Research Article

Received: September 19, 2019; Accepted: April 23, 2020

การศึกษาเปรียบเทียบตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี

Comparative Study to Determine Optimal Models for Forecasting the Number of Tap Water Users of the Provincial Waterworks Authority Pathum Thani Branch

คชินทร์ โกกนุทาภรณ์*

หลักสูตรคณิตศาสตร์ประยุกต์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ ในพระบรมราชูปถัมภ์ ตำบลคลองหนึ่ง อำเภอคลองหลวง จังหวัดปทุมธานี 13180

Kachin Goganutapon*

Program in Applied Mathematics, Faculty of Science and Technology, Valaya Alongkorn Rajabhat University under the Royal Patronage, Khlong Nueng, Khlong Luang, Pathum Thani 13180

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาหาตัวแบบพยากรณ์ที่เหมาะสมสำหรับการพยากรณ์ จำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี ในการศึกษาครั้งนี้ได้นำเทคนิคการพยากรณ์มาช่วยในการวิเคราะห์ข้อมูล ซึ่งประกอบด้วยวิธีสมการแนวโน้มเชิงเส้น วิธีสมการแนวโน้มกำลังสอง วิธีสมการแนวโน้มกำลังสาม วิธีสมการแนวโน้มเอ็กซ์โพเนนเชียล และวิธีสมการแนวโน้มกำลัง ข้อมูลที่ใช้เป็นข้อมูลทุติยภูมิที่รวบรวมจากกองศูนย์ข้อมูลและแผนเทคโนโลยีสารสนเทศ การประปาส่วนภูมิภาค สาขาปทุมธานี ลักษณะข้อมูลจำแนกเป็นรายเดือน ระหว่างเดือนมกราคม พ.ศ. 2555 ถึงเดือนกรกฎาคม พ.ศ. 2562 จำนวน 91 ค่า ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด ข้อมูลชุดที่ 1 ตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึงเดือนธันวาคม พ.ศ. 2561 จำนวน 84 ค่า สำหรับเปรียบเทียบหาวิธีการพยากรณ์ที่เหมาะสมที่สุด โดยใช้เกณฑ์พิจารณาค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAD) และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (MAPE) ที่ต่ำที่สุด จากนั้นเลือกวิธีการพยากรณ์ที่เหมาะสมที่สุดคำนวณหาช่วงการพยากรณ์ล่วงหน้า โดยใช้ข้อมูลชุดที่ 2 คือ ตั้งแต่เดือนมกราคม พ.ศ. 2562 ถึงเดือนกรกฎาคม พ.ศ. 2562 จำนวน 7 ค่า โดยใช้เกณฑ์ค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (MAPE) ที่ต่ำที่สุด ผลการวิจัยพบว่าวิธีการพยากรณ์ที่มีความเหมาะสมที่สุด คือ การพยากรณ์โดยวิธีสมการแนวโน้มกำลังสาม โดยมีตัวแบบพยากรณ์ คือ $\hat{Y}_3(t) = -0.007t^3 - 1.01t^2 + 243.173t + 43423.657$ เมื่อ t แทนช่วงเวลา จากรูปแบบดังกล่าว นำมาคำนวณหาช่วงการพยากรณ์ล่วงหน้าที่เหมาะสมที่สุด 3 เดือน 5 เดือน และ 7 เดือน พบว่าวิธีนี้เหมาะสำหรับการพยากรณ์ล่วงหน้า 7 เดือน

***ผู้รับผิดชอบบทความ :** kachin@vru.ac.th

คำสำคัญ : การพยากรณ์; จำนวนผู้ใช้น้ำประปา; วิธีสมการแนวโน้มกำลังสาม; ค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์; ค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย

Abstract

The objective of this study was to determine optimal models for forecasting the number of tap water users of the provincial waterworks authority (PWA) Pathum Thani branch. In this study, we applied 5 forecasting techniques to analyze the data. These include linear trend analysis method, quadratic trend analysis method, cubic trend analysis method, exponential trend analysis method, and power trend analysis method. Our raw data were secondary data, which were taken from the data center and information technology planning division, the PWA Pathum Thani branch. All 91 values were separated on a monthly basis between from January 2012 and July 2019, which were used and separated into 2 sets. The first set with 84 values were separated on a monthly basis between from January 2012 and December 2018 for comparing and finding the most suitable forecasting method via criteria of the lowest mean absolute deviation (MAD) and mean absolute percent error (MAPE). Then the selected suitable method was used to determine the most suitable forecasting period by the second group, which contained 7 values and were separated on a monthly basis between from January 2019 and July 2019. The lowest MAPE was used as the criteria of each period. The result indicated that cubic trend analysis method was the best method, with the forecasting model: $\hat{Y}_3(t) = -0.007t^3 - 1.01t^2 + 243.173t + 43423.657$. When t represents time from the method, it was implemented for forecasting 3, 5 and 7 months, and showed that the method was suitable for advance 7 months.

Keywords: forecasting; number of new students; cubic trend analysis method; mean absolute deviation; mean absolute percent error

1. บทนำ

ข้อมูลจากสำนักงานสถิติแห่งชาติระบุว่า ปัจจุบันประเทศไทยมีจำนวนประชากรประมาณ 66 ล้านคน เฉพาะกรุงเทพมหาครมีประชากรอาศัยอยู่มากกว่า 8 ล้านคน จึงเริ่มมีการขยายเมืองไปยังพื้นที่รอบนอกเพิ่มมากขึ้น แต่เมื่อพิจารณาถึงความต้องการใช้น้ำของประชาชนในพื้นที่กรุงเทพฯ และปริมณฑล ทำให้มีข้อกังวลว่าน้ำมีปริมาณเพียงพอต่อความต้องการที่มากขึ้นจากปริมาณคนที่เพิ่มมากขึ้นหรือไม่ ซึ่งยังไม่รวมประชากรแฝง การบริหารจัดการที่มีอยู่เพียงพอแล้วหรือยัง โดยข้อมูลจากกระทรวงพลังงาน ระบุว่าคนกรุงเทพฯ ใช้น้ำเฉลี่ยวันละ 200 ลิตร ขณะที่คนต่างจังหวัดใช้เพียงวันละ 50 ลิตร ซึ่งความต้องการใช้สอยน้ำเพื่อบริโภค อาบน้ำ ซักผ้า และอื่น ๆ จะต่างกันไปตามฐานะและความเป็นอยู่ของผู้บริโภค ตลอดจนลักษณะการใช้สอยของอาคารและสภาพของ

ดินฟ้าอากาศ [1] จากปัญหาข้างต้นผู้วิจัยจึงสนใจศึกษาจำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี โดยการใช้วิธีการพยากรณ์ ซึ่งการพยากรณ์เป็นกระบวนการคาดคะเนแนวโน้ม หรือรูปแบบของการเกิดเหตุการณ์ต่าง ๆ ในอนาคต โดยอาศัยข้อมูลที่มีอยู่ ดังนั้นการพยากรณ์จึงมีบทบาทสำคัญเป็นอย่างมากในการวางแผนและการตัดสินใจเกี่ยวกับการดำเนินงานในด้านต่าง ๆ ในปัจจุบัน ทั้งระยะสั้นและระยะยาว การพยากรณ์ยังเป็นเครื่องมืออย่างหนึ่งที่จะช่วยให้ได้ข้อมูลในอนาคตเพื่อประกอบการวางแผน จึงเป็นที่ยอมรับกันว่าการพยากรณ์มีบทบาทสำคัญทั้งในงานของภาครัฐและภาคเอกชน ปัจจุบันตัวแบบพยากรณ์ที่ใช้เทคนิคการวิเคราะห์อนุกรมเวลาหลายวิธีได้ถูกนำมาใช้ในการพยากรณ์จำนวนผู้ใช้น้ำประปาสอดคล้องงานวิจัยของ Sakulsri [2] และ Srichai และคณะ [3] การคำนวณวิธีการพยากรณ์แต่ละวิธีที่มีประสิทธิภาพพิจารณาจากค่าความคลาดเคลื่อนที่น้อยที่สุดของการพยากรณ์ซึ่งการเลือกวิธีการพยากรณ์ที่เหมาะสมควรพิจารณาจากหลาย ๆ ปัจจัย ได้แก่ ช่วงการพยากรณ์ที่ต้องการเวลาที่ใช้ในการพยากรณ์ลักษณะของข้อมูลอนุกรมเวลาและขนาดของอนุกรมเวลา [4] เพื่อให้ได้วิธีการพยากรณ์ที่เหมาะสม และง่ายต่อการนำไปใช้สำหรับพยากรณ์จำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี จึงนำมาสู่ความสนใจของผู้วิจัยที่จะศึกษาเพื่อหาวิธีการพยากรณ์ที่เหมาะสม โดยเลือกวิธีการพยากรณ์ที่ง่ายต่อการเข้าใจ และวิธีที่เป็นที่นิยมใช้ นำมาเปรียบเทียบกัน 5 วิธี ได้แก่ วิธีสมการแนวโน้มเชิงเส้น วิธีสมการแนวโน้มกำลังสอง วิธีสมการแนวโน้มกำลังสาม วิธีสมการแนวโน้มเอ็กซ์โปเนนเชียล และวิธีสมการแนวโน้มกำลัง เพื่อให้ได้เครื่องมือที่ให้ผลการทำนายที่แม่นยำ สามารถนำไปเป็นข้อมูลเบื้องต้นในการวิเคราะห์สถานการณ์การผลิตน้ำประปา เพื่อเป็นการวางแผนในการผลิตน้ำประปาให้เพียงพอต่อความต้องการการใช้น้ำประปาในจังหวัดปทุมธานีต่อไป

2. อุปกรณ์และวิธีการ

2.1 การจัดเตรียมข้อมูล

ข้อมูลที่นำมาวิเคราะห์ คือ จำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี กองศูนย์ข้อมูลและแผนเทคโนโลยีสารสนเทศ การประปาส่วนภูมิภาค สาขาปทุมธานี [5] ลักษณะข้อมูลจำแนกเป็นรายเดือน ระหว่างเดือนมกราคม พ.ศ. 2555 ถึงเดือนกรกฎาคม พ.ศ. 2561 จำนวน 91 ค่า ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด ข้อมูลชุดที่ 1 ตั้งแต่เดือนมกราคม 2555 ถึงเดือนธันวาคม พ.ศ. 2561 จำนวน 84 ค่า สำหรับเปรียบเทียบหาวิธีการพยากรณ์ที่เหมาะสมที่สุด จากนั้นเลือกวิธีการพยากรณ์ที่เหมาะสมที่สุด คำนวณหาช่วงการพยากรณ์ล่วงหน้าโดยใช้ข้อมูลชุดที่ 2 คือ ตั้งแต่เดือนมกราคม พ.ศ. 2562 ถึงเดือนกรกฎาคม พ.ศ. 2562 จำนวน 7 ค่า

2.2 การศึกษาความเคลื่อนไหวของอนุกรมเวลา

เพื่อพิจารณาเบื้องต้นว่าอนุกรมเวลานั้น ๆ มีลักษณะเป็นแบบใด โดยทดสอบองค์ประกอบของแนวโน้มและความแปรผันตามฤดูกาล โดยพิจารณาจากแผนภาพการกระจาย การทดสอบแนวโน้มแบบเครื่องหมาย (sign test) และการทดสอบความแปรผันตามฤดูกาลด้วยวิธีการทดสอบของครัสคาล-วัลลิส (Kruskal-Wallis test)

2.2.1การทดสอบแนวโน้มแบบเครื่องหมาย

เป็นการทดสอบแนวโน้มที่ตัวทดสอบสถิติได้จากการพิจารณาค่าสังเกตในอนุกรมเวลาที่อยู่ในตำแหน่งติดกันว่ามีความต่างกันทางบวกหรือลบ นั่นคือ พิจารณาว่าค่าสังเกต Y_t มากว่าหรือน้อยกว่ากว่า

ค่าสังเกต Y_{t-1} หรือพิจารณาเครื่องหมายของผลต่างครั้งที่หนึ่งของค่าสังเกต $(Y_t - Y_{t-1})$ ถ้าจำนวนผลต่างเป็นบวกมากแสดงว่าอนุกรมเวลามีแนวโน้มขึ้น และถ้าเป็นลบมากแสดงว่าอนุกรมเวลามีแนวโน้มลง กำหนด V เป็นจำนวนครั้งที่หนึ่งค่าของค่าสังเกตเป็นบวก [4] ขั้นตอนของการทดสอบแบบเครื่องหมายมีดังนี้

(1) กำหนด H_0: อนุกรมเวลาไม่มีแนวโน้ม กับ H_1: อนุกรมเวลามีแนวโน้มขึ้นหรือลง

(2) สำหรับอนุกรมเวลาขนาดใหญ่ หรือ $n \geq 20$ ตัวสถิติ V มีการแจกแจงประมาณแบบปกติมีค่าเฉลี่ย μ_V และส่วนเบี่ยงเบนมาตรฐาน σ_V ใช้ตัวทดสอบสถิติ $Z = \frac{V-\mu_V}{\sigma_V}$ เมื่อ $\mu_V = \frac{n}{2}$ และ $\sigma_V = \sqrt{\frac{n}{4}}$ ที่ระดับนัยสำคัญ α มีช่วงวิกฤติ $CR: |Z| \geq Z_{\alpha/2}$ ซึ่ง $Z_{\alpha/2}$ เป็นค่าวิกฤติ

2.2.2 การทดสอบความแปรผันตามฤดูกาลด้วยวิธีการทดสอบของครัสคาล-วัลลิส

การทดสอบของครัสคาล-วัลลิสนำมาประยุกต์กับการทกสอบอนุกรมเวลาที่จำกัดค่าแนวโน้มแล้วมีความแปรผันตามฤดูกาลหรือไม่ การทดสอบนี้จะใช้ลำดับของข้อมูลอนุกรมเวลาที่จำกัดแนวโน้มแล้ว การทดสอบตามวิธีนี้มีขั้นตอนดังต่อไปนี้

(1) หาข้อมูลอนุกรมเวลาที่จำกัดแนวโน้มทั้งหมด แล้วเรียงลำดับข้อมูลดังกล่าวจากค่าต่ำสุด (ให้เป็นอันดับที่ 1) และเพิ่มขึ้นเรื่อย ๆ จนครบข้อมูลทุกตัว ซึ่งจะเท่ากับ n เมื่อ n คือจำนวนข้อมูลทั้งหมด

(2) ตั้งสมมติฐานหลักและสมมุติฐานทางเลือก ดังนั้น H_0: อนุกรมเวลาไม่มีความผันแปรตามฤดูกาล กับ H_1: อนุกรมเวลามีความผันแปรตามฤดูกาล

(3) ตัวสถิติที่ใช้ทดสอบ $H = \frac{12}{n(n+1)}\left[\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(n+1)$ เมื่อ K แทนจำนวนฤดูกาล R_i แทนผลรวมของอันดับข้อมูลในฤดูกาลที่ i, n_i แทนจำนวนข้อมูลในฤดูกาลที่ i โดยที่ $n = \sum_{i=1}^{n} n_i$

(4) การตัดสินใจที่จะยอมรับหรือปฏิเสธสมมติฐานหลัก (H_0) ถ้า $n_1 \leq K, K \leq 5, n \leq 15$ จะใช้ตาราง Kruskal-Wallis แต่ถ้า $n_i > 5$ จะใช้ตารางการแจกแจงความน่าจะเป็นไคกำลังสองโดยปฏิเสธ (H_0) เมื่อ $H > \chi^2_{\alpha,K-1}$ [6]

2.3 การวิเคราะห์ข้อมูลทางสถิติ

การวิเคราะห์ข้อมูลในงานวิจัยนี้เพื่อเปรียบเทียบวิธีการพยากรณ์อนุกรมเวลา 5 วิธี สร้างตัวแบบพยากรณ์ด้วยวิธีสมการแนวโน้มเส้นตรง วิธีสมการแนวโน้มกำลังสอง วิธีสมการแนวโน้มกำลังสาม วิธีสมการแนวโน้มเอ็กโปแนนเชียล และวิธีสมการแนวโน้มกำลัง โดยวิเคราะห์ข้อมูลโดยโปรแกรมสำเร็จรูปทางสถิติ การดำเนินการวิเคราะห์แสดงรายละเอียดดังนี้

2.3.1 การพยากรณ์โดยวิธีสมการแนวโน้มเส้นตรง (linear trend equations method)

ตัวแบบสมการแนวโน้มเส้นตรงอยู่ในรูป $Y = a + bX$ (1) จะได้สมการปกติดังนี้

$$\sum_{i=1}^{n} y_i = an + b\sum_{i=1}^{n} x_i$$
$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2$$

เมื่อ a และ b เป็นค่าคงที่ Y แทนค่าแนวโน้มของข้อมูล X แทนเวลาของข้อมูล n แทนจำนวนข้อมูลอนุกรมเวลาทั้งหมด [6]

2.3.2 การพยากรณ์โดยวิธีสมการแนวโน้มกำลังสอง (quadratic trend equations method)

ตัวแบบสมการแนวโน้มกำลังสองอยู่ในรูป $Y = a + bX + cX^2$ (2) จะได้สมการปกติดังนี้

$$\sum_{i=1}^{n} y_i = an + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x^2$$
$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i^3$$
$$\sum_{i=1}^{n} x_i^2 y_i = a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4$$

เมื่อ a, b และ c เป็นค่าคงที่ Y แทน ค่าแนวโน้มของข้อมูล X แทน เวลาของข้อมูล n แทน จำนวนข้อมูลอนุกรมเวลาทั้งหมด [6]

2.3.3 การพยากรณ์โดยวิธีสมการแนวโน้มกำลังสาม (cubic trend equations method)

ตัวแบบสมการแนวโน้มกำลังสามอยู่ในรูป $Y = a + bX + cX^2 + dX^3$ (3) จะได้สมการปกติดังนี้

$$\sum_{i=1}^{n} y_i = an + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x_1^2 + d\sum_{i=1}^{n} x_i^3$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i^3 + d\sum_{i=1}^{n} x_i^4$$

$$\sum_{i=1}^{n} x_i^2 y_i = a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4 + d\sum_{i=1}^{n} x_i^5$$

$$\sum_{i=1}^{n} x_i^3 y_i = a\sum_{i=1}^{n} x_i^3 + b\sum_{i=1}^{n} x_i^4 + c\sum_{i=1}^{n} x_i^5 + d\sum_{i=1}^{n} x_i^6$$

เมื่อ a, b, c และ d เป็นค่าคงที่ Y แทน ค่าแนวโน้มของข้อมูล X แทน เวลาของข้อมูล n แทน จำนวนข้อมูลอนุกรมเวลาทั้งหมด [6]

2.3.4 การพยากรณ์โดยวิธีสมการแนวโน้มเอ็กโปแนนเชียล (exponential trend equations method)

รูปแบบแนวโน้มเอ็กซ์โปแนนเชียล คือ $Y_t = \beta_0\beta_1^t\varepsilon_t$ สร้างตัวแบบสมการแนวโน้มเอ็กซ์โปแนนเชียล $\hat{Y}_t = b_0 b_1^t$ (4) ด้วยวิธีกำลังน้อยที่สุดทำโดยแปลงรูปแนวโน้มเอ็กซ์โปแนนเชียลเป็นรูปแบบแนวโน้มเส้นตรง $Y_t^{/} = \acute{\beta}_0 + \beta_1 t + \acute{\varepsilon}_1$ ซึ่ง $\acute{Y}_t = lnY_t, \acute{\beta}_0 = ln\beta_0, \beta_1 = ln\beta_1$ และ $\acute{\varepsilon}_t = ln\varepsilon_t$ หา $\acute{b}_0$ และ $\acute{b}_1$ ซึ่งเป็นค่าประมาณของ $\acute{\beta}_0$ และ $\acute{\beta}_1$ จาก $\acute{b}$ และ $\acute{b_1}$ ที่ได้หา b_0 และ b_1 จาก $b_0 = exp(\acute{b}_0)$ และ $b_1 = exp(\acute{b}_1)$ ส่วนวิธีกำลังสองน้อยที่สุดไม่เชิงเส้นตรงหา b_0 และ b_1 ที่ทำให้ผลรวมกำลังสองของความคลาดเคลื่อน หรือ SSE มีค่าน้อยที่สุด โดยกำหนดค่าเริ่มต้นให้กับ b_0 และ b_1 จากการพล๊อตของอนุกรมเวลา การคำนวณค่า b_0 และ b_1 จะทำหลายรอบจนกว่าจะให้ค่า b_0 และ b_1 ที่คงที่และให้ค่า SEE มีค่าน้อยที่สุด [4]

2.3.5 การพยากรณ์โดยวิธีสมการแนวโน้มกำลัง (power trend equations method)

รูปแบบสมการแนวโน้มกำลัง $Y_t = \beta_0 t^{\beta_1}\varepsilon_t$ สร้างตัวแบบสมการแนวโน้มกำลัง $\hat{Y}_t = b_0 t^{b_1}$ (5)

ในทำนองเดียวกันจากรูปแบบสมการแนวโน้มเอ็กโปแนนเชียลด้วยวิธีกำลังสองน้อยที่สุด ทำโดยแปลงรูปแบบแนวโน้มกำลังเป็นรูปแบบสมการแนวโน้มเส้นตรง $Y_t^{/} = \acute{\beta}_0 + \beta_1\acute{t} + \acute{\varepsilon}_1$ ซึ่ง $\acute{Y}_t = lnY_t, \acute{\beta}_0 = ln\beta_0, \acute{t} = lnt$ และ $\acute{\varepsilon}_t = ln\varepsilon_t$ ประมาณ $\acute{\beta}_0$ และ $\acute{\beta}_1$ ด้วย $\acute{b}_0$ และ b_1 ตามลำดับและหา b_0 จาก $b_0 = exp(\acute{b}_0)$ กรณีไม่แปลงเป็นรูปแบบสมการแนวโน้มเส้นตรงจะใช้วิธีกำลังสองน้อยที่สุดไม่เชิงเส้นตรง [4]

2.4 การเปรียบเทียบความแม่นยำของค่าพยากรณ์

การวิจัยครั้งนี้ได้ศึกษาเปรียบเทียบตัวแบบพยากรณ์ที่เหมาะสมกับอนุกรมเวลาจำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี โดยการเปรียบเทียบค่า MAD และ MAPE จากวิธีการพยากรณ์ทั้ง 5 วิธี ตัวแบบพยากรณ์ที่มีค่า MAD และ RMSE ต่ำที่สุดจัดเป็นตัวแบบที่มีความเหมาะสมกับอนุกรมเวลาชุดนี้มากที่สุด เนื่องจากให้ค่าพยากรณ์ที่มีความแตกต่างกับข้อมูลจริงน้อยที่สุด เกณฑ์ MAD และ MAPE [6] แสดงดังนี้

2.4.1 ค่าเฉลี่ยความคลาดเคลื่อน (mean absolute deviation, MAD)

$$MAD = \frac{1}{n}\sum_{t=1}^{n}\left|Y_t - \hat{Y}_t\right| \quad (6)$$

2.4.2 ค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (mean absolute percentage error, MAPE)

$$MAPE = \frac{100}{n}\sum_{t=1}^{n_2}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \quad (7)$$

เมื่อ Y_t แทน ค่าของข้อมูลจริง ณ เวลา t, $\hat{Y}_t$ แทน ค่าพยากรณ์ ณ เวลา t, n แทน จำนวนข้อมูล

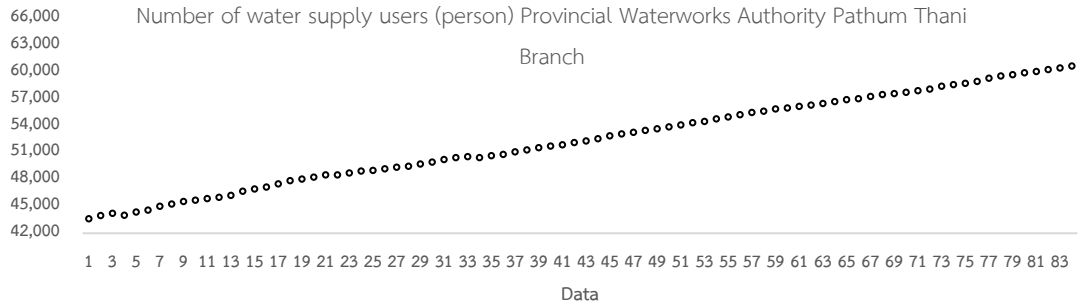

Figure 1 The movement of the number of tap water users of the PWA Pathum Thani branch

3. ผลการวิจัย

3.1 ผลการวิเคราะห์ลักษณะการเคลื่อนไหวของอนุกรมเวลา

การพิจารณาลักษณะการเคลื่อนไหวของอนุกรมเวลาจำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี ตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึงเดือนกรกฎาคม พ.ศ. 2561 นำมาสร้างแผนภาพการกระจาย ดังรูปที่ 1

จากภาพที่ 1 พบว่าอนุกรมเวลาจำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี มีการเคลื่อนไหวของแนวโน้มแต่ไม่มีอิทธิพลของความแปรผันแปรตามฤดูกาล

3.1.1 การตรวจสอบแนวโน้ม

การตรวจสอบแนวโน้มด้วยวิธีการทดสอบเครื่องหมาย (sign test) กำหนดสมมติฐานหลักและรองดังนี้ H_0: อนุกรมเวลาไม่มีแนวโน้ม กับ H_1: อนุกรมเวลามีแนวโน้มขึ้นหรือลง เมื่อจำนวนเครื่องหมายของผลต่างที่เป็นบวก (V) คือ 88 ค่าตัวทดสอบสถิติ คือ $Z = 8.91$ เมื่อ $\mu_V = 45.5, \sigma = 4.77$ ที่ระดับนัยสำคัญ 0.05 มีช่วงวิกฤติ $CR: |Z| \geq Z_{\alpha/2} = 1.96$ พบว่า $Z > Z_{\alpha/2}$ จึงปฏิเสธ H_0 สรุปได้ว่าอนุกรมเวลามีแนวโน้มขึ้นหรือลง

3.1.2 การตรวจสอบความแปรผันตามฤดูกาล

การตรวจสอบความแปรผันตามฤดูกาลด้วยวิธีการทดสอบของครัสคาล-วัลลิส กำหนดสมมติฐานหลักและรองดังนี้ H_0: อนุกรมเวลาไม่มีส่วนประกอบความแปรผันแปรตามฤดูกาล กับ H_1: อนุกรมเวลามีส่วนประกอบความแปรผันแปรตามฤดูกาล ค่าตัวทดสอบสถิติ คือ $H = -275.209$ เมื่อที่ระดับนัยสำคัญ 0.05 มีช่วงวิกฤติ $CR: |H| \geq \chi^2_{.05,11} = 26.7569$ พบว่า $H < \chi^2_{.05,11}$ จึงยอมรับ H_0 สรุปได้ว่าอนุกรมเวลาไม่มีส่วนประกอบความแปรผันแปรตามฤดูกาล

3.2 ผลการพยากรณ์

3.2.1 ผลการพยากรณ์โดยวิธีสมการแนวโน้มเส้นตรง

เมื่อใช้โปรแกรมสำเร็จรูปทางสถิติคำนวณค่าพยากรณ์ โดยวิธีสมการแนวโน้มเส้นตรง พบว่าค่า $\text{MAD} = 149.28$ และค่า $\text{MAPE} = 0.3008$ ดังรูปที่ 2 ตัวแบบพยากรณ์สมการแนวโน้มเส้นตรง คือ $\hat{Y}_1(t) = 201.407t + 43813.353$ (8)

เมื่อ $\hat{Y}_1(t)$ แทน ค่าพยากรณ์จำนวนผู้ใช้น้ำประปาด้วยวิธีสมการแนวโน้มเส้นตรง ณ เวลา t เมื่อ t แทน ช่วงเวลาที่พยากรณ์ $t = 58$ (เดือนมกราคม ปี พ.ศ. 2562)

3.3 ผลการพยากรณ์โดยวิธีสมการแนวโน้มกำลังสอง

เมื่อใช้โปรแกรมสำเร็จรูปทางสถิติคำนวณค่าพยากรณ์ โดยวิธีสมการแนวโน้มกำลังสอง พบว่าค่า $\text{MAD} = 137.223$ และค่า $\text{MAPE} = 0.2726$ ตัวแบบพยากรณ์สมการแนวโน้มกำลังสอง คือ $\hat{Y}_2(t) = -0.142t^2 + 213.499t + 43640.037$ (9) เมื่อ $\hat{Y}_2(t)$ แทนค่าพยากรณ์จำนวนผู้ใช้น้ำประปา โดยวิธีสมการแนวโน้มกำลังสอง ณ เวลา t เมื่อ t แทนช่วงเวลาที่พยากรณ์ $t = 58$ (เดือน มกราคม ปี พ.ศ. 2562) แสดงความสัมพันธ์ระหว่างค่าจริงและค่าพยากรณ์ดังรูปที่ 3

3.4 ผลการพยากรณ์โดยวิธีสมการแนวโน้มกำลังสาม

เมื่อใช้โปรแกรมสำเร็จรูปทางสถิติคำนวณค่าพยากรณ์ โดยวิธีสมการแนวโน้มกำลังสามพบว่าค่า $\text{MAD} = 131.3655$ และค่า $\text{MAPE} = 0.2564$ ตัวแบบพยากรณ์ โดยวิธีสมการแนวโน้มกำลังสาม คือ $\hat{Y}_3(t) = -0.007t^3 - 1.01t^2 + 243.17t + 43423.66$ (10)

เมื่อ $\hat{Y}_3(t)$ แทน ค่าพยากรณ์จำนวนผู้ใช้น้ำประปา โดยวิธีสมการแนวโน้มกำลังสาม ณ เวลา t เมื่อ t แทน ช่วงเวลาที่พยากรณ์ $t = 58$ (เดือน มกราคม ปี พ.ศ. 2562) แสดงความสัมพันธ์ระหว่างค่าจริงและค่าพยากรณ์ดังรูปที่ 4

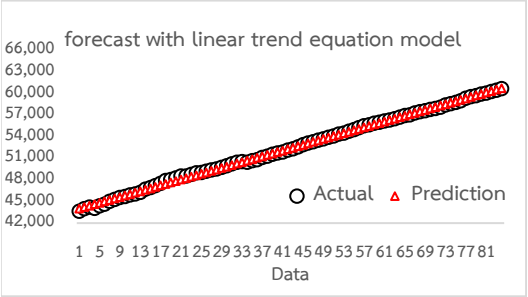

Figure 2 The relationship between the number of real water supply users and the predicted values $\hat{Y}_1(t)$

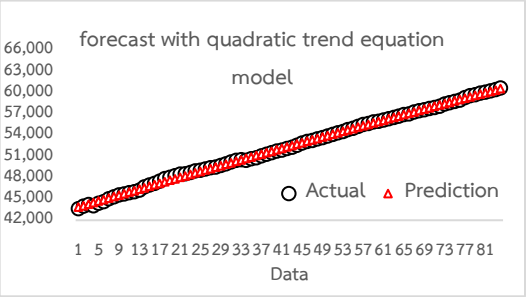

Figure 3 The relationship between the number of real water supply users and the predicted values $\hat{Y}_2(t)$

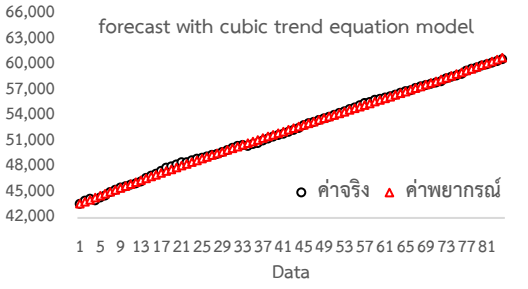

Figure 4 The relationship between the number of real water supply users and the predicted values $\hat{Y}_3(t)$

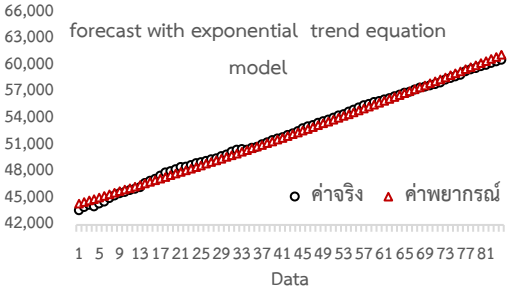

Figure 5 The relationship between the number of real water supply users and the predicted values $\hat{Y}_4(t)$

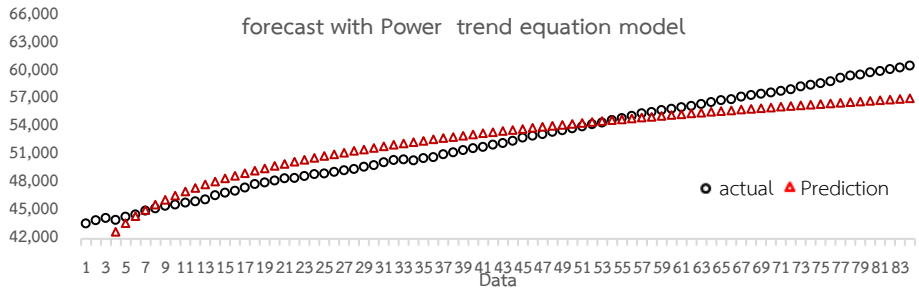

Figure 6 The relationship between the number of real water supply users and the predicted values $\hat{Y}_5(t)$

Table 1 Forecasting accuracy

Forecasting models	MAPE	MAD
Linear trend equations method: $\hat{Y}_1(t)$	0.3008	149.2802
Quadratic trend equations method: $\hat{Y}_2(t)$	0.2726	137.2230
Cubic trend equations method: $\hat{Y}_3(t)$	0.2564	131.3655
Exponential trend equations method: $\hat{Y}_4(t)$	0.5546	2826517
Power trend equation model: $\hat{Y}_5(t)$	2.9610	1540.8224

3.5 ผลการพยากรณ์โดยวิธีสมการแนวโน้มเอ็กโปแนนเชียล

เมื่อใช้โปรแกรมสำเร็จรูปทางสถิติคำนวณค่าพยากรณ์ โดยวิธีสมการแนวโน้มเอ็กโปแนนเชียล พบว่าค่า $\text{MAD} = 282.6517$ และค่า $\text{MAPE} = 0.5546$ ตัวแบบพยากรณ์สมการแนวโน้มเอ็กโปแนนเชียล คือ $\hat{Y}_4(t) = 4423.861(1.003879193)^t$ (11) เมื่อ $\hat{Y}_4(t)$ แทน ค่าพยากรณ์จำนวนผู้ใช้น้ำประปา โดยวิธีสมการแนวโน้มเอ็กโปแนนเชียล ณ เวลา t เมื่อ t แทน ช่วงเวลาที่พยากรณ์ $t = 58$ (เดือนมกราคม ปี พ.ศ. 2562) แสดงความสัมพันธ์ระหว่างค่าจริงและค่าพยากรณ์ดังรูปที่ 5

3.6 ผลการพยากรณ์โดยวิธีสมการแนวโน้มกำลัง

เมื่อใช้โปรแกรมสำเร็จรูปทางสถิติคำนวณค่าพยากรณ์ โดยวิธีสมการแนวโน้มกำลัง พบว่าค่า $\text{MAD} = 1540.82$ และค่า $\text{MAPE} = 2.9610$ ตัวแบบพยากรณ์สมการแนวโน้มกำลัง คือ $\hat{Y}_5(t) = 37510.23t^{0.095}$ (12) เมื่อ $\hat{Y}_5(t)$ แทน ค่าพยากรณ์จำนวนผู้ใช้น้ำประปา โดยวิธีสมการแนวโน้มกำลัง ณ เวลา t เมื่อ t แทน ช่วงเวลาที่พยากรณ์ $t = 58$ (เดือนมกราคม ปี พ.ศ. 2562) แสดงความสัมพันธ์ระหว่างค่าจริงและค่าพยากรณ์ดังรูปที่ 6

3.7 ผลการวัดความคลาดเคลื่อนของการพยากรณ์

เมื่อได้ผลการพยากรณ์ทั้ง 5 วิธี แล้ว ขั้นตอนนี้เป็นการเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์ โดยใช้เกณฑ์พิจารณาค่า MAPE และค่า MAD ที่ต่ำที่สุด ซึ่งจะพิจารณาที่ค่า MAPE ก่อน ถ้าหากมีค่าเท่ากัน จึงจะพิจารณาที่ค่า MAD เป็นลำดับถัดไป ผลการเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์แสดงดังตารางที่ 1

ตารางที่ 1 พบว่าการพยากรณ์ด้วยวิธีสมการแนวโน้มกำลังสาม ซึ่งมีตัวแบบพยากรณ์ คือ $\hat{Y}_3(t)$ เป็นวิธีที่มีความแม่นยำมากที่สุด เนื่องจากมีค่า MAPE และค่า MAD ต่ำที่สุด

3.8 ช่วงการพยากรณ์ล่วงหน้าที่เหมาะสม

การพิจารณาหาช่วงการพยากรณ์ที่เหมาะสม โดยกำหนดช่วงพยากรณ์เป็น 3 ช่วง คือ ช่วงเวลาล่วงหน้า 3, 5 และ 7 เดือน โดยใช้เกณฑ์การพิจารณาค่า MAPE ที่ต่ำที่สุด ดังแสดงในตารางที่ 2

Table 2 Comparison of suitable forecasting intervals using cubic trend equation model

Forecast period (months)	MAPE
3	0.1767
5	0.1773
7	0.1638

ตารางที่ 2 พบว่าวิธีการพยากรณ์ผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี โดยวิธีสมการแนวโน้มกำลังสาม ซึ่งมีตัวแบบพยากรณ์ คือ $\hat{Y}_3(t) = -0.007t^3 - 1.01t^2 + 243.173t + 43423.657$ เหมาะสำหรับการพยากรณ์ล่วงหน้า 7 เดือน เนื่องจากมีค่า MAPE ที่ต่ำที่สุด คือ 0.1638

4. วิจารณ์

การศึกษาเปรียบเทียบวิธีการพยากรณ์ 5 วิธี คือ วิธีสมการแนวโน้มเชิงเส้น วิธีสมการแนวโน้มกำลังสอง วิธีสมการแนวโน้มกำลังสาม วิธีสมการแนวโน้มเอ็กโปแนนเชียล และวิธีสมการแนวโน้มกำลังของข้อมูลจำนวนผู้ใช้น้ำประปา ของการประปาส่วนภูมิภาค สาขาปทุมธานี ตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึงเดือนกรกฎาคม พ.ศ. 2562 พบว่าอนุกรมเวลามีการเคลื่อนไหวของแนวโน้มแต่ไม่มีอิทธิพลของความแปรผันแปรตามฤดูกาล ซึ่งวิธีการพยากรณ์โดยวิธีสมการแนวโน้มกำลังสามเป็นวิธีที่เหมาะสมกับลักษณะของข้อมูลมากที่สุด แต่ผลการศึกษาของพิราวรรณ และคณะ [7] ที่ศึกษาการเปรียบเทียบตัวแบบการพยากรณ์ปริมาณการผลิตน้ำมันดิบในประเทศไทย ซึ่งลักษณะเคลื่อนไหวของอนุกรมเวลา มีการเคลื่อนไหวของแนวโน้มแต่ไม่มีอิทธิพลของฤดูกาล ผลการศึกษาพบว่าวิธีบ็อกซ์-เจนกินเหมาะสมกับการพยากรณ์ข้อมูลนี้มากที่สุด สาเหตุที่ผลการศึกษาได้วิธีการพยากรณ์ที่ต่างกันเพราะว่าลักษณะข้อมูลต่างกัน แม้ว่าข้อมูลทั้งสองจะมีการเคลื่อนไหวของข้อมูลในลักษณะเดียวกัน

5. สรุป

การวิจัยครั้งนี้พบว่าวิธีการพยากรณ์ที่มีความเหมาะสมที่สุด คือ การพยากรณ์โดยวิธีสมการแนวโน้มกำลังสาม โดยมีตัวแบบพยากรณ์ คือ $\hat{Y}_3(t) = -0.007t^3 - 1.01t^2 + 243.173t + 43423.657$ เมื่อ t แทนช่วงเวลาที่พยากรณ์ $t = 58$ (เดือนมกราคม ปี พ.ศ. 2562) ซึ่งมีค่า MAPE และค่า MAD ต่ำที่สุด จากรูปแบบดังกล่าวนำมาคำนวณหาช่วงการพยากรณ์ล่วงหน้าที่เหมาะสมที่สุด 3, 5 และ 7 เดือน พบว่าวิธีนี้เหมาะสำหรับการพยากรณ์ล่วงหน้า 7 เดือน รองลงมา คือ 5 และ 3 เดือน ตามลำดับ ซึ่งพบว่าค่า MAPE ของช่วงการพยากรณ์ล่วงหน้าทั้ง 3 ช่วง มีค่าไม่ต่างกันมาก จึงได้ว่าช่วงการพยากรณ์ทั้ง 3 ช่วง มีความเหมาะสมกับการพยากรณ์ โดยช่วงการพยากรณ์ล่วงหน้า 7 เดือน มีความเหมาะสมมากที่สุด

6. ข้อเสนอแนะ

แม้ว่า ตัวแบบพยากรณ์ โดยวิธีสมการแนวโน้ม

กำลังสามพยากรณ์จำนวนผู้ใช้น้ำประปาของการประปาส่วนภูมิภาค สาขาปทุมธานี ถึงจะให้ค่า MAPE และค่า MAD ต่ำที่สุด แต่จำนวนผู้ใช้น้ำประปาอาจไม่ได้ขึ้นอยู่กับปัจจัยเวลาเพียงอย่างเดียว ดังนั้นการศึกษาครั้งต่อไปผู้วิจัยควรพิจารณาปัจจัยอื่น ๆ ในการสร้างตัวแบบพยากรณ์ด้วย เช่น จำนวนประชากรในจังหวัดปทุมธานี และจำนวนหมู่บ้านจัดสรรภายในจังหวัดปทุมธานี รวมถึงควรพิจารณาวิธีการพยากรณ์อื่น ๆ เช่น วิธีค่าเฉลี่ยเคลื่อนที่อย่างง่าย หรือวิธีการแยกส่วนประกอบอนุกรมเวลา อีกทั้งเมื่อมีจำนวนผู้ใช้น้ำประปาที่เป็นปัจจุบันมากขึ้น ผู้วิจัยควรนำมาปรับปรุงตัวแบบพยากรณ์ เพื่อให้ได้ตัวแบบพยากรณ์ที่มีความเหมาะสมสำหรับการพยากรณ์ค่าในอนาคตต่อไป

7. References

[1] Sakulsri, T., Urban People Use 200 Liters of Water per Day, Inadequate Water Supply Risk in the Next 20 years, Available Sources: https://www.bltbangkok.com, September 13, 2019. (in Thai)

[2] Srichai, N., Kuayrakan, S. and Suwanprasit, C., 2016, Water use and water demand modeling for hotel and tourism business Patong, Phuket province, Songklanakarin J. Soc. Sci. Human. 22(2): 255-292. (in Thai)

[3] Supapim, N. and Kaenmanee, S., 2012, Forecasting future demand for water consumption using the ARIMA modal and GARCH modal, KKU Res. J. 11(1): 45-55. (in Thai)

[4] Taesombat, S., 2006, Quantitative Forecasting, Kasetsart University, Bangkok, 15 p. (in Thai)

[5] Data Center and Information Technology Planning Division, Provincial Waterworks Authority Phathum Thani Branch, Available Sources: https://www.pwa.co.th/province/branch/5540227, September 13, 2019. (in Thai)

[6] Ket-iam, S., 2003, Forecasting Techniques, Department of Mathematics faculty of Science, Thaksin University, Songkhla, 8 p. (in Thai)

[7] Noosen, P., Payakkapong, N. and Supapakorn, T., A comparison of quantity production of petroleum forecasting models in Thailand, Thai Sci. Technol. J. 23(3): 377-384. (in Thai)