

การเปรียบเทียบวิธีพยากรณ์ของอนุกรมเวลารูปแบบตามฤดูกาลที่มีเลขศูนย์จำนวนมาก

ธรรศกรณ์ เศวตสุทธิพันธ์^{*1} และ พีรยุทธ์ ชาญเศรษฐิกุล²

มหาวิทยาลัยเกษตรศาสตร์ 50 ถนนงามวงศ์วาน แขวงลาดยาว เขตจตุจักร กรุงเทพฯ 10900

Received: 20 April 2021; Revised: 7 June 2021; Accepted: 14 June 2021

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์อนุกรมเวลาของข้อมูลจำนวนพื้นที่การเกิดไฟป่ารายวันในจังหวัดเชียงใหม่ซึ่งจัดอยู่ในประเภทของอนุกรมเวลารูปแบบตามฤดูกาล และข้อมูลยังมีค่าที่เป็นศูนย์เป็นช่วงกว้างในหลายช่วงเวลาซึ่งอาจส่งผลให้มีความแปรปรวนที่สูง โดยข้อมูลที่ใช้ประกอบด้วย 2 ส่วน ได้แก่ จำนวนพื้นที่การเกิดไฟป่า และสภาพภูมิอากาศซึ่งนำมาใช้เป็นปัจจัยที่อาจส่งผลต่อจำนวนพื้นที่ไฟป่า และนำข้อมูลที่ได้มาสร้างตัวแบบจำลองพยากรณ์ที่เหมาะสม โดยจะเปรียบเทียบวิธีพยากรณ์ทั้งหมด 6 วิธี ได้แก่ การพยากรณ์ด้วยสมการถดถอยแบบพหุคูณที่มีตัวแปรจัดประเภท, การพยากรณ์ด้วยสมการถดถอยแบบแบบโพลีโนเมียลหลายตัวแปร, การพยากรณ์ด้วยอนุกรมฟูเรียร์แบบตัดทอน, การพยากรณ์ด้วยวิธี Holt-Winters แบบผลบวก, การพยากรณ์ตามแนวคิดของบ็อกซ์-เจนกินส์ด้วยวิธี SARIMAX, และการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียม เพื่อวัดประสิทธิภาพของตัวแบบจำลองจะแบ่งข้อมูลสำหรับทดสอบออกเป็น 4 ช่วง คือ 3 เดือน, 6 เดือน, 1 ปี และ 1 ปี 9 เดือน จากนั้นเปรียบเทียบด้วยตัวแบบพยากรณ์ทั้ง 6 แบบด้วยค่าคลาดเคลื่อน RMSE ซึ่งผลการทดลองพบว่า การพยากรณ์ด้วยสมการถดถอยแบบพหุคูณแบบมีตัวแปรจัดประเภทให้ค่า RMSE ต่ำสุดสำหรับข้อมูลทดสอบในช่วงเวลา 3 เดือน, 6 เดือน และ 1 ปี และ การพยากรณ์ด้วยวิธี SARIMAX ให้ค่า RMSE ต่ำสุดสำหรับข้อมูลทดสอบในช่วงเวลา 1 ปี 9 เดือน

คำสำคัญ: ข้อมูลอนุกรมเวลา, การพยากรณ์, ไฟป่า

* Corresponding author. E-mail: thassakorn.sa@ku.th

¹ นักศึกษาปริญญาโท ภาควิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

² รองศาสตราจารย์ ภาควิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

A Comparison of Forecasting Methods for Seasonal Time-series with Many Zeros

Thassakorn Sawetsuthipan^{*1} and Peerayuth Charnsethikul²

Kasetsart University, 50 Ngamwongwan Rd., Lad Yao, Chatuchak, Bangkok, 10900, Thailand

Received: 20 April 2021; Revised: 7 June 2021; Accepted: 14 June 2021

Abstract

The research aims to analyze the time series of daily wildfires area in Chiang Mai which are classified as seasonality, and the data also has a wide range of zero values over a number of periods which can result in high variance. The data used consists of two parts: the number of wildfires area and the climate which can be used as a factor that can affect the number of wildfire area. The obtained data used to create the appropriate forecasting model by comparing the six forecasting methods which are multiple regression with categorical variable, multivariate polynomial regression, the truncated fourier series, Holt-Winters's additive method, SARIMAX methods by Box-Jenkins, and artificial neural network; in addition, to measure the efficiency of the model. The test data is divided into four phases: 3 months, 6 months, 1 year and 1 year 6 months, and compare with the six forecast methods by root mean square error (RMSE). The results showed that the Multiple Regression with Categorical Variable provided the lowest RMSE values for test data over a period of 3 months, 6 months, and 1 year, and SARIMAX method provided the lowest RMSE values for test data over a period of 1 year and 9 months.

Keywords: time series data, forecasting, wildfires

* Corresponding author. E-mail: thassakorn.sa@ku.th

¹ Graduate Student in Faculty of engineering, Kasetsart University

² Associate Professor in Faculty of engineering, Kasetsart University

1. คำนำ

“ไฟป่า” จัดเป็นภัยพิบัติอย่างหนึ่งที่มีสาเหตุการเกิดได้ทั้งจากธรรมชาติ และมนุษย์ ซึ่งมักเกิดขึ้นบริเวณทางตอนบนของประเทศ เช่น ภาคเหนือและภาคตะวันออกเฉียงเหนือ โดยจะเกิดในช่วงฤดูแล้งระหว่างช่วงเดือนกุมภาพันธ์ถึงเดือนพฤษภาคม [1] ซึ่งในบางเดือนอาจจะไม่พบการเกิดไฟป่าตลอดทั้งเดือน หรือเป็นเวลาหลายเดือน ส่งผลให้การเกิดไฟป่าที่จัดเป็นข้อมูลอนุกรมเวลาที่มีรูปแบบตามฤดูกาลที่มีเลขศูนย์จำนวนมาก การพยากรณ์สถานการณ์ไฟป่าในปัจจุบันจะใช้ปัจจัยสภาพอากาศประกอบกับแผนที่แสดงจุดความร้อน (Hotspot Maps) ซึ่งเป็นการติดตามสถานการณ์จากดาวเทียมAqua โดยระบบ MODIS โดยในช่วงเดือนกุมภาพันธ์ พื้นที่ทางภาคเหนือจัดอยู่ในกลุ่มที่มีดัชนีการเกิดไฟป่า (Fire Weather Index; FWI) ในระดับสูง ไปจนถึงรุนแรง ปัจจัยหลักที่มีผลต่อพฤติกรรมไฟป่าที่ประกอบด้วย ลักษณะเชื้อเพลิง ลักษณะอากาศ และลักษณะภูมิประเทศ [2] จากข้อมูลสถิติการเกิดไฟป่าของกรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช ส่วนควบคุมไฟป่า ตั้งแต่วันที่ 1 มกราคม พ.ศ. 2552 ถึง วันที่ 10 กันยายน พ.ศ. 2562 พบว่าภาคเหนือมีจำนวนการเกิดไฟป่าถึง 37,306 ครั้ง ซึ่งคิดเป็น 67.04 % ของอัตราการการเกิดไฟป่าในประเทศไทย โดยจังหวัดที่มีการเกิดไฟป่ามากที่สุดคือ จังหวัดเชียงใหม่ ซึ่งมีจำนวนการเกิดไฟป่าสูงถึง 14,306 ครั้ง คิดเป็น 25.71 % ของอัตราการเกิดไฟป่าในประเทศไทย

งานวิจัยนี้จะแบ่งข้อมูลอนุกรมเวลาสำหรับสร้างตัวแบบจำลองของวิธีพยากรณ์ แบ่งออกเป็น 2 ประเภทตามข้อมูลที่ใช้ ได้แก่

1) ข้อมูลอนุกรมเวลาที่ไม่ใช่ปัจจัยทางสภาพอากาศมาเกี่ยวข้อง จะทำการเปรียบเทียบวิธีการพยากรณ์ 2 วิธี คือ วิธีอนุกรมฟูเรียร์แบบตัดตอน (The Truncated Fourier series) และ วิธี Holt-Winters แบบผลบวก (Holt-Winters’s additive method)

2) ข้อมูลอนุกรมเวลาที่ใช้ปัจจัยทางสภาพอากาศมาเกี่ยวข้อง จะทำการเปรียบเทียบวิธีการพยากรณ์ 4 วิธี โดยการใช้การเรียนรู้ของเครื่องจักร คือ วิธีการวิเคราะห์การถดถอยแบบพหุคูณที่มีตัวแปรจัดประเภท (Multiple Regression with Categorical Variables), วิธีการวิเคราะห์การถดถอยแบบโพลีโนเมียลแบบหลายตัวแปร

(Multivariate Polynomial Regression), วิธีการใช้แบบจำลอง SARIMAX (Seasonal Autoregressive Integrated Moving Average with Explanatory Variable) และวิธีโครงข่ายประสาทเทียม (Artificial Neural Network)

2. ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

2.1 วิธี Holt-Winters แบบผลบวก

เป็นวิธีการพยากรณ์ที่นิยมใช้ในการพยากรณ์ข้อมูลอนุกรมเวลาที่มีรูปแบบเป็นฤดูกาล ถูกคิดค้นโดยไฮลท์ วินเทอร์ [3-4] มีการใช้ค่าปรับเรียบ 3 ค่าได้แก่ ค่าสัมประสิทธิ์ของการปรับระดับให้เรียบ (Coefficient for The Level Smoothing; α), ค่าสัมประสิทธิ์ของการปรับแนวโน้มให้เรียบ(Coefficient for The Trend Smoothing; β) และค่าสัมประสิทธิ์ของการปรับฤดูกาลให้เรียบ (Coefficient for The Seasonal Smoothing; γ) โดยมีรูปแบบตามสมการที่ 1 – 5

$$s_0 = x_0 \quad (1)$$

$$s_t = \alpha(x_t - c_{t-L}) + (1 - \alpha)(s_{t-1} + b_{t-1}) \quad (2)$$

$$b_t = \beta(s_t - s_{t-1}) + (1 - \beta)b_{t-1} \quad (3)$$

$$c_t = \gamma(x_t - s_{t-1} - b_{t-1}) + (1 - \gamma)c_{t-L} \quad (4)$$

$$F_{t+m} = s_t + mb_t + c_{t-L+1+(m-1)mod L} \quad (5)$$

เมื่อ s_t คือ ค่าระดับของอนุกรมเวลา ณ เวลา t

b_t คือ ค่าแนวโน้มของอนุกรมเวลา ณ เวลา t

c_t คือ ค่าฤดูกาลของอนุกรมเวลา ณ เวลา t

x_t คือ ค่าอุปสงค์ ณ เวลา t

α, β และ γ คือ ค่าสัมประสิทธิ์การปรับเรียบ โดยที่ $0 \leq \alpha \leq 1, 0 \leq \beta \leq 1$ และ $0 \leq \gamma \leq 1$

L คือ จำนวนคาบฤดูกาลของอนุกรมเวลา

m คือ จำนวนเวลาที่ต้องการพยากรณ์ไปข้างหน้า

F_{t+m} คือ ค่าพยากรณ์ ณ เวลา t+m

ซึ่งในการศึกษาครั้งนี้ใช้วิธี Holt-Winters แบบผลบวกเนื่องจากข้อมูลมีเลขศูนย์เป็นจำนวนมาก หากใช้วิธี Holt-Winters แบบผลคูณ อาจส่งผลให้ไม่สามารถคำนวณค่าสัมประสิทธิ์การปรับเรียบได้

2.2 อนุกรมฟูเรียร์แบบตัดทอน

อนุกรมฟูเรียร์ หรือ อนุกรมฟูรีเยถูกคิดค้นโดย โฌแซ็ฟ ฟูรีเย นิยมใช้กับอนุกรมที่มีลักษณะเป็นสัญญาณที่มีคาบเป็นระยะ ๆ มีลักษณะเป็นอนุกรมอนันต์ ซึ่งอนุกรมฟูเรียร์แบบตัดทอนสามารถใช้ประมาณค่าของฟังก์ชันในช่วงที่จำกัดได้โดยใช้จำนวน N ตัว ซึ่งการวิเคราะห์โดยใช้อนุกรมฟูเรียร์มีความสามารถในการตามข้อมูลรูปแบบปรากฏการณ์ตามฤดูกาลได้ จึงนำลักษณะเฉพาะนี้มาใช้ในการพยากรณ์ โดยมีรูปแบบเป็นฟังก์ชันตรีโกณมิติตามสมการที่ 6

$$f_N(t) = a_0 + \sum_{n=1}^N [a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right)] \quad (6)$$

เมื่อ a_n คือ ค่าสัมประสิทธิ์แทนขนาดของเทอมโคไซน์
 b_n คือ ค่าสัมประสิทธิ์แทนขนาดของเทอมไซน์
 T คือ จำนวนคาบเวลาของสัญญาณ

2.3 วิธีการวิเคราะห์การถดถอย (Regression Analysis)

เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม โดยสร้างสมการทางคณิตศาสตร์ในการพยากรณ์ค่าของตัวแปรตาม ในงานวิจัยนี้ได้ทำการใช้การวิเคราะห์การถดถอยเพื่อใช้ในการศึกษาความสัมพันธ์ของปัจจัยทางสภาพอากาศซึ่งเป็นตัวแปรอิสระ กับจำนวนพื้นที่การเกิดไฟป่าที่เป็นตัวแปรตาม สำหรับใช้ในการพยากรณ์ 2 ประเภท ได้แก่

1) การวิเคราะห์การถดถอยแบบพหุคูณที่มีตัวแปรจัดประเภทเป็นการวิเคราะห์การถดถอยแบบเชิงเส้น โดยศึกษาความสัมพันธ์ของตัวแปรมากกว่า 2 ตัวขึ้นไป มีรูปแบบสมการถดถอยดังสมการที่ 7

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \varepsilon \quad (7)$$

เมื่อ Y คือ ค่าของตัวแปรตาม

β คือ ค่าสัมประสิทธิ์ของสมการถดถอย

X คือ ค่าของตัวแปรอิสระ

2) การวิเคราะห์ด้วยการถดถอยแบบโพลีโนเมียลแบบหลายตัวแปร เป็นการวิเคราะห์การถดถอยแบบเส้นโค้ง โดยศึกษาความสัมพันธ์ของตัวแปรมากกว่า 2 ตัวขึ้นไป มีรูปแบบสมการถดถอยดังสมการที่ 8

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1^2 + \beta_4 X_2^2 + \beta_5 X_1 X_2 + \dots + \varepsilon \quad (8)$$

เมื่อ Y คือ ค่าของตัวแปรตาม

β คือ ค่าสัมประสิทธิ์ของสมการถดถอย

X คือ ค่าของตัวแปรอิสระ

2.4 ตัวแบบจำลอง SARIMAX

เป็นแบบจำลองที่ถูกพัฒนามาจากแบบจำลอง ARIMA ของ Box และ Jenkins [5] ซึ่งเป็นวิเคราะห์อนุกรมเวลาเพื่อใช้ในการพยากรณ์จากการสังเกตข้อมูลในอดีตและวิเคราะห์แนวโน้มในอนาคต ซึ่งตัวแบบจำลอง SARIMAX มีการใช้พารามิเตอร์ของฤดูกาลและปัจจัยภายนอกมาเกี่ยวข้องเพื่อใช้ในการพยากรณ์ โดยมีรูปแบบดังสมการที่ 9

$$\theta(L)^p \theta(L^s)^p \Delta^d \Delta_s^D y_t = \phi(L)^q \phi(L^s)^q \Delta^d \Delta_s^D \varepsilon_t + \sum_{i=1}^n \beta_i X_t^i \quad (9)$$

เมื่อ p คือ จำนวนระยะห่างของข้อมูลในอดีตจากปัจจุบัน

q คือ จำนวนระยะห่างของค่าคลาดเคลื่อนในอดีตจากปัจจุบัน

P คือ จำนวนระยะห่างของข้อมูลฤดูกาล s ในอดีตจากปัจจุบัน

Q คือ จำนวนระยะห่างของค่าคลาดเคลื่อนของฤดูกาล s ในอดีตจากปัจจุบัน

n คือ จำนวนตัวแปรปัจจัยภายนอก

Θ คือ พารามิเตอร์ของค่าเฉลี่ยเคลื่อนที่

θ คือ พารามิเตอร์ของค่าเฉลี่ยเคลื่อนที่ของฤดูกาล s

Φ คือ พารามิเตอร์ของ Autoregressive Parameter

- ∅ คือ พารามิเตอร์ของ Autoregressive Parameter ของฤดูกาล s
- L คือ ตัวดำเนินการล่าช้า (Lag Operator)
- L^s คือ ตัวดำเนินการล่าช้าของฤดูกาล s
- Δ คือ ตัวดำเนินการรวม (Integration Operator)
- Δ_s คือ ตัวดำเนินการรวมของฤดูกาล s
- β_i คือ ค่าสัมประสิทธิ์ของตัวแปรอิสระตัวที่ i
- X_tⁱ คือ ค่าของตัวแปรอิสระตัวที่ i ณ เวลา t
- y_t คือ ค่าสังเกตของอนุกรมเวลา ณ เวลา t
- ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t

2.5 วิธีโครงข่ายประสาทเทียม

เป็นศาสตร์ด้านปัญญาประดิษฐ์ที่ใช้ในการหาคำตอบ โดยเลียนแบบจากสมองของมนุษย์ เป็นการเรียนรู้ผ่านตัวอย่างโดยการป้อนข้อมูลนำเข้า และข้อมูลออก โดยความแม่นยำของค่าประมาณที่ออกมาจะขึ้นอยู่กับจำนวนโหนด (Node) ซึ่งเป็นหน่วยประมวลผลย่อยจำนวนชั้นของโครงข่าย (Layer) ค่าน้ำหนักของเส้นประสาท (Weight) รวมถึงข้อมูลนำเข้าทำให้ตัวแบบจำลองมีความแตกต่างได้อย่างไม่จำกัด มีรูปแบบสมการดังสมการที่ 10 – 11

$$a_i^k = g(\sum_i^n \sum_j^m \theta_{ij}^{k-1} x_j) \quad (10)$$

$$h_\theta(x) = a_i^n = g(\sum_i^n \sum_j^m \theta_{ij}^{n-1} x_j) \quad (11)$$

- เมื่อ θ_{ij} คือ ค่าน้ำหนักที่ถูกส่งผ่านจากโหนด j ในชั้นก่อนหน้าไปโหนด i ในชั้นปัจจุบัน
- X คือ ค่าเวกเตอร์ของข้อมูลนำเข้า
- a คือ ค่าเวกเตอร์ของข้อมูลเอาต์พุต
- n คือ ลำดับชั้นปัจจุบัน
- h_θ(x) คือ ค่าพยากรณ์ที่เลือกจากข้อมูลเอาต์พุต

2.6 การเปรียบเทียบความแม่นยำของตัวแบบพยากรณ์ด้วยค่า RMSE (Root Mean Square Error)

ในการเปรียบเทียบค่าพยากรณ์จะพิจารณาจากเกณฑ์รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error, RMSE) เป็นเกณฑ์ในการเปรียบเทียบความแม่นยำของตัวแบบพยากรณ์ ตัวแบบใด

ที่มีค่า RMSE ที่ต่ำที่สุดจะเป็นตัวแบบพยากรณ์ที่ดีที่สุด มีสมการดังสมการที่ 12

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^n (P_i - O_i)^2} \quad (12)$$

- เมื่อ RMSE คือ รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย
- P_i คือ ค่าประมาณการของอนุกรมเวลาตัวที่ i
- O_i คือ ค่าสังเกตของอนุกรมเวลาตัวที่ i
- n คือ จำนวนข้อมูลที่ใช้ในการคาดการณ์

2.7 ตำราและงานวิจัยที่เกี่ยวข้อง

ความสัมพันธ์ระหว่างปัจจัยทางสภาพอากาศกับการเกิดไฟฟ้าในจังหวัดเชียงใหม่ พบว่าความชื้นสัมพัทธ์เป็นปัจจัยเดียวที่สร้างสมการถดถอยได้ [6] การพยากรณ์อนุกรมเวลานั้น ปัจจัยอื่น ๆ ที่คาดว่าจะมีอิทธิพลกับต่อตัวแปรที่ต้องการพยากรณ์จะทำให้ความแม่นยำของการพยากรณ์มากกว่าตัวแปรพยากรณ์ตัวแปรเดียว [7] ซึ่ง A.E. Hessl ได้ทำการวิเคราะห์แล้วพบว่า อุณหภูมิ ความชื้นสัมพัทธ์ และความเร็วลม เป็นปัจจัยทางสภาพอากาศที่มีผลต่อการเกิดไฟฟ้า [8] ทั้งนี้เนื่องจากประเทศไทยมีลักษณะภูมิอากาศที่มีความชื้นสูง เนื่องจากอยู่ในเขตร้อนชื้นอีกทั้งพื้นที่การเกิดไฟฟ้าในประเทศไทยส่วนใหญ่เกิดจากการเผา ซึ่งเป็นปัจจัยที่เกิดขึ้นโดยมนุษย์ และข้อมูลที่น่ามาวิเคราะห์เป็นถูกรวมเป็นข้อมูลรายเดือนและรายปี จึงทำให้ผลการวิเคราะห์แตกต่างกัน [6] การพยากรณ์ด้วยสมการถดถอยเชิงเส้นนั้นอาจไม่เหมาะสมกับอนุกรมเวลารูปแบบฤดูกาล เนื่องจากไม่สามารถจับความเคลื่อนไหวจากส่วนประกอบแนวโน้ม และส่วนประกอบของฤดูกาลได้ [9] การใส่ตัวแปรหุ่นฤดูกาลไปจะช่วยให้ประสิทธิภาพให้กับสมการถดถอยในการจับการเคลื่อนไหวของส่วนประกอบฤดูกาล [10] ปัจจุบันมีการใช้แบบจำลองพยากรณ์มากมายที่เหมาะสมสำหรับการพยากรณ์อนุกรมรูปแบบฤดูกาลมาเปรียบเทียบกัน เช่น ตัวแบบจำลอง SARIMA, SARIMAX, Holt-Winters แบบผลคูณ และแบบผลบวก รวมถึงศาสตร์ด้านปัญญาประดิษฐ์อย่างวิธีโครงข่ายประสาทเทียม ในส่วน

ของการวิเคราะห์ด้วยอนุกรมฟูเรียร์ส่วนมากจะใช้คาดการณ์ความต้องการในการใช้ไฟฟ้า และทางด้านวิศวกรรมทางแผ่นดินไหว ไม่ค่อยนิยมใช้ในการพยากรณ์ยอดขายในรูปแบบฤดูกาล [11] การพยากรณ์นั้นการใช้ปัจจัยภายนอกที่มีความสัมพันธ์กับอนุกรมเวลาให้ผลลัพธ์ที่ดีกว่าเมื่อเทียบกับการพยากรณ์โดยไม่มีปัจจัยภายนอกมาเกี่ยวข้อง [12]

3. วิธีการดำเนินงานวิจัย

3.1 ขั้นตอนการรวบรวมข้อมูล

โดยข้อมูลที่ใช้ประกอบด้วย 2 ส่วน ได้แก่ จำนวนพื้นที่ที่เกิดไฟฟ้า และสภาพภูมิอากาศซึ่งนำมาใช้เป็นปัจจัยที่อาจส่งผลกระทบต่อจำนวนพื้นที่ไฟฟ้า ซึ่งได้เก็บรวบรวมจากข้อมูลสถิติการเกิดไฟฟ้าของกรมอุทยานแห่งชาติ สัตว์ป่า และพืชพันธุ์ ส่วนควบคุมไฟฟ้า และข้อมูลสถิติสภาพภูมิอากาศของกรมอุตุนิยมวิทยา ตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนกันยายน ปี พ.ศ. 2562 รวมทั้งสิ้น 3925 ข้อมูล ซึ่งข้อมูลปัจจัยภายนอกอย่างสภาพอากาศที่นำมาใช้ ได้แก่ อุณหภูมิอากาศ (Temperature) มีหน่วยเป็นองศาฟาเรนไฮต์ (°F), อุณหภูมิจุดน้ำค้าง (Dew Point) มีหน่วยเป็นองศาฟาเรนไฮต์ (°F), ค่าความชื้น (Humidity) มีหน่วยเป็นเปอร์เซ็นต์ (%), ความเร็วลม (Wind Speed) มีหน่วยเป็นเมตรต่อชั่วโมง (mph) และค่าความกดอากาศ (Pressure) มีหน่วยเป็นปรอท (Hg) โดยข้อมูลแต่ละตัวจะแบ่งออกเป็น 3 ค่า คือ ค่าสูงสุด (Max) ค่าเฉลี่ย (Average) และค่าต่ำสุด (Min) เพื่อนำมาใช้ในการพยากรณ์จำนวนพื้นที่ที่ถูกไฟฟ้า ซึ่งมีหน่วยเป็นจำนวนไร่ ซึ่งข้อมูลสภาพอากาศแสดงตัวอย่างดังตารางที่ 1

ตารางที่ 1 ตัวอย่างข้อมูลสภาพอากาศรายวัน

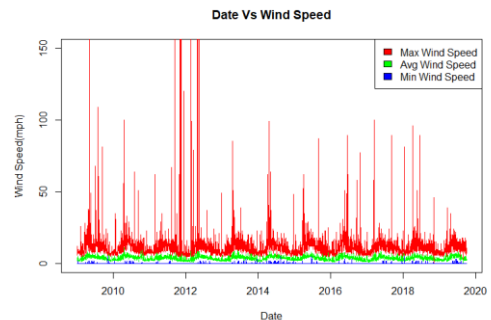
Date	Max Temp (F)	Avg Temp (F)	Min Temp (F)	Max Dew Point(F)	Avg Dew Point(F)	Min Dew Point(F)
1/1/2009	82	72.1	63	63	59.8	55
2/1/2009	77	71	66	63	61	57
3/1/2009	82	72.1	66	63	60.2	55
4/1/2009	82	71	61	73	59.2	54
5/1/2009	82	70.8	59	61	57.4	50
6/1/2009	82	70.5	61	61	57.7	50
7/1/2009	84	70.2	63	59	56.4	48
8/1/2009	82	69.4	59	59	55.5	50

3.2 การทำความสะอาดข้อมูล (Data Cleansing)

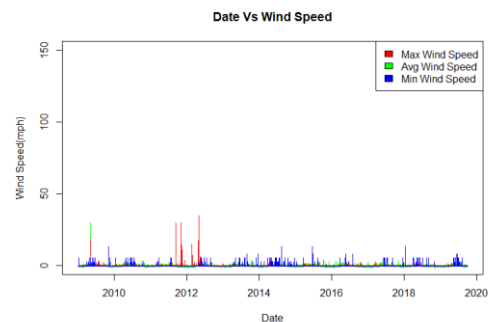
เนื่องจากข้อมูลที่ได้เก็บรวบรวมมาไม่ครบถ้วน และข้อมูลบางอย่างที่ผิดพลาด เช่น การมีข้อมูลบางค่าที่สูญหาย (Missing Value) หรือข้อมูลบางตัวที่อยู่นอกกลุ่ม (Outlier) จึงได้นำทฤษฎีการทำความสะอาดข้อมูลซึ่งเป็นการลบการในการแก้ไขข้อมูลที่ผิดพลาด เช่นการใส่ค่าเฉลี่ย (Mean) ลงในค่าของข้อมูลที่สูญหาย การทำข้อมูลให้อยู่ในรูปแบบเดียวกัน (Standardizing) ซึ่งเป็นการแก้ไขข้อมูลที่มีความกว้างที่ไม่เท่ากันให้มาอยู่ในรูปของการแจกแจงแบบปกติมาตรฐาน คำนวณได้จากสมการที่ 13 และตัวอย่างกราฟข้อมูลก่อนและหลังทำ Standardized แสดงดังรูปที่ 1-2

$$Z = \frac{X - \mu}{\sigma} \tag{13}$$

เมื่อ X คือ ข้อมูลที่ต้องการทำ Standardized
 μ คือ ค่าเฉลี่ยของชุดข้อมูล
 σ คือ ค่าเบี่ยงเบนมาตรฐาน



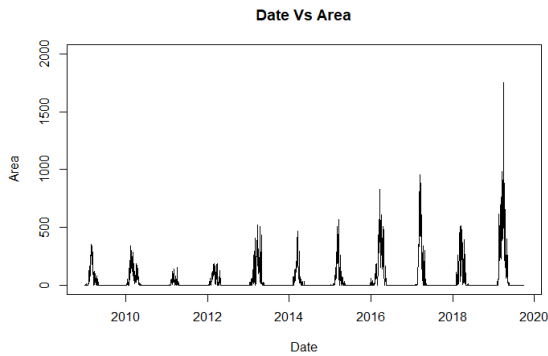
รูปที่ 1 ตัวอย่างกราฟข้อมูลของความเร็วลมก่อนทำ Standardized



รูปที่ 2 ตัวอย่างกราฟข้อมูลของความเร็วลมหลังทำ Standardized

3.3 การวิเคราะห์ข้อมูล

เมื่อนำข้อมูลที่ได้จากการทำความสะอาดเรียบร้อยแล้วจะทำมาแสดงข้อมูลด้วยกราฟเพื่อหารูปแบบของอนุกรมเวลาดังรูปที่ 3



รูปที่ 3 กราฟข้อมูลจำนวนพื้นที่ไฟป่า

จากรูปที่ 3 จะพบว่าอนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากฤดูกาล และนำข้อมูลมาวิเคราะห์หาค่าเฉลี่ยและความแปรปรวนพบว่าค่าเฉลี่ย (μ) เท่ากับ 40.2377 และ ค่าเบี่ยงเบนมาตรฐาน (σ) เท่ากับ 115.1595 พบว่ามีค่า 3σ มากกว่า μ [13] จัดเป็นข้อมูลที่มีความแปรปรวนสูงอาจทำให้ยากต่อการคาดการณ์

3.4 การแบ่งข้อมูลและการตรวจสอบไขว้

ข้อมูลอนุกรมเวลาจะถูกทำการแบ่งข้อมูลเป็น 2 ชุดข้อมูล ข้อมูลชุดแรกคือข้อมูลฝึกสอน (Training Data) จะถูกนำมาใช้เพื่อสร้างตัวแบบพยากรณ์ โดยมีการแบ่งข้อมูล 8 ปีแรกของข้อมูลทั้งหมด และข้อมูลที่เหลือจะถูกแบ่งเป็นข้อมูลทดสอบ (Test Data) เพื่อวัดประสิทธิภาพของตัวแบบจำลอง โดยแบ่งการทดสอบเป็น 4 ช่วงได้แก่ 3 เดือน, 6 เดือน, 1 ปี และ 1 ปี 9 เดือน

สำหรับข้อมูลอนุกรมเวลาที่ใช้การเรียนรู้ของเครื่องจักร จะมีการทำการตรวจสอบไขว้ (Cross Validation) ซึ่งเป็นเครื่องมือในการทดสอบแบบจำลองของข้อมูลก่อนนำไปพยากรณ์ข้อมูลทดสอบ โดยใช้ 2 วิธีคือ

1) K-Fold Cross Validation วิธีนี้จะใช้การแบ่งข้อมูลของออกเป็น K ส่วนเท่า ๆ กันโดยการสุ่ม เพื่อนำมาวัดประสิทธิภาพของตัวแบบพยากรณ์ ในแต่ละรอบของการทดสอบจะมีการวัดค่าความคลาดเคลื่อน และ ความแม่นยำ ก่อนนำไปใช้ในการพยากรณ์ข้อมูลทดสอบ สำหรับงานวิจัยนี้จะใช้จำนวน K คือ 5

2) Repeat K-Fold Cross Validation วิธีนี้มีหลักการเดียวกับวิธี K-Fold Cross Validation แต่จะมีการทำซ้ำ ตามจำนวนรอบที่กำหนด สำหรับงานวิจัยนี้จะใช้จำนวน K คือ 10 และ รอบการทำซ้ำ คือ 5

3.5 การสร้างตัวแบบพยากรณ์

3.5.1 การพยากรณ์ด้วยวิธี Holt-Winters แบบผลบวก

การสร้างตัวแบบจำลองพยากรณ์ด้วยวิธี Holt-Winters แบบผลบวกโดยมีการกำหนดจำนวนคาบของฤดูกาลอยู่ที่ $L = 365$ ซึ่งจะใช้โปรแกรม Excel ในการสร้าง และใช้โปรแกรม OpenSolver ในการคำนวณหาค่าพารามิเตอร์ต่าง ๆ เพื่อให้ได้ค่า RMSE ที่ต่ำที่สุดแสดงตัวอย่างดังรูปที่ 4 สรุปค่า RMSE ที่ได้จากการพยากรณ์ได้ดังตารางที่ 2

Cell Reference	Value	Cell Reference	Value	Cell Reference	Value
L=	365	alpha	0.200679	Gamma	0
Error^2	1	beta	0	b	0
33.46387	2	0	5.786524791	2.358672	4.54589
42.40211	3	0	7.904579413	2.220316	4.373204
73.8778	4	0	8.595119544	1.503538	4.18543
75.49825	5	0	8.88896722	0.941751	3.995807
90.96119	6	0	9.337377949	1.112323	3.97925
75.25246	7	0	8.674817533	2.46402	3.97737
13.28321	8	0	1.33720394	0.462124	3.97297
144.8434	9	0	11.03509051	3.09432	3.310045
84.22961	10	0	9.215726081	3.24353	3.108716
61.96168	11	0	7.98518157	6.79404	2.936891
140.5919	12	10	-1.857145878	3.146883	3.195926
69.3968	13	0	8.962794613	1.408791	3.01323
0.178102	14	6	6.422021551	4.172032	3.00401
84.19957	15	0	9.176032396	1.940286	2.803947
20.49462	16	2	6.434832745	1.892133	2.704261
42.58856	17	0	6.556413892	0.672491	2.561047
3.100389	18	5	2.133727673	1.279664	2.596238
47.32952	19	0	6.879601502	2.802277	2.448344
27.6	20	0	5.151970308	2.144332	2.334572
20.30463	21	0	4.790540509	1.826485	2.236788
16.49394	22	0	4.061272816	1.659434	2.148064
14.466	23	0	3.800980816	1.550362	2.064971
13.07063	24	0	3.615332395	1.473663	1.989869
11.96919	25	0	3.499651709	1.410205	1.918408
7.73814	26	6	3.32061206	1.307843	1.868943

รูปที่ 4 ตัวอย่างการคำนวณพารามิเตอร์สำหรับวิธี Holt-Winters

ตารางที่ 2 ค่า RMSE ของวิธี Holt-Winters แบบผลบวก

Variable	RMSE
RMSE 3 เดือน	198.1840408
RMSE 6 เดือน	153.3521299
RMSE 1 ปี	107.9897828
RMSE 1 ปี 9 เดือน	201.5325591

3.5.2 การพยากรณ์ด้วยอนุกรมฟูเรียร์แบบตัดทอน

การสร้างตัวแบบจำลองพยากรณ์ด้วยอนุกรมฟูเรียร์แบบตัดทอนโดยมีการกำหนดจำนวนคาบของสัญญาณเวลาอยู่ที่ $T = 365$ ซึ่งจะใช้โปรแกรม Excel ในการสร้าง และใช้โปรแกรม OpenSolver ในการคำนวณหาค่าพารามิเตอร์ต่าง ๆ เพื่อให้ได้ค่า RMSE ที่ต่ำที่สุด สรุปค่า RMSE ที่ได้จากการพยากรณ์ได้ดังตารางที่ 3

ตารางที่ 4 ค่าสัมประสิทธิ์สหสัมพันธ์ของสมการถดถอยแบบเต็มรูประหว่างปัจจัยสภาพอากาศกับพื้นที่ไฟป่า

Coefficients:	Estimate	Pr(> t)
(Intercept)	2.57E+01	8.41E-08
Max.Temperature...F.	-7.73E+00	0.000128
Avg.Temperature...F.	1.84E+02	< 2.00E-16
Min.Temperature...F.	-2.46E+01	4.89E-09
Max.Dew.Point...F.	5.60E+00	0.051747
Avg.Dew.Point...F.	-2.61E+02	< 2.00E-16
Min.Dew.Point...F.	2.93E+01	2.42E-07
Max.Humidity....	-2.70E+01	< 2.00E-16
Avg.Humidity....	2.16E+02	< 2.00E-16
Min.Humidity....	-7.69E+00	0.046249
Max.Wind.Speed...mph.	-2.41E-01	0.844327
Avg.Wind.Speed...mph.	2.37E-01	0.870545
Min.Wind.Speed...mph.	-9.93E-03	0.992593
Max.Pressure..Hg.	-3.48E-01	0.712637
Avg.Pressure..Hg.	-1.93E+00	0.350948
Min.Pressure..Hg.	1.68E+00	0.254858
Month02	2.28E+01	7.39E-06
Month03	1.03E+02	< 2.00E-16
Month04	-7.22E+00	0.284141
Month05	-2.64E+01	0.000165
Month06	-7.49E+00	0.299014
Month07	-3.46E+00	0.628467
Month08	-2.95E+00	0.674405
Month09	-1.51E+00	0.827491
Month10	4.36E+00	0.483658
Month11	4.19E+00	0.443781
Month12	5.95E+00	0.21058

รูปที่ 5 ตัวอย่างการคำนวณพารามิเตอร์สำหรับอนุกรมฟูเรียร์แบบตัดทอน

ตารางที่ 3 ค่า RMSE ของอนุกรมฟูเรียร์แบบตัดทอน

Variable	RMSE
RMSE 3 เดือน	186.3560544
RMSE 6 เดือน	140.9722406
RMSE 1 ปี	99.27195697
RMSE 1 ปี 9 เดือน	147.2478356

3.5.3 การพยากรณ์ด้วยการถดถอยแบบพหุคูณที่มีตัวแปรจัดประเภท

การสร้างตัวแบบจำลองพยากรณ์ด้วยการถดถอยแบบพหุคูณที่มีตัวแปรจัดประเภทจากโปรแกรม R studio โดยใช้ตัวแปรจัดประเภทในรูปแบบของเดือน จะแบ่งการสร้างสมการออกเป็น 4 แบบคือ การสร้างการถดถอยแบบเต็มรูป (Full Model) แสดงดังตารางที่ 4, การสร้างการถดถอยแบบลดรูป (Reduced Model), การสร้างการถดถอยแบบลดรูปแบบมี K-Fold Cross Validation และการสร้างการถดถอยแบบลดรูปแบบมี Repeated K-Fold Cross Validation

จากตารางที่ 4 เนื่องจากมีปัจจัยทางสภาพอากาศ และตัวแปรจัดประเภทบางตัวมีค่า p-value ที่มากกว่า 0.05 จึงทำการลดรูปด้วยวิธี Stepwise เพื่อคัดเลือกและตัดตัวแปรโดยการตรวจสอบระดับนัยสำคัญ และยังคงตัวแปรจัดประเภทไว้เหมือนเดิมสำหรับสมการถดถอยแบบลดรูป และทำการลดรูปด้วยวิธี Stepwise ของตัวแปรจัดประเภทสำหรับสมการถดถอยแบบลดรูปโดยมีการตรวจสอบไขว้ แสดงดังตารางที่ 5 และสรุปค่า RMSE ที่ได้จากการพยากรณ์ได้ดังตารางที่ 6

ตารางที่ 5 ค่าสัมประสิทธิ์สหสัมพันธ์ของสมการถดถอยแบบลดรูปโดยมีการตรวจสอบไขว้ระหว่างปัจจัยสภาพอากาศกับพื้นที่ไผ่ป่า

Coefficients:	Estimate	Pr(> t)
(Intercept)	2.26E+01	<2.00E-16
Max.Temperature....F.	-7.88E+00	9.09E-05
Avg.Temperature....F.	1.85E+02	<2.00E-16
Min.Temperature....F.	-2.46E+01	2.57E-09
Max.Dew.Point....F.	5.46E+00	5.68E-02
Avg.Dew.Point....F.	-2.63E+02	<2.00E-16
Min.Dew.Point....F.	2.99E+01	8.09E-08
Max.Humidity....	-2.68E+01	<2.00E-16
Avg.Humidity....	2.18E+02	<2.00E-16
Min.Humidity....	-8.14E+00	3.15E-02
Month02	2.54E+01	9.68E-09
Month03	1.07E+02	<2.00E-16
Month05	-2.17E+01	4.32E-09
Month10	6.32E+00	8.15E-02
Month11	5.61E+00	1.31E-01
Month12	6.79E+00	9.40E-02

ตารางที่ 6 ค่า RMSE ของการถดถอยแบบพหุคูณที่มีตัวแปรจัดประเภท

Method	RMSE			
	3 เดือน	6 เดือน	1 ปี	1 ปี 9 เดือน
Full model	91.98635	76.80512	55.18481	144.1421
Reduced model	92.113	76.89221	55.25577	144.0143
Reduced model K5	92.24924	76.77547	55.16708	143.8217
Reduced model RcvK10	92.24924	76.77547	55.16708	143.8217

3.5.4 การพยากรณ์ด้วยการถดถอยแบบโพลีโนเมียลแบบหลายตัวแปร

การสร้างตัวแบบจำลองพยากรณ์ด้วยการถดถอยแบบโพลีโนเมียลแบบหลายตัวแปรจะแบ่งการสร้างสมการออกเป็น 4 แบบจากการถดถอยแบบพหุคูณ ที่มีตัวแปรจัดประเภทโดยมีการเพิ่มสมการกำลังสอง (Quadratic) และส่วนที่เป็นปฏิสัมพันธ์ระหว่างตัวแปรอิสระ 2 ตัว (interaction) สรุปค่า RMSE ที่ได้จากการพยากรณ์ได้ดังตารางที่ 7

ตารางที่ 7 ค่า RMSE ของการถดถอยแบบโพลีโนเมียลแบบหลายตัวแปร

Method	RMSE			
	3 เดือน	6 เดือน	1 ปี	1 ปี 9 เดือน
Full model	93.33195	78.93754	55.88681	145.9292
Reduced model	93.99207	79.06667	56.09014	145.0891
Reduced model with K5	93.59851	78.94547	56.09811	144.828
Reduced model with RcvK10	93.59851	78.94547	56.09811	144.828

3.5.5 การพยากรณ์ด้วยตัวแบบจำลอง SARIMAX

การพยากรณ์ด้วยวิธี SARIMAX เป็นเทคนิคทางอนุกรมเวลาที่ต้องทดสอบความนิ่งของอนุกรม (Stationary test) ก่อนนำมาสร้างแบบจำลอง การศึกษาครั้งนี้ทำการทดสอบความนิ่ง ด้วยวิธี ADF unit root ของ Dickey and Fuller [14] โดยกำหนดสมมติฐานหลักในการทดสอบ (Null Hypothesis: H0) คือ อนุกรมเวลาไม่มีความนิ่ง ผลการทดสอบดังรูปที่ 6 พบว่าค่า P-value สามารถปฏิเสธสมมติฐานหลักที่ระดับนัยสำคัญ 0.05 ได้

Augmented Dickey-Fuller Test
data: Area
Dickey-Fuller = -6.577, Lag order = 14, p-value = 0.01
alternative hypothesis: stationary

รูปที่ 6 ผลการทดสอบความนิ่งของอนุกรมด้วยวิธี ADF unit root

จากนั้นนำมาสร้างแบบจำลอง SARIMAX โดยใช้ปัจจัยภายนอกจากสมการถดถอยแบบลดรูปด้วยโปรแกรม R studio โดยมีวิธีคัดเลือกตัวแบบจำลองจากค่า Akaike

information criterion (AIC) ที่น้อยที่สุด โดยคัดเลือกมาทั้งหมด 6 ตัวแบบ สรุปค่า RMSE ที่ได้จากการพยากรณ์ได้ดังตารางที่ 8

ตารางที่ 8 ค่า RMSE ของการแบบจำลอง SARIMAX

Method	RMSE			
	3 เดือน	6 เดือน	1 ปี	1 ปี 9 เดือน
SARIMAX(3,0,1)(0,1,0)	181.6291	135.3868	95.39092	141.4235
SARIMAX(3,0,2)(0,1,0)	181.6845	135.4304	95.42068	141.4462
SARIMAX(4,0,1)(0,1,0)	181.6222	135.3809	95.38683	141.4179
SARIMAX(5,0,1)(0,1,0)	181.6247	135.3828	95.38794	141.418
SARIMAX(4,0,2)(0,1,0)	181.6175	135.3775	95.38458	141.4177
SARIMAX(5,0,5)(0,1,0)	181.5213	135.2967	95.33204	141.3259

3.5.6 การพยากรณ์ด้วยวิธีโครงข่ายประสาทเทียม

ในการสร้างแบบจำลองด้วยวิธีโครงข่ายประสาทเทียมจะใช้โปรแกรม R studio โดยมีข้อมูลนำเข้าเป็นปัจจัยสภาพอากาศจากสมการถดถอยแบบเต็มรูปแบบและแบบลดรูป เพื่อมาเปรียบเทียบผลลัพธ์ที่ได้ โดยกำหนดค่าพารามิเตอร์ของแบบจำลองคือ จำนวนโหนดเท่ากับ 5, 10, 15, 20, 25, 30, 35 และ 40 และ อัตราการเรียนรู้ (Learning Rate) เท่ากับ 0.01, 0.05, 0.1, 0.2, 0.4, 0.8 และ 1 โดยแบบจำลองจะถูกเลือกจากค่าพารามิเตอร์ที่ทำให้ค่า RMSE ต่ำสุดมาใช้ สรุปค่าพารามิเตอร์ และ RMSE ที่ได้จากการพยากรณ์ได้ดังตารางที่ 9

ตารางที่ 9 ค่า RMSE ของวิธีโครงข่ายประสาทเทียม

Neural Network Method (size,decay)	RMSE			
	3 เดือน	6 เดือน	1 ปี	1 ปี 9 เดือน
Full Model (30,0.05)	106.3965	87.55564	62.2416	147.4387
Full Model K5 (30,0.4)	97.2116	86.10667	64.74964	150.9985
Full Model RcvK10 (35,1)	100.4768	84.23596	60.65833	150.1992
Reduced Model (30,0.4)	102.2462	84.34666	60.16941	145.2222
Reduced Model K5 (25,0.8)	97.52495	81.41748	58.18625	143.1615
Reduced Model RcvK10 (35,1)	98.30117	84.77458	59.98173	145.9034

3.5.7 การพัฒนาแบบจำลอง

หลังจากสร้างแบบจำลองเพื่อใช้ในการพยากรณ์ผลลัพธ์บางค่าที่ได้ออกมาจะมีค่าน้อยกว่า 0 ซึ่งจากข้อมูลพื้นที่ไฟป่าจะเห็นว่าไม่มีค่าใดที่มีค่าน้อยกว่า 0 จึงต้องทำการพัฒนาแบบจำลอง โดยวิธีที่ใช้คือการใช้ฟังก์ชันที่เรียกว่า Rectified Linear Unit (ReLU) โดยมีรูปแบบสมการดังสมการที่ 14

$$f(x) = \max(0, x) = \begin{cases} 0 & \text{เมื่อ } x \leq 0 \\ x & \text{เมื่อ } x > 0 \end{cases} \quad (14)$$

เมื่อ x คือ ค่าพยากรณ์

4. สรุปผลการดำเนินงานและข้อเสนอแนะ

งานวิจัยนี้ได้ทำการศึกษาการเปรียบเทียบวิธีพยากรณ์ของอนุกรมเวลารูปแบบฤดูกาลที่มีเลขศูนย์จำนวนมากโดยใช้ข้อมูลจำนวนพื้นที่การเกิดไฟป่า และข้อมูลสภาพอากาศ ด้วยการสร้างแบบจำลองทั้ง 6 แบบ ได้แก่ วิธีอนุกรมฟูเรียร์แบบตัดทอน, วิธี Holt-Winters แบบผลบวก, วิธีการวิเคราะห์การถดถอยแบบพหุคูณที่มีตัวแปรจัดประเภท, วิธีการวิเคราะห์การถดถอยแบบโพลีโนเมียลแบบหลายตัวแปร, วิธีการใช้แบบจำลอง SARIMAX และวิธีโครงข่ายประสาทเทียม

จากทดลองพบว่าสมการถดถอยแบบเต็มรูปแบบให้ค่า RMSE ที่ต่ำสุดในการพยากรณ์ช่วง 3 เดือน เท่ากับ 91.98635 สมการถดถอยแบบลดรูปโดยใช้การตรวจสอบไขว้ ให้ค่า RMSE ต่ำสุดในการพยากรณ์ช่วง 6 เดือน เท่ากับ 76.77547 และ 1 ปี เท่ากับ 55.16708 และพบว่ารูปแบบวิธีของการตรวจสอบไขว้ที่แตกต่างกันให้ผลลัพธ์ที่เหมือนกันในการสร้างสมการถดถอยแบบเต็มรูปแบบและแบบลดรูป ในส่วนของการพยากรณ์ในช่วง 1 ปี 9 เดือนนั้น แบบจำลอง SARIMAX (5,0,5)(0,1,0) ให้ค่า RMSE ต่ำที่สุดเท่ากับ 141.3259 จากผลการทดลองทำให้ทราบว่าวิธีการพยากรณ์ที่ใช้ปัจจัยภายนอกที่อาจส่งผลกระทบต่ออนุกรมเวลาให้ผลลัพธ์ที่ดีกว่าวิธีพยากรณ์โดยไม่ใช้ปัจจัยภายนอกมาเกี่ยวข้อง

งานวิจัยนี้สามารถนำไปใช้เพื่อเป็นแนวทางในการพยากรณ์จำนวนพื้นที่ไฟป่าในแต่ละจังหวัดเนื่องจากมีรูปแบบอนุกรมที่คล้ายคลึงกัน ทั้งนี้ผลลัพธ์ของวิธีการ

พยากรณ์อาจได้ผลที่แตกต่างกันเนื่องจากปัจจัยด้านสภาพอากาศของแต่ละจังหวัดนั้นมีความแตกต่างกัน อีกทั้งข้อมูลปัจจัยทางสภาพอากาศบางค่ามีการสูญหายซึ่งอาจส่งผลกระทบต่อค่าคลาดเคลื่อนของการพยากรณ์ ในการพยากรณ์เนื่องจากการวัดความแม่นยำของตัวแบบพยากรณ์นั้นเป็นการใช้ข้อมูลทดสอบในการวัดซึ่งเป็นการทราบค่าของปัจจัยทางสภาพอากาศที่นำมาใช้ล่วงหน้า หากนำมาใช้จริงต้องทำการพยากรณ์ปัจจัยทางสภาพอากาศก่อน ซึ่งในการพยากรณ์ปัจจัยทางสภาพอากาศในระยะยาวนั้น อาจส่งผลให้เกิดความคลาดเคลื่อนต่อการพยากรณ์จำนวนพื้นที่ไฟป่า

เอกสารอ้างอิง

[1] สมิทธ ธรรมสโรช, “เทคโนโลยีภูมิสารสนเทศเพื่อศึกษากฎพิบัติทางธรรมชาติในประเทศไทย,” *วารสารวิทยาศาสตร์บูรพา*, ปีที่ 19, ฉบับที่ 2, น. 182, 2557.

[2] สำนักป้องกัน ปราบปราม และควบคุมไฟป่า, “การพยากรณ์สถานการณ์ไฟป่า,” [ออนไลน์]. แหล่งที่มา: <https://www.dnp.go.th/forestfire/firepredict/Fab%2052/9-16Fab52.htm>. [วันที่ เข้า ถึง 2 กุมภาพันธ์ 2564].

[3] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *International Journal of Forecasting*, vol. 20, no.1, pp. 5-10, 2004.

[4] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Management Science*, vol.6, no. 3, pp. 324-342, 1960.

[5] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, "Time Series Analysis; Forecasting and Control. 3rd Edition", Prentice Hall, Englewood Cliff, New Jersey, 1994.

[6] P. Thanadolmethaphorn, et al., “การวิเคราะห์ผลกระทบจากการเปลี่ยนแปลงสภาพภูมิอากาศต่อศักยภาพการเกิดไฟป่าในจังหวัดเชียงใหม่โดยใช้แบบจำลองการถดถอย,” *The Journal of King*

Mongkut's University of Technology North Bangkok, vol. 28, no. 4, 2018.

[7] นิฉา แก้วหาวงษ์ , “การพยากรณ์ปริมาณการใช้ไฟฟ้าของประเทศไทยโดยใช้ตัวแบบ SARIMA และตัวแบบการถดถอยที่มีความคลาดเคลื่อนเป็นตัวแบบ ARMA,” *Thai Journal of Science and Technology*, ปีที่ 4, ฉบับที่ 1, น. 34, 2558

[8] A. E. Hesel, “Pathway for climate change effects on fire: models, data, and uncertainties,” *Progress in Physical Geography: Earth and Environment*, vol. 35, no. 3, pp. 393-407, 2011.

[9] รณชัย ชื่นธวัช, "การวิเคราะห์อนุกรมเวลาของข้อมูลหน่วยจำหน่ายไฟฟ้า เพื่อค้นหาตัวแบบการพยากรณ์ที่เหมาะสมด้วยภาษาอาร์," *วารสาร มทร.อีสาน ฉบับวิทยาศาสตร์และเทคโนโลยี*, ปีที่ 9, ฉบับที่ 3 , 2559.

[10] C. Jatuporn and P. Sukprasert, “Forecasting models for rubber production and export quantity of Thailand,” *Khon Kaen Agriculture Journal*, vol. 44, no. 2, pp. 219-228, 2016

[11] A. Fumi et al., "Fourier analysis for demand forecasting in a fashion company," *International Journal of Engineering Business Management*, vol. 5, 2013

[12] P. Kritharas, "Developing a SARIMAX model for monthly wind speed forecasting in the UK," Ph.D. thesis, Dept Eng. ,Loughborough Univ. , Loughborough, England, 2013.

[13] สถาพร โอภาสานนท์, “การพยากรณ์อุปสงค์โดยใช้เทคนิคการพยากรณ์มากกว่าหนึ่งวิธี,” *วารสารบริหารธุรกิจ คณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์*, ปีที่ 33, ฉบับที่ 127, 2553.

[14] D. Dickey and W. A. Fuller., “Likelihood ratio statistics for autoregressive time series with a unit root,” *Econometrica*, vol. 49, pp. 1057-1072, 1981.