การเปรียบเทียบตัวแบบ SARIMA และตัวแบบผสม SARIMA-SVR เพื่อการพยากรณ์ปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา

Comparison of SARIMA and Hybrid SARIMA-SVR Models for Forecasting Water Distribution Volume in Mueang District, Chachoengsao Province

นันทพร บุญสุข* รุ้งรวี อำนาจตระกูล และ วัชรากร พาหะนิชย์

Nantaporn Boonsuk* Rungrawee Amnarttrakul and Watcharakorn Pahanich

สาขาคณิตศาสตร์และสถิติประยุกต์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏราชนครินทร์
ฉะเชิงเทรา 24000 ประเทศไทย

Department of Mathematics and Applied Statistics, Faculty of Science and Technology,
Rajabhat Rajanagarindra University, Chachoengsao 24000, Thailand

*E-mail: nantaporn.boo@rru.ac.th

Received: Jun 18, 2026
Revised: Jul 25, 2026
Accepted: Aug 03, 2026

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบในการพยากรณ์ปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา ซึ่งเป็นพื้นที่เศรษฐกิจที่มีความต้องการใช้น้ำเติบโตอย่างต่อเนื่อง โดยทำการเปรียบเทียบระหว่างตัวแบบ Seasonal Autoregressive Integrated Moving Average (SARIMA) และตัวแบบผสม SARIMA ร่วมกับตัวแบบ SVR (SARIMA-SVR) ข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลปริมาณน้ำจำหน่ายรายเดือน ตั้งแต่เดือน มกราคม พ.ศ. 2558 ถึง ธันวาคม พ.ศ. 2568 รวมทั้งสิ้น 132 เดือน โดยแบ่งข้อมูลออกเป็น 2 ชุด คือ ข้อมูลสำหรับสร้างตัวแบบ (training data) 120 เดือน และข้อมูลสำหรับทดสอบการพยากรณ์ (testing data) 12 เดือน ประเมินประสิทธิภาพของตัวแบบด้วยเกณฑ์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) และเกณฑ์ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) ผลการศึกษาพบว่า ตัวแบบผสม SARIMA-SVR2 ให้ความแม่นยำในการพยากรณ์ปริมาณน้ำจำหน่ายมากที่สุด ภายใต้ข้อมูลและตัวแบบที่ศึกษา โดยมีค่า MAE เท่ากับ 20,260.07 ลูกบาศก์เมตร และ MAPE ร้อยละ 2.35 มีความแม่นยำในการพยากรณ์ร้อยละ 97.65 ตัวแบบผสม SARIMA-SVR ให้ความแม่นยำในการพยากรณ์ปริมาณน้ำจำหน่ายดีกว่าตัวแบบ SARIMA โดยมีค่า MAE เท่ากับ 20,260.07 ลูกบาศก์เมตร และ MAPE ร้อยละ 2.35 ทั้งนี้เนื่องจากตัวแบบ SARIMA สามารถตรวจจับแนวโน้มและฤดูกาลที่เป็นเชิงเส้นได้ดี ในขณะที่ตัวแบบ SVR ช่วยลดความคลาดเคลื่อนในส่วนที่ไม่เป็นเชิงเส้นได้อย่างมีประสิทธิภาพ ผลลัพธ์จากงานวิจัยนี้สามารถนำไปใช้เป็นแนวทางสนับสนุนการตัดสินใจสำหรับหน่วยงานที่เกี่ยวข้องในการวางแผนบริหารจัดการการสำรองน้ำประปา และการพัฒนาโครงสร้างพื้นฐานด้านระบบสาธารณูปโภคในเขตอำเภอเมือง จังหวัดฉะเชิงเทราได้

คำสำคัญ: การพยากรณ์ปริมาณน้ำจำหน่าย ตัวแบบ SARIMA ตัวแบบผสม SARIMA-SVR จังหวัดฉะเชิงเทรา

Abstract

This research aimed to study and compare the efficiency of forecasting models for water distribution volume in Chachoengsao Province, an economic zone experiencing continuous growth in water demand. The study compared between a standalone Seasonal Autoregressive Integrated Moving Average (SARIMA) model and a hybrid model combining SARIMA with Support Vector Regression (SARIMA-SVR). The dataset used in this study comprised monthly water distribution volumes spanning 132 months, from January 2015 to December 2025. This dataset was partitioned into two subsets: a training dataset of 120 months for model development and a testing

dataset of 12 months for out-of-sample forecast validation. Predictive performance was evaluated using Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The results revealed that the SARIMA-SVR2 hybrid model provided the highest accuracy in forecasting water distribution volume under the studied data and model, with a MAE of 20,260.07 cubic meters and a MAPE of 2.35%, achieving a forecasting accuracy of 97.65%. This superior performance is attributed to the SARIMA component's ability to effectively capture linear trends of seasonal patterns and to SVR component's ability to efficiently minimize errors derived from nonlinear residuals. The findings of this study can serve as a guideline to support decision-making for relevant agencies in planning and managing water reserves and developing public utility infrastructure in Mueang District, Chachoengsao Province.

Keywords: Water distribution forecasting, SARIMA model, Hybrid SARIMA-SVR model, Chachoengsao Province

1. บทนำ

การประปาส่วนภูมิภาคจังหวัดฉะเชิงเทรา ผลิตน้ำประปาเพื่อให้บริการแก่ประชาชนโดยพึ่งพาแหล่งน้ำดิบหลักจากแม่น้ำบางปะกง แม้ว่าที่ผ่านมาจะมีการพัฒนาและปรับปรุงระบบการผลิตอย่างต่อเนื่อง แต่ยังคงต้องเผชิญกับปัญหาวิกฤตการณ์น้ำทะเลหนุนในช่วงฤดูแล้ง ซึ่งส่งผลกระทบต่อคุณภาพและทำให้ปริมาณน้ำดิบไม่เพียงพอต่อการผลิตในบางช่วงเวลา ประกอบกับการขยายตัวอย่างรวดเร็วของภาคครัวเรือนและภาคอุตสาหกรรมในเขตพื้นที่เศรษฐกิจส่งผลให้ความต้องการใช้น้ำประปามีแนวโน้มสูงขึ้นอย่างต่อเนื่อง ดังนั้น การมีแผนยุทธศาสตร์เพื่อรองรับความต้องการใช้น้ำที่มีประสิทธิภาพจึงเป็นสิ่งจำเป็นอย่างยิ่ง การพยากรณ์ปริมาณน้ำจำหน่ายที่มีความแม่นยำสูงจึงถือเป็นกลไกสำคัญในการบริหารจัดการทรัพยากรน้ำ เพื่อเตรียมความพร้อมและสำรองน้ำประปาให้เพียงพอต่อความต้องการในอนาคต การพยากรณ์เชิงสถิตินับเป็นกระบวนการทางคณิตศาสตร์ที่นำข้อมูลและพฤติกรรมในอดีตมาคาดการณ์แนวโน้มที่จะเกิดขึ้น ซึ่งช่วยลดความเสี่ยงและเป็นข้อมูลเชิงประจักษ์ประกอบการตัดสินใจเชิงนโยบาย

จากการทบทวนวรรณกรรมที่เกี่ยวข้องพบว่า ตัวแบบอนุกรมเวลาดั้งเดิมตามระเบียบวิธีของบ็อกซ์-เจนกินส์ (Box-Jenkins) เช่น ตัวแบบ ARIMA และ SARIMA ได้รับความนิยมอย่างแพร่หลายในการพยากรณ์ปริมาณความต้องการใช้น้ำ เนื่องจากสามารถจับพฤติกรรมแนวโน้มและลักษณะฤดูกาลของข้อมูลได้เป็นอย่างดี อย่างไรก็ตาม เมื่อข้อมูลอนุกรมเวลามีลักษณะความไม่เป็นเชิงเส้น (Nonlinear) ประสิทธิภาพในการพยากรณ์ของตัวแบบดังกล่าวจะลดลง [1], [2] ในช่วงที่ผ่านมา เทคนิคการเรียนรู้ด้วยเครื่อง (Machine learning) โดยเฉพาะตัวแบบซัพพอร์ตเวกเตอร์รีเกรสชัน (Support vector regression: SVR) จึงถูกนำมาประยุกต์ใช้เพื่อจัดการกับโครงสร้างข้อมูลที่ไม่เป็นเชิงเส้น ซึ่งงานวิจัยจำนวนมากชี้ให้เห็นว่า SVR สามารถให้ความแม่นยำที่สูงกว่าตัวแบบเดี่ยวในหลายบริบท และยังมีความทนทานต่อภาวะ Overfitting บนชุดข้อมูลที่มีขนาดจำกัดได้ดี [3] ทว่าจำเป็นต้องอาศัยการปรับแต่งไฮเปอร์พารามิเตอร์อย่างระมัดระวัง ด้วยเหตุนี้ การพัฒนาตัวแบบผสม (Hybrid model) ระหว่าง SARIMA และ SVR จึงเป็นแนวทางที่มุ่งผสานจุดแข็งของทั้งสองตัวแบบ โดยให้ SARIMA ทำหน้าที่ตรวจจับโครงสร้างเชิงเส้นตามกรอบเวลา และให้ SVR จัดการกับส่วนเหลือที่ไม่เป็นเชิงเส้น (Nonlinear residuals) ซึ่งงานวิจัยส่วนใหญ่ยืนยันว่าตัวแบบผสมมักให้ความแม่นยำสูงกว่าตัวแบบเดี่ยว [1], [4], [5] แต่ผลลัพธ์สุดท้ายยังคงขึ้นอยู่กับกระบวนการเตรียมข้อมูลและลักษณะเฉพาะของอนุกรมเวลานั้น ๆ

ด้วยเหตุผลดังกล่าว ผู้วิจัยจึงมีความสนใจที่จะศึกษาและพัฒนาตัวแบบพยากรณ์ที่เหมาะสม แม่นยำ และเอื้อต่อการนำไปประยุกต์ใช้จริงสำหรับพยากรณ์ปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา โดยนำตัวแบบที่เป็นที่นิยมและมีประสิทธิภาพมาทำการศึกษาเปรียบเทียบ ได้แก่ ตัวแบบเดี่ยว SARIMA และตัวแบบผสม SARIMA-SVR เพื่อให้ได้เครื่องมือในการทำนายที่มีความคลาดเคลื่อนต่ำที่สุด ซึ่งจะสามารถใช้เป็นข้อมูลสารสนเทศพื้นฐานในการวางแผนบริหารจัดการและสร้างความมั่นคงด้านทรัพยากรน้ำประปาในจังหวัดฉะเชิงเทราได้อย่างยั่งยืนต่อไป

2. วัสดุ อุปกรณ์และวิธีการวิจัย

2.1. ข้อมูลที่ใช้ในงานวิจัย

ข้อมูลที่ใช้ในงานวิจัยนี้คือ ปริมาณน้ำจำหน่ายรายเดือนของการประปาส่วนภูมิภาค สาขาฉะเชิงเทรา ในเขตพื้นที่อำเภอเมืองฉะเชิงเทรา ย้อนหลัง 11 ปี ตั้งแต่เดือนมกราคม พ.ศ. 2558 ถึงเดือนธันวาคม พ.ศ. 2568 รวมทั้งสิ้น 132 ค่า

ซึ่งรวบรวมจากเว็บไซต์ของการประปาส่วนภูมิภาค [6] ในการวิเคราะห์ผู้วิจัยได้ทำการแบ่งข้อมูลออกเป็น 2 ชุด ดังนี้ ชุดที่ 1 ข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2558 ถึงเดือนธันวาคม พ.ศ. 2567 จำนวน 120 ค่า ใช้เป็นชุดข้อมูลสำหรับสร้างและศึกษาวิธีการพยากรณ์อนุกรมเวลา (Training set) และชุดที่ 2 ข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2568 ถึงเดือนธันวาคม พ.ศ. 2568 จำนวน 12 ค่า ใช้เป็นชุดข้อมูลทดสอบสำหรับเปรียบเทียบความแม่นยำของการพยากรณ์ (Testing set)

2.2. แนวคิดและทฤษฎีที่เกี่ยวข้อง

2.2.1. ระเบียบวิธีของบ็อกซ์-เจนกินส์ (Box-Jenkins)

เป็นวิธีการสร้างสมการพยากรณ์อนุกรมเวลาโดยการกำหนดรูปแบบตัวแบบ ARIMA ซึ่งระเบียบวิธีนี้สามารถอธิบายพฤติกรรมการเคลื่อนไหวของอนุกรมเวลาที่มีลักษณะสหสัมพันธ์ในตัวเอง (Autocorrelation) และข้อมูลมีความนิ่ง (Stationary) ทั้งนี้ ตัวแบบดังกล่าวสามารถนำไปประยุกต์ใช้กับข้อมูลที่มีพฤติกรรมการเคลื่อนไหวได้ทุกประเภท และมักให้ผลการพยากรณ์ระยะสั้นที่มีความแม่นยำสูงกว่าวิธีการพยากรณ์รูปแบบอื่น [7] โดยกระบวนการวิเคราะห์ตามระเบียบวิธีของบ็อกซ์-เจนกินส์ ประกอบด้วย 5 ขั้นตอนหลักดังต่อไปนี้

ขั้นตอนที่ 1: การตรวจสอบความนิ่งของข้อมูล (Identification of stationarity)

ดำเนินการตรวจสอบอนุกรมเวลาเพื่อพิจารณาว่าข้อมูลมีความนิ่งหรือไม่ โดยพิจารณาจากกราฟอนุกรมเวลา (Time series plot) หรือพิจารณาจากพฤติกรรมของกราฟฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation function: ACF) และกราฟฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (Partial autocorrelation function: PACF) ของอนุกรมเวลา y_t ซึ่งอนุกรมเวลาที่มีความนิ่งจะมีลักษณะเส้นกราฟที่เคลื่อนไหวขึ้นลงรอบค่าเฉลี่ยศูนย์อย่างสม่ำเสมอ ในกรณีที่อนุกรมเวลาไม่มีความนิ่ง ผู้วิจัยจำเป็นต้องแปลงข้อมูลให้เป็นอนุกรมเวลาชุดใหม่ที่มีความนิ่ง ด้วยวิธีการกำจัดแนวโน้ม (Trend) กำจัดอิทธิพลฤดูกาล (Seasonality) หรือการปรับความแปรปรวนของข้อมูลให้คงที่ เช่น การหาผลต่าง (Differencing) การหาผลต่างฤดูกาล (Seasonal differencing) หรือการแปลงข้อมูลด้วยลอการิทึม (Log transformation) ทั้งนี้ การเลือกวิธีแปลงข้อมูลจะขึ้นอยู่กับลักษณะเฉพาะของอนุกรมเวลาที่ศึกษา

ขั้นตอนที่ 2: การกำหนดรูปแบบตัวแบบที่คาดว่าเหมาะสม (Model identification)

ทำการพิจารณากำหนดรูปแบบการพยากรณ์ ARIMA(p,d,q)(P,D,Q)$_L$ ที่คาดว่ามีความเหมาะสม โดยวิเคราะห์และเปรียบเทียบลักษณะการตัดลงหรือลดลงแบบเอกซ์โพเนนเชียลของกราฟ ACF และ PACF

ขั้นตอนที่ 3: การประมาณค่าพารามิเตอร์ (Parameter estimation)

ดำเนินการประมาณค่าพารามิเตอร์ของรูปแบบตัวแบบพยากรณ์ที่กำหนดไว้ในขั้นตอนก่อนหน้า โดยใช้ระเบียบวิธีกำลังสองน้อยที่สุด (Least squares method) หรือวิธีภาวะน่าจะเป็นสูงสุด (Maximum likelihood estimation)

ขั้นตอนที่ 4: การตรวจสอบความเหมาะสมของตัวแบบ (Diagnostic checking)

ทำการตรวจสอบความเหมาะสมของตัวแบบที่พัฒนาขึ้น โดยพิจารณาจากกราฟ ACF และ PACF ของค่าความคลาดเคลื่อนส่วนเหลือ (e_t) หรือพิจารณาจากการทดสอบความอิสระของค่าความคลาดเคลื่อนด้วยสถิติทดสอบ Box-Ljung เมื่อพบว่าตัวแบบมีความเหมาะสม จึงจะนำรูปแบบดังกล่าวไปสร้างเป็นสมการพยากรณ์ แต่หากพบว่าตัวแบบยังไม่มีความเหมาะสม ผู้วิจัยต้องกลับไปเริ่มต้นกำหนดรูปแบบตัวแบบใหม่ในขั้นตอนที่ 2

ขั้นตอนที่ 5: การพยากรณ์ (Forecasting)

ภายหลังจากการตรวจสอบในขั้นตอนที่ 4 แล้วนำตัวแบบที่ผ่านเกณฑ์ความเหมาะสมมาใช้ในการสร้างสมการพยากรณ์และประมาณค่าพารามิเตอร์ เพื่อใช้สำหรับคาดการณ์ค่าของตัวแปรในอนาคตต่อไป

งานวิจัยนี้ได้ประยุกต์ใช้ฟังก์ชัน auto.arima() จากแพ็กเกจ forecast ในโปรแกรมภาษา R เพื่อค้นหาอันดับของตัวแบบที่เหมาะสม และระบุอันดับ (p,d,q)(P,D,Q)$_L$

2.2.2. ซัพพอร์ตเวกเตอร์รีเกรสชัน (Support vector regression model: SVR)

ซัพพอร์ตเวกเตอร์รีเกรสชัน (Support vector regression : SVR) เป็นวิธีการหนึ่งที่ใช้พยากรณ์ค่าที่ได้รับความนิยมอย่างมาก และมีการใช้อย่างแพร่หลาย เนื่องจากค่าที่พยากรณ์นั้นมีความแม่นยำสูง เป็นการดัดแปลงมาจาก

วิธีซัพพอร์ตเวกเตอร์แมชชีน เพื่อพยากรณ์ค่าต่อเนื่อง โดยมีเป้าหมายเพื่อสร้างฟังก์ชัน $f(x)$ ที่มีความคลาดเคลื่อนจากข้อมูลจริงไม่เกินค่า ε (Epsilon) [5] โดยมีสมการดังนี้

$$f(x) = w \cdot \phi x + b \qquad (1)$$

โดยที่

w = ค่าน้ำหนักของซัพพอร์ตเวกเตอร์

ϕx = ปริภูมิคุณลักษณะที่มีมิติที่สูง ซึ่งถูกแปลงค่าแบบไม่เป็นเส้นตรงจากปริภูมิข้อมูลนำเข้า

b = ค่าคงที่

โดยค่าสัมประสิทธิ์ w และ b จะถูกประมาณค่าด้วยการหาค่าต่ำสุดของฟังก์ชันความเสี่ยง (Risk function) [8] ดังต่อไปนี้

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\left(L\left|y_i - f(x_i)\right|\right) \qquad (2)$$

โดยที่

$$L\left(\left|y_i - f(x_i)\right|\right) = \begin{cases} \left|y_i - f(x_i)\right| - \varepsilon, & \left|y_i - f(x_i)\right| \geq \varepsilon \\ 0 & , \ other \end{cases}$$

และ $y_i \leq f(x_i) + \varepsilon$

$y_i \geq f(x_i) - \varepsilon$

$\frac{1}{2}\|w\|^2$ จะระบุถึงความราบเรียบ (Flatness) ของฟังก์ชัน โดยที่ C และ ε คือพารามิเตอร์ของตัวแบบ โดยค่า C ถูกกำหนดให้เป็นค่าคงที่เพื่อประเมินความสมดุลระหว่างค่าความคลาดเคลื่อนเชิงประจักษ์ (Empirical error) กับความราบเรียบของตัวแบบ และ $L\left|y_i - f(x_i)\right|$ ระบุถึง ε-insensitive loss function ซึ่งฟังก์ชันจะมีค่าเป็นศูนย์ หากความแตกต่างระหว่างค่าที่พยากรณ์ได้กับค่าจริงมีค่าน้อยกว่า ε ต้องเพิ่มตัวแปรที่ไม่สามารถหลีกเลี่ยงได้ เพื่อหาค่าที่เหมาะสมที่สุด [9] ในกรณีนี้ จึงจำเป็นต้องใช้ตัวแปรช่วย (Slack variables) สองตัว สำหรับเงื่อนไข $f(x_i) - y_i > \varepsilon$ และ $y_i - f(x_i) > \varepsilon$ โดยตัวแปรช่วย นี้แทนด้วยสัญลักษณ์ ξ_i และ ξ_i^* เพื่อให้สมการฟังก์ชันความเสี่ยง (Risk function) เปลี่ยนรูปไปดังต่อไปนี้ [8], [10]

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\left(\xi_i - \xi_i^*\right) \qquad (3)$$

โดยที่

$f(x_i) - y_i \leq \varepsilon + \xi_i$

$y_i - f(x_i) \leq \varepsilon + \xi_i^*$

$\xi_i, \xi_i^* \geq 0, \ i = 1, 2, ..., m$

i = ดัชนีของข้อมูลพยากรณ์

m = จำนวนของซัพพอร์ตเวกเตอร์

C = ค่าคงที่ซึ่งระบุถึงน้ำหนักของฟังก์ชันความเสี่

w = ส่วนที่ทำให้ฟังก์ชันมีความราบเรียบมากที่สุดเท่าที่จะเป็นไปได้

$\sum_{i=1}^{m}\left(\xi_i - \xi_i^*\right) = \varepsilon$ -insensitive loss function

ทั้งนี้ เพื่อให้ได้ระนาบเกิน (Hyperplane) ที่ดีที่สุด สมการดังกล่าวสามารถแสดงออกมาได้โดยใช้หลักการตัวคูณลากรานจ์ (Lagrangian multipliers principle) ซึ่งสอดคล้องกับการหาค่าเหมาะสมที่สุดแบบโปรแกรมเชิงเส้นโค้ง (Quadratic programming optimization) ดังนี้ [8], [10], [11]

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\left(\xi_i - \xi_i^*\right) + \sum_{i=1}^{m} a_i^*\left(y_i - w \cdot x_i - \varepsilon - \xi_i^*\right) + \sum_{i=1}^{m} a_i\left(w \cdot x_i - y_i - \varepsilon - \xi_i^*\right) - \sum_{i=1}^{m}\left(\lambda_i \xi_i + \lambda_i^* \xi_i^*\right) \qquad (4)$$

ตามเงื่อนไขของ Karush-Kuhn-Tucker (KKT) ผลลัพธ์จากการลดรูปตัวคูณลากรานจ์จะทำให้ได้ค่าพารามิเตอร์ w ซึ่งระบุได้ดังต่อไปนี้

$$w = \sum_{i=1}^{m}\left(a_i^* - a_i\right)x_i \qquad (5)$$

ด้วยการใช้หลักการตัวคูณลากรานจ์ และการหาค่าสูงสุดของฟังก์ชันกำลังสอง (L_p) ซึ่งเปลี่ยนจากปัญหาปฐมภูมิ (Primal problem) ให้กลายเป็นปัญหาคู่กัน (Dual problem) จะได้สมการดังต่อไปนี้

$$\max -\varepsilon\sum_{i=1}^{m}(a_i^* - a_i) + \sum_{i=1}^{m} y_i(a_i^* - a_i) - \frac{1}{2}\sum_{i,j=1}^{m}(a_i^* - a_i)(a_j^* - a_j)(a_i^* - a_i)k(x_i a_j) \qquad (6)$$

$$\sum_{i=1}^{m} a_i = \sum_{i=1}^{m} a_i^*$$

เมื่อ $0 \leq a_i \leq C, i = 1, 2, ..., m$

$0 \leq a_i^* \leq C, i = 1, 2, ..., m$

โดยที่ a_i และ a_i^* คือตัวคูณลากรานจ์ ดังนั้นจึงได้สมการสำหรับตัวแบบ SVR ดังต่อไปนี้

$$f(x) = \sum_{i=1}^{m} (a_i - a_i^*) K(x, x_i) + b \quad (7)$$

โดยที่

$K(x, x_i)$ = ฟังก์ชันเคอร์เนล

b = ค่าความเอนเอียง (Bias) ซึ่งสามารถประมาณค่าได้โดยใช้เงื่อนไข KKT โดยค่าประมาณของ b จะคำนวณได้จาก $b = y_i - w \cdot x_i - b - \varepsilon$ สำหรับเงื่อนไข $0 \le a_i \le C$ และ $b = -y_i + w \cdot x_i - \varepsilon$ สำหรับเงื่อนไข $0 \le a_i^* \le C$

ในงานวิจัยนี้ ฟังก์ชันเคอร์เนลที่ใช้คือ RBF (Radial basis function) ซึ่ง Pai and Lin [3] ได้เสนอแนะให้ใช้ฟังก์ชันเคอร์เนล RBF เนื่องจากมีประสิทธิภาพที่ดีกว่าในการประมาณค่าพฤติกรรมของข้อมูลที่ไม่เป็นเส้นตรง โดยฟังก์ชันเคอร์เนลดังกล่าวสามารถระบุได้ดังนี้ [3]

$$K(|x - x_i|) = \exp\left\{ -\frac{|x - x_i|^2}{2\sigma^2} \right\} \quad (8)$$

ดังนั้น สมการของตัวแบบ SVR ที่ใช้ฟังก์ชันเคอร์เนล RBF จึงสรุปได้ดังต่อไปนี้

$$f(x) = \sum_{i=1}^{m} (a_i - a_i^*) \exp\left\{ -\frac{|x - x_i|^2}{2\sigma^2} \right\} + b \quad (9)$$

โดยที่ $\sigma > 0$ และประกอบด้วยพารามิเตอร์ 3 ตัว ได้แก่ σ (sigma), ε (epsilon) และ C (cost) ซึ่งค่าของพารามิเตอร์ที่แตกต่างกันจะส่งผลให้ผลการพยากรณ์แตกต่างกันด้วย ดังนั้นเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด จึงจำเป็นต้องประมาณค่าพารามิเตอร์ที่เหมาะสมที่สุด (Optimal parameters) ซึ่งสามารถทำได้โดยการปรับจูน (Tuning) หรือการหาค่าที่เหมาะสมที่สุดของพารามิเตอร์ด้วยวิธีการหาค่าที่เหมาะสมที่สุด (Optimization methods)

ในงานวิจัยชิ้นนี้ดำเนินการเขียนโปรแกรมโดยใช้โปรแกรมภาษาอาร์ซึ่งมีฟังก์ชัน svm () ใน Package e1071 ที่เผยแพร่โดย Meyer et al. [12] ก่อนที่จะนำ SVR ไปใช้กับการพยากรณ์อนุกรมเวลาผู้วิจัยได้ปรับชุด ข้อมูลให้อยู่ในรูปแบบของอินพุตเวกเตอร์และทาร์เก็ต เป็นค่าสังเกตย้อนหลัง 12 คาบเวลาจากค่าพยากรณ์ และค่าความคลาดเคลื่อนที่ได้จากตัวแบบการพยากรณ์ SARIMA และปริมาณน้ำจำหน่ายจริง จะต้องมีการกำหนดเคอร์เนลฟังก์ชันและกำหนดค่าพารามิเตอร์ที่เหมาะสม และประเมินความเหมาะสมด้วยการวัดค่า MSE (Mean square error) งานวิจัยนี้ใช้ฟังก์ชัน tune() ค้นหาเคอร์เนลฟังก์ชันและกำหนดค่าพารามิเตอร์ที่เหมาะสม โดยฟังก์ชัน tune() จะค้นหาค่าที่เหมาะสมแบบกริดเสิร์จ (Grid search) ตลอดช่วงของพารามิเตอร์ที่กำหนดไว้โดยกำหนดค่าไว้ที่ 0-1 และให้เพิ่มค่อย ๆ เพิ่มค่าที่ละ 0.001 และกำหนดค่า Cost ไว้ที่ 1-16 [4] จะได้พารามิเตอร์ที่เหมาะสม 3 ตัว ได้แก่ γ (gamma), ε (epsilon) และ C (cost) โดยที่ $\gamma = \frac{1}{2\sigma^2}$

2.2.3. ตัวแบบผสม (Hybrid models)

จากวิธีการพยากรณ์โดยใช้ตัวแบบ SARIMA และตัวแบบ SVR ข้างต้นจะเห็นได้ว่าเป็นวิธีการพยากรณ์ที่เหมาะกับการพยากรณ์ข้อมูลอนุกรมเวลาต่างกันกล่าวคือ ตัวแบบ SARIMA เหมาะสำหรับการวิเคราะห์อนุกรมเวลาที่เป็นเชิงเส้นตรง ส่วนตัวแบบ SVR เป็นตัวแบบที่ใช้ได้ดีกับข้อมูลอนุกรมเวลาที่ไม่เป็นเชิงเส้นตรง ดังนั้นหากผู้พยากรณ์เลือกใช้วิธีการพยากรณ์วิธีใดวิธีหนึ่งก็จะทำให้ค่าพยากรณ์ที่ได้ขาดความแม่นยำ จากปัญหาดังกล่าว Zhang [1] จึงได้เสนอตัวแบบผสมขึ้นโดยรวมค่าพยากรณ์แบบเป็นเชิงเส้นตรงและไม่เป็นเชิงเส้นตรงเข้าด้วยกัน ค่าพยากรณ์ที่ได้จะมีความแม่นยำมากขึ้น สมการพยากรณ์ดังปรากฏในสมการที่ (10)

$$y_t = L_t + N_t + \varepsilon_t \quad (10)$$

โดยที่

y_t = ค่าสังเกตของข้อมูลอนุกรมเวลา ณ เวลา t

L_t = ข้อมูลในส่วนที่เป็นเชิงเส้นตรงในเทอมของค่าพารามิเตอร์ ณ เวลา t

N_t = ข้อมูลส่วนที่ไม่เป็นเชิงเส้นตรงในเทอมของค่าพารามิเตอร์ ณ เวลา t

ε_t = ค่าความคลาดเคลื่อน ณ เวลา t โดยคำนวณจากตัวแบบ SARIMA ดังสมการที่ (11)

$$\varepsilon_t = y_t - L_t \quad (11)$$

โดยที่

L_t = ค่าพยากรณ์ข้อมูลส่วนที่เป็นฟังก์ชันเชิงเส้นตรงที่ได้จาก ตัวแบบ SARIMA ณ เวลา t

เมื่อได้ส่วนเหลือจากสมการที่ (11) จะนำไปพยากรณ์ข้อมูลโดยใช้วิธีการ SVR

ขั้นตอนการสร้างตัวแบบผสม

ขั้นตอนที่ 1 เตรียมข้อมูล นำข้อมูลปริมาณการใช้น้ำรายเดือนของจังหวัดฉะเชิงเทรามาทำการแบ่งชุดข้อมูลออกเป็น 2 ส่วนตามลำดับเวลา ได้แก่ ชุดข้อมูลสำหรับเรียนรู้ (Training Set) เพื่อใช้ในการสร้างตัวแบบและปรับจูนพารามิเตอร์ และชุดข้อมูลสำหรับทดสอบ (Testing set) ซึ่งเป็นข้อมูลปี พ.ศ. 2568 (จำนวน 12 เดือน) เพื่อใช้ประเมินประสิทธิภาพการพยากรณ์ไปข้างหน้าโดยเฉพาะ

ขั้นตอนที่ 2 สร้างตัวแบบ SARIMA

ขั้นตอนที่ 3 คำนวณค่าความคลาดเคลื่อนตกค้าง

ขั้นตอนที่ 4 สร้างตัวแปรตามช่วงเวลาย้อนหลัง นำข้อมูลปริมาณการใช้น้ำจริง ค่าประมาณการจากตัวแบบ SARIMA และค่าความคลาดเคลื่อนตกค้าง มาทำการสร้างเป็นตัวเข้าย้อนหลัง ณ ช่วงเวลา t-1 (เพื่อจับอิทธิพลความต่อเนื่องระยะสั้นระหว่างเดือนที่ติดกัน) และ t-12 (เพื่อจับอิทธิพลของรอบพฤติกรรมฤดูกาลรายปี) เพื่อนำไปใช้เป็นตัวแปรต้นในการฝึกตัวแบบการเรียนรู้ของเครื่อง (Machine learning)

ขั้นที่ 5 ปรับจูนพารามิเตอร์และการสร้างตัวแบบ SVR ดำเนินการปรับจูนไฮเปอร์พารามิเตอร์ที่สำคัญของตัวแบบ Support Vector Regression (SVR) ได้แก่ ค่าความคลาดเคลื่อนที่ยอมรับได้ (ε) ค่า Cost (C) และค่าพารามิเตอร์ของเคอร์เนล (γ) โดยกำหนดกลยุทธ์การทวนสอบแบบย้อนหลังตามลำดับเวลา (Walk-forward validation) ภายในชุดข้อมูลเรียนรู้เพื่อป้องกันปัญหาข้อมูลรั่วไหล (Data leakage) จากนั้นทำการสร้างตัวแบบผสม SARIMA-SVR เพื่อทำการพยากรณ์โครงสร้างไม่เป็นเชิงเส้น

ขั้นตอนที่ 6 ทดสอบประสิทธิภาพของตัวแบบ ดำเนินการคัดเลือกตัวแบบผสมที่มีรูปแบบตัวแปรเข้าและชุดไฮเปอร์พารามิเตอร์ที่เหมาะสมที่สุดจากการประเมินค่าความคลาดเคลื่อนของชุดข้อมูลเรียนรู้ (Training set) โดยใช้เกณฑ์ค่า MAPE ที่ต่ำที่สุดเป็นหลักร่วมกับค่า MAE ที่ต่ำที่สุดเป็นเกณฑ์ตัดสินรอง จากนั้นนำตัวแบบผสมที่ดีที่สุดดังกล่าวไปทดสอบพยากรณ์ปริมาณการใช้น้ำจริงล่วงหน้า 12 เดือนถัดไปกับชุดข้อมูลทดสอบ (Testing set ประจำปี พ.ศ. 2568) และคำนวณค่า MAPE และ MAE เพื่อใช้เป็นดรรชนีประเมินความแม่นยำของการพยากรณ์ไปข้างหน้า

ในงานวิจัยนี้พิจารณาตัวแบบผสม 5 ตัวแบบ [4]

ตัวแบบที่ 1 SARIMA-SVR1:

$$y_t = L_t + f(\varepsilon_{t-1}, \varepsilon_{t-12}) \quad (12)$$

ตัวแบบที่ 2 SARIMA-SVR2:

$$y_t = f(y_{t-1}, y_{t-12}, \varepsilon_t) \quad (13)$$

ตัวแบบที่ 3 SARIMA-SVR3:

$$y_t = f(y_{t-1}, y_{t-12}, L_t) \quad (14)$$

ตัวแบบที่ 4 SARIMA-SVR4:

$$y_t = f(y_{t-1}, y_{t-12}, L_{t-1}, L_{t-12}) \quad (15)$$

ตัวแบบที่ 5 SARIMA-SVR5:

$$y_t = f(y_{t-1}, y_{t-12}, L_{t-1}, L_{t-12}, \varepsilon_t) \quad (16)$$

โดยที่

f = ฟังก์ชันการคำนวณของ SVR ที่ทำหน้าที่หาความสัมพันธ์แบบไม่เป็นเชิงเส้น

$\varepsilon_{t-1},...,\varepsilon_{t-12}$ = ค่าความคลาดเคลื่อน ย้อนหลังไปตั้งแต่ 1 ถึง 12 คาบเวลาโดยคำนวณจากตัวแบบ SARIMA

y_t = ค่าพยากรณ์ ณ เวลา t

N_t = ค่าพยากรณ์ที่ได้จากตัวแบบ SVR

y_t = ค่าสังเกตของอนุกรมเวลา ณ เวลา t

$y_{t-1},..., y_{t-12}$ = ค่าสังเกตย้อนหลังไปตั้งแต่ 1 ถึง 12 คาบเวลา

$\hat{L}_{t-1},..., \hat{L}_{t-12}$ = ค่าพยากรณ์ที่ได้จากตัวแบบ SARIMA ณ เวลา $t-1$ ถึง $t-12$

2.2.4. การวัดความคลาดเคลื่อนของการพยากรณ์

ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย

$$MAE = \frac{\sum_{t=1}^{n} \left| y_t - y_t \right|}{n} \quad (17)$$

ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย

$$MAPE = \frac{\sum_{t=1}^{n} \left| 1 - y_t / y_t \right|}{n} \times 100 \quad (18)$$

โดยที่ MAE และ MAPE มีค่าน้อยแสดงว่าการพยากรณ์มีความแม่นยำสูง

2.2.5. การเลือกตัวแบบที่เหมาะสม

การเปรียบเทียบประสิทธิภาพความคลาดเคลื่อนในการพยากรณ์จะพิจารณาจากวิธีที่ให้ค่า MAE และ MAPE ต่ำที่สุด ซึ่งถือเป็นวิธีการพยากรณ์ที่มีความแม่นยำที่สุด จากนั้นจึงนำวิธีดังกล่าวไปพยากรณ์ปริมาณน้ำจำหน่ายในอนาคต พร้อมทั้งทดสอบความแม่นยำของตัวแบบอีกครั้งด้วยข้อมูลชุดที่ 2 (ชุดข้อมูลทดสอบ) โดยใช้เกณฑ์ MAPE เป็นดัชนีชี้วัด

3. ผลการวิจัย

3.1. ผลการศึกษาความเคลื่อนไหวของข้อมูลอนุกรมเวลา

ลักษณะการเคลื่อนไหวของอนุกรมเวลาจากปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา (หน่วย: ลบ.ม./เดือน) จำนวน 120 ค่า พบว่า อนุกรมเวลาของปริมาณน้ำจำหน่ายประกอบไปด้วยแนวโน้มและความผันแปลของฤดูกาลดังแสดงใน Figure 1

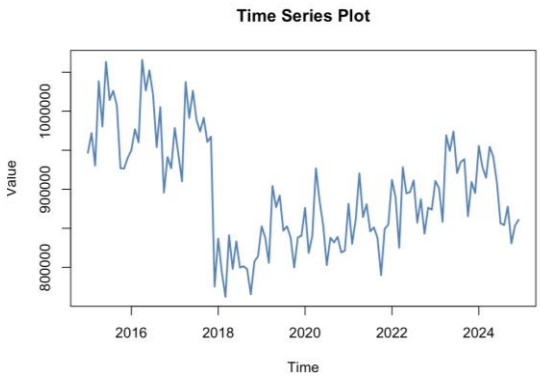


Figure 1 The characteristics of the time series data for water distribution volume in Mueang District, Chachoengsao Province, from January 2015 to December 2024

3.2. ผลการวิเคราะห์ข้อมูลค่าสถิติพื้นฐาน

ผลการวิเคราะห์สถิติพื้นฐานปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา พบว่า ปริมาณน้ำจำหน่ายให้กับประชาชนโดยเฉลี่ยเท่ากับ 895,686.21 ลบ.ม.ต่อเดือน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 68,632.53 ลบ.ม.ต่อเดือน ปริมาณน้ำจำหน่ายสูงสุดเท่ากับ 1,065,714 ลบ.ม.ต่อเดือน และต่ำสุด 762,222 ลบ.ม.ต่อเดือน ดังแสดงใน Table 1

Table 1 Descriptive statistics of water distribution volume (m^3/month) in Mueang District, Chachoengsao Province

Data	All	Training	Testing
N	132	120	12
Mean	895,686.21	897,193.69	876,471.33
SD	68,632.53	71,503.03	24,210.63
Maximum	1,065,714	1,065,714	911,183
Minimum	762,222	762,222	834,966

SD = standard deviation

3.3. ผลการพยากรณ์โดยใช้วิธีบอกซ์-เจนกินส์

ปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา พบว่า ก่อนแปลงข้อมูล ค่า ADF = -2.3472 และ ค่า P–value = 0.4211 ข้อมูลยังไม่คงที่ หลังจากแปลงข้อมูลแล้ว ค่า ADF = -4.641392 และ ค่า P-value = 0.01 ซึ่งมีลักษณะที่คงที่แล้ว ดังแสดงใน Table 2 และ เมื่อนำไปเขียนกราฟหลังจากแปลงข้อมูลด้วยการหาผลต่างอันดับที่ 1จะเห็นได้ชัดว่าข้อมูลมีลักษณะคงที่ ดังแสดงใน Figure 2

Table 2 ADF test results for water distribution volume in Mueang District, Chachoengsao Province

Before Transformation		After Transformation	
ADF	P-value	ADF	P-value
-2.3742	0.4211	-4.6414	0.01

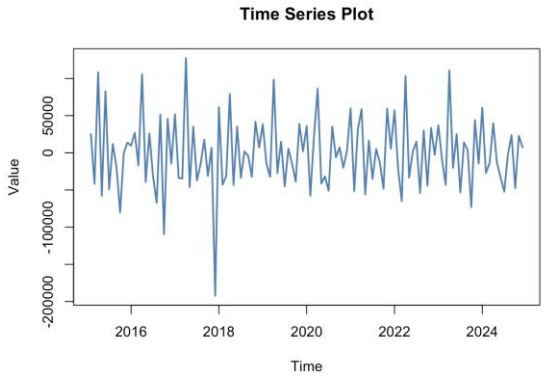


Figure 2 The stationary characteristics of the time series data for water distribution volume in Mueang District, Chachoengsao Province

การค้นหาอันดับของตัวแบบที่เหมาะสมที่สุด ซึ่งผลการวิเคราะห์และรูปแบบของตัวแบบ SARIMA ที่ได้ ดังแสดงใน Table 3

Table 3 Parameter estimation of the SARIMA model for water distribution volume in Mueang District, Chachoengsao Province

Model	ARIMA(1,1,0)(2,0,0)$_{12}$		
	AR1	SAR1	SAR2
Estimate	-0.3424	0.4178	0.2616
Std.Error	0.0873	0.0864	0.0933
Z-value	-3.9207	4.8356	2.8031
P-value	<0.0001***	<0.0001***	0.0051**

การตรวจสอบความเหมาะสมของตัวแบบ โดยพิจารณาตามข้อตกลงเบื้องต้นของตัวแบบ SARIMA ซึ่งมีทั้งหมด 4 ข้อ ได้แก่

1) ความคลาดเคลื่อนมีค่าเฉลี่ยเท่ากับศูนย์ผลจากการทดสอบพบว่ามีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับศูนย์ที่ t-value = -0.1711 และ P-value = 0.8644 ดังแสดงใน Table 4

2) ความคลาดเคลื่อนมีการแจกแจงแบบปกติ ผลจากการทดสอบพบว่า ความคลาดเคลื่อนมีการแจกแจงแบบปกติที่ KS-value = 0.0874 และ P-value = 0.3190 ดังแสดงใน Table 4 และ Figure 3

3) ความคลาดเคลื่อนเป็นอิสระต่อกัน โดยใช้ Ljung-Box Q-statistics พบว่า ความคลาดเคลื่อนเป็นอิสระต่อกัน ค่า Q-statistics = 25.049 และ P-value = 0.2451 ดังแสดงใน Table 5 และ Figure 4

4) ความคลาดเคลื่อนมีค่าความแปรปรวนคงที่ พบว่าความคลาดเคลื่อนมีค่าความแปรปรวนคงที่ ดังแสดงใน Figure 5 ดังนั้นสรุปได้ว่าตัวแบบที่ได้มีความเหมาะสมที่จะนำไปพยากรณ์

Table 4 Assumption testing of the SARIMA model for water distribution volume in Mueang District, Chachoengsao Province

t-value	P-value	KS-value	P-value
-0.1711	0.8644	0.0874	0.3190

The residuals exhibit a normal distribution with a mean of zero.

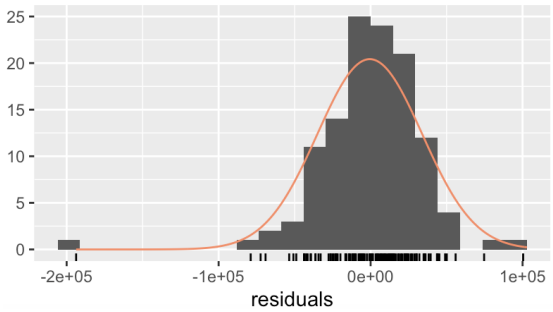


Figure 3 Histogram of normally distributed residuals for water distribution volume in Mueang District, Chachoengsao Province

3.4. ผลการพยากรณ์ด้วยตัวแบบผสมระหว่างวิธีบอกซ์-เจนกินส์ และวิธีซัพพอร์ตเวกเตอร์รีเกรสชัน

ผลการวิเคราะห์และคัดเลือกเคอร์เนลฟังก์ชันที่เหมาะสมสำหรับตัวแบบ SVR บ่งชี้ว่า เคอร์เนลแบบเรเดียลเบสิส (RBF) ให้ประสิทธิภาพดีที่สุดในการพยากรณ์ปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา โดยพารามิเตอร์สำคัญของเคอร์เนล RBF ซึ่งประกอบด้วยค่า C, ε และ γ ที่ผ่านการปรับแต่งอย่างเหมาะสม เพื่อนำไปใช้เป็นฐานในการพยากรณ์และเปรียบเทียบประสิทธิภาพของตัวแบบในขั้นตอนถัดไป ดังแสดงใน Table 6

Table 5 Ljung-Box Q-statistics for autocorrelation testing of residuals for water distribution volume in Mueang District, Chachoengsao Province

Q	Degree of freedom	P-value
25.049	21	0.2451

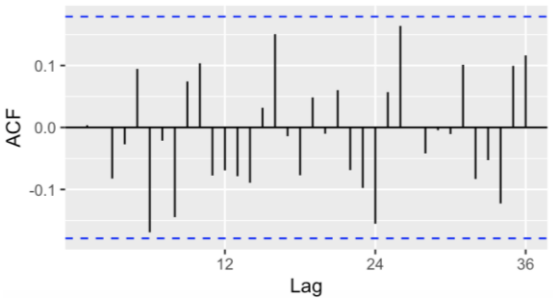


Figure 4 Autocorrelation function (ACF) plot of residuals for the water distribution volume model in Mueang District, Chachoengsao Province

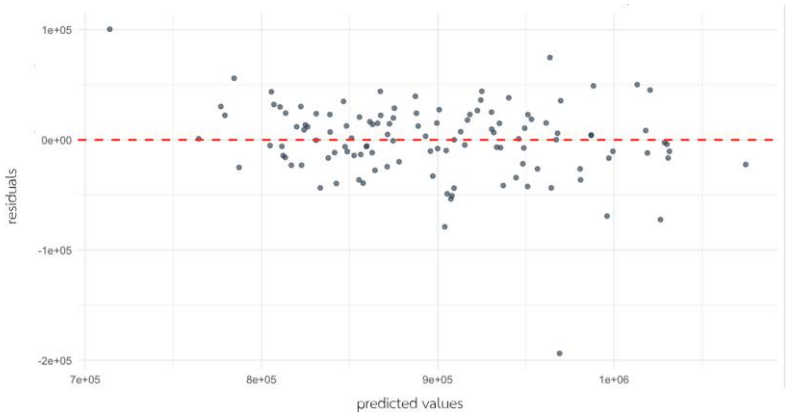


Figure 5 Scatter plot of residuals versus predicted values

Table 6 Optimal parameters of the hybrid model

Model	C	γ	ε
SARIMA-SVR1	1	0.001	0.235
SARIMA-SVR2	4	0.1	0.024
SARIMA-SVR3	16	0.1	0.196
SARIMA-SVR4	7	0.1	0.235
SARIMA-SVR5	16	0.01	0.146

ในการคัดเลือกตัวแบบพยากรณ์ปริมาณน้ำจำหน่ายในเขตจังหวัดฉะเชิงเทรา ผู้วิจัยได้นำผลการพยากรณ์ของตัวแบบเดี่ยว (SARIMA) และตัวแบบผสม (SARIMA-SVR1 ถึง SARIMA-SVR5) มาเปรียบเทียบความคลาดเคลื่อนเชิงสถิติจากการประเมินค่าความคลาดเคลื่อนของชุดข้อมูลเรียนรู้ (Training set) โดยใช้เกณฑ์ค่า MAPE ที่ต่ำที่สุดเป็นหลักร่วมกับค่า MAE ที่ต่ำที่สุดเป็นเกณฑ์ตัดสินรอง พบว่า ตัวแบบ SARIMA-SVR2 ให้ค่าความคลาดเคลื่อนต่ำที่สุด จึงเป็นตัวแบบที่เหมาะสมที่สุดสำหรับใช้พยากรณ์ปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา ดังแสดงใน Table 7

Table 7 Forecasting errors of the six models

Model	MAE	MAPE
SARIMA	24277.32	2.78
SARIMA-SVR1	24204.74	2.78
SARIMA-SVR2	15096.26	1.72
SARIMA-SVR3	25763.00	2.95
SARIMA-SVR4	24001.38	2.74
SARIMA-SVR5	16155.66	1.84

นำตัวแบบผสมที่ดีที่สุดดังกล่าวไปทดสอบพยากรณ์ปริมาณการใช้น้ำจริงล่วงหน้า 12 เดือนถัดไปกับชุดข้อมูลทดสอบ (Testing set ประจำปี พ.ศ. 2568) พบว่า ตัวแบบพยากรณ์ให้ค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์ (MAE) เท่ากับ 20,260.07 ลูกบาศก์เมตรต่อเดือน และมีค่าความคลาดเคลื่อนเปอร์เซ็นต์เฉลี่ยสัมบูรณ์ (MAPE) เท่ากับ 2.35% ดังแสดงใน Table 8

เมื่อพิจารณากราฟเปรียบเทียบระหว่างปริมาณน้ำจำหน่ายจริงและค่าพยากรณ์รายเดือน พบว่า เส้นกราฟค่าพยากรณ์สามารถจับทิศทาง (Trend) และความผันแปรโดยรวมของข้อมูลจริงได้ดี โดยเฉพาะในช่วงครึ่งปีหลัง (เดือนสิงหาคม ถึง ธันวาคม) ที่ค่าพยากรณ์มีความใกล้เคียงกับค่าจริงเป็นอย่างมาก อย่างไรก็ตาม ตัวแบบพยากรณ์ยังคงมีข้อจำกัดในการพยากรณ์ค่าที่ลดลงอย่างรวดเร็วผิดปกติในบางคาบเวลา เช่น ในเดือนมีนาคมและเดือนตุลาคม ซึ่งเป็นจุดที่ข้อมูลจริงมีปริมาณน้ำจำหน่ายต่ำกว่าแนวโน้มปกติอย่างเห็นได้ชัด ดังแสดงใน Figure 6

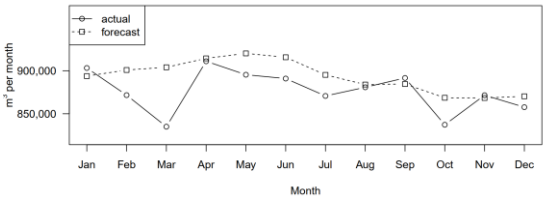


Figure 6 Actual versus predicted values of water distribution volume in Mueang District, Chachoengsao Province

Table 8 Comparison between actual and predicted values of water distribution volume in Mueang District, Chachoengsao Province for 12 months

Month	Actual (m^3/month)	Predicted (m^3/month)	Error (m^3/month)
January 2025	903,347	893,882.3	9,464.7
February 2025	871,730	901,024.6	-29,294.6
March 2025	834,966	904,160.3	-69,194.3
April 2025	911,183	914,554.8	-3,371.8
May 2025	895,489	920,353.2	-24,864.2
June 2025	891,046	915,849.1	-24,803.1
July 2025	870,779	895,334.8	-24,555.8
August 2025	880,858	884,065.6	-3,207.6
September 2025	891,722	884,657.1	7,064.9
October 2025	837,219	868,632.2	-31,413.2
November 2025	871,629	868,412.3	3,216.7
December 2025	857,688	870,357.9	-12,669.9
MAE (m^3/month)	20260.07		
MAPE (%)	2.35		

4. อภิปรายผลการวิจัย

จากผลการศึกษาพบว่า ตัวแบบผสม SARIMA-SVR (โดยเฉพาะ SARIMA-SVR2) ให้ประสิทธิภาพในการพยากรณ์สูงกว่าตัวแบบเดี่ยว SARIMA เนื่องจากโครงสร้างตัวแปรเข้าของ SARIMA-SVR2 ที่มีการผสานข้อมูลเชิงเส้นและส่วนเหลือย้อนหลัง (t-1, t-12) อย่างสมดุล ทำให้ SVR เรียนรู้รูปแบบความคลาดเคลื่อนได้ดีกว่าตัวแบบอื่น แสดงว่า ปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา มีรูปแบบพฤติกรรมของข้อมูลที่มีความซับซ้อนและมีความสัมพันธ์แบบไม่เป็นเชิงเส้นปะปนอยู่ ซึ่งเกิดจากปัจจัยภายนอก เช่น การขยายตัวของภาคอุตสาหกรรม การเพิ่มขึ้นของประชากร และปัจจัยด้านฤดูกาล การใช้ตัวแบบเดี่ยว SARIMA จึงมีข้อจำกัดเนื่องจากพัฒนาขึ้นภายใต้สมมติฐานเชิงเส้น อย่างไรก็ตาม เมื่อนำระเบียบวิธีของบ็อกซ์-เจนกินส์มาประยุกต์ร่วมกับซัพพอร์ตเวกเตอร์รีเกรสชัน (SVR) โดยให้ SVR ทำหน้าที่เรียนรู้และปรับลดความคลาดเคลื่อนในส่วนที่เหลือ ซึ่งไม่เป็นเชิงเส้น (Nonlinear residuals) จึงส่งผลให้ตัวแบบผสมสามารถเพิ่มความแม่นยำและลดความคลาดเคลื่อนในการพยากรณ์ลงได้อย่างมีประสิทธิภาพ ซึ่งผลการศึกษานี้สอดคล้องกับงานวิจัยของ Zhang [1], Sutthison [4] และ Albarr and Kusumawati [5] ที่พบว่าตัวแบบผสมช่วยเพิ่มประสิทธิภาพการพยากรณ์อนุกรมเวลาได้ดีกว่าการใช้ตัวแบบเดี่ยว

เมื่อพิจารณากราฟเปรียบเทียบค่าจริงและค่าพยากรณ์ดังแสดงใน Figure 6 แม้ว่าในภาพรวมตัวแบบผสมจะสามารถจับทิศทางหลักได้ดี แต่ยังคงปรากฏความคลาดเคลื่อนในบางคาบเวลา โดยเฉพาะในเดือนมีนาคมและเดือนตุลาคม พ.ศ. 2568 ซึ่งเป็นช่วงเวลาที่ปริมาณน้ำจำหน่ายจริงดิ่งลดลงต่ำกว่าแนวโน้มปกติอย่างมาก ความคลาดเคลื่อนดังกล่าวอาจเนื่องมาจากเดือนมีนาคมและตุลาคมเป็นช่วงน้ำทะเลหนุนสูงและเป็นช่วงลอยต่อของฤดูกาล ซึ่งส่งผลให้กิจกรรมทางการค้าหรือภาคครัวเรือนลดการใช้น้ำลงชั่วคราว พฤติกรรมที่ผันผวนอย่างรวดเร็วและกะทันหันเช่นนี้ถือเป็นข้อจำกัดของตัวแบบพยากรณ์ในการปรับตัวตามข้อมูลที่มีการเปลี่ยนแปลงอย่างรุนแรงในระยะสั้น สำหรับการวิจัยในอนาคต ควรพัฒนาตัวแบบผสม SARIMAX-SVR ที่มีการนำตัวแปรภายนอก (Exogenous variables) เช่น ปัจจัยน้ำทะเลหนุน สภาพอากาศ และดัชนีฤดูกาล เข้ามาร่วมพยากรณ์เพื่อลดความคลาดเคลื่อนในช่วงเวลาผันผวน

5. บทสรุป

งานวิจัยนี้ดำเนินงานเพื่อศึกษาและเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์ปริมาณน้ำจำหน่ายในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา โดยใช้ข้อมูลรายเดือนตั้งแต่ปี พ.ศ. 2558 ถึง พ.ศ. 2568 รวม 132 เดือน แบ่งเป็นชุดข้อมูลสร้างตัวแบบ 120 เดือน และชุดข้อมูลทดสอบ 12 เดือน ผลการศึกษาสรุปได้ว่า ตัวแบบผสม SARIMA-SVR2 (ซึ่งใช้เคอร์เนลฟังก์ชันแบบ RBF) เป็นตัวแบบที่เหมาะสมและมีความแม่นยำมากที่สุด ภายใต้ข้อมูลและตัวแบบที่ศึกษา โดยให้ค่าความคลาดเคลื่อนเปอร์เซ็นต์เฉลี่ยสัมบูรณ์ (MAPE) ต่ำที่สุดเพียงร้อยละ 2.35 และมีค่า MAE เท่ากับ 20,260.07 ลูกบาศก์เมตร ซึ่งถือว่าตัวแบบมีความแม่นยำในการพยากรณ์อยู่ในระดับสูง [13] ผลลัพธ์เชิงตัวเลขและตัวแบบพยากรณ์ที่พัฒนาขึ้นในงานวิจัยนี้ สามารถนำไปใช้เป็นแนวทางในการสนับสนุนการตัดสินใจเชิงกลยุทธ์สำหรับหน่วยงานการประปาและองค์กรปกครองส่วนท้องถิ่นในเขตอำเภอเมือง จังหวัดฉะเชิงเทรา เพื่อใช้ในการวางแผนสำรองน้ำประปา การบริหารจัดการความมั่นคงด้านทรัพยากรน้ำให้สอดคล้องกับการเติบโตของเขตเศรษฐกิจ ตลอดจนการวางโครงสร้างพื้นฐานระบบสาธารณูปโภคเพื่อรองรับความต้องการใช้น้ำในอนาคต

6. กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณ สถาบันวิจัยและพัฒนา และคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏราชนครินทร์ ที่ให้การสนับสนุนการทำวิจัยในครั้งนี้

7. References

[1] Zhang, G.P. 2003. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing.** 50: 159-175.

[2] Box, G.E.P. and et al. 2008. **Time Series Analysis: Forecasting and Control. 4th edition.** Hoboken: John Wiley and Sons, Inc.

[3] Pai, P.F. and Lin, C.S. 2005. A hybrid ARIMA and support vector machines model in stock price forecasting. **Omega.** 33(6): 497-505.

[4] Sutthison, T. 2020. Appropriate models for forecasting of water supply consumption in Ubonratchathani Province. **Burapha Science Journal.** 25(1): 28-50. (*in Thai*)

[5] Albarr, H. and Kusumawati, R. 2023. Hybrid autoregressive integrated moving average-support vector regression for stock price forecasting. **Jurnal Matematika, Sains, dan Teknologi.** 24(2): 1-17.

[6] Provincial Waterworks Authority. 2025. **Data Water Consumption.** https://www.pwa.co.th/province/. Accessed 20 February 2026. (*in Thai*)

[7] Bowerman, B.L., O'Connell, R.T. and Koehler, A.B. 2005. **Forecasting, Time series and Regression: An Applied Approach. 4th edition.** Belmont: Thomson Brooks/Cole.

[8] Vapnik, V.N. 1995. **The Nature of Statistical Learning Theory.** Berlin: Springer-Verlag.

[9] Ojemakinde, B.T. 2006. **Support Vector Regression for Non-Stationary Time Series.** M.Sc. Thesis, University of Tennessee, Knoxville.

[10] Scholkopf, B. and Smola, A.J. 2001. **Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond.** Cambridge: MIT Press.

[11] Awad, M. and Khanna, R. 2015. **Efficient Learning Machines.** Berkeley: Apress.

[12] Meyer, D. and et al. 2025. **e1071: Misc functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien.** https://CRAN.R-project.org/package=e1071. Accessed 22 April 2026.

[13] Lewis, C.D. 1982. **Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting.** Oxford: Butterworth-Heinemann.