

# การหาตัวแบบที่เหมาะสมเพื่อพยากรณ์ปริมาณการใช้น้ำประปาในจังหวัดอุบลราชธานี

## Appropriate Models for Forecasting of Water Supply Consumption in Ubonratchathani Province

ธนกร สุทธิสนธิ

Thanakon Sutthison\*

สาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี

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

Received : 25 April 2019

Revised : 2 July 2019

Accepted : 21 August 2019

### บทคัดย่อ

ปริมาณการใช้น้ำประปาของประชาชนในจังหวัดอุบลราชธานี ปัจจุบันมีแนวโน้มเพิ่มมากขึ้น การพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนโดยใช้ตัวแบบที่เหมาะสมและมีความแม่นยำจึงมีความสำคัญและเป็นประโยชน์อย่างยิ่งในวางแผนและหามาตรการรองรับกับสถานการณ์การใช้น้ำประปาที่เพิ่มขึ้น การวิจัยครั้งนี้ มีวัตถุประสงค์ เพื่อสร้างตัวแบบที่เหมาะสมกับอนุกรมเวลาปริมาณการใช้น้ำประปาของประชาชนในจังหวัดอุบลราชธานี จำนวน 4 สาขา ได้แก่ สาขาอุบลราชธานี สาขาพิบูลมังสาหาร สาขาเดชอุดม และสาขาเขมราฐ โดยใช้ข้อมูลจากเว็บไซต์ของการประปาส่วนภูมิภาค ตั้งแต่เดือนมกราคม พ.ศ. 2547 ถึงเดือนธันวาคม พ.ศ. 2560 จำนวน 168 ค่า ข้อมูลถูกแบ่งออกเป็น 2 ชุด คือ ข้อมูลชุดที่ 1 ตั้งแต่เดือนมกราคม พ.ศ. 2547 ถึงเดือนธันวาคม พ.ศ. 2559 จำนวน 156 ค่า ใช้สำหรับสร้างตัวแบบการพยากรณ์แบบอนุกรมเวลาเชิงเดี่ยวตามวิธีการของบอซ-เจนกินส์ และตัวแบบผสมระหว่างเทคนิคของบอซ-เจนกินส์กับเทคนิคซัพพอร์ตเวกเตอร์เกรสชัน ข้อมูลชุดที่ 2 ตั้งแต่เดือนมกราคม ถึงเดือนธันวาคม พ.ศ. 2560 จำนวน 12 ค่า ใช้สำหรับเปรียบเทียบความแม่นยำของการพยากรณ์แต่ละสาขา โดยใช้เกณฑ์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) และเกณฑ์ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percent Error : MAPE) ผลการวิจัย พบว่า ตัวแบบผสมมีความแม่นยำในการพยากรณ์มากกว่าตัวแบบการพยากรณ์เชิงเดี่ยว เนื่องจากมีค่า MAE และ MAPE ต่ำสุด ดังนั้น ตัวแบบผสมสามารถใช้เป็นเครื่องมือในการพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนในจังหวัดอุบลราชธานีได้อย่างเหมาะสม และสามารถใช้ประกอบการตัดสินใจสำหรับการวางแผนการจัดการน้ำประปาให้เพียงพอต่อความต้องการของประชาชนในอนาคตได้

**คำสำคัญ :** การใช้น้ำประปา, การพยากรณ์, บอซ - เจนกินส์, ตัวแบบผสม, จังหวัดอุบลราชธานี

\*Corresponding author. E-mail : thanakon.sutthison@gmail.com

## Abstract

At present, water supply consumption in Ubonratchathani province is increasing. Forecasting water consumption of people using appropriate and accurate model is very important and useful for planning the measures to support the increased use of water supply. The purpose of this research was to construct a model suitable for the time series of water consumption of people in Ubonratchathani province in 4 stations, including Ubonratchathani Station, Phipun mangsahan Station, Det Udom Station, and Khemarat Station. The data of 168 values were obtained from the website of Provincial Waterworks Authority from January 2004 to December 2017. The data were divided into two sets. Set 1 was the data from January 2004 to December 2016, of which 156 values were used for constructing univariate time series based on the method of Box – Jenkins and the hybrid model of the technique of Box – Jenkins and support vector regression. Set 2 was the data from January to December 2017, of which 12 values were used for the comparison of the accuracy of forecasting of each station by using Mean Absolute Error (MAE) and Mean Absolute Percent Error (MAPE). The results showed that the hybrid model was more accurate in forecasting than a single model due to the lowest MAE and MAPE values. The hybrid model can be used as a tool to forecast the amount of water supply of people in Ubon Ratchathani Province appropriately. It can also be used for decision making for planning water supply management to meet the needs of people in the future.

**Keywords:** water supply consumption, forecasting, Box – Jenkins, hybrid forecasting, Ubonratchathani Province

## บทนำ

ปัจจุบันความต้องการใช้น้ำประปาของประชาชนในจังหวัดอุบลราชธานี เป็นวิกฤตหนึ่งที่สำคัญอย่างยิ่งโดยเฉพาะในช่วงฤดูแล้งประชาชนมีน้ำไม่เพียงพอต่อความต้องการ เนื่องจากจำนวนคนใช้น้ำกับปริมาณน้ำประปาที่ผลิตไม่เพียงพอและปัจจุบันการขยายตัวทางด้านเศรษฐกิจ การขยายตัวของชุมชนซึ่งเป็นชุมชนเมืองมากขึ้นความต้องการใช้น้ำประปาก็จะมากขึ้นตามไปด้วย ดังนั้นการวางแผนเพื่อเตรียมรองรับการใช้น้ำของประชาชนจึงเป็นสิ่งสำคัญอย่างยิ่ง การพยากรณ์เชิงสถิติเป็นอีกเทคนิคหนึ่งที่มีการนำมาใช้ประโยชน์และแก้ปัญหาเพื่อวางแผนตัดสินใจอย่างหลากหลาย เช่น การพยากรณ์การใช้น้ำประปาของประชาชน (Tongpitak *et al.*, 2017) การพยากรณ์ความเร็วลม (Yang & Wang, 2018) การพยากรณ์การใช้พลังงานไฟฟ้า (Cheng *et al.*, 2019; de Oliveira & Cyrino Oliveira, 2018) เป็นต้น จากการศึกษางานวิจัยดังกล่าวพบว่านักวิจัยแต่ละท่านพยายามสร้างตัวแบบและนำเสนอตัวแบบเพื่อพยากรณ์ข้อมูลอนุกรมเวลาให้มีความแม่นยำมากที่สุด จากการทบทวนเอกสารพบว่า การพยากรณ์อนุกรมเวลาแบ่งเป็น 2 กลุ่มใหญ่ ได้แก่ (1) วิธีการแบบเชิงเส้น (Linear approach) เป็นวิธีการพยากรณ์ที่นิยมและใช้อย่างแพร่หลาย ได้แก่ วิธีของบอซ-เจนกินส์ ตัวแบบที่ใช้ได้แก่ ARIMA และ SARIMA แต่มีข้อเสีย คือ ไม่สามารถวิเคราะห์ข้อมูลอนุกรมเวลาแบบไม่เป็นเชิงเส้นตรงได้ และ (2) วิธีการแบบไม่เป็นเชิงเส้นตรง (Non -Linear approach) วิธีที่

นิยมใช้ ได้แก่ วิธีโครงข่ายประสาทเทียม (Artificial Neural Network) วิธีซัพพอร์ตเวกเตอร์รีเกรสชัน (Support vector regression model: SVR) ซึ่งมีความยืดหยุ่นสูงสามารถนำไปใช้ได้หลากหลายโดยเฉพาะข้อมูลที่มีความซับซ้อนสูงสามารถพยากรณ์ข้อมูลในอนาคตได้ทั้งระยะสั้นและระยะยาว แต่ไม่สามารถวิเคราะห์ข้อมูลที่มีลักษณะการเปลี่ยนแปลงตามฤดูกาลและแนวโน้มในอนุกรมเวลาแบบเชิงเส้นได้ (Junsagoon, 1972) จากวิธีการพยากรณ์อนุกรมเวลาทั้งแบบเชิงเส้นตรงและไม่เป็นเชิงเส้นตรง ต่างก็มีข้อดีและข้อด้อยแตกต่างกัน จึงเกิดแนวคิดการผสมผสานวิธีการพยากรณ์ทั้งสองแบบเข้าด้วยกันโดยเรียกชื่อว่า ตัวแบบผสม (Hybrid Models) เพื่อเป็นการแก้ปัญหาในส่วนที่วิธีการพยากรณ์แบบเชิงเส้นตรงและวิธีการพยากรณ์แบบไม่เป็นเชิงเส้นตรงไม่สามารถอธิบายได้ ทำให้ได้ค่าพยากรณ์ที่มีความแม่นยำมากขึ้น (Zhang, 2003) ด้วยเหตุนี้จึงมีงานวิจัยต่าง ๆ ทำการพัฒนาตัวแบบผสมเพื่อพยากรณ์ข้อมูลอนุกรมเวลาอย่างหลากหลาย และพบว่า ตัวแบบผสมให้ค่าความแม่นยำของการพยากรณ์สูงกว่าการใช้ตัวแบบการพยากรณ์เชิงเดี่ยว (Barman & dev Choudhury, 2019; Cheng *et al.*, 2019) จากเหตุผลดังกล่าว เพื่อเพิ่มความแม่นยำในการพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนในจังหวัดอุบลราชธานี ผู้วิจัยจึงพัฒนาวิธีการพยากรณ์แบบผสมขึ้น โดยจะรวมอนุกรมเวลาแบบเชิงเส้นตรงและอนุกรมเวลาแบบไม่เป็นเชิงเส้นเข้าด้วยกัน การพยากรณ์อนุกรมเวลาแบบเชิงเส้นตรง ผู้วิจัยจะเลือกใช้วิธีการพยากรณ์ตามวิธีของบ็อก – จินกินส์ โดยใช้ตัวแบบ SARIMA เนื่องจากตัวแบบดังกล่าวมีประสิทธิภาพในการพยากรณ์อนุกรมเวลาที่มีความสัมพันธ์เชิงเส้นตรงได้ดีและมีการประยุกต์ใช้อย่างหลากหลายในปัจจุบัน (Tongpitak *et al.*, 2017; Mao *et al.*, 2018; Arumugam & Saranya, 2018; Curceac *et al.*, 2019) แต่ตัวแบบ SARIMA มีข้อกำหนดในการสร้างตัวแบบที่เหมาะสมจำนวนมาก ประกอบกับข้อมูลอนุกรมเวลาในปัจจุบันมีความสัมพันธ์ที่ซับซ้อนจึงทำให้ค่าพยากรณ์ที่ได้มีความคลาดเคลื่อนสูง ส่วนวิธีการพยากรณ์แบบไม่เป็นเชิงเส้นตรงจะเลือกใช้วิธีซัพพอร์ตเวกเตอร์รีเกรสชัน เนื่องจากเป็นวิธีการพยากรณ์ที่มีความแม่นยำและมีประสิทธิภาพสูง โดยมีจุดเด่นในการสร้างฟังก์ชันการทำนายจากลักษณะข้อมูลที่มีความสัมพันธ์ไม่เป็นเชิงเส้นตรงได้ดีและมีข้อกำหนดไม่มาก และเป็นที่ยอมรับในปัจจุบัน (Yaslan & Bican, 2017; Chen *et al.*, 2017; Xu, Chan, & Zhang, 2019) ซึ่งเหมาะที่จะนำมาใช้สำหรับการสร้างความสัมพันธ์เพื่ออธิบายลักษณะข้อมูลอนุกรมเวลาที่มีความสัมพันธ์ไม่เป็นเส้นตรงของส่วนเหลือ (Residuals) จากตัวแบบ SARIMA (Sujavinyasup, 2019) ตัวแบบผสมที่พัฒนาขึ้นในงานวิจัยนี้มีทั้งหมด 5 ตัวแบบ คือ SARIMASVR1 SARIMASVR2 SARIMASVR3 SARIMASVR4 และ SARIMASVR5 จากนั้นพิจารณาเลือกตัวแบบที่เหมาะสมเพื่อนำไปใช้ในการพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนในอนาคต โดยใช้เกณฑ์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย และเกณฑ์ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ผลการวิจัยที่ได้จะเป็นประโยชน์ให้แก่หน่วยงานที่เกี่ยวข้องใช้ประกอบการดำเนินการวางแผนการจัดการทรัพยากรน้ำได้อย่างมีประสิทธิภาพและเพียงพอต่อความต้องการของประชาชนต่อไป

## วิธีดำเนินการวิจัย

### 1. ข้อมูลที่ใช้ในการศึกษา

ปริมาณการใช้น้ำประปาของประชาชน คือ ปริมาณน้ำประปาที่ทางการประปาส่วนภูมิภาคในจังหวัดอุบลราชธานี ประกอบด้วย 4 สาขา ได้แก่ สาขาอุบลราชธานี สาขาพิบูลมังสาหาร สาขาเดชอุดม และสาขาเขมราฐ ผลิตและจำหน่ายให้กับประชาชน ซึ่งงานวิจัยครั้งนี้ทำการเก็บรวบรวมข้อมูลอนุกรมเวลาปริมาณการใช้น้ำประปาของประชาชน โดยรวม จาก 3 ประเภทใหญ่ๆ คือ 1) ที่อยู่

อาศัยและอื่น ๆ 2) ราชการและธุรกิจขนาดเล็ ก และ 3) รัฐวิสาหกิจอุตสาหกรรมและธุรกิจขนาดใหญ่ของจังหวัดอุบลราชธานี 4 สาขา ได้แก่ สาขาอุบลราชธานี สาขาพิบูลมังสาหาร สาขาเดชอุดม และสาขาเขมราฐ (หน่วย: ลบ.ม./เดือน) ย้อนหลัง 14 ปี คือตั้งแต่เดือน มกราคม พ.ศ.2547 ถึงเดือนธันวาคม พ.ศ.2560 จำนวน 168 ค่า จากเว็บไซต์ของการประปาส่วนภูมิภาค (Provincial Waterworks Authority, 2018) เมื่อได้ข้อมูลผู้วิจัยทำการแบ่งข้อมูลออกเป็น 2 ชุด ชุดที่ 1 คือข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2547 ถึงเดือนธันวาคม พ.ศ. 2559 จำนวน 156 ค่า ใช้สำหรับศึกษาวิธีการพยากรณ์อนุกรมเวลาจำนวน 6 ตัวแบบ ผู้วิจัยได้ประยุกต์ใช้ โปรแกรมภาษาอาร์ (R-Language) (R Core Team, 2017) ในการวิเคราะห์ข้อมูล การเลือกใช้วิธีการพยากรณ์ที่เหมาะสม ผู้วิจัยใช้การเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์ 2 เกณฑ์ คือเกณฑ์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย และเกณฑ์ ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย เนื่องจากมีการประยุกต์ใช้ในการพยากรณ์อย่างหลากหลายและง่ายต่อการพิจารณา (Tongpitak *et al.*, 2017; Yang & Wang,2018; Cheng *et al.*, 2019; de Oliveira & Oliveira, 2018) ตัวแบบการพยากรณ์ใด ที่ให้ค่า ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย และเกณฑ์ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด แสดงว่า เป็นวิธีการพยากรณ์ ที่มีความแม่นยำที่สุด ข้อมูลชุดที่ 2 คือ ข้อมูลตั้งแต่เดือนมกราคม ถึงเดือนธันวาคม พ.ศ. 2560 จำนวน 12 ค่า ใช้สำหรับการ เปรียบเทียบความแม่นยำของการพยากรณ์ของแต่ละสาขา

2. แนวคิดและทฤษฎีที่เกี่ยวข้อง

2.1 ตัวแบบการพยากรณ์โดยวิธีของบอกรี -เจนกินส์ (Box –Jenkins Method) ตัวแบบการพยากรณ์ตามวิธีของ บอกรี – เจนกินส์ เป็นวิธีการพยากรณ์หนึ่งที่มีความถูกต้องแม่นยำสูง เนื่องจากมีการพิจารณาลักษณะของอนุกรมเวลาว่ามี สหสัมพันธ์กันอย่างไร การพยากรณ์ข้อมูลจะพิจารณาจากรูปแบบหลัก 2 รูปแบบ ได้แก่ Autoregressive (AR) คือ รูปแบบที่ กำหนดว่าค่าพยากรณ์ที่เวลาใด ๆ ขึ้นอยู่กับค่าสังเกตก่อนหน้า และรูปแบบ Moving average (MA) คือ รูปแบบที่กำหนดว่า ค่าพยากรณ์ที่เวลาใด ๆ ขึ้นอยู่กับความคลาดเคลื่อนก่อนหน้า เงื่อนไขที่สำคัญของการใช้ตัวแบบบอกรี-เจนกินส์ คือ อนุกรม เวลาที่นำมาพยากรณ์จะต้องมีสถานะคงที่ (Stationary)แต่โดยทั่วไปอนุกรมเวลาที่เก็บรวบรวมมาจะมีการเคลื่อนไหว เนื่องมาจากแนวโน้มและฤดูกาล ซึ่งเป็นลักษณะของอนุกรมเวลาที่ไมคงที่ (Non - Stationary)โดยมีตัวแบบทั่วไปคือ SARIMA(p,d,q)(P,D,Q)s (Box *et al.*, 2008) แสดงดังสมการที่ (1) – (6)

$$\phi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^Dz_t = \delta + \theta_q(B)\Theta_q(B^s)\varepsilon_t \tag{1}$$

$$\delta = \mu\phi_p(B)\Phi_p(B^s) \tag{2}$$

$$\phi_p(B) = 1 - \phi_1B - \phi_2B^2 - \dots - \phi_pB^p \tag{3}$$

$$\Phi_p(B^s) = 1 - \Phi_1B^s - \Phi_2B^{2s} - \dots - \Phi_pB^{ps} \tag{4}$$

$$\theta_q(B) = 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q \tag{5}$$

$$\Theta_q(B^s) = 1 - \Theta_1B^s - \Theta_2B^{2s} - \dots - \Theta_qB^{qs} \tag{6}$$

เมื่อ  $Z_t$  แทน อนุกรมเวลา ณ เวลา  $t$  และ  $\varepsilon_t$  คือ ความคลาดเคลื่อน ซึ่งมีการแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับศูนย์มีความแปรปรวนคงที่และเป็นอิสระกัน ( $\varepsilon_t \sim NID(0, \sigma^2)$ ) และ  $t$  แทน คาบเวลาซึ่งมีค่าตั้งแต่ 1 ถึง  $n$ ,  $s$  แทน ความยาวของคาบฤดูกาลใน 1 ฤดูกาล,  $d$  และ  $D$  แทน ลำดับที่ของการหาผลต่างภายในฤดูกาลและผลต่างระหว่างฤดูกาล,  $B$  แทน ตัวดำเนินการถอยหลัง (Backward operator) โดยที่  $B^s Z_t = Z_{t-s}$  และ  $BZ_t = Z_{t-1}$ ,  $\phi_p(B)$  แทน ตัวดำเนินการสหสัมพันธ์ในตนเองแบบไม่มีฤดูกาลอันดับที่  $p$ ,  $\Phi_p(B^s)$  แทน ตัวดำเนินการสหสัมพันธ์ในตัวเองแบบมีฤดูกาลอันดับที่  $P$ ,  $\theta_q(B)$  แทน ตัวดำเนินการเฉลี่ยเคลื่อนที่แบบภายในฤดูกาลอันดับที่  $q$ ,  $\Theta_q(B^s)$  แทน ตัวดำเนินการเฉลี่ยเคลื่อนที่ระหว่างฤดูกาลอันดับที่  $Q$  การสร้างตัวแบบ SARIMA ต้องวิเคราะห์อนุกรมเวลาเพื่อกำหนดรูปแบบ SARIMA  $(p,d,q)(P,D,Q)_s$  การดำเนินการผู้วิจัยเขียนคำสั่งโดยใช้โปรแกรมภาษาอาร์มีฟังก์ชัน `auto.arima()` ซึ่งอยู่ใน Package “forecast” (R Core Team, 2019) ที่ใช้ในการระบุรูปแบบ ARIMA  $(p,d,q)(P,D,Q)_s$  ที่เหมาะสม ขั้นตอนการสร้างตัวแบบ คือ 1) พิจารณาอนุกรมเวลาว่าคงที่หรือไม่ โดยพิจารณาจากกราฟของอนุกรมเวลาชุดนั้น หากกราฟอนุกรมเวลาชุดนั้นมีแนวโน้มเป็นส่วนประกอบ กล่าวคือ ข้อมูลลักษณะเพิ่มขึ้นหรือลดลงอย่างเป็นสัดส่วนกับเวลาและมีการกระจายไม่สม่ำเสมอมีลักษณะบานออกเป็นปากแตร แสดงว่า ข้อมูลอนุกรมเวลาชุดนั้นมีสถานะไม่คงที่นั่นคือมีค่าเฉลี่ยไม่คงที่และความแปรปรวนไม่คงที่งานวิจัยนี้จะทำการทดสอบด้วย Augmented Dickey – Fuller (ADF) หากผลการทดสอบไม่มีนัยสำคัญทางสถิติแสดงว่าข้อมูลอนุกรมเวลาชุดนั้นมีลักษณะไม่คงที่ เมื่อทราบว่าคุณข้อมูลมีลักษณะไม่คงที่จะต้องทำการแปลงข้อมูลให้มีลักษณะคงที่ โดยการหาลอการิทึมของข้อมูลก่อนและจึงหาผลต่างของข้อมูลชุดดังกล่าว เพื่อให้ข้อมูลอนุกรมเวลามีลักษณะคงที่ก่อนที่จะนำไปกำหนดตัวแบบ 2) กำหนดตัวแบบพยากรณ์ เนื่องจากข้อมูลอนุกรมเวลาปริมาณการใช้น้ำประปาของประชาชน มีแนวโน้มและฤดูกาลเป็นส่วนประกอบ ดังนั้นตัวแบบที่เหมาะสมในการพยากรณ์ คือ ตัวแบบ SARIMA การใช้ตัวแบบจะต้องระบุอันดับ  $(p,d,q)$  และอันดับ  $(P,D,Q)_s$  ในงานวิจัยนี้ผู้วิจัยใช้โปรแกรมภาษาอาร์มีฟังก์ชัน `auto.arima()` ซึ่งอยู่ใน Package “forecast” สำหรับกำหนดอันดับ  $(p,d,q)$  และอันดับ  $(P,D,Q)_s$  ที่เหมาะสมต่อไป 3) ประมาณค่าพารามิเตอร์และตรวจสอบความเหมาะสมของตัวแบบ เมื่อได้ตัวแบบที่เหมาะสมก่อนที่จะนำตัวแบบไปใช้ในการพยากรณ์จะต้องมีการตรวจสอบตัวแบบ (Diagnostics Checking) ซึ่งมีทั้งหมด 4 ข้อ คือ 3.1) ทดสอบความคลาดเคลื่อนมีค่าเฉลี่ยเท่ากับศูนย์ โดยใช้สถิติทดสอบ  $t$ -test 3.2) ทดสอบความคลาดเคลื่อนมีการแจกแจงแบบปกติโดยใช้สถิติทดสอบ Kolmogorov – Smirnov 3.3) ทดสอบความเป็นอิสระกันของความคลาดเคลื่อน โดยใช้สถิติทดสอบ Ljung – Box Q-statistics 3.4) ความคลาดเคลื่อนมีความแปรปรวนคงที่โดยพิจารณาจากการพล็อตกราฟดูลักษณะการกระจายระหว่างความคลาดเคลื่อนกับค่าพยากรณ์หากกราฟมีการกระจายไม่มีรูปแบบแสดงว่า ความคลาดเคลื่อนมีความแปรปรวนคงที่ (ดังภาพที่ 6) 4) นำตัวแบบที่เหมาะสมไปใช้ในการพยากรณ์ปริมาณการใช้น้ำประปาล่วงหน้าต่อไป

## 2.2 ซัพพอร์ตเวกเตอร์เรเกรสชัน (Support vector regression model: SVR)

ซัพพอร์ตเวกเตอร์เรเกรสชัน พัฒนาขึ้นในปี ค.ศ. 1995 โดย Vapnik เป็นตัวแบบที่มีการใช้กระบวนการลดความเสี่ยงเชิงโครงสร้างให้มีค่าต่ำที่สุด ด้วยกระบวนการปรับรูปแบบของข้อมูลที่มีมิติต่ำ (Low dimension data set) บนพื้นที่ข้อมูลนำเข้า (Input space) ให้อยู่ในรูปแบบของข้อมูลที่มีมิติสูง (High dimension data set) บนพื้นที่ข้อมูลคุณลักษณะ (Feature space) โดยใช้ฟังก์ชันในการปรับรูปแบบข้อมูลเรียกว่า ฟังก์ชันเคอร์เนล (kernel function) เป็นวิธีการพยากรณ์ที่มีความ

ถูกต้องสูง และมีการประยุกต์ใช้ในการพยากรณ์อย่างแพร่หลาย (Chen *et al.*, 2017; Mohammadi *et al.*, 2015; Xu & Zhang, 2019; Yaslan & Bican, 2017) โดยมีฟังก์ชันการประมาณค่า ดังสมการที่ (7)

$$f(x) = w^T \phi(x) + b \tag{7}$$

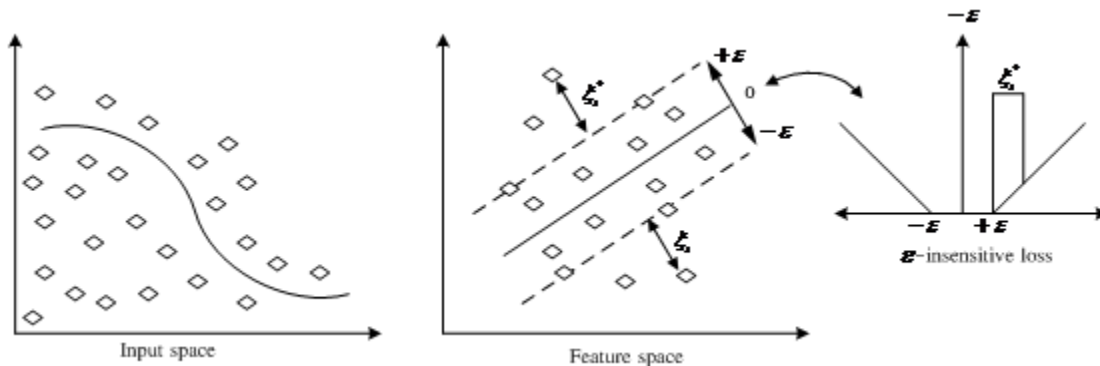
เมื่อ  $w$  คือ น้ำหนักของเวกเตอร์ และ  $b$  คือ ความเอนเอียง (bias) การกำหนดค่าของ  $w$  และ  $b$  ทำได้โดยการหาค่าต่ำสุดจากสมการที่ 8

$$R(C) = \frac{1}{2} \|w\|^2 + C \cdot \frac{1}{n} \sum_{i=1}^N |y_i - f(x)|_\epsilon \tag{8}$$

การใช้เทคนิค SVR ทำนายค่าเอาต์พุตจากอินพุตเวกเตอร์ จะมีการสร้างท่อเอปซิลอน (Epsilon tube) โดยใช้ฟังก์ชันสูญเสีย (Loss function) ในหลายรูปแบบ เช่น Quadratic, Laplacian, Huber และ  $\epsilon$ -insensitive แต่ที่นิยมใช้คือ  $\epsilon$ -insensitive ที่นำเสนอโดย Vapnik (1998) เนื่องจากเป็นฟังก์ชันสูญเสียที่นิยมใช้ในงานวิจัยอย่างแพร่หลาย เช่น งานวิจัยของ Yashlan & Bican (2017) ได้ประยุกต์ใช้ฟังก์ชันสูญเสียแบบ  $\epsilon$ -insensitive เพื่อใช้ในการพยากรณ์การใช้พลังงานไฟฟ้าของประเทศไทย เป็นต้น ฟังก์ชันสูญเสียแบบ  $\epsilon$ -insensitive สามารถแสดงได้ดังสมการที่ (9)

$$|y_i - f(x)|_\epsilon = \begin{cases} 0 & \text{if } |y_i - f(x)| \leq \epsilon \\ |y_i - f(x)|_\epsilon & \text{otherwise} \end{cases} \tag{9}$$

กำหนดให้  $\zeta_i$  และ  $\zeta_i'$  แทน ตัวแปรที่ทำให้เกิดความยืดหยุ่น (Slack variable) และอยู่ห่างจากระยะขอบของเอปซิลอน ( $\epsilon$ -tube) ดังภาพที่ 1



**ภาพที่ 1** รูปแบบแสดงการทำงานของวิธีซัพพอร์ตเวกเตอร์เรกเรชันโดยใช้ฟังก์ชันการสูญเสียแบบ  $\epsilon$ -insensitive ที่มา : Yaslan & Bican, 2017 : 54

การเพิ่มตัวแปร  $\xi_i$  และ  $\xi_i^*$  ซึ่งเป็นการวัดส่วนเบี่ยงเบนทั้งหมดของชุดฝึกสอนที่อยู่นอกขอบระนาบหรือคลาดเคลื่อนของข้อมูลจากขอบระนาบบนและล่าง ดังนั้นการเพิ่ม  $\xi_i$  และ  $\xi_i^*$  จะทำให้สมการที่ 8 ถูกแปลงให้อยู่ในเงื่อนไขบังคับ (Constrained) ดังสมการที่ (11)

$$\text{Minimize} = \frac{1}{2} \|w\|^2 + c \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (10)$$

$$\text{Subject to} \begin{cases} y_i - w\phi_i - b \leq \xi_i + \varepsilon \\ -y_i + w\phi_i + b \leq \xi_i^* + \varepsilon \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (11)$$

การแก้ปัญหของสมการที่ (10) โดยมีเงื่อนไขตามสมการที่ (11) ปัญหานี้สามารถแก้ได้โดยแปลงเป็นปัญหาควมคู่ (Dual Problem) ด้วยตัวคูณลากรองจ์ (Lagrange multipliers) แสดงดังสมการที่ (12) และ เงื่อนไขข้อบังคับดังสมการที่ (13)

$$\begin{aligned} \text{Maximize } (\alpha_i, \alpha_i^*) &= \sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) \\ &\quad - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k(x_i, x_j) \end{aligned} \quad (12)$$

$$\text{Subject to} \begin{cases} \sum_{i=1}^N (\alpha_i + \alpha_i^*) = 0 \\ 0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \\ 0 \leq \alpha_i^* \leq C, \quad i = 1, \dots, N \end{cases} \quad (13)$$

เมื่อ  $x_i, x_j$  คือ ข้อมูลอินพุต  $\alpha_i, \alpha_i^*$  คือ ตัวคูณลากรองจ์  $C$  คือ ค่าคงที่  $N$  คือ จำนวนซัพพอร์ตเวกเตอร์ซึ่งมีอินพุตเป็นซัพพอร์ตเวกเตอร์ จะมี  $\alpha_i, \alpha_i^* > 0$  ส่วนอินพุตเวกเตอร์ที่ไม่ใช่ซัพพอร์ตเวกเตอร์ จะมี  $\alpha_i, \alpha_i^* = 0$  และหลังจากที่คำนวณค่า  $\alpha_i$  และ  $\alpha_i^*$  จากชุดข้อมูลชุดที่ 1 ดังนั้นเราสามารถสร้างสมการ SVR เพื่อใช้ทำนายค่าเอาต์พุตจากอินพุตเวกเตอร์ได้ดังสมการที่ (14)

$$f(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (14)$$

เมื่อ  $\alpha_i$  และ  $\alpha_i^*$  คือ ตัวคูณลากรางจ์,  $k(x_i, x)$  คือ เคอร์เนลฟังก์ชันที่ใช้ในการแปลงข้อมูลให้มีมิติที่สูงขึ้น (High dimensional feature space) โดยเคอร์เนลฟังก์ชันที่นิยมใช้ในวิธีซัพพอร์ตเวกเตอร์แมชชีน เช่น 1) ลินีเยอร์เคอร์เนล (Linear Kernel) :  $k(x_i, x) = x_i^T x$  2) โพลีโนเมียลเคอร์เนล (Polynomial kernel):  $k(x_i, x) = (1 + x_i \cdot x)^d$  3) เกาส์เซียนเคอร์เนล (Gaussian(RBF) kernel) :  $k(x_i, x) = \exp(-\frac{\|x_i - x\|^2}{2\sigma^2})$  และ  $b$  คือ ความเอนเอียง (bias)

### 2.3 ตัวแบบผสม (Hybrid models)

จากวิธีการพยากรณ์โดยใช้ตัวแบบ SARIMA และตัวแบบ SVR ข้างต้นจะเห็นได้ว่าเป็นวิธีการพยากรณ์ที่เหมาะสมกับการพยากรณ์ข้อมูลอนุกรมเวลาต่างกันกล่าวคือ ตัวแบบ SARIMA เหมาะสำหรับการวิเคราะห์อนุกรมเวลาที่เป็นเชิงเส้นตรง ส่วนตัวแบบ SVR เป็นตัวแบบที่ใช้ได้ดีกับข้อมูลอนุกรมเวลาที่ไม่เป็นเชิงเส้นตรง ดังนั้นหากผู้พยากรณ์เลือกใช้วิธีการพยากรณ์วิธีใดวิธีหนึ่งก็จะทำให้ค่าพยากรณ์ที่ได้ขาดความแม่นยำ จากปัญหาดังกล่าว Zhang (2003) จึงได้เสนอตัวแบบผสมขึ้น โดยรวมค่าพยากรณ์แบบเป็นเชิงเส้นตรงและไม่เป็นเชิงเส้นตรงเข้าด้วยกัน ค่าพยากรณ์ที่ได้จะมีความแม่นยำมากขึ้น สมการพยากรณ์ดังปรากฏในสมการที่ (15)

$$y_t = L_t + N_t + \varepsilon_t \quad (15)$$

เมื่อ  $y_t$  แทน ค่าสังเกตของข้อมูลอนุกรมเวลา ณ เวลา  $t$ ,  $L_t$  แทน ข้อมูลในส่วนที่เป็นเชิงเส้นตรงในเทอมของค่าพารามิเตอร์ ณ เวลา  $t$ ,  $N_t$  แทน ข้อมูลส่วนที่ไม่เป็นเชิงเส้นตรงในเทอมของค่าพารามิเตอร์ ณ เวลา  $t$ ,  $\varepsilon_t$  แทน ค่าความคลาดเคลื่อน ณ เวลา  $t$  โดยคำนวณจากตัวแบบ SARIMA ดังสมการที่ (16)

$$\varepsilon_t = y_t - \hat{L}_t \quad (16)$$

โดยที่  $\hat{L}_t$  แทน ค่าพยากรณ์ข้อมูลส่วนที่เป็นฟังก์ชันเชิงเส้นตรงที่ได้จากตัวแบบ SARIMA ณ เวลา  $t$ , เมื่อได้ส่วนเหลือจากสมการที่ (16) จะนำไปพยากรณ์ข้อมูลโดยใช้วิธีการ SVR ซึ่งในการวิจัยครั้งนี้ได้เสนอวิธีการพยากรณ์รวมทั้งหมด 5 ตัวแบบ คือ

$$\text{ตัวแบบที่ 1: SARIMASVR1 } \varepsilon_t = f(\varepsilon_{t-1}, \varepsilon_{t-12}) + e_t \quad (17)$$

เมื่อ  $f$  คือ ฟังก์ชันไม่เป็นเชิงเส้นตรงซึ่งหาได้จากตัวแบบ SVR และ  $e_t$  คือ ค่าความคลาดเคลื่อนสุ่มซึ่งหาได้จากการนำเอาค่าจริงลบค่าพยากรณ์ ดังนั้น

$$\text{ค่าพยากรณ์รวมของตัวแบบที่ 1 คือ } \hat{y}_t = \hat{L}_t + \hat{N}_t \quad (18)$$



เมื่อ  $\hat{y}_t$  คือ ค่าพยากรณ์ที่ได้จากตัวแบบซัพพอร์ตเวกเตอร์เรกเรชัน,  $\hat{y}_t$  คือค่าพยากรณ์ ณ เวลา  $t$ ,  $y_t$  แทน ค่าสังเกตของข้อมูลอนุกรมเวลา ณ เวลา  $t$ ,  $y_{t-1}, \dots, y_{t-12}$  แทน ค่าสังเกตค่าหลังไปตั้งแต่ 1 ถึง 12 คาบเวลา และ  $\hat{y}_{t-1}, \dots, \hat{y}_{t-12}$  คือ ค่าพยากรณ์ที่ได้จากตัวแบบ SARIMA ณ เวลา  $t-1$  ถึง  $t-12$

$$\text{ตัวแบบที่ 2: SARIMASVR2 } \hat{y}_t = f(y_{t-1}, y_{t-12}, \varepsilon_t) + e_t \quad (19)$$

$$\text{ตัวแบบที่ 3: SARIMASVR3 } \hat{y}_t = f(y_{t-1}, y_{t-12}, \hat{L}_t) + e_t \quad (20)$$

$$\text{ตัวแบบที่ 4: SARIMASVR4 } \hat{y}_t = f(y_{t-1}, y_{t-12}, \hat{L}_{t-1}, \hat{L}_{t-12}) + e_t \quad (21)$$

$$\text{ตัวแบบที่ 5: SARIMASVR5 } \hat{y}_t = f(y_{t-1}, y_{t-12}, \hat{L}_{t-1}, \hat{L}_{t-12}, \varepsilon_t) + e_t \quad (22)$$

#### 2.4 การวัดความคลาดเคลื่อนของการพยากรณ์

##### (1) ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย

$$MAE = \sum_{t=1}^n |y_t - \hat{y}_t| / n \quad (23)$$

##### (2) ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย

$$MAPE = 100 \sum_{t=1}^n |1 - \hat{y}_t / y_t| / n \quad (24)$$

โดยที่ค่า MAE และ MAPE ยิ่งน้อยก็ยิ่งแสดงว่าการพยากรณ์มีความแม่นยำสูง

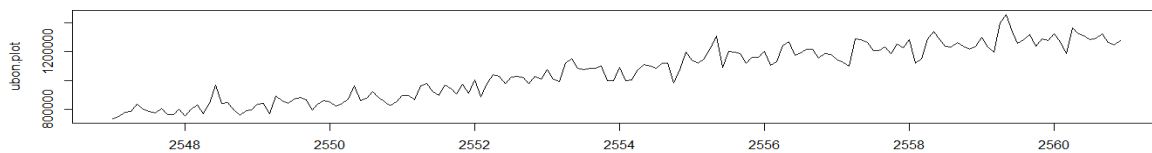
#### 2.5 การเลือกตัวแบบที่เหมาะสม

จากการเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์ทั้ง 6 วิธี เมื่อทราบว่าวิธีใดมีค่า MAE, MAPE ต่ำสุดจะถือว่าเป็นวิธีการพยากรณ์ที่มีความแม่นยำที่สุด จากนั้นจะนำวิธีการพยากรณ์นั้นไปพยากรณ์ปริมาณความต้องการใช้น้ำประปาของประชาชน ทั้ง 4 สาขา จากนั้นทำการเปรียบเทียบข้อมูลชุดที่ 2 ข้อมูลชุดทดสอบโดยพิจารณาความแม่นยำของการพยากรณ์ จากค่า MAPE เพื่อเป็นการตรวจสอบความแม่นยำของตัวแบบอีกครั้งหนึ่ง ดังตารางที่ 8 ตามลำดับ

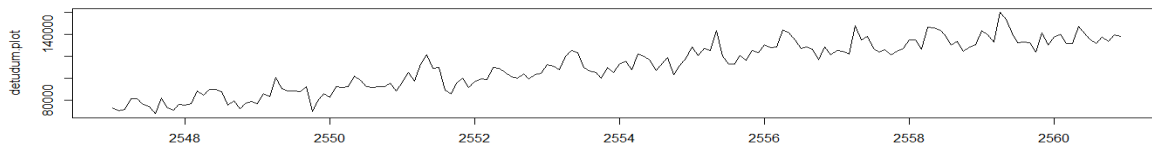
**ผลการวิจัย**

1. ผลการศึกษาความเคลื่อนไหวของข้อมูลอนุกรมเวลา

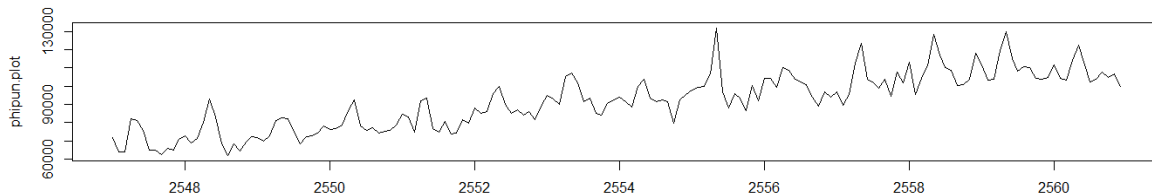
จากการพิจารณาลักษณะการเคลื่อนไหวของอนุกรมเวลาปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานี 4 สาขา ได้แก่ สาขาอุบลราชธานี สาขาเดชอุดม สาขาพิบูลมังสาหาร และสาขาเขมราฐ (หน่วย: ลบ.ม./เดือน) ย้อนหลัง 14 ปี คือตั้งแต่เดือนมกราคม พ.ศ. 2547 ถึงเดือนธันวาคม พ.ศ. 2560 จำนวน 168 ค่า ดังภาพที่ 2 พบว่า อนุกรมเวลาทั้ง 4 ชุด มีแนวโน้มและความผันแปรตามฤดูกาลเป็นส่วนประกอบ กล่าวคือ ปริมาณการใช้น้ำประปามีแนวเพิ่มเมื่อเวลาเปลี่ยนแปลงไป



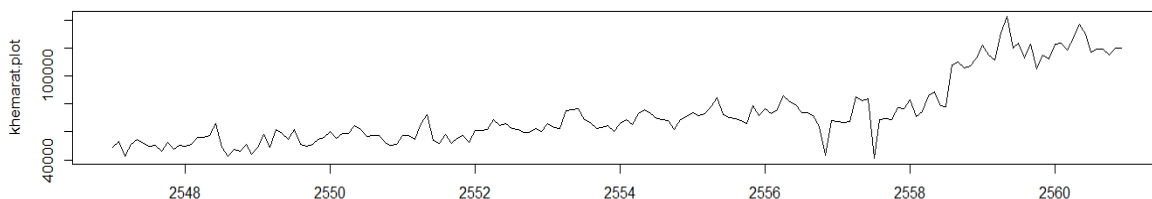
สาขาอุบลราชธานี



สาขาเดชอุดม



สาขาพิบูลมังสาหาร



สาขาเขมราฐ

**ภาพที่ 2** ลักษณะการเคลื่อนไหวของอนุกรมเวลาปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานี 4 สาขา ตั้งแต่เดือนมกราคม พ.ศ. 2547 ถึง เดือนธันวาคม พ.ศ. 2560

## 2. ผลการวิเคราะห์ข้อมูลค่าสถิติพื้นฐาน

ผลการวิเคราะห์สถิติพื้นฐานปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานี 4 สาขา พบว่า สาขาอุบลราชธานี มีปริมาณน้ำจำหน่ายให้กับประชาชนโดยเฉลี่ยต่อเดือนเท่ากับ 1,055,467.70 ลบ.ม.ต่อเดือน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 183,812.51 ลบ.ม.ต่อเดือน ปริมาณน้ำจำหน่ายสูงสุดเท่ากับ 1,457,410 ลบ.ม.ต่อเดือน และต่ำสุด 730,820 ลบ.ม.ต่อเดือน สาขาเดชอุดม มีปริมาณน้ำจำหน่ายให้กับประชาชนโดยเฉลี่ยต่อเดือนเท่ากับ 110,963.33 ลบ.ม.ต่อเดือน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 22,054.67 ลบ.ม.ต่อเดือน ปริมาณน้ำจำหน่ายสูงสุดเท่ากับ 159,471 ลบ.ม.ต่อเดือน ต่ำสุด 67,869 ลบ.ม.ต่อเดือน สาขาพิบูลมังสาหาร มีปริมาณน้ำจำหน่ายให้กับประชาชนโดยเฉลี่ยต่อเดือนเท่ากับ 91,576.67 ลบ.ม.ต่อเดือน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 15,456.68 ลบ.ม.ต่อเดือน ปริมาณน้ำจำหน่ายสูงสุดเท่ากับ 131,931 ลบ.ม.ต่อเดือน ต่ำสุด 61,747 ลบ.ม.ต่อเดือน สาขาเขมราฐ มีปริมาณน้ำจำหน่ายให้กับประชาชนโดยเฉลี่ยต่อเดือนเท่ากับ 72,865.10 ลบ.ม.ต่อเดือน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 23,570.19 ลบ.ม.ต่อเดือน ปริมาณน้ำจำหน่ายสูงสุดเท่ากับ 142,369 ลบ.ม.ต่อชั่วโมง ต่ำสุด 41,246 ลบ.ม.ต่อเดือน รายละเอียดดังตารางที่ 1

**ตารางที่ 1** ค่าสถิติพื้นฐานปริมาณน้ำจำหน่ายให้กับประชาชนในเขตพื้นที่รับผิดชอบ ทั้ง 4 สาขา (หน่วย: ลบ.ม.ต่อเดือน)

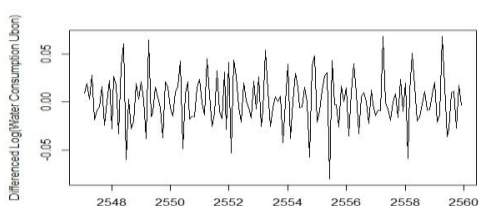
สาขา	ข้อมูล	จำนวน	ค่าเฉลี่ย	ค่าสูงสุด	ค่ามัธยฐาน	ค่าต่ำสุด	ส่วนเบี่ยงเบนมาตรฐาน
อุบลราชธานี	All sample	168	1,055,467.70	1,457,410	1,084,791	730,820	183,812.51
	Training	156	1,037,540.13	1,457,410	1,035,040.5	730,820	178,116.49
	Testing	12	1,288,526.08	1,364,064	1,288,258	1,188,811	45,778.77
เดชอุดม	All sample	168	110,963.33	159,471	112,908	67,869	22,054.67
	Training	156	108,978.94	159,471	109,570	67,869	21,611.89
	Testing	12	136,760.33	146,788	137,127	131,172	4,608.45
พิบูลมังสาหาร	All sample	168	91,576.60	131,931	92,487	61,747	15,456.68
	Training	156	90,328.57	131,931	91,533.50	61,747	15,252.92
	Testing	12	107,801	122,510	105,941.50	100,072	6,288.76
เขมราฐ	All sample	168	72,865.10	142,369	66,425	41,246	23,570.19
	Training	156	69,047.61	142,369	65,572.50	41,246	19,762.13
	Testing	12	122,492.42	136,934	120,011	115,193	6,146.90

3. ผลพยากรณ์โดยใช้วิธีบ็อกซ์ – เจนกินส์

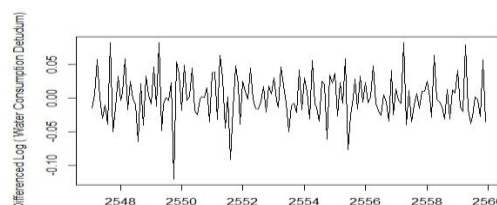
3.1 จากการพิจารณาข้อมูลปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานี 4 สาขา พบว่า ข้อมูลมีลักษณะไม่คงที่โดยการพิจารณาจากกราฟ (ภาพที่ 2) และผลการทดสอบความคงที่ของข้อมูลด้วย Augmented Dickey – Fuller (ADF) (ตารางที่ 2) เมื่อทราบว่าข้อมูลอนุกรมเวลาทั้ง 4 ชุด มีลักษณะไม่คงที่ ผู้วิจัยจึงทำการแปลงข้อมูลโดยการหาลอการิทึมก่อนแล้วจึงหาผลต่างของชุดข้อมูลก่อนนำไปสร้างตัวแบบ SARIMA ผลการวิเคราะห์ข้อมูล ดังปรากฏในตารางที่ 2 และภาพที่ 3 พบว่า หลังแปลงข้อมูลดังกล่าว ข้อมูลอนุกรมเวลาปริมาณการใช้น้ำของประชาชนทั้ง 4 สาขา มีลักษณะคงที่และจะนำไปใช้ในการกำหนดตัวแบบในลำดับต่อไป

ตารางที่ 2 ค่า ADF ปริมาณการใช้น้ำประปาของประชาชนทั้ง 4 สาขา

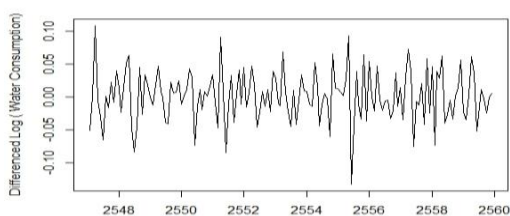
สาขา	ก่อนแปลงข้อมูล		หลังแปลงข้อมูล	
	ADF	p-value	ADF	p-value
อุบลราชธานี	0.5441	0.7877	-15.1418	0.01
เดชอุดม	0.3948	0.7402	-11.3815	0.01
พิบูลมังสาหาร	-0.0762	0.590	-12.3712	0.01
เขมราฐ	0.3943	0.74	-13.0023	0.01



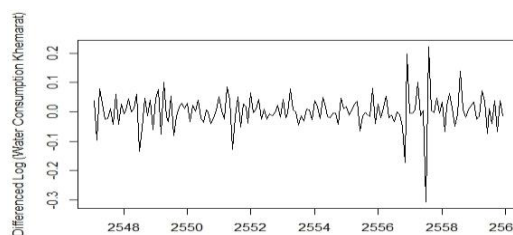
สาขาอุบลราชธานี



สาขาเดชอุดม



สาขาพิบูลมังสาหาร



สาขาเขมราฐ

ภาพที่ 3 ลักษณะการเคลื่อนไหวของอนุกรมเวลาปริมาณการใช้น้ำประปาของประชาชนที่มีลักษณะคงที่

3.2 กำหนดรูปแบบของตัวแบบ SARIMA ต้องระบุอันดับ (p,d,q) และอันดับ (P,D,Q)<sub>s</sub> ซึ่งโปรแกรมภาษาอาร์ มีฟังก์ชัน auto.arima() ซึ่งอยู่ใน Package “forecast” สำหรับกำหนดอันดับ (p,d,q) และอันดับ (P,D,Q)<sub>s</sub> ที่เหมาะสม ผลจากการใช้ฟังก์ชัน auto.arima() กำหนดรูปแบบของตัวแบบของ SARIMA ผลการวิเคราะห์ข้อมูลดังปรากฏตารางที่ 3

**ตารางที่ 3** การประมาณของค่าตัวแบบ SARIMA ปริมาณการใช้น้ำประปาของประชาชนทั้ง 4 สาขา

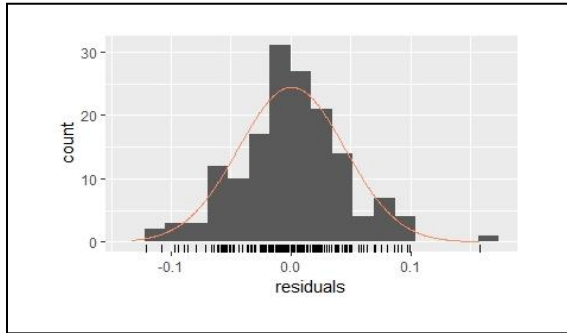
สาขา		Estimate	Std. Error	z-value	p-value	
อุบลราชธานี	AR(1)	0.1938	0.0945	2.0508	0.0402	
	MA(1)	-0.9251	0.0446	-20.7434	<0.0001	
	SARIMA(1,1,1)(0,0,2) <sub>12</sub>	SMA(1)	0.2368	0.0820	2.8874	0.0039
		SMA(2)	0.3153	0.0859	3.6694	0.0002
เดชอุดม	AR(1)	-0.7903	0.0980	-8.0636	<0.0001	
	MA(1)	0.1194	0.0551	2.1682	0.0301	
	SARIMA(1,1,2)(2,0,0) <sub>12</sub>	MA(2)	-0.8350	0.0618	-13.5082	<0.0001
		SAR(1)	0.3697	0.0500	7.3998	<0.0001
	SAR(2)	0.3849	0.0820	4.6939	<0.0001	
พิบูลมังสาหาร	AR(1)	-0.4325	0.0750	-5.7670	<0.0001	
	SARIMA(1,1,0)(2,0,0) <sub>12</sub>	SAR(1)	0.4247	0.0732	5.8018	<0.0001
		SAR(2)	0.3996	0.0788	5.0685	<0.0001
เขมราฐ	AR(1)	-0.5729	0.0758	-7.5553	<0.0001	
	SARIMA(0,1,1)(0,0,1) <sub>12</sub>	MA(1)	0.1895	0.0812	2.3334	0.0196

3.3 การตรวจสอบความเหมาะสมของตัวแบบ โดยพิจารณาตามข้อตกลงเบื้องต้นของตัวแบบ SARIMA ซึ่งมีทั้งหมด 4 ข้อ ได้แก่ 1) ความคลาดเคลื่อนมีค่าเฉลี่ยเท่ากับศูนย์ ผลการตรวจสอบ พบว่า ตัวแบบสาขาอุบลราชธานี มีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 0 (t-test = 0.1656, p-value = 0.8687) สาขาเดชอุดม มีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 0 (t-test = 1.2676, p-value = 0.2068) สาขาพิบูลมังสาหาร มีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 0 (t-test = 0.0288, p-value = 0.9771) สาขาเขมราฐ มีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 0 (t-test = 1.158, p-value = 0.2487) (ดังตารางที่ 4) 2) ความคลาดเคลื่อนมีการแจกแจงแบบปกติ ผลการตรวจสอบ พบว่า ความคลาดเคลื่อนมีการแจกแจงแบบปกติทั้ง 4 สาขา (สาขาอุบลราชธานี KS -value = 0.0498, p-value = 0.8335, สาขาเดชอุดม KS -value = 0.0386, p-value = 0.9742 สาขาพิบูลมังสาหาร KS -value = 0.05847, p-value = 0.2146, สาขาเขมราฐ KS -value = 0.0947, p-value = 0.122) (ดังตารางที่ 4 และภาพที่ 4) 3) ความคลาดเคลื่อนเป็นอิสระต่อกัน โดยใช้ Ljung – Box Q-statistics พบว่า ความคลาดเคลื่อนเป็นอิสระต่อกัน ทั้ง 4 สาขา (สาขาอุบลราชธานี Q-statistics = 23.938, p-value = 0.0909, สาขาเดชอุดม Q-statistics = 22.244, p-value = 0.1335,

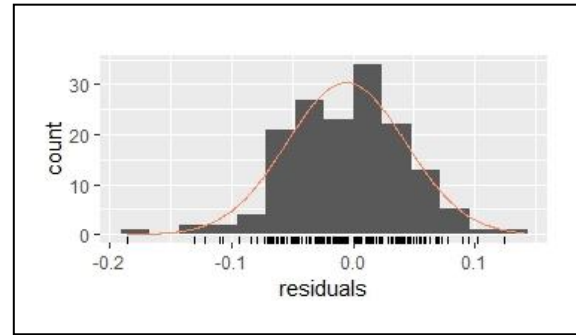
สาขาพิบูลมังสาหาร Q-statistics = 21.225, p-value = 0.1433 และสาขาเขมราฐ Q-statistics = 26.14, p-value = 0.0521) (ดังตารางที่ 5 และภาพที่ 5) และ 4) ความคลาดเคลื่อนมีค่าความแปรปรวนคงที่ (ดังภาพที่ 6) ซึ่งพบว่า ความคลาดเคลื่อนมีค่าความแปรปรวนคงที่ทั้ง 4 สาขา เนื่องจากความคลาดเคลื่อนมีการกระจายคงที่รอบค่าใดค่าหนึ่ง ดังนั้นสามารถสรุปได้ว่าตัวแบบที่ได้มีความเหมาะสมที่จะนำไปพยากรณ์

**ตารางที่ 4** ตรวจสอบตัวแบบ SARIMA ปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานีทั้ง 4 สาขา

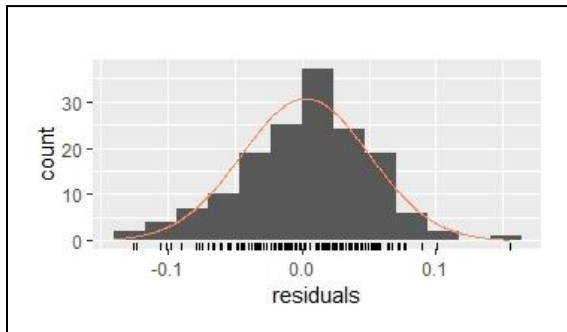
สาขา	ค่าความคลาดเคลื่อน มีค่าเฉลี่ยเท่ากับ 0		ค่าความคลาดเคลื่อน มีการแจกแจงแบบปกติ	
	t-value	p-value	KS -value	p-value
อุบลราชธานี SARIMA(1,1,1)(0,0,2) <sub>12</sub>	0.1656	0.8687	0.0498	0.8335
เดชอุดม SARIMA(1,1,2)(2,0,0) <sub>12</sub>	1.2676	0.2068	0.0386	0.9742
พิบูลมังสาหาร SARIMA(1,1,0)(2,0,0) <sub>12</sub>	0.0288	0.9771	0.0585	0.2146
เขมราฐ SARIMA(0,1,1)(0,0,1) <sub>12</sub>	1.1580	0.2487	0.0947	0.122



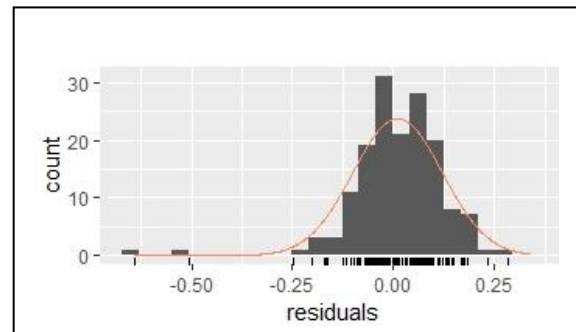
สาขาอุบลราชธานี



สาขาเดชอุดม



สาขาพิบูลมังสาหาร

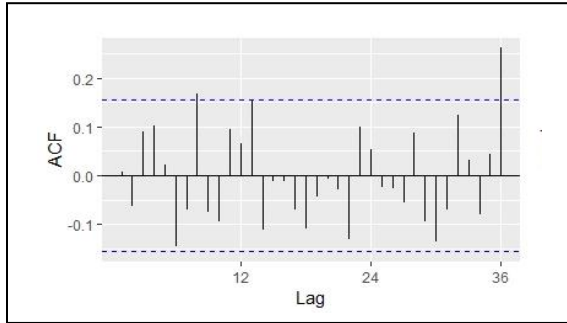


สาขาเขมราฐ

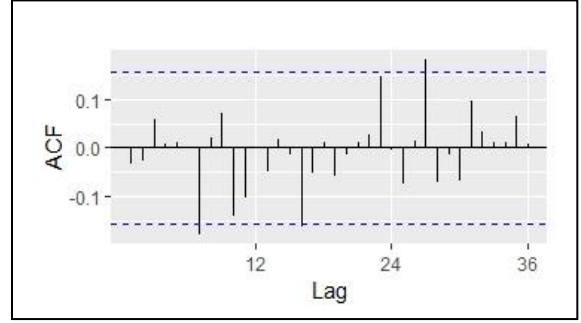
ภาพที่ 4 แผนภูมิแท่งแสดงการแจกแจงแบบปกติของความคลาดเคลื่อนปริมาณการใช้น้ำประปาของประชาชน  
ในเขตจังหวัดอุบลราชธานีทั้ง 4 สาขา

ตารางที่ 5 Ljung – Box Q-statistics ใช้ทดสอบ Autocorrelation ของความคลาดเคลื่อนของปริมาณการใช้น้ำประปา  
ของประชาชนในเขตจังหวัดอุบลราชธานีทั้ง 4 สาขา

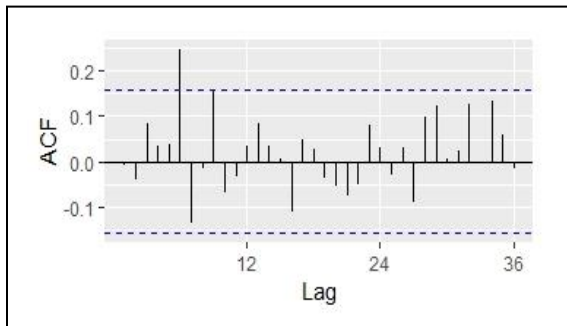
สาขา	Ljung – Box Q-statistics		
	Q-statistics	Degree of freedom	p-value
อุบลราชธานี	23.938	16	0.0909
เดชอุดม	22.244	16	0.1335
พิบูลมังสาหาร	21.225	16	0.1433
เขมราฐ	26.140	16	0.0521



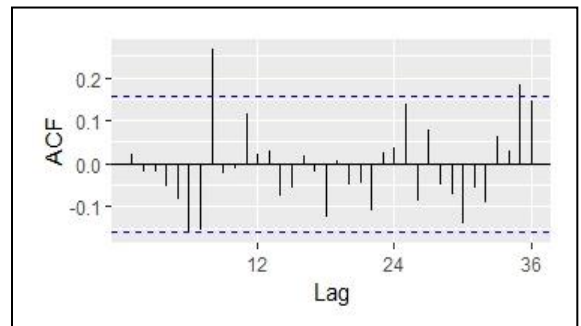
สาขาอุบลราชธานี



สาขาเดชอุดม

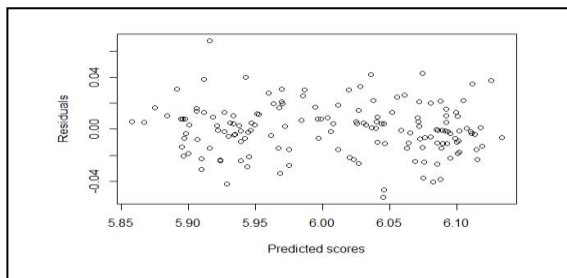


สาขาพิบูลมังสาหาร

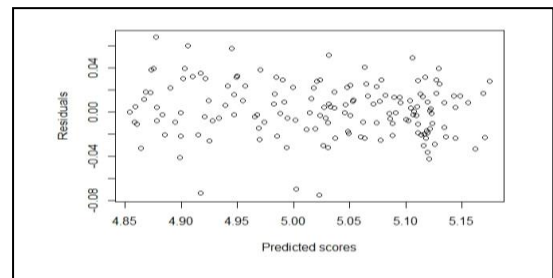


สาขาเขมราฐ

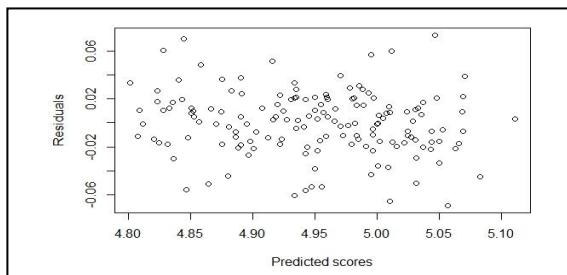
ภาพที่ 5 กราฟ Autocorrelation functions ความคลาดเคลื่อนของตัวแบบปริมาณการใช้น้ำประปาของประชาชน ในเขตจังหวัดอุบลราชธานี ทั้ง 4 สาขา



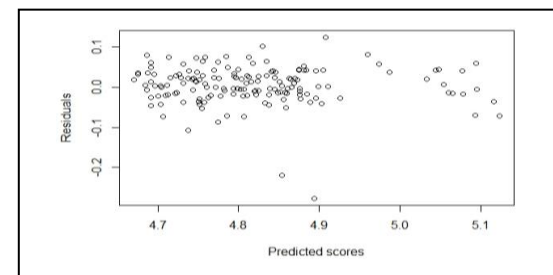
สาขาอุบลราชธานี



สาขาเดชอุดม



สาขาพิบูลมังสาหาร



สาขาเขมราฐ

ภาพที่ 6 แผนภาพการกระจายระหว่างความคลาดเคลื่อนกับค่าพยากรณ์



#### 4. ผลพยากรณ์โดยใช้ตัวแบบผสมระหว่างวิธีบ็อกซ์ – เจนกินส์และวิธีซัพพอร์ตเวกเตอร์รีเกรสชัน

การพยากรณ์ด้วยตัวแบบผสมระหว่างวิธีบ็อกซ์ – เจนกินส์และวิธีซัพพอร์ตเวกเตอร์รีเกรสชัน ซึ่งผู้วิจัยได้เสนอทั้งหมด 5 ตัวแบบ คือ SARIMASVR1, SARIMASVR2, SARIMASVR3, SARIMASVR4 และ SARIMASVR5 รายละเอียดดังปรากฏในขั้นตอนที่ 2.3 ในการสร้างตัวแบบผสมดังกล่าวผู้วิจัยดำเนินการเขียนโปรแกรมโดยใช้โปรแกรมภาษาอาร์ซึ่งมีฟังก์ชัน svm () ใน Package e1071 ที่เผยแพร่โดย Meyer *et al* (2015) ก่อนที่จะนำ SVR ไปใช้กับการพยากรณ์อนุกรมเวลาผู้วิจัยได้ปรับชุดข้อมูลให้อยู่ในรูปแบบของอินพุตเวกเตอร์และทาร์เก็ต ( $D = \{(x_i, y_i)\}_{i=1}^n$ ) ตามตัวแบบที่นำเสนอในขั้นตอนที่ 2.3 เนื่องจากอนุกรมเวลาที่ใช้ในการศึกษามีคาบเวลาเป็นรายเดือนและจำนวนคาบเวลาใน 1 ฤดูกาลเท่ากับ 12 คาบเวลาตามรอบปฏิทิน ดังนั้นผู้วิจัยจึงเลือกอินพุตเวกเตอร์ เป็นค่าสังเกตย้อนหลัง 12 คาบเวลาจากค่าพยากรณ์ และค่าความคลาดเคลื่อนที่ได้จากตัวแบบการพยากรณ์ SARIMA และปริมาณการใช้น้ำจริงของประชาชนซึ่งแต่ละตัวแบบก็必将มีความแตกต่างกันที่จำนวนอินพุตเวกเตอร์ตามลำดับ

การสร้างตัวแบบ SVR จะต้องมีการกำหนดเคอร์เนลฟังก์ชันและกำหนดค่าพารามิเตอร์ที่เหมาะสม งานวิจัยนี้ใช้ฟังก์ชัน tune() ค้นหาเคอร์เนลฟังก์ชันและกำหนดค่าพารามิเตอร์ที่เหมาะสม โดยฟังก์ชัน tune() จะค้นหาค่าที่เหมาะสมแบบกริดเสิร์จ (Grid search) ตลอดช่วงของพารามิเตอร์ที่กำหนดไว้ งานวิจัยนี้กำหนดค่าไว้ที่ 0 -1 และให้เพิ่มค่อย ๆ เพิ่มค่าทีละ 0.001 และกำหนดค่า Cost ไว้ที่ 1-16 และประเมินความเหมาะสมด้วยการวัดค่า MSE (Mean Square Error) ผลลัพธ์การค้นหาเคอร์เนลฟังก์ชันที่เหมาะสมจากฟังก์ชัน tune() พบว่า เคอร์เนลแบบเรเดียลเบซิส (Radial Basis Function: RBF) เป็นเคอร์เนลฟังก์ชันที่เหมาะสมในการนำไปใช้พยากรณ์ปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานี ทั้ง 4 สาขา การใช้เคอร์เนลแบบเรเดียลเบซิสจะต้องมีการกำหนดค่าพารามิเตอร์ Cost,  $\epsilon$  และ  $\gamma$  ผลจากการใช้ฟังก์ชัน tune() ได้ค่าที่เหมาะสมดังปรากฏในตารางที่ 6 เมื่อได้เคอร์เนลฟังก์ชันและค่าพารามิเตอร์ที่เหมาะสม ผู้วิจัยได้ดำเนินการนำค่าดังกล่าวไปใช้ในการพยากรณ์เพื่อเปรียบเทียบตัวแบบและเลือกตัวแบบที่เหมาะสมในการพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานีทั้ง 4 สาขาต่อไป

**ตารางที่ 6** ค่าพารามิเตอร์ที่เหมาะสมของตัวแบบผสม 5 ตัวแบบ

ตัวแบบ	สาขาอุบลราชธานี			สาขาเดชอุดม			สาขาพิบูลมังสาหาร			สาขาเขมราฐ		
	Cost	$\epsilon$	$\gamma$	Cost	$\epsilon$	$\gamma$	Cost	$\epsilon$	$\gamma$	Cost	$\epsilon$	$\gamma$
SARIMASVR1	1	0.098	0.083	1	0.647	0.083	1	0.541	0.083	1	0.771	0.083
SARIMASVR2	3	0.016	0.077	3	0.022	0.077	9	0.070	0.077	4	0.055	0.077
SARIMASVR3	2	0.184	0.077	1	0.310	0.077	2	0.187	0.077	1	0.072	0.077
SARIMASVR4	2	0.206	0.040	3	0.328	0.040	2	0.146	0.040	2	0.048	0.040
SARIMASVR5	2	0.020	0.038	11	0.030	0.038	7	0.024	0.038	3	0.008	0.038

### 5. การเลือกตัวแบบที่เหมาะสมในการพยากรณ์

เมื่อได้ผลการพยากรณ์จากตัวแบบทั้ง 6 ตัวแบบที่สร้างจากข้อมูลชุดที่ 1 ประกอบด้วยตัวแบบการพยากรณ์แบบเดี่ยว 1 ตัวแบบ คือ ตัวแบบ SARIMA และตัวแบบผสม 5 ตัวแบบคือตัวแบบ SARIMASVR1 ตัวแบบ SARIMASVR2 ตัวแบบ SARIMASVR3 ตัวแบบ SARIMASVR4 และตัวแบบ SARIMASVR5 เพื่อได้ตัวแบบที่เหมาะสมในการพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานีทั้ง 4 สาขาผู้วิจัยได้ทำการเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์ก่อนที่จะนำไปใช้ในการพยากรณ์ในอนาคต โดยใช้เกณฑ์พิจารณาจากค่า MAPE และค่า MAE ที่ต่ำที่สุด โดยจะพิจารณาค่า MAPE เป็นลำดับแรก หากมีค่าเท่ากันจึงจะพิจารณาที่ค่า MAE เป็นลำดับถัดไปและผลการเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์แสดงดังตารางที่ 7

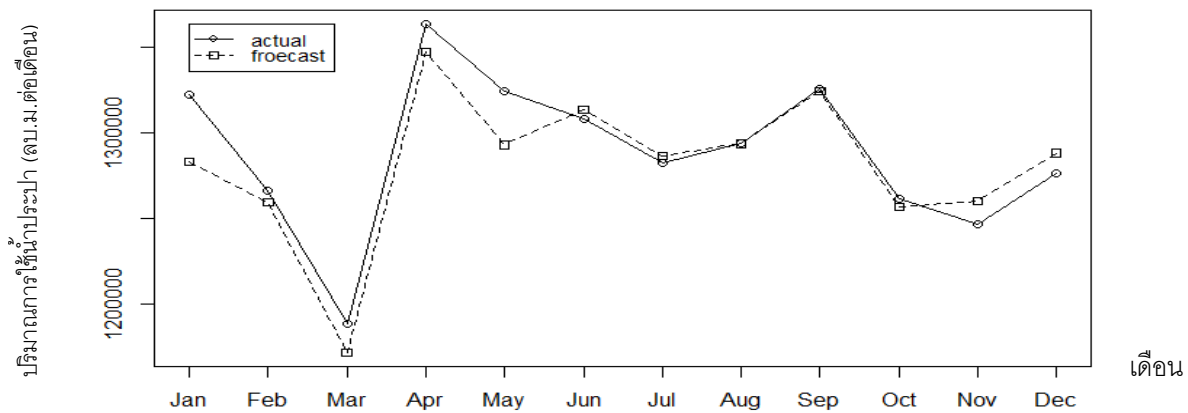
ตารางที่ 7 ค่าความคลาดเคลื่อนของการพยากรณ์ทั้ง 6 ตัวแบบ

ตัวแบบ	สาขาอุบลราชธานี		สาขาเดชอุดม		สาขาพิบูลมังสาหาร		สาขาเขมราฐ	
	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE
SARIMA	0.1624	0.0195	0.3769	0.0189	0.4130	0.0204	0.7009	0.0337
SARIMASVR1	0.2014	0.0123	0.2509	0.0129	0.3039	0.0153	0.3177	0.0161
SARIMASVR2	0.0959	0.0058	0.1818	0.0093	0.3672	0.0185	0.9507	0.0485
SARIMASVR3	0.1631	0.0099	0.1835	0.0094	0.3347	0.0168	1.0487	0.0535
SARIMASVR4	0.1474	0.0090	0.1941	0.0099	0.3425	0.0172	1.2906	0.0658
SARIMASVR5	0.0704	0.0043	0.1159	0.0059	0.2721	0.0137	1.3751	0.0701

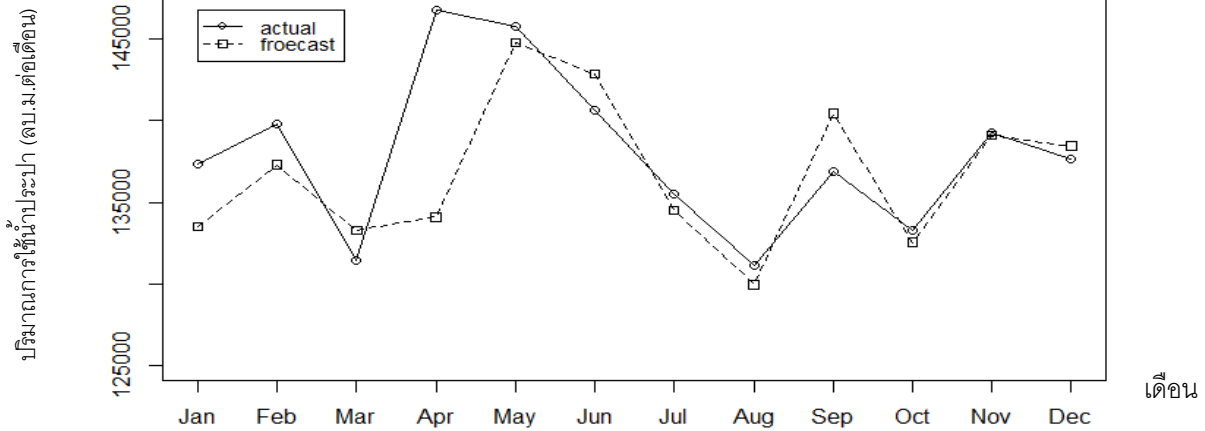
จากการพิจารณาผลการเปรียบเทียบจากตารางที่ 7 พบว่า ตัวแบบ SARIMASVR5 เป็นตัวแบบที่เหมาะสมในการพยากรณ์ปริมาณการใช้น้ำประปาของสาขาอุบลราชธานี สาขาเดชอุดม และสาขาพิบูลมังสาหาร และตัวแบบ SARIMASVR1 เป็นตัวแบบที่เหมาะสมในการพยากรณ์ปริมาณการใช้น้ำประปาของสาขาเขมราฐ เมื่อได้ตัวแบบที่เหมาะสมผู้วิจัยทำการพยากรณ์ล่วงหน้า 12 เดือนเพื่อเปรียบเทียบกับข้อมูลชุดที่ 2 ชุดทดสอบ เพื่อพิจารณาความแม่นยำของตัวแบบอีกครั้งหนึ่งผลปรากฏดังตารางที่ 8 และทำการพล็อตกราฟแสดงค่าจริงกับค่าพยากรณ์ดังภาพที่ 7 – 10

**ตารางที่ 8** การเปรียบเทียบค่าจริงและค่าพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนในเขตจังหวัดอุบลราชธานี  
จำนวน 12 เดือน

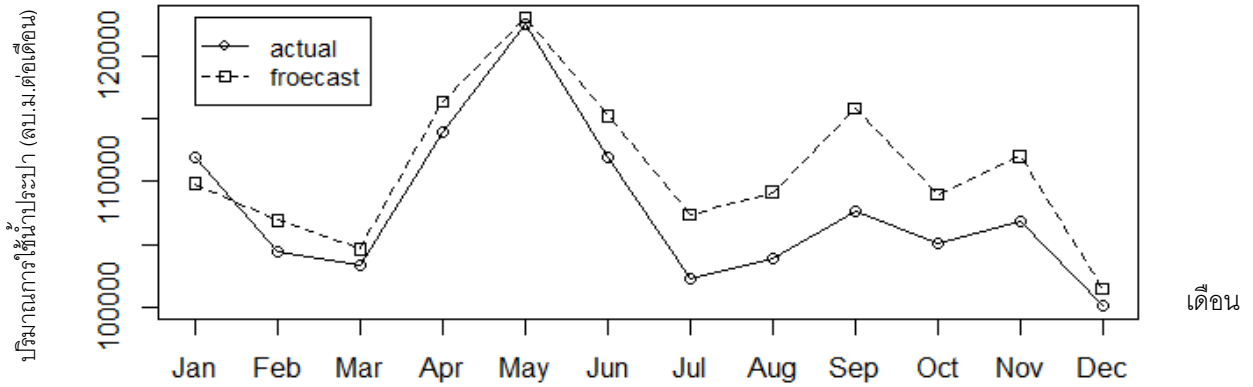
เดือน	ปริมาณการใช้น้ำประปาของประชาชน							
	สาขาอุบลราชธานี		สาขาเดชอุดม		สาขาพิบูลมังสาหาร		สาขาเขมราฐ	
	ค่าจริง (ลบ.ม./เดือน)	ค่าพยากรณ์ (ลบ.ม./เดือน)	ค่าจริง (ลบ.ม./เดือน)	ค่าพยากรณ์ (ลบ.ม./เดือน)	ค่าจริง (ลบ.ม./เดือน)	ค่าพยากรณ์ (ลบ.ม./เดือน)	ค่าจริง (ลบ.ม./เดือน)	ค่าพยากรณ์ (ลบ.ม./เดือน)
ม.ค. 2560	1,322,246	1,283,135	137,338	133,491.5	111,853	109834.7	122,891	120360.8
ก.พ. 2560	1,266,338	1,259,602	139,771	137,302.2	104,333	106884.1	123,556	122026.9
มี.ค. 2560	1,188,811	1,171,457	131,413	133,278.4	103,328	104604.8	118,626	126521.7
เม.ย. 2560	1,364,064	1,347,490	146,788	134,100	113,985	116358.8	126,787	130691.4
พ.ค. 2560	1,324,593	1,293,182	145,748	144,779.2	122,510	123060.6	136,934	137819.8
มิ.ย. 2560	1,308,383	1,313,839	140,613	142,849.7	111,911	115213.0	130,116	123813.4
ก.ค. 2560	1,282,595	1,286,549	135,530	134,515.0	102,262	107313.0	117,167	132900.0
ส.ค. 2560	1,293,921	1,293,816	131,172	129,966.7	103,795	109146.2	119,348	122820.0
ก.ย. 2560	1,326,154	1,324,599	136,916	140,430.9	107,680	115823.9	119,269	126344.2
ต.ค. 2560	1,261,939	1,256,710	133,296	132,547.7	105,021	108986.2	115,193	117770.7
พ.ย. 2560	1,246,705	1,260,179	139,225	139,094.6	106,862	112003.5	119,882	117134.2
ธ.ค. 2560	1,276,564	1,288,312	137,649	138,446.6	100,072	101462.2	120,140	119621.1
MAE	12725.583		2623.808		3426.300		4597.700	
MAPE	0.98351		1.86293		3.21682		3.81279	



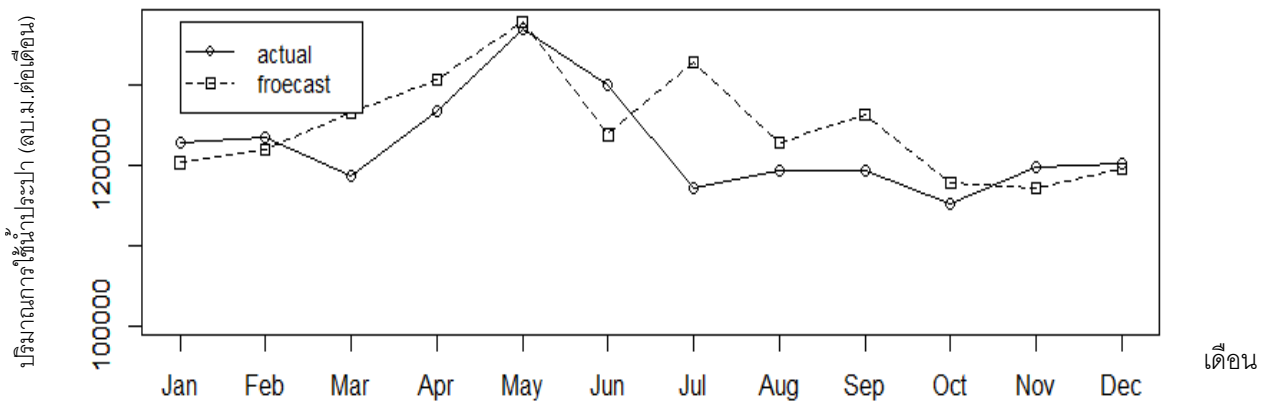
**ภาพที่ 7** ค่าจริงและค่าพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนสาขาอุบลราชธานี



ภาพที่ 8 ค่าจริงและค่าพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนสาขาเกษตร



ภาพที่ 9 ค่าจริงและค่าพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนสาขาปิโตรเคมี



ภาพที่ 10 ค่าจริงและค่าพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนสาขาบริการ

## วิจารณ์ผลการวิจัย

จากผลการวิจัยพบว่า ตัวแบบผสม SARIMASVR5 เป็นตัวแบบที่มีความเหมาะสมในการพยากรณ์ปริมาณการใช้น้ำประปาของประชาชน 3 สาขา คือ สาขาอุบลราชธานี สาขาพิบูลมังสาหาร สาขาเดชอุดม ส่วนตัวแบบ SARIMASVR1 เป็นตัวแบบที่มีความเหมาะสมในการพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนสาขาเขมราฐ เนื่องจากมีค่า MAPE ต่ำสุด คือ มีค่าเท่ากับ 0.98351, 1.86293, 3.21682 และ 3.81279 และมีค่า MAE เท่ากับ 12725.583, 2623.808, 3426.300 และ 4597.700 ตามลำดับ ผลการศึกษาสอดคล้องกับการศึกษาของ Yang and Wang (2018) ; Zhang (2003) ; Yashlan and Bican (2017) ; Cheng, Ding, Zhou, & Ding (2019); de Marcos, Bello, & Reneses (2019) ; de Oliveira & Cyrino Oliveira (2018) ทั้งนี้ เพราะตัวแบบผสมได้รวมค่าพยากรณ์ที่เป็นเส้นตรงและไม่เป็นเส้นตรงเข้าด้วยกันจึงให้ค่าพยากรณ์ที่มีความแม่นยำมากกว่า

## สรุปผลการวิจัย

จากการศึกษาเพื่อหาตัวแบบที่เหมาะสมเพื่อพยากรณ์ปริมาณการใช้น้ำประปาของประชาชนในจังหวัดอุบลราชธานี พบว่า ตัวแบบผสมจะให้ค่าพยากรณ์ที่มีความแม่นยำมากกว่าตัวแบบการพยากรณ์แบบเดี่ยว ทั้งนี้อาจเป็นเพราะว่าตัวแบบผสมได้รวมค่าพยากรณ์ที่เป็นเส้นตรงและไม่เป็นเส้นตรงเข้าด้วยกันจึงให้ค่าพยากรณ์ที่มีความแม่นยำมากกว่า ถึงแม้ค่าพยากรณ์ที่ได้จะมีความแม่นยำสูงแต่ข้อควรระวังในการนำตัวแบบไปพยากรณ์และนำค่าพยากรณ์ไปใช้ประโยชน์ ผู้ใช้ต้องพิจารณาปัจจัยต่าง ๆ ร่วมด้วย เช่น ข้อมูลประชากรผู้ใช้น้ำประปา สภาพแวดล้อมอื่นที่อาจจะมีผลกระทบต่อปริมาณการใช้น้ำประปาของประชาชนในแต่ละสาขา เป็นต้น ทั้งนี้เพราะว่าค่าพยากรณ์ที่ได้จากตัวแบบที่นำเสนอเป็นเพียงสารสนเทศหนึ่งที่ช่วยประกอบการตัดสินใจในการวางแผนการจัดการทรัพยากรน้ำให้เพียงพอต่อความต้องการของประชาชนที่มีแนวโน้มเพิ่มมากขึ้นในอนาคตเท่านั้น

## กิตติกรรมประกาศ

ขอขอบพระคุณสถาบันวิจัยและพัฒนา มหาวิทยาลัยราชภัฏอุบลราชธานีที่สนับสนุนทุนการวิจัยและขอขอบคุณรองศาสตราจารย์ ดร.เผ่าไทย วงศ์เหลาและรองศาสตราจารย์ ดร.จำลอง วงษ์ประเสริฐ อาจารย์ประจำสาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี ที่ให้คำปรึกษาให้แก่งคิดและสร้างแรงบันดาลใจในการทำงานวิจัยขึ้นนี้จนสำเร็จลุล่วงด้วยดี

## เอกสารอ้างอิง

- Arumugam, P., & Saranya, R. (2018). Outlier Detection and Missing Value in Seasonal ARIMA Model Using Rainfall Data\*. *Materials Today: Proceedings*, 5(1), 1791–1799. <https://doi.org/10.1016/j.matpr.2017.2017.11.277>
- Barman, M., & Dev Choudhury, N. B. (2019). Season specific approach for short-term load forecasting based on hybrid FA-SVM and similarity concept. *Energy*, 174, 886–896. <https://doi.org/10.1016/j.energy.2019.03.010>
- Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Ljung, G.M., (2008). *Time Series Analysis: Forecasting and Control*, 4th ed. John Wiley and Sons, Inc., New Jersey.

- Chen, Y., Xu, P., Chu, Y., Li, W., Wu, Y., Ni, L., Wang, K. (2017). Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. *Applied Energy*, 195, 659–670. <https://doi.org/10.1016/j.apenergy.2017.03.034>
- Cheng, H., Ding, X., Zhou, W., & Ding, R. (2019). A hybrid electricity price forecasting model with Bayesian optimization for German energy exchange. *International Journal of Electrical Power & Energy Systems*, 110, 653–666. <https://doi.org/10.1016/j.ijepes.2019.03.056>
- Cheng, Y., Zhang, H., Liu, Z., Chen, L., & Wang, P. (2019). Hybrid algorithm for short-term forecasting of PM2.5 in China. *Atmospheric Environment*, 200, 264–279. <https://doi.org/10.1016/j.atmosenv.2018.12.025>
- Curceac, S., Ternynck, C., Ouarda, T. B. M. J., Chebana, F., & Niang, S. D. (2019). Short-term air temperature forecasting using Nonparametric Functional Data Analysis and SARMA models. *Environmental Modelling & Software*, 111, 394–408. <https://doi.org/10.1016/j.envsoft.2018.09.017>
- de Marcos, R. A., Bello, A., & Reneses, J. (2019). Electricity price forecasting in the short term hybridising fundamental and econometric modelling. *Electric Power Systems Research*, 167, 240–251. <https://doi.org/10.1016/j.epsr.2018.10.034>
- de Oliveira, E. M., & Cyrino Oliveira, F. L. (2018). Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy*, 144, 776–788. <https://doi.org/10.1016/j.energy.2017.12.049>
- Junsagoon, S. (2015). Time Series Forecasting : Linear Approach, Non – Linear Approach and Hybrid Models. *EAU HERITAGE JOURNAL Science and Technology*, 9(2) ,50 – 63. (in Thai)
- Mao, Q., Zhang, K., Yan, W., & Cheng, C. (2018). Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. *Journal of Infection and Public Health*, 11(5), 707–712. <https://doi.org/10.1016/j.jiph.2018.04.009>
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C.C., & Lin, C.C. (2015). Misc Functions of the Department of Statistics, Probability Theory Group. <http://cran.r-project.org/web/packages/e1071/index.html>, 5 April 2018.
- Mohammadi, K., Shamshirband, S., Tong, C. W., Arif, M., Petković, D., & Ch, S. (2015). A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation. *Energy Conversion and Management*, 92, 162–171. <https://doi.org/10.1016/j.enconman.2014.12.050>
- Provincial Waterworks Authority. Data water consumption. 2018. Retrieved February 20, 2018, from: <https://www.pwa.co.th/province/> (in Thai)

- R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria . Retrieved February 20, 2018, from <https://www.R-project.org/>.
- R Core Team. (2019). forecast: Forecasting Functions for Time Series and Linear Models. R Foundation for Statistical Computing, Vienna, Austria. Retrieved February 20, 2018, from <https://cran.r-project.org/web/packages/forecast/>
- Sujjaviriyasup, T. (2019). Hybrid Model of Linear and Nonlinear Model for Forecasting Annual Molumes of International Rice Exports. *Burapha Science Journal*, 24(2) , 517 – 531. (in Thai)
- Tongpitak, N., Matchima, K., Sutthiaion, T. and Thepchim, S. (2017). Forecasting of the Water Consumption of the Resident Residing in the Urban Areas of Muang Ubon and Warinchamra Districts. *JOURNAL OF INDUSTRIAL TECHNOLOGY UBON RATCHATANI RAJABHAT UNIVERSITY*, 7(1), 58 – 74. (in Thai)
- Vapnik, V. (1998). The Nature of Statistical Learning Theory. Springer-Verlag.
- Xu, S., Chan, H. K., & Zhang, T. (2019). Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach. *Transportation Research Part E: Logistics and Transportation Review*, 122, 169–180. <https://doi.org/10.1016/j.tre.2018.12.005>
- Yang, Z., & Wang, J. (2018). A hybrid forecasting approach applied in wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm. *Energy*, 160, 87–100. <https://doi.org/10.1016/j.energy.2018.07.005>
- Yaslan, Y., & Bican, B. (2017). Empirical mode decomposition based denoising method with support vector regression for time series prediction: A case study for electricity load forecasting. *Measurement*, 103, 52–61. <https://doi.org/10.1016/j.measurement.2017.02.007>
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 50, 159 – 175. [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)