



บทความวิจัย

ตัวแบบผสมสำหรับพยากรณ์ข้อมูลอนุกรมเวลาปริมาณหน่วยจำหน่ายไฟฟ้าบ้านอยู่อาศัยรายเดือนของประเทศไทย

ธนกร สุทธิสนธิ*

สาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี

ภาณุเดช เพ็ญความสุข

สาขาวิชาการพัฒนาชุมชน คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 08 8125 5873 อีเมล: sutthison.statubru04@gmail.com DOI: 10.14416/j.kmutnb.2024.02.003

รับเมื่อ 5 พฤษภาคม 2565 แก้ไขเมื่อ 1 สิงหาคม 2565 ตอรับเมื่อ 11 สิงหาคม 2565 เผยแพร่ออนไลน์ 20 กุมภาพันธ์ 2567

© 2024 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาตัวแบบที่เหมาะสมในการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณหน่วยจำหน่ายไฟฟ้าบ้านอยู่อาศัยรายเดือนของประเทศไทย เก็บรวบรวมข้อมูลจากการไฟฟ้าส่วนภูมิภาค ตั้งแต่เดือนมกราคม 2553 ถึงเดือนกุมภาพันธ์ 2565 จำนวน 146 ค่า จำนวน 4 ชุด ได้แก่ ภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคกลาง และภาคใต้ การสร้างตัวแบบการพยากรณ์มีทั้งหมด 5 ตัวแบบ ได้แก่ ตัวแบบ SARIMA ตัวแบบ ANN ตัวแบบ SARIMA – ANN ตัวแบบ SARIMA – ANN – REG และตัวแบบ Propose ซึ่งเป็นตัวแบบที่ผู้วิจัยนำเสนอ การสร้างตัวแบบได้ประยุกต์ใช้วิธีการวิเคราะห์องค์ประกอบของข้อมูลอนุกรมเวลาเชิงประจักษ์ (Empirical Mode Decomposition; EMD) เพื่อลดสัญญาณแกว่งไกวของข้อมูลก่อนนำไปสร้างตัวแบบผสม SARIMA – ANN – REG การดำเนินการวิจัยทำการแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดที่ 1 ข้อมูลฝึกฝน (Training Data Set) ตั้งแต่ เดือนมกราคม 2553 ถึงเดือนธันวาคม 2562 จำนวน 120 ค่า สำหรับสร้างตัวแบบการพยากรณ์ และชุดที่ 2 ข้อมูลทดสอบ (Testing Data Set) ตั้งแต่ เดือนมกราคม 2563 ถึงเดือนกุมภาพันธ์ 2565 จำนวน 26 ค่าสำหรับเปรียบเทียบความแม่นยำของตัวแบบการพยากรณ์ ด้วยเกณฑ์การวัดประสิทธิภาพ 3 เกณฑ์ ได้แก่ ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย ร้อยละของความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ผลการวิจัยพบว่าตัวแบบผสมที่พัฒนาขึ้นมีประสิทธิภาพและมีความแม่นยำดีกว่าตัวแบบ SARIMA ตัวแบบ ANN ตัวแบบ SARIMA – ANN และตัวแบบ SARIMA – ANN – REG ในทุกเกณฑ์วัดประสิทธิภาพการพยากรณ์ สรุปได้ว่า ตัวแบบผสมที่พัฒนาขึ้นเป็น ตัวแบบที่เหมาะสมสำหรับการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณหน่วยจำหน่ายไฟฟ้าบ้านอยู่อาศัยรายเดือนของประเทศไทย

คำสำคัญ: ตัวแบบ SARIMA ตัวแบบ ANN ตัวแบบผสม ประสิทธิภาพการพยากรณ์

การอ้างอิงบทความ: ธนกร สุทธิสนธิ และ ภาณุเดช เพ็ญความสุข, “ตัวแบบผสมสำหรับพยากรณ์ข้อมูลอนุกรมเวลาปริมาณหน่วยจำหน่ายไฟฟ้าบ้านอยู่อาศัยรายเดือนของประเทศไทย,” *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 34, ฉบับที่ 2, หน้า 1–18, เลขที่บทความ 242-146035, เม.ย.-มิ.ย. 2567.



Hybrid Model for Forecasting Time Series Data of Monthly Household Electrical Distribution Units in Thailand

Thanakon Sutthison*

Applied Statistics Program, Faculty of Science, Ubon Ratchathani Rajabhat University, Ubon Ratchathani, Thailand

Panudet Pienkhawsook

Community Development Program, Faculty of Humanities and Social Sciences, Ubon Ratchathani Rajabhat University, Ubon Ratchathani, Thailand

* Corresponding Author, Tel. 08 8125 5873, E-mail: sutthison.statubru04@gmail.com DOI: 10.14416/j.kmutnb.2024.02.003

Received 5 May 2022; Revised 1 August 2022; Accepted 11 August 2022; Published online: 20 February 2024

© 2024 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

Abstract

The purpose of this research was to develop a model suitable for forecasting the time series data of monthly household electrical distribution units in Thailand. Data were collected from the Provincial Electricity Authority from January 2010 to February 2022. There were a total of 146 values, consisting of 4 data sets, namely Northern, Northeastern, Central, and Southern Thailand. Five forecasting models were constructed: SARIMA, ANN, SARIMA-ANN, SARIMA-ANN-REG, and Propose, which was the model presented by the researcher. An empirical mode decomposition was used to reduce the fast oscillation of the data before being used to create a SARIMA-ANN-REG hybrid model. The data were divided into two sets. The first set contained 120 values of the training data set from January 2010 to December 2019, which were used for constructing the forecasting models. The second set consisted of 26 values of the testing data set from January 2020 to February 2022. The forecasting accuracy was compared by three performance measurement criteria: Mean Absolute Error (MAE), Root Mean Square error (RMSE), and Mean Absolute Percent Error (MAPE). The results showed that the developed hybrid model had better forecasting performance and accuracy than SARIMA, ANN, SARIMA-ANN, and SARIMA-ANN-REG in all forecasting performance measurement criteria. It can be concluded that the developed model is suitable for forecasting the time series data of monthly household electrical distribution units in Thailand.

Keywords: SARIMA Model, ANN Model, Hybrid Model, Forecasting Performance

Please cite this article as: T. Sutthison and P. Pienkhawsook, "Hybrid model for forecasting time series data of monthly household electrical distribution units in Thailand," *The Journal of KMUTNB*, vol. 34, no. 2, pp. 1-18, ID. 242-146035, Apr.-Jun. 2024 (in Thai).

1. บทนำ

ไฟฟ้าเป็นแหล่งพลังงานที่เป็นปัจจัยสำคัญในการขับเคลื่อนเศรษฐกิจและการพัฒนาประเทศ [1] การใช้พลังงานไฟฟ้าสะท้อนให้เห็นถึงการพัฒนาเศรษฐกิจ สังคมของประเทศนั้นอย่างต่อเนื่อง ดังนั้นการรักษาความมั่นคงเรื่องพลังงานไฟฟ้าจึงเป็นเรื่องสำคัญที่ทุกประเทศต้องคำนึง [2], [3] การขาดแคลนพลังงานไฟฟ้าในบางช่วงเวลาไฟดับหรือไฟตกถือเป็นค่าใช้จ่ายมหาศาลต่อเศรษฐกิจในประเทศ [4] การพยากรณ์ปริมาณการใช้พลังงานไฟฟ้าในอนาคตว่ามีปริมาณเพิ่มขึ้นหรือลดลงเท่าใดจึงมีความสำคัญอย่างยิ่งต่อการวางแผน การดำเนินการผลิตพลังงานไฟฟ้าให้ตรงตามความต้องการไฟฟ้าและการเจริญเติบโตทางเศรษฐกิจภายในประเทศ [5] การพยากรณ์ปริมาณการใช้พลังงานไฟฟ้าให้ถูกต้องแม่นยำขึ้นอยู่กับปัจจัยหลายอย่าง เช่น ความผันผวนของเศรษฐกิจโลก ราคาพลังงาน ประเด็นการเมือง ปัญหาโลกร้อน และผลกระทบด้านสิ่งแวดล้อม ปัญหาโรคระบาด เป็นต้น ปัจจัยเหล่านี้ล้วนมีอิทธิพลต่อปริมาณการใช้พลังงานไฟฟ้าทั้งสิ้น และการคาดการณ์การเปลี่ยนแปลงของปัจจัยที่กล่าวมาข้างต้นไม่สามารถทำได้ ทำให้การพยากรณ์ปริมาณการใช้พลังงานไฟฟ้ามีความคลาดเคลื่อนสูง [1] ด้วยเหตุนี้การพัฒนาตัวแบบการพยากรณ์ปริมาณการใช้พลังงานไฟฟ้าที่มีความแม่นยำสูงจึงเป็นความท้าทายอย่างยิ่ง

จากการทบทวนเอกสารพบว่า มีนักวิจัยจำนวนมากได้เสนอแนะและพัฒนาตัวแบบสำหรับพยากรณ์ปริมาณการใช้พลังงานไฟฟ้าหลากหลาย เช่น ตัวแบบ SARIMA เป็นตัวแบบที่นิยมนำมาใช้ในการพยากรณ์ปริมาณการใช้พลังงานไฟฟ้าของประชาชน [3], [6], [7] ตัวแบบนี้ให้ค่าพยากรณ์ที่มีความแม่นยำเมื่อข้อมูลอนุกรมเวลาที่มีลักษณะคงที่ (Stationary) ข้อมูลอนุกรมเวลาเป็นเชิงเส้นตรง (Linear) และเป็นการพยากรณ์ระยะสั้น จากข้อจำกัดนี้จึงเป็นสาเหตุสำคัญทำให้ค่าพยากรณ์มีความคลาดเคลื่อนสูง ตัวแบบการเรียนรู้ของเครื่อง (Machine Learning) เช่น โครงข่ายประสาทเทียม (Artificial Neural Network; ANN) [6] วิธีซัพพอร์ตเวกเตอร์รีเกรสชัน (Support Vector Regression; SVR) [8] เป็นตัวแบบที่มีความยืดหยุ่นมี

ความแม่นยำสูงใช้ได้กับข้อมูลอนุกรมเวลาที่ไม่เป็นเชิงเส้น (Non-linear) รวมทั้งไม่มีข้อจำกัดเรื่องข้อตกลงเบื้องต้นทางสถิติ (Assumption) ในการสร้างตัวแบบพยากรณ์แต่มีข้อเสีย คือ ถ้าข้อมูลที่นำมาใช้ในการพยากรณ์มีจำนวนน้อยจะทำให้ค่าพยากรณ์มีความคลาดเคลื่อนสูง เพราะการพยากรณ์ตามแนวทางนี้มีการแบ่งข้อมูลออกเป็นสองชุด คือ ชุดข้อมูลฝึกฝน (Training Data Set) และชุดข้อมูลทดสอบ (Testing Data Set) ดังนั้น หากเลือกวิธีการพยากรณ์วิธีใดวิธีหนึ่งเพียงวิธีเดียวพยากรณ์ข้อมูลอนุกรมเวลาซึ่งมักมีโครงสร้างที่มีส่วนประกอบทั้งแบบเชิงเส้นตรงและแบบที่ไม่ใช่เชิงเส้นตรงจะไม่เหมาะสมเนื่องจากไม่สามารถอธิบายลักษณะการเคลื่อนไหวของข้อมูลอนุกรมเวลาได้ครบถ้วน จึงเกิดแนวคิดที่จะรวมจุดเด่นของทั้งสองตัวแบบเข้าด้วยกัน เรียกว่า ตัวแบบผสม (Hybrid Model) และ Zhang [9] เป็นคนแรกที่เสนอตัวแบบการพยากรณ์แบบผสมระหว่างตัวแบบ ARIMA กับตัวแบบ ANN เพื่อพยากรณ์ข้อมูลอนุกรมเวลาที่มียอดประกอบเชิงเส้นตรงและไม่ใช่เชิงเส้นตรงให้มีความแม่นยำมากขึ้น นักวิจัยจำนวนมากนำแนวคิดดังกล่าวไปประยุกต์ใช้พยากรณ์ข้อมูลอนุกรมเวลาพบว่า ตัวแบบผสมสามารถพยากรณ์ข้อมูลอนุกรมเวลาได้แม่นยำมากกว่าตัวแบบ ARIMA และตัวแบบ ANN เพียงตัวเดียว [10] อย่างไรก็ตามถึงแม้วิธีการพยากรณ์ ตัวแบบผสมตามวิธีของ Zhang [9] จะมีความแม่นยำสูงแต่ขั้นตอนการสร้างตัวแบบจะทำการวิเคราะห์แยกกันระหว่างการวิเคราะห์องค์ประกอบที่เป็นเชิงเส้นตรง (Linear Component) ด้วยตัวแบบ ARIMA และการวิเคราะห์องค์ประกอบที่ไม่เป็นเชิงเส้นตรง (Non - Linear Component) ด้วยตัวแบบ ANN จากนั้นนำค่าพยากรณ์มารวมกันในขั้นตอนสุดท้ายซึ่งไม่ได้วิเคราะห์ไปพร้อมกันทำให้ไม่สะดวกในการใช้งานเพื่อแก้ปัญหาที่ Khairalla และคณะ [11] ได้เสนอตัวแบบผสมระหว่างตัวแบบ ARIMA และตัวแบบ ANN โดยอาศัยวิธีการวิเคราะห์การถดถอย (Regression Analysis) โดยไม่จำเป็นต้องวิเคราะห์แยกเป็นสองขั้นตอนต่อเนื่องกันผลการวิจัยพบว่า ค่าพยากรณ์มีความแม่นยำมากกว่าตัวแบบผสม ARIMA – ANN ตามแนวคิดของ Zhang [9] และจะนำแนวคิดดังกล่าวมาประยุกต์ใช้ในการวิจัยครั้งนี้ ถึงแม้



การพยากรณ์ข้อมูลอนุกรมเวลาด้วยตัวแบบผสมจะให้ค่าพยากรณ์ที่มีความแม่นยำสูง แต่ปัจจุบันข้อมูลอนุกรมเวลาปริมาณการใช้พลังงานไฟฟ้ามีความซับซ้อนอาจเกิดจากปัจจัยหลายอย่าง ทำให้ไม่สามารถระบุความสัมพันธ์ที่ชัดเจนให้กับข้อมูลอนุกรมเวลาได้ ความซับซ้อนอาจปรากฏออกมาในลักษณะของความถี่หรือสัญญาณแกว่งไกว (Fluctuation) พร้อมกับข้อมูลอนุกรมเวลา นักวิจัยหลายท่านเชื่อว่าหากลดสัญญาณแกว่งไกวดังกล่าวลงแล้วนำข้อมูลอนุกรมเวลาชุดนั้นไปพยากรณ์ทำให้ได้ค่าพยากรณ์ที่มีความแม่นยำมากกว่าข้อมูลอนุกรมเวลาที่ยังไม่ได้ลดสัญญาณแกว่งไกว วิธีการลดสัญญาณแกว่งไกวของข้อมูลอนุกรมเวลาปริมาณการใช้พลังงานไฟฟ้าที่นิยมใช้ คือ วิธีการวิเคราะห์องค์ประกอบของข้อมูลอนุกรมเวลาเชิงประจักษ์ (Empirical Mode Decomposition; EMD) [2], [5], [6], [10] โดยวิธีการนี้ช่วยเพิ่มประสิทธิภาพให้กับตัวแบบการพยากรณ์ได้

จากข้อค้นพบที่กล่าวมาข้างต้นเพื่อให้ได้ค่าพยากรณ์ที่มีความแม่นยำ ทีมผู้วิจัยจึงสนใจที่จะนำเสนอตัวแบบผสมเพื่อใช้ในการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณหน่วยจ่ายไฟฟ้าบ้านอยู่อาศัยรายเดือนของประเทศไทย โดยจะนำแนวคิดการสร้างตัวแบบผสมของ Zhang [9] เป็นฐานคิดแต่จะปรับเปลี่ยนตัวแบบ ARIMA เป็นตัวแบบ SARIMA สำหรับพยากรณ์อนุกรมเวลาเชิงเส้นตรง เนื่องจากข้อมูลอนุกรมเวลาปริมาณการใช้พลังงานไฟฟ้ามีฤดูกาลเป็นส่วนประกอบ หากใช้ตัวแบบ ARIMA พยากรณ์ข้อมูลอนุกรมเวลาดังกล่าวอาจไม่เหมาะสม ส่วนการพยากรณ์อนุกรมเวลาไม่เป็นเชิงเส้นตรงยังคงใช้ตัวแบบ ANN เช่นเดิม การรวมค่าพยากรณ์ระหว่างตัวแบบเชิงเส้นตรงและไม่เป็นเชิงเส้นตรงจะใช้แนวคิดของ Khairalla และคณะ [11] ที่รวมค่าพยากรณ์ตัวแบบผสมโดยอาศัยวิธีการวิเคราะห์การถดถอย ซึ่งมีความแม่นยำมากกว่าวิธีการพยากรณ์แบบผสมของ Zhang [9] และก่อนที่นำข้อมูลอนุกรมเวลาไปใช้สร้างตัวแบบพยากรณ์ จะทำการลดสัญญาณแกว่งไกวให้กับข้อมูลอนุกรมเวลาด้วยวิธี EMD ซึ่งจะทำให้ได้ค่าพยากรณ์ที่มีความแม่นยำมากขึ้น ผลการวิจัยนี้เป็นประโยชน์ในทางปฏิบัติโดยเฉพาะหน่วยงานที่เกี่ยวข้องสามารถนำ ตัวแบบการพยากรณ์ที่ทีมผู้วิจัย

นำเสนอใช้เป็นแนวทางในการเลือกใช้ตัวแบบการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณหน่วยจ่ายไฟฟ้าบ้านอยู่อาศัยรายเดือนของประชาชนในอนาคตได้

2. วัสดุ อุปกรณ์และวิธีการวิจัย

2.1 การจัดเตรียมข้อมูล

ข้อมูลที่ใช้ในการทดลองครั้งนี้เป็นข้อมูลอนุกรมเวลาปริมาณหน่วยจ่ายไฟฟ้าบ้านอยู่อาศัยของการไฟฟ้าส่วนภูมิภาค [12] เป็นข้อมูลทุติยภูมิจากรายงานสถานการณ์การจำหน่ายไฟฟ้ารายเดือนของประเทศไทย จำนวน 4 ชุด ชุดละ 146 ค่า ตั้งแต่เดือนมกราคม 2553 ถึงเดือนกุมภาพันธ์ 2565 แบ่งเป็นภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคกลาง และภาคใต้ ดังแสดงในรูปที่ 4 ทีมผู้วิจัยแบ่งข้อมูลออกเป็น 2 ชุด คือ ข้อมูลชุดที่ 1 ข้อมูลฝึกฝน (Training Data Set) ตั้งแต่เดือนมกราคม 2553 ถึงเดือนธันวาคม 2562 จำนวน 120 ค่า ใช้สำหรับสร้างตัวแบบ และข้อมูลชุดที่ 2 ข้อมูลทดสอบ (Testing Data Set) ตั้งแต่เดือนมกราคม 2563 ถึงเดือนกุมภาพันธ์ 2565 จำนวน 26 ค่า ใช้สำหรับการทดสอบตัวแบบ เกณฑ์ในการแบ่งข้อมูลเพื่อวัดประสิทธิภาพการพยากรณ์ (Cross-Validation) ใช้เกณฑ์ร้อยละ 80 และ 20 [13] ตามลำดับ

2.2 รูปแบบการพยากรณ์ มีรายละเอียดดังนี้

2.2.1 วิธี EMD เป็นวิธีการแยกความถี่ของข้อมูลอนุกรมเวลาที่ไม่เป็นเชิงเส้นตรงและมีลักษณะเคลื่อนไหวไม่คงที่ (Non-Stationary) ออกจากกันเป็นการลดสัญญาณแกว่งไกวของข้อมูลอนุกรมเวลาให้มีลักษณะราบเรียบมากขึ้น ปัจจุบันมีการนำวิธี EMD มาประยุกต์ใช้สำหรับพยากรณ์การใช้พลังงานไฟฟ้าอย่างหลากหลาย [2], [5], [6] นำเสนอโดย Huang และคณะ [14] มีหลักการ คือ ทำการแยกออกข้อมูลอนุกรมเวลาออกเป็นฟังก์ชันภายใน (Intrinsic Mode Function; IMF) โดยมีเงื่อนไข คือ 1) ให้จำนวนจุดปลายสุด (Extrema Points) ซึ่งจะประกอบด้วย จุดสูงสุดและต่ำสุด และจำนวนจุดตัดแกนศูนย์ จะต้องมิต่างกัน หรือมีค่าต่างกันไม่เกินหนึ่ง 2) ณ จุดใด ๆ ของสัญญาณค่าเฉลี่ยของขอบบน

และขอบล่างจะมีค่าเท่ากับศูนย์ กระบวนการทำงานเพื่อสกัดชุดฟังก์ชันภายใน (IMFs) ของการแยกองค์ประกอบด้วยวิธี EMD ประกอบด้วย 2 ขั้นตอน คือ

1) การตรวจจับจุดปลายสุดของสัญญาณ (Extrema Detection) โดยการหาจุดสูงสุดและจุดต่ำสุดของแต่ละช่วงสัญญาณย่อย

2) การประมาณค่าขอบเขตพื้นผิว (Surface Interpolation) ของสัญญาณโดยการหาค่าประมาณของสัญญาณขอบบนและสัญญาณขอบล่าง แล้วคำนวณหาสัญญาณเฉลี่ย ดังสมการที่ (1) จากนั้นนำสัญญาณต้นฉบับลบออกด้วยสัญญาณเฉลี่ยจะได้สัญญาณผลลัพธ์ จากนั้นตรวจสอบว่าสัญญาณผลลัพธ์นั้นเป็นฟังก์ชันภายในหรือไม่ ถ้าผลปรากฏว่าสัญญาณผลลัพธ์ที่ได้ไม่เป็นฟังก์ชันภายในจะแทนที่สัญญาณต้นฉบับด้วยสัญญาณผลลัพธ์และทำซ้ำกระบวนการดังกล่าวข้างต้น หากสัญญาณผลลัพธ์ที่ได้เป็นฟังก์ชันภายในให้จัดเก็บสัญญาณผลลัพธ์นั้น และทำการสกัดฟังก์ชันภายใน ลำดับถัดไปจากสัญญาณต้นฉบับใหม่ที่ได้ออกจากการนำสัญญาณต้นฉบับเดิมลบด้วยฟังก์ชันภายในแล้วทำกระบวนการข้างต้นซ้ำจนสัญญาณต้นฉบับเป็นสัญญาณฟังก์ชันเดียว

$$m(t) = \frac{e_{\max} + e_{\min}}{2} \quad (1)$$

เมื่อ $m(t)$ แทน สัญญาณเฉลี่ย e_{\max} , e_{\min} แทนสัญญาณขอบบน และสัญญาณขอบล่าง ดังสมการที่ (2)

$$x(t) = \sum_{i=1}^n h_i(t) + r_N(t) \quad (2)$$

เมื่อ $x(t)$ แทน สัญญาณต้นฉบับ $h(t)$; $i = 1, 2, 3, \dots, N$ เป็นชุดฟังก์ชันภายใน (IMFs) และ $r_N(t)$ คือ ส่วนเหลือ (Residual)

2.2.2 การลดสัญญาณแกว่งไกวของข้อมูล

จากการวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธี EMD ในสมการที่ (2) จะได้ฟังก์ชันภายใน (IMF_{*s*}) ของข้อมูล $h_i(t)$ จากนั้นนำ $h_i(t)$ มาหาค่าเฉลี่ยโดยไม่รวมส่วนเหลือ ($r_N(t)$) แสดง

ดังสมการที่ (3)

$$aIMF = \frac{\sum_{i=1}^n h_i(t)}{n} \quad (3)$$

นำ $aIMF$ ลบสัญญาณต้นฉบับ แสดงดังสมการที่ (4)

$$Y_n(t) = x(t) - aIMF \quad (4)$$

เมื่อ $Y_n(t)$ คือ ชุดข้อมูลใหม่ที่มีการลดข้อมูลสัญญาณแกว่งไกว

2.2.3 ตัวแบบ SARIMA (Seasonal Autoregressive Integrated Moving Average; SARIMA)

ตัวแบบ SARIMA เป็นตัวแบบการพยากรณ์ที่มีความแม่นยำสูง พัฒนาขึ้นโดยบ็อกซ์-เจนกินส์ [15] สามารถเขียนสมการตัวแบบ SARIMA ดังสมการที่ (5)–(10)

$$\phi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^D Y_t = \delta + \theta_q(B)\Theta_q(B^s)\varepsilon_t \quad (5)$$

$$\delta = \mu\phi_p(B)\Theta_p(B^s) \quad (6)$$

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (7)$$

$$\Phi_p(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps} \quad (8)$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (9)$$

$$\Theta_q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_q B^{qs} \quad (10)$$

เมื่อ Y_t แทน อนุกรมเวลา ณ เวลา t

ε_t แทน ความคลาดเคลื่อนที่มีการแจกแจงปรกติมีค่าเฉลี่ยเท่ากับศูนย์มีความแปรปรวนคงที่และเป็นอิสระกัน ($\varepsilon_t \sim NID(0, \sigma^2)$)

t แทน คาบเวลาซึ่งมีค่าตั้งแต่ 1 ถึง n

s แทน ความยาวของคาบฤดูกาลใน 1 ฤดูกาล



d และ D แทน ลำดับที่ของการหาผลต่างภายในฤดูกาล และผลต่างระหว่างฤดูกาล

B แทน ตัวดำเนินการถอยหลัง (Backward Operator) โดยที่ $B^s Y_t = Y_{t-s}$ และ $BY_t = Y_{t-1}$

$\phi_p(B)$ แทน ตัวดำเนินการสหสัมพันธ์ในตัวแบบไม่มีฤดูกาล อันดับที่ p

$\Phi_p(B^s)$ แทน ตัวดำเนินการสหสัมพันธ์ในตัวแบบมีฤดูกาล อันดับที่ P

$\theta_q(B)$ แทน ตัวดำเนินการเฉลี่ยเคลื่อนที่แบบภายในฤดูกาล อันดับที่ q

$\Theta_Q(B^s)$ แทน ตัวดำเนินการเฉลี่ยเคลื่อนที่ระหว่างฤดูกาล อันดับที่ Q

การสร้างตัวแบบ SARIMA จะต้องทำการกำหนดรูปแบบ SARIMA (p,d,q)(P,D,Q)s การดำเนินการครั้งนี้ ทีมผู้วิจัยเขียนคำสั่งโดยใช้โปรแกรมภาษาอาร์สร้างฟังก์ชันเพื่อกำหนดพารามิเตอร์ที่เหมาะสม โดยเพิ่มอันดับพารามิเตอร์ของตัวแบบ p, d, P, D งานวิจัยนี้กำหนดอันดับของพารามิเตอร์ทั้ง 4 ตัวไว้ไม่เกิน 2 เนื่องจากหากกำหนดไว้สูงเกินไปอาจได้ตัวแบบไม่เหมาะสม [16] การเลือกตัวแบบจะนำตัวแบบที่เป็นไปได้ทั้งหมดไปพยากรณ์ล่วงหน้าและดำเนินการคำนวณค่าวัดประสิทธิภาพของตัวแบบแต่ละตัว การเลือกตัวแบบที่เหมาะสมจะพิจารณา จากเกณฑ์รากของความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error; RMSE) ตัวแบบใดที่มีค่า RMSE ต่ำสุดจะเป็นตัวแบบที่เหมาะสมสำหรับนำไปพยากรณ์ การสร้างตัวแบบมี 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 พิจารณานุกรมเวลาว่าคงที่หรือไม่ โดยพิจารณาจากกราฟของข้อมูลนุกรมเวลาชุดนั้น หากกราฟของข้อมูลนุกรมเวลาชุดนั้นมีแนวโน้มและฤดูกาลเป็นส่วนประกอบ แสดงว่า ข้อมูลนุกรมเวลาชุดนั้นมีสถานะไม่คงที่ กล่าวคือ มีค่าเฉลี่ยและความแปรปรวนไม่คงที่ งานวิจัยนี้ทดสอบความไม่คงที่ของข้อมูลด้วย Augmented Dickey – Fuller (ADF) หากผลการทดสอบไม่มีนัยสำคัญทางสถิติ แสดงว่าข้อมูลนุกรมเวลาชุดนั้นมีลักษณะไม่คงที่ เมื่อทราบว่ามีข้อมูลมีลักษณะไม่คงที่ที่จะต้องทำการแปลงข้อมูลให้มีลักษณะคงที่ ทั้งนี้ข้อมูลที่ศึกษามีความแปรปรวนไม่คงที่ ผู้วิจัยจึง

เลือกแปลงข้อมูลโดยการหาลอการิทึมและหาผลต่างของข้อมูลเพื่อกำจัดแนวโน้ม (d=1) และฤดูกาล (D=1) เพื่อให้ข้อมูลนุกรมเวลามีลักษณะคงที่ก่อนที่จะนำไปกำหนดตัวแบบ

ขั้นตอนที่ 2 กำหนดตัวแบบการพยากรณ์

ขั้นตอนที่ 3 ประเมินค่าพารามิเตอร์และตรวจสอบความเหมาะสมของตัวแบบ เมื่อได้ตัวแบบที่เหมาะสมก่อนที่จะนำตัวแบบไปใช้ในการพยากรณ์จะต้องมีการตรวจสอบตัวแบบ (Diagnostics Checking) มีทั้งหมด 4 ข้อ คือ 1) ทดสอบความคลาดเคลื่อนมีค่าเฉลี่ยเท่ากับศูนย์ โดยใช้สถิติทดสอบ t-test 2) ทดสอบความคลาดเคลื่อนมีการแจกแจงปกติ โดยใช้สถิติทดสอบ Kolmogorov – Smirnov 3) ทดสอบความเป็นอิสระกันของความคลาดเคลื่อนโดยใช้สถิติทดสอบ Ljung – Box Q- statistics และ 4) ความคลาดเคลื่อนมีความแปรปรวนคงที่โดยพิจารณาจากกราฟพล็อตกราฟดูลักษณะการกระจายระหว่างความคลาดเคลื่อนกับค่าพยากรณ์หากกราฟมีการกระจายไม่มีรูปแบบแสดงว่าความคลาดเคลื่อนมีความแปรปรวนคงที่

ขั้นตอนที่ 4 นำตัวแบบที่เหมาะสมไปใช้ในการพยากรณ์

2.2.4 ตัวแบบโครงข่ายประสาทเทียม (Artificial Neural Network Model; ANN)

ตัวแบบ ANN เป็นเทคนิคการเรียนรู้โดยใช้เครื่อง (Machine learning) [17] การประมวลผลของโครงข่ายประสาทเทียมจะเกิดขึ้นในหน่วยประมวลผลย่อย (Node) เมื่อมีข้อมูลนำเข้า (Input Data) เข้ามายังโครงข่ายข้อมูลนำเข้า แต่ละค่าจะถูกนำมาคูณกับค่าถ่วงน้ำหนัก (Weight) แล้วนำมารวมกันเพื่อส่งผ่านไปยังฟังก์ชันการแปลง (Transfer Function) หรือฟังก์ชันกระตุ้น (Activation Function) และส่งข้อมูลผลลัพธ์ไปยังหน่วยประมวลผลย่อยถัดไป [16] โดยโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น (Multi-layer Perceptron; MLP) จัดเป็นตัวแบบ ANN ที่นิยมใช้ในการพยากรณ์ข้อมูลนุกรมเวลาโครงสร้างของตัวแบบมีทั้งหมด 3 ชั้น ได้แก่ Input layer, Hidden Layer และ Output Layer แต่ละชั้นจะประกอบด้วยโหนดที่มีจำนวนต่างกัน และโหนดในชั้น Input Layer จะเชื่อมไปข้างหน้า

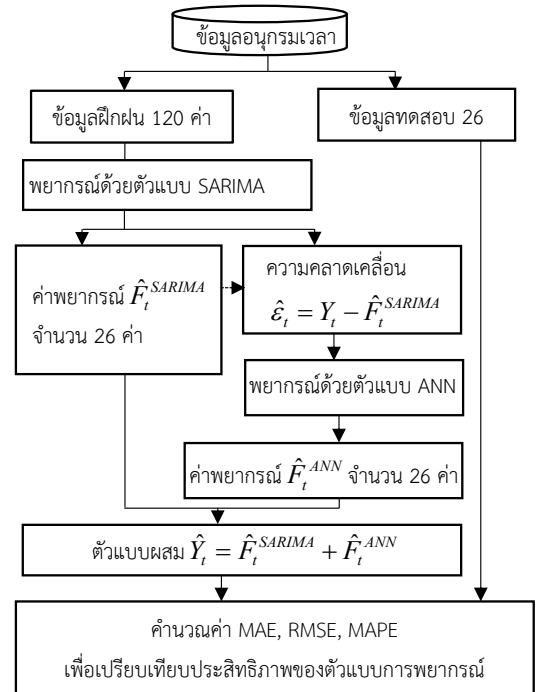
กับทุกโหนดในชั้นถัดไปตามลำดับจนถึง Output Layer ตัวแบบ ANN [9] แสดงดังสมการที่ (11)

$$y_i = \alpha + \sum_{j=1}^n \alpha_j \cdot f\left(\sum_{i=1}^m \beta_{ij} y_{i-1} + \beta_j\right) + \varepsilon_i \quad (11)$$

เมื่อ Y_t คือ ค่าสังเกตของข้อมูลอนุกรมเวลา ณ เวลา t , α คือ เวกเตอร์ของค่าถ่วงน้ำหนักระหว่างหน่วยประมวลผลย่อย จำนวน n หน่วยใน Hidden Layer กับหน่วยประมวลผลย่อยใน Output Layer, β_{ij} คือ ค่าถ่วงน้ำหนักระหว่างหน่วยประมวลผลย่อยแต่ละหน่วยใน Input Layer ซึ่งมีทั้งหมด m หน่วย กับหน่วยประมวลผลย่อยใน Hidden Layer โดยที่ $i = 1, 2, \dots, m, j = 1, 2, \dots, n, f$ คือ ฟังก์ชันกระตุ้นงานวิจัยนี้ใช้ฟังก์ชันกระตุ้น Sigmoid Logistic Function เนื่องจากทดลองใช้แล้วพบว่า มีความเหมาะสมและมีความแม่นยำมากที่สุด โดยมีสมการคือ $f(x) = 1/(1 + e^{-x})$ และ ε_t คือ ความคลาดเคลื่อน ณ เวลา t การกำหนดค่าพารามิเตอร์ของโครงข่ายประสาทเทียมในการเรียนรู้ในครั้งนี้เป็นแบบแพร่กลับ (Backpropagation Learning) และทำการปรับปรุงองคกอลิเทียมด้วยฟังก์ชัน Levenberg- Marquardt เพื่อให้การปรับค่าน้ำหนักและค่าเอนเอียงในเส้นเชื่อมแต่ละชั้นของโครงข่ายให้มีความรวดเร็วและได้ค่าที่มีความเหมาะสม จากนั้นทำการปรับค่าผลรวมของนิวรอนในชั้นซ่อนด้วยฟังก์ชันถ่ายโอนคือ ฟังก์ชันซิกมอยด์แบบโลจิสติกส์ (Sigmoid Logistic Function) มีผลลัพธ์อยู่ในช่วง $[-1, 1]$ มีอัตราการเรียนรู้ (Learning Rate) เท่ากับ 0.0001 จำนวนรอบในการฝึกฝนการเรียนรู้ (Epoch) สูงสุด 1,000 รอบ และใช้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error; MSE) เป็นเกณฑ์ในการเลือกนิวรอนในชั้นซ่อนที่เหมาะสมให้กับโครงข่าย

2.2.5 ตัวแบบผสม SARIMA-ANN (Hybrid Model; SARIMA-ANN)

ตัวแบบผสมตามแนวคิดของ Zhang [9] มีข้อสมมติว่าข้อมูลอนุกรมเวลาที่นำวิเคราะห์มี 2 องค์ประกอบ ได้แก่ องค์ประกอบที่เป็นเชิงเส้นตรง (Linear Component) และองค์ประกอบที่ไม่เป็นเชิงเส้นตรง (Non-linear Component) การพยากรณ์องค์ประกอบที่เป็นเชิงเส้นตรง Zhang [9]



รูปที่ 1 ตัวแบบผสม SARIMA – ANN

ประยุกต์ใช้ตัวแบบ ARIMA แต่งานวิจัยนี้จะประยุกต์ใช้ตัวแบบ SARIMA เนื่องจากข้อมูลที่นำมาศึกษามีฤดูกาลเป็นส่วนประกอบ และนำมารวมกับค่าพยากรณ์จากตัวแบบ ANN ซึ่งเป็นตัวแบบที่พยากรณ์ข้อมูลอนุกรมเวลาที่ไม่เป็นเชิงเส้นตรงได้ดี โดยตัวแบบผสมตามแนวคิดของ Zhang [9] แสดงดังรูปที่ 1 สามารถเขียนได้ในรูปสมการเชิงคณิตศาสตร์ดังสมการที่ (12)

$$Y_t = F_t^{SARIMA} + F_t^{ANN} + \varepsilon_t \quad (12)$$

เมื่อ Y_t แทน ค่าสังเกตของข้อมูลอนุกรมเวลา ณ เวลา t

F_t^{SARIMA} แทน ข้อมูลส่วนที่เป็นฟังก์ชันเชิงเส้นตรงที่พยากรณ์ด้วยตัวแบบ SARIMA

F_t^{ANN} แทน ข้อมูลส่วนที่ไม่เป็นฟังก์ชันเชิงเส้นตรงที่พยากรณ์ด้วยตัวแบบ ANN

ε_t แทน ความคลาดเคลื่อน ณ เวลา t ชั้นตอนและวิธี



การสร้างตัวแบบ มีรายละเอียดดังนี้

ขั้นตอนที่ 1 จัดเตรียมข้อมูลตามขั้นตอนที่ 2.1

ขั้นตอนที่ 2 สร้างและวิเคราะห์ข้อมูลอนุกรมเวลา ชุดที่ 1 ด้วยตัวแบบ SARIMA เพื่อพยากรณ์ข้อมูลส่วนที่เป็น ฟังก์ชันเชิงเส้นตรง แทนด้วย \hat{F}_t^{SARIMA} จำนวน 26 ค่า

ขั้นตอนที่ 3 คำนวณค่าส่วนเหลือ (Residual) จากผล การพยากรณ์ด้วยตัวแบบ SARIMA โดยที่ $\varepsilon_t = Y_t - \hat{F}_t^{SARIMA}$

ขั้นตอนที่ 4 นำส่วนเหลือที่ได้จากตัวแบบ SARIMA ซึ่งเป็นข้อมูลในส่วนที่ไม่เป็นฟังก์ชันเชิงเส้นตรง พยากรณ์ด้วย ตัวแบบ ANN แทนด้วย \hat{F}_t^{ANN} จำนวน 26 ค่า

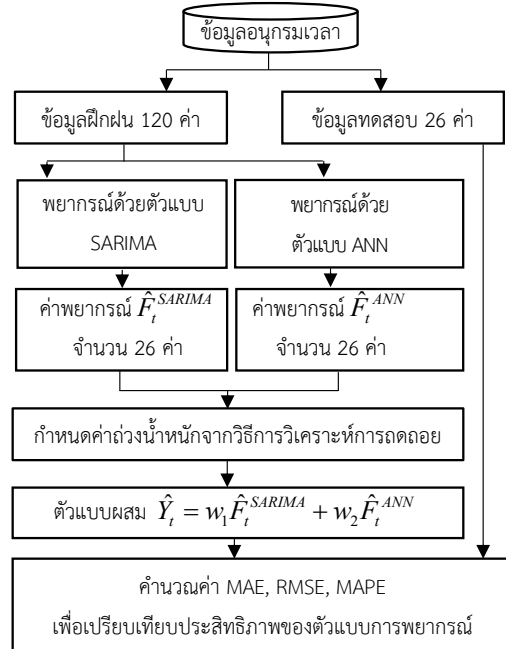
ขั้นตอนที่ 4 รวมค่าพยากรณ์ในขั้นตอนที่ 2 และ ขั้นตอนที่ 4 จะได้ค่าพยากรณ์รวมของตัวแบบ SARIMA-ANN และเปรียบเทียบประสิทธิภาพของตัวแบบกับข้อมูลชุดที่ 2 สมการพยากรณ์ของตัวแบบ แสดงดังสมการที่ (13)

$$\hat{Y}_t = \hat{F}_t^{SARIMA} + \hat{F}_t^{ANN} \quad (13)$$

2.2.6 ตัวแบบผสม SARIMA-ANN โดยใช้วิธีการถดถอย (Hybrid Model: SARIMA-ANN-REG)

จากแนวคิดการสร้างตัวแบบผสม SARIMA-ANN ตามแนวคิดของ Zhang [9] ในหัวข้อที่ 2.2.5 พบว่า การสร้าง ตัวแบบการพยากรณ์แยกวิเคราะห์ระหว่างองค์ประกอบที่เป็นเชิงเส้นตรงด้วยตัวแบบ SARIMA และการวิเคราะห์องค์ ประกอบที่ไม่เป็นเชิงเส้นตรงด้วยตัวแบบ ANN จากนั้นนำ ค่าพยากรณ์มารวมกันในขั้นตอนสุดท้ายซึ่งไม่ได้วิเคราะห์ พร้อมกันทำให้ไม่สะดวกในการใช้งาน ด้วยเหตุนี้ Khairalla และคณะ [11] จึงเสนอตัวแบบผสมแบบใหม่ขึ้นโดยรวม ค่าพยากรณ์องค์ประกอบเชิงเส้นตรงที่จากตัวแบบ ARIMA แต่งานวิจัยนี้จะใช้ตัวแบบ SARIMA เนื่องจากข้อมูลที่ศึกษา มีฤดูกาลเป็นส่วนประกอบ และค่าพยากรณ์องค์ประกอบ ที่ไม่เป็นเชิงเส้นตรงด้วยตัวแบบ ANN โดยอาศัยวิธีการ วิเคราะห์การถดถอย (Regression Analysis) ซึ่งไม่จำเป็นต้องวิเคราะห์แยกเป็นสองขั้นตอนต่อเนื่องกัน แสดงดังรูปที่ 2 สามารถเขียนได้ในรูปสมการเชิงคณิตศาสตร์ ดังสมการที่ (14)

$$Y_t = w_1 F_t^{SARIMA} + w_2 F_t^{ANN} + \varepsilon_t \quad (14)$$



รูปที่ 2 ตัวแบบผสม SARIMA – ANN - REG

เมื่อ Y_t แทน ค่าสังเกตของข้อมูลอนุกรมเวลา ณ เวลา t , F_t^{SARIMA} แทน ข้อมูลส่วนที่เป็นฟังก์ชันเชิงเส้นตรงที่ พยากรณ์ด้วยตัวแบบ SARIMA, F_t^{ANN} แทน ข้อมูลส่วนที่ไม่เป็นฟังก์ชันเชิงเส้นตรงที่พยากรณ์ด้วยตัวแบบ ANN, w_1 แทน ค่าถ่วงน้ำหนักของข้อมูลส่วนที่เป็นฟังก์ชันเชิงเส้นตรง คำนวณได้จากการวิเคราะห์การถดถอย w_2 แทน ค่าถ่วงน้ำหนักของข้อมูลส่วนที่เป็นฟังก์ชันไม่เป็นเชิงเส้นตรง คำนวณได้จากการวิเคราะห์การถดถอย และ ε_t แทน ความคลาดเคลื่อน ณ เวลา t

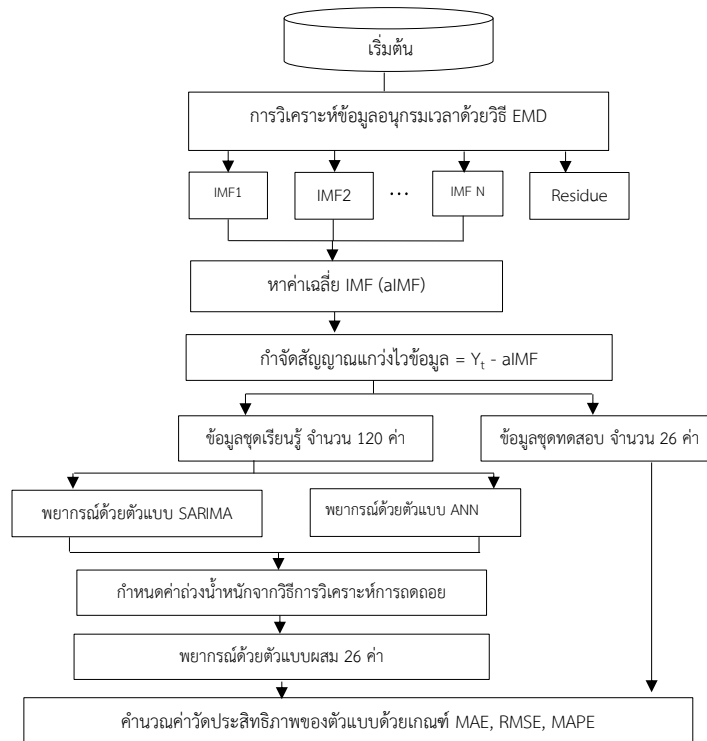
ขั้นตอนการสร้างตัวแบบผสมมีรายละเอียดดังนี้

ขั้นตอนที่ 1 จัดเตรียมข้อมูลตามขั้นตอนที่ 2.1

ขั้นตอนที่ 2 สร้างและวิเคราะห์ข้อมูลอนุกรมเวลา ชุดที่ 1 ด้วยตัวแบบ SARIMA เพื่อพยากรณ์ข้อมูลส่วนที่เป็น ฟังก์ชันเชิงเส้นตรง คือ \hat{F}_t^{SARIMA} จำนวน 26 ค่า

ขั้นตอนที่ 3 สร้างและวิเคราะห์ข้อมูลอนุกรมเวลา ชุดที่ 1 ด้วยตัวแบบ ANN เพื่อพยากรณ์ข้อมูลส่วนที่ไม่เป็น ฟังก์ชันเชิงเส้นตรง คือ \hat{F}_t^{ANN} จำนวน 26 ค่า

ขั้นตอนที่ 4 นำค่า \hat{F}_t^{SARIMA} และ \hat{F}_t^{ANN} มาเป็นตัวแปร



รูปที่ 3 ขั้นตอนการสร้างตัวแบบ Propose

อิสระในสมการถดถอย และให้ค่าสังเกต Y_t เป็น ตัวแปรตาม
 ขั้นตอนที่ 5 ใช้ค่าสัมประสิทธิ์การถดถอยมาเป็นค่า
 ถ่วงน้ำหนัก w_1 และ w_2 และคำนวณหาค่าพยากรณ์รวมของ
 ตัวแบบ SARIMA – ANN – REG และเปรียบเทียบ
 ประสิทธิภาพของตัวแบบกับข้อมูลชุดที่ 2 สมการพยากรณ์
 ของตัวแบบ แสดงดังสมการที่ (15)

$$\hat{Y}_t = w_1 \hat{F}_t^{SARIMA} + w_2 \hat{F}_t^{ANN} \quad (15)$$

2.2.7 ตัวแบบ Propose

จากแนวคิดการสร้างตัวแบบผสมของ Zhang [9] ใน
 หัวข้อที่ 2.2.5 ซึ่งมีข้อบกพร่อง คือ ต้องทำการวิเคราะห์ใน
 สองขั้นตอนต่อเนื่องกันโดยไม่ได้วิเคราะห์ไปพร้อมกัน และ
 แนวคิดการสร้างตัวแบบผสมของ Khairalla และคณะ [11]
 ในหัวข้อที่ 2.2.6 ได้เสนอตัวแบบผสมโดยรวมค่าพยากรณ์
 จากแนวคิดของ Zhang [9] โดยอาศัยวิธีการวิเคราะห์

การถดถอย (Regression Analysis) จากสองแนวคิดดังกล่าว
 ผู้วิจัยจึงนำเสนอตัวแบบผสมแบบใหม่ โดยนำเทคนิควิธี EMD
 มาทำการกรองข้อมูลอนุกรมเวลาเพื่อลดสัญญาณแกว่งไว
 ของข้อมูลก่อนทำการพยากรณ์ โดยผู้วิจัยจะพัฒนาต่อยอด
 จากแนวคิดของ Khairalla และคณะ [11] รายละเอียดการ
 สร้างตัวแบบแสดงดังรูปที่ 3 สามารถเขียนตัวแบบได้ ดัง
 สมการที่ (16)

$$Y_t = w_1 F_t^{EMDSARIMA} + w_2 F_t^{EMDANN} + \varepsilon_t \quad (16)$$

เมื่อ Y_t แทน ค่าสังเกตของข้อมูลอนุกรมเวลา ณ เวลา
 t ที่กรองด้วยวิธี EMD, $F_t^{EMDSARIMA}$ แทน ข้อมูลส่วนที่เป็น
 ฟังก์ชันเชิงเส้นตรงที่พยากรณ์ด้วยตัวแบบ SARIMA, F_t^{EMDANN}
 แทน ข้อมูลส่วนที่ไม่เป็นฟังก์ชันเชิงเส้นตรงที่พยากรณ์ด้วย
 ตัวแบบ ANN

w_1 แทน ค่าถ่วงน้ำหนักของข้อมูลส่วนที่เป็นฟังก์ชัน



เชิงเส้นตรง คำนวณได้จากการวิเคราะห์การถดถอย

w_2 แทน ค่าถ่วงน้ำหนักของข้อมูลส่วนที่เป็นฟังก์ชัน

ไม่เป็นเชิงเส้นตรง คำนวณได้จากการวิเคราะห์การถดถอย

ε_t แทน ความคลาดเคลื่อน ณ เวลา t

ขั้นตอนการสร้างตัวแบบผสมมีรายละเอียดดังนี้

ขั้นตอนที่ 1 วิเคราะห์ข้อมูลด้วยวิธี EMD (แสดง

ตัวอย่างดังรูปที่ 5 ข้อมูลอนุกรมเวลาปริมาณหน่วยจำหน่าย

ไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนของภาคเหนือ และ

ชุดข้อมูลอื่น ๆ วิเคราะห์ในลักษณะเดียวกัน) เพื่อลดสัญญาณ

แกว่งไวกของข้อมูลอนุกรมเวลาตามขั้นตอนที่ 2.2.1 และ

ขั้นตอนที่ 2.2.2

ขั้นตอนที่ 2 จัดเตรียมข้อมูลตามขั้นตอนที่ 2.1

ขั้นตอนที่ 3 วิเคราะห์ข้อมูลอนุกรมเวลาด้วยตัวแบบ

SARIMA เพื่อพยากรณ์ข้อมูลส่วนที่เป็นฟังก์ชันเชิงเส้นตรง

คือ $\hat{F}_t^{EMDSARIMA}$ จำนวน 26 ค่า

ขั้นตอนที่ 4 วิเคราะห์ข้อมูลอนุกรมเวลาด้วยตัวแบบ

ANN เพื่อพยากรณ์ข้อมูลส่วนที่ไม่เป็นฟังก์ชันเชิงเส้นตรง คือ

\hat{F}_t^{EMDANN} จำนวน 26 ค่า

ขั้นตอนที่ 5 นำค่า $\hat{F}_t^{EMDSARIMA}$ และ \hat{F}_t^{EMDANN} เป็น

ตัวแปรอิสระในสมการถดถอย และให้ค่าสังเกต Y_t เป็น

ตัวแปรตาม

ขั้นตอนที่ 6 ใช้ค่าสัมประสิทธิ์การถดถอยมาเป็น

ค่าถ่วงน้ำหนัก w_1 และ w_2 และคำนวณหาค่าพยากรณ์รวม

ของตัวแบบ Propose และเปรียบเทียบประสิทธิภาพของ

ตัวแบบกับข้อมูลชุดที่ 2 สมการพยากรณ์ของตัวแบบ แสดง

ดังสมการที่ (17)

$$\hat{Y}_t = w_1 \hat{F}_t^{EMDSARIMA} + w_2 \hat{F}_t^{EMDANN} \quad (17)$$

2.2.8 เกณฑ์การประเมินประสิทธิภาพของตัวแบบ

การประเมินประสิทธิภาพของตัวแบบงานวิจัยนี้มี

ทั้งหมด 3 เกณฑ์ [6] ดังนี้

1) ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean

Absolute Error; MAE) มีสูตรทางคณิตศาสตร์ แสดงดัง

สมการที่ (18)

$$MAE = \sum_{t=1}^n |Y_t - \hat{Y}_t| / n \quad (18)$$

2) รากของความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error; $RMSE$) มีสูตรทางคณิตศาสตร์ แสดงดังสมการที่ (19)

$$RMSE = \sqrt{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2 / n} \quad (19)$$

3) ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error; $MAPE$) มีสูตรทางคณิตศาสตร์แสดงดังสมการที่ (20)

$$MAPE = 100 \times \sum_{t=1}^n |1 - \hat{Y}_t / Y_t| / n \quad (20)$$

โดยที่ค่า MAE , $RMSE$ และ $MAPE$ มีค่าน้อย แสดงว่าตัวแบบมีประสิทธิภาพสูง

3.ผลการวิจัย

3.1 ผลการวิเคราะห์ลักษณะความเคลื่อนไหวของข้อมูลอนุกรมเวลา

จากการพิจารณาลักษณะการเคลื่อนไหวของข้อมูล

อนุกรมเวลาปริมาณหน่วยจำหน่ายไฟฟ้าประเภทบ้านอยู่

อาศัยรายเดือนของประเทศไทย จำนวน 4 ชุด ได้แก่ ภาคเหนือ

ภาคตะวันออกเฉียงเหนือ ภาคกลางและภาคใต้ เป็นข้อมูล

ย้อนหลัง 12 ปี 2 เดือน ตั้งแต่เดือนมกราคม พ.ศ. 2553 ถึง

เดือนกุมภาพันธ์ พ.ศ. 2565 จำนวน 146 ค่า ดังรูปที่ 4 พบว่า

อนุกรมเวลาทั้ง 4 ชุด มีแนวโน้มและความผันแปรตาม

ฤดูกาลเป็นส่วนประกอบเนื่องจากลักษณะของข้อมูลมีค่า

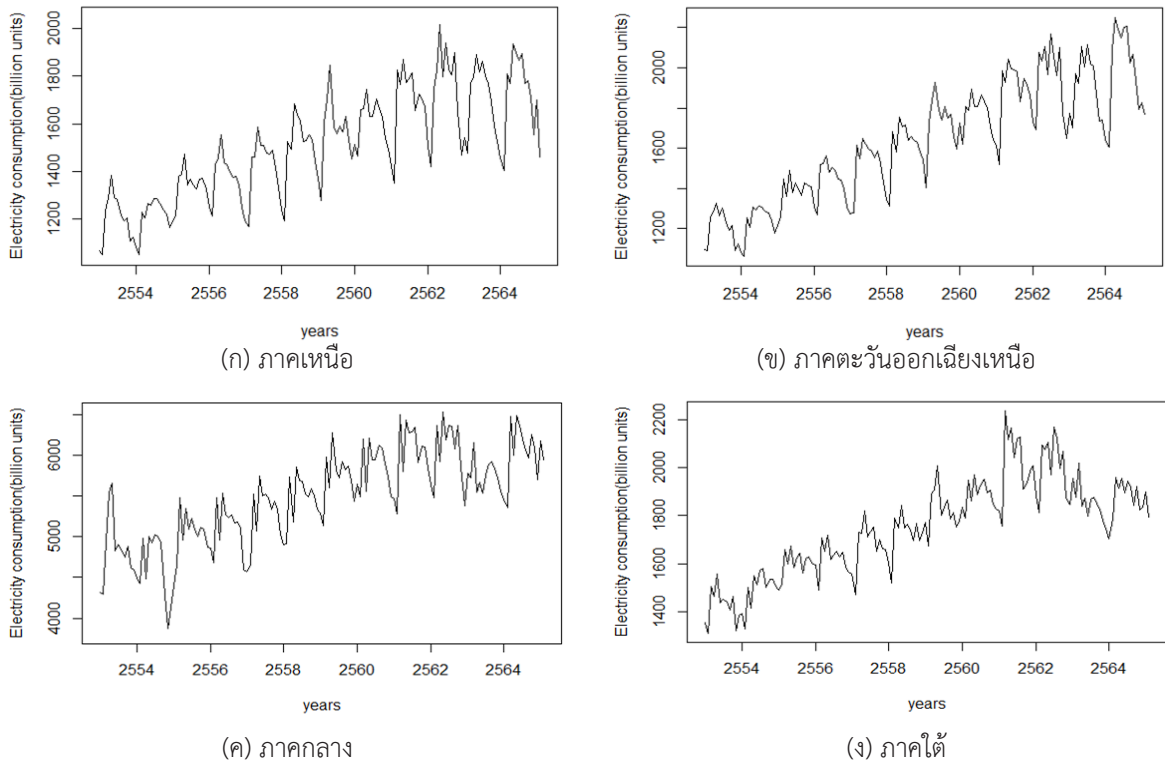
เพิ่มขึ้นและลดลงในช่วงเวลาเดียวกันเมื่อเวลาเปลี่ยนแปลงไป

3.3 ผลพยากรณ์โดยใช้ตัวแบบ SARIMA

3.3.1 จากการพิจารณาข้อมูลปริมาณหน่วยจำหน่าย

ไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนของประเทศไทยทั้ง

4 ชุด (ดังรูปที่ 4) พบว่า กราฟมีแนวโน้มและฤดูกาลเป็น



รูปที่ 4 ลักษณะการเคลื่อนไหวอนุกรมเวลาปริมาณหน่วยจำหน่ายไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนของประเทศไทย ทั้ง 4 ชุม ประกอบด้วย (ก) ภาคเหนือ (ข) ภาคตะวันออกเฉียงเหนือ (ค) ภาคกลาง และ (ง) ภาคใต้

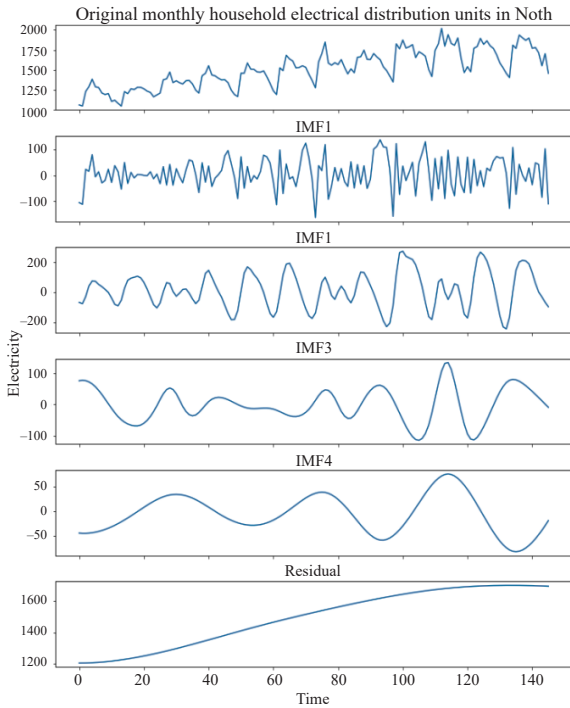
ส่วนประกอบโดยมีลักษณะการกระจายไม่สม่ำเสมอกราฟ มีลักษณะเพิ่มขึ้นและลดลงตามเวลาที่เปลี่ยนไป แสดงว่า ข้อมูลมีลักษณะไม่คงที่กล่าวคือ มีค่าเฉลี่ยไม่คงที่และความแปรปรวนไม่คงที่ สอดคล้องกับผลการทดสอบความคงที่ของข้อมูลด้วย ADF (ตารางที่ 1) พบว่า ค่า ADF ของข้อมูลทั้ง 4 ชุม ไม่มีนัยสำคัญทางสถิติ ($p\text{-value} > 0.05$) เมื่อทราบว่าคุณสมบัติอนุกรมเวลาทั้ง 4 ชุม มีลักษณะไม่คงที่ ผู้วิจัยจึงแปลงข้อมูลโดยการหาลอการิทึมเนื่องจากข้อมูลมีความแปรปรวนไม่คงที่ จากนั้นทำการหาผลต่างของชุดข้อมูลเพื่อกำจัดแนวโน้มและฤดูกาล โดยกำหนดอันดับของการหาผลต่างแนวโน้มและฤดูกาลเท่ากับ 1 ($d=1, D=1$) ก่อนนำไปสร้างตัวแบบ SARIMA ผลการวิเคราะห์ข้อมูล ADF (ตารางที่ 1) เมื่อมีการแปลงข้อมูลพบว่า หลังแปลงข้อมูลอนุกรมเวลาทั้ง 4 ชุม มีลักษณะคงที่ โดยมีค่า ADF มีนัยสำคัญทางสถิติทั้ง 4 ชุม ($p\text{-value} < 0.05$) ทีมผู้วิจัยจึงนำชุดข้อมูลไปใช้ในการ

กำหนดตัวแบบในขั้นตอนต่อไป

ตารางที่ 1 ค่า ADF ปริมาณหน่วยจำหน่ายไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนของประเทศไทย ทั้ง 4 ชุม

ชุดข้อมูล	ก่อนแปลงข้อมูล		หลังแปลงข้อมูล	
	ADF	p-value	ADF	p-value
ภาคเหนือ	0.9015	0.9001	-5.4213	0.01
ภาคตะวันออกเฉียงเหนือ	1.2771	0.9341	-5.7898	0.01
ภาคกลาง	1.9881	0.9721	-5.9087	0.01
ภาคใต้	2.1002	0.8821	-4.8211	0.01

3.3.2 กำหนดรูปแบบของตัวแบบ SARIMA ผู้วิจัยเขียนคำสั่งด้วยโปรแกรมภาษารัสสร้างฟังก์ชัน เพื่อกำหนดพารามิเตอร์ที่เหมาะสม โดยเลือกอันดับ p, d, P, D แบบวนซ้ำจากนั้นนำตัวแบบที่เป็นไปได้ทั้งหมดพยากรณ์ค่าในอนาคต



รูปที่ 5 การวิเคราะห์ข้อมูลอนุกรมเวลาเชิงประจักษ์ปริมาณหน่วยจำหน่ายไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนของภาคเหนือ

และคำนวณค่า RMSE ตัวแบบที่ให้ค่า RMSE ต่ำสุดเป็นตัวแบบที่มีความแม่นยำมากที่สุด ผลการเลือก ตัวแบบ แสดงดังตารางที่ 2 และนำไปพยากรณ์ต่อไป

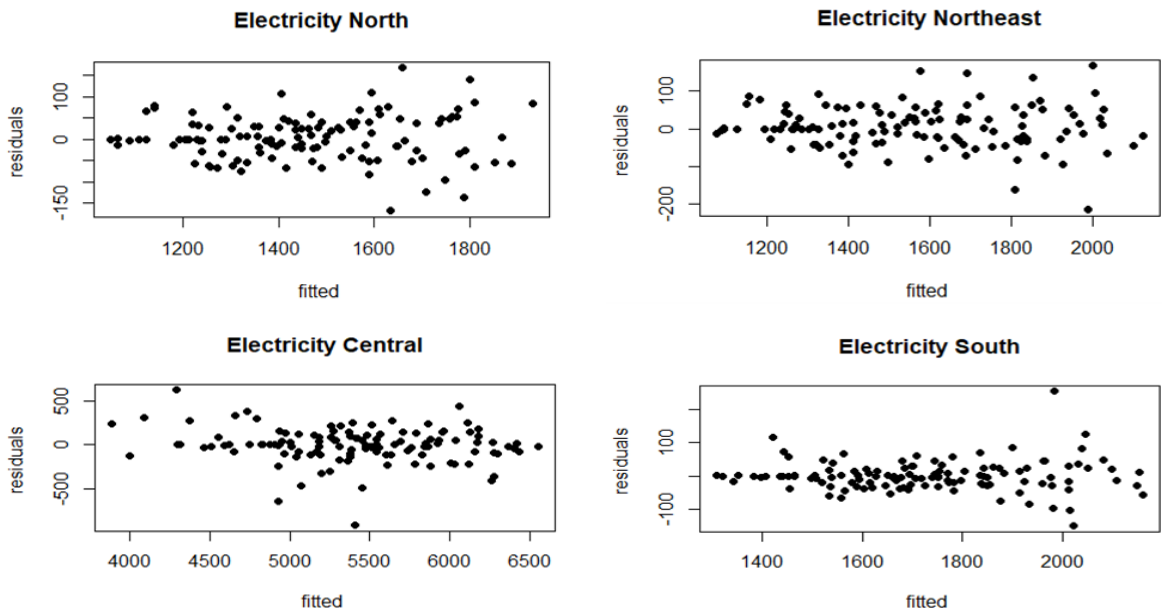
ตารางที่ 2 การประมาณของค่าพารามิเตอร์ของตัวแบบ SARIMA ของข้อมูลทั้ง 4 ชุด

ชุดข้อมูล		Estimate	Std. Error	z-value	p-value
ภาคเหนือ SARIMA (1,1,1) (0,1,1)12	AR1	0.4933	0.0919	5.3648	<0.001
	MA1	-0.9999	0.03499	-28.5760	<0.001
	SMA1	-0.6787	0.0882	-7.6931	<0.001
ภาคตะวันออกเฉียงเหนือ SARIMA (1,1,1) (0,1,1)12	AR1	0.4464	0.0967	4.6140	<0.001
	MA1	-0.9999	0.0403	-24.8172	<0.001
	SMA1	-0.6754	0.0994	-6.7929	<0.001
ภาคกลาง SARIMA (0,1,0) (0,1,1)12	SMA1	-0.6518	0.1053	-6.1905	<0.001
ภาคใต้ SARIMA (1,1,1) (0,1,1)12	AR1	0.5751	0.0851	6.7562	<0.001
	MA1	-1.0000	0.0552	-18.1174	<0.001
	SMA1	-0.75839	0.1045	-7.2576	<0.001

3.3.3 การตรวจสอบความเหมาะสมของตัวแบบ พิจารณาตามข้อตกลงเบื้องต้นของตัวแบบ SARIMA พบว่าตัวแบบ SARIMA ของข้อมูลทั้ง 4 ชุดที่กำหนดขึ้นมีความเหมาะสมที่จะนำไปใช้ในการพยากรณ์ เนื่องจากผลการตรวจสอบตัวแบบผ่านข้อตกลงเบื้องต้นทุกข้อ (ตารางที่ 3 และรูปที่ 6)

3.4 ผลพยากรณ์โดยใช้ตัวแบบ ANN

การกำหนดตัวแบบ ANN ผู้วิจัยประยุกต์ใช้โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น (Multi - Layer Perceptron) ซึ่งโปรแกรมภาษาอาร์มีฟังก์ชัน MLP () ซึ่งอยู่ใน Package “neuralnet” สำหรับกำหนดข้อมูลนำเข้า (Input) และชั้นซ่อน (Hidden) การพยากรณ์ด้วยตัวแบบ ANN จะต้องมีการแปลงข้อมูลปริมาณหน่วยจำหน่ายไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนทั้ง 4 ชุด ให้เป็นชุดข้อมูล $D = \{(x_i, y_i)\}_{i=1}^n$ การแปลงข้อมูลจะทำการจัดเรียงข้อมูลอนุกรมเวลาให้อยู่ในรูปค่าสังเกตแล้วไปคอลัมน์ รูปที่ 7 เป็นข้อมูลย้อนหลังตั้งแต่คอลัมน์ที่ 1 ถึง คอลัมน์ที่ $m-1$ ใช้เป็นข้อมูลป้อนเข้า (Input Data) และคอลัมน์ที่ m เป็นคอลัมน์เป้าหมาย (Target Data) เพื่อให้ตัวแบบเรียนรู้และสร้างฟังก์ชันการทำนายลักษณะของข้อมูล โดยแบ่งเป็นชุดข้อมูลการเรียนรู้และชุดข้อมูลทดสอบ ตามหัวข้อที่ 2.1 การใช้ค่าสังเกตแล้ว งานวิจัยนี้เลือกใช้จำนวนคอลัมน์ตั้งแต่ 2-12



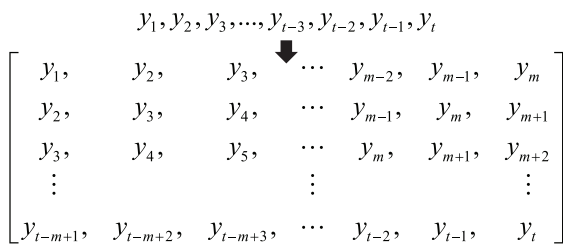
รูปที่ 6 แผนภาพการกระจายระหว่างความคลาดเคลื่อนกับค่าพยากรณ์จากตัวแบบ SARIMA

คอลัมน์ เนื่องจากผู้วิจัยเก็บข้อมูลมาเป็นรายเดือน จำนวน คาบเวลาใน 1 ฤดูกาลเท่ากับ 12 คาบ ตามรอบปฏิทิน หาก ใช้ข้อมูลย้อนหลังหลายช่วงเวลามากเกินไปค่าพยากรณ์ที่ได้ อาจจะไม่เหมาะสม [18] การวิจัยครั้งนี้ผู้วิจัยได้กำหนดโหนด ซึ่งเป็นข้อมูลป้อนเข้าเท่ากับ 11 คอลัมน์ ข้อมูลเป้าหมาย 1 คอลัมน์ จากข้อมูลชุดฝึกฝน 80% ในการฝึกการเรียนรู้ให้ กับโครงข่ายและทำการปรับจำนวนนิวรอนในชั้นซ่อนจาก

1-10 นิวรอนเพื่อหานิวรอนในชั้นซ่อนที่เหมาะสม โดยพบว่า โครงข่ายที่เหมาะสมของข้อมูลทั้ง 4 ชุด คือ โครงข่ายที่มี จำนวนนิวรอน 7 นิวรอน (ANN-7) เนื่องจากมีค่า RMSE ต่ำสุด เมื่อทดสอบประสิทธิภาพและความแม่นยำของ โครงข่ายประสาทเทียมเรียบร้อยแล้ว ผู้วิจัยทำการพยากรณ์ และคำนวณค่าวัดประสิทธิภาพการพยากรณ์ตามเกณฑ์ที่ กำหนด รายละเอียดแสดงดังตารางที่ 4

ตารางที่ 3 ตรวจสอบความคลาดเคลื่อนของตัวแบบ SARIMA ของข้อมูลทั้ง 4 ชุด

ชุดข้อมูล	ค่าเฉลี่ยเท่ากับ 0		การแจกแจงปกติ		ความเป็นอิสระกัน		
	t-value	p-value	KS -value	p-value	Ljung - Box Q-statistics		
					Q-statistics	df	p-value
ภาคเหนือ SARIMA(1,1,1)(0,1,1)12	-0.5465	0.5858	0.058	0.200	19.798	19	0.4068
ภาคตะวันออกเฉียงเหนือ SARIMA(1,1,1)(0,1,1)12	0.9262	0.3562	0.054	0.200	15.621	19	0.6824
ภาคกลาง SARIMA(0,1,0)(0,1,1)12	-0.20651	0.8367	0.048	0.200	16.146	19	0.5028
ภาคใต้ SARIMA(1,1,1)(0,1,1)12	0.0514	0.9591	0.052	0.200	29.332	19	0.0609



รูปที่ 7 การแปลงข้อมูลสำหรับสร้างตัวแบบ ANN

3.5 การพยากรณ์ด้วยตัวแบบผสม

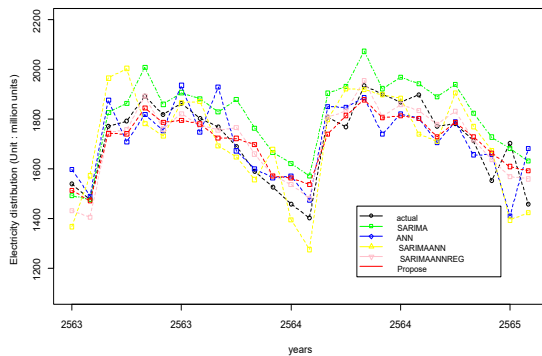
ผู้วิจัยได้ดำเนินการสร้างและพยากรณ์ด้วยตัวแบบผสมตามหัวข้อที่ 2.2.5 ตัวแบบผสม SARIMA-ANN (Hybrid Model; SARIMA-ANN) หัวข้อที่ 2.2.6 ตัวแบบผสม SARIMA-ANN โดยใช้วิธีการถดถอย (Hybrid Model; SARIMA-ANN-REG) และหัวข้อที่ 2.2.7 ตัวแบบ Propose จากนั้นคำนวณค่าวัดประสิทธิภาพของตัวแบบด้วยเกณฑ์ที่กำหนดและเปรียบเทียบความแม่นยำของตัวแบบ ผลการพยากรณ์แสดงดังตารางที่ 4 และรูปที่ 8

ตารางที่ 4 เปรียบเทียบประสิทธิภาพของตัวแบบการพยากรณ์ข้อมูลอนุกรมเวลาทั้ง 4 ชุด

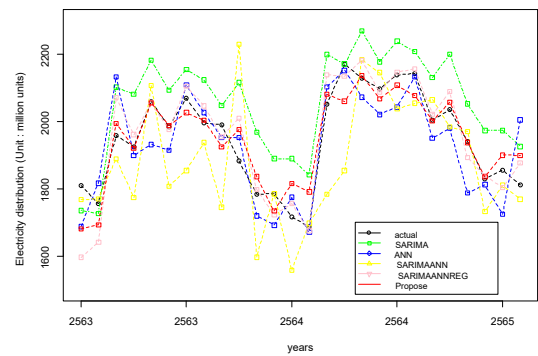
ชุดข้อมูล	ตัวแบบ		MAE	RMSE	MAPE
ภาคเหนือ	SARIMA	ข้อมูลฝึกฝน	40.2911	52.9617	2.6761
		ข้อมูลทดสอบ	102.3563	117.0285	0.0614
	ANN	ข้อมูลฝึกฝน	14.81035	21.4563	0.0104
		ข้อมูลทดสอบ	83.02812	104.8123	0.0491
	SARIMA-ANN	ข้อมูลฝึกฝน	12.7828	23.6874	0.0089
		ข้อมูลทดสอบ	95.9909	120.8323	0.0571
	SARIMA-ANN-REG	ข้อมูลฝึกฝน	14.7704	21.2723	0.0103
		ข้อมูลทดสอบ	52.2585	63.3858	0.0319
	Propose	ข้อมูลฝึกฝน	15.0694	21.9558	0.0105
		ข้อมูลทดสอบ	36.8414	45.5295	0.0223
ภาคตะวันออกเฉียงเหนือ	SARIMA	ข้อมูลฝึกฝน	42.2898	57.0434	2.6339
		ข้อมูลทดสอบ	113.997	134.5277	0.0613
	ANN	ข้อมูลฝึกฝน	19.4629	28.3004	0.0119
		ข้อมูลทดสอบ	90.0447	106.7405	0.0465
	SARIMA-ANN	ข้อมูลฝึกฝน	8.4841	13.5133	0.0058
		ข้อมูลทดสอบ	119.4112	164.3126	0.0600
	SARIMA-ANN-REG	ข้อมูลฝึกฝน	19.4777	28.2803	0.0120
		ข้อมูลทดสอบ	59.5759	72.9275	0.0315
	Propose	ข้อมูลฝึกฝน	22.3322	29.8006	0.0138
		ข้อมูลทดสอบ	45.7825	59.3136	0.0243
ภาคกลาง	SARIMA	ข้อมูลฝึกฝน	134.2612	200.4097	2.5680
		ข้อมูลทดสอบ	309.753	378.3268	0.0522
	ANN	ข้อมูลฝึกฝน	48.0917	67.6811	0.0086
		ข้อมูลทดสอบ	366.5042	461.9065	0.0624
	SARIMA-ANN	ข้อมูลฝึกฝน	37.7527	77.4437	0.0073
		ข้อมูลทดสอบ	275.2587	343.2777	0.0464

ตารางที่ 4 เปรียบเทียบประสิทธิภาพของตัวแบบการพยากรณ์ข้อมูลอนุกรมเวลาทั้ง 4 ชุด (ต่อ)

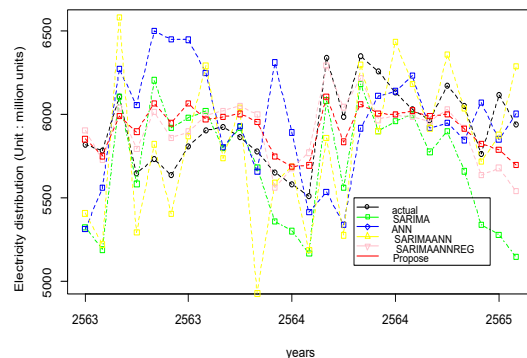
ชุดข้อมูล	ตัวแบบ		MAE	RMSE	MAPE
ภาคกลาง (ต่อ)	SARIMA-ANN-REG	ข้อมูลฝึกฝน	48.5443	67.3634	0.0087
		ข้อมูลทดสอบ	198.7011	239.2149	0.0337
	Propose	ข้อมูลฝึกฝน	51.2166	65.5601	0.0095
		ข้อมูลทดสอบ	160.461	187.9241	0.0271
ภาคใต้	SARIMA	ข้อมูลฝึกฝน	30.4329	46.4341	1.70154
		ข้อมูลทดสอบ	247.4068	258.2104	0.1333
	ANN	ข้อมูลฝึกฝน	19.1674	25.1689	0.0111
		ข้อมูลทดสอบ	170.6039	197.4117	0.0921
	SARIMA-ANN	ข้อมูลฝึกฝน	11.9318	19.0480	0.0071
		ข้อมูลทดสอบ	94.19038	118.8568	0.0509
	SARIMA-ANN-REG	ข้อมูลฝึกฝน	17.4301	23.1700	0.0101
		ข้อมูลทดสอบ	46.2149	60.6099	0.0248
Propose	ข้อมูลฝึกฝน	13.3209	17.4069	0.0078	
	ข้อมูลทดสอบ	33.5263	43.5202	0.0178	



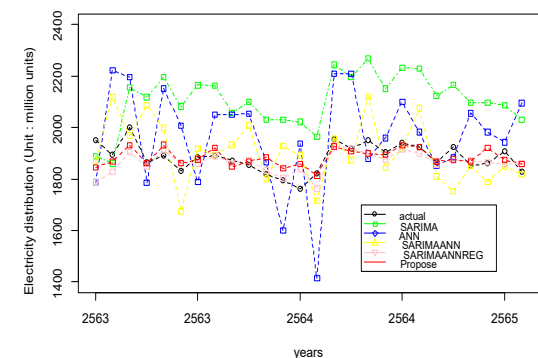
(ก) ภาคเหนือ



(ข) ภาคตะวันออกเฉียงเหนือ



(ค) ภาคกลาง



(ง) ภาคใต้

รูปที่ 8 เปรียบเทียบค่าจริงกับค่าพยากรณ์ข้อมูลอนุกรมเวลาปริมาณหน่วยจำหน่ายไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนของประเทศไทย ทั้ง 4 ชุด (ก) ภาคเหนือ (ข) ภาคตะวันออกเฉียงเหนือ (ค) ภาคกลาง และ (ง) ภาคใต้

ชนกร สุทธิสนธิ์ และ ภาณุเดช เพียรความสุข, “ตัวแบบผสมสำหรับพยากรณ์ข้อมูลอนุกรมเวลาปริมาณหน่วยจำหน่ายไฟฟ้าบ้านอยู่อาศัยรายเดือนของประเทศไทย.”



4. อภิปรายผลและสรุป

จากการนำเสนอวิธีการสร้างตัวแบบผสมเพื่อลดความคลาดเคลื่อนของค่าพยากรณ์ เมื่อเปรียบเทียบกับตัวแบบ SARIMA ตัวแบบ ANN ตัวแบบผสม SARIMA – ANN ตามแนวคิดของ Zhang [9] และตัวแบบผสม SARIMA-ANN-REG ตามแนวคิดของ Khairalla และคณะ [11] กรณีที่ยังไม่มีการกรองข้อมูล ตารางที่ 4 รูปที่ 8 พบว่า ตัวแบบ Propose เป็นตัวแบบผสมที่ทีมผู้วิจัยพัฒนาขึ้น โดยประยุกต์ใช้วิธี EMD ทำการกรองข้อมูลก่อนนำไปสร้างตัวแบบผสมด้วยตัวแบบ SARIMA-ANN-REG มีประสิทธิภาพและมีความแม่นยำมากกว่าตัวแบบ SARIMA ตัวแบบ ANN ตัวแบบผสม SARIMA – ANN และตัวแบบผสม SARIMA-ANN-REG ในทุกชุดข้อมูลและทุกเกณฑ์การวัดประสิทธิภาพ ข้อค้นพบจากการวิจัยนี้บ่งชี้ว่า การกรองข้อมูลด้วยวิธี EMD เพื่อลดสัญญาณแกว่งไวของข้อมูลก่อนนำไปสร้างตัวแบบการพยากรณ์จะทำให้ค่าพยากรณ์มีความแม่นยำมากขึ้นและมีความแม่นยำมากกว่าข้อมูลที่ยังไม่มีการกรองข้อมูล [2], [5], [6], [10] อย่างไรก็ตาม หากมองในแง่ประสิทธิภาพการพยากรณ์ตัวแบบผสม SARIMA-ANN-REG ตามแนวคิดของ Khairalla และคณะ [11] เป็นตัวแบบที่น่าสนใจเนื่องจากเป็นตัวแบบที่มีความแม่นยำรองจากตัวแบบผสมที่ทีมผู้วิจัยนำเสนอทั้ง 4 ชุด ตัวแบบนี้เป็นตัวแบบผสมที่แก้ข้อบกพร่องของตัวแบบผสมตามแนวคิดของ Zhang [9] ทำให้ได้ตัวแบบที่มีประสิทธิภาพมากกว่า เมื่อพิจารณาตัวแบบอนุกรมเวลาเชิงเดี่ยว 2 ตัวแบบ ได้แก่ ตัวแบบ SARIMA และตัวแบบ ANN พบว่า ตัวแบบ ANN มีประสิทธิภาพและมีความแม่นยำมากกว่าตัวแบบ SARIMA ในชุดข้อมูลปริมาณหน่วยจำหน่ายไฟฟ้าบ้านอยู่อาศัยรายเดือนภาคเหนือ ภาคตะวันออกเฉียงเหนือ และภาคใต้ ส่วนภาคกลางพบว่า ตัวแบบ SARIMA มีประสิทธิภาพมากกว่า ตัวแบบ ANN ตามลำดับ ในส่วนของการประยุกต์ใช้งานหากต้องการได้ค่าพยากรณ์ที่มีความแม่นยำสูงแนะนำให้เลือกใช้ตัวแบบผสมที่ทีมผู้วิจัยนำเสนอ ทั้งนี้เพราะเป็นตัวแบบที่มีประสิทธิภาพและแม่นยำมากที่สุด เนื่องจากตัวแบบผสมมีความซับซ้อนสูง กรณีที่ผู้ใช้งานต้องการตัวแบบที่ไม่ซับซ้อนมากอาจเลือกใช้ตัวแบบ

อนุกรมเวลาเชิงเดี่ยว เช่น ตัวแบบ SARIMA และตัวแบบ ANN ที่มีความซับซ้อนน้อยกว่าตัวแบบผสมที่ทีมผู้วิจัยพัฒนาขึ้น งานวิจัยนี้สามารถนำไปประยุกต์ใช้กับข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวคล้ายกันเพื่อเพิ่มประสิทธิภาพการพยากรณ์ ข้อพึงระวังในการใช้ตัวแบบ คือ ข้อมูลที่นำมาศึกษาต้องมีฤดูกาลเป็นส่วนประกอบ และต้องมีการตรวจสอบข้อตกลงเบื้องต้นของตัวแบบ SARIMA ก่อนการพยากรณ์ทุกครั้ง เพื่อสร้างความน่าเชื่อถือของตัวแบบการพยากรณ์ในการศึกษาครั้งนี้ทีมผู้วิจัยพิจารณาเฉพาะปัจจัยด้านเวลาที่มีผลต่อปริมาณหน่วยจำหน่ายไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนของประเทศไทยเท่านั้น หากสามารถนำมาปัจจัยอื่นมาร่วมพิจารณา หรือนำมาเป็นตัวแปรนำเข้าเพื่อการเรียนรู้และปรับปรุงตัวแบบจะทำให้ตัวแบบผสมมีความแม่นยำมากขึ้น

5. กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณการไฟฟ้าส่วนภูมิภาคที่อนุเคราะห์ข้อมูลปริมาณหน่วยจำหน่ายไฟฟ้าประเภทบ้านอยู่อาศัยรายเดือนของประเทศไทยสำหรับการศึกษานี้ และขอขอบคุณสาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี ที่อนุเคราะห์สถานที่สำหรับการทำวิจัย

เอกสารอ้างอิง

- [1] P. Suksawang, S. Suphachan, and K. Kaewnuch, "Electricity consumption forecasting in thailand using hybrid model sarima and gaussian process with combine kernel function technique," *International Journal of Energy Economics and Policy*, vol. 8, no. 4, pp. 98–109, 2018.
- [2] Y. Yaslan and B. Bican, "Empirical mode decomposition based denoising method with support vector regression for time series prediction: A case study for electricity load forecasting," *Measurement*, vol. 103, pp. 52–61, 2017.

- [3] G.-F. Fan, X. Wei, Y.-T. Li, and W.-C. Hong, "Forecasting electricity consumption using a novel hybrid model," *Sustainable Cities and Society*, vol. 61, pp. 1–16, 2020.
- [4] P. Du, J. Wang, W. Yang, and T. Niu, "Multi-step ahead forecasting in electrical power system using a hybrid forecasting system," *Renewable Energy*, vol. 122, pp. 533–550, 2018.
- [5] Z. Liu, X. Wang, Q. Zhang, and C. Huang, "Empirical mode decomposition-based hybrid ensemble model for electrical energy consumption forecasting of the cement grinding process," *Measurement*, no. 138, pp. 314–324, 2019.
- [6] T. Sutthison and S. Haruay, "The development of a hybrid model for forecasting time series data of monthly household electrical distribution units of people in the Northeast, Thailand," *The Journal of KMUTNB*, vol. 30, no.4, pp. 630– 644, 2020 (in Thai).
- [7] M.-R. Kazemzadeh, A. Amjadian, and T. Amraee, "A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting," *Energy*, vol. 204, pp. 1–19, 2020.
- [8] P. Jiang, R. Li, N. Liu, and Y. Gao, "A novel composite electricity demand forecasting framework by data processing and optimized support vector machine," *Applied Energy*, vol. 260, pp. 1–15, 2020.
- [9] P. G. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [10] Ü. Ç. Büyüksahin and S. Ertekin, "Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition," *Neurocomputing*, vol. 361, pp. 151–163, 2019.
- [11] M. Khairalla, X. Ning and N. T. AL-Jallad, "Hybrid forecasting scheme for financial time-series data using neural network and statistical methods," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 9, pp. 310–327, 2017.
- [12] Provincial Electricity Authority. (2022, March). Electricity distribution units Situation. Electricity Economy Division. Bangkok, Thailand [Online]. (in Thai). Available: <http://peaoc.pea.co.th/ped/>
- [13] A. Gholamy, V. Kreinovich, and O. Kosheleva, "Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation," *Departmental Technical Reports (CS)*, no. 1209, 2018.
- [14] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of The Royal Society a Mathematical, Physical and Engineering Sciences*, 1998, pp. 903–995.
- [15] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G.M. Ljung, *Time Series Analysis: Forecasting and Control*, 4th ed. John Wiley and Sons, Inc., New Jersey, 2008.
- [16] N. Kulvanich, "Improving prediction accuracy of time series data using ARIMA-ANN hybrid model and regression analysis," *Thai Science and Technology Journal (TSTJ)*, vol. 28, no. 12, pp. 2101 – 2112, 2020 (in Thai).
- [17] W. Liu, W. D. Liu, and J. Gu, "Forecasting oil production using ensemble empirical mode



decomposition based Long Short-Term Memory neural network,” *Journal of Petroleum Science and Engineering*, vol. 189, pp. 1–14, 2020.

[18] T. Sujjaviriyasup, “Artificial neural network model for forecasting monthly price of maize in Thailand,” *Srinakharinwirot Science Journal*, vol. 34, no. 1, pp. 91–107, 2018 (in Thai).