

ตัวแบบการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบในประเทศไทย Forecasting Model for the Number of Patients with Pneumonia in Thailand

วารางคณา กীরตวิบูลย์

Warangkhana Keerativibool

สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ

Department of Mathematics and Statistics, Faculty of Science, Thaksin University

E-mail: warang27@gmail.com.com

บทคัดย่อ

วัตถุประสงค์ของการวิจัยครั้งนี้ คือ การสร้างตัวแบบการพยากรณ์ที่เหมาะสมที่สุดของจำนวนผู้ป่วยโรคปอดอักเสบในประเทศไทย โดยใช้ข้อมูลจากเว็บไซต์ของระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต ตั้งแต่ไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 48 ค่า ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด ชุดที่ 1 คือ ข้อมูลตั้งแต่ไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2556 จำนวน 44 ค่า สำหรับการสร้างตัวแบบการพยากรณ์ด้วยวิธีการทางสถิติ 3 วิธี ได้แก่ วิธีบอซซ์-เจนกินส์ วิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก และวิธีการพยากรณ์รวม ชุดที่ 2 คือ ข้อมูลตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า สำหรับการตรวจสอบความแม่นยำของตัวแบบการพยากรณ์ด้วยเกณฑ์เปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ต่ำที่สุด ผลการวิจัยพบว่า จากวิธีการพยากรณ์ทั้งหมดที่ได้ศึกษา วิธีการพยากรณ์รวมเป็นวิธีที่มีความเหมาะสมกับอนุกรมเวลาชุดนี้มากที่สุด ซึ่งมีตัวแบบการพยากรณ์เป็น

$$\hat{Y}_t = 0.234904 \hat{Y}_{1t} + 0.765096 \hat{Y}_{2t}$$

เมื่อ \hat{Y}_{1t} และ \hat{Y}_{2t} แทนค่าพยากรณ์เชิงเดี่ยว ณ เวลา t จากวิธีบอซซ์-เจนกินส์ และวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก ตามลำดับ

เมื่อใช้วิธีการพยากรณ์รวมในการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ พบว่า จำนวนผู้ป่วยโรคปอดอักเสบมีแนวโน้มเพิ่มขึ้น อย่างไรก็ตาม ค่าพยากรณ์ของไตรมาสที่ 1 อาจมีค่าต่ำกว่าค่าจริง ดังนั้นผู้ที่นำตัวแบบการพยากรณ์ไปใช้ประโยชน์ควรระมัดระวัง และเมื่อมีข้อมูลที่เป็นปัจจุบันควรนำมาปรับปรุงตัวแบบ รวมถึงควรพิจารณาอนุกรมเวลารายเดือนและรายสัปดาห์ประกอบการสร้างตัวแบบการพยากรณ์ด้วย

คำสำคัญ : ปอดอักเสบ บอซซ์-เจนกินส์ การปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลัง การพยากรณ์รวม
เปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย

Abstract

The purpose of this research was to construct the most suitable forecasting model for the number of patients with pneumonia in Thailand. The data gathered from the website of Social and Quality of Life Database System during the first quarter, 2003 to the fourth quarter, 2014 (48 values) were used and divided into two categories. The first category had 44 values, which were the data during the first quarter, 2003 to the fourth quarter, 2013 for the modeling by the methods of Box-Jenkins, Winters' additive exponential smoothing, and combined forecasting. The second category had 4 values, which were the data from all four quarters in 2014 for checking the accuracy of the forecasting models via the criterion of the lowest mean absolute percentage error. The results showed that for all forecasting methods that had been studied, combined forecasting method was the most suitable for this time series and the forecasting model was

$$\hat{Y}_t = 0.234904 \hat{Y}_{1t} + 0.765096 \hat{Y}_{2t}$$

where \hat{Y}_{1t} and \hat{Y}_{2t} represented the single forecasts at time t from Box-Jenkins and Winters' additive exponential smoothing, respectively.

When using the combined forecasting method to predict the number of patients with pneumonia, we found that the number of pneumonia cases increased. However, the predictions of the first quarter may be lower than the actual value, so who used the forecasting model should be careful and if there were the current time series data, the model should be updated. Including should be considering the time series of monthly and weekly in order to construct the forecasting model.

Keywords : Pneumonia, Box-Jenkins, Exponential Smoothing, Combined Forecasting, Mean Absolute Percentage Error.

บทนำ

ปอดอักเสบ เป็นโรคระบบทางเดินหายใจที่สามารถพบได้ตลอดทั้งปี โดยตั้งแต่ปี 2552 เป็นต้นมา มีการรายงานโรคปอดอักเสบมากขึ้น เนื่องจากเป็นโรคที่มีความสำคัญ และเป็นโรคหนึ่งที่ถูกกำหนดให้เป็นโรคที่มีอาการสัมพันธ์กับโรคทางเดินหายใจเฉียบพลันรุนแรง (SARS) และโรคไข้หวัดนก^{1, 2} ร้อยละ 8-10 ของผู้ป่วยโรคปอดอักเสบจะมีการติดเชื้อเฉียบพลันระบบหายใจ จึงเป็นสาเหตุการตายอันดับหนึ่งของโรคติดเชื้อในเด็กอายุต่ำกว่า 5 ปี ซึ่งสาเหตุการเกิดโรคอาจเกิดได้ทั้งจากไวรัส แบคทีเรีย และเชื้อรา แตกต่างกันไป ตามแต่ละกลุ่มอายุ มีทั้งการเกิดโรคจากการติดเชื้อและไม่ได้เกิดจากการติดเชื้อ แต่โดยทั่วไปมักพบโรคปอดอักเสบที่เกิดจากการติดเชื้อมากกว่า ชนิดของปอดอักเสบจำแนกได้หลายแบบปัจจุบันนิยมจำแนกตามสภาพแวดล้อมที่เกิดปอดอักเสบ เป็นปอดอักเสบในชุมชน (Community-Acquired Pneumonia: CAP) หมายถึง ปอดอักเสบที่เกิดจากการติดเชื้อที่เกิดขึ้นนอกโรงพยาบาลโดยไม่รวมปอดอักเสบที่เกิดขึ้นหลังจำหน่ายผู้ป่วยออกจากโรงพยาบาลภายในเวลาไม่เกิน 2 สัปดาห์ และปอดอักเสบในโรงพยาบาล (Nosocomial Pneumonia หรือ Hospital-Acquired Pneumonia: HAP) หมายถึง ปอดอักเสบจากการติดเชื้อที่เกิดขึ้นหลังจากผู้ป่วยนอนรักษาในโรงพยาบาลแล้วอย่างน้อย 48-72 ชั่วโมง³ เนื่องจากโรคปอดอักเสบเป็นโรคที่มีปัจจัยเสี่ยงหลายด้าน ดังนั้นแนวคิดด้านการรายงานสถานการณ์โรคจึงเปลี่ยนไปสู่การสร้างสรรความคิดโดยการรวบรวมและวิเคราะห์ข้อมูลจากแหล่งข้อมูลต่างๆ เพื่อประโยชน์ในการป้องกันและเตือนภัยในเหตุการณ์ต่างๆ มากขึ้นซึ่งนำไปสู่การทำนาย (Prediction) หรือการพยากรณ์ (Forecast) ที่อาจเกิดขึ้นในอนาคตรวมทั้งการประเมินความเสี่ยงทางสุขภาพ (Health Risk Assessment) ที่ใช้เป็นเครื่องมือ

คาดการณ์ความเสี่ยงที่มีโอกาสเกิดขึ้น ด้วยเหตุผลดังกล่าว ผู้วิจัยจึงมีความสนใจที่จะศึกษาการสร้างตัวแบบการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบโดยใช้วิธีการทางสถิติทั้งหมด 3 วิธี ได้แก่ วิธีบอซซ์-เจนกินส์ วิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก และวิธีการพยากรณ์รวมเพื่อเป็นเครื่องมือช่วยให้นักวิจัยที่เกี่ยวข้องสามารถตัดสินใจในการกำหนดนโยบาย กำหนดแผนงาน และวางกลยุทธ์ในการป้องกันผลกระทบที่อาจเกิดขึ้นต่อไป

วิธีการศึกษา

การวิจัยครั้งนี้ดำเนินการสร้างตัวแบบการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ โดยใช้โปรแกรม SPSS (Statistical Package for Social Sciences) รุ่น 17 ซึ่งข้อมูลที่ใช้ในการศึกษาได้มาจากเว็บไซต์ของระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต⁴ เป็นอนุกรมเวลาจำนวนผู้ป่วยโรคปอดอักเสบ (คน) ตั้งแต่ไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 48 ค่า ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด ชุดที่ 1 ตั้งแต่ไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2556 จำนวน 44 ค่า สำหรับการสร้างตัวแบบการพยากรณ์ด้วยวิธีการทางสถิติ 3 วิธีที่มีความเหมาะสมกับอนุกรมเวลานี้มากที่สุดได้แก่วิธีบอซซ์-เจนกินส์วิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก และวิธีการพยากรณ์รวม เนื่องจากได้พิจารณาจากค่าเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) ของข้อมูลชุดที่ 1 แล้วพบว่า วิธีเหล่านี้เป็นวิธีที่มีความเหมาะสมกับอนุกรมเวลานี้มากกว่าวิธีการพยากรณ์อื่นๆ ข้อมูลชุดที่ 2 ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า นำมาใช้สำหรับการตรวจสอบความแม่นยำ (Accuracy) ของตัวแบบการพยากรณ์ด้วยเกณฑ์เปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ต่ำที่สุด

1. การสร้างตัวแบบการพยากรณ์โดยวิธีบอกซ์-เจนกินส์ (Box-Jenkins Method)

วิธีบอกซ์-เจนกินส์ เป็นวิธีการพยากรณ์ที่มีความถูกต้องสูง เนื่องจากได้กำหนดตัวแบบโดยการตรวจสอบคุณสมบัติของฟังก์ชันอัตโนมัติสหสัมพันธ์ (Auto correlation Function: ACF) และฟังก์ชันอัตโนมัติสหสัมพันธ์บางส่วน (Partial Auto correlation Function: PACF) ซึ่งพิจารณาภายใต้อนุกรมเวลาที่คงที่ (Stationary) หรืออนุกรมเวลาที่มีค่าเฉลี่ยและความแปรปรวนคงตัว⁴ ตัวแบบทั่วไป (General Model) ของวิธีบอกซ์-เจนกินส์ คือ Seasonal Autoregressive Integrated Moving Average: SARIMA(p, d, q)(P, D, Q)_s แสดงดังสมการที่ (1)⁵

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

เมื่อ Y_t แทนอนุกรมเวลา ณ เวลา t

ε_t แทนอนุกรมเวลาของค่าคลาดเคลื่อนที่มีการแจกแจงปกติและเป็นอิสระกัน ด้วยค่าเฉลี่ยเท่ากับศูนย์ และความแปรปรวนคงตัวทุกช่วงเวลา

$\delta = \mu\phi_p(B)\Phi_p(B^s)$ แทนค่าคงตัวโดยที่แทนค่าเฉลี่ยของอนุกรมเวลาที่คงที่

$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ แทนตัวดำเนินการอัตโนมัติสหสัมพันธ์แบบไม่มีฤดูกาลอันดับที่ p (Non-Seasonal Autoregressive Operator of Order p : AR(p))

$\Phi_p(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}$ แทนตัวดำเนินการอัตโนมัติสหสัมพันธ์แบบมีฤดูกาลอันดับที่ P (Seasonal Autoregressive Operator of Order P : SAR(P))

$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ แทนตัวดำเนินการเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาลอันดับที่ q (Non-Seasonal Moving Average Operator of Order q : MA(q))

$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}$$

แทนตัวดำเนินการเฉลี่ยเคลื่อนที่แบบมีฤดูกาลอันดับที่ Q (Seasonal Moving Average Operator of Order Q : SMA(Q))

t แทนช่วงเวลา ซึ่งมีค่าตั้งแต่ 1 ถึง n โดยที่ n แทนจำนวนข้อมูลในอนุกรมเวลาชุดที่ 1

s แทนจำนวนฤดูกาล

d และ D แทนลำดับที่ของผลต่างและผลต่างฤดูกาล ตามลำดับ

B แทนตัวดำเนินการย้อนหลัง (Backward Operator) โดยที่

ขั้นตอนการสร้างตัวแบบการพยากรณ์โดยวิธีบอกซ์-เจนกินส์แสดงรายละเอียดดังนี้

1. พิจารณาอนุกรมเวลาว่าคงที่หรือไม่ โดยพิจารณาจากกราฟของอนุกรมเวลาเทียบกับเวลา (Y_t, t) กราฟ ACF และ PACF หากพบว่าอนุกรมเวลาไม่คงที่ (Non-Stationary) ต้องแปลงอนุกรมเวลาให้คงที่ก่อนที่จะทำขั้นตอนต่อไป เช่น การแปลงข้อมูลด้วยผลต่างหรือผลต่างฤดูกาล (Difference or Seasonal Difference) การแปลงข้อมูลด้วยลอการิทึมสามัญหรือลอการิทึมธรรมชาติ (Common Logarithm or Natural Logarithm) การแปลงข้อมูลด้วยเลขยกกำลัง เช่น ยกกำลัง 0.5 (Square Root Transformation) หรือยกกำลัง 2 (Square Transformation) เป็นต้น⁶

2. กำหนดตัวแบบการพยากรณ์ที่เป็นไปได้จากกราฟ ACF และ PACF ของอนุกรมเวลาที่คงที่ นั่นคือ กำหนดค่า p, q, P และ Q พร้อมทั้งประมาณค่าพารามิเตอร์ของตัวแบบ

3. ตัดพารามิเตอร์ที่ไม่มีนัยสำคัญออกจากตัวแบบการพยากรณ์ครั้งละ 1 ตัว จากนั้นจึงกำหนดตัวแบบการพยากรณ์และประมาณค่าพารามิเตอร์ใหม่จนกว่าจะได้ตัวแบบการพยากรณ์ที่ประกอบด้วยพารามิเตอร์ที่มีนัยสำคัญทั้งหมด

4. คัดเลือกตัวแบบการพยากรณ์ที่มีค่าเกณฑ์สารสนเทศเบย์เซียน (Bayesian Information Criterion: BIC) ที่ต่ำที่สุด มีค่าสถิติ Ljung-Box Q

ที่ไม่มีนัยสำคัญ และอนุกรมเวลาของค่าคลาดเคลื่อนจากการพยากรณ์มีการแจกแจงปกติ ตรวจสอบโดยใช้การทดสอบโคลโมโกรอฟ-สเมียร์นอฟ (Kolmogorov-Smirnov's Test) มีการเคลื่อนไหวเป็นอิสระกัน ตรวจสอบโดยพิจารณาจากกราฟ ACF และ PACF ของค่าคลาดเคลื่อน มีค่าเฉลี่ยเท่ากับศูนย์ ตรวจสอบโดยใช้การทดสอบที (t-Test) มีความแปรปรวนคงตัวทุกช่วงเวลา ตรวจสอบโดยใช้การทดสอบของเลวีเนียนภายใต้การใช้ค่ามัธยฐาน (Levene's Test based on Median)

5. พยากรณ์อนุกรมเวลา โดยใช้ตัวแบบการพยากรณ์ที่เหมาะสมที่สุดจากขั้นตอนที่ 4

2. การสร้างตัวแบบการพยากรณ์โดยวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก (Winters' Additive Exponential Smoothing Method)

การปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์จัดเป็นเทคนิคการวิเคราะห์ข้อมูลโดยนำอนุกรมเวลาจากอดีตมาวิเคราะห์หาค่าประมาณเพื่อกำหนดตัวแบบจำลองสำหรับการพยากรณ์ค่าอนาคต วิธีนี้มีความเหมาะสมกับอนุกรมเวลาที่มีแนวโน้มเชิงเส้นและความผันแปรตามฤดูกาลซึ่งแบ่งออกได้เป็น 2 กรณี คือ การปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก (Winters' Additive Exponential Smoothing) ควรใช้กับการพยากรณ์อนุกรมเวลาที่มีความผันแปรตามฤดูกาลคงตัว กล่าวคือ ความผันแปรตามฤดูกาลมีค่าไม่เพิ่มขึ้นและไม่ลดลงตามเวลาที่เปลี่ยนแปลงไปและการปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ (Winters' Multiplicative Exponential Smoothing) ควรใช้กับการพยากรณ์อนุกรมเวลาที่มีความผันแปรตามฤดูกาลเพิ่มขึ้นหรือลดลงตามเวลาที่เปลี่ยนแปลงไป สำหรับการวิจัยครั้งนี้ได้ใช้วิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก เนื่องจากอนุกรมเวลาจำนวน

ผู้ป่วยโรคปอดอักเสบของข้อมูลชุดที่ 1 ในช่วงไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2556 มีความผันแปรตามฤดูกาลคงตัวตามเวลาที่เปลี่ยนแปลงไป (แสดงรายละเอียดในภาพที่ 1) ตัวแบบแสดงดังสมการที่ (2) และตัวแบบการพยากรณ์แสดงดังสมการที่ (3)⁸

$$Y_t = (\beta_0 + \beta_1 t) + S_t + \varepsilon_t \quad (2)$$

$$\hat{Y}_{t+m} = (a_t + b_t m) + \hat{S}_t \quad (3)$$

เมื่อ Y_t แทนอนุกรมเวลา ณ เวลา t

β_0 , β_1 และ S_t แทนระดับของอนุกรมเวลา ความชันของแนวโน้ม และความผันแปรตามฤดูกาลตามลำดับ

ε_t แทนอนุกรมเวลาของค่าคลาดเคลื่อนที่มีการแจกแจงปกติและเป็นอิสระกัน ด้วยค่าเฉลี่ยเท่ากับศูนย์ และความแปรปรวนคงตัวทุกช่วงเวลา

\hat{Y}_{t+m} แทนค่าพยากรณ์ ณ เวลา $t+m$ โดยที่ m แทนจำนวนช่วงเวลาที่ต้องการพยากรณ์ไปข้างหน้า

a_t , b_t และ \hat{S}_t แทนค่าประมาณ ณ เวลา t ของ β_0 , β_1 และ S_t ตามลำดับ

$$\text{โดยที่ } a_t = \alpha (Y_t - \hat{S}_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \gamma (a_t - a_{t-1}) + (1-\gamma) b_{t-1}$$

$$\hat{S}_t = \delta (Y_t - a_t) + (1-\delta) \hat{S}_{t-s}$$

α , γ และ δ แทนค่าคงตัวการปรับให้เรียบ โดยที่ $0 < \alpha < 1$, $0 < \gamma < 1$ และ $0 < \delta < 1$

t แทนช่วงเวลา ซึ่งมีค่าตั้งแต่ 1 ถึง n โดยที่ n แทนจำนวนข้อมูลในอนุกรมเวลาชุดที่ 1 s แทนจำนวนฤดูกาล

เมื่อได้ตัวแบบการพยากรณ์แล้วจะดำเนินการตรวจสอบคุณลักษณะของค่าคลาดเคลื่อนจากการพยากรณ์ตามขั้นตอนที่ 4 ของการสร้างตัวแบบการพยากรณ์โดยวิธีบ็อกซ์-เจนกินส์

3. การสร้างตัวแบบการพยากรณ์โดยวิธีการพยากรณ์รวม (Combined Forecasting Method)

การพยากรณ์รวมเป็นวิธีการประยุกต์ที่มีการรวมค่าพยากรณ์จากวิธีการพยากรณ์เชิงเดี่ยวตั้งแต่ 2 วิธีขึ้นไป เพื่อให้ได้ค่าพยากรณ์ใหม่ที่มีค่าคลาดเคลื่อนน้อยที่สุด สามารถใช้ได้ทั้งกรณีที่วิธีการพยากรณ์เชิงเดี่ยวมีความเหมาะสมกับอนุกรมเวลามากกว่า 1 วิธี ในที่นี้ได้พิจารณาวิธีการพยากรณ์เชิงเดี่ยว 2 วิธี คือ วิธีบ็อกซ์-เจนกินส์ และวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก เนื่องจากค่าพยากรณ์ของวิธีการเหล่านี้มีความสัมพันธ์กันอย่างมากกับอนุกรมเวลาชุดที่ 1 (สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient: r) ระหว่างจำนวนผู้ป่วยโรคปอดอักเสบในอนุกรมเวลาชุดที่ 1 กับค่าพยากรณ์โดยวิธีบ็อกซ์-เจนกินส์มีค่าเท่ากับ 0.8357 และสัมประสิทธิ์สหสัมพันธ์ระหว่างจำนวนผู้ป่วยโรคปอดอักเสบในอนุกรมเวลาชุดที่ 1 กับค่าพยากรณ์โดยวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก มีค่าเท่ากับ 0.8780) ดังนั้นตัวแบบของวิธีการพยากรณ์รวมที่ใช้ในการวิจัยครั้งนี้คือ

$$\hat{Y}_t = w_1 \hat{Y}_{1t} + w_2 \hat{Y}_{2t} \quad (4)$$

เมื่อ \hat{Y}_t แทนค่าพยากรณ์รวม ณ เวลา t
 \hat{Y}_{1t} และ \hat{Y}_{2t} แทนค่าพยากรณ์เชิงเดี่ยว ณ เวลา t จากวิธีบ็อกซ์-เจนกินส์ และวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวกตามลำดับ

w_1 และ w_2 แทนค่าถ่วงน้ำหนักของวิธีบ็อกซ์-เจนกินส์ และวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก ตามลำดับโดยที่

$$w_1 = \frac{b_1}{b_1 + b_2} \text{ และ } w_2 = \frac{b_2}{b_1 + b_2} \quad (5)$$

b_1 และ b_2 แทนค่าสัมประสิทธิ์การถดถอยจากวิธีกำลังสองน้อยที่สุด (Least Squares Method)¹⁰ ของวิธีบ็อกซ์-เจนกินส์ และวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก ตามลำดับ เมื่อกำหนดให้ค่าพยากรณ์เชิงเดี่ยวจากทั้ง 2 วิธีเป็นตัวแปรอิสระ และจำนวนผู้ป่วยโรคปอดอักเสบเป็นตัวแปรตาม ซึ่งค่า b_1 และ b_2 จะคำนวณจากจำนวนข้อมูลพยากรณ์ในอนุกรมเวลาชุดที่ 1 ณ ที่นี้คือ 39 ค่า เนื่องจากการแปลงข้อมูลด้วยผลต่างและผลต่างฤดูกาลลำดับที่ 1 ของวิธีบ็อกซ์-เจนกินส์ ทำให้ไม่มีค่าพยากรณ์ 5 ค่าแรก (จำนวนฤดูกาลเท่ากับ 4)

เมื่อได้ตัวแบบการพยากรณ์แล้วจะดำเนินการตรวจสอบคุณลักษณะของค่าคลาดเคลื่อนจากการพยากรณ์ตามขั้นตอนที่ 4 ของการสร้างตัวแบบการพยากรณ์โดยวิธีบ็อกซ์-เจนกินส์

4. การตรวจสอบความแม่นยำของตัวแบบการพยากรณ์

การวิจัยครั้งนี้ได้ตรวจสอบความแม่นยำของตัวแบบการพยากรณ์จากวิธีการทั้งหมด 3 วิธี คือ วิธีบ็อกซ์-เจนกินส์ วิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก และวิธีการพยากรณ์รวม โดยทำการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบของข้อมูลชุดที่ 2 คือ อนุกรมเวลาตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า ได้ค่าความแตกต่างระหว่างข้อมูลจริงกับค่าพยากรณ์ (Error: e_t) เพื่อคำนวณค่าเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) มีสูตรแสดงดังสมการที่ (6)⁸ โดยวิธีการพยากรณ์ใดมีค่า

เปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด คือ วิธีที่มีความแม่นยำในการพยากรณ์มากที่สุด

$$MAPE = \frac{100}{n_2} \sum_{t=1}^{n_2} \left| \frac{e_t}{Y_t} \right| \quad (6)$$

เมื่อ $e_t = Y_t - \hat{Y}_t$ แทนค่าคลาดเคลื่อนจากการพยากรณ์ ณ เวลา t

Y_t แทนอนุกรมเวลา ณ เวลา t

\hat{Y}_t แทนค่าพยากรณ์ ณ เวลา t

t แทนช่วงเวลา ซึ่งมีค่าตั้งแต่ 1 ถึง n_2 โดยที่

n_2 แทนจำนวนข้อมูลในอนุกรมเวลาชุดที่ 2

5. การพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ

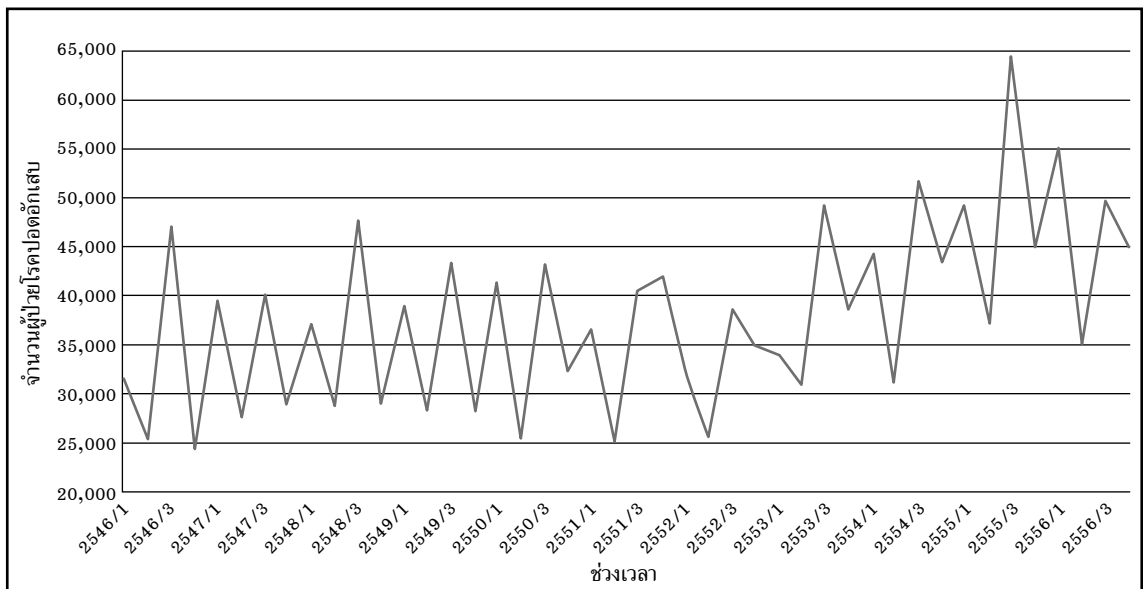
จากการตรวจสอบความแม่นยำของตัวแบบการพยากรณ์ทั้ง 3 วิธี คือ วิธีบอชซ์-เจนกินส์ วิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์ แบบบวก และวิธีการพยากรณ์รวม เมื่อทราบว่าตัวแบบ

การพยากรณ์ใดมีค่าเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุดจะใช้ตัวแบบการพยากรณ์นั้น สำหรับการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2558 ต่อไป

ผลการศึกษา

1. ผลการสร้างตัวแบบการพยากรณ์โดยวิธีบอชซ์-เจนกินส์

จากการพิจารณาลักษณะการเคลื่อนไหวของอนุกรมเวลาชุดที่ 1 คือ จำนวนผู้ป่วยโรคปอดอักเสบ ตั้งแต่ไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2556 จำนวน 44 ค่า ดังภาพที่ 1 พบว่า อนุกรมเวลาชุดนี้มีแนวโน้มคงตัวในช่วงปี 2546 ถึงปี 2551 และตั้งแต่ปี 2552 เป็นต้นมา พบว่า แนวโน้มเพิ่มขึ้นเล็กน้อย แต่ยังคงมีความผันแปรตามฤดูกาลคงตัว กล่าวคือ ความผันแปรตามฤดูกาลมีค่าไม่เพิ่มขึ้นและไม่ลดลงตามเวลาที่เปลี่ยนแปลงไป



ภาพที่ 1 ลักษณะการเคลื่อนไหวของอนุกรมเวลาจำนวนผู้ป่วยโรคปอดอักเสบตั้งแต่ไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2556

จากกราฟ ACF และ PACF ดังภาพที่ 2 พบว่า อนุกรมเวลายังไม่คงที่ ดังนั้นผู้วิจัยจึงแปลงข้อมูลด้วยผลต่างและผลต่างฤดูกาลลำดับที่ 1 ($d=1, D=1, s=4$) ได้กราฟ ACF และ PACF ของอนุกรมเวลาที่แปลงข้อมูลแล้ว แสดงดังภาพที่ 3 ซึ่งพบว่า อนุกรมเวลามีลักษณะคงที่ จึงกำหนดตัวแบบการพยากรณ์ที่เป็นไปได้ พร้อมกับประมาณค่าพารามิเตอร์ ดังแสดงในตารางที่ 1 โดยตัวแบบการพยากรณ์ที่มีค่า BIC ต่ำที่สุด และมีค่าสถิติ Ljung-Box Q ไม่มีนัยสำคัญที่ระดับ 0.05 คือตัวแบบ SARIMA(1,1,0)(1,1,0)₄ ไม่มีพจน์ค่าคงตัว เมื่อตรวจสอบคุณลักษณะของค่าคลาดเคลื่อนจากการพยากรณ์ พบว่า ค่าคลาดเคลื่อนมีการแจกแจงปกติ (Kolmogorov-Smirnov $Z = 0.520$, $p\text{-value} = 0.949$) มีการเคลื่อนไหวเป็นอิสระกัน (แสดงรายละเอียดในภาพที่ 4 ซึ่งพบว่า ค่าสัมประสิทธิ์อัตโนมัติสัมพันธ์และสัมประสิทธิ์อัตโนมัติสัมพันธ์บางส่วนของค่าคลาดเคลื่อนตกอยู่ในขอบเขตความเชื่อมั่นร้อยละ 95) มีค่าเฉลี่ยเท่ากับศูนย์ ($t = -0.649$, $p\text{-value} = 0.520$) และมีความแปรปรวนคงตัวทุกช่วงเวลา (Levene Statistic = 0.774, $p\text{-value} = 0.517$) ดังนั้นตัวแบบ SARIMA (1, 1, 0)(1, 1, 0)₄ ไม่มีพจน์ค่าคงตัว มีความเหมาะสม ซึ่งจากสมการที่ (1) สามารถเขียนเป็นตัวแบบได้ดังนี้

$$(1-\phi_1 B)(1-\Phi_1 B^4)(1-B)(1-B^4)Y_t = \varepsilon_t$$

$$(1-\Phi_1 B^4 - \phi_1 B + \phi_1 \Phi_1 B^5)(1-B^4 - B + B^5)Y_t = \varepsilon_t$$

$$(1 - B^4 - B + B^5 - \Phi_1 B^4 + \Phi_1 B^8 + \Phi_1 B^5 - \Phi_1 B^9 - \phi_1 B + \phi_1 B^5 + \phi_1 B^2 - \phi_1 B^6 + \phi_1 \Phi_1 B^5 - \phi_1 \Phi_1 B^9 - \phi_1 \Phi_1 B^6 + \phi_1 \Phi_1 B^{10}) Y_t = \varepsilon_t$$

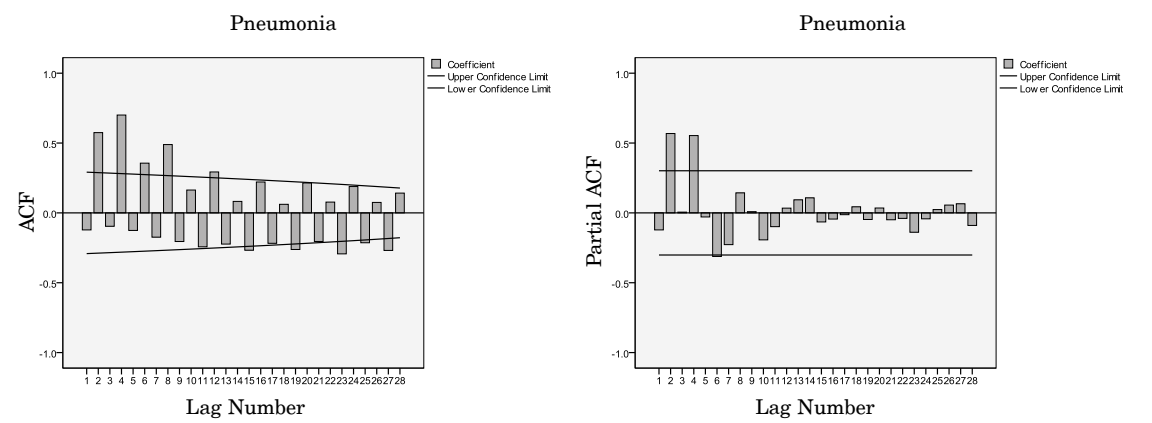
$$Y_t = (1+\phi_1) Y_{t-1} - \phi_1 Y_{t-2} + (1+\Phi_1) Y_{t-4} - (1+\phi_1) Y_{t-5} - (1+\phi_1) \Phi_1 Y_{t-5} + (1+\Phi_1) \phi_1 Y_{t-6} - \Phi_1 Y_{t-8} + (1+\phi_1) \Phi_1 Y_{t-9} - \phi_1 \Phi_1 Y_{t-10} + \varepsilon_t$$

$$Y_t = (1+\phi_1)(Y_{t-1} - Y_{t-5}) - \phi_1 Y_{t-2} + (1+\Phi_1) Y_{t-4} - (1+\phi_1) \Phi_1 (Y_{t-5} - Y_{t-9}) + (1+\Phi_1) \phi_1 Y_{t-6} - \Phi_1 Y_{t-8} - \phi_1 \Phi_1 Y_{t-10} + \varepsilon_t$$

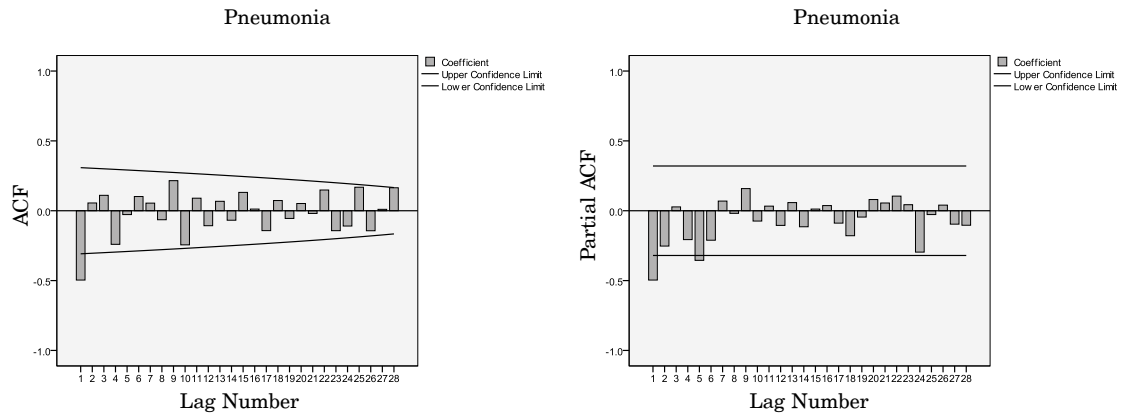
เมื่อแทนค่าประมาณพารามิเตอร์จากตารางที่ 1 จะได้ตัวแบบการพยากรณ์แสดงดังนี้

$$\hat{Y}_t = 0.342865(Y_{t-1} - Y_{t-5}) + 0.657135Y_{t-2} + 0.423567Y_{t-4} + 0.197639(Y_{t-5} - Y_{t-9}) - 0.278341Y_{t-6} + 0.576433Y_{t-8} - 0.378794Y_{t-10}$$

เมื่อ \hat{Y}_t แทนค่าพยากรณ์ ณ เวลา t
 \hat{Y}_{t-j} แทนอนุกรมเวลา ณ เวลา $t - j$



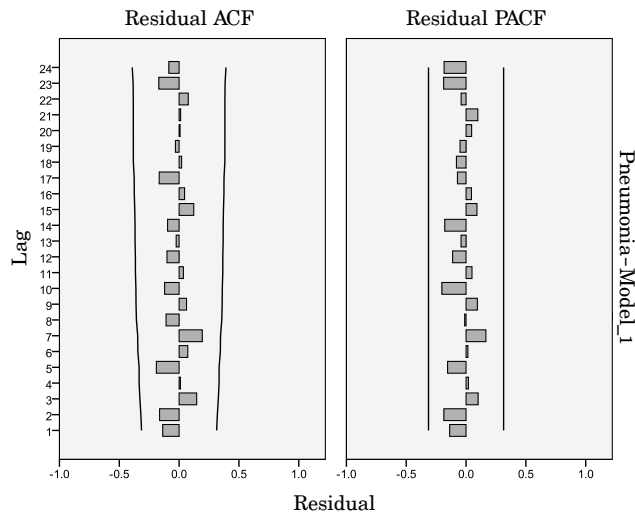
ภาพที่ 2 กราฟ ACF และ PACF ของอนุกรมเวลาจำนวนผู้ป่วยโรคปอดอักเสบ



ภาพที่ 3 กราฟ ACF และ PACF ของอนุกรมเวลาจำนวนผู้ป่วยโรคปอดอักเสบเมื่อแปลงข้อมูลด้วยผลต่างและผลต่างฤดูกาลลำดับที่ 1

ตารางที่ 1 ค่าประมาณพารามิเตอร์ ค่า BIC และค่าสถิติ Ljung-Box Q ของตัวแบบ SARIMA (p, d, q) (P, D, Q)_s

		SARIMA(p, d, q)(P, D, Q) _s		
ค่าประมาณพารามิเตอร์		SARIMA (1, 1, 1)(1, 1, 0) ₄	SARIMA (1, 1, 1)(1, 1, 0) ₄ ไม่มีพจน์ค่าคงตัว	SARIMA (1, 1, 0)(1, 1, 0) ₄ ไม่มีพจน์ค่าคงตัว
ค่าคงตัว	ค่าประมาณ	-87.7895		
	p-value	0.718	-	-
AR(1):	ค่าประมาณ	-0.443162	-0.437699	-0.657135
	ϕ_1 p-value	0.074	0.074	0.000
MA(1):	ค่าประมาณ	0.401957	0.408046	
	Θ_1 p-value	0.124	0.113	-
SAR(1):	ค่าประมาณ	-0.586537	-0.589691	-0.576433
	Φ_1 p-value	0.001	0.001	0.001
BIC		17.618	17.501	17.447
Ljung-Box Q (ณ lag 18)		9.467	9.542	12.855
p-value		0.852	0.847	0.683



ภาพที่ 4 กราฟ ACF และ PACF ของค่าคลาดเคลื่อนจากการพยากรณ์ โดยวิธีบอกซ์-เจนกินส์ ที่มีตัวแบบ SARIMA (1, 1, 0)(1, 1, 0) ไม่มีพจน์ค่าคงตัว

2. ผลการสร้างตัวแบบการพยากรณ์ โดยวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของ วินเทอร์แบบบวก

จากการสร้างตัวแบบการพยากรณ์ โดยวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของ วินเทอร์แบบบวกพบว่า BIC มีค่าเท่ากับ 17.099 และมีค่าสถิติ Ljung-Box Q ไม่มีนัยสำคัญ ที่ระดับ 0.05 (Ljung-Box Q ณ lag 18=15.522, p-value=0.415) เมื่อตรวจสอบคุณลักษณะ ของค่าคลาดเคลื่อนจากการพยากรณ์ พบว่า ค่า คลาดเคลื่อนมีการแจกแจงปกติ (Kolmogorov-Smirnov Z = 0.531, p-value = 0.941) มีการ เคลื่อนไหวเป็นอิสระกัน (แสดงรายละเอียดใน ภาพที่ 5 ซึ่งพบว่าค่าสัมประสิทธิ์อัตโนมัติและ สัมประสิทธิ์อัตโนมัติบางส่วนของค่าคลาด เคลื่อนตกอยู่ในขอบเขตความเชื่อมั่นร้อยละ 95) มีค่าเฉลี่ยเท่ากับศูนย์ ($t = -0.003$, p-value = 0.998)

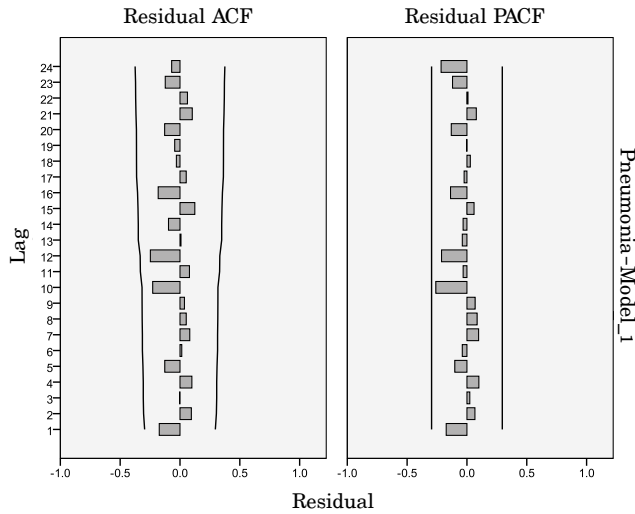
และมีความแปรปรวนคงตัวทุกช่วงเวลา (Levene Statistic = 0.697, p-value = 0.560) ดังนั้น ตัวแบบการพยากรณ์ที่ได้มีความเหมาะสม ตัวแบบ การพยากรณ์แสดงดังนี้

$$\hat{Y}_{t+m} = 47,131.268895 + 355.118517(m) + \hat{S}_t \quad (8)$$

เมื่อ \hat{Y}_{t+m} แทนค่าพยากรณ์ ณ เวลา $t + m$ โดยที่ $m = 1$ ถึง 4 (ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า)

S_t แทนดัชนีฤดูกาล รายละเอียดแสดงดัง ตารางที่ 2 ซึ่งสามารถอธิบายได้ว่า จำนวนผู้ป่วย โรคปอดอักเสบของไตรมาสที่ 1 และ 3 ของทุกปี มีค่ามากกว่าไตรมาสอื่น ๆ เนื่องจากมีค่าดัชนีฤดูกาล มากกว่า 0

α , γ และ δ มีค่าเท่ากับ 0.299307, 0.000006 และ 0.000007 ตามลำดับ



ภาพที่ 5 กราฟ ACF และ PACF ของค่าคลาดเคลื่อนจากการพยากรณ์ โดยวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก

ตารางที่ 2 ดัชนีฤดูกาลของอนุกรมเวลาจำนวนผู้ป่วยโรคปอดอักเสบ จากวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก

ไตรมาสที่	ดัชนีฤดูกาล	ไตรมาสที่	ดัชนีฤดูกาล
1	2,675.1099	3	8,914.4797
2	-8,680.2052	4	-2,909.3806

3. ผลการสร้างตัวแบบการพยากรณ์โดยวิธีการพยากรณ์รวม

จากการประมาณค่าสัมประสิทธิ์การถดถอยของวิธีพยากรณ์เดี่ยวทั้ง 2 วิธี ด้วยวิธีกำลังสองน้อยที่สุดได้ว่า $b_1 = 0.233453$ และ $b_2 = 0.760371$ ดังนั้น จากสมการที่ (5) สามารถคำนวณค่าถ่วงน้ำหนักของแต่ละวิธีการพยากรณ์เชิงเดี่ยวได้เป็น $w_1 = 0.234904$ และ $w_2 = 0.765096$ เพราะฉะนั้นตัวแบบการพยากรณ์รวมเขียนได้ดังนี้

$$\hat{Y}_t = 0.234904\hat{Y}_{1t} + 0.765096\hat{Y}_{2t} \quad (9)$$

เมื่อ \hat{Y}_t แทนค่าพยากรณ์รวม ณ เวลา t
 \hat{Y}_{1t} และ \hat{Y}_{2t} แทนค่าพยากรณ์เชิงเดี่ยว ณ เวลา t จากวิธีบอกซ์-เจนกินส์ และวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก ตามลำดับ

ผลการตรวจสอบคุณลักษณะของค่าคลาดเคลื่อนจากการพยากรณ์โดยใช้ตัวแบบการพยากรณ์รวม พบว่าค่าคลาดเคลื่อนมีการแจกแจงปกติ (Kolmogorov-Smirnov $Z = 0.432$, $p\text{-value} = 0.992$) มีการเคลื่อนไหวเป็นอิสระกัน (จาก Runs Test: $Z = 0.004$, $p\text{-value} = 0.997$)

มีค่าเฉลี่ยเท่ากับศูนย์ ($t = 0.082$, $p\text{-value} = 0.935$) และมีความแปรปรวนคงตัวทุกช่วงเวลา (Levene Statistic = 0.760, $p\text{-value} = 0.524$) ดังนั้น ตัวแบบการพยากรณ์รวมที่ได้มีความเหมาะสม

4. ผลการตรวจสอบความแม่นยำของตัวแบบการพยากรณ์

จากการใช้ตัวแบบการพยากรณ์โดยวิธีบอกซ์-เจนกินส์ ในสมการที่ (7) โดยวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวกในสมการที่ (8) และโดยวิธีการพยากรณ์รวม

ในสมการที่ (9) สำหรับการพยากรณ์ข้อมูลชุดที่ 2 คือ จำนวนผู้ป่วยโรคปอดอักเสบ ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า ได้ค่าพยากรณ์และค่าเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยแสดงดังตารางที่ 3 ผลการตรวจสอบพบว่า จากวิธีการพยากรณ์ทั้งหมดที่ได้ศึกษา วิธีการพยากรณ์รวมเป็นวิธีที่มีความแม่นยำในการพยากรณ์มากที่สุด เนื่องจากให้ค่าพยากรณ์ที่มีความแตกต่างกับข้อมูลจริงน้อยที่สุด หรือมีค่าเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด จึงมีความเหมาะสมสำหรับการพยากรณ์

ตารางที่ 3 ค่าจริงและค่าพยากรณ์ของจำนวนผู้ป่วยโรคปอดอักเสบ (คน) ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 และค่าเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย

ช่วงเวลา	จำนวนผู้ป่วยโรคปอดอักเสบ	จำนวนผู้ป่วยโรคปอดอักเสบ จากการพยากรณ์โดยวิธี		
		บอกซ์-เจนกินส์	วินเทอร์	พยากรณ์รวม
ไตรมาสที่ 1 ปี 2557	60,625	47,390	50,161	49,510
ไตรมาสที่ 2 ปี 2557	38,264	35,437	39,161	38,286
ไตรมาสที่ 3 ปี 2557	56,840	55,062	57,111	56,630
ไตรมาสที่ 4 ปี 2557	45,083	43,171	45,642	45,062
MAPE		9.1471	5.3303	4.7024

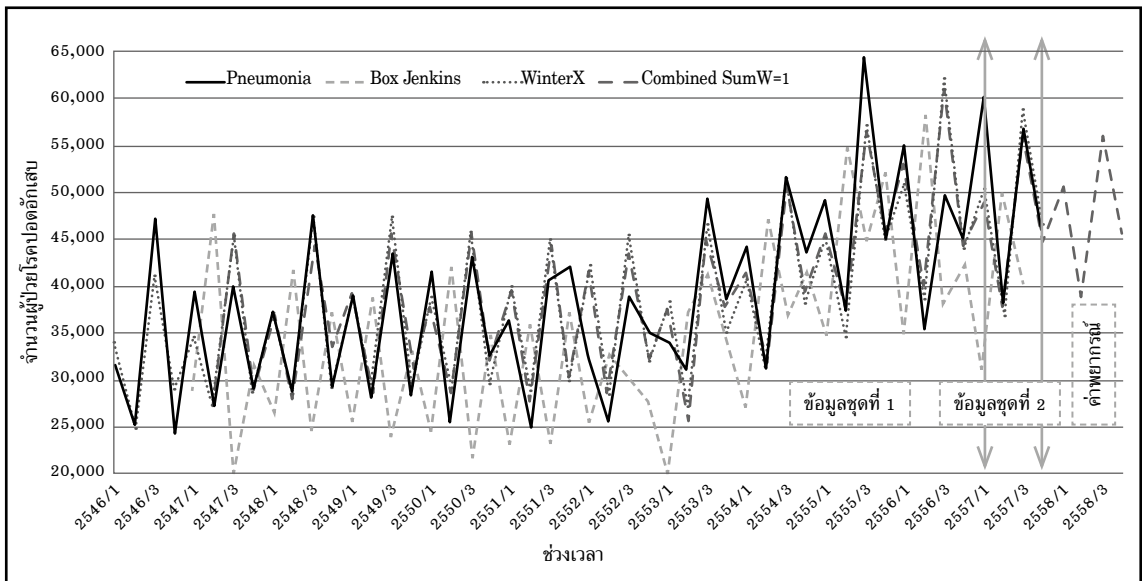
5. ผลการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ

จากการตรวจสอบความแม่นยำของตัวแบบการพยากรณ์ซึ่งพบว่า วิธีการพยากรณ์รวมมีความ

เหมาะสมกับอนุกรมเวลาชุดนี้มากที่สุด ดังนั้นผู้วิจัยจึงเลือกใช้วิธีการดังกล่าวในการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2558 รายละเอียดแสดงดังตารางที่ 4 และภาพที่ 6

ตารางที่ 4 ค่าพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบ (คน) ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2558

ช่วงเวลา	ค่าพยากรณ์	ช่วงเวลา	ค่าพยากรณ์
ไตรมาสที่ 1 ปี 2558	51,038	ไตรมาสที่ 3 ปี 2558	56,414
ไตรมาสที่ 2 ปี 2558	38,860	ไตรมาสที่ 4 ปี 2558	45,875



ภาพที่ 6 การเปรียบเทียบอนุกรมเวลาจำนวนผู้ป่วยโรคปอดอักเสบและค่าพยากรณ์จากวิธีการทางสถิติ 3 วิธี

สรุปและอภิปรายผล

การวิจัยครั้งนี้ได้นำเสนอวิธีการสร้างและคัดเลือกตัวแบบการพยากรณ์ที่เหมาะสมกับอนุกรมเวลาจำนวนผู้ป่วยโรคปอดอักเสบ โดยใช้อุณหภูมิต่ำรายไตรมาสจากเว็บไซต์ของระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต ตั้งแต่ไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 48 ค่า ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด ข้อมูลชุดที่ 1 ตั้งแต่ไตรมาสที่ 1 ปี 2546 ถึงไตรมาสที่ 4 ปี 2556 จำนวน 44 ค่า สำหรับการสร้างตัวแบบการพยากรณ์ด้วยวิธีการทางสถิติ 3 วิธี ได้แก่ วิธีบอกซ์-เจนกินส์ วิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก และวิธีการพยากรณ์รวม ข้อมูลชุดที่ 2 ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า สำหรับการตรวจสอบความแม่นยำของตัวแบบการพยากรณ์ด้วยเกณฑ์เปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ต่ำที่สุด ผลการวิจัยพบว่า จากวิธีการพยากรณ์ทั้งหมดที่ได้ศึกษา วิธีการพยากรณ์รวมเป็นวิธีที่มีความแม่นยำในการพยากรณ์มากที่สุด เนื่องจากให้ค่าพยากรณ์ที่มีความแตกต่างกับข้อมูลจริงน้อยที่สุด หรือมีค่าเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์

เฉลี่ยต่ำที่สุด จึงมีความเหมาะสมกับการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบในอนาคตต่อไป โดยตัวแบบการพยากรณ์รวม แสดงดังนี้

$$\hat{Y}_t = 0.234904 \hat{Y}_{1t} + 0.765096 \hat{Y}_{2t}$$

เมื่อ \hat{Y}_t และ \hat{Y}_{2t} แทนค่าพยากรณ์เชิงเดี่ยว ณ เวลา t จากวิธีบอกซ์-เจนกินส์ และวิธีปรับให้เรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก ตามลำดับ

ผลการศึกษาในครั้งนี้มีความสอดคล้องกับการศึกษาในอดีตที่พบว่า การพยากรณ์รวมเป็นวิธีที่มีความแม่นยำในการพยากรณ์มากที่สุด^{9, 11-14} แต่ก็มีความขัดแย้งกับการศึกษาบางงานที่พบว่า การพยากรณ์รวมไม่ได้เป็นวิธีที่ดีที่สุด¹⁵⁻¹⁶ เนื่องจากการพยากรณ์รวมจะให้ค่าพยากรณ์ที่มีความถูกต้องน่าเชื่อถือ เมื่อมีการกำหนดค่าถ่วงน้ำหนักที่เหมาะสม ซึ่งการกำหนดค่าถ่วงน้ำหนักสามารถทำได้หลายวิธี เช่น วิธีกำลังสองน้อยที่สุด ที่ได้ใช้ในการวิจัยครั้งนี้ การถ่วงน้ำหนักด้วยการผกผันของรากที่สองของผลรวมของค่าคลาดเคลื่อนกำลังสอง (Inverse of Root Sum Squares Error) การถ่วงน้ำหนักด้วย

การผกผันของผลรวมของค่าคลาดเคลื่อนสัมบูรณ์ (Inverse of Sum Absolute Error) และการถ่วงน้ำหนักด้วยสัดส่วนของค่าจากเวกเตอร์เจาะจงของการวิเคราะห์ตัวประกอบหลัก (Proportion of the Value in the Eigen-vector from the Principal Component Analysis)¹⁷ โดยน้ำหนักถ่วงจากแต่ละวิธีจะมีความเหมาะสมกับข้อมูลที่แตกต่างกัน

เมื่อใช้วิธีการพยากรณ์รวมในการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2558 พบว่า จำนวนผู้ป่วยโรคปอดอักเสบมีแนวโน้มเพิ่มขึ้น โดยไตรมาสที่ 1 ถึง 4 ของปี 2558 มีค่าประมาณจำนวนผู้ป่วยโรคปอดอักเสบเป็น 51,038, 38,860, 56,414 และ 45,875 คน ตามลำดับ อย่างไรก็ตาม จากผลการใช้ตัวแบบการพยากรณ์ทั้ง 3 วิธี สำหรับการพยากรณ์ข้อมูลชุดที่ 2 คือ จำนวนผู้ป่วยโรคปอดอักเสบ ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 พบว่า ค่าพยากรณ์ของไตรมาสที่ 1 จะมีค่าต่ำกว่าค่าจริง ดังนั้นผู้นำตัวแบบการพยากรณ์ไปใช้ประโยชน์ควรระมัดระวังกับค่าพยากรณ์ในไตรมาสที่ 1 ของแต่ละปีด้วย และการเปลี่ยนแปลงของจำนวนผู้ป่วยโรคปอดอักเสบอาจเกิดจากปัจจัยอื่นๆ ดังนั้นเมื่อมีข้อมูลที่เป็นปัจจุบัน ผู้วิจัยควรนำมาปรับปรุงตัวแบบ เพื่อให้ได้ตัวแบบการพยากรณ์ที่มีความเหมาะสมสำหรับการพยากรณ์ในอนาคตต่อไป รวมถึงข้อมูลการระบาดโดยทั่วไปมักจะรายงานเป็นรายเดือนหรือรายสัปดาห์เพื่อใช้ในการเฝ้าระวัง แต่ในที่นี้จำนวนและอัตราผู้ป่วยด้วยโรคเฝ้าระวัง ปี พ.ศ. 2546-2557 ที่ได้จากระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต² เป็นอนุกรมเวลารายไตรมาสซึ่งเป็นช่วงเวลาที่ความกว้างมากกว่า จึงทำให้เห็นรูปแบบการเปลี่ยนแปลงวงรอบ (Cyclical Trend) ที่ชัดเจนกว่า อย่างไรก็ตาม ช่วงเวลาที่กว้างอาจส่งผลต่อความความแม่นยำของตัวแบบการพยากรณ์ ดังนั้นการศึกษาครั้งต่อไปควรพิจารณาอนุกรมเวลารายเดือนและรายสัปดาห์ประกอบการสร้างตัวแบบการพยากรณ์ด้วย

เอกสารอ้างอิง

1. สำนักงานป้องกันควบคุมโรคที่ 5 นครราชสีมา. สถานการณ์โรคปอดอักเสบ. [ออนไลน์].2551 เข้าถึงได้จาก: http://www.dpck5.com/hot/situation_pneu.htm. (วันที่ค้นข้อมูล: 14 เมษายน 2558)
2. ระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต. จำนวนและอัตราผู้ป่วยด้วยโรคเฝ้าระวัง ปี พ.ศ. 2546-2557. [ออนไลน์]. 2558 เข้าถึงได้จาก: http://social.nesdb.go.th/SocialStat/StatReport_Final.aspx?reportid=303&template=1R2C&yeartype=M&subcatid=17. (วันที่ค้นข้อมูล: 9 เมษายน 2558)
3. สำนักโรคระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข. โรคปอดอักเสบ. [ออนไลน์]. เข้าถึงได้จาก:<http://www.boe.moph.go.th/fact/Pneumonia.htm>.(วันที่ค้นข้อมูล: 14 เมษายน 2558)
4. ทรงศิริ แต่สมบัติ. การพยากรณ์เชิงปริมาณ. กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์; 2549.
5. Box, GEP, Jenkins, GM, Reinsel, GC. Time Series Analysis: Forecasting and Control. 3rded. New Jersey: Prentice Hall; 1994.
6. Bowerman, BL, O'Connell, RT. Forecasting and Time Series: An Applied Approach. 3rded. California: Duxbury Press; 1993.
7. Winters, P. Forecasting sale by exponentially weighted moving average. Management Science 1960; 6(3): 324-42.
8. สมเกียรติเกิดตุเอี่ยม. เทคนิคการพยากรณ์. พิมพ์ครั้งที่ 2. สงขลา: มหาวิทยาลัยทักษิณ; 2548.
9. มุกดา แม้นมินทร์. อนุกรมเวลาและการพยากรณ์. กรุงเทพฯ: โพรพรินติ้ง; 2549.

10. Montgomery, DC, Peck, EA, Vining, GG. Introduction to Linear Regression Analysis. 4thed. New York: Wiley; 2006.
11. ยี่งยง แสนเดช, นิดา ชาญบรรยง, ประสิทธิ์ พยัคฆพงษ์. การศึกษาตัวแบบการพยากรณ์ ปริมาณการส่งออกกุ้งสดแช่แข็ง. วารสาร มหาวิทยาลัยศรีนครินทรวิโรฒ, 2554; 3(2): 32-44.
12. วรางคณา กীরติวิบูลย์. ตัวแบบพยากรณ์จำนวน นักท่องเที่ยวต่างชาติที่มาท่องเที่ยวในประเทศไทย. วารสารวิทยาศาสตร์ มศว, 2556; 29(2) : 9-26.
13. วรางคณา กীরติวิบูลย์. การเปรียบเทียบวิธีการพยากรณ์ระหว่างวิธีบอซซ์-เจนกินส์ วิธีการทำให้เรียบแบบเอกซ์โพเนนเชียลที่มีฤดูกาลอย่างง่าย และวิธีการพยากรณ์รวม สำหรับการพยากรณ์อุณหภูมิเฉลี่ยต่อเดือนในเขตกรุงเทพมหานคร. วารสารวิทยาศาสตร์บูรพา, 2556; 18(2): 149-60.
14. วรางคณา กীরติวิบูลย์. ตัวแบบพยากรณ์ ปริมาณการส่งออกกาแฟคั่วและบด. วารสาร วิทยาศาสตร์ มศว, 2557; 30(1): 55-73.
15. วรางคณา กীরติวิบูลย์. ตัวแบบพยากรณ์มูลค่า การส่งออกกุ้งสดแช่แข็งของประเทศไทย. วารสารวิทยาศาสตร์ มช, 2558; 43(1): 148-62.
16. วราฤทธิ์ พานิชกิจโกศลกุล. การเปรียบเทียบวิธีการพยากรณ์ราคาทองคำรูปพรรณ รายวันระหว่างวิธีการพยากรณ์ของโฮลต์ วิธีการพยากรณ์ของบอซซ์-เจนกินส์ และวิธีการพยากรณ์รวม. วารสารมหาวิทยาลัยนเรศวร, 2549; 14(2): 9-16.
17. วรางคณา กীরติวิบูลย์, เจ๊ะอัสฮัน มาหิเละ. ตัวแบบพยากรณ์ความเร็วลมตามแนวชายฝั่ง อำเภอบ้านลาด จังหวัดนครศรีธรรมราช. วารสารวิจัย มช, 2556; 18(1) : 32-50.