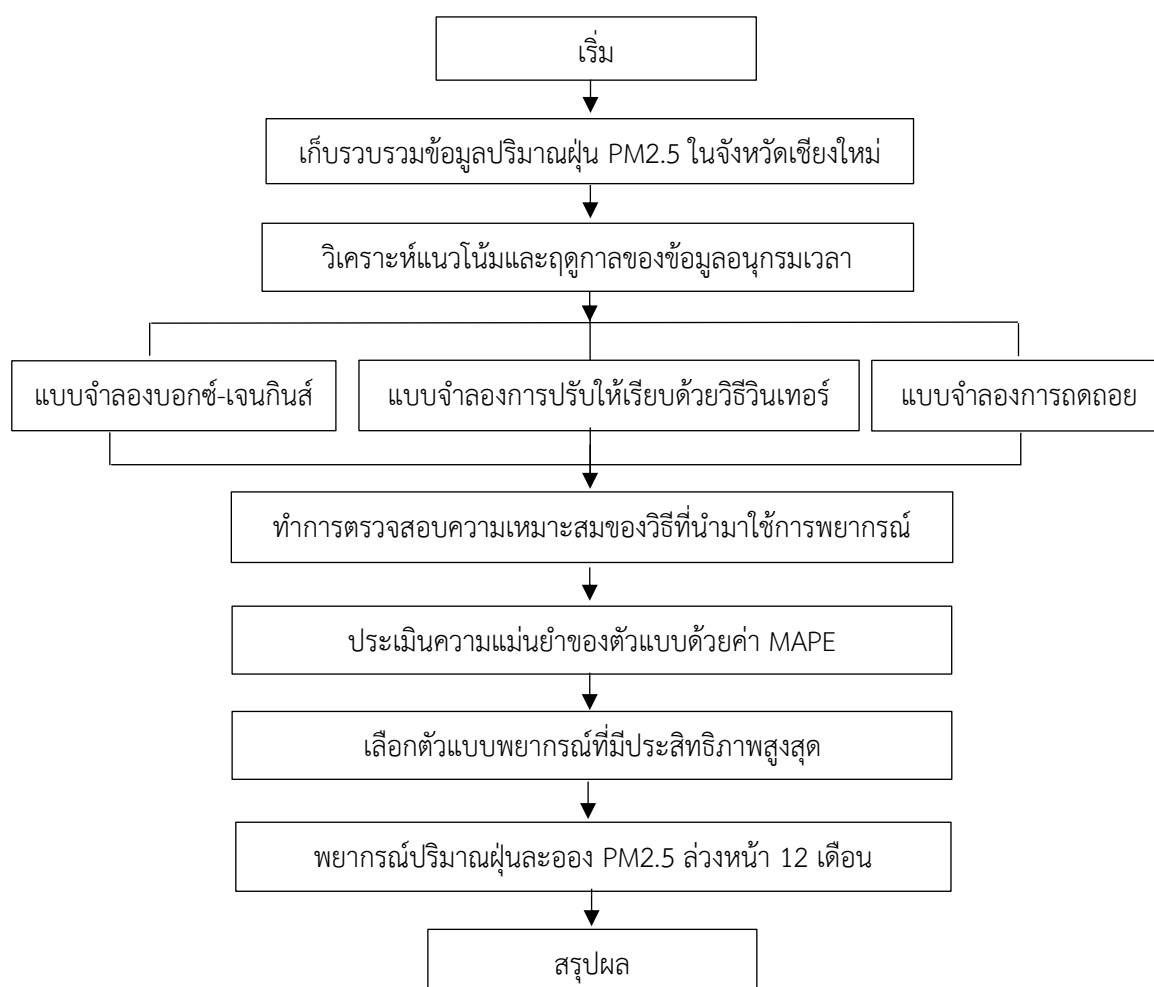
MAPE	ระดับความแม่นยำ
< 10%	แม่นยำสูง
10% – 20%	แม่นยำดี
20% - 50%	แม่นยำใช้ได้
มากกว่า 50%	ไม่มีความแม่นยำ

3.4 การพยากรณ์ PM2.5

นำแบบจำลองที่มีเหมาะสมและมีความแม่นยำสูงสุดไปใช้ในการคาดการณ์ค่าเฉลี่ยรายเดือนของค่าฝุ่นละออง PM2.5 ล่วงหน้า 12 เดือน สำหรับปี พ.ศ. 2568

4. การดำเนินการวิจัย

การศึกษาค้นคว้าครั้งนี้พัฒนาแบบจำลองการทำนายปริมาณฝุ่นละออง PM2.5 ในเขตจังหวัดเชียงใหม่ มีวิธีดำเนินงานวิจัยดังนี้



รูปที่ 1 ขั้นตอนวิธีดำเนินการวิจัย

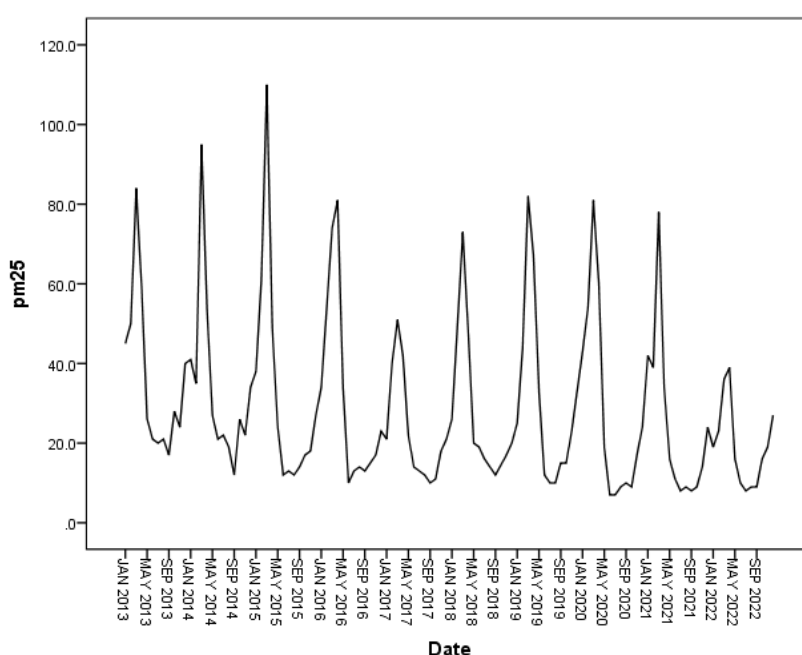
ประชากร คือ ข้อมูลค่าฝุ่นละออง PM2.5 เฉลี่ย 24 ชั่วโมง ในจังหวัดเชียงใหม่

กลุ่มตัวอย่าง คือ ข้อมูลค่าฝุ่นละออง PM2.5 เฉลี่ย 24 ชั่วโมง ณ จุดตรวจวัดในเขต ต.ศรีภูมิ อ.เมือง จ.เชียงใหม่ ที่จัดเก็บบันทึกรายเดือนในฐานะข้อมูลของกรมควบคุมมลพิษ ตั้งแต่ พ.ศ. 2556 ถึง 2567 รวมทั้งสิ้นจำนวน 144 เดือน ข้อมูลที่ศึกษาจะถูกแบ่งเป็น 2 ชุด ได้แก่ ชุดฝึกฝน (Training Data Set) จำนวน 120 เดือน โดยใช้ข้อมูลตั้งแต่ พ.ศ. 2556 - 2565 เพื่อใช้ในการเรียนรู้รูปแบบความสัมพันธ์และแนวโน้มของข้อมูล โดยข้อมูลชุดนี้จะถูกนำมาสร้างแบบจำลองด้วยเทคนิคทางสถิติ 3 วิธี ได้แก่ แบบจำลองบอซ-เจนกินส์ แบบจำลองการปรับให้เรียบด้วยวิธีของวินเทอร์ และแบบจำลองการถดถอยเชิงเส้น พร้อมตรวจสอบความเหมาะสมและคุณสมบัติตามข้อสมมติของตัวแบบ ข้อมูลชุดที่ 2 จะถูกนำมาใช้เป็นชุดทดสอบ (Test Data Set) จำนวน 24 เดือน โดยใช้ข้อมูลตั้งแต่ พ.ศ. 2566 - 2567

ในการศึกษาครั้งนี้แบบจำลองที่พัฒนาขึ้นจะดำเนินการโดยใช้โปรแกรมวิเคราะห์ข้อมูลทางสถิติ ได้แก่ Minitab และ SPSS หลังจากได้แบบจำลองทั้งสามแล้ว จะประเมินประสิทธิภาพของแบบจำลองด้วยค่า MAPE ในการประเมินความแม่นยำของแต่ละวิธี ซึ่งเป็นเกณฑ์ที่สามารถใช้ได้กับข้อมูลที่มีหน่วยต่างกัน และไม่ถูกรบกวนจากขนาดของข้อมูล เมื่อได้แบบจำลองที่มีประสิทธิภาพสูงสุดแล้ว จะนำไปประยุกต์ใช้ในการพยากรณ์ปริมาณฝุ่น PM2.5 ล่วงหน้า 12 เดือน ของปีพ.ศ. 2568 ต่อไป

5. ผลการวิจัย

ข้อมูลระดับฝุ่น PM2.5 ในพื้นที่จังหวัดเชียงใหม่ ตั้งแต่ พ.ศ. 2556 - 2565 มีลักษณะการเคลื่อนไหวที่ไม่คงที่ ดังแสดงในรูปที่ 2 จากผลวิเคราะห์หาค่าเฉลี่ย 28.80 มคก./ลบ.ม. มีค่าต่ำสุด 7 มคก./ลบ.ม. และค่าสูงสุดเท่ากับ 110 มคก./ลบ.ม. จากผลการทดสอบสถิติ Kolmogorov-Smirnov เพื่อตรวจสอบการแจกแจง พบว่าข้อมูลไม่เป็นไปตามรูปแบบการแจกแจงปกติ (KS-Statistic = 0.192, P-Value < 0.01) จึงได้ดำเนินการแปลงข้อมูลให้อยู่ในรูปลอการิทึมธรรมชาติ $\ln(Y)$ เพื่อปรับลักษณะการกระจายของข้อมูลให้สอดคล้องกับสมมติฐานการแจกแจงปกติ (KS-Statistic = 0.071, P-Value < 0.139)



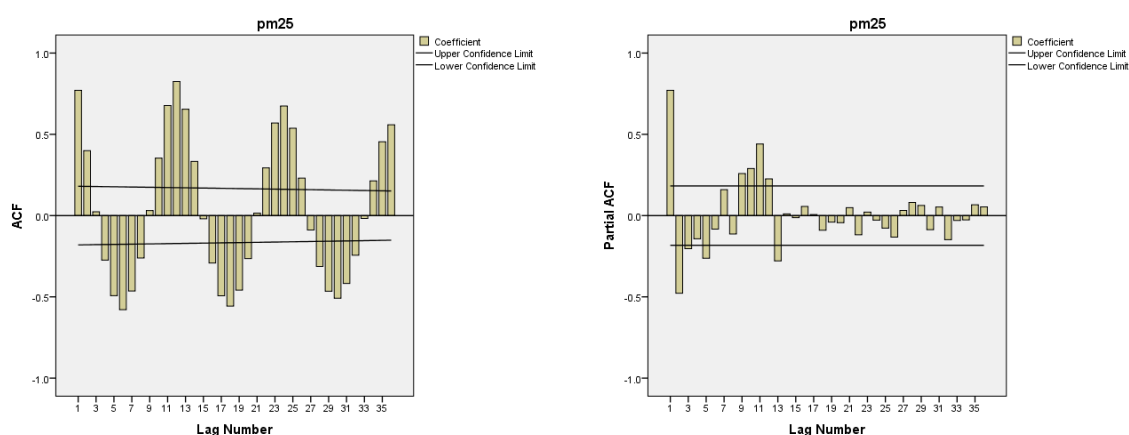
รูปที่ 2 การเคลื่อนไหวของฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ ระหว่างปี พ.ศ. 2556 – 2565

ตารางที่ 2 การตรวจสอบการแจกแจงของข้อมูลฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่

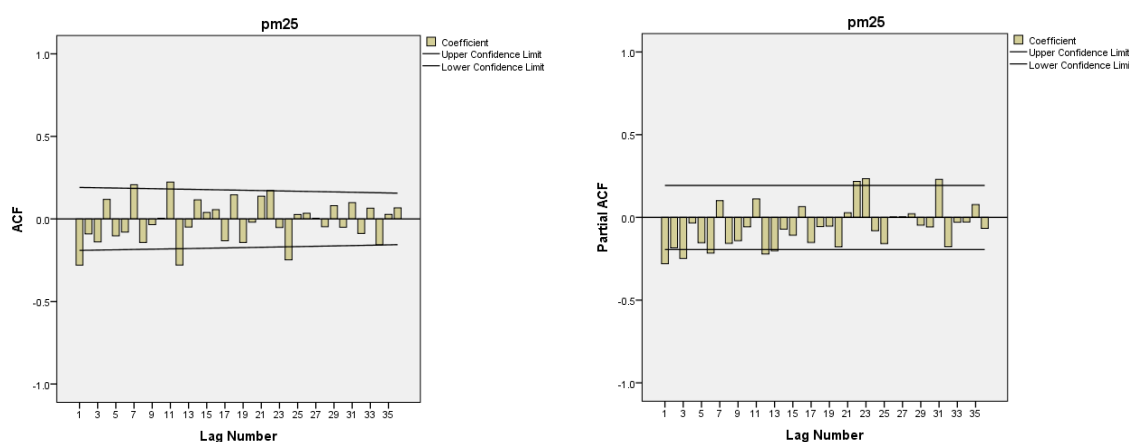
ข้อมูล	Min	Max	Mean	Standard Deviation	KS-Statistic	P-value
PM2.5	7	110	28.796	21.303	0.192	< 0.01
ln(PM2.5)	1.946	4.700	3.130	0.669	0.071	0.139

5.1 แบบจำลองบอซ-เจนกินส์

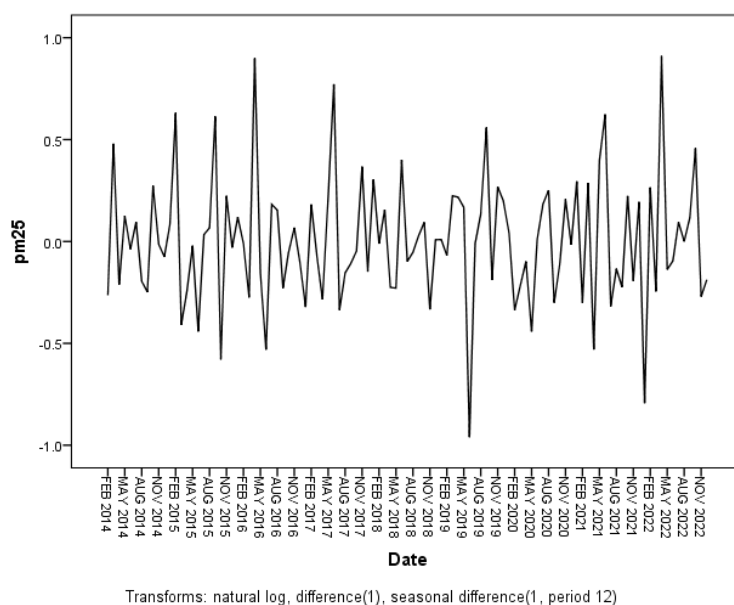
ข้อมูลฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ ระหว่างปี พ.ศ. 2556 ถึง 2565 หลังจากถูกแปลงเป็น $\ln(Y)$ แล้วพบว่าอนุกรมเวลายังมีลักษณะไม่คงที่ โดยสะท้อนให้เห็นถึงแนวโน้มที่ต่อเนื่องและรูปแบบการเปลี่ยนแปลงตามฤดูกาลอย่างเด่นชัด จากกราฟ ACF และ PACF ในรูปที่ 3 ชี้ให้เห็นว่าถึงลักษณะไม่คงที่ของข้อมูลอนุกรมเวลา จึงทำการแปลงข้อมูลเพิ่มเติมด้วยการหาผลต่างอันดับที่ 1 ($d=1$) และผลต่างฤดูกาลอันดับที่ 1 ($D=1$) เพื่อขจัดแนวโน้มและความผันแปรตามฤดูกาล ซึ่งผลลัพธ์จากการแปลง (รูปที่ 4 และ 5) แสดงให้เห็นว่าอนุกรมเวลามีลักษณะคงที่มากยิ่งขึ้นและเหมาะสมต่อการนำไปสร้างแบบจำลองเชิงอนุกรมเวลา หลังจากกำหนดรูปแบบที่เหมาะสมและคัดเลือกแบบจำลองที่ดี 3 ตัวแบบ โดยใช้เกณฑ์ AIC เป็นการคัดเลือกจึงได้ตัวแบบดังนี้ SARIMA (1,1,0)(0,1,1)₁₂, SARIMA (0,1,1)(2,1,0)₁₂ และ SARIMA (0,1,1)(3,1,0)₁₂ ตามลำดับ ดังผลวิเคราะห์ในตารางที่ 3



รูปที่ 3 กราฟ ACF และ PACF ของฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ เมื่อแปลงข้อมูลในรูป $\ln(Y)$



รูปที่ 4 กราฟ ACF และ PACF ของฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ เมื่อแปลงข้อมูลในรูป $\ln(Y)$ และทำการหาผลต่างและผลต่างฤดูกาลอันดับที่ 1



รูปที่ 5 การเคลื่อนไหวของฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ รายเดือนตั้งแต่พ.ศ. 2556 – 2565
เมื่อแปลงข้อมูลในรูป $\ln(Y)$ และทำการหาผลต่างและผลต่างฤดูกาลอันดับที่ 1

ตารางที่ 3 ค่าประมาณพารามิเตอร์ของแบบจำลองบอกซ์-เจนกินส์ เมื่อแปลงข้อมูลในรูป $\ln(Y)$

Model	Type	Parameter Estimation			
		Estimate	SE	t-test	P-Value
SARIMA	AR (1)	-0.3426	0.0921	-3.720	<0.001
$(1,1,0)(0,1,1)_{12}$	SMA (12)	0.8539	0.0741	11.521	<0.001
Ljung-Box Q (at Lag 18th) = 24.54 (P-Value = 0.078), KS-Statistic = 0.073 (P-Value = 0.618)					
Levene-Statistic = 1.658 (P-Value = 0.095, t-test = 0.21 ($H_0: \mu_e=0$, P-Value = 0.836))					
SARIMA	SAR (12)	-0.5092	0.0864	-5.894	<0.001
$(0,1,1)(2,1,0)_{12}$	SAR (24)	-0.7132	0.0905	-7.885	<0.001
	MA (1)	0.5197	0.0850	6.114	<0.001
Ljung-Box Q (at Lag 18th) = 20.51 (P-Value = 0.153), KS-Statistic = 0.069 (P-Value = 0.681)					
Levene-Statistic = 1.745 (P-Value = 0.075), t-test = 0.043 ($H_0: \mu_e=0$, P-Value = 0.966)					
SARIMA	SAR (12)	-0.6095	0.1028	-5.926	<0.001
$(0,1,1)(3,1,0)_{12}$	SAR (24)	-0.7862	0.0977	-8.046	<0.001
	SAR (36)	-0.2895	0.1124	-2.576	0.011
	MA (1)	0.4782	0.0887	5.390	<0.001
Ljung-Box Q (at Lag 18th) = 20.13 (P-Value = 0.126), KS-Statistic = 0.055 (P-Value = 0.980)					
Levene-Statistic = 1.199 (P-value = 0.298), t-test = 0.215 ($H_0: \mu_e=0$, P-Value = 0.830)					

จากการวิเคราะห์แบบจำลอง SARIMA ในตารางที่ 3 ได้ตัวแบบที่เหมาะสมที่สุด 3 ตัวแบบ ได้แก่ SARIMA (1,1,0)(0,1,1)₁₂ SARIMA (0,1,1)(2,1,0)₁₂ และ SARIMA (0,1,1)(3,1,0)₁₂ ผลวิเคราะห์การประมาณค่าพารามิเตอร์แสดงให้เห็นว่าทั้ง 3 ตัวแบบ ให้ค่าพารามิเตอร์ที่มีนัยสำคัญทางสถิติ (P-Value < 0.01) เพื่อประเมินความเหมาะสมของแบบจำลอง จึงได้ดำเนินการตรวจสอบข้อสมมติของค่าความคลาดเคลื่อนในหลายด้านที่ระดับนัยสำคัญ 0.01 พบว่า ผลการทดสอบ Ljung-Box ที่ lag 18 ของทั้งสามตัว แสดงว่าค่าความคลาดเคลื่อนไม่มีอัตสัมพันธ์เหลืออยู่ มีความเป็นอิสระกัน (P-Value > 0.01) นอกจากนี้ผลวิเคราะห์ KS-Test บ่งชี้ว่า มีการแจกแจงใกล้เคียงกับการแจกแจงปกติ (P-Value > 0.01) อีกทั้งยังผ่านการทดสอบความแปรปรวนคงที่ด้วย Levene's Test (P-Value > 0.01) และค่าเฉลี่ยไม่แตกต่างจากศูนย์ (P-Value > 0.01) ดังนั้นแบบจำลองบอกซ์-เจนกินส์มีความเหมาะสมทั้งในด้านความสามารถในการอธิบายโครงสร้างของข้อมูล และการคงสมมติฐานของแบบจำลองอนุกรมเวลา จึงสามารถนำไปประยุกต์ใช้ในการพยากรณ์ข้อมูลในอนาคต

5.2 แบบจำลองการปรับให้เรียบด้วยวิธีของวินเทอร์

จากผลวิเคราะห์ในตารางที่ 4 พบว่า แบบจำลองการปรับให้เรียบด้วยวิธีของวินเทอร์ ภายหลังการแปลงข้อมูลให้อยู่ในรูป $\ln(Y)$ ได้ค่าประมาณพารามิเตอร์ค่าคงที่ปรับให้เรียบสำหรับระดับ (α) เท่ากับ 0.4289 (P-Value < 0.01) ค่าคงที่ปรับให้เรียบสำหรับฤดูกาล (δ) เท่ากับ 0.3377 (P-Value < 0.01) แสดงว่าทั้งองค์ประกอบแนวโน้มและฤดูกาลมีบทบาทในการอธิบายความเคลื่อนไหวของข้อมูลอย่างมีนัยสำคัญที่ระดับ 0.01 และได้ค่าคงที่ปรับให้เรียบสำหรับความชัน (γ) เท่ากับ 0.0005 (P-Value > 0.01) แสดงให้เห็นว่าไม่มีการเปลี่ยนแปลงของอัตราการเพิ่มหรือลดในข้อมูล แม้ว่าพารามิเตอร์บางตัวจะไม่มีนัยสำคัญในระดับที่กำหนด แต่ภาพรวมของแบบจำลองยังคงแสดงถึงความเหมาะสมในระดับหนึ่ง ทั้งนี้จากการตรวจสอบสมมติฐานของแบบจำลองที่ระดับนัยสำคัญ 0.01 พบว่า การทดสอบ Ljung-Box ที่ Lag 18 ให้ค่า Q = 26.738 (P-Value = 0.031) กล่าวคือแบบจำลองไม่มีปัญหาในเรื่องความเป็นอิสระของค่าความคลาดเคลื่อน และพบว่าค่าความคลาดเคลื่อนมีการแจกแจงใกล้เคียงกับการแจกแจงปกติ โดยให้ค่า KS = 0.063 (P-Value = 0.728) ขณะที่ผลการทดสอบความแปรปรวนคงที่ด้วย Levene's Test (P-Value = 0.080) และค่าเฉลี่ยของค่าความคลาดเคลื่อนเท่ากับศูนย์ (t = 0.215, P-Value = 0.830)

ตารางที่ 4 ค่าประมาณพารามิเตอร์ของแบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ เมื่อแปลงข้อมูลในรูป $\ln(Y)$

Model	Parameter	Parameter Estimation			
		Estimate	SE	t-test	P-Value
Winter	α	0.4289	0.0671	6.391	<0.001
	γ	0.0005	0.0170	0.032	0.979
	δ	0.3377	0.0883	3.824	<0.001
Ljung-Box Q (at Lag 18 th) = 26.738 (P-Value = 0.031), KS-Statistic = 0.063 (P-Value = 0.728)					
Levene-Statistic = 1.713 (P-Value = 0.080), t-test = 0.215 (H ₀ : $\mu_e=0$, P-Value = 0.830)					

5.3 แบบจำลองด้วยวิธีการถดถอยเชิงเส้น

ผลการวิเคราะห์แบบจำลองด้วยวิธีการถดถอยเชิงเส้นภายหลังการแปลงข้อมูลในรูป $\ln(Y)$ ในตารางที่ 5 แสดงให้เห็นว่าตัวแปรอิสระสามารถสามารถอธิบายและทำนายค่าฝุ่นละออง PM2.5 ได้อย่างมีนัยสำคัญ ($F = 71.077$, $p\text{-value} < 0.001$) และแบบจำลองสามารถอธิบายความผันแปรของค่าฝุ่นละออง PM2.5 ได้ถึง 87.6% ($R^2 = 0.876$) ได้แบบจำลองดังนี้

$$\ln(Y_t) = 3.6252 - 0.0052t + 0.1244M_{1t} + 0.4367M_{2t} + 0.9652M_{3t} + 0.6201M_{4t} - 0.1866M_{5t} \\ - 0.7555M_{6t} - 0.8133M_{7t} - 0.7911M_{8t} - 0.8394M_{9t} - 0.5860M_{10t} - 0.3619M_{11t}$$

เมื่อกำหนดให้ t คือ คาบเวลาที่ 1, 2, 3, ... ($t=1$ คือ เดือนมกราคม พ.ศ. 2556) และ

$$M_{it} = \begin{cases} 1 & \text{เมื่อข้อมูลอนุกรมเวลา ณ เวลา } t \text{ เกิดในเดือนที่ } i; i = 1, 2, \dots, 11 \\ 0 & \text{เมื่อข้อมูลอนุกรมเวลา ณ เวลา } t \text{ ไม่ได้เกิดในเดือนที่ } i; i = 1, 2, \dots, 11 \end{cases}$$

ตารางที่ 5 ค่าประมาณพารามิเตอร์ของแบบจำลองด้วยวิธีการถดถอยเชิงเส้น เมื่อแปลงข้อมูลในรูป $\ln(Y)$

Model	Estimate	SE	t-test	P-Value	VIF
Constant	3.6252	0.0851	42.609	<0.001	-
t	-0.0052	0.0006	-8.301	<0.001	1.010
M_{1t}	0.1244	0.1055	1.179	0.241	1.841
M_{2t}	0.4367	0.1055	4.140	<0.001	1.840
M_{3t}	0.9652	0.1055	9.152	<0.001	1.839
M_{4t}	0.6201	0.1054	5.882	<0.001	1.837
M_{5t}	-0.1866	0.1054	-1.771	0.079	1.836
M_{6t}	-0.7555	0.1054	-7.169	<0.001	1.836
M_{7t}	-0.8133	0.1054	-7.719	<0.001	1.835
M_{8t}	-0.7911	0.1053	-7.510	<0.001	1.834
M_{9t}	-0.8394	0.1053	-7.969	<0.001	1.834
M_{10t}	-0.5860	0.1053	-5.564	<0.001	1.834
M_{11t}	-0.3619	0.1053	-3.436	<0.001	1.833

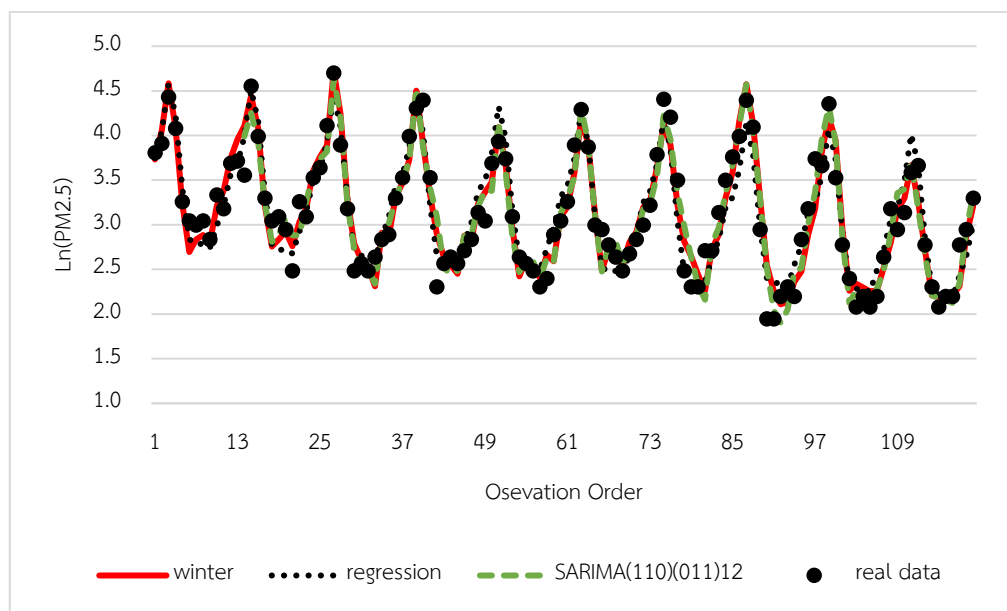
$R = 0.889$, $R^2 = 0.876$, $F = 71.077$ (P-Value <0.001), KS-Statistic = 0.074 (P-Value = 0.109)

Levene-Statistic = 0.717 (P-Value = 0.720), Durbin -Watson = 1.100 (P-Value <0.001)

t-test = -3.177e-16 ($H_0: \mu_e=0$, P-Value = 0.999)

จากผลวิเคราะห์พบว่า สัมประสิทธิ์การถดถอยของ $t=-0.0052$ (P-Value <0.001) แสดงถึงแนวโน้มลดลงของระดับฝุ่นละออง PM2.5 เมื่อเวลาผ่านไป ขณะที่ตัวแปรฤดูกาลในช่วงเดือนกุมภาพันธ์-เมษายน มีค่าสัมประสิทธิ์เป็นบวกแสดงให้เห็นถึงอิทธิพลของฤดูกาลที่มีต่อระดับฝุ่นละออง PM2.5 อย่างมีนัยสำคัญ โดยส่งผลให้ค่าฝุ่นละอองสูงขึ้นอย่างชัดเจนเมื่อเทียบกับช่วงเดือนอื่นของปี เพื่อประเมินความเหมาะสมของแบบจำลองถดถอยเชิงเส้น ได้ตรวจสอบสมมติฐานสำคัญ พบว่า ค่าความคลาดเคลื่อนมีการแจกแจงใกล้เคียงปกติ (KS=0.074, p-value =

0.109), มีความแปรปรวนคงที่ (Levene=0.717, P-Value = 0.720) มีค่าเฉลี่ยไม่แตกต่างจากศูนย์ ($t=-3.177e-16$, P-Value = 0.999) อย่างไรก็ตามจากผลการวิเคราะห์พบว่า ค่าความคลาดเคลื่อนมีความสัมพันธ์กันเชิงบวก (Durbin Watson = 1.100, P-Value <0.001) แบบจำลองนี้จึงยังไม่เหมาะสมตามหลักของการถดถอยเชิงเส้น



รูปที่ 6 ค่าพยากรณ์ฝุ่นละออง PM2.5 ในช่วงปี พ.ศ. 2556 - 2567

รูปที่ 6 แสดงผลการพยากรณ์ค่าฝุ่นละออง PM2.5 รายเดือนในช่วงปี พ.ศ. 2556 ถึง 2565 โดยใช้ข้อมูลจากชุดฝึกฝน (Train Data Set) เพื่อพัฒนาแบบจำลองทางสถิติ จากผลการวิเคราะห์พบว่า แบบจำลองทั้งสามแบบที่นำมาเปรียบเทียบให้ผลการพยากรณ์ที่ใกล้เคียงกัน แต่แต่ละเทคนิคมีความสามารถในการจับแนวโน้มและรูปแบบของข้อมูลได้ในระดับที่สอดคล้องกัน เพื่อให้สามารถระบุแบบจำลองที่มีประสิทธิภาพสูงสุดและเหมาะสมต่อการนำไปใช้ในเชิงปฏิบัติ จึงได้ดำเนินการประเมินประสิทธิภาพเชิงเปรียบเทียบของแต่ละแบบจำลองในขั้นตอนถัดไป

5.4 การประเมินประสิทธิภาพของแบบจำลอง

เมื่อได้แบบจำลองการพยากรณ์ค่าฝุ่นละออง PM2.5 บนสเกลของลอการิทึมธรรมชาติแล้ว การแปลงค่าพยากรณ์กลับสู่สเกลดั้งเดิมเพื่อให้สามารถตีความได้ในบริบทของความเข้มข้นของมลพิษในอากาศ จำเป็นต้องคำนึงถึงความเอนเอียงที่เกิดจากการแปลงกลับด้วยวิธี Smearing Estimate [15] โดยมีหลักการคำนวณดังนี้: ให้ y_t แทนค่าพยากรณ์ของ $\ln(Y_t)$ และให้ θ แทนค่า Smearing Factor

$$\theta = \frac{1}{n} \sum_{t=1}^n \exp(e_t) \quad \text{โดยที่ } e_t = \ln(Y_t) - y_t$$

ดังนั้นค่าพยากรณ์เมื่อแปลงกลับมาจะประมาณได้จาก

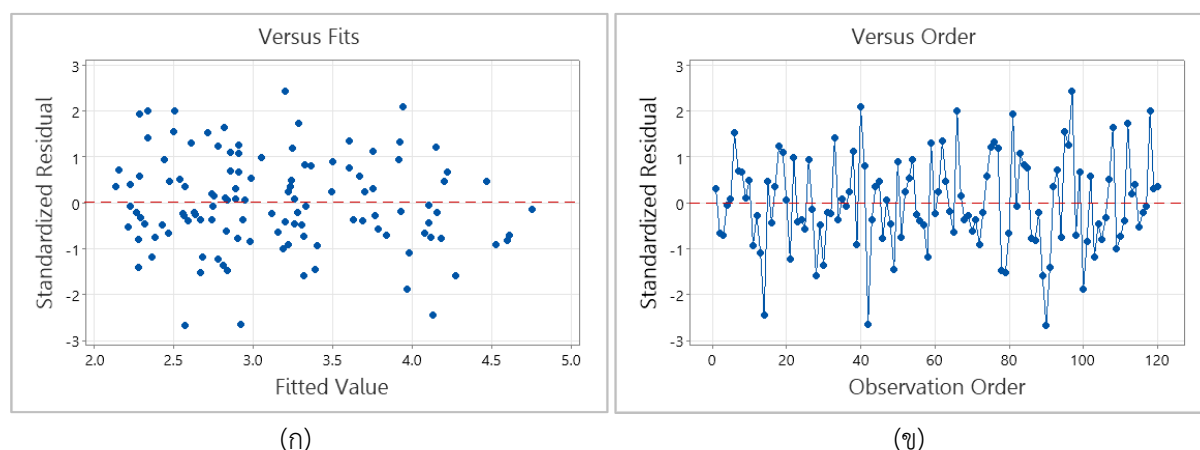
$$Y_t = \theta \exp(\hat{y}_t)$$

กระบวนการดังกล่าวช่วยให้ค่าพยากรณ์ที่แปลงกลับมีความแม่นยำและลดความเอนเอียง และไม่จำเป็นต้องสมมติให้ค่าความคลาดเคลื่อนมีการแจกแจงปกติ ในการประเมินประสิทธิภาพของแบบจำลอง ได้ดำเนินการเปรียบเทียบค่าพยากรณ์ฝุ่นละออง PM2.5 รายเดือนในช่วงปี พ.ศ. 2566 ถึง 2567 โดยใช้ข้อมูลชุดทดสอบ (Test Data Set) เพื่อทดสอบความแม่นยำของแบบจำลองที่ได้จากการฝึกด้วยข้อมูลในอดีต หลังการแปลงค่าพยากรณ์ที่อยู่ในรูปของลอการิทึมธรรมชาติกลับสู่หน่วยจริง โดยใช้วิธี Smearing Estimate พบว่า แบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ ให้ผลการพยากรณ์ที่แม่นยำที่สุด โดยมีค่า MAPE เท่ากับ 19.15 รองลงมาคือ ARIMA (0,1,1)(0,1,1)₁₂ (MAPE = 22.23) และวิธีการถดถอยเชิงเส้น (MAPE = 22.44) ตามลำดับ ดังแสดงในตารางที่ 6

ตารางที่ 6 ผลวิเคราะห์ความแม่นยำของแบบจำลอง

ตัวแบบ	บอกซ์-เจนกินส์			วินเทอร์	การถดถอย
	ARIMA (1,1,0)(0,1,1) ₁₂	ARIMA (0,1,1)(2,1,0) ₁₂	ARIMA (0,1,1)(3,1,0) ₁₂		
MAPE	22.23	22.97	22.78	19.15	24.44

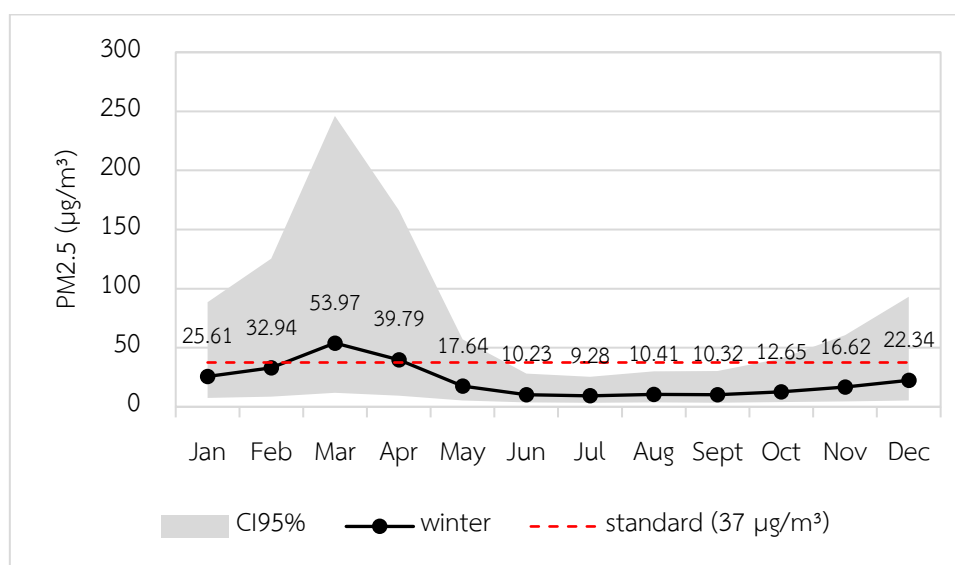
ผลการวิเคราะห์นี้บ่งชี้ว่าแบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ สามารถจับแนวโน้มและรูปแบบของข้อมูลฝุ่นละออง PM2.5 ได้ผลลัพธ์ที่ดีที่สุดเมื่อเทียบกับแบบจำลองอื่น ๆ อย่างไรก็ตามแม้ว่าแบบจำลองด้วยวิธีบอกซ์-เจนกินส์จะมีความแม่นยำรองลงมา แต่ก็ยังถือว่ามีประสิทธิภาพใกล้เคียงกันยังคงให้ผลลัพธ์ที่อยู่ในเกณฑ์ยอมรับได้ในทางกลับกันแบบจำลองด้วยวิธีการถดถอยเชิงเส้นมีค่า MAPE สูงที่สุด ซึ่งอาจสะท้อนข้อจำกัดในการอธิบายตัวแปรแฝงหรือปัจจัยภายนอกที่ส่งผลต่อปริมาณฝุ่นละออง PM2.5 ได้ไม่ครบถ้วน



รูปที่ 7 แผนภาพการกระจายของค่าความคลาดเคลื่อน เมื่อพยากรณ์พยากรณ์ด้วยแบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ ($\alpha = 0.4289$, $\gamma = 0.0005$, $\delta = 0.3377$) เมื่อแปลงข้อมูลในรูป $\ln(Y)$

รูปที่ 7 (ก) แสดงแผนภาพการกระจายระหว่างค่าคลาดเคลื่อนกับค่าพยากรณ์ ซึ่งพบการกระจายอย่างสม่ำเสมอของค่าคลาดเคลื่อน มีความคงที่ของความแปรปรวน ในขณะที่รูปที่ 7 (ข) แสดงแผนภาพการกระจายระหว่างค่าคลาดเคลื่อนกับลำดับเวลา พบว่าค่าคลาดเคลื่อนกระจายอย่างเป็นอิสระรอบค่าศูนย์ในทุกช่วงเวลา สะท้อน

ถึงลักษณะของความเป็นอิสระและไม่มีรูปแบบที่ชัดเจนของความคลาดเคลื่อน ซึ่งสอดคล้องกับข้อสมมติของแบบจำลองที่ดี



รูปที่ 8 การพยากรณ์ค่าฝุ่นละออง PM2.5 ปี พ.ศ. 2568 โดยใช้แบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ ($\alpha = 0.4289$, $\gamma = 0.0005$, $\delta = 0.3377$)

เมื่อประยุกต์ใช้แบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ในการพยากรณ์ค่าฝุ่นละออง PM2.5 สำหรับปี พ.ศ. 2568 พบว่าค่าพยากรณ์รายเดือนหลังจากการแปลงกลับเป็นค่าปกติได้แสดงผลค่าฝุ่นละออง PM2.5 ตามที่ปรากฏในรูปที่ 8 ซึ่งในช่วงเดือนมีนาคมและเมษายน ค่าฝุ่นละออง PM2.5 ที่ทำนายออกมาสูงกว่าเกณฑ์ที่กรมควบคุมมลพิษตั้งไว้ (≤ 37.5 มคก./ลบ.ม.) ซึ่งมีภาวะเสี่ยงต่อโรคและการเจ็บป่วยของประชาชน

6. สรุป

แบบจำลองที่มีประสิทธิภาพสามารถคาดการณ์ปริมาณฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ได้แม่นยำสูงสุด คือแบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ ซึ่งให้ค่าคลาดเคลื่อนจากข้อมูลจริงน้อยกว่าวิธีบ็อกซ์-เจนกินส์ และวิธีการถดถอยเชิงเส้น ทั้งนี้ผลการวิเคราะห์สอดคล้องกับการศึกษาของอินทนิตร สุขเกษม [16] ซึ่งเสนอว่าแบบจำลองการปรับให้เรียบมีความเหมาะสมในการพยากรณ์ค่าฝุ่น PM2.5 รายชั่วโมงและข้อมูลเฉลี่ย 24 ชั่วโมงในเขตเทศบาลนครนครราชสีมา ผลลัพธ์นี้ชี้ให้เห็นว่าแบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์สามารถจับลักษณะเฉพาะของข้อมูล PM2.5 ได้ดีกว่า โดยเฉพาะในช่วงฤดูร้อนของทุกปี ซึ่งค่าฝุ่นละออง PM2.5 มักพุ่งสูงกว่ามาตรฐานที่กำหนดไว้ วิธีวินเทอร์สามารถปรับน้ำหนักของข้อมูลในแต่ละช่วงเวลาและฤดูกาลได้อย่างยืดหยุ่น มีความได้เปรียบมากกว่าวิธีอื่น โดยเฉพาะเมื่อเปรียบเทียบกับวิธีถดถอยเชิงเส้น ที่มีข้อจำกัดในการอธิบายความผันแปรตามฤดูกาลผ่านตัวแปรหุ่นจำลอง ซึ่งอาจไม่สะท้อน พฤติกรรมจริงของข้อมูลได้ครบถ้วน ในขณะที่แบบจำลองด้วยวิธีบ็อกซ์-เจนกินส์ให้การทำนายที่แม่นยำใกล้เคียงกับวิธีวินเทอร์ แต่จุดอ่อนคือวิธีความซับซ้อนของการคำนวณกว่าวิธีอื่น ๆ

7. ข้อเสนอแนะ

ในการศึกษาครั้งนี้ ได้ใช้ แบบจำลองการปรับให้เรียบด้วยวิธีวินเทอร์ ในการพยากรณ์ค่าฝุ่นละออง PM2.5 ในจังหวัดเชียงใหม่ โดยพบว่าพารามิเตอร์ที่ใช้สำหรับการปรับแนวโน้มในแบบจำลองนั้นไม่มีนัยสำคัญทางสถิติ อย่างไรก็ตาม ภาพรวมของแบบจำลองยังคงสามารถให้ผลการพยากรณ์ค่าฝุ่นละออง PM2.5 ได้อย่างแม่นยำในระดับที่น่าพึงพอใจ อย่างไรก็ตาม มีความเป็นไปได้ที่แบบจำลองอาจเกิดปัญหาจากการ Overfitting ซึ่งอาจทำให้การทำนายในข้อมูลใหม่ไม่แม่นยำเท่าที่ควร เพื่อปรับปรุงประสิทธิภาพของแบบจำลองในอนาคต ควรพิจารณาเพิ่มจำนวนข้อมูลฝึกหัด (Train Data Set) เพื่อให้แบบจำลองได้เรียนรู้จากข้อมูลที่หลากหลายและครอบคลุมมากยิ่งขึ้น ซึ่งจะช่วยลดความเสี่ยงจากการ Overfitting และเพิ่มความแม่นยำในการทำนาย นอกจากนี้ยังควรทบทวนและตรวจสอบข้อมูลที่อาจเป็นค่าผิดปกติ (Outliers) หรือข้อมูลที่ขาดหาย (Missing Data) ซึ่งอาจส่งผลกระทบต่อความแม่นยำของผลลัพธ์การพยากรณ์ อีกหนึ่งแนวทางที่สามารถนำมาใช้ในการปรับปรุงแบบจำลองคือ การพิจารณาสร้างแบบจำลองในระดับรายวัน ซึ่งจะช่วยเพิ่มจำนวนข้อมูลและความละเอียดในการทำนาย รวมทั้งช่วยให้สามารถจับลักษณะการเปลี่ยนแปลงที่มีความผันผวนในระยะสั้นได้ดียิ่งขึ้น ทั้งนี้ การปรับปรุงดังกล่าวจะช่วยให้แบบจำลองมีความเสถียรและสามารถทำนายค่าฝุ่นละออง PM2.5 ได้แม่นยำยิ่งขึ้นในอนาคต

สำหรับการพัฒนาแบบจำลองอื่น ๆ ในอนาคตสำหรับฝุ่นละออง PM2.5 ในเชียงใหม่ อาจพิจารณาการใช้ตัวแบบพยากรณ์แบบผสม (Hybrid Model) เพื่อเพิ่มความแม่นยำและความยืดหยุ่นในการพยากรณ์ เนื่องจากค่าฝุ่นละออง PM2.5 ได้รับอิทธิพลจากหลายปัจจัย การเพิ่มประสิทธิภาพในการพยากรณ์สามารถทำได้โดยการนำข้อมูลด้านสิ่งแวดล้อมเข้ามาประกอบ เช่น อุณหภูมิ ความชื้น ทิศทางและความเร็วลม จุดความร้อน (Hotspots) รวมถึงข้อมูลจากประเทศเพื่อนบ้านที่อาจส่งผลกระทบต่อมลภาวะในภาคเหนือ นอกจากนี้การนำเทคนิค Machine Learning ควรถูกนำมาเปรียบเทียบกับวิธีการทางสถิติแบบดั้งเดิม เพื่อประเมินความแม่นยำของการพยากรณ์ ตลอดจนการพัฒนาแอปพลิเคชันหรือระบบแจ้งเตือนล่วงหน้า เพื่อสนับสนุนการวางแผนนโยบายและการดำเนินการของหน่วยงานที่เกี่ยวข้อง ตลอดจนเพิ่มขีดความสามารถของประชาชนในการรับมือกับจากฝุ่นละออง PM2.5 สำหรับงานวิจัยในอนาคตควรขยายขอบเขตไปสู่ระดับตำบลหรืออำเภอ เพื่อให้ได้ข้อมูลเชิงลึกสำหรับการจัดการปัญหาในแต่ละพื้นที่ได้ตรงจุด รวมถึงการใช้ ข้อมูลดาวเทียม เพื่อเพิ่มความครอบคลุมและความละเอียดของข้อมูล พร้อมกันนี้ภาครัฐควรพิจารณาดำเนินมาตรการเชิงรุกในช่วงที่ระดับฝุ่นละอองอยู่ในเกณฑ์วิกฤต อาทิ การสนับสนุนการใช้พลังงานสะอาดที่ไม่ส่งผลกระทบต่อสิ่งแวดล้อม การควบคุมยานพาหนะที่มีอัตราการปล่อยมลพิษสูง การบังคับใช้ข้อกำหนดในการจำกัดการเผาในที่โล่ง และการจำกัดกิจกรรมกลางแจ้งในพื้นที่เสี่ยง นอกจากนี้ ควรมีการจัดสรรทรัพยากรด้านสาธารณสุขอย่างมีประสิทธิภาพ เพื่อรองรับภาวะฉุกเฉินจากผลกระทบด้านสุขภาพ

8. เอกสารอ้างอิง

- [1] สำนักงานส่งเสริมและสนับสนุนวิชาการ 9, “PM2.5 ภัยไร้เสียง สถานการณ์ฝุ่นควัน PM2.5 ภาคเหนือตอนบน,” [ออนไลน์]. Available: https://olderlampang.dop.go.th/download/knowledge/th1716527042-130_0.pdf. (เข้าถึงเมื่อ: 10 พฤศจิกายน 2567).
- [2] อิดารัตน์ ผลพิบูลย์, อิศริย์ธิดา ชัยสวัสดิ์ และอนุวัตร รุ่งพิสุทธิพงษ์, “ภัยในหน้าหนาวจากฝุ่นละอองขนาดเล็ก (PM2.5),” วารสารวิชาการมหาวิทยาลัยอีสเทิร์นเอเซีย ฉบับวิทยาศาสตร์และเทคโนโลยี, ปีที่ 8, ฉบับที่ 1, หน้า 40-46, มกราคม – มิถุนายน, 2557.
- [3] กรมประชาสัมพันธ์, “ปัญหาฝุ่น PM2.5 ในภาคเหนือของไทย,” [ออนไลน์]. Available: <https://www.prd.go.th/th/content/category/detail/id/31/iid/340962> (เข้าถึงเมื่อ: 1 พฤศจิกายน 2567).

- [4] สำนักงานสิ่งแวดล้อมและควบคุมมลพิษที่ 1, “สรุปสถานการณ์ฝุ่นละออง PM2.5,” [ออนไลน์]. Available: <https://epo01.pcd.go.th/th/news/detail/173594/>. (เข้าถึงเมื่อ: 1 พฤศจิกายน 2567).
- [5] กรมควบคุมมลพิษ, “สถานการณ์และการจัดการปัญหามลพิษทางอากาศและเสียงของประเทศไทย ปี 2566,” [ออนไลน์]. Available: https://www.pcd.go.th/wp-content/uploads/2024/09/pcdnew-2024-09-23_07-36-26_390200.pdf. (เข้าถึงเมื่อ: 1 พฤษภาคม 2567).
- [6] วัฒนา ชยธวัช, “การพยากรณ์ค่าเฉลี่ยรายเดือนฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมครอนในจังหวัดเชียงใหม่ ด้วยวิธีแยกส่วนประกอบแบบดั้งเดิม,” *วารสารวิทยาศาสตร์และเทคโนโลยีชุมชน*, ปีที่ 2, ฉบับที่ 5, หน้า 16-30, กันยายน – ตุลาคม, 2567.
- [7] ปรัชญา สิงหรวงศ์, “การพยากรณ์ฝุ่น PM2.5 ล่วงหน้าในจังหวัดเชียงใหม่ โดยใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว,” *สารนิพนธ์วิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยศรีนครินทรวิโรฒ*, กรุงเทพมหานคร, 2563.
- [8] R. J. Hyndman, and G. Athanasopoulos, *Forecasting: principles and practice*, 3rd edition. Melbourne, Australia: OTexts, 2021.
- [9] สุมิตรา เมืองขวา, “การพยากรณ์อนุกรมเวลาโดยวิธีบ็อกซ์-เจนกินส์,” *วารสารและวารสารเวชศาสตร์เขตเมือง*, ปีที่ 63 ฉบับเพิ่มเติม, หน้า 185 – 192, กรกฎาคม 2562.
- [10] รณชัย ชื่นธวัช, “การพยากรณ์ค่าความเข้มข้นของฝุ่นละอองขนาดเล็กไม่เกิน 10 ไมครอนด้วยเครือข่ายประสาทเทียมร่วมกับรูปแบบออโตรีเกรสซีฟ,” *วารสารวิทยาศาสตร์บูรพา*, ปีที่ 22, ฉบับที่ 3, หน้า 519 – 537, กันยายน – ธันวาคม, 2560.
- [11] วรางคณา เรียนสุทธิ, “การคัดเลือกตัวแบบพยากรณ์ราคามะม่วงเขียวเสวย,” *วารสารศรีนครินทรวิโรฒวิจัยและพัฒนา สาขามนุษยศาสตร์และสังคมศาสตร์*, ปีที่ 12, ฉบับที่ 23, หน้า 52-62, มกราคม - มิถุนายน, 2563
- [12] C. Wongoutong, “Improvement of the Holt-Winters Multiplicative Method with a New Initial Value Settings Method,” *Thailand Statistician*, vol. 19, no.2, pp. 280–293, April, 2021.
- [13] ศิริลักษณ์ สุวรรณวงศ์, “เทคนิคการพยากรณ์เชิงปริมาณ: การวิเคราะห์อนุกรมเวลา,” พิมพ์ครั้งที่ 1. นครปฐม: มหาวิทยาลัยมหิดล, 2556.
- [14] เปรมพร เขมาวุฒม์ และ ชมพร ชาวบางแก้ว, “การศึกษาค่าปรับเรียบที่เหมาะสมสำหรับข้อมูลที่ไม่มีแนวโน้มและไม่มีฤดูกาล,” *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 31, ฉบับที่ 2, หน้า 245 – 256, เมษายน - มิถุนายน, 2564.
- [15] N. Duan, “Smearing estimate: a nonparametric retransformation method,” *Journal of the American Statistical Association*, vol. 78, no.383, pp. 605-610, September, 1983.
- [16] อินทร์ฉัตร สุขเกษม, “การพยากรณ์ปริมาณฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมครอน (PM2.5) จังหวัดนครราชสีมา,” รายงานการวิจัย, สำนักงานป้องกันควบคุมโรคที่ 9, กรมควบคุมโรค, นครราชสีมา, 2563.