



เทคนิคการพยากรณ์แบบผสมของซัพพอร์ตเวกเตอร์แมชชีนและขั้นตอนวิธีเชิงพันธุกรรมสำหรับการพยากรณ์ปริมาณความต้องการพลังงานไฟฟ้าสูงสุดรายปีของประเทศไทย

ธรณินทร์ สัจจวิริยทรัพย์*

สาขาวิศวกรรมโลจิสติกส์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยหอการค้าไทย

* ผู้พิมพ์ประสานงาน โทรศัพท์ 08-5993-4296 อีเมล: thoranin_suj@utcc.ac.th DOI: 10.14416/j.kmutnb.2017.07.002

รับเมื่อ 4 กรกฎาคม 2559 ตอรับเมื่อ 15 กันยายน 2559 เผยแพร่ออนไลน์ 3 กรกฎาคม 2560

© 2017 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

บทคัดย่อ

ไฟฟ้าถือเป็นแหล่งพลังงานที่สำคัญในการขับเคลื่อนการเจริญเติบโตทางเศรษฐกิจของประเทศไทยซึ่งถูกใช้เพิ่มสูงขึ้นอย่างต่อเนื่องในช่วงหลายสิบปีที่ผ่านมา เพื่อการบริหารจัดการระบบการจ่ายไฟฟ้าให้มีเสถียรภาพนั้น ความต้องการพลังงานไฟฟ้าสูงสุดที่เกิดขึ้นในอนาคตซึ่งมีลักษณะไม่แน่นอนมีบทบาทสำคัญในการบริหารจัดการ การคาดการณ์ปริมาณความต้องการใช้ในอนาคตจึงต้องอาศัยตัวแบบพยากรณ์ที่เหมาะสมเพื่อให้ค่าพยากรณ์ที่ได้มีความเหมาะสมและเป็นประโยชน์ในการวางแผนและบริหารจัดการได้ ตัวแบบการพยากรณ์แบบผสมของซัพพอร์ตเวกเตอร์แมชชีนและขั้นตอนวิธีเชิงพันธุกรรมถูกพัฒนาขึ้นในงานวิจัยนี้เพื่อพยากรณ์ปริมาณความต้องการพลังงานไฟฟ้าสูงสุด โดยแรงจูงใจของการพัฒนาตัวแบบผสมนี้เกิดจากการใช้ความสามารถในการสร้างตัวแบบที่ซับซ้อนของซัพพอร์ตเวกเตอร์แมชชีนเพื่อทำนายค่าในอนาคต แต่เนื่องจากประสิทธิภาพของซัพพอร์ตเวกเตอร์แมชชีนขึ้นอยู่กับ การเลือกค่าพารามิเตอร์ที่เหมาะสม ขั้นตอนวิธีการเชิงพันธุกรรมจึงถูกนำมาใช้ประโยชน์ในการหาค่าที่เหมาะสมของซัพพอร์ตเวกเตอร์แมชชีน นอกจากนี้ตัวแบบผสมที่นำเสนอได้ถูกเปรียบเทียบกับประสิทธิภาพการพยากรณ์กับตัวแบบดั้งเดิม ได้แก่ ตัวแบบอาร์มาและซัพพอร์ตเวกเตอร์แมชชีนโดยปราศจากการหาค่าพารามิเตอร์ที่เหมาะสมที่สุดด้วยการประเมินประสิทธิภาพ 6 วิธี จากผลการวิจัยพบว่า ตัวแบบผสมที่นำเสนอมีประสิทธิภาพในการพยากรณ์ที่ถูกต้องและแม่นยำกว่าตัวแบบดั้งเดิม โดยสามารถลดค่าความคลาดเคลื่อนของตัวแบบอาร์มาและซัพพอร์ตเวกเตอร์แมชชีนได้อย่างน้อยร้อยละ 7.78 และ 5.72 ตามลำดับ

คำสำคัญ: ตัวแบบการพยากรณ์ผสม, เทคนิคการเลือกค่าพารามิเตอร์ที่เหมาะสม, การเรียนรู้ของเครื่อง, ซัพพอร์ตเวกเตอร์แมชชีน, วิธีการเชิงพันธุกรรม

การอ้างอิงบทความ: ธรณินทร์ สัจจวิริยทรัพย์, "เทคนิคการพยากรณ์แบบผสมของซัพพอร์ตเวกเตอร์แมชชีนและขั้นตอนวิธีเชิงพันธุกรรมสำหรับการพยากรณ์ปริมาณความต้องการพลังงานไฟฟ้าสูงสุดรายปีของประเทศไทย," วารสารวิชาการพระจอมเกล้าพระนครเหนือ, ปีที่ 27, ฉบับที่ 3, หน้า 453-465, ก.ค.-ก.ย. 2560

Hybrid Model of Support Vector Machine and Genetic Algorithm for Forecasting the Annual Peak Electricity Demand of Thailand

Thorarin Sujjaviriyasup*

Department of Logistics Engineering , Faculty of Engineering, University of the Thai Chamber of Commerce, Bangkok, Thailand

* Corresponding Author, Tel. 08-5993-4296, E-mail: thorarin_suj@utcc.ac.th DOI: 10.14416/j.kmutnb.2017.07.002

Received 4 July 2016; Accepted 15 September 2016; Published online: 3 July 2017

© 2017 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

Abstract

One of several energy resources is electricity, which plays a significant role in driving the economic growth of Thailand and has been increasingly used in past decades continuously. In order to gain a reliable and efficient system of power supply handling, the future peak electricity demand, which is inevitably uncertain, will play a crucial role. Therefore, the peak electricity demand has to be predicted by using the most appropriate forecasting models. In this research, a hybrid model of the Support Vector Machine (SVM) and Genetic Algorithm (GA) is proposed in order to forecast future peak demand. The motivation of the proposed model was to develop a complex model that takes advantage of the capability of the SVM model to formulate a predictive complex model. Even though the SVM model is more attractive in terms of forecasting, its performance relies heavily on the appropriate selection of the SVM hyper-parameters. In this regard, the GA was introduced to search for the optimal hyper-parameters of the SVM model. In order to evaluate the performance of the proposed model, it was compared with traditional single forecasting models (i.e., ARIMA and SVM) based on six measures of forecast accuracy. The empirical results indicated that the proposed model outperforms those compared models. Furthermore, the proposed model is able to reduce the errors of the ARIMA model and the SVM model at at least 7.78% and 5.72%, respectively.

Keywords: Hybrid Model, Optimal Parameter Selection Technique, Machine Learning, Support Vector Machine, Genetic Algorithm

1. บทนำ

จากการเจริญเติบโตทางเศรษฐกิจของประเทศไทย ในช่วงหลายสิบปีที่ผ่านมา ส่งผลให้เกิดความต้องการใช้พลังงานอย่างมากในการขับเคลื่อนเศรษฐกิจ ไฟฟ้าเป็นแหล่งพลังงานที่สำคัญประเภทหนึ่งที่ถูกใช้ในทุกภาคส่วนของเศรษฐกิจ จากความสำคัญของพลังงานไฟฟ้า การบริหารจัดการด้านการจ่ายไฟฟ้าที่ล้มเหลวจะส่งผลกระทบต่ออย่างมากต่อการพัฒนาทางเศรษฐกิจของประเทศ ดังในหลายประเทศ [1], [2] ที่ประสบปัญหาดังกล่าวจากการบริหารจัดการด้านพลังงานไฟฟ้าที่ไม่มีประสิทธิภาพ เพื่อให้การบริหารจัดการด้านพลังงานไฟฟ้าเป็นไปอย่างมีประสิทธิภาพ ปริมาณความต้องการพลังงานไฟฟ้าสูงสุดที่จะเกิดขึ้นในอนาคตจึงมีบทบาทที่สำคัญในการวางแผนการผลิตและการบริหารการจ่ายไฟฟ้า เพื่อให้ระบบการผลิตและการจ่ายไฟฟ้าเกิดเสถียรภาพ แต่ปริมาณความต้องการใช้ที่จะเกิดขึ้นในอนาคตนั้นมีรูปแบบที่ไม่คงที่และไม่แน่นอน ตัวแบบการพยากรณ์ต่างๆ จึงเข้ามามีบทบาทสำคัญเพื่อประมาณค่าปริมาณความต้องการและเป็นข้อมูลสารสนเทศที่มีประโยชน์ในการวางแผนและบริหารจัดการ ถึงแม้ว่าจะมีข้อถกเถียงเกี่ยวกับประสิทธิภาพของการใช้ตัวแบบการพยากรณ์ในการคาดการณ์ปริมาณความต้องการไฟฟ้าที่จะเกิดขึ้น [1] แต่นักวิจัยก็ยังสนับสนุนการใช้การพยากรณ์เพื่อใช้เป็นแนวทางในการบริหารจัดการที่จะดำเนินการ สำหรับตัวแบบพยากรณ์ที่ใช้ในการคาดการณ์สามารถแบ่งได้หลายประเภท เช่น ตัวแบบทางสถิติ ตัวแบบการเรียนรู้ของเครื่อง และตัวแบบผสม เป็นต้น ตัวแบบทางสถิติเป็นตัวแบบที่ง่ายที่สุด ซึ่งอาศัยการวิเคราะห์แนวโน้มของข้อมูลในอดีตเพื่อสร้างความสัมพันธ์ในการพยากรณ์ค่าในอนาคต ตัวแบบอาร์มา (ARIMA) เป็นตัวแบบหนึ่งในตัวแบบทางสถิติ [1], [3], [4] ซึ่งมีความโดดเด่นและถูกใช้ในการพยากรณ์ในหลายๆ งานวิจัย แต่ตัวแบบอาร์มายังคงมีข้อจำกัดที่สำคัญ [5] คือ ความไม่ยืดหยุ่นในการใช้งานและความถูกต้องที่ยังคงไม่มากนักในบางสถานการณ์ เช่น ข้อมูลมีลักษณะไม่เชิงเส้นตรง เป็นต้น จึงมีความสนใจ

พัฒนาตัวแบบการเรียนรู้ของเครื่องอย่างกว้างขวาง [6]–[8] เพื่อเพิ่มความถูกต้องในการพยากรณ์และมีความยืดหยุ่นในการใช้งานมากขึ้น โดยจุดเด่นของตัวแบบการเรียนรู้ของเครื่องมีดังต่อไปนี้ ประการแรก ตัวแบบการเรียนรู้ของเครื่องมีความยืดหยุ่นในการพยากรณ์ไม่เชิงเส้นตรง โดยใช้ความสามารถในการแปลงข้อมูลจากไม่เชิงเส้นตรงไปสู่ข้อมูลที่เป็นเชิงเส้นตรง ซึ่งมีคุณลักษณะที่สอดคล้องกับลักษณะข้อมูลไม่เชิงเส้นตรง เพื่อให้สามารถพยากรณ์ได้ในหลายลักษณะข้อมูล ประการที่สอง ตัวแบบการเรียนรู้ของเครื่องถูกสร้างขึ้นจากการใช้คุณลักษณะข้อมูลในการสร้างตัวแบบจึงไม่จำเป็นต้องใช้สมมุติฐานมากเหมือนตัวแบบอาร์มา และประการที่สาม ตัวแบบการเรียนรู้ของเครื่องมีความยืดหยุ่นในการใช้งานซึ่งสามารถปรับเปลี่ยนไปตามการใช้งานและระดับความถูกต้องที่ต้องการ จากจุดเด่นข้างต้นตัวแบบการเรียนรู้ของเครื่องจึงได้รับความสนใจและถูกนำมาใช้ในการพยากรณ์ในงานวิจัยต่างๆ โดยซัพพอร์ตเวกเตอร์แมชชีนเป็นตัวแบบหนึ่งในตัวแบบการเรียนรู้ของเครื่องต่างๆ ซึ่งซัพพอร์ตเวกเตอร์แมชชีนมีจุดเด่นที่น่าสนใจและถูกใช้ในการพยากรณ์ [6]–[8] คือ ซัพพอร์ตเวกเตอร์แมชชีนถูกสร้างจากหลักการลดความเสี่ยงเชิงโครงสร้างให้น้อยที่สุด (Principle of Structural Risk Minimization) จากหลักการนี้ จึงทำให้ได้ตัวแบบที่เหมาะสมที่สุดและให้ผลการพยากรณ์ที่เหมาะสมที่สุดเพียงค่าเดียว ดังนั้นซัพพอร์ตเวกเตอร์แมชชีนจึงประสบความสำเร็จในการพยากรณ์ [9]–[11] และมีประสิทธิภาพดีกว่าตัวแบบโครงข่ายประสาทเทียมแบบดั้งเดิมและตัวแบบทางสถิติต่างๆ ถึงแม้ว่าซัพพอร์ตเวกเตอร์แมชชีนจะได้รับความสนใจและถูกพัฒนาในการพยากรณ์มากขึ้นอย่างต่อเนื่อง แต่ประสิทธิภาพของซัพพอร์ตเวกเตอร์แมชชีนขึ้นอยู่กับ การเลือกค่าพารามิเตอร์ต่างๆ ให้เหมาะสม [12], [13] การนำขั้นตอนวิธีเชิงพันธุกรรมซึ่งเป็นเมตาฮีวิริสติกส์อย่างหนึ่งในการค้นหาพารามิเตอร์ที่เหมาะสมและลดระยะเวลาในการค้นหาให้น้อยลง เพื่อค้นหาพารามิเตอร์ที่เหมาะสมที่สุดสำหรับตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

ที่จะใช้ทำนายค่าในอนาคต ซึ่งผลของการนำเมตาฮีวิวิสติค มาใช้ในการค้นหาค่าพารามิเตอร์ที่เหมาะสมของซัพพอร์ตเวกเตอร์แมชชีนนั้น แสดงให้เห็นว่าการใช้งานของตัวแบบผสมให้ค่าความถูกต้องของการพยากรณ์ที่ดีกว่าซัพพอร์ตเวกเตอร์แมชชีนแบบดั้งเดิมโดยตัวแบบพยากรณ์แบบดั้งเดิม ได้แก่ ตัวแบบอาร์มาและซัพพอร์ตเวกเตอร์แมชชีน โดยปราศจากการหาค่าพารามิเตอร์ที่เหมาะสมที่สุด [12], [13]

งานวิจัยนี้จึงได้เสนอตัวแบบผสมระหว่างซัพพอร์ตเวกเตอร์แมชชีนและขั้นตอนวิธีการเชิงพันธุกรรมและเปรียบเทียบประสิทธิภาพการทำนายกับตัวแบบดั้งเดิม ได้แก่ ตัวแบบอาร์มาและซัพพอร์ตเวกเตอร์แมชชีนโดยปราศจากการหาค่าพารามิเตอร์ที่เหมาะสมที่สุดด้วยเกณฑ์ประเมิน 6 เกณฑ์ สำหรับการพยากรณ์ปริมาณความต้องการพลังงานไฟฟ้าสูงสุด และสามารถเป็นข้อมูลที่เป็นประโยชน์ต่อการวางแผนบริหารจัดการได้ ซึ่งตัวแบบพยากรณ์ต่างๆ ถูกพัฒนาจากการเขียนชุดคำสั่งของภาษา R (R Programming) ซึ่งมีไลบรารีต่างๆ ที่ได้ถูกพัฒนาบางส่วนเพื่อช่วยในการประยุกต์ใช้งานได้สะดวกและรวดเร็วมากยิ่งขึ้น

2. ข้อมูลและวิธีการวิจัย

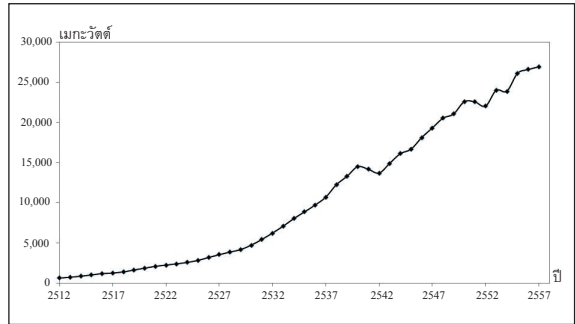
2.1 ข้อมูล

ข้อมูลที่ใช้ในการวิจัยนี้ เป็นข้อมูลปริมาณความต้องการพลังงานไฟฟ้าสูงสุดรายปีของประเทศไทย ซึ่งสามารถเข้าไปดาวน์โหลดได้ที่เว็บไซต์ของการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย [14] โดยข้อมูลรายปีตั้งแต่ปี พ.ศ. 2512 จนถึงปี พ.ศ. 2557 ที่ถูกแสดงในรูปที่ 1

2.2 วิธีการวิจัย

2.2.1 ตัวแบบอาร์มา

ตัวแบบอาร์มาเป็นตัวแบบทางสถิติตัวแบบหนึ่งที่มีความโดดเด่นในการพยากรณ์ข้อมูลที่มีลักษณะเชิงเส้นตรง ซึ่งเป็นรูปแบบทั่วไปของตัวแบบอามา (ARMA) สำหรับข้อมูลที่ไม่มีอยู่ในสภาวะคงที่ (Non-stationary State)



รูปที่ 1 ปริมาณความต้องการพลังงานไฟฟ้าสูงสุดรายปีสมการทางคณิตศาสตร์ของตัวแบบอาร์มา ดังที่ถูกระบุแสดงในสมการที่ (1)

$$\left(1 - \sum_{i=1}^p \phi_i B^i\right) (1 - B)^d (y_t - \mu) = \left(1 - \sum_{j=1}^q \theta_j B^j\right) \varepsilon_t \quad (1)$$

เมื่อ y_t และ ε_t คือ ค่าจริงและค่าความคลาดเคลื่อน ณ เวลาที่ t ตามลำดับ B และ μ คือ ตัวแปลงค่าอนุกรมเวลาย้อนกลับ (Backward Shift Operator) และค่าเฉลี่ย ตามลำดับตัวแปร p และ q คือ ลำดับของอาร์มา โดยตัวแบบอาร์มาทั่วไป คือ ARIMA(p, d, q)

ขั้นตอนการพยากรณ์ด้วยตัวแบบอาร์มามี 3 ขั้นตอน ดังนี้

1. ขั้นตอนการระบุตัวแบบ ซึ่งในขั้นตอนนี้ชุดข้อมูลอนุกรมเวลาจะถูกพิจารณาในสภาวะคงที่ (Stationary State) ของชุดข้อมูลอนุกรมเวลา ถ้าชุดข้อมูลไม่อยู่ในสภาวะคงที่ จะต้องใช้การแปลงค่าอนุกรมเวลาของตัวแปลงค่าอนุกรมเวลาแบบย้อนกลับ เพื่อปรับให้ชุดข้อมูลอนุกรมเวลามีสภาวะคงที่ ซึ่งสภาวะคงที่นี้มีความสำคัญในการสร้างตัวแบบอาร์มา โดยลำดับของตัวแบบอาร์มาสามารถพิจารณาได้จากค่าฟังก์ชันความสัมพันธ์กันเอง (Autocorrelation Function: ACF) และค่าฟังก์ชันความสัมพันธ์กันเองบางส่วน (Partial Autocorrelation Function: PACF) เพื่อเป็นจุดเริ่มต้นในการกำหนดลำดับของตัวแบบอาร์มา แต่ถ้าลักษณะของค่าฟังก์ชันความสัมพันธ์กันเองและค่าฟังก์ชันความสัมพันธ์กันบางส่วนลู่เข้าสู่ค่าศูนย์

อย่างรวดเร็วทั้งคู่ การระบุลำดับจะมีความซับซ้อน จึงต้องอาศัยเกณฑ์การพิจารณาอื่น เช่น เกณฑ์อะไคอินโฟเมชัน (Akaike Information Criterion: AIC) เป็นต้น เพื่อกำหนดลำดับของตัวแบบอาร์มา

2. ขั้นตอนการประมาณค่าพารามิเตอร์ของตัวแบบอาร์มา ซึ่งขั้นตอนนี้จะประมาณค่าพารามิเตอร์จากการเลือกค่าพารามิเตอร์ที่ให้ค่าความคลาดเคลื่อนน้อยที่สุด

3. ขั้นตอนการทดสอบตัวแบบ ซึ่งในขั้นตอนนี้ตัวแบบที่ได้จะถูกทดสอบตามข้อกำหนดต่างๆ หากตัวแบบที่ได้ยังไม่เหมาะสมจะต้องดำเนินการในข้อที่ 2 และ 3 ใหม่

ในการหาตัวแบบอาร์มาที่เหมาะสมของงานวิจัยนี้ใช้เกณฑ์อะไคอินโฟเมชันสำหรับขนาดตัวอย่างจำกัด (AICc) ที่มีค่าน้อยที่สุด ซึ่งถูกแสดงดังสมการที่ (2)

$$AICc = AIC + \frac{2k(k+1)}{n-k-1} \quad (2)$$

$$AIC = 2k - 2 \ln(L) \quad (3)$$

เมื่อ n และ k คือ ขนาดตัวอย่างและจำนวนพารามิเตอร์ตามลำดับ

ในการพยากรณ์ด้วยตัวแบบอาร์มาในงานวิจัยนี้ตัวแบบอาร์มาที่เหมาะสมโดยใช้เกณฑ์อะไคอินโฟเมชันสำหรับขนาดตัวอย่างจำกัดที่มีค่าน้อยที่สุดนั้น ลำดับของตัวแบบอาร์มาจะเปลี่ยนแปลงไปตามการปรับเปลี่ยนชุดข้อมูลการสร้างตัวแบบที่เพิ่มเติมค่าที่ถูกเฉลี่ยแล้วเข้ามารวมกับชุดข้อมูลที่มีอยู่เดิม ดังนั้นจึงใช้ตัวแบบอาร์มาที่ไม่ได้ระบุลำดับของตัวแบบอาร์มาแทนตัวแบบอาร์มาที่ระบุลำดับ (ARIMA(p, d, q)) ดังนั้นจึงใช้สัญลักษณ์ ARIMA ในส่วนของผลการวิจัย เพื่อให้การนำเสนออยู่ในรูปแบบที่กระชับ

2.2.2 ซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีน คือ ตัวแบบการเรียนรู้ของเครื่องตัวแบบหนึ่งที่ถูกพัฒนาขึ้นโดย Vapnik ซึ่งอาศัยหลักการเรียนรู้ทางสถิติและการลดความเสี่ยงเชิง

โครงสร้างที่น้อยที่สุดและถูกใช้ในการวิจัยหลายลักษณะ เช่น การจำแนกกลุ่มและการพยากรณ์ เป็นต้น สมการทางคณิตศาสตร์ของซัพพอร์ตเวกเตอร์แมชชีนสำหรับการพยากรณ์เชิงเส้นตรงดังสมการที่ (4)

$$f(x) = (w, x) + b \quad x \in \mathbb{R}^d, \quad b \in \mathbb{R} \quad (4)$$

โดย $f(x)$ คือ ค่าตัวแปรตามของชุดข้อมูลป้อนเข้า (Input Data) นอกจากนี้ w และ b คือ เวกเตอร์ถ่วงน้ำหนัก (Weight Vector) และค่าเอนเอียง (Bias) ตามลำดับ และ (\cdot, \cdot) คือ การคูณแบบดอท (Dot Product) และ x คือ เวกเตอร์ของชุดข้อมูลป้อนเข้า ซึ่งมีขนาดเท่ากับ d

ความเรียบของสมการที่ (4) สามารถทำให้เรียบที่สุดด้วยวิธีการหาค่าที่เหมาะสมที่สุด ดังสมการที่ (5)

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^T (\xi_i + \xi_i^*) \quad (5)$$

สมการข้อบ่งชี้คือ $y_i - (w \cdot x_i) - b \leq \varepsilon + \xi_i$ และ $(w \cdot x_i) + b - y_i \leq \varepsilon + \xi_i^*$ เมื่อ $\xi_i, \xi_i^*, C \geq 0$ โดยค่า C ซึ่งถูกใช้ในการพิจารณาความเรียบและค่าของการเบี่ยงเบนที่ออกนอกขอบเขตที่กำหนด และตัวแปร ξ_i และ ξ_i^* คือ ระยะห่างของจุดที่อยู่นอกขอบเขตที่กำหนด ซึ่งถูกใช้เพื่อให้ครอบคลุมกรณีของค่าสังเกตที่อยู่นอกขอบเขตที่กำหนด ซึ่งการเลือกค่า C ที่เหมาะสมจะส่งผลให้ประสิทธิภาพการพยากรณ์ของซัพพอร์ตเวกเตอร์แมชชีนดีที่สุด จากสมการที่ (5) การหาค่าที่เหมาะสมมีความซับซ้อน การหาค่าด้วยวิธีการของลากรางจ์ (Lagrange Multipliers) ถูกนำมาใช้เพื่อหาค่าที่เหมาะสมในรูปแบบคู่ (Dual Formulation) ซึ่งซัพพอร์ตเวกเตอร์แมชชีนที่มีฟังก์ชันการทำนายเชิงเส้นตรง ดังแสดงในรูปแบบของสมการที่ (6)

$$f(x) = \sum_{i=1}^T (\alpha_i - \alpha_i^*) (x, x_i) + b \quad (6)$$

โดย α_i และ α_i^* คือ ค่าของลากรางจ์ และ (\cdot, \cdot) คือ การคูณแบบดอท (Dot Product) และ b คือ ค่าเอนเอียง และ

ตัวแปร x_i คือ ข้อมูลป้อนเข้าลำดับที่ i

แต่ปัญหาการพยากรณ์ส่วนมากมักพบลักษณะแบบไม่เชิงเส้นตรง การแปลงข้อมูลจากข้อมูลแบบไม่เชิงเส้นตรง ไปสู่ข้อมูลที่เป็นเชิงเส้นตรงด้วยการใช้ฟังก์ชันเคอร์เนล (Kernel Function) โดยฟังก์ชันเคอร์เนลสำหรับข้อมูลที่ไม่เป็นเชิงเส้นตรง คือ เรเดียลเบสิคฟังก์ชัน (Radial Basis Function) ซึ่งเป็นฟังก์ชันที่มีประสิทธิภาพและนิยมใช้ในการพยากรณ์ข้อมูลที่ไม่เป็นเชิงเส้นตรง [15], [16] ในการพยากรณ์ด้วยซัพพอร์ตเวกเตอร์แมชชีนในงานวิจัยนี้ใช้ฟังก์ชันเรเดียลเบสิคสำหรับการพยากรณ์ข้อมูลไม่เชิงเส้นตรงและฟังก์ชันเชิงเส้นตรงสำหรับข้อมูลเชิงเส้นตรง โดยมีสมการที่ (7) และ (8)

$$\text{ฟังก์ชันเชิงเส้นตรง } K(x, x_i) = x^T x_i \quad (7)$$

$$\text{ฟังก์ชันเรเดียลเบสิค } K(x, x_i) = \exp(-\gamma \|x - x_i\|^2) \quad (8)$$

โดยค่า γ คือ ค่าพารามิเตอร์ของฟังก์ชันเรเดียลเบสิค

ในงานวิจัยนี้ ข้อมูลอนุกรมเวลาจะถูกจัดเรียงใหม่ m คอลัมน์ โดยใช้ข้อมูลย้อนหลัง $m - 1$ ค่า เพื่อสร้างตัวแบบพยากรณ์ค่าที่ m ดังแสดงในรูปที่ 2

ดังนั้นจึงใช้สัญลักษณ์ SVM(m) ในส่วนของผลการวิจัย เพื่อให้การนำเสนออยู่ในรูปแบบที่กระชับ

2.2.3 ขั้นตอนวิธีเชิงพันธุกรรม

ขั้นตอนวิธีเชิงพันธุกรรมเป็นหนึ่งในวิธีการค้นหาคำตอบของวิธีการเมตาฮิวริสติก ซึ่งอาศัยหลักการจากทฤษฎีวิวัฒนาการจากชีววิทยา โดยมีขั้นตอนการดำเนินการดังนี้

1. กำหนดกลุ่มประชากรเริ่มต้น
2. การคัดเลือกพันธุกรรมที่ดีที่สุดเพื่อเป็นต้นแบบในการสร้างประชากรรุ่นต่อไป
3. สร้างประชากรรุ่นใหม่ ด้วยวิธีการไขว้เบลีี่ยน (Crossover) และการกลายพันธ์ (Mutation)

2.2.4 ตัวแบบผสมซัพพอร์ตเวกเตอร์แมชชีนและขั้นตอนวิธีเชิงพันธุกรรม

y_1	y_2	y_3	y_4	y_5	...	y_{t-3}	y_{t-2}	y_{t-1}	y_t
-------	-------	-------	-------	-------	-----	-----------	-----------	-----------	-------

จัดเรียงชุดข้อมูลให้อยู่ในรูปแบบ m คอลัมน์

y_1	y_2	y_3	...	y_m
y_2	...	y_{m-1}	y_m	y_{m+1}
y_3	...	y_m	y_{m+1}	y_{m+2}
\vdots	\vdots	\vdots	\vdots	\vdots
y_{t-m+1}	...	y_{t-2}	y_{t-1}	y_t

รูปที่ 2 การจัดรูปแบบข้อมูลสำหรับการพยากรณ์โดยซัพพอร์ตเวกเตอร์แมชชีน

วัตถุประสงค์ของการพัฒนาตัวแบบผสมนี้ คือ การนำจุดเด่นของแต่ละตัวแบบเพื่อช่วยหรือลดข้อด้อยของตัวแบบอื่น ตัวแบบผสมนี้จึงเพิ่มโอกาสในการพยากรณ์ค่าในสถานการณ์ต่างๆ และเพิ่มความถูกต้องให้มากยิ่งขึ้น โดยขั้นตอนการดำเนินการของตัวแบบผสมดังต่อไปนี้

1. กำหนดค่าพารามิเตอร์เริ่มต้นแบบสุ่มจำนวนหนึ่งของเคอร์เนลฟังก์ชันของตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

เช่น การสร้างค่าประชากรขนาด 10 ค่า โดยมีขอบเขตของประชากรมีค่าระหว่าง 0 ถึง 20 เพื่อใช้เป็นค่าพารามิเตอร์ของเคอร์เนลฟังก์ชัน ดังนี้

1.807448, 16.813212, 5.502158, 19.307602, 4.89785, 12.095677, 10.106531, 16.157176, 11.873357, 2.624342

2. แบ่งข้อมูลอนุกรมเวลาออกเป็นสองส่วน คือ 60% ของข้อมูลทั้งหมดเพื่อการสร้างตัวแบบและ 40% ของข้อมูลทั้งหมดถูกใช้ในการทดสอบความถูกต้องของการพยากรณ์

3. ข้อมูล 60% แรกถูกนำมาสร้างตัวแบบพยากรณ์ของซัพพอร์ตเวกเตอร์แมชชีนเพื่อพยากรณ์ค่าที่ถัดไปหนึ่งค่าของชุดข้อมูลทดสอบ เมื่อค่าถัดไปค่าหนึ่งนั้น

ถูกเฉลี่ย ค่าดังกล่าวจะถูกรวมเข้ากับชุดข้อมูลการสร้างตัวแบบเพื่อสร้างตัวแบบการพยากรณ์สำหรับการทำนายค่าที่ถัดไปหนึ่งค่าจนกระทั่งพยากรณ์ค่าสุดท้ายของชุดข้อมูลทดสอบ

4. คำนวณค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error: MAPE)

5. เลือกค่าตัวแปรที่ดีที่สุดเพื่อสร้างค่าพารามิเตอร์รุ่นใหม่ โดยนำค่าประชากรที่ได้จากข้อที่ 1 มาเป็นต้นแบบเพื่อสร้างประชากรรุ่นใหม่ โดยแปลงค่าให้อยู่ในรูป Binary Code เช่น ค่า 1.807448 สามารถแปลงค่าได้เป็น 1 1 0 1 1 1 0 0 1 0 1 0 0 0 1 0 1 1 0 0 0 เพื่อสร้างประชากรรุ่นใหม่ด้วยการไขว้เปลี่ยนและการกลายพันธุ์ตั้งข้อที่ 6

6. สร้างพารามิเตอร์รุ่นใหม่โดยการไขว้เปลี่ยนที่ความน่าจะเป็น 0.8 และการกลายพันธุ์ที่ความน่าจะเป็น 0.1 เพื่อใช้เป็นพารามิเตอร์ของตัวแบบพยากรณ์ของซัพพอร์ตเวกเตอร์แมชชีน เช่น ประชากรรุ่นต่อมาจากการไขว้เปลี่ยนและการกลายพันธุ์ ดังนี้

12.095677, 12.000552, 1.807448, 10.106531, 19.307602, 16.629044, 10.106531, 11.968482, 16.813212, 16.341343

เพื่อใช้เป็นค่าพารามิเตอร์ของเคอเนลฟังก์ชันของตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

7. ดำเนินการข้อ 2 ถึง 6 จนกระทั่งถึงจุดกำหนดให้สิ้นสุด

8. แสดงค่าพารามิเตอร์ที่เหมาะสมที่สุด ซึ่งให้ค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์ที่น้อยที่สุด

ในงานวิจัยนี้ สัญลักษณ์ SVM(m)-GA แทนตัวแบบผสมระหว่างซัพพอร์ตเวกเตอร์แมชชีนและขั้นตอนวิธีเชิงพันธุกรรมในส่วนของผลการวิจัย เพื่อให้การนำเสนออยู่ในรูปแบบที่กระชับ

2.2.5 การประเมินประสิทธิภาพของตัวแบบการพยากรณ์

เพื่อประเมินประสิทธิภาพของตัวแบบการพยากรณ์ต่างๆ ชุดข้อมูลอนุกรมเวลาทั้งหมด 46 ค่าสังเกตจะถูกแบ่งออกเป็นสองชุด คือ ชุดข้อมูลที่ใช้ในการสร้าง

ตัวแบบและชุดข้อมูลสำหรับการทดสอบตัวแบบ โดยชุดข้อมูลที่ใช้ในการสร้างตัวแบบมีขนาดเท่ากับ 60% ของชุดอนุกรมเวลาทั้งหมด (ข้อมูลค่าที่ 1 ถึงค่าที่ 27) และชุดข้อมูลสำหรับการทดสอบมีขนาดเท่ากับ 40% ของชุดอนุกรมเวลาทั้งหมด (ค่าที่ 28 ถึงค่าที่ 46 ซึ่งเท่ากับ 19 - fold cross - validation) โดยข้อมูลในชุดทดสอบจะถูกปิดไว้ เพื่อทดสอบประสิทธิภาพการทำนายของตัวแบบพยากรณ์ในการทำนายค่าที่จะเกิดขึ้นในอนาคตถัดไปหนึ่งค่า และเมื่อค่าถัดไปในอนาคตดังกล่าวถูกเฉลี่ยค่าดังกล่าวจะถูกนำมารวมกับชุดข้อมูลในการสร้างตัวแบบเพื่อทำนายค่าถัดไปที่จะเกิดขึ้นในอนาคต ซึ่งการพยากรณ์นี้จะถูกดำเนินการไปเรื่อยๆ จนกระทั่งถึงค่าสุดท้ายของชุดข้อมูลทดสอบ โดยการประเมินประสิทธิภาพของตัวแบบพยากรณ์แต่ละตัวแบบใช้วิธีการประเมิน 6 วิธี มีสมการดังนี้

1. ค่าเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Error: MAE) ดังสมการที่ (9)

$$MAE = \text{mean}(|e_i|) \quad (9)$$

2. ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error: RMSE) ดังสมการที่ (10)

$$RMSE = \sqrt{\text{mean}(e_i^2)} \quad (10)$$

3. ค่าเฉลี่ยของเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error: MAPE) ดังสมการที่ (11)

$$MAPE = \text{mean}(|p_i|) \quad (11)$$

โดย $p_i = \frac{(y_i - \hat{y}_i)}{y_i} \times 100$ นอกจากนี้ y_i และ \hat{y}_i คือ ค่าจริงและค่าที่พยากรณ์ ตามลำดับ

4. ค่ามัธยฐานของเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์ (Median Absolute Percentage Error: MdAPE) ดังสมการที่ (12)

$$\text{MdAPE} = \text{median}(|p_i|) \quad (12)$$

โดย $p_i = \frac{(y_i - \hat{y}_i)}{y_i} \times 100$ นอกจากนี้ y_i และ \hat{y}_i คือ ค่าจริงและค่าที่พยากรณ์ ตามลำดับ

5. ค่าเฉลี่ยของรากที่สองของความคลาดเคลื่อนกำลังสอง (Root Mean Square Percentage Error: RMSPE) ดังสมการที่ (13)

$$\text{RMSPE} = \sqrt{\text{mean}(p_i^2)} \quad (13)$$

โดย $p_i = \frac{(y_i - \hat{y}_i)}{y_i} \times 100$ นอกจากนี้ y_i และ \hat{y}_i คือ ค่าจริงและค่าที่พยากรณ์ ตามลำดับ

6. ค่าสัมประสิทธิ์การตัดสินใจ (R-squared) ดังสมการที่ (14)

$$r^2 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (14)$$

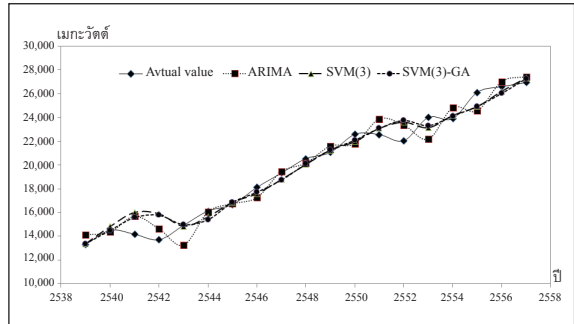
โดย x_i และ y_i คือ ค่าใดๆ ของตัวแปรแต่ละตัว นอกจากนี้ \bar{x} และ \bar{y} คือ เฉลี่ยของแต่ละตัวแปร

ตัวแบบที่เหมาะสมที่สุด คือ ตัวแบบที่ให้ค่าความคลาดเคลื่อนและความแปรปรวนน้อยที่สุด แต่ต้องให้ค่าสัมประสิทธิ์การตัดสินใจมากที่สุด ซึ่งตัวแบบที่ให้ค่าความคลาดเคลื่อนน้อยที่สุดแสดงว่า ตัวแบบดังกล่าวสามารถให้ค่าความถูกต้องของการพยากรณ์มากที่สุด และการที่มีค่าสัมประสิทธิ์การตัดสินใจที่มีค่าสูงจะแสดงให้เห็นว่าตัวแบบนั้นสามารถอธิบายตัวแปรทำนายได้ดี

3. ผลการวิจัย

จากผลการพยากรณ์ด้วยตัวแบบพยากรณ์ต่างๆ ที่ได้นำเสนอในงานวิจัยนี้ เพื่อพยากรณ์ปริมาณความต้องการพลังงานไฟฟ้าสูงสุดถูกนำเสนอในรูปแบบที่ 3

การพยากรณ์ด้วยตัวแบบ ARIMA ที่ลำดับต่างๆ ถูกนำมาใช้สร้างตัวแบบพยากรณ์ ซึ่งตัวแบบที่ให้ค่า AICc ที่น้อยที่สุดจะเป็นตัวแบบที่เหมาะสมที่สุดในการพยากรณ์ค่าที่จะเกิดขึ้นในอนาคต ดังแสดงในรูปแบบที่ 4



รูปที่ 3 ค่าพยากรณ์ของแต่ละตัวแบบและค่าจริงของปริมาณความต้องการไฟฟ้าสูงสุด

ARIMA (2, 2, 2)	: 1e+20
ARIMA (0, 2, 0)	: 322.3362
ARIMA (1, 2, 0)	: 321.3614
ARIMA (0, 2, 1)	: 322.1375
ARIMA (2, 2, 0)	: 323.7797
ARIMA (1, 2, 1)	: 323.8669
ARIMA (2, 2, 1)	: 326.5718
ARIMA (1, 2, 0)	: 321.3614
Best model: ARIMA (1, 2, 0)	
Series: data	
ARIMA (1, 2, 0)	
Coefficients:	
arl	
0.5428	
s.e.	0.2747
sigma^2 estimated as 18413: log likelihood=-158.41	
AIC=320.82	AICc=321.36
	BIC=323.25

รูปที่ 4 การประมาณค่าพารามิเตอร์และการหาตัวแบบ ARIMA ที่เหมาะสมที่สุด

จากตัวอย่างที่แสดงในรูปแบบที่ 4 เป็นการหาตัวแบบ ARIMA ที่เหมาะสมที่สุดสำหรับชุดข้อมูล 27 ค่าแรก เพื่อใช้พยากรณ์ค่าที่ 28 นั้นตัวแบบ ARIMA หลายๆ ตัวแบบถูกสร้างและคำนวณค่า AICc ซึ่งตัวแบบที่เหมาะสมที่สุดคือ ARIMA(1,2,0) ที่มีค่า AICc เท่ากับ 321.3614 ซึ่งเป็นค่าที่น้อยที่สุดเมื่อเทียบกับตัวแบบอื่น โดยมีค่าสัมประสิทธิ์และค่าวิเคราะห์ต่างๆ ของตัวแบบ ARIMA ดังที่แสดงในรูปแบบ



```

ARIMA (2, 2, 2)           : 1e+20
ARIMA (0, 2, 0)          : 357.5619
ARIMA (1, 2, 0)          : 357.1477
ARIMA (0, 2, 1)          : 357.1463
ARIMA (1, 2, 1)          : 359.6159
ARIMA (0, 2, 2)          : 359.6076
ARIMA (1, 2, 2)          : 358.7398
ARIMA (0, 2, 1)          : 357.1463

Best model: ARIMA (0, 2, 1)

Series: data
ARIMA (0, 2, 1)

Coefficients:
      mal
      -0.3043
s.e.   0.1637

sigma2 estimated as 27433: log likelihood=-176.32
AIC=356.65 AICc=357.15 BIC=359.24

```

รูปที่ 5 ตัวแบบ ARIMA ที่เหมาะสมหลังจากมีการปรับชุดข้อมูลสำหรับการสร้างตัวแบบ

แต่การปรับเปลี่ยนชุดข้อมูลสำหรับการสร้างตัวแบบ ARIMA เมื่อค่าในชุดข้อมูลทดสอบที่เฉลี่ยแล้วจะถูกรวมเข้ากับชุดข้อมูลสำหรับการสร้างตัวแบบเดิม เพื่อเป็นชุดข้อมูลสำหรับการสร้างตัวแบบพยากรณ์ใหม่ ดังนั้น ตัวแบบ ARIMA ที่ดีที่สุดจึงมีการเปลี่ยนแปลง ดังแสดงในรูปที่ 5

ตัวแบบ ARIMA ที่เหมาะสำหรับชุดข้อมูล 29 ค่าแรก ซึ่งเกิดจากการรวมค่าที่ 29 ซึ่งถูกเฉลี่ยในขั้นตอนการประเมินประสิทธิภาพของตัวแบบ ซึ่งค่าดังกล่าวถูกรวมเข้ากับชุดข้อมูล 28 ค่าแรก พบว่าตัวแบบ ARIMA(1,2,0) ที่เดิมเหมาะสมที่สุดสำหรับชุดข้อมูลในการสร้างตัวแบบ 27 ค่าแรกนั้นไม่เหมาะสำหรับชุดข้อมูลสำหรับการสร้างตัวแบบใหม่ ดังแสดงในรูปที่ 5 จึงใช้ตัวแบบ ARIMA โดยไม่ระบุลำดับแทนตัวแบบ ARIMA(p,d,q) ที่มีการระบุลำดับ

สำหรับการพยากรณ์ด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนนั้น ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนต่างๆ ถูกสร้าง

จากการจัดรูปแบบของชุดข้อมูลสำหรับสร้างตัวแบบที่มีคอลัมน์ตั้งแต่ 2 ถึง 10 คอลัมน์ของข้อมูลย้อนหลัง โดยใช้ฟังก์ชันเคอร์เนลทั้งเชิงเส้นตรงและเรเดียลเบสิค โดยตัวแบบที่เหมาะสมที่สุดคือ ตัวแบบ SVM(3) ที่ใช้ฟังก์ชันเคอร์เนลแบบเชิงเส้นตรง ซึ่งให้ค่าน้อยที่สุดในทุกเกณฑ์ประเมินเมื่อเทียบกับตัวแบบ SVM(m) โดย m มีค่าตั้งแต่ 2 ถึง 10

ตัวแบบการพยากรณ์แบบผสมระหว่างซัพพอร์ตเวกเตอร์แมชชีนและขั้นตอนวิธีเชิงพันธุกรรมที่เหมาะสมในงานวิจัยนี้ ถูกสร้างจากการนำตัวแบบซัพพอร์ตเวกเตอร์แมชชีนเพื่อเรียนรู้และสร้างตัวแบบพยากรณ์จากการจัดรูปแบบของชุดข้อมูลสำหรับสร้างตัวแบบที่มีคอลัมน์ตั้งแต่ 2 ถึง 10 คอลัมน์ โดยใช้ฟังก์ชันเคอร์เนลทั้งเชิงเส้นตรงและเรเดียลเบสิค ซึ่งนำขั้นตอนวิธีเชิงพันธุกรรมในการค้นหาค่าพารามิเตอร์ที่เหมาะสมที่สุดของซัพพอร์ตเวกเตอร์แมชชีน โดยตัวแบบที่เหมาะสมที่สุดคือ ตัวแบบ SVM(3)-GA ที่ใช้ฟังก์ชันเคอร์เนลแบบเชิงเส้นตรง ซึ่งให้ค่าน้อยที่สุดในทุกเกณฑ์ประเมิน โดยค่าพารามิเตอร์ที่เหมาะสมของซัพพอร์ตเวกเตอร์แมชชีน คือ C เท่ากับ 0.187529

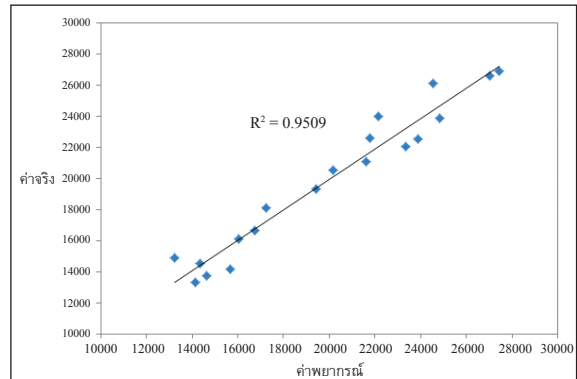
ผลสรุปของการประเมินประสิทธิภาพของตัวแบบพยากรณ์ต่างๆ ดังแสดงในตารางที่ 1 และ 2

ตารางที่ 1 ค่าการประเมินประสิทธิภาพของตัวแบบพยากรณ์ต่างๆ

เกณฑ์การประเมิน	ตัวแบบพยากรณ์		
	ARIMA	SVM(3)	SVM(3)-GA
MAE	831.73	667.82	626.91
RMSE	1000.98	889.03	838.18
MAPE	4.36%	3.65%	3.37%
MdAPE	3.90%	2.39%	2.21%
RMSPE	5.40%	5.44%	4.98%
R-Squared	0.951	0.963	0.967

ตารางที่ 2 ค่าพยากรณ์ของตัวแบบพยากรณ์ต่างๆ

ค่าจริง	ตัวแบบพยากรณ์		
	ARIMA	SVM(3)	SVM(3)-GA
13,310.90	14,141.97	13,373.97	13,373.97
14,506.30	14,353.90	14,820.65	14,497.80
14,179.90	15,685.70	16,012.36	15,578.28
13,712.40	14,628.36	15,899.36	15,775.09
14,918.30	13,244.90	14,857.91	14,986.18
16,126.40	16,055.32	15,648.06	15,396.13
16,681.10	16,749.43	16,867.44	16,858.17
18,121.40	17,258.35	17,594.65	17,711.75
19,325.80	19,438.31	18,781.61	18,730.77
20,537.50	20,164.52	20,045.94	20,087.12
21,064.00	21,604.48	21,252.50	21,274.34
22,586.10	21,784.17	21,970.95	22,087.01
22,568.20	23,885.01	23,083.19	23,092.35
22,044.90	23,363.06	23,607.10	23,750.03
24,009.90	22,173.98	23,143.72	23,318.79
23,900.21	24,833.29	24,176.04	24,127.61
26,121.10	24,563.71	24,921.09	24,940.32
26,598.14	27,032.34	26,212.55	26,027.36
26,942.10	27,440.39	27,337.80	27,281.66



รูปที่ 6 ค่า R-Squared ของตัวแบบ ARIMA

```
Call:
lm(formula = Y ~ X, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1528.41  -830.13   -70.08    654.74   1940.55

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.302e+02  1.103e+03   0.299   0.768
X           9.804e-01  5.405e-02  18.140 1.47e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1052 on 17 degrees of freedom
Multiple R-squared:  0.9509, Adjusted R-squared:  0.948
F-statistic: 329 on 1 and 17 DF, p-value: 1.467e-12

> sum(fit$residuals^2)/(19-2)
[1] 1107192
```

โดยกราฟแสดงผลของค่า R-Squared ของตัวแบบพยากรณ์ต่างๆ ถูกนำเสนอในรูปแบบที่ 6, 8, และ 10

จากกราฟที่ถูกแสดงในรูปแบบที่ 6 ลักษณะการกระจายตัวของค่าพยากรณ์ออกจากค่าจริงซึ่งอยู่บนเส้นตรงค่อนข้างมาก ซึ่งค่าความแปรปรวนของตัวแบบเชิงเส้นตรงของชุดข้อมูลที่อยู่ในรูปที่ 6 ถูกแสดงในรูปแบบของสมการของสมการถดถอยเชิงเส้นตรง ดังรูปที่ 7

จากผลการวิเคราะห์เชิงเส้นที่ถูกแสดงในรูปแบบที่ 7 ค่าพยากรณ์และค่าจริงมีความสัมพันธ์กันร้อยละ 95.1 และมีความแปรปรวนของตัวแบบ (Mean Square Error: MSE) เท่ากับ 1,107,192 ซึ่งค่านี้จะแสดงถึงการกระจายตัวของค่าพยากรณ์ที่เบี่ยงเบนจากค่าจริง

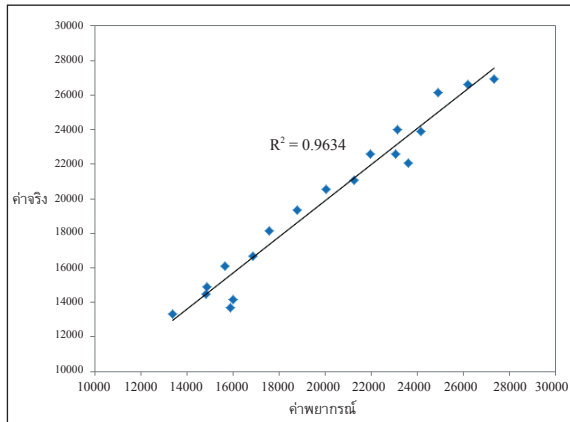
จากลักษณะความสัมพันธ์ของค่าพยากรณ์และค่าจริงที่ถูกแสดงในรูปแบบที่ 8 มีลักษณะเชิงเส้นตรง การวิเคราะห์

รูปที่ 7 การวิเคราะห์ความสัมพันธ์เชิงเส้นตรงของค่าพยากรณ์และค่าจริงของตัวแบบ ARIMA

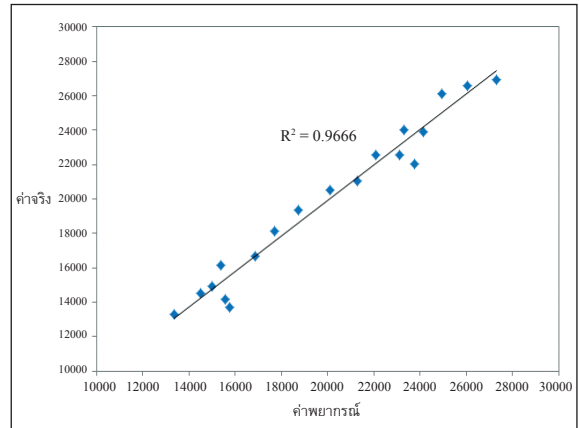
สมการถดถอยจึงถูกนำมาร่วมพิจารณา เพื่อคำนวณค่าความแปรปรวนเชิงเส้นตรงของความสัมพันธ์ โดยความสัมพันธ์ของค่าพยากรณ์และค่าจริงมีความสัมพันธ์กันร้อยละ 96.34 และมีความแปรปรวนของการกระจายตัวของค่าพยากรณ์ที่เบี่ยงเบนออกจากค่าจริงเท่ากับ 825,590 ซึ่งถูกแสดงในรูปแบบที่ 9

สำหรับการวิเคราะห์ความสัมพันธ์เชิงเส้นตรงของค่าพยากรณ์และค่าจริงของรูปที่ 10 ถูกแสดงในรูปแบบที่ 11

จากผลลัพธ์ที่ได้แสดงให้เห็นว่าค่าพยากรณ์ที่ได้จากตัวแบบผสมมีค่าใกล้เคียงกับค่าจริงมากกว่าของตัวแบบพยากรณ์สองตัวแบบก่อนหน้านี้ ซึ่งส่งผลให้



รูปที่ 8 ค่า R-Squared ของตัวแบบ SVM(3)



รูปที่ 10 ค่า R-Squared ของตัวแบบ SVM(3)-GA

```
Call:
lm(formula = Y ~ X, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1878.7  -436.7   227.5   685.5  1100.3

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.027e+03  1.009e+03  -1.018   0.323
X             1.045e+00  4.943e-02  21.144  1.2e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 908.6 on 17 degrees of freedom
Multiple R-squared:  0.9634, Adjusted R-squared:  0.9612
F-statistic: 447.1 on 1 and 17 DF, p-value: 1.203e-13

> sum(fit$residuals^2)/(19-2)
[1] 825590.4
```

รูปที่ 9 การวิเคราะห์ความสัมพันธ์เชิงเส้นตรงของค่าพยากรณ์และค่าจริงของตัวแบบ SVM(3)

```
Call:
lm(formula = Y ~ X, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1832.2  -396.5   250.5   552.4  1094.2

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -776.20494  951.07058  -0.816   0.426
X             1.03459    0.04664  22.184  5.47e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 867.5 on 17 degrees of freedom
Multiple R-squared:  0.9666, Adjusted R-squared:  0.9646
F-statistic: 492.1 on 1 and 17 DF, p-value: 5.465e-14

> sum(fit$residualsa^2)/(19-2)
[1] 752513.8
```

รูปที่ 11 การวิเคราะห์ความสัมพันธ์เชิงเส้นตรงของค่าพยากรณ์และค่าจริงของตัวแบบ SVM(3)-GA

ความสัมพันธ์ระหว่างค่าทั้งสองมีค่าร้อยละ 96.66 และมีความแปรปรวนเท่ากับ 752,514 ซึ่งน้อยที่สุดเมื่อเทียบกับสองตัวแบบก่อนหน้านี้

4. อภิปรายผลและสรุป

จากการประเมินประสิทธิภาพของตัวแบบพยากรณ์ต่างๆ โดยอาศัยเกณฑ์การประเมินทั้ง 6 เกณฑ์ พบว่าตัวแบบผสม SVM(3)-GA ที่นำเสนอให้ผลความคลาดเคลื่อนน้อยที่สุดเมื่อเปรียบเทียบกับตัวแบบ ARIMA และ SVM(3) ซึ่งสนับสนุนการใช้ตัวแบบผสมเพื่อลดความเสี่ยงจากการใช้

ตัวแบบพยากรณ์ที่ไม่เหมาะสมในการพยากรณ์ปริมาณความต้องการพลังงานไฟฟ้าที่สูงสุดรายปีและให้ค่าความถูกต้องมากขึ้น โดยขั้นตอนเชิงพันธุกรรมสามารถถูกใช้ประโยชน์ในการค้นหาค่าพารามิเตอร์ที่เหมาะสมสำหรับตัวแบบการพยากรณ์ซัพพอร์ตเวกเตอร์แมชชีน ซึ่งการใช้พารามิเตอร์ของตัวแบบ SVM(3) และ SVM(3)-GA มีจำนวนพารามิเตอร์เท่ากัน คือ ค่า C เท่านั้น เนื่องจากฟังก์ชันที่เหมาะสม คือ ฟังก์ชันเชิงเส้นตรง แต่การใช้การปรับค่าพารามิเตอร์ด้วยการเพิ่มหรือลดค่าพารามิเตอร์นั้น จะทำให้เสียเวลาในการค้นหานั้นและอาจไม่สามารถการันตี

ค่าที่เหมาะสมที่สุดในขอบเขตการค้นหาได้ในทางกลับกัน วิธีการเชิงพันธุกรรมสามารถหาค่าพารามิเตอร์ที่เหมาะสม และสามารถปรับค่าที่เหมาะสมที่สุดภายในขอบเขตค้นหาที่กำหนดได้ โดยตัวแบบ SVM(3) มีพารามิเตอร์ C เท่ากับ 1 แต่ตัวแบบ SVM(3)-GA มีค่าพารามิเตอร์ C เท่ากับ 0.187529 หรือกล่าวอีกนัยหนึ่งได้ว่าการเลือกค่าพารามิเตอร์ที่ไม่เหมาะสมจะส่งผลให้ประสิทธิภาพการพยากรณ์ของซัพพอร์ตเวกเตอร์แมชชีนต้อยลง นอกจากนี้ ตัวแบบผสมสามารถสร้างความมั่นใจในการนำตัวแบบผสมดังกล่าวไปใช้ในการพยากรณ์ปริมาณความต้องการไฟฟ้าสูงสุดรายปีในหลายๆ ชุดข้อมูล

เอกสารอ้างอิง

- [1] A. Hussain, M. Rahman, and J. A. Memon, “Forecasting electricity consumption in Pakistan: the way forward,” *Energy Policy*, vol. 90, pp. 73–80, March 2016.
- [2] H. H. Istepanian, “Iraq’s electricity crisis,” *The Electricity Journal*, vol. 27, pp. 51–69, May 2014.
- [3] R. G. Kavasseri and K. Seetharaman, “Day-ahead wind speed forecasting using f -ARIMA models,” *Renewable Energy*, vol. 34, pp. 1388–1393, May 2009.
- [4] V. S. Ediger and S. Akar, “ARIMA forecasting of primary energy demand by fuel in Turkey,” *Energy Policy*, vol. 35, pp. 1701–1708, March 2007.
- [5] M. Khashei and M. Bijari, “A novel hybridization of artificial neural networks and ARIMA models for time series forecasting,” *Applied Soft Computing*, vol. 11, pp. 2664–2675, March 2011.
- [6] F. Kaytez, M. C. Taplamacioglu, E. Cam, and F. Hardalac, “Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines,” *International Journal of Electrical Power & Energy Systems*, vol. 67, pp. 431–438, May 2015.
- [7] J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer, “Short-term load forecasting in a non-residential building contrasting models and attributes,” *Energy and Buildings*, vol. 92, pp. 322–330, April 2015.
- [8] G. Oğcu, O. F. Demirel, and S. Zaim, “Forecasting electricity consumption with neural networks and support vector regression,” *Procedia - Social and Behavioral Sciences*, vol. 58, pp. 1576–1585, October 2012.
- [9] K. Mohammadi, S. Shamshirband, C. W. Tong, M. Arif, D. Petković, and S. Ch, “A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation,” *Energy Conversion and Management*, vol. 92, pp. 162–171, March 2015.
- [10] J. Che and J. Wang, “Short-term electricity prices forecasting based on support vector regression and auto-regressive integrated moving average modeling,” *Energy Conversion and Management*, vol. 51, pp. 1991–1917, October 2010.
- [11] W.-C. Hong, “Electric load forecasting by support vector model,” *Applied Mathematical Modelling*, vol. 33, pp. 2444–2454, May 2009.
- [12] F. Zhang, C. Deb, S. E. Lee, J. Yang, and K. W. Shah, “Time series forecasting for building energy consumption using weighted support vector regression with differential evolution optimization technique,” *Energy and Buildings*, vol. 126, pp. 94–103, August 2016.
- [13] P. J. García Nieto, E. García-Gonzalo, F. Sánchez Lasheras, and F. J. de Cos Juez, “Hybrid



- PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability,” *Reliability Engineering & System Safety*, vol. 138, pp. 219–231, June 2015.
- [14] Peak electricity demand: Electricity Generating Authority of Thailand [online]. Available: http://www.egat.co.th/index.php?option=com_content&view=article&layout=edit&id=353&Itemid=200 (in Thai).
- [15] K. Kandananond, “Applying 2k factorial design to assess the performance of ann and svm methods for forecasting stationary and non-stationary time series,” *Procedia Computer Science*, vol. 22, pp. 60–69, 2013.
- [16] R. Schaback, “Limit problems for interpolation by analytic radial basis functions,” *Journal of Computational and Applied Mathematics*, vol. 212, pp. 127–149, March 2008.