ตัวแบบพยากรณ์จำนวนผู้ว่างงานในประเทศไทย

FORECASTING MODEL FOR THE NUMBER OF UNEMPLOYED PERSON IN THAILAND

*วรางคณา กีรติวิบูลย์**

Warangkhana Keerativibool*

สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ วิทยาเขตพัทลุง

Department of Mathematics and Statistics, Faculty of Science, Thaksin University, Phatthalung Campus.

***Corresponding author, E-mail:** warang27@gmail.com*

บทคัดย่อ

วัตถุประสงค์ของการวิจัยครั้งนี้ คือ การศึกษาและพยากรณ์จำนวนผู้ว่างงานในประเทศไทย โดยใช้ข้อมูลจากเว็บไซต์ของระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต ตั้งแต่ไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 52 ค่า ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด ชุดที่ 1 คือ ข้อมูลตั้งแต่ไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2556 จำนวน 48 ค่า สำหรับการสร้างตัวแบบพยากรณ์ด้วยวิธีการทางสถิติ 2 วิธี ได้แก่ วิธีบอกซ์-เจนกินส์ และวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ ชุดที่ 2 คือ ข้อมูลตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า สำหรับการตรวจสอบความแม่นของตัวแบบพยากรณ์ด้วยเกณฑ์เปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย และเกณฑ์รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำที่สุด ผลการวิจัยพบว่า วิธีบอกซ์-เจนกินส์มีความแม่นมากที่สุด อย่างไรก็ตาม ค่าพยากรณ์ของทั้ง 2 วิธี มีความน่าเชื่อถือ เนื่องจากไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

คำสำคัญ: ผู้ว่างงาน บอกซ์-เจนกินส์ การปรับเรียบด้วยเส้นโค้งเลขชี้กำลัง เปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย

Abstract

The purpose of this research was to study and forecast the number of unemployed person in Thailand. The data gathered from the website of Social and Quality of Life Database System during the first quarter, 2002 to the fourth quarter, 2014 (52 values) were used and divided into two categories. The first category had 48 values, which were the data during the first quarter, 2002 to the fourth quarter, 2013 for the modeling by the methods of Box-Jenkins and Winters' multiplicative exponential smoothing. 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 criteria of the lowest mean absolute percentage error and root mean squared error. The results showed that, Box-Jenkins method was the most accurate. However, the forecast values of two methods were reliable because there was no statistically significant difference.

Keywords: Unemployed Person, Box-Jenkins, Exponential Smoothing, Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE)

บทนำ

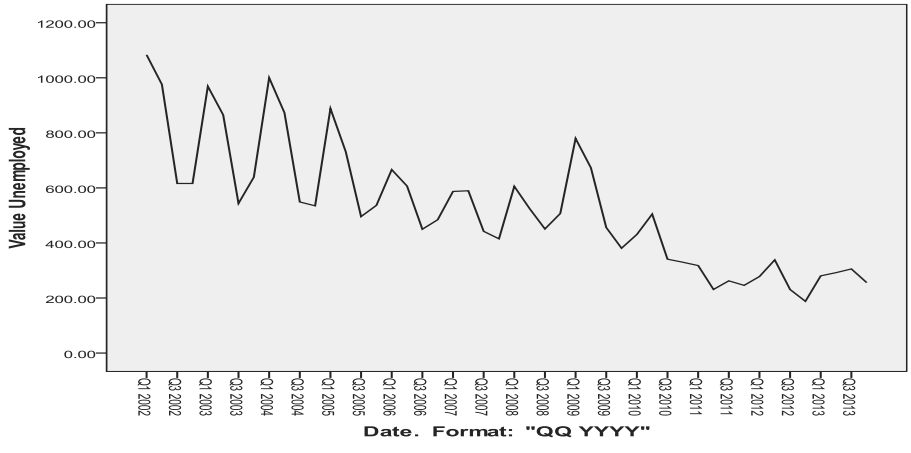
ประชากรเป็นทรัพยากรกำลังคนของชาติที่มีบทบาทสำคัญต่อความสำเร็จหรือความล้มเหลวในการพัฒนาประเทศ ดังนั้นจึงมีความจำเป็นอย่างยิ่งที่จะต้องวิเคราะห์สถานการณ์ด้านประชากรกำลังคน แรงงาน การมีงานทำ และการว่างงาน เพื่อกำหนดเป็นแนวทางการพัฒนาและการแก้ไขปัญหาต่างๆ [1] ปัจจุบันสำนักงานสถิติแห่งชาติเป็นหน่วยงานหนึ่งที่รับผิดชอบในการเก็บรวบรวมข้อมูลด้านเศรษฐกิจและสังคมของประเทศด้วยวิธีการทำสำมะโน และการสำรวจข้อมูลทั้งในระดับประเทศ ระดับภาค และระดับจังหวัด ซึ่งโครงการสำรวจที่สำคัญและต้องเก็บรวบรวมข้อมูลทุกเดือน คือ การสำรวจภาวะการทำงานของประชากร โดยมีวัตถุประสงค์เพื่อประมาณจำนวน ลักษณะของกำลังแรงงาน และการว่างงานภายในประเทศ จากวิธีการประมาณค่ายอดรวมของตัวอย่าง 3 เดือนรวมกันแล้วใช้สูตรการประมาณค่าประชากร ซึ่งวิธีการดังกล่าวต้องรอให้มีข้อมูลครบทั้ง 3 เดือน ทำให้เกิดความล่าช้าในการแสดงผลของตัวเลขอัตราการว่างงานในแต่ละเดือน ส่งผลให้ไม่ทันต่อความต้องการของหลายหน่วยงานที่จะนำตัวเลขอัตราการว่างงานไปใช้ประโยชน์ รวมถึงจากรายงานของกองวิจัยตลาดแรงงาน กรมการจัดหางาน [2] และระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต [3] เกี่ยวกับแนวโน้มสถานการณ์กำลังแรงงาน การมีงานทำ และการว่างงานในปี 2557 พบว่าจำนวนผู้ว่างงานในปัจจุบันจะยังคงมีความผันผวนค่อนข้างสูง การได้ทราบถึงจำนวนผู้ว่างงานในอนาคตย่อมส่งผลดีต่อการวางแผนพัฒนาประเทศ ด้วยเหตุผลและความจำเป็นดังกล่าว ผู้วิจัยจึงได้เล็งเห็นถึงความสำคัญในความต้องการของผู้ใช้ และได้ศึกษาหารูปแบบการพยากรณ์จำนวนผู้ว่างงานในประเทศไทยด้วยการนำข้อมูลจำนวนผู้ว่างงานในอดีตมาศึกษาการเคลื่อนไหวเพื่อสร้างตัวแบบพยากรณ์ที่เหมาะสมที่สุด โดยผลการทบทวนวรรณกรรมในอดีตเกี่ยวกับการพยากรณ์ผู้ว่างงาน พบว่า กัญญ์ลภัส มหิพันธุ์ [4] ได้ใช้แบบจำลองโครงข่ายประสาทเทียม (Artificial Neural Network Model: ANN) และวิธีบอกซ์-เจนกินส์ (Box-Jenkins Method) ในการพยากรณ์ ผลการศึกษาพบว่า วิธีบอกซ์-เจนกินส์มีความเหมาะสมมากกว่าแบบจำลองโครงข่ายประสาทเทียม ดังนั้นการศึกษาครั้งนี้ผู้วิจัยจึงยังคงใช้วิธีบอกซ์-เจนกินส์ในการพยากรณ์ และเพิ่มเติมวิธีการพยากรณ์อื่นๆ อีก 1 วิธี นั่นคือ วิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ จากนั้นจึงคัดเลือกวิธีการพยากรณ์ที่เหมาะสมเพียงวิธีเดียวสำหรับใช้ในการพยากรณ์จำนวนผู้ว่างงานให้ทันสมัยตรงตามความต้องการของผู้ใช้ ซึ่งตัวเลขอัตราการว่างงานในประเทศยังเป็นเครื่องชี้ภาวะเศรษฐกิจได้อีกด้วย เช่น ในช่วงภาวะเศรษฐกิจตกต่ำ อัตราการว่างงานค่อนข้างสูง และในช่วงภาวะเศรษฐกิจขยายตัว อัตราการว่างงานจะลดลง อีกทั้งผลการพยากรณ์จำนวนผู้ว่างงานยังเป็นประโยชน์ต่อหน่วยงานต่างๆ ทั้งภาครัฐ ภาคธุรกิจเอกชน ภาคประชาชน และภาคส่วนต่างๆ สำหรับใช้ในการวางแผนและการกำหนดนโยบายด้านแรงงานต่อไป

วัตถุประสงค์ของการวิจัย

วัตถุประสงค์ของการวิจัยครั้งนี้ คือ การศึกษาและพยากรณ์จำนวนผู้ว่างงานในประเทศไทยด้วยวิธีการทางสถิติ 2 วิธี ได้แก่ วิธีบอกซ์-เจนกินส์ และวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ โดยใช้ข้อมูลจากเว็บไซต์ของระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต ตั้งแต่ไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 52 ค่า

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ดำเนินการสร้างตัวแบบพยากรณ์จำนวนผู้ว่างงาน โดยใช้โปรแกรมสำเร็จรูปทางสถิติ ซึ่งข้อมูลที่ใช้ในการศึกษาได้มาจากเว็บไซต์ของระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต [3] เป็นอนุกรมเวลาจำนวนผู้ว่างงานตั้งแต่ไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 52 ค่า ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด ชุดที่ 1 ตั้งแต่ไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2556 จำนวน 48 ค่า สำหรับการสร้างตัวแบบพยากรณ์ด้วยวิธีการทางสถิติ 2 วิธี ได้แก่ วิธีบอกซ์-เจนกินส์ และวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ เนื่องจากทั้ง 2 วิธีการพยากรณ์มีความเหมาะสมกับอนุกรมเวลาชุดนี้มากที่สุด เพราะจากการพิจารณาลักษณะการเคลื่อนไหวของอนุกรมเวลาดังภาพที่ 1 พบว่าอนุกรมเวลาชุดนี้มีทั้งส่วนประกอบของแนวโน้มและความผันแปรตามฤดูกาล โดยที่แนวโน้มมีลักษณะลดลง และความผันแปรตามฤดูกาลมีลักษณะไม่คงที่ กล่าวคือ ความผันแปรตามฤดูกาลมีค่าลดลงตามเวลาที่เปลี่ยนแปลงไป อีกทั้งได้พิจารณาจากค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) และค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error: RMSE) ของข้อมูลชุดที่ 1 แล้วพบว่า วิธีการพยากรณ์ทั้ง 2 นี้เป็นวิธีที่มีค่า MAPE และค่า RMSE ต่ำกว่าวิธีการพยากรณ์อื่นๆ เช่น วิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก และวิธีการพยากรณ์รวม ข้อมูลชุดที่ 2 ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า นำมาใช้สำหรับการตรวจสอบความแม่น (Accuracy) ของตัวแบบพยากรณ์ด้วยเกณฑ์ MAPE และเกณฑ์ RMSE ที่ต่ำที่สุด



ภาพที่ 1 ลักษณะการเคลื่อนไหวของอนุกรมเวลาจำนวนผู้ว่างงาน ตั้งแต่ไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2556

1. การสร้างตัวแบบพยากรณ์โดยวิธีบอกซ์-เจนกินส์ (Box-Jenkins Method)

วิธีบอกซ์-เจนกินส์ เป็นวิธีการพยากรณ์ที่มีความถูกต้องสูง เนื่องจากได้กำหนดตัวแบบโดยการตรวจสอบคุณสมบัติของฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation Function: ACF) และฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (Partial Autocorrelation Function: PACF) ซึ่งพิจารณาภายใต้อนุกรมเวลาที่คงที่ (Stationary) หรืออนุกรมเวลาที่มีค่าเฉลี่ยและความแปรปรวนคงที่ [5] ตัวแบบทั่วไป (General Model) ของวิธีบอกซ์-เจนกินส์ คือ Seasonal Autoregressive Integrated Moving Average: SARIMA (p, d, q) $(P, D, Q)_s$ แสดงดังสมการที่ (1) [6]

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

เมื่อ Y_t แทนอนุกรมเวลา ณ เวลา t

ε_t แทนอนุกรมเวลาของความคลาดเคลื่อนที่มีการแจกแจงปกติและเป็นอิสระกัน ด้วยค่าเฉลี่ยเท่ากับศูนย์ และความแปรปรวนคงที่ทุกช่วงเวลา

$\delta = \mu\phi_p(B)\Phi_P(B^s)$ แทนค่าคงที่ โดยที่ μ แทนค่าเฉลี่ยของอนุกรมเวลาที่คงที่

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

$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \Phi_P B^{Ps}$ แทนตัวดำเนินการสหสัมพันธ์ในตัวเองแบบมีฤดูกาลอันดับที่ P (Seasonal Autoregressive Operator of Order P: SAR(P))

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

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

B แทนตัวดำเนินการถอยหลัง (Backward Operator) โดยที่ $B^s Y_t = Y_{t-s}$

ขั้นตอนการสร้างตัวแบบพยากรณ์โดยวิธีบอกซ์-เจนกินส์ แสดงรายละเอียดดังนี้

1) พิจารณาอนุกรมเวลาว่าคงที่หรือไม่ โดยพิจารณาจากกราฟของอนุกรมเวลาเทียบกับเวลา (Y_t, t) กราฟ ACF และ PACF หากพบว่าอนุกรมเวลาไม่คงที่ (Non-Stationary) ต้องแปลงอนุกรมเวลาให้คงที่ก่อนที่จะทำขั้นตอนต่อไป เช่น การแปลงข้อมูลด้วยการหาผลต่างหรือผลต่างฤดูกาล (Difference or Seasonal Difference) การแปลงข้อมูลด้วยลอการิทึมสามัญหรือลอการิทึมธรรมชาติ (Common Logarithm or Natural Logarithm) การแปลงข้อมูลด้วยเลขยกกำลัง เช่น ยกกำลัง 0.5 (Square Root Transformation) หรือยกกำลัง 2 (Square Transformation) เป็นต้น [7]

2) กำหนดตัวแบบพยากรณ์ที่เป็นไปได้จากกราฟ ACF และ PACF ของอนุกรมเวลาที่คงที่ นั่นคือ กำหนดค่า p, q, P และ Q พร้อมทั้งประมาณค่าพารามิเตอร์ของตัวแบบ

3) ตัดพารามิเตอร์ที่ไม่มีนัยสำคัญออกจากตัวแบบพยากรณ์ครั้งละ 1 ตัว จากนั้นจึงกำหนดตัวแบบพยากรณ์และประมาณค่าพารามิเตอร์ใหม่จนกว่าจะได้ตัวแบบพยากรณ์ที่ประกอบด้วยพารามิเตอร์ที่มีนัยสำคัญทั้งหมด

4) คัดเลือกตัวแบบพยากรณ์ที่มีค่าเกณฑ์สารสนเทศเบย์เซียน (Bayesian Information Criterion: BIC) ที่ต่ำที่สุด มีค่าสถิติ Ljung-Box Q ที่ไม่มีนัยสำคัญ และอนุกรมเวลาของความคลาดเคลื่อนจากการพยากรณ์มีการแจกแจงปกติ ตรวจสอบโดยใช้การทดสอบโคลโมโกรอฟ-สเมียร์นอฟ (Kolmogorov-Smirnov's Test) มีการเคลื่อนไหวเป็นอิสระกัน ตรวจสอบโดยพิจารณาจากกราฟ ACF และ PACF ของความคลาดเคลื่อน มีค่าเฉลี่ยเท่ากับศูนย์ ตรวจสอบโดยใช้การทดสอบที (t-Test) มีความแปรปรวนคงที่ทุกช่วงเวลา ตรวจสอบโดยใช้การทดสอบของเลวีนภายใต้การใช้ค่ามัธยฐาน (Levene's Test based on Median)

5) พยากรณ์อนุกรมเวลา โดยใช้ตัวแบบพยากรณ์ที่เหมาะสมที่สุดจากขั้นตอนที่ 4

2. การสร้างตัวแบบพยากรณ์โดยวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ (Winters' Multiplicative Exponential Smoothing Method)

การพยากรณ์โดยวิธีการปรับเรียบ (Smoothing Method) คือ การพยากรณ์โดยใช้ค่าสังเกตจากอดีตส่วนหนึ่งหรือทั้งหมดในการสร้างสมการพยากรณ์ ซึ่งน้ำหนักที่ให้กับค่าสังเกตแต่ละค่าจะแตกต่างกัน สำหรับวิธีการปรับเรียบนั้นมีวิธีการหลายวิธี และการใช้งานจะขึ้นอยู่กับลักษณะของอนุกรมเวลา [8] โดยวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์ แบ่งออกเป็น 2 กรณี คือ การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบบวก (Winters' Additive Exponential Smoothing) ควรใช้กับการพยากรณ์อนุกรมเวลาที่มีแนวโน้มเชิงเส้นและมีความผันแปรตามฤดูกาลคงที่ กล่าวคือความผันแปรตามฤดูกาลมีค่าไม่เพิ่มขึ้นและไม่ลดลงตามเวลาที่เปลี่ยนแปลงไป และการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ (Winters' Multiplicative Exponential Smoothing) ควรใช้กับการพยากรณ์อนุกรมเวลาที่มีแนวโน้มเชิงเส้นและมีความผันแปรตามฤดูกาลเพิ่มขึ้นหรือลดลงตามเวลาที่เปลี่ยนแปลงไป [9] สำหรับการวิจัยครั้งนี้ได้ใช้วิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ เนื่องจากอนุกรมเวลาจำนวนผู้ว่างงานของข้อมูลชุดที่ 1 ในช่วงไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2556 มีความผันแปรตามฤดูกาลลดลงตามเวลาที่เปลี่ยนแปลงไป (แสดงรายละเอียดในภาพที่ 1) ตัวแบบแสดงดังสมการที่ (2) และตัวแบบพยากรณ์แสดงดังสมการที่ (3) [10]

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

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

เมื่อ Y_t แทนอนุกรมเวลา ณ เวลา t

β_0, β_1, และ S_t แทนพารามิเตอร์ของตัวแบบแสดงระยะตัดแกน ความชันของแนวโน้ม และความผันแปรตามฤดูกาล ตามลำดับ

ε_t แทนอนุกรมเวลาของความคลาดเคลื่อนที่มีการแจกแจงปกติและเป็นอิสระกัน ด้วยค่าเฉลี่ยเท่ากับศูนย์ และความแปรปรวนคงที่ทุกช่วงเวลา

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

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

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

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

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

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

t แทนช่วงเวลา ซึ่งมีค่าตั้งแต่ 1 ถึง n โดยที่ n แทนจำนวนข้อมูลในอนุกรมเวลาชุดที่ 1

s แทนจำนวนฤดูกาล

เมื่อได้ตัวแบบพยากรณ์แล้วจะดำเนินการตรวจสอบคุณลักษณะของความคลาดเคลื่อนจากการพยากรณ์ตามขั้นตอนที่ 4 ของการสร้างตัวแบบพยากรณ์โดยวิธีบอกซ์-เจนกินส์

3. การตรวจสอบความแม่นของตัวแบบพยากรณ์

การวิจัยครั้งนี้ได้ตรวจสอบความแม่นของตัวแบบพยากรณ์จากวิธีการทั้งหมด 2 วิธี คือ วิธีบอกซ์-เจนกินส์ และวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ โดยทำการพยากรณ์จำนวนผู้ว่างงานของข้อมูลชุดที่ 2 คืออนุกรมเวลาตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า ได้ค่าความแตกต่างระหว่างข้อมูลจริงกับค่าพยากรณ์ (Error: e_t) เพื่อคำนวณค่า MAPE และค่า RMSE มีสูตรแสดงดังสมการที่ (4) [10] โดยวิธีการพยากรณ์ใดมีค่า MAPE และค่า RMSE ต่ำที่สุด คือ วิธีที่มีความแม่นในการพยากรณ์มากที่สุด

$$MAPE = \frac{100}{n_2}\sum_{t=1}^{n_2}\left|\frac{e_t}{Y_t}\right| \text{ และ } RMSE = \sqrt{\frac{1}{n_2}\sum_{t=1}^{n_2}e_t^2} \quad (4)$$

เมื่อ $e_t = Y_t - \hat{Y}_t$ แทนความคลาดเคลื่อนจากการพยากรณ์ 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

4. การพยากรณ์จำนวนผู้ว่างงาน

จากการตรวจสอบความแม่นของตัวแบบพยากรณ์ทั้ง 2 วิธี คือ วิธีบอกซ์-เจนกินส์ และวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ เมื่อทราบว่าตัวแบบพยากรณ์ใดมีค่า MAPE และค่า RMSE ของข้อมูลชุดที่ 2 ต่ำที่สุด จะใช้ตัวแบบพยากรณ์นั้นสำหรับการพยากรณ์จำนวนผู้ว่างงานตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2558 ต่อไป

ผลการวิจัย

1. ผลการสร้างตัวแบบพยากรณ์โดยวิธีบอกซ์-เจนกินส์

จากกราฟ ACF และ PACF ดังภาพที่ 2 พบว่า อนุกรมเวลายังไม่คงที่ ดังนั้นผู้วิจัยจึงแปลงข้อมูลด้วยการหาผลต่างและผลต่างฤดูกาลลำดับที่ 1 (d = 1, D = 1, s = 4) ได้กราฟ ACF และ PACF ของอนุกรมเวลาที่แปลงข้อมูลแล้ว แสดงดังภาพที่ 3 ซึ่งพบว่า อนุกรมเวลามีลักษณะคงที่ จึงกำหนดตัวแบบพยากรณ์ที่เป็นไปได้ พร้อมกับประมาณค่าพารามิเตอร์ดังแสดงในตารางที่ 1 โดยตัวแบบพยากรณ์ที่มีค่า BIC ต่ำที่สุด และมีค่าสถิติ Ljung-Box Q ไม่มีนัยสำคัญที่ระดับ 0.01 คือ ตัวแบบ SARIMA $(0,1,0)(0,1,1)_4$ ไม่มีพจน์ค่าคงที่ เมื่อตรวจสอบคุณลักษณะของความคลาดเคลื่อนจากการพยากรณ์ พบว่า ความคลาดเคลื่อนมีการแจกแจงปกติ (Kolmogorov-Smirnov Z = 0.467, p-value = 0.981) มีการเคลื่อนไหวเป็นอิสระกัน (แสดงรายละเอียดในภาพที่ 4 ซึ่งพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองและสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนของความคลาดเคลื่อนตกอยู่ในขอบเขตความเชื่อมั่นร้อยละ 99) มีค่าเฉลี่ยเท่ากับศูนย์(t=0.352,p-value=0.727) และมีความแปรปรวนคงที่ทุกช่วงเวลา (Levene Statistic = 1.132, p-value = 0.348) ดังนั้นตัวแบบ SARIMA$(0, 1, 0)(0, 1, 1)_4$ ไม่มีพจน์ค่าคงที่ มีความเหมาะสม ซึ่งจากสมการที่ (1) สามารถเขียนเป็นตัวแบบได้ดังนี้

$$(1-B)(1-B^4)Y_t = (1-\Theta_1 B^4)\varepsilon_t$$

$$(1-B^4-B+B^5)Y_t = \varepsilon_t - \Theta_1 \varepsilon_{t-4}$$

$$Y_t = Y_{t-1} + Y_{t-4} - Y_{t-5} + \varepsilon_t - \Theta_1 \varepsilon_{t-4}$$

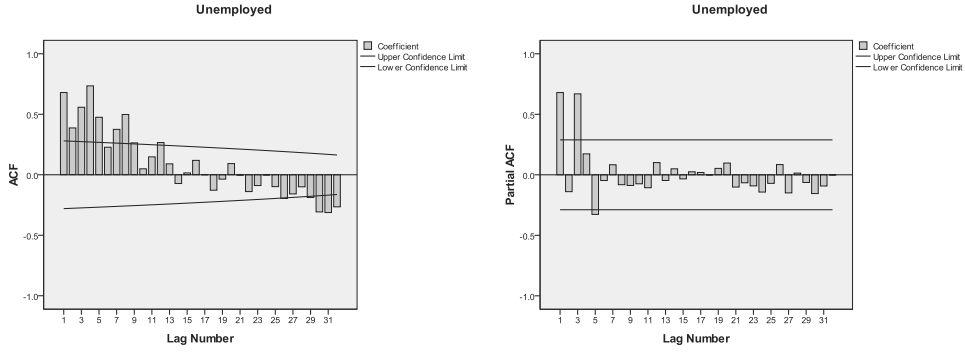
เมื่อแทนค่าประมาณพารามิเตอร์จากตารางที่ 1 จะได้ตัวแบบพยากรณ์แสดงดังนี้

$$\hat{Y}_t = Y_{t-1} + Y_{t-4} - Y_{t-5} - 0.424174 e_{t-4} \quad (5)$$

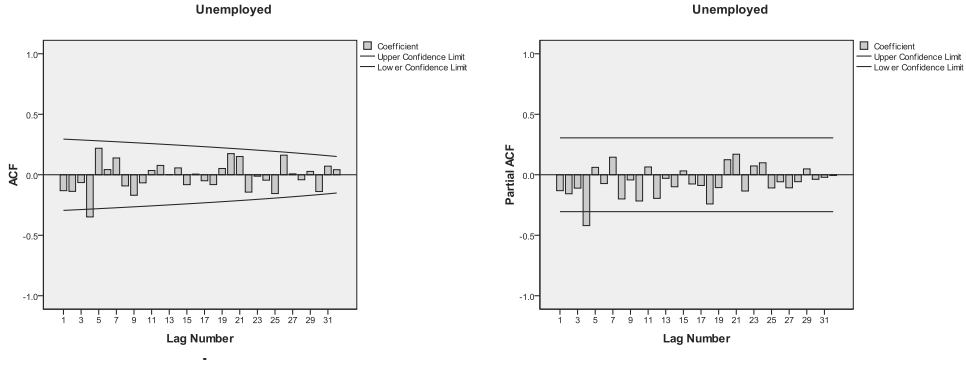
เมื่อ $\hat{Y}_t$ แทนค่าพยากรณ์ ณ เวลา t

Y_{t-j} แทนอนุกรมเวลา ณ เวลา t – j

e_{t-4} แทนความคลาดเคลื่อนจากการพยากรณ์ ณ เวลา t – 4



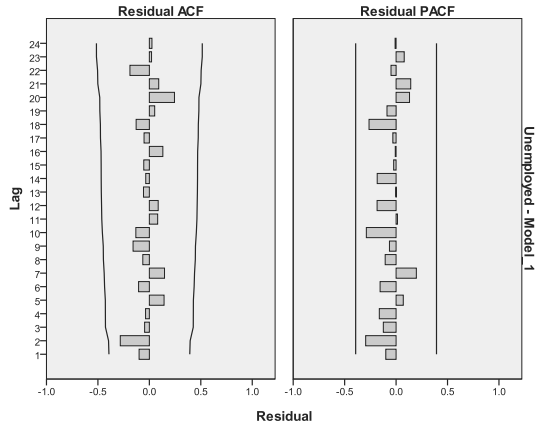
ภาพที่ 2 กราฟ ACF และ PACF ของอนุกรมเวลาจำนวนผู้ว่างงาน



ภาพที่ 3 กราฟ ACF และ PACF ของอนุกรมเวลาจำนวนผู้ว่างงาน
เมื่อแปลงข้อมูลด้วยการหาผลต่างและผลต่างฤดูกาลลำดับที่ 1

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

ค่าประมาณพารามิเตอร์		SARIMA(p, d, q)(P, D, Q)$_s$		
		SARIMA (0, 1, 0)(1, 1, 1)$_4$	**SARIMA (0, 1, 0)(0, 1, 1)$_4$**	**SARIMA (0, 1, 0)(0, 1, 1)$_4$ ไม่มีพจน์ค่าคงที่**
ค่าคงที่	ค่าประมาณ	2.481830	2.503156	-
	p-value	0.775	0.770	
SAR(1): Φ_1	ค่าประมาณ	-0.021920	-	-
	p-value	0.955		
SMA(1): Θ_1	ค่าประมาณ	0.409017	0.424883	0.424174
	p-value	0.266	0.007	0.006
BIC		9.292	9.181	9.071
Ljung-Box Q (ณ lag 18)		13.907	13.909	13.910
p-value		0.606	0.674	0.673



ภาพที่ 4 กราฟ ACF และ PACF ของความคลาดเคลื่อนจากการพยากรณ์ โดยวิธีบอกซ์-เจนกินส์ ที่มีตัวแบบ SARIMA(0, 1, 0)(0, 1, 1)$_4$ ไม่มีพจน์ค่าคงที่

2. ผลการสร้างตัวแบบพยากรณ์โดยวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ

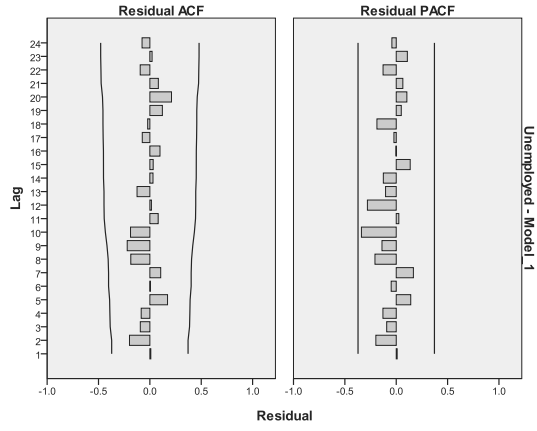
จากการสร้างตัวแบบพยากรณ์โดยวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ พบว่า BIC มีค่าเท่ากับ 8.576 และมีค่าสถิติ Ljung-Box Q ไม่มีนัยสำคัญที่ระดับ 0.01 (Ljung-Box Q ณ lag 18 = 15.346, p-value = 0.427) เมื่อตรวจสอบคุณลักษณะของความคลาดเคลื่อนจากการพยากรณ์ พบว่า ความคลาดเคลื่อนมีการแจกแจงปกติ (Kolmogorov-Smirnov Z = 0.666, p-value = 0.768) มีการเคลื่อนไหวเป็นอิสระกัน (แสดงรายละเอียดในภาพที่ 5 ซึ่งพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองและสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนของความคลาดเคลื่อนตกอยู่ในขอบเขตความเชื่อมั่นร้อยละ 99) มีค่าเฉลี่ยเท่ากับศูนย์ (t = -0.249, p-value = 0.805) และมีความแปรปรวนคงที่ทุกช่วงเวลา (Levene Statistic = 2.926, p-value = 0.046) ดังนั้นตัวแบบพยากรณ์ที่ได้มีความเหมาะสม ตัวแบบพยากรณ์แสดงดังนี้

$$\hat{Y}_{t+m} = (335.334558 - 13.070342\,m)\,\hat{S}_t \qquad (6)$$

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

$\hat{S}_t$ แทนดัชนีฤดูกาล รายละเอียดแสดงดังตารางที่ 2 ซึ่งสามารถอธิบายได้ว่า จำนวนผู้ว่างงานของไตรมาสที่ 1 และ 2 ของทุกปี มีค่ามากกว่าไตรมาสอื่นๆ เนื่องจากมีค่าดัชนีฤดูกาลมากกว่า 1

α, γ และ δ มีค่าเท่ากับ 0.739606, 0.001, 0.884896 ตามลำดับ



ภาพที่ 5 กราฟ ACF และ PACF ของความคลาดเคลื่อนจากการพยากรณ์ โดยวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ

ตารางที่ 2 ดัชนีฤดูกาลของอนุกรมเวลาจำนวนผู้ว่างงาน จากวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ

ไตรมาสที่	ดัชนีฤดูกาล	ไตรมาสที่	ดัชนีฤดูกาล
1	1.013429	3	0.877447
2	1.052850	4	0.762794

3. ผลการตรวจสอบความแม่นของตัวแบบพยากรณ์

จากการใช้ตัวแบบพยากรณ์โดยวิธีบอกซ์-เจนกินส์ ในสมการที่ (5) และโดยวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ ในสมการที่ (6) สำหรับการพยากรณ์ข้อมูลชุดที่ 2 คือ จำนวนผู้ว่างงานตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า ได้ค่าพยากรณ์ ค่า MAPE และค่า RMSE แสดงดังตารางที่ 3 ผลการตรวจสอบพบว่าวิธีบอกซ์-เจนกินส์เป็นวิธีที่มีความแม่นในการพยากรณ์มากที่สุด เนื่องจากให้ค่าพยากรณ์ที่มีความแตกต่างกับข้อมูลจริงน้อยที่สุด หรือมีค่า MAPE และค่า RMSE ต่ำที่สุด โดยมีค่า MAPE เท่ากับ 8.7136 หมายความว่าความผิดพลาดโดยเฉลี่ยในการพยากรณ์จำนวนผู้ว่างงานด้วยวิธีบอกซ์-เจนกินส์มีค่าประมาณร้อยละ 8.71 และมีค่า RMSE เท่ากับ 30.4322 หมายความว่า ความผิดพลาดโดยเฉลี่ยในการพยากรณ์จำนวนผู้ว่างงานด้วยวิธีบอกซ์-เจนกินส์มีค่าประมาณ 30.4322 พันคน หรือ 30,432 คน อย่างไรก็ตาม ค่าพยากรณ์ของทั้ง 2 วิธี มีความน่าเชื่อถือ เนื่องจากค่าพยากรณ์ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($t = 0.92$, p-value = 0.3938)

ตารางที่ 3 ค่าจริงและค่าพยากรณ์ของจำนวนผู้ว่างงาน (พันคน) ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 ค่า MAPE และค่า RMSE

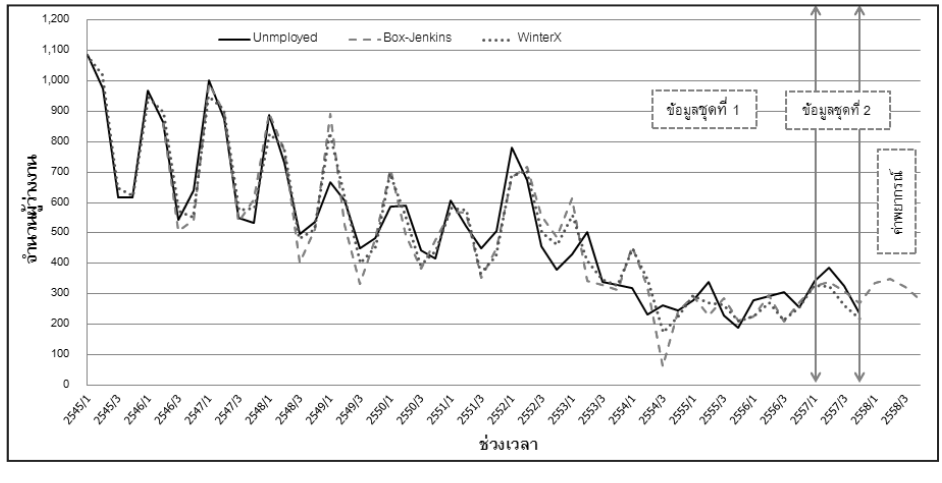
ช่วงเวลา	จำนวนผู้ว่างงาน	จำนวนผู้ว่างงาน จากการพยากรณ์โดยวิธี	
		บอกซ์-เจนกินส์	วินเทอร์
ไตรมาสที่ 1 ปี 2557	341.12	325.41	326.59
ไตรมาสที่ 2 ปี 2557	385.70	338.30	325.53
ไตรมาสที่ 3 ปี 2557	326.62	310.18	259.83
ไตรมาสที่ 4 ปี 2557	237.27	267.94	215.91
MAPE		**8.7136**	12.3277
RMSE		**30.4322**	46.7673

4. ผลการพยากรณ์จำนวนผู้ว่างงาน

จากการตรวจสอบความแม่นของตัวแบบพยากรณ์ ซึ่งพบว่า วิธีบอกซ์-เจนกินส์มีความเหมาะสมกับอนุกรมเวลาชุดนี้มากที่สุด ดังนั้นผู้วิจัยจึงเลือกใช้วิธีการดังกล่าวในการพยากรณ์จำนวนผู้ว่างงานตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2558 รายละเอียดแสดงดังตารางที่ 4 และภาพที่ 6

ตารางที่ 4 ค่าพยากรณ์จำนวนผู้ว่างงาน (พันคน) ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2558

ช่วงเวลา	ค่าพยากรณ์	ช่วงเวลา	ค่าพยากรณ์
ไตรมาสที่ 1 ปี 2558	337.45	ไตรมาสที่ 3 ปี 2558	322.23
ไตรมาสที่ 2 ปี 2558	350.34	ไตรมาสที่ 4 ปี 2558	279.98



ภาพที่ 6 การเปรียบเทียบอนุกรมเวลาจำนวนผู้ว่างงาน และค่าพยากรณ์จากวิธีการทางสถิติ 2 วิธี

สรุปและอภิปรายผล

การวิจัยครั้งนี้ได้นำเสนอวิธีการสร้างและคัดเลือกตัวแบบพยากรณ์ที่เหมาะสมกับอนุกรมเวลาจำนวนผู้ว่างงาน โดยใช้อนุกรมเวลารายไตรมาสจากเว็บไซต์ของระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต ตั้งแต่ไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 52 ค่า ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ชุด ข้อมูลชุดที่ 1 ตั้งแต่ไตรมาสที่ 1 ปี 2545 ถึงไตรมาสที่ 4 ปี 2556 จำนวน 48 ค่า สำหรับการสร้างตัวแบบพยากรณ์ด้วยวิธีการทางสถิติ 2 วิธี ได้แก่ วิธีบอกซ์-เจนกินส์ และวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของวินเทอร์แบบคูณ ข้อมูลชุดที่ 2 ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2557 จำนวน 4 ค่า สำหรับการตรวจสอบความแม่นของตัวแบบพยากรณ์ด้วยเกณฑ์ MAPE และเกณฑ์ RMSE ที่ต่ำที่สุด ผลการวิจัยพบว่า วิธีบอกซ์-เจนกินส์เป็นวิธีที่มีความแม่นในการพยากรณ์มากที่สุด เนื่องจากให้ค่าพยากรณ์ที่มีความแตกต่างกับข้อมูลจริงน้อยที่สุด หรือมีค่า MAPE และค่า RMSE ต่ำที่สุด จึงมีความเหมาะสมกับการพยากรณ์จำนวนผู้ว่างงานในอนาคตต่อไป อย่างไรก็ตาม ค่าพยากรณ์ของทั้ง 2 วิธี มีความน่าเชื่อถือ เนื่องจากไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ผลการศึกษาในครั้งนี้มีความสอดคล้องกับการศึกษาในอดีตที่พบว่า วิธีบอกซ์-เจนกินส์เป็นวิธีที่มีความแม่นในการพยากรณ์มากที่สุด [11-12] แต่ก็มีความขัดแย้งกับการศึกษาบางงานที่พบว่าวิธีบอกซ์-เจนกินส์ไม่ได้เป็นวิธีที่ดีที่สุด [13-14] เนื่องจากการเคลื่อนไหวของข้อมูลอนุกรมเวลาที่นำมาวิเคราะห์มีความแตกต่าง สำหรับการศึกษาครั้งต่อไป ผู้วิจัยควรพิจารณาวิธีการพยากรณ์อื่นๆ เพิ่มเติม เช่น การพยากรณ์ด้วยแบบจำลองโครงข่ายประสาทเทียม เพื่อให้ได้ค่าพยากรณ์ที่มีความแม่นในการพยากรณ์มากยิ่งขึ้น

จากการใช้วิธีบอกซ์-เจนกินส์ในการพยากรณ์จำนวนผู้ว่างงานตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 4 ปี 2558 ดังตารางที่ 4 และภาพที่ 6 พบว่า จำนวนผู้ว่างงานมีแนวโน้มลดลง โดยไตรมาสที่ 1 ถึง 4 ของปี 2558 มีค่าประมาณของจำนวนผู้ว่างงานเป็น 337.45, 350.34, 322.23 และ 279.98 พันคน ตามลำดับ อย่างไรก็ตาม จำนวนผู้ว่างงานมีการเปลี่ยนแปลงอยู่เสมอ ดังนั้นเมื่อมีข้อมูลที่เป็นปัจจุบัน ผู้วิจัยควรนำมาปรับปรุงตัวแบบ เพื่อให้ได้ตัวแบบพยากรณ์ที่มีความเหมาะสมสำหรับการพยากรณ์ค่าในอนาคตต่อไป

เอกสารอ้างอิง

[1] สำนักงานคณะกรรมการพัฒนาการเศรษฐกิจและสังคมแห่งชาติ. (2558). *ประชากร กำลังคน การมีงานทำ และค่าจ้าง.* สืบค้นเมื่อ 11 เมษายน 2558, จาก http://www.nesdb.go.th/portals/0/news/plan/p4/m2_5.doc

[2] กองวิจัยตลาดแรงงาน กรมการจัดหางาน. (2557). *แนวโน้มสถานการณ์กำลังแรงงาน การมีงานทำและการว่างงาน ปี 2557.* สืบค้นเมื่อ 11 เมษายน 2558, จาก http://www.sskcat.ac.th/doc/labormarket57.pdf

[3] ระบบฐานข้อมูลด้านสังคมและคุณภาพชีวิต. (2558). *ผู้ว่างงาน จำแนกตามระดับการศึกษาที่สำเร็จทั่วราชอาณาจักร ปี พ.ศ. 2545-2557.* สืบค้นเมื่อ 12 เมษายน 2558, จาก http://social.nesdb.go.th/SocialStat/StatReport_Final.aspx?reportid=411&template=2R2C&yeartype=M&subcatid=9

[4] กัญญ์ลภัส มหิพันธุ์. (2555). *การพยากรณ์อัตราการว่างงานในประเทศไทย.* วิทยานิพนธ์ปริญญามหาบัณฑิต. มหาวิทยาลัยมหาสารคาม.

[5] ทรงศิริ แต้สมบัติ. (2549). *การพยากรณ์เชิงปริมาณ.* กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์.

[6] Box, G.E.P., Jenkins, G.M.; and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control.* 3rd ed. New Jersey: Prentice Hall.

[7] Bowerman, B.L.; and O'Connell, R.T. (1993). *Forecasting and Time Series: An Applied Approach.* 3rd ed. California: Duxbury Press.

[8] วรางคณา กีรติวิบูลย์. (2557). การพยากรณ์ปริมาณการส่งออกยางคอมปาวด์. *วารสารวิทยาศาสตร์ มศว.* 30(2): 41-56.

[9] Winters, P. (1960). Forecasting Sale by Exponentially Weighted Moving Average. *Management Science.* 6(3): 324-342.

[10] สมเกียรติ เกตุเอี่ยม. (2548). *เทคนิคการพยากรณ์.* พิมพ์ครั้งที่ 2. สงขลา: มหาวิทยาลัยทักษิณ.

[11] วรางคณา กีรติวิบูลย์. (2556). **ตัวแบบพยากรณ์ราคาขายปลีกน้ำมันแก๊สโซฮอล์ 91 ในเขตกรุงเทพมหานครและปริมณฑล.** *วารสารมหาวิทยาลัยทักษิณ (จากงานประชุมวิชาการระดับชาติ มหาวิทยาลัยทักษิณ ครั้งที่ 23 ปี 2556).* 16(3): 1-10.

[12] วรางคณา กีรติวิบูลย์. (2557). ตัวแบบพยากรณ์ปริมาณน้ำฝนจังหวัดสงขลา. *วารสารมหาวิทยาลัยทักษิณ.* 17(1): 40-48.

[13] วรางคณา กีรติวิบูลย์. (2557). การเปรียบเทียบวิธีการพยากรณ์ราคาไข่ไก่ด้วยวิธีบอกซ์-เจนกินส์ และวิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของโฮลต์. *วารสารมหาวิทยาลัยทักษิณ.* 17(2): 35-43.

[14] วรางคณา กีรติวิบูลย์. (2558). ตัวแบบพยากรณ์มูลค่าการส่งออกกุ้งสดแช่แข็งของประเทศไทย. *วารสารวิทยาศาสตร์ มข.* 43(1): 148-162.