

## การศึกษาเปรียบเทียบเพื่อหาตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนผู้ป่วยที่เป็นโรคเฝ้าระวังทางระบาดวิทยาในกรุงเทพมหานคร

ดาว สงวรรังศิริกุล<sup>1</sup> ھرรษา เชี่ยวอนันตวานิช<sup>2</sup> และ มณีรัตน์ แสงเกษม<sup>2</sup>  
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี แขวงบางมด เขตทุ่งครุ กรุงเทพมหานคร 10140

### บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาหาตัวแบบพยากรณ์ที่เหมาะสมสำหรับการพยากรณ์ ประชากรที่เป็นโรคเฝ้าระวังทางระบาดวิทยาในกรุงเทพมหานคร ในการศึกษาครั้งนี้ได้นำเทคนิคการพยากรณ์มาช่วยในการวิเคราะห์ข้อมูลซึ่งประกอบด้วย วิธีเฉลี่ยเคลื่อนที่แบบง่าย (Simple Moving Average method) วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบง่าย (Simple Exponential Smoothing method) วิธีบอกซ์เจนกินส์ (Box-Jenkins method) วิธีสัดส่วนกับแนวโน้ม (Ratio-To-Trend method) และวิธีการปรับให้เรียบแบบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ (Exponential Smoothing Holt-Winter method) ทั้งนี้พิจารณาวิธีการพยากรณ์ที่เหมาะสมจากค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสัมบูรณ์ (MAPE) ที่ต่ำที่สุด ข้อมูลที่ใช้เป็นข้อมูลทุติยภูมิที่รวบรวมจากสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข กรุงเทพมหานคร ลักษณะข้อมูลจำแนกเป็นรายเดือนระหว่างเดือนมกราคม พ.ศ. 2551 ถึงเดือน ธันวาคม พ.ศ. 2555 จำนวน 9 โรคและระหว่างเดือนมกราคม พ.ศ. 2548 ถึง เดือนธันวาคม พ.ศ. 2555 จำนวน 1 โรค จากการศึกษาเปรียบเทียบข้อมูลอนุกรมเวลาที่นำมาวิเคราะห์โดยทั้ง 5 วิธี พบว่าสำหรับข้อมูลอนุกรมเวลาที่ไม่มีแนวโน้มและฤดูกาลวิธีการพยากรณ์ที่เหมาะสมที่สุดกับข้อมูลอนุกรมเวลาส่วนใหญ่ คือ วิธีเฉลี่ยเคลื่อนที่แบบง่าย ส่วนข้อมูลอนุกรมเวลาที่มีแนวโน้มและฤดูกาล วิธีการพยากรณ์ที่เหมาะสมที่สุดกับข้อมูลอนุกรมเวลาส่วนใหญ่ คือ วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์

**คำสำคัญ** : วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ / วิธีบอกซ์เจนกินส์

\* Corresponding author : E-mail: daow.sag@kmutt.ac.th

1 อาจารย์ สาขาวิชาสถิติ ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์

2 นักศึกษา สาขาวิชาสถิติ ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์

## **A Comparative Study to Determine Optimal Models for Forecasting the Number of Patients Having Epidemiological-Surveillance Diseases in Bangkok**

**Daow Sanguanrungsirikul<sup>1\*</sup> Hansa Chiewanantavanich<sup>2</sup> and Maneerat Sangkasem<sup>2</sup>**

King Mongkut's University of Technology Thonburi, Bang Mod, ThungKhru, Bangkok 10140

### **Abstract**

The objective of this study was to determine optimal models for forecasting the population of Epidemiological-Surveillance Diseases in Bangkok. In this study we applied 5 forecasting techniques to analyze the data. These include Simple Moving Average method, Simple Exponential Smoothing method, Box-Jenkins method, Ratio-To-Trend method and Exponential Smoothing Holt-Winter method. The suitable forecasting models were chosen by considering the smallest value of MAPE. Our raw data were secondary data, which were taken from the Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health. All 10 diseases data were separated on a monthly basis between from January 2008 and December 2012 and between January 2005 and December 2012. Through the comparative study of the 5 forecasting methods, the results showed that the Simple Moving Average method is the most appropriate forecasting method for almost all time-series data with no trend and seasonal characteristics. On the other hand, the Exponential Smoothing Holt-Winters is the most appropriate forecasting method for almost all time-series data exhibiting trend and seasonal characteristics.

**Keywords :** Exponential Smoothing Holt-Winter method / Box-Jenkins method

---

\* Corresponding author : E-mail: daow.sag@kmutt.ac.th

<sup>1</sup> Lecturer, Statistics Program, Department of Mathematics, Faculty of Science.

<sup>2</sup> Student, Statistics Program, Department of Mathematics, Faculty of Science.

## 1. บทนำ

การพยากรณ์ เป็นกระบวนการคาดคะเนแนวโน้มหรือรูปแบบของการเกิดเหตุการณ์ต่างๆ ในอนาคต โดยอาศัยข้อมูลที่มีอยู่ ดังนั้นการพยากรณ์จึงมีบทบาทสำคัญเป็นอย่างมากในการวางแผน และการตัดสินใจเกี่ยวกับการดำเนินงานในด้านต่างๆ ในปัจจุบัน กรมควบคุมโรค กระทรวงสาธารณสุข ได้เห็นความสำคัญของการพยากรณ์โรค จึงกำหนดให้ การพยากรณ์โรคและภัยสุขภาพ เป็นตัวชี้วัดระดับกรมบรรจุไว้ในแผนยุทธศาสตร์ โดยมีเป้าหมายหลักที่จะพัฒนานักวิชาการของทุกหน่วยงานที่เกี่ยวข้องทั้งในส่วนกลางและภูมิภาค ให้สามารถพยากรณ์โรคและภัยสุขภาพล่วงหน้าได้ โดยใช้ข้อมูลจากแหล่งข้อมูลต่างๆ ที่ได้รับรวบรวมไว้ ใช้เทคนิคและรูปแบบการประมวลผลอย่างง่าย รวมถึงใช้หลักวิชาการวิเคราะห์พื้นที่เสี่ยงต่อการเกิดโรค ทั้งโรคที่เกิดขึ้นตามฤดูกาล โรคที่เกิดขึ้นเฉพาะพื้นที่ และโรคระบาดที่เกิดขึ้นได้ตลอดทั้งปี ซึ่งผลพยากรณ์ที่ได้จะทำให้ทราบว่าโรคจะเกิดขึ้นมากน้อยแค่ไหน เกิดขึ้นกับใคร ที่ไหน เมื่อไรและอย่างไร มีปัจจัยใดบ้างที่เกี่ยวข้อง และมีวิธีการป้องกันควบคุมอย่างไร อีกทั้งยังได้ข้อมูลสำหรับใช้ในการวางแผนป้องกันและการรับมือกับโรคที่จะเกิดขึ้น นอกจากนี้ ยังมีแผนที่จะดำเนินการจัดตั้งศูนย์พยากรณ์โรคที่สำนักระบาดวิทยา และจัดทำคลังข้อมูลสำหรับการพยากรณ์โรค โดยรวบรวมข้อมูลที่เกี่ยวข้องกับการพยากรณ์โรคไว้ในแหล่งเดียวกัน และเปิดโอกาสให้ทุกหน่วยงานสามารถเข้าถึงบริการเพื่อขอใช้และแลกเปลี่ยนข้อมูลร่วมกัน รวมถึงสนับสนุนและให้บริการในด้านอื่นๆ เช่น เครื่องมือ อุปกรณ์ และโปรแกรมคอมพิวเตอร์ สำเร็จรูป เพื่อนำไปสู่เป้าหมายสำคัญคือการลดโรคของประชาชน [1]

หลายปีที่ผ่านมาเทคนิคการพยากรณ์โดยใช้หลักการทางสถิติถูกพัฒนาอย่างต่อเนื่อง เนื่องจากการพยากรณ์มีประโยชน์อย่างมากในการวางแผนและการตัดสินใจในการดำเนินงาน ทั้งระยะสั้นและระยะยาวและการพยากรณ์ยังเป็นเครื่องมืออย่างหนึ่งที่จะช่วยให้ได้ข้อมูลในอนาคตเพื่อประกอบการวางแผน จึงเป็นที่ยอมรับกันว่าการพยากรณ์มีบทบาทสำคัญทั้งในงานของภาครัฐและภาคเอกชน [2] ในปัจจุบัน ตัวแบบพยากรณ์ที่ใช้เทคนิคการวิเคราะห์อนุกรมเวลา หลายวิธีได้ถูกนำมาใช้ในการ

พยากรณ์จำนวนผู้ป่วยทางระบาดวิทยา [3,4,5,6,7] การคำนวณวิธีการพยากรณ์แต่ละวิธีที่มีประสิทธิภาพพิจารณาจากค่าความคลาดเคลื่อนที่น้อยที่สุดของการพยากรณ์ ซึ่งการเลือกวิธีการพยากรณ์ที่เหมาะสมควรพิจารณาจากหลายๆ ปัจจัยได้แก่ ช่วงการพยากรณ์ที่ต้องการเวลาที่ใช้ในการพยากรณ์ ลักษณะของข้อมูลอนุกรมเวลา และขนาดของอนุกรมเวลา [8] ซึ่งความง่ายของวิธีการพยากรณ์ในการที่ผู้ปฏิบัติจะนำไปใช้ต่อก็เป็นปัจจัยที่สำคัญอีกปัจจัยหนึ่ง หากผู้ที่เกี่ยวข้องขาดความรู้ความเข้าใจในวิธีการพยากรณ์ที่นำไปใช้ การพยากรณ์ที่ทำขึ้นก็อาจไม่มีความหมายหรือถูกนำไปประยุกต์ใช้อย่างไม่ถูกต้อง

ดังนั้น เพื่อให้ได้วิธีการพยากรณ์ที่เหมาะสม และง่ายต่อการนำไปใช้สำหรับพยากรณ์จำนวนผู้ป่วยที่เป็นโรคเฝ้าระวังทางระบาดวิทยา จึงนำมาสู่ความสนใจของผู้วิจัยที่จะทำการศึกษาเพื่อหาตัวแบบสมการพยากรณ์ที่เหมาะสม โดยเลือกวิธีการพยากรณ์ที่ง่ายต่อการเข้าใจและวิธีที่เป็นที่นิยมใช้นำมาทำการเปรียบเทียบกัน 5 วิธี ได้แก่ วิธีเฉลี่ยเคลื่อนที่แบบง่าย วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบง่าย วิธีบอซซ์เจนกินส์ วิธีสัดส่วนกับแนวโน้ม และวิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ เพื่อให้ได้เครื่องมือที่ให้ผลการทำนายที่แม่นยำ สามารถนำไปเป็นข้อมูลเบื้องต้น ประอบการวางแผนป้องกัน และควบคุมโรคที่ต้องเฝ้าระวังทางระบาดวิทยาต่อไป

## 2. วิธีการดำเนินการวิจัย

### 2.1 การจัดเตรียมข้อมูล

การวิจัยนี้มุ่งเน้นเพื่อศึกษาวิธีการพยากรณ์อนุกรมเวลากับข้อมูลจำนวนผู้ป่วยที่เป็นโรคที่ต้องเฝ้าระวังทางระบาดวิทยา กรุงเทพมหานคร โดยเลือกเฉพาะโรคที่ต้องเฝ้าระวังทางระบาดวิทยา 10 อันดับโรค ที่มีอัตราป่วยมากที่สุดในปีพุทธศักราช 2555 ได้แก่

1. โรคอุจจาระร่วงเฉียบพลัน (Acute Diarrhea)
2. โรคไข้เลือดออก (D.H.F.: Dengue Hemorrhagic Fever)
3. โรคไข้หวัดใหญ่ (Influenza)
4. โรคตาแดง (Conjunctivitis)
5. โรคปอดอักเสบ (Pneumonia)

6. โรคมือเท้าและปาก (Hand Foot Mouth disease)
7. โรคสุกใส (Chickenpox)
8. โรคอาหารเป็นพิษ (Food poisoning)
9. โรคแผลในปาก (Herpangina)
10. โรควัณโรคปอด (Pulmonary Tuberculosis)

ข้อมูลที่น่าสนใจมาทำการวิเคราะห์เป็นข้อมูลจำนวนผู้ป่วยย้อนหลังจากการรายงานของสถานพยาบาลกรุงเทพมหานคร ของกองควบคุมการระบาดในแต่ละโรครายเดือน ตั้งแต่ เดือนมกราคม 2551 จนถึง เดือนธันวาคม 2555 โดยแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดที่ 1 ใช้สำหรับการวิเคราะห์เพื่อสร้างตัวแบบพยากรณ์เป็นข้อมูลตั้งแต่ เดือนมกราคม พ.ศ. 2551 ถึง เดือนธันวาคม พ.ศ. 2554 และชุดที่ 2 เป็นข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2555 ใช้เป็นข้อมูลทดสอบสำหรับการตรวจสอบผลของตัวแบบและสมการพยากรณ์ที่สร้างขึ้นด้วยข้อมูลชุดที่ 1 (ยกเว้นโรคไข้เลือดออก ที่ใช้ข้อมูลชุดที่ 1 ตั้งแต่ เดือนมกราคม พ.ศ. 2548 จนถึง เดือนธันวาคม พ.ศ. 2554 และ ชุดที่ 2 ตั้งแต่ เดือนมกราคม พ.ศ. 2555 จนถึง เดือนธันวาคม พ.ศ. 2555 เนื่องจากมีการเก็บข้อมูลย้อนหลังถึง พ.ศ. 2548 เพียงโรคเดียว)

## 2.2 การศึกษาความเคลื่อนไหวของอนุกรมเวลา

เพื่อพิจารณาเบื้องต้นว่าอนุกรมเวลานั้นๆ มีลักษณะเป็นแบบใดทำการทดสอบองค์ประกอบของแนวโน้มและความแปรผันตามฤดูกาลโดยพิจารณาจากกราฟแสดงความสัมพันธ์ระหว่าง ค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติ ( $r_k$ ) และช่วงเวลาห่าง  $k$  แบ่งวิธีการศึกษาการพยากรณ์ออกเป็น 2 ประเภท ตามความเหมาะสมของลักษณะข้อมูลดังนี้

2.2.1 ข้อมูลอนุกรมเวลาที่ไม่มียอดประกอบแนวโน้มและฤดูกาลจะทำการเปรียบเทียบวิธีการพยากรณ์ 3 วิธี คือ วิธีเฉลี่ยเคลื่อนที่แบบง่าย วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบง่าย และวิธีบอกซ์-เจนกินส์

2.2.2 ข้อมูลอนุกรมเวลาที่มีองค์ประกอบแนวโน้มและฤดูกาลจะทำการเปรียบเทียบวิธีการพยากรณ์ 3 วิธี คือ วิธีสัดส่วนกับแนวโน้ม วิธีการปรับให้เรียบแบบ

เอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ และวิธีบอกซ์-เจนกินส์

## 2.3 การวิเคราะห์ข้อมูลทางสถิติ

การวิเคราะห์ในงานวิจัยนี้เปรียบเทียบวิธีการพยากรณ์อนุกรมเวลา 5 วิธี สร้างตัวแบบพยากรณ์ด้วยวิธีสัดส่วนกับแนวโน้ม วิเคราะห์ข้อมูลโดยใช้โปรแกรม Microsoft Excel 2010 วิธีค่าเฉลี่ยเคลื่อนที่แบบง่าย วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบง่าย วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์และวิธีบอกซ์-เจนกินส์วิเคราะห์ข้อมูลโดยใช้โปรแกรม Minitab 15 มีวิธีการดำเนินการวิเคราะห์ที่แสดงรายละเอียดดังนี้

### 2.3.1 วิธีค่าเฉลี่ยเคลื่อนที่แบบง่าย (Simple Moving Average: SMA) [8]

เป็นวิธีการสร้างสมการพยากรณ์จากการเฉลี่ยค่าสังเกตล่าสุดจำนวน  $k$  ค่า โดยให้น้ำหนักของข้อมูลเท่ากันและไม่กำหนดจำนวนค่าสังเกต ( $k$ ) ที่นำมาเฉลี่ยแน่นอนแต่จะเลือก  $k$  ที่ทำให้ค่าพยากรณ์ใกล้เคียงกับค่าจริงมากที่สุดโดยดูจากค่า  $k$  ที่ทำให้ค่าคลาดเคลื่อนโดยรวมต่ำที่สุดวิธีนี้เหมาะกับข้อมูลที่ไม่มีแนวโน้มและฤดูกาล มีสมการพยากรณ์หนึ่งช่วงเวลาล่วงหน้าที่เวลา  $t$  ดังนี้

$$\hat{Y}_t(1) = \frac{(Y_t + Y_{t-1} + \dots + Y_{t-k+1})}{k}$$

เมื่อ  $Y_t$  คือ ข้อมูลจริง ณ เวลา  $t$

$k$  คือ จำนวนข้อมูลที่ใช้ในการหาค่าเฉลี่ย

$\hat{Y}_t(1)$  คือ ค่าพยากรณ์ ณ เวลา  $t+1$

### 2.3.2 วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบง่าย (Simple Exponential Smoothing) [8]

เป็นวิธีการสร้างสมการพยากรณ์โดยการหาค่าเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนักโดยให้น้ำหนักของข้อมูลในปัจจุบันมากที่สุด และน้ำหนักจะลดหลั่นกันไปแบบเอ็กซ์โพเนนเชียลสำหรับค่าของข้อมูลที่ห่างไกลออกไป โดยมีค่าสัมประสิทธิ์ปรับให้เรียบ  $\alpha$  มีค่าอยู่ระหว่าง 0 ถึง 1 วิธีนี้เหมาะสำหรับข้อมูลที่ไม่มีแนวโน้มและฤดูกาลมีสมการพยากรณ์หนึ่งช่วงเวลาล่วงหน้าที่เวลา  $t$  ดังนี้

$$\begin{aligned}\hat{Y}_t(1) &= \alpha Y_t + (1-\alpha)\hat{Y}_{t-1}(1) \\ &= \hat{Y}_{t-1}(1) + \alpha(Y_t - \hat{Y}_{t-1}(1))\end{aligned}$$

$$\text{จะได้ } = \hat{Y}_{t-1}(1) + \alpha e_t$$

เมื่อ  $Y_t$  คือ ข้อมูลจริง ณ เวลา  $t$

$\hat{Y}_t(1)$  คือ ค่าพยากรณ์ ณ เวลา  $t+1$

$\hat{Y}_{t-1}(1)$  คือ ค่าพยากรณ์ที่ผ่านมาล่าสุด

$e_t$  คือ ผลต่างระหว่างค่าจริงกับค่าพยากรณ์

### 2.3.3 วิธีสกัดส่วนกับแนวโน้ม (Ratio-To-Trend) [8]

เป็นวิธีการสร้างสมการพยากรณ์ที่เน้นการแยกส่วนประกอบของอนุกรมเวลาออกจากกันและอธิบายแต่ละส่วนประกอบในเทอมของสมการหรือแผนแบบจากแต่ละส่วนประกอบที่แยกออกมาได้จะทำให้เห็นลักษณะการเคลื่อนไหวของอนุกรมเวลาแต่ละส่วนและนำไปสร้างสมการพยากรณ์เพื่อใช้ในการพยากรณ์ต่อไปวิธีนี้เหมาะกับข้อมูลที่มีแนวโน้มและฤดูกาล มี 2 รูปแบบ คือ วิธีสกัดส่วนกับแนวโน้มสำหรับรูปแบบบวกและวิธีสกัดส่วนกับแนวโน้มสำหรับรูปแบบคูณ

2.3.3.1 วิธีสกัดส่วนกับแนวโน้มสำหรับรูปแบบบวก  
กำหนดรูปแบบแนวโน้มฤดูกาลแบบบวกที่มีแนวโน้มเส้นตรง

$$Y_t = \beta_0 + \beta_1 t + S_t + \varepsilon_t$$

เมื่อ  $\beta_0$  เป็นค่าคงที่  $\beta_1$  เป็นอัตราการเปลี่ยนแปลงของค่าสังเกตต่อหนึ่งช่วงเวลา  $S_t$  เป็นค่าวัดอิทธิพลฤดูกาลโดยที่  $S_t = S_i$  สำหรับ  $i=1,2,\dots,L$   $t$  เป็นเวลาในฤดูกาลที่  $i$  และมี  $\sum s_i = 0$  การสร้างสมการพยากรณ์ที่ประกอบด้วยสมการแนวโน้มและค่าวัดอิทธิพลของฤดูกาล มีขั้นตอนดังนี้

ขั้นตอนที่ 1 จากค่าเฉลี่ยอนุกรมเวลาในแต่ละปี นำมาสร้างสมการแนวโน้มด้วยวิธีกำลังสองน้อยที่สุดโดยมี  $b_1$  เป็นอัตราการเปลี่ยนแปลงของค่าเฉลี่ยต่อปี จากนั้น

ทำการแปลงค่า  $b_1$  ให้เป็นอัตราการเปลี่ยนแปลงของค่าเฉลี่ยต่อฤดูกาล คือ  $b_1^* = \frac{b_1}{L}$

ขั้นตอนที่ 2 หาค่าเฉลี่ยฤดูกาลที่  $i$  ( $\bar{Y}_i$ ) ซึ่ง  $\bar{Y}_i$  ที่ได้เป็นค่าเฉลี่ยที่ยังไม่ได้ปรับแนวโน้มและฤดูกาล จากนั้นทำการปรับแนวโน้มออกจากค่าเฉลี่ยฤดูกาลที่  $i$  โดยการลบค่าเฉลี่ยแต่ละฤดูกาลด้วยขนาดของแนวโน้มจะได้ค่าเฉลี่ยที่ปรับแนวโน้มฤดูกาลที่  $i$  ( $\bar{Y}_i(adj)$ ) กล่าวคือ

$$\bar{Y}_i(adj) = \bar{Y}_i - (i-1)b_1^* ; i = 1,2,\dots,L$$

ขั้นตอนที่ 3 หาค่าวัดอิทธิพลของฤดูกาลโดยการลบค่าเฉลี่ย  $\bar{Y}_i(adj)$  ด้วย  $\bar{Y}(adj)$

เมื่อ  $\bar{Y}(adj) = \frac{\sum_{i=1}^L \bar{Y}_i(adj)}{L}$  จะได้ค่าวัดอิทธิพลฤดูกาล  $\hat{S}_i = \bar{Y}_i(adj) - \bar{Y}(adj) ; i = 1,2,\dots,L$

ขั้นตอนที่ 4 จากอนุกรมเวลาที่ปรับฤดูกาลแล้ว (ได้จากการลบค่าสังเกตจากอนุกรมเวลาเดิมด้วยค่าวัดอิทธิพลของฤดูกาล ( $\hat{S}_i$ ) ในฤดูกาล และปีที่ตรงกัน) สร้างสมการแนวโน้มด้วยวิธีกำลังสองน้อยที่สุดได้สมการแนวโน้มคือ

$$\hat{T} = b_0 + b_1 t$$

ขั้นตอนที่ 5 สร้างสมการพยากรณ์จาก  $\hat{Y}_t = \hat{T} + \hat{S}_t = b_0 + b_1 t + \hat{S}_t ; i = 1,2,\dots,L$  เมื่อ  $\hat{Y}_t$  เป็นค่าพยากรณ์ ณ เวลา  $t$  ในฤดูกาลที่  $i$   $b_0$  เป็นค่าคงที่  $b_1$  เป็นอัตราการเปลี่ยนแปลงของค่าสังเกตต่อหนึ่งช่วงเวลา  $\hat{S}_t$  เป็นค่าวัดอิทธิพลของฤดูกาล  $t$  เป็นเวลาในฤดูกาลที่  $i$  และมี  $\sum s_i = 0$

2.3.3.2 วิธีสกัดส่วนกับแนวโน้มสำหรับรูปแบบคูณ  
กำหนดรูปแบบแนวโน้มฤดูกาลรูปแบบคูณที่มีแนวโน้มแบบเอ็กซ์โพเนนเชียล

$$Y_t = \beta_0 \beta_1^t S_t \varepsilon_t$$

เมื่อ  $\beta_0$  เป็นค่าคงที่  $\beta_1$  เป็นอัตราการเปลี่ยนแปลงของค่าสังเกตต่อหนึ่งช่วงเวลา  $S_i$  เป็นค่าสถิติพลฤดูกาลโดยที่  $S_i = S_i$  สำหรับ  $i = 1, 2, \dots, L$   $t$  เป็นเวลาในฤดูกาลที่  $i$  และมี  $\sum s_i = L$

การสร้างสมการพยากรณ์ ทำโดยการสร้างอนุกรมเวลาใหม่ โดยการแปลงค่าสังเกตเดิมเป็นค่าลอการิทึมของค่าสังเกต ( $\ln Y_t$ ) จากนั้นจะทำทำนองเดียวกันกับกรณีรูปแบบแนวโน้มฤดูกาลแบบบวก ตั้งแต่ขั้นตอนที่ 1 ถึงขั้นตอนที่ 5 และได้สมการพยากรณ์รูปแบบบวก แปลงกลับให้เป็นสมการพยากรณ์สำหรับรูปแบบแนวโน้มฤดูกาลแบบคูณ โดยการเทคเอ็กซ์โพเนนเชียล กลับเข้าไปในสมการพยากรณ์ จะได้สมการพยากรณ์คือ

$$\hat{Y}_t = b_0 b_1^i \hat{S}_i ; i = 1, 2, \dots, 12$$

เมื่อ  $\hat{Y}_t$  เป็นค่าพยากรณ์ ณ เวลา  $t$  ในฤดูกาลที่  $i$   $b_0$  เป็นค่าคงที่  $b_1$  เป็นอัตราการเปลี่ยนแปลงของค่าสังเกตต่อหนึ่งช่วงเวลา  $\hat{S}_i$  เป็นค่าสถิติพลของฤดูกาล

ในการวิจัยนี้เลือกรูปแบบบวกหรือรูปแบบคูณพิจารณาจากการเคลื่อนไหวของอนุกรมเวลากว่าคือ ถ้าการเคลื่อนไหวของอนุกรมเวลาแกว่งออกจากเส้นแนวโน้มที่เวลาเดียวกันแต่ต่างปีกันไม่แตกต่างกัน (ความแปรปรวนของข้อมูลอนุกรมเวลาไม่แตกต่างกัน) จะใช้สมการพยากรณ์รูปแบบบวก

ถ้าการเคลื่อนไหวของอนุกรมเวลาแกว่งออกจากเส้นแนวโน้มที่เวลาเดียวกันแต่ต่างปีกันแตกต่างกัน (ความแปรปรวนของข้อมูลอนุกรมเวลาแตกต่างกัน) จะใช้สมการพยากรณ์รูปแบบคูณ

**2.3.4 วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ (Exponential Smoothing Holt-Winter method) [8]**

เป็นวิธีการสร้างสมการพยากรณ์สำหรับอนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากแนวโน้มและฤดูกาลมีทั้งรูปแบบแนวโน้มฤดูกาลแบบบวกและรูปแบบแนวโน้มฤดูกาลแบบคูณ โดยใช้ค่าปรับให้เรียบ 3 ค่า ได้แก่  $\alpha$   $\gamma$  และ  $\delta$  ที่มีค่าอยู่ระหว่าง 0 ถึง 1 โดยที่  $\alpha$  เป็นค่าปรับน้ำหนักสำหรับแนวโน้ม  $\gamma$  เป็นค่าปรับน้ำหนักสำหรับค่า

ความลาดชัน (slope) และ  $\delta$  เป็นค่าปรับน้ำหนักสำหรับฤดูกาล ตามลำดับ

**2.3.4.1 รูปแบบแนวโน้มฤดูกาลแบบบวก**

มีรูปแบบสมการพยากรณ์  $p$  ช่วงเวลาล่วงหน้าที่เวลา  $t$  เป็น

$$\hat{Y}_t(p) = \hat{T}_{t+p}(t) + \hat{S}_{t+p}(t) ; p = 1, 2, \dots$$

ซึ่ง  $\hat{T}_{t+p}(t)$  เป็นค่าแนวโน้มที่เวลา  $t + p$  เมื่อใช้ค่าสังเกต  $t$  ค่า มีค่าเท่ากับ  $\hat{T}_t(t) + p\hat{\beta}_1(t)$  และ  $\hat{S}_{t+p}(t)$  มีค่าเท่ากับ  $\hat{S}_i(t)$  จะได้

$$\hat{Y}_t(p) = \hat{T}_t(t) + p\hat{\beta}_1(t) + \hat{S}_i(t)$$

โดยที่  $\hat{T}_t(t)$  เป็นค่าแนวโน้ม ณ เวลา  $t$

$p\hat{\beta}_1$  เป็นค่าความลาดชัน ณ เวลา  $t$

$\hat{S}_i(t)$  เป็นค่าสถิติพลของฤดูกาลที่  $i$  และใช้ค่าสังเกต  $t$  ค่า

ค่า  $\hat{T}_t(t)$ ,  $p\hat{\beta}_1$  และ  $\hat{S}_i(t)$  หาได้จาก

$$\hat{T}_t(t) = \alpha(Y_t - \hat{S}_i(t-1)) + (1-\alpha)\hat{T}_t(t-1)$$

$$\hat{\beta}_1(t) = \gamma(\hat{T}_t(t) - \hat{T}_t(t-1)) + (1-\gamma)\hat{\beta}_1(t-1)$$

$$\hat{S}_i(t) = \delta(Y_t - \hat{T}_t(t)) + (1-\delta)\hat{S}_i(t-1)$$

**2.3.4.2 รูปแบบแนวโน้มฤดูกาลแบบคูณ**

มีรูปแบบสมการพยากรณ์  $p$  ช่วงเวลาล่วงหน้าที่เวลา  $t$  เป็น

$$\hat{Y}_t(p) = \hat{T}_{t+p}(t)\hat{S}_{t+p}(t) ; p = 1, 2, \dots$$

ความหมายของ  $\hat{T}_{t+p}(t)$  และ  $\hat{S}_{t+p}(t)$  เป็นทำนองเดียวกับกรณีรูปแบบแนวโน้มฤดูกาลแบบบวก จะได้

$$\hat{Y}_t(p) = (\hat{T}_t(t) + p\hat{\beta}_1(t))\hat{S}_i(t)$$

ค่า  $\hat{T}_i(t)$ ,  $\hat{\beta}_i(t)$  และ  $\hat{S}_i(t)$  หาได้จาก

$$\hat{T}_i(t) = \frac{\alpha Y_t}{\hat{S}_i(t-1)} + (1-\alpha)\hat{T}_i(t-1)$$

$$\hat{\beta}_i(t) = \gamma(\hat{T}_i(t) - \hat{T}_i(t-1)) + (1-\gamma)\hat{\beta}_i(t-1)$$

$$\hat{S}_i(t) = \frac{\delta Y_t}{\hat{T}_i(t)} + (1-\delta)\hat{S}_i(t-1)$$

การสร้างสมการพยากรณ์ต้องทำการกำหนดค่าเริ่มต้นของ ค่า  $\hat{T}_i(t)$ ,  $\hat{\beta}_i(t)$  และ  $\hat{S}_i(t)$  และกำหนดค่าปรับให้เรียบ  $\alpha$ ,  $\gamma$  และ  $\delta$  ที่เหมาะสมที่มีค่าอยู่ระหว่าง 0 ถึง 1 และให้ค่า MSE ต่ำที่สุด

ในการวิจัยนี้เลือกรูปแบบบวกรูปแบบคูณพยากรณ์จากการเคลื่อนไหวของอนุกรมเวลาเช่นเดียวกันกับวิธีสัดส่วนกับแนวโน้ม

### 2.3.5 วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins)

เป็นวิธีการสร้างสมการพยากรณ์อนุกรมเวลา โดยกำหนดรูปแบบ ARIMA ซึ่งเป็นรูปแบบที่อธิบายลักษณะการเคลื่อนไหวของอนุกรมเวลาที่ค่าสังเกตมีสหสัมพันธ์กันแบบอัตโนมัติ และเป็นสเตชันนารี สามารถใช้ได้กับข้อมูลที่มีลักษณะการเคลื่อนไหวทุกประเภท และให้ผลการพยากรณ์ค่อนข้างแม่นยำกว่าวิธีการพยากรณ์อื่นๆ สำหรับการพยากรณ์ระยะสั้น [8]

วิธีการของบ็อกซ์-เจนกินส์ มีขั้นตอนดังนี้

ขั้นตอนที่ 1 ตรวจสอบข้อมูลเพื่อพิจารณาว่าอนุกรมเวลาเป็นสเตชันนารีหรือไม่ โดยพิจารณาจากกราฟอนุกรมเวลาหรือพิจารณาจากกราฟฟังก์ชันสหสัมพันธ์แบบอัตโนมัติ (Autocorrelation Function: ACF) และฟังก์ชันสหสัมพันธ์แบบอัตโนมัติบางส่วน (Partial Autocorrelation Function: ACF) ของอนุกรมเวลา  $\{Y_t\}$  (ถ้าสเตชันนารีจะได้กราฟที่มีลักษณะขึ้นลงรอบศูนย์อย่างสม่ำเสมอ) ในกรณีที่อนุกรมเวลาไม่สเตชันนารี ต้องแปลงให้เป็นอนุกรมเวลาชุดใหม่ ที่เป็นสเตชันนารีทำได้โดยการกำจัดแนวโน้ม กำจัดฤดูกาล และแปลงอนุกรมเวลาให้มีค่าความแปรปรวนคงที่ เช่นการหาผลต่าง การหา

ผลต่างฤดูกาล การแปลงข้อมูลด้วยลอการิทึม เป็นต้น ซึ่งการแปลงอนุกรมเวลาขึ้นอยู่กับลักษณะของอนุกรมเวลาที่สนใจศึกษา [9]

ขั้นตอนที่ 2 พิจารณากำหนดรูปแบบการพยากรณ์ ARIMA(p,d,q)(P,D,Q)L ที่คาดว่าเหมาะสมโดยพิจารณาจากกราฟ ACF และ PACF

ขั้นตอนที่ 3 ประเมินค่าพารามิเตอร์ของรูปแบบที่กำหนดด้วยวิธีกำลังสองน้อยที่สุด

ขั้นตอนที่ 4 ตรวจสอบความเหมาะสมของรูปแบบที่กำหนดโดยพิจารณาจากกราฟ ACF และ PACF ของ  $e_t$  หรือจากการทดสอบสหสัมพันธ์แบบอัตโนมัติของ Box-Ljung เมื่อพบว่ารูปแบบมีความเหมาะสมจะสร้างสมการพยากรณ์จากรูปแบบดังกล่าว แต่หากพบว่ารูปแบบที่กำหนดไม่เหมาะสมต้องเริ่มต้นกำหนดรูปแบบในขั้นตอนที่ 2 ใหม่

ขั้นตอนที่ 5 เมื่อผ่านขั้นตอนที่ 4 จะได้สมการพยากรณ์และค่าประมาณพารามิเตอร์สำหรับการพยากรณ์

## 2.4 เกณฑ์การเปรียบเทียบประสิทธิภาพวิธีการพยากรณ์ [8]

เกณฑ์ที่ใช้ในการพิจารณาเปรียบเทียบประสิทธิภาพของวิธีการพยากรณ์พิจารณาจากค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error : MAPE) ซึ่งเป็นการวัดความคลาดเคลื่อนระหว่างค่าจริง ( $Y_t$ ) และค่าพยากรณ์ ( $\hat{Y}_t$ ) ค่า MAPE นี้ไม่มีหน่วยและมีค่าเป็นบวกเสมอ คำนวณได้ดังนี้

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{e_t}{Y_t} \right| = \frac{100}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|$$

โดยที่  $e_t$  เป็นผลต่างของค่าจริงกับค่าพยากรณ์ ณ เวลา  $t$ ,  $Y_t$  เป็นค่าจริงของข้อมูลอนุกรมเวลา ณ เวลา  $t$ ,  $\hat{Y}_t$  เป็นค่าพยากรณ์ของข้อมูลอนุกรมเวลา ณ เวลา  $t$ ,  $n$  เป็นจำนวนช่วงเวลาทั้งหมด

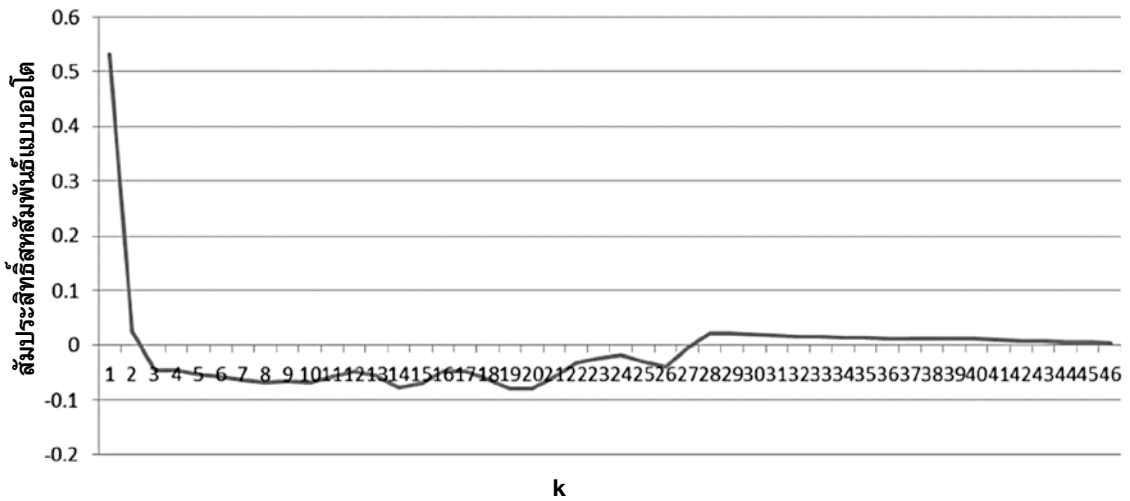
วิธีการพยากรณ์ที่ให้ค่า MAPE ต่ำที่สุดจะเป็นวิธีการพยากรณ์ที่มีความแม่นยำที่สุด

**3. ผลการวิจัย**

**3.1 ผลการวิเคราะห์ลักษณะการเคลื่อนไหวของอนุกรมเวลา**

ผลการวิเคราะห์ลักษณะการเคลื่อนไหวของอนุกรมเวลาทดสอบองค์ประกอบแนวโน้มและความ

ผันแปรของฤดูกาลโดยพิจารณาจากกราฟแสดงความสัมพันธ์ระหว่างค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่าระยะห่าง  $k$  พบว่าข้อมูลอนุกรมเวลาที่ไม่มีองค์ประกอบแนวโน้มและฤดูกาลคือข้อมูลอนุกรมเวลาของโรคอาหารเป็นพิษ โรคฉี่หนู โรคปอด โรคตาแดง และโรคมือเท้าปาก เนื่องจากค่า  $r_k$  มีค่าสูงเมื่อ  $k$  มีค่าน้อยและเท่ากับ 0 เมื่อ  $k$  มีค่ามาก [10] เช่น ตัวอย่างกราฟของโรคตาแดง ดังรูปที่ 1

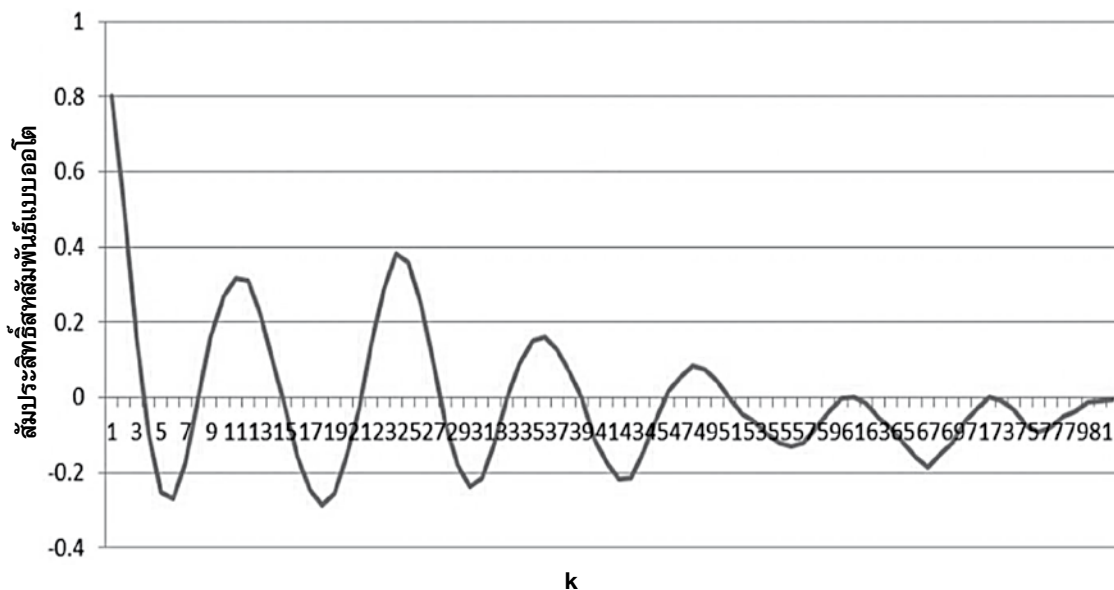


รูปที่ 1 กราฟแสดงความสัมพันธ์ระหว่างค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่า k ของโรคตาแดง

ข้อมูลอนุกรมเวลาที่มีองค์ประกอบแนวโน้มและฤดูกาลคือข้อมูลอนุกรมเวลาของโรคไข้เลือดออก โรคไข้หวัดใหญ่ โรคปอดอักเสบ โรคสุกใส โรคแผลในปาก และ โรคอุจจาระร่วงเฉียบพลัน เนื่องจากค่า  $r_k$  มีค่ามาก

และเป็นบวกเมื่อ  $k$  มีค่าน้อยและลดลงช้าๆ เมื่อ  $k$  มีค่าเพิ่มขึ้น และ  $r_k$  มีค่ามากสำหรับ  $k = 12, 24, 36, \dots$  [10] เช่น ตัวอย่างกราฟของโรคไข้เลือดออก ดังรูปที่ 2





รูปที่ 2 กราฟแสดงความสัมพันธ์ระหว่างค่าสัมประสิทธิ์สหสัมพันธ์แบบออโตและค่า k ของโรคไข้เลือดออก

### 3.2 ผลการวิเคราะห์เปรียบเทียบวิธีการพยากรณ์

ผลการวิเคราะห์เปรียบเทียบวิธีการพยากรณ์

#### 3.2.1 ข้อมูลอนุกรมเวลาที่ไม่มีแนวโน้มและ

เพื่อหารูปแบบการพยากรณ์ที่เหมาะสมที่สุดสำหรับข้อมูลอนุกรมเวลาที่ไม่มีแนวโน้มและฤดูกาล แสดงดังตารางที่ 1

ฤดูกาล

ตารางที่ 1 แสดงการเปรียบเทียบค่า MAPE ของข้อมูลอนุกรมเวลาที่ไม่มีแนวโน้มและฤดูกาลจากวิธีการพยากรณ์ด้วยวิธีเฉลี่ยเคลื่อนที่แบบง่าย วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบง่าย และวิธีบอซซ์เจนกินส์

โรคเฝ้าระวัง ทางระบาดวิทยา	ค่า MAPE ของวิธีการพยากรณ์		
	เฉลี่ยเคลื่อนที่ แบบง่าย	เอ็กซ์โพเนนเชียล แบบง่าย	บอซซ์เจนกินส์
1. โรคอาหารเป็นพิษ	20.678* (k=10)	23.178 ( $\alpha=0.16$ )	21.264 (ARIMA(1,0,0)) ( $\hat{\theta}_0 = 5.71, \hat{\phi}_1 = 0.444$ )
2. โรคฉี่หนู	14.512* (k=5)	16.399 ( $\alpha=0.2$ )	18.013 (ARIMA(1,0,0)) ( $\hat{\theta}_0 = 205.13, \hat{\phi}_1 = 0.604$ )
3. โรคตาแดง	15.957* (k=4)	16.82 ( $\alpha=0.6$ )	45.353 (ARIMA(1,0,1)) ( $\hat{\theta}_0 = 5.362, \hat{\phi}_1 = 0.529,$ $\hat{\theta}_1 = -0.559$ )
4. โรคมือเท้าปาก	36.961 (k=12)	36.840* ( $\alpha=0.04$ )	37.864 (ARIMA(0,0,1)) ( $\hat{\theta}_0 = 5.106, \hat{\phi}_1 = -0.715$ )

หมายเหตุ : \* หมายถึง ค่า MAPE ต่ำที่สุด

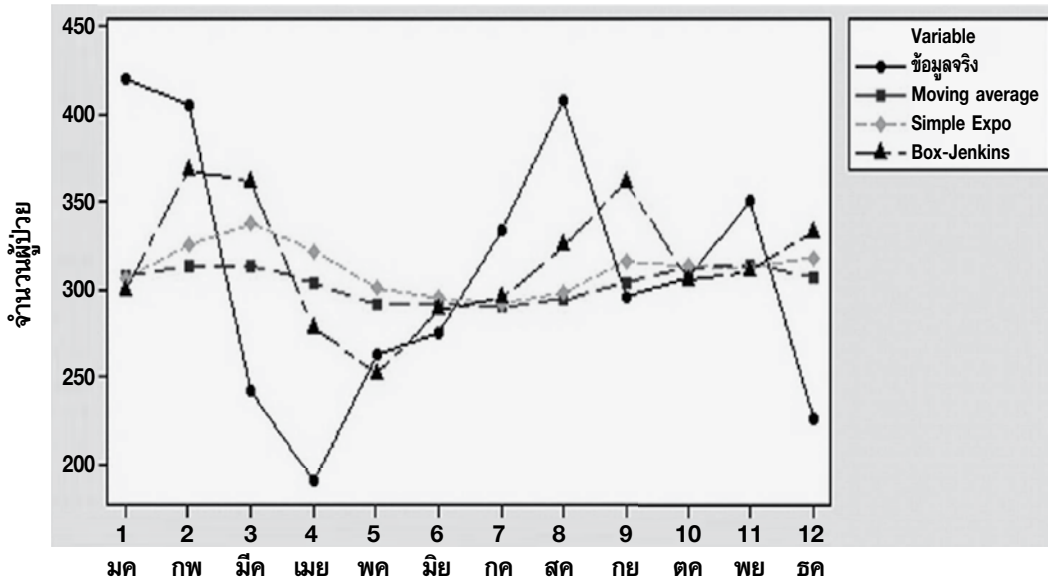
จากตารางที่ 1 พบว่าวิธีเฉลี่ยเคลื่อนที่แบบง่ายให้ค่า MAPE ต่ำที่สุดเท่ากับ 20.678 14.512 และ 15.957 สำหรับข้อมูลโรคอาหารเป็นพิษ โรควัดโรคปอดและโรคตาแดง ตามลำดับ วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบง่ายให้ค่า MAPE ต่ำที่สุด เท่ากับ 36.840 สำหรับโรคมือเท้าปาก โดยมีตัวแบบสมการพยากรณ์ของวิธีการพยากรณ์ที่ให้ค่า MAPE ต่ำที่สุดของโรคต่างๆ ดังนี้

**โรคอาหารเป็นพิษ**

กำหนดค่าสังเกตล่าสุดเป็น 10 ค่า ซึ่งเป็นค่าที่ให้ค่า MSE ต่ำที่สุดได้ตัวแบบสมการพยากรณ์ คือ

$$\hat{Y}_t(1) = \frac{(Y_t + Y_{t-1} + \dots + Y_{t-9})}{10} ; t = 60, 61, \dots$$

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการพยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 3



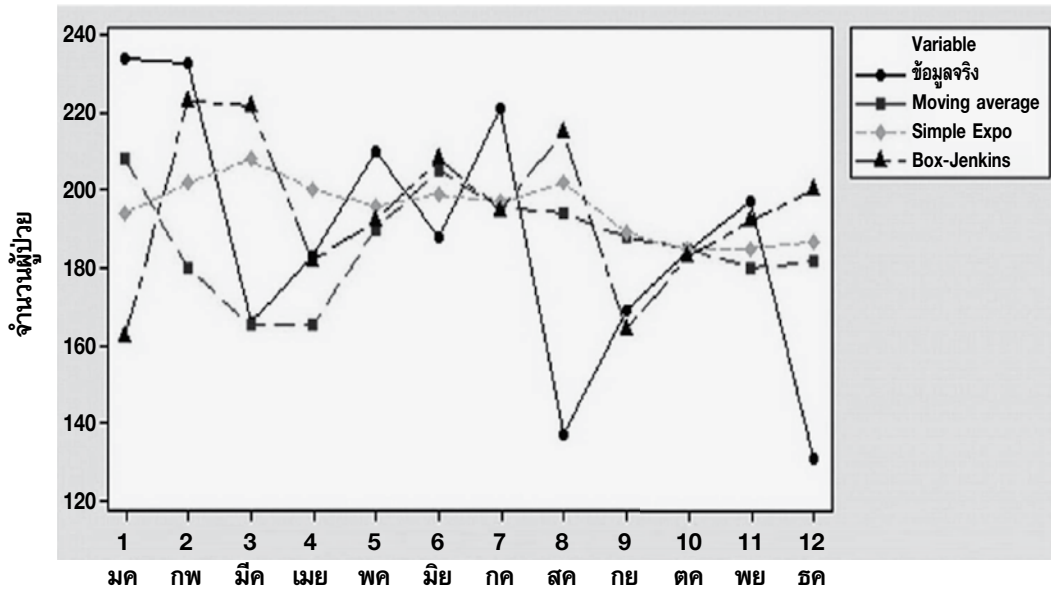
รูปที่ 3 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคอาหารเป็นพิษ

**โรควัดโรคปอด**

กำหนดค่าสังเกตล่าสุดเป็น 5 ค่า ซึ่งเป็นค่าที่ให้ค่า MSE ต่ำที่สุดได้ตัวแบบสมการพยากรณ์ คือ

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการพยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 4

$$\hat{Y}_t(1) = \frac{(Y_t + Y_{t-1} + \dots + Y_{t-4})}{5} ; t = 60, 61, \dots$$



รูปที่ 4 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรควัดโรคปอด

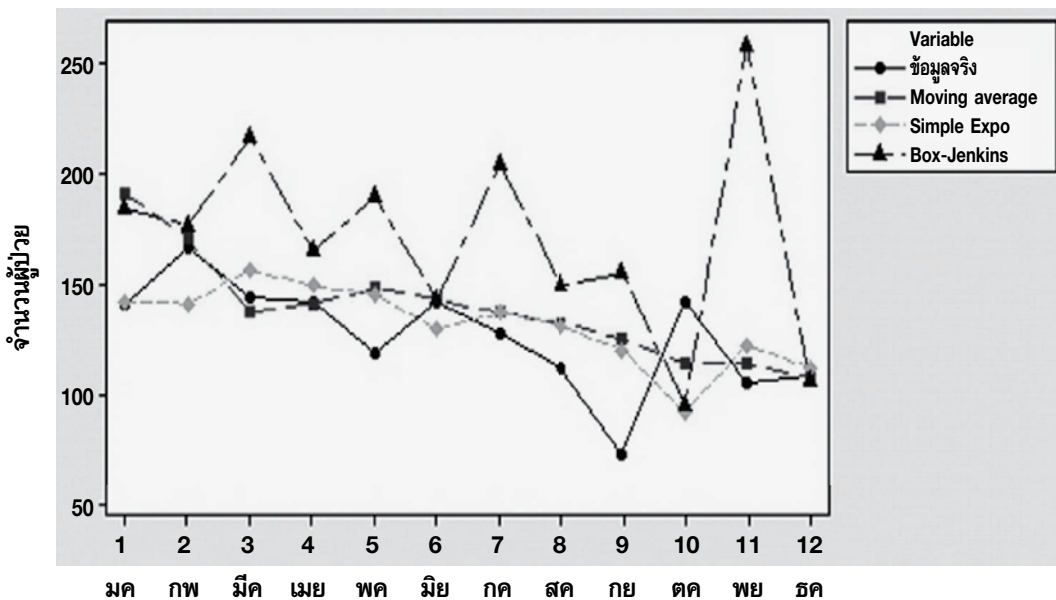
**โรคตาแดง**

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการ

กำหนดค่าสังเกตล่าสุดเป็น 4 ค่าซึ่งเป็นค่าที่ให้ค่า  
MSE ต่ำที่สุดได้ตัวแบบสมการพยากรณ์ คือ

พยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 5

$$\hat{Y}_t(1) = \frac{(Y_t + Y_{t-1} + \dots + Y_{t-3})}{4} ; t = 60, 61, \dots$$



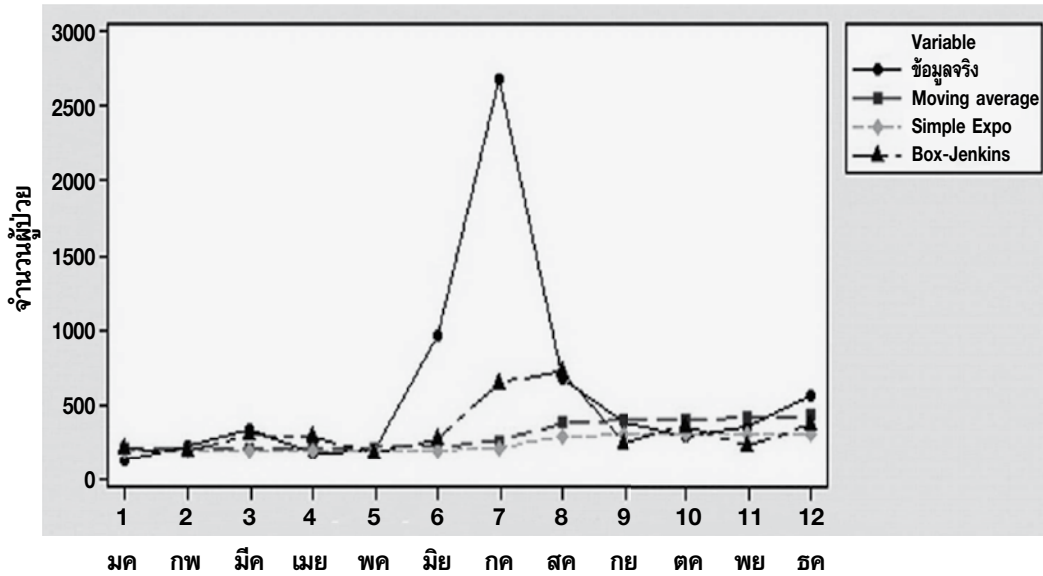
รูปที่ 5 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคตาแดง

**โรคมือเท้าปาก**

กำหนดค่าคงที่ปรับให้เรียบ  $\alpha$  มีค่าเท่ากับ 0.04 ซึ่งเป็นค่าที่ให้ค่า MSE ต่ำที่สุด ได้ตัวแบบสมการพยากรณ์ คือ

$$\hat{Y}_t(1) = \hat{Y}_{t-1}(1) + 0.04e_t ; t = 60,61,\dots$$

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการพยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 6



รูปที่ 6 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคมือเท้าปาก

**3.2.2 ข้อมูลอนุกรมเวลาที่มีแนวโน้มและฤดูกาล**  
ผลการวิเคราะห์เปรียบเทียบวิธีการพยากรณ์

เพื่อหารูปแบบการพยากรณ์ที่เหมาะสมที่สุดสำหรับข้อมูลอนุกรมเวลาที่มีแนวโน้มและฤดูกาล แสดงดังตารางที่ 2

ตารางที่ 2 แสดงการเปรียบเทียบค่า MAPE ของข้อมูลอนุกรมเวลาที่มีแนวโน้มและฤดูกาลจากวิธีการพยากรณ์ด้วยวิธี สัดส่วนกับแนวโน้มวิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ และวิธีบอซเจนกินส์

โรคเฝ้าระวัง ทางระบาดวิทยา	ค่า MAPE ของวิธีการพยากรณ์		
	สัดส่วนกับแนวโน้ม	โฮลท์-วินเทอร์	บอซเจนกินส์
1. โรคไข้เลือดออก	41.402 ( $b_0=463.76, b_1=1.01$ )	=24.769 ( $\alpha = 0.5, \gamma = 0.5, \delta = 0.5$ )	17.585* (ARIMA(0,1,0)(1,1,0) <sub>12</sub> ) ( $\hat{\theta}_0 = 0.003, \hat{\phi}_{12} = -0.678$ )
2. โรคไข้หวัดใหญ่	52.150 ( $b_0=307.48, b_1=1.02$ )	39.226* ( $\alpha = 0.4, \gamma = 0.4, \delta = 0.4$ )	66.090 (ARIMA(0,1,0)(1,1,1) <sub>12</sub> ) ( $\hat{\theta}_0 = 0.984, \hat{\phi}_{12} = -0.15,$ $\hat{\theta}_{12} = 0.979$ )
3. โรคปอดอักเสบ	27.929 ( $b_0=431.89, b_1=1.01$ )	11.667* ( $\alpha = 0.08, \gamma = 0.08, \delta = 0.08$ )	18.420 (ARIMA(0,1,1)(0,1,0) <sub>12</sub> ) ( $\hat{\theta}_0 = -0.006, \hat{\theta}_1 = -0.542$ )
4. โรคสุกใส	35.398 ( $b_0=219.48, b_1=1.01$ )	25.706* ( $\alpha = 0.3, \gamma = 0.3, \delta = 0.3$ )	27.411 (ARIMA(1,1,0)(0,1,0) <sub>12</sub> ) ( $\hat{\theta}_0 = -0.023, \hat{\phi}_1 = -0.267$ )
5. โรคแผลในปาก	41.404 ( $b_0=93.37, b_1=1.01$ )	27.395* ( $\alpha = 0.1, \gamma = 0.1, \delta = 0.1$ )	36.611 (ARIMA(1,1,0)(0,1,0) <sub>12</sub> ) ( $\hat{\theta}_0 = 0.038, \hat{\phi}_1 = -0.061$ )
6. โรคอุจจาระร่วง เฉียบพลัน	13.483 ( $b_0=2922.9, b_1=16.67$ )	12.240* ( $\alpha = 0.1, \gamma = 0.1, \delta = 0.1$ )	18.486 (ARIMA(1,1,0)(0,1,0) <sub>12</sub> ) ( $\hat{\theta}_0 = -0.016, \hat{\phi}_1 = -0.298$ )

หมายเหตุ : \* หมายถึง ค่า MAPE ที่ต่ำที่สุด

จากตารางที่ 2 พบว่า วิธีบอซเจนกินส์ให้ค่า MAPE ต่ำที่สุดเท่ากับ 17.585 สำหรับข้อมูลโรคไข้เลือดออก วิธีปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ ให้ค่า MAPE ต่ำที่สุดเท่ากับ 39.226 11.667 25.706 27.395 และ 12.240 สำหรับข้อมูลโรคไข้หวัดใหญ่ โรคปอดอักเสบ โรคสุกใส โรคแผลในปาก และโรคอุจจาระร่วงเฉียบพลัน ตามลำดับ โดยมีตัวแบบสมการพยากรณ์ของวิธีการพยากรณ์ที่ให้ค่า MAPE ต่ำที่สุดของโรคต่างๆ ดังนี้

#### โรคไข้เลือดออก

ในการสร้างสมการพยากรณ์ทำการแปลงอนุกรมเวลาให้เป็นสเตชันนารีโดยการหาผลต่างเพื่อกำจัด

แนวโน้มและหาผลต่างฤดูกาลเพื่อกำจัดอิทธิพลฤดูกาลและใช้ฟังก์ชันลอการิทึม เพื่อให้อนุกรมเวลามีความแปรปรวนคงที่ การประมาณค่าพารามิเตอร์มีนัยสำคัญที่ระดับ 0.05 (p-value = 0.000) และมีค่าสถิติ Box-Ljung ไม่มีนัยสำคัญที่ระดับ 0.05 (Ljung-Box Q(18)=15.288, df=17, p-value = 0.575) ได้รูปแบบการพยากรณ์ที่เหมาะสมคือ ARIMA(0,1,0)(1,1,0)<sub>12</sub> มีรูปแบบสมการพยากรณ์เป็น

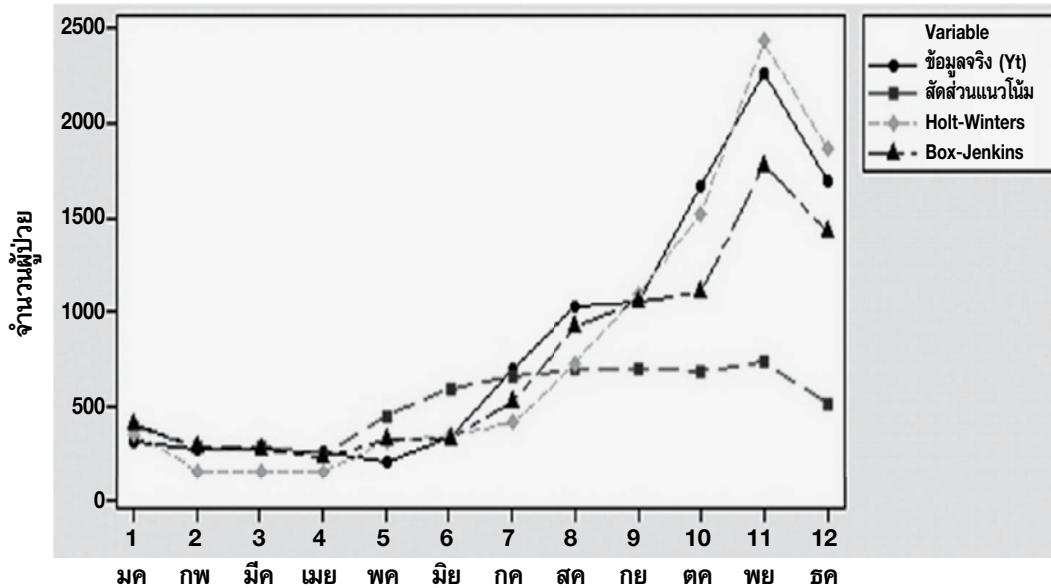
$$Z_t = \theta_0 + \phi_{12}Z_{t-12} + \varepsilon_t$$

$$Z_t = \ln(Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13})$$

มีค่าประมาณพารามิเตอร์ของรูปแบบ คือ  $\hat{\theta}_0$  เท่ากับ 0.003 และ  $\hat{\phi}_{12}$  เท่ากับ -0.678 ได้สมการพยากรณ์ คือ

$$\hat{Z}_t = 0.003 + (-0.678)Z_{t-12} ; t = 97, 98, \dots$$

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการพยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 7



รูปที่ 7 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคไข้เลือดออก

**โรคไข้หวัดใหญ่**

กำหนดค่าคงที่ปรับให้เรียบ  $\alpha$  มีค่าเท่ากับ 0.4  
 ค่าคงที่ปรับให้เรียบสำหรับแนวโน้ม  $\gamma$  มีค่าเท่ากับ 0.4

ค่าคงที่ปรับให้เรียบสำหรับฤดูกาล  $\delta$  มีค่าเท่ากับ 0.4 ซึ่งเป็นค่าปรับให้เรียบที่ให้ค่า MSE ต่ำที่สุด และได้ค่าดัชนีฤดูกาลรายเดือน ดังตารางที่ 3

ตารางที่ 3 ค่าดัชนีฤดูกาลรายเดือน (โรคไข้หวัดใหญ่)

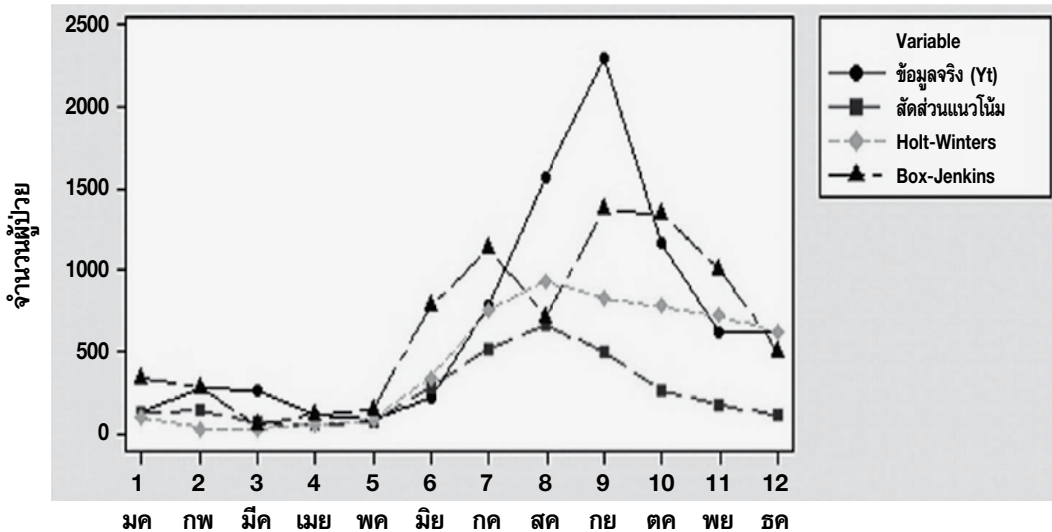
เดือน	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.
$\hat{S}_t$	0.644	1.119	0.607	0.233	0.201	0.666
เดือน	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
$\hat{S}_t$	2.507	1.748	0.693	0.384	0.311	1.873

รูปแบบสมการพยากรณ์คือ

$$\hat{Y}_t = (3,986.18 + p(569.900))\hat{S}_i$$

$$p = 1, 2, \dots, i = 1, 2, \dots, 12$$

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการ  
พยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 8



รูปที่ 8 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคไข้หวัดใหญ่

### โรคปอดอักเสบ

กำหนดค่าคงที่ปรับให้เรียบ  $\alpha$  มีค่าเท่ากับ 0.08  
ค่าคงที่ปรับให้เรียบสำหรับแนวโน้ม  $\gamma$  มีค่าเท่ากับ 0.08

ค่าคงที่ปรับให้เรียบสำหรับฤดูกาล  $\delta$  มีค่าเท่ากับ 0.08 ซึ่ง  
เป็นค่าปรับให้เรียบที่ให้ค่า MSE ต่ำที่สุด และได้ค่าดัชนี  
ฤดูกาลรายเดือน ดังตารางที่ 4

ตารางที่ 4 ค่าดัชนีฤดูกาลรายเดือน (โรคปอดอักเสบ)

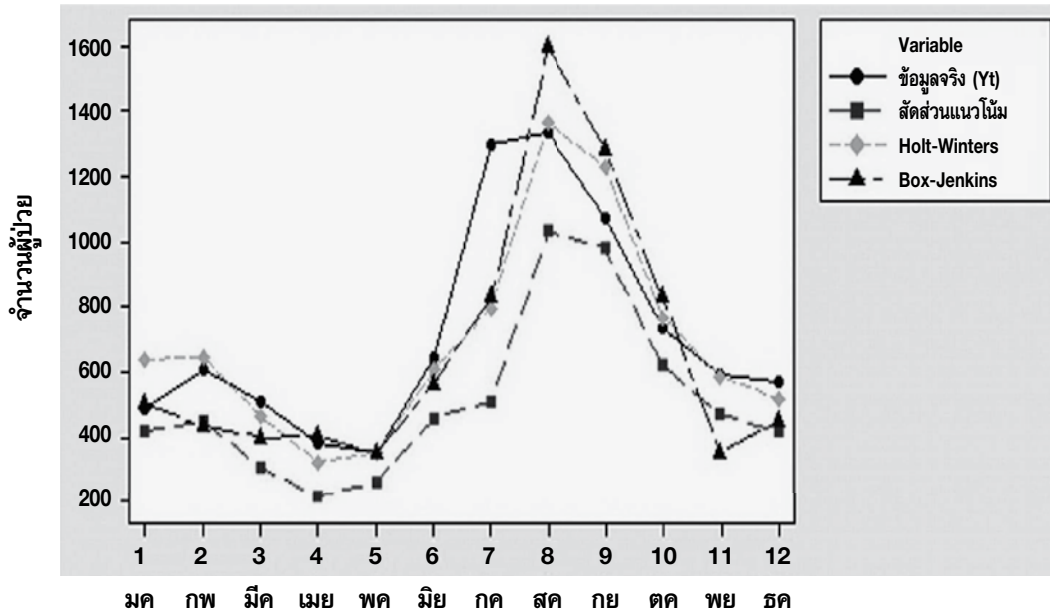
เดือน	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.
$\hat{S}_i$	0.844	0.890	0.656	0.457	0.489	0.854
เดือน	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
$\hat{S}_i$	1.182	1.845	1.656	1.056	0.814	0.731

รูปแบบสมการพยากรณ์คือ

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการ  
พยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 9

$$\hat{Y}_t = (714.085 + p(-0.347))\hat{S}_t$$

$$p = 1, 2, \dots, i = 1, 2, \dots, 12$$



รูปที่ 9 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคปอดอักเสบ

**โรคสุกใส**

กำหนดค่าคงที่ปรับให้เรียบ  $\alpha$  มีค่าเท่ากับ 0.3  
ค่าคงที่ปรับให้เรียบสำหรับแนวโน้ม  $\gamma$  มีค่าเท่ากับ 0.3

ค่าคงที่ปรับให้เรียบสำหรับฤดูกาล  $\delta$  มีค่าเท่ากับ 0.3 ซึ่ง  
เป็นค่าปรับให้เรียบที่ให้ค่า MSE ต่ำที่สุด และได้ค่าดัชนี  
ฤดูกาลรายเดือน ดังตารางที่ 5

ตารางที่ 5 ค่าดัชนีฤดูกาลรายเดือน (โรค สุกใส)

เดือน	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.
$\hat{S}_t$	1.626	1.955	2.010	1.508	1.328	1.076
เดือน	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
$\hat{S}_t$	0.730	0.699	0.579	0.620	0.506	0.642

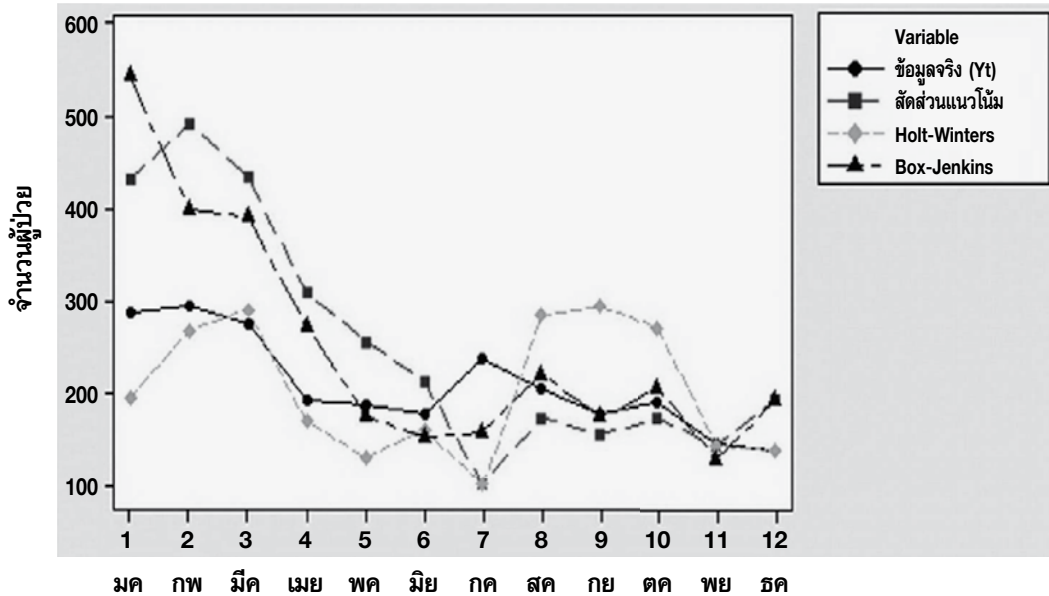


รูปแบบสมการพยากรณ์คือ

$$\hat{Y}_i = (298.628 + p(9.200))\hat{S}_i$$

$$p = 1, 2, \dots, i = 1, 2, \dots, 12$$

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการ  
พยากรณ์ทั้ง 3 วิธีกับข้อมูลจริงแสดงดังรูปที่ 10



รูปที่ 10 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคสุกใส

### โรคแผลในปาก

กำหนดค่าคงที่ปรับให้เรียบ  $\alpha$  มีค่าเท่ากับ 0.1  
ค่าคงที่ปรับให้เรียบสำหรับแนวโน้ม  $\gamma$  มีค่าเท่ากับ 0.1

ค่าคงที่ปรับให้เรียบสำหรับฤดูกาล  $\delta$  มีค่าเท่ากับ 0.1 ซึ่ง  
เป็นค่าปรับให้เรียบที่ให้ค่า MSE ต่ำที่สุด และได้ค่าดัชนี  
ฤดูกาลรายเดือน ดังตารางที่ 6

ตารางที่ 6 ค่าดัชนีฤดูกาลรายเดือน (โรคแผลในปาก)

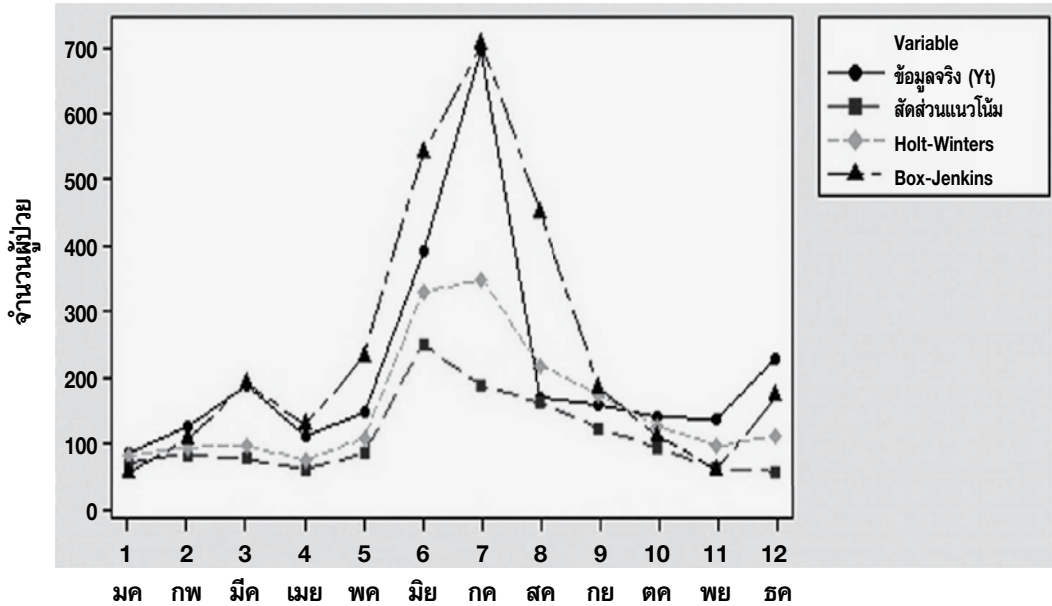
เดือน	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.
$\hat{S}_i$	0.625	0.758	0.800	0.543	0.744	2.147
เดือน	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
$\hat{S}_i$	2.358	1.192	0.971	0.721	0.550	0.634

รูปแบบสมการพยากรณ์คือ

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการ  
พยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 11

$$\hat{Y}_t = (211.718 + p(4.968))\hat{S}_t$$

$$p = 1, 2, \dots, i = 1, 2, \dots, 12$$



รูปที่ 11 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคแผลในปาก

**โรคอุจจาระร่วงเฉียบพลัน**

กำหนดค่าคงที่ปรับให้เรียบ  $\alpha$  มีค่าเท่ากับ 0.1  
ค่าคงที่ปรับให้เรียบสำหรับแนวโน้ม  $\gamma$  มีค่าเท่ากับ 0.1

ค่าคงที่ปรับให้เรียบสำหรับฤดูกาล  $\delta$  มีค่าเท่ากับ 0.1 ซึ่ง  
เป็นค่าปรับให้เรียบที่ให้ค่า MSE ต่ำที่สุด และได้ค่าดัชนี  
ฤดูกาลรายเดือน ดังตารางที่ 7

ตารางที่ 7 ค่าดัชนีฤดูกาลรายเดือน (โรคอุจจาระร่วงเฉียบพลัน)

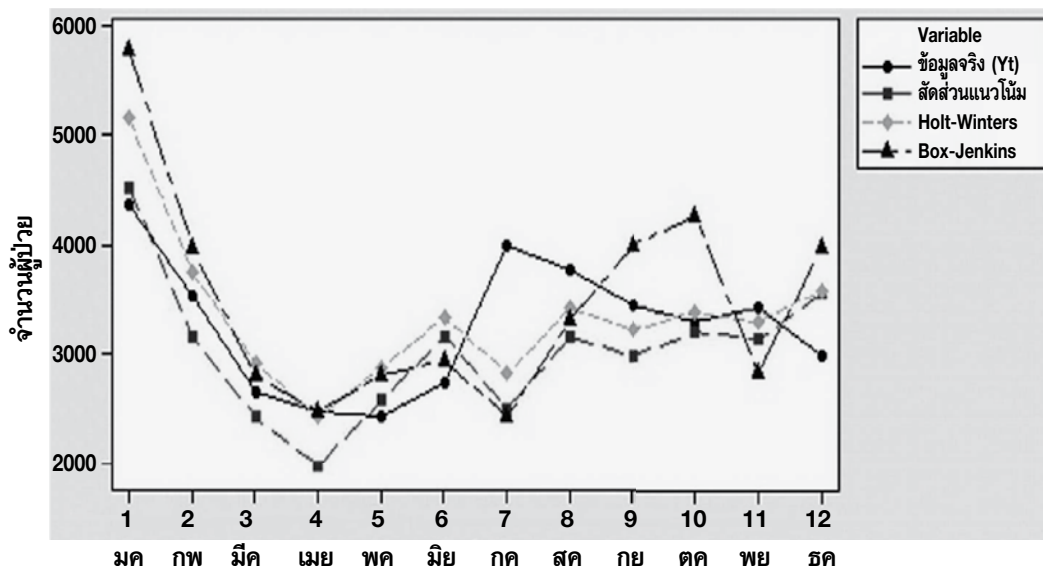
เดือน	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.
$\hat{S}_i$	1413.16	156.13	-652.49	-1068.30	-648.12	-137.31
เดือน	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
$\hat{S}_i$	-394.49	29.07	-205.23	-82.28	-137.19	92.49

รูปแบบสมการพยากรณ์คือ

$$\hat{Y}_t = 3,377.290 + p(-17.896) + \hat{S}_t$$

$$p = 1, 2, \dots, i = 1, 2, \dots, 12$$

ผลจากการเปรียบเทียบค่าพยากรณ์จากวิธีการพยากรณ์ทั้ง 3 วิธีกับข้อมูลจริง แสดงดังรูปที่ 12



รูปที่ 12 กราฟแสดงการเปรียบเทียบค่าพยากรณ์จากตัวแบบสมการพยากรณ์กับข้อมูลจริงของโรคอุจจาระร่วงเฉียบพลัน

#### 4. สรุปและอภิปรายผล

จากการศึกษาเปรียบเทียบวิธีการพยากรณ์อนุกรมเวลาที่ผ่านมาวิเคราะห์ทั้ง 5 วิธี โดยแบ่งวิธีการศึกษาตามลักษณะองค์ประกอบของข้อมูลได้ 2 ประเภท คือ ข้อมูลโรคที่ไม่มีแนวโน้มและฤดูกาลจำนวน 4 โรค ได้แก่ โรคอาหารเป็นพิษ โรคฉี่หนูโรคปอด โรคตาแดง โรคมือเท้าปาก ทำการเปรียบเทียบโดยใช้ วิธีเฉลี่ยเคลื่อนที่แบบง่าย วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบง่ายและวิธีบอซ-เจนกินส์ กับข้อมูลโรคที่มีแนวโน้มและฤดูกาลจำนวน 6 โรค ได้แก่ โรคไข้เลือดออก โรคไข้หวัดใหญ่ โรคปอดอักเสบ โรคสกุสโร โรคแผลในปาก โรคอุจจาระร่วงเฉียบพลัน ทำการเปรียบเทียบโดยใช้วิธีสัดส่วนกับแนวโน้ม วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโบลท์-วินเทอร์ และ วิธีบอซ-เจนกินส์ โดยพิจารณาจากค่า MAPE ที่ต่ำที่สุดพบว่าวิธีการที่ให้รูปแบบพยากรณ์ที่เหมาะสมที่สุดสำหรับ

ข้อมูลโรคที่ไม่มีแนวโน้มและฤดูกาลส่วนใหญ่คือ วิธีเฉลี่ยเคลื่อนที่แบบง่าย วิธีการที่ให้รูปแบบพยากรณ์ที่เหมาะสมที่สุดสำหรับข้อมูลโรคที่มีแนวโน้มและฤดูกาลส่วนใหญ่คือ วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโบลท์-วินเทอร์

อย่างไรก็ตามแม้ว่าวิธีการพยากรณ์ต่างๆ ที่นำมาทำการศึกษเปรียบเทียบบางวิธีอาจไม่ได้ให้ค่า MAPE ที่ต่ำที่สุด แต่ให้ค่า MAPE ที่ใกล้เคียงกันแสดงว่าเป็นวิธีการพยากรณ์ที่มีประสิทธิภาพใกล้เคียงกันเราจึงอาจเลือกใช้วิธีการพยากรณ์วิธีใดก็ได้ที่ผู้เกี่ยวข้องในการปฏิบัติการเห็นว่าสะดวกและสามารถศึกษาทำความเข้าใจได้ง่ายในขั้นตอนของการวิเคราะห์ เช่นโรคมือเท้าปาก จะใช้วิธีเฉลี่ยเคลื่อนที่แบบง่ายหรือวิธีเอ็กซ์โพเนนเชียลแบบง่าย ค่าพยากรณ์ที่ได้ก็อาจไม่แตกต่างกันมากนัก เป็นต้น แต่ถ้าค่า MAPE มีค่าแตกต่างกันมากก็ควรเลือกใช้วิธีการพยากรณ์ที่ให้ค่า MAPE ต่ำที่สุด แม้ว่าจะเป็นวิธีที่ยาก

ต่อการทำความเข้าใจก็ตาม เพราะจะทำให้ได้ค่าพยากรณ์ที่แม่นยำกว่า เช่น โรคไข้เลือดออก ควรใช้วิธีบอกซ์-เจนกินส์ ซึ่งสอดคล้องกับงานวิจัยในอดีต [4,5,11] เป็นต้น

## 5. ข้อเสนอแนะ

1. ผลการศึกษาวิจัยนี้สามารถนำไปใช้ประโยชน์ในการวางแผนและป้องกันรวมถึงการควบคุมโรคที่ต้องเฝ้าระวังทางระบาดวิทยาที่จะเกิดขึ้นในอนาคต

2. ตัวแบบการพยากรณ์หรือวิธีการพยากรณ์อาจเปลี่ยนแปลงได้ตามความเหมาะสม เมื่อเปลี่ยนข้อมูลโรคที่ต้องเฝ้าระวังทางระบาดวิทยา

3. จากรูปแบบสมการพยากรณ์ที่เหมาะสมในการวิจัยนี้สามารถนำไปประยุกต์ใช้กับอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวที่คล้ายคลึงกัน

4. เมื่อพิจารณาค่า MAPE ที่ได้จากการวิจัยนี้พบว่ายังมีค่าค่อนข้างสูงอาจเนื่องมาจากจำนวนข้อมูลผู้ป่วยย้อนหลังที่นำมาทำการวิเคราะห์ยังมีไม่มากพอ เนื่องจากสำนักระบาดวิทยา กรุงเทพมหานคร ทำการเก็บข้อมูลในอดีตได้เพียง 5 ปี ซึ่งเป็นข้อจำกัดของการวิจัยในครั้งนี้ และไม่ได้มีการนำปัจจัยอื่นๆ เช่น การระบาดของโรคที่ผิดปกติในบางปีมาพิจารณาด้วย ซึ่งเป็นข้อดีของการพยากรณ์โดยใช้เทคนิคอนุกรมเวลา ที่ทำการศึกษการเปลี่ยนแปลงของข้อมูลอันเนื่องมาจากเวลาเท่านั้น ในการวิจัยครั้งต่อไปควรเพิ่มข้อมูลที่นำมาทำการวิเคราะห์ให้มากขึ้น และอาจศึกษาวิธีการพยากรณ์ที่แตกต่างจากการศึกษาครั้งนี้เพื่อที่จะได้รูปแบบการพยากรณ์ที่เหมาะสมกับข้อมูลต่อไป เช่น การวิเคราะห์การถดถอย ซึ่งเป็นรูปแบบที่จะนำตัวแปรที่คาดว่าจะมีผลต่อจำนวนประชากรที่ป่วยเป็นโรคไข้หวัดใหญ่มาพิจารณาเพิ่มเติม นอกเหนือจากตัวแปรเวลาเพียงอย่างเดียว เพื่อให้ผลของการพยากรณ์มีความถูกต้องแม่นยำน่าเชื่อถือมากยิ่งขึ้น

## 6. กิตติกรรมประกาศ

ผู้วิจัยใคร่ขอขอบคุณสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข กรุงเทพมหานคร ที่ได้ให้ความอนุเคราะห์ข้อมูลสำหรับการวิจัยและขอขอบคุณ อาจารย์ธนศ จิตต์สุภาพรรณ ภาควิชาคณิตศาสตร์ สาขาสถิติ

คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ที่ได้คำปรึกษาแนะนำงานวิจัยสำเร็จลุล่วงด้วยดี

## 7. เอกสารอ้างอิง

1. Suthachana, S., Kaewnorkkao, W., Suangtho, P., Sayumpurujinun, S. and Kongyu, S., 2012, "Forecasting the situation of influenza in Thailand in 2012", Workshop Conference in Analyzed and Forecasting, Bureau of Epidemiology Department of Disease control Ministry of Public Health, 12-14 september 2012, Narai Hotel Bangkok, (In Thai).
2. Manmin, M., 2006, Time Series and Forecasting. Prakaiphruk Printing, Bangkok, (In Thai).
3. Yangkao, K., Srakeaw, K. and Riabporn, V., 2013, "Dengue hemorrhagic Fever : A Forecast Report in 9th service network area for 2013", Weekly Epidemiological Surveillance Report, Vol. 44, 577-584, (In Thai).
4. Promprou, S., Jaroensutasinee, M. and Jaroensutasinee, K., 2006, "Forecasting Dengue Haemorrhagic Fever Cases in Southern Thailand using ARIMA Models", Dengue Bulletin, Vol. 30, 99-106.
5. Wongkoon, S., Pollar, M., Jaroensutasinee, M. and Jaroensutasinee, K., 2008, "Predicting DHF Incidence in Northern Thailand Using Time Series Analysis Technique", International Journal of Biological and Life Sciences, Vol.4, No.3, 117-121.
6. Weisent, J., Seaver, W., Odol, A. and Rohrbach, B., 2010, "Comparison of Three Time-Series Models for Predicting ampylobacteriosis Risk", Epidemiol. Infect., Vol. 138, 898-906.
7. Al-Sakkaf, A. and Jones, G., 2013 "Comparison of Time Series Models for Predicting Campylobacteriosis Risk in New Zealand", Zoonoses Public Health, Apr 2. doi: 10.1111/zph.12046.
8. Taesombut, S., 2006, Quantitative Forecasting,

1<sup>st</sup> ed., Kasetsart University Printing, Bangkok (In Thai).

9. Bowerman, B.L., O' Connell, R.T. and Koehler, A.B., 2005. Forecasting, Time series and Regression: an Applied Approach, 4<sup>th</sup> ed., Thomson Books, The United States of America.

10. Getaim, S., 2005, Forecasting Techniques, 2<sup>nd</sup> ed. Educational Administration Group, Thaksin

University, (In Thai).

11. Gharbi, M., Quenel, P., Gustave. J., Cas-sadou, S., Ruche, G.L., Girdary, L. and Marrama, L., 2011, "Time Series Analysis of Dengue Incidence in Guadeloupe, French West Indies: Forecasting Models Using Climate Variables as Predictors", BMC Infectious Diseases, Vol.11, No.166, 1-13.

