j 4504

วารสารวิทยาศาสตร์และเทคโนโลยี ปีที่ 22 ฉบับที่ 1 มกราคม - มีนาคม 2557

การพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวในประเทศไทยโดยวิธีบอกซ์-เจนกินส์และวิธีการของวินเตอร์

A Forecasting Methods for the Number of International Tourists in Thailand: Box-Jenkins Method and Winter's Method

ลักขณา เศาธยะนันท์, สุณี ทวีสกุลวัชระ, ยุพิน กาญจนะศักดิ์ดา และบุญหญิง สมร่าง*

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหอการค้าไทย

ถนนวิภาวดีรังสิต เขตดินแดง กรุงเทพฯ 10400

Luckhana Saothayanun, Sunee Taweesakulvatchara, Yupin Kanjanasakda and Boonying Somrang*

School of Science and Technology, University of the Thai Chamber of Commerce,

Vibhavadi Rangsit Road, Din Daeng, Bangkok 10400

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวในประเทศไทย 2 วิธี คือ วิธีบอกซ์-เจนกินส์ และวิธีการของวินเตอร์ การเปรียบเทียบพิจารณาจากรากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (root mean square error, RMSE) และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (mean absolute percent error, MAPE) ผลการวิจัยพบว่าการพยากรณ์โดยวิธีการของวินเตอร์ให้ค่า RMSE และ MAPE ต่ำกว่าวิธีบอกซ์-เจนกินส์ในทุกช่วงเวลา

คำสำคัญ : นักท่องเที่ยวต่างชาติ; การพยากรณ์; การวิเคราะห์อนุกรมเวลา

Abstract

This research aims to compare two forecasting methods for the number of international tourists in Thailand. The methods are Box–Jenkins and Winter's methods. Results show that Winter's method gives Root Mean Square Error (RMSE) and Mean Absolute Percent Error (MAPE) lower than Box-Jenkins method for all durations.

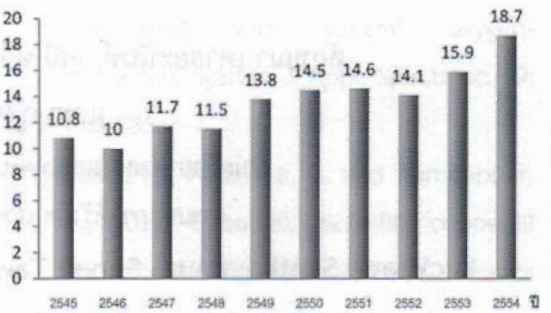
Keywords: international tourists; forecasting; time series analysis

***ผู้รับผิดชอบบทความ :** boonying_som@utcc.ac.th

1. บทนำ

อุตสาหกรรมท่องเที่ยวถือว่าเป็นอุตสาหกรรมที่มีความสำคัญในการพัฒนาเศรษฐกิจและสังคมของประเทศไทย การท่องเที่ยวเป็นแหล่งที่มาของเงินตราต่างประเทศที่สำคัญที่สุด รายได้จากอุตสาหกรรมท่องเที่ยวที่ได้มาในรูปของเงินตราต่างประเทศมีส่วนช่วยสร้างเสถียรภาพให้กับดุลการชำระเงิน นอกจากนี้อุตสาหกรรมการท่องเที่ยวยังเป็นตัวกระตุ้นให้เกิดการขยายตัวทางเศรษฐกิจซึ่งนำไปสู่การจ้างงาน สร้างอาชีพ เป็นการกระจายรายได้และการลงทุนในธุรกิจที่เกี่ยวเนื่องกันมากมายทั้งที่เกี่ยวข้องโดยตรง ได้แก่ ด้านการขนส่ง ที่พัก อาหาร และธุรกิจนำเที่ยว ส่วนที่เกี่ยวข้องทางอ้อม ได้แก่ การผลิตสินค้า เป็นต้น จะเห็นได้ว่าในหลายประเทศ อุตสาหกรรมท่องเที่ยวเป็นอุตสาหกรรมที่มีความสำคัญอยู่ในระดับที่ 1-3 ของ 10 อันดับแรกในอุตสาหกรรมสำคัญของประเทศนั้น ๆ และเป็นที่เชื่อกันว่าความสำคัญของการท่องเที่ยวจะมีมากขึ้นในอนาคต จะเห็นได้จากการที่รัฐบาลมีนโยบายส่งเสริมอย่างต่อเนื่อง เนื่องจากอุตสาหกรรมการท่องเที่ยวมีบทบาทสำคัญทั้งในด้านเศรษฐกิจ สังคม และการเมือง ในช่วงระยะเวลา 10 ปี ที่ผ่านมา (2545–2554) จำนวนนักท่องเที่ยวชาวต่างชาติที่เดินทาง เข้ามาในประเทศไทย มีแนวโน้มสูงขึ้นอย่างต่อเนื่อง จาก 10.8 ล้านคน ในปี 2545 เป็น 14.6 ล้านคน ในปี 2551 และลดลงเป็น 14.1 ล้านคน ในปี 2552 และในปีต่อ ๆ มาจำนวนนักท่องเที่ยวก็เพิ่มขึ้นเป็น 15.9 ล้านคน และ 18.7 ล้านคน ในปี 2553 และ 2554 ตาม ลำดับ ดังรูปที่ 1 [1]

และในปี 2555 มีนักท่องเที่ยวเข้ามาในประเทศไทย 22.3 ล้านคน คาดว่าในปี 2556 จะมีนักท่องเที่ยวเพิ่มขึ้นเป็น 24 ล้านคน ขึ้นไป คิดเป็นเฉลี่ย 2 ล้านคนต่อเดือน [2] นอกจากนี้ผลการสำรวจของ Master Global Destination Cities Index 2013 ยกให้กรุงเทพมหานครได้รับตำแหน่งชนะเลิศอันดับหนึ่ง สุดยอดเมืองท่องเที่ยวของโลก [3] ดังนั้นการคาดการณ์จำนวนนักท่องเที่ยวจึงมีความสำคัญต่อหน่วยงานต่าง ๆ ทั้งภาครัฐและเอกชนที่เกี่ยวข้องกับอุตสาหกรรมการท่องเที่ยว สามารถนำมาใช้ประโยชน์ในการวางแผนและการดำเนินงานเพื่อรองรับจำนวนนักท่องเที่ยวที่เพิ่มขึ้น



รูปที่ 1 จำนวนนักท่องเที่ยวชาวต่างชาติ พ.ศ. 2545-2554

จากการศึกษางานวิจัยเกี่ยวกับการพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวในประเทศไทยของอัครพงศ์และปวีณา [4] งานวิจัยนี้ใช้ข้อมูลรายเดือนระหว่างเดือนมกราคม 2528 ถึงเดือนธันวาคม 2548 สร้างสมการทำนายจำนวนนักท่องเที่ยวต่างชาติที่เดินทางเข้ามาท่องเที่ยวในประเทศไทยตั้งแต่ปี พ.ศ. 2550-2554 ด้วยวิธี SARIMA intervention ซึ่งแบบจำลองดังกล่าวให้ผลการทำนายที่มีค่า root mean square error (RMSE) ที่ต่ำ ผลการทำนายจำนวนนักท่องเที่ยวต่างชาติที่จะเดินทางเข้ามาท่องเที่ยวในประเทศไทยปี 2554 มีประมาณ 17 ล้านคน และได้รับรายได้จากนักท่องเที่ยวต่างชาติประมาณ 6.89 แสนล้านบาท โดยนักท่องเที่ยวจากโอเชียเนียและตะวันออกกลางมีอัตราการขยายตัวมากที่สุด รองลงมา ได้แก่ ตลาด

ยุโรป อเมริกา และเอเชียใต้ ส่วนเอเชียตะวันออกและอัฟริกาจะมีอัตราการขยายตัวต่ำที่สุดประมาณร้อยละ 4-5 ต่อปี [4]

ไอริณ [5] ได้สร้างสมการพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่จะมาท่องเที่ยวในประเทศไทยด้วยเทคนิคการวิเคราะห์อนุกรมเวลา และได้ทำการพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่มาท่องเที่ยวในประเทศไทยเป็นรายเดือน ระหว่างปี พ.ศ. 2538-2547

ประเสริฐ และคณะ [6] ได้ศึกษาการพยากรณ์จำนวนนักท่องเที่ยวชาวต่างชาติที่เดินทางเข้ามาในประเทศไทยในระหว่างปี ค.ศ. 2006-2010 โดยใช้วิธีการพยากรณ์ SARIMA, ARIMA, Holt-Winter-Additive-Multiplicative-No seasonal, Neural network, Var, GMM ซึ่งในการศึกษาครั้งนี้พบว่าแบบจำลองที่ดีที่สุดที่ใช้ตัวแปรเดียวคือ SARIMA (0,1,1)(0,1,4) และในแบบจำลองที่ใช้มากกว่าหนึ่งตัวแปรคือ VAR

เมตตา [7] ได้พัฒนาระบบพยากรณ์จำนวนนักท่องเที่ยวโดยใช้ โครงข่ายใยประสาทเทียมผ่านเครือข่ายอินเทอร์เน็ตระบบพัฒนาเป็นเว็บแอพพลิเคชันด้วยโปรแกรมภาษา PHP ทำงานบนระบบปฏิบัติการ Microsoft Windows 2000 และระบบการจัดการฐานข้อมูล ORACLE9I โดยลักษณะของระบบพยากรณ์จะพยากรณ์ข้อมูลล่วงหน้าเป็นรายเดือน เพื่อให้ผู้ใช้สามารถนำข้อมูลที่ได้จากการพยากรณ์ไปช่วยในการวางแผนงานเรื่องการบริหารงานกับนักท่องเที่ยว โดยทราบจำนวนได้ล่วงหน้าจึงเกิดความสะดวกและทันเวลา และสรุปได้ว่าการทำงานของระบบที่พัฒนาขึ้นมีประเมินคุณภาพอยู่ในเกณฑ์ดีและสามารถที่จะนำไปใช้ในงานจริงได้อย่างเหมาะสม

คณะผู้วิจัยได้เล็งเห็นว่าการขยายตัวของนักท่องเที่ยวต่างชาติมีความสำคัญต่อระบบเศรษฐกิจและสังคมของประเทศไทย และเนื่องจากการพยากรณ์ด้วยวิธีบอกซ์-เจนกินส์เป็นวิธีที่นิยมและใช้กันอย่างแพร่หลายเหมาะสำหรับการพยากรณ์ในระยะสั้น และจำนวนนักท่องเที่ยวมีองค์ประกอบของฤดูกาลมาเกี่ยวข้อง จึงสนใจที่จะพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทาง มาท่องเที่ยวในประเทศไทยโดยใช้ตัวแบบพยากรณ์บอกซ์-เจนกินส์เปรียบเทียบกับวิธีการของวินเตอร์

2. วิธีดำเนินการวิจัย

การสร้างตัวแบบพยากรณ์จำนวนนักท่องเที่ยวต่างชาติ มีขั้นตอนดังนี้คือ

2.1 การจัดเตรียมข้อมูล

ข้อมูลที่นำมาใช้ในการสร้างตัวแบบพยากรณ์เป็นจำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวประเทศไทยรายเดือน ตั้งแต่เดือนมกราคม 2540 ถึงเดือนกันยายน 2555 จำนวน 189 เดือน ที่มาของแหล่งข้อมูล คือ เว็ปไซต์ของธนาคารแห่งประเทศไทย [8]

2.2 การวิเคราะห์ข้อมูล

วิเคราะห์ข้อมูลโดยเปรียบเทียบค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (root mean square error, RMSE) และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (mean absolute percent error, MAPE) ของวิธีพยากรณ์ 2 วิธี คือ วิธีพยากรณ์ของบอกซ์-เจนกินส์ และวิธีการพยากรณ์ของวินเตอร์

2.2.1 วิธีพยากรณ์ของบอกซ์-เจนกินส์ [9] เป็นวิธีที่ใช้สำหรับเลือกรูปแบบที่เหมาะสมให้กับอนุกรมเวลา โดยพิจารณาจากสหสัมพันธ์ระหว่าง Y ที่คาบเวลา t (Y_t) และที่คาบเวลาที่ผ่านมา (Y_{t-1}, Y_{t-2}, ...) เมื่อได้ตัวแบบที่เหมาะสมแล้วจะใช้ตัวแบบนี้ในการพยากรณ์ Y_{t+1}, Y_{t+2}, ... ในอนาคต อนุกรมเวลาที่จะกำหนดรูปแบบโดยวิธีบอกซ์-เจนกินส์จะต้องเป็น

อนุกรมเวลาที่อยู่ในสภาวะนิ่ง (stationary data series) เท่านั้น ซึ่งหมายถึงคงที่ในค่าเฉลี่ย ความแปรปรวน และความแปรปรวนร่วม ไม่แปรผันตามเวลา

ดังนั้นขั้นตอนของวิธีพยากรณ์ของบอกซ์-เจนกินส์ที่สำคัญประกอบด้วย 5 ขั้นตอน ได้แก่

(1) ตรวจสอบสภาวะนิ่งโดยพิจารณาจากกราฟของอนุกรมเวลาและจากกราฟฟังก์ชันอัตสหสัมพันธ์ตัวอย่าง (sample autocorrelation function, SACF) แทนด้วย r_k

(2) ถ้าตรวจสอบแล้วพบว่าอนุกรมเวลาไม่อยู่ในสภาวะคงที่ในค่าเฉลี่ยจะทำการแปลงเป็นอนุกรมเวลาชุดใหม่ที่มีลักษณะคงที่ในค่าเฉลี่ย โดยการหาผลต่างของอนุกรมเวลาและถ้าไม่คงที่ในความแปรปรวนจะแปลงอนุกรมเวลาด้วยลอการิทึมธรรมชาติหรือรากที่สอง เป็นต้น

(3) กำหนดตัวแบบที่คาดว่าจะเหมาะสมให้กับอนุกรมเวลาโดยพิจารณาจากกราฟ SACF (sample autocorrelation function, SACF) และ SPACF (sample partial autocorrelation function, SPACF) แทนด้วย r_{kk}

(4) ประมาณค่าพารามิเตอร์ในสมการที่เลือกไว้ด้วยวิธีกำลังสองน้อยที่สุด (ordinary least squares method)

(5) ตรวจสอบความเหมาะสมของตัวแบบจากกราฟ SACF และ SPACF ของส่วนตกค้าง (residuals: $e_t = Y_t - \hat{Y}_t$) ด้วยวิธีบอกซ์-เจนกินส์จะได้ตัวแบบอนุกรมเวลาที่เรียกว่าตัวแบบ ARIMA (autoregressive integrated moving average) และตัวแบบที่ศึกษาในครั้งนี้เป็นตัวแบบ ARIMA (p,d,q)(P,D,Q)s มีรูปแบบดังนี้

$$\varphi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Y_t = \delta + \theta_q(B)\Theta_Q(B^s)a_t$$

โดยที่ $\varphi_p(B) = 1 - \varphi_1 B - \varphi_2 B^2 - ... - \varphi_p B^p$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q$$

$$\Phi_P(B^s) = 1 - \Phi_s B^s - \Phi_{2s} B^{2s} - ... - \Phi_{Ps} B^{Ps}$$

$$\Theta_Q(B^s) = 1 - \Theta_s B^s - \Theta_{2s} B^{2s} - ... - \Theta_{Qs} B^{Qs}$$

โดยที่ $\varphi_1,...,\varphi_p$ คือ สัมประสิทธิ์การถดถอย (autoregressive coefficient)

$\theta_1,...,\theta_q$ คือ สัมประสิทธิ์ค่าเฉลี่ยเคลื่อนที่ (moving average coefficient)

$\Phi_s,...,\Phi_{Ps}$ คือ สัมประสิทธิ์การถดถอยในส่วนที่มีฤดูกาล (seasonal autoregressive coefficients)

$\Theta_s,...,\Theta_{Qs}$ คือ สัมประสิทธิ์ค่าเฉลี่ยเคลื่อนที่ในส่วนที่มีฤดูกาล (seasonal moving average coefficients)

B คือ ตัวดำเนินการถอยหลังเวลา (backward shift operator) นั่นคือ $B^m Y_t = Y_{t-m}$

d คือ จำนวนครั้งของการทำผลต่างเพื่อให้อนุกรมเวลา $\{Y_t\}$ เป็นอนุกรมเวลาที่อยู่ในสภาวะคงที่ในค่าเฉลี่ย

p คือ อันดับของตัวแบบการถดถอย

q คือ อันดับของตัวแบบค่าเฉลี่ยเคลื่อนที่

s คือ คาบเวลาของฤดูกาล

P คือ อันดับในส่วนของกระบวนการ AR ในส่วนที่มีฤดูกาล

Q คือ อันดับในส่วนของกระบวนการ MA ในส่วนที่มีฤดูกาล

D คือ จำนวนครั้งทำผลต่างอนุกรมเวลา ในส่วนที่มีฤดูกาล

a_t คือ ตัวแปรสุ่มอิสระและมีการแจกแจงปกติ ซึ่งมีค่าเฉลี่ยเท่ากับศูนย์และความแปรปรวนคงที่ให้เท่ากับ σ_a^2 เรียก a_t ว่ากระตุกสุ่ม (random shock)

2.2.2 วิธีพยากรณ์ของวินเตอร์ [10] สำหรับอนุกรมเวลา $\{Y_t\}$ มีแนวโน้มเชิงเส้น มีองค์ประกอบฤดูกาลและการเปลี่ยนแปลงหรือการแกว่ง

ของฤดูกาลไม่คงที่ แปรผันตามเวลาหรือแปรผันตามระดับของอนุกรมเวลา กล่าวคือ การแกว่งจะมากขึ้นขณะที่ระดับของอนุกรมเพิ่มขึ้น ในกรณีนี้ควรเลือกใช้วิธีวินเตอร์เชิงคูณ (multiplicative Winter's method) ซึ่งมีตัวแบบอนุกรมเวลาดังนี้

$$Y_t = (\mu + \beta_t)s_t + a_t, t = 1,2,...,n$$

โดยที่ μ เป็นพารามิเตอร์แทนระดับ หรือค่าเฉลี่ยของอนุกรมเวลา

β เป็นพารามิเตอร์แทนความชันของอนุกรมเวลา

s_t เป็นพารามิเตอร์แทนองค์ประกอบฤดูกาลของอนุกรมเวลา และ

a_t เป็นค่าคลาดเคลื่อนสุ่มหรือองค์ประกอบไม่ปกติ ซึ่งมีข้อสมมติเบื้องต้นคือมีค่าเฉลี่ยเท่ากับศูนย์ ความแปรปรวนคงที่ ไม่มีสหสัมพันธ์กันและมีการแจกแจงปกติ

ตัวแบบพยากรณ์ค่าจริง Y_{t+l} แบบจุดที่เวลา $t+l$ จากเวลาปัจจุบัน t $(t = 1,2,...,l)$ เป็นดังนี้

$$\hat{Y}_t(l) = (\hat{\mu}_t + l\hat{\beta}_t)\hat{s}_{t+l-m},\ l = 1,2,...,m$$

$$\hat{Y}_t(l) = (\hat{\mu}_t + l\hat{\beta}_t)\hat{s}_{t+l-km}, l = (k-1)m+1,(k-1)m+2,...,km;\ k = 2,3,...$$

โดยปรับค่าประมาณของพารามิเตอร์องค์ประกอบต่าง ๆ ดังนี้

$\hat{\mu}_t$ เป็นค่าประมาณระดับ ซึ่ง $\hat{\mu}_t = \alpha(Y_t / \hat{s}_{t-m}) + (1-\alpha)(\hat{\mu}_{t-1} + \hat{\beta}_{t-1})$

α $(0 < \alpha < 1)$ แทนค่าคงที่ปรับให้เรียบสำหรับปรับระดับ (หรือค่าเฉลี่ย) ของอนุกรมเวลา

$\hat{\beta}_t$ เป็นค่าประมาณแนวโน้ม ซึ่ง $\hat{\beta}_t = \gamma(\hat{\mu}_t - \hat{\mu}_{t-1}) + (1-\gamma)\hat{\beta}_{t-1}$

γ $(0 < \gamma < 1)$ แทนค่าคงที่ปรับแนวโน้มหรือความชันของอนุกรมเวลา

$\hat{s}_t$ เป็นค่าประมาณฤดูกาล ซึ่ง $\hat{s}_t = \delta(Y_t / \hat{\mu}_t) + (1-\delta)\hat{s}_{t-m}$

δ $(0 < \delta < 1)$ แทนค่าคงที่ปรับให้เรียบสำหรับฤดูกาลของอนุกรมเวลา

2.3 การเปรียบเทียบวิธีพยากรณ์

การเปรียบเทียบวิธีพยากรณ์ทั้ง 2 วิธี จะพิจารณาจากค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (root mean square error, RMSE) และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (mean absolute percent error, MAPE)

โดยที่ $e_t = Y_t - \hat{Y}_t$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$$

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

3. ผลการวิจัย

3.1 วิธีบอกซ์-เจนกินส์

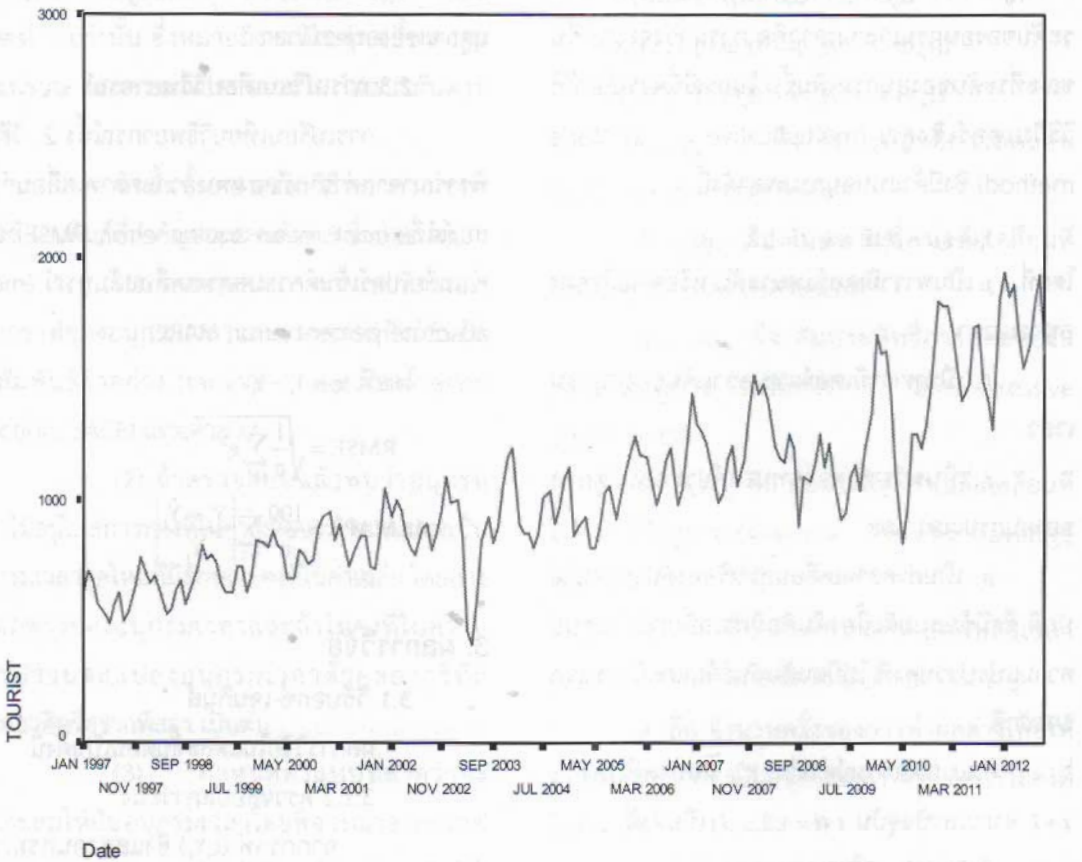
ผลการวิจัยในแต่ละขั้นตอนเป็นดังนี้

3.1.1 ตรวจสอบสภาวะนิ่ง

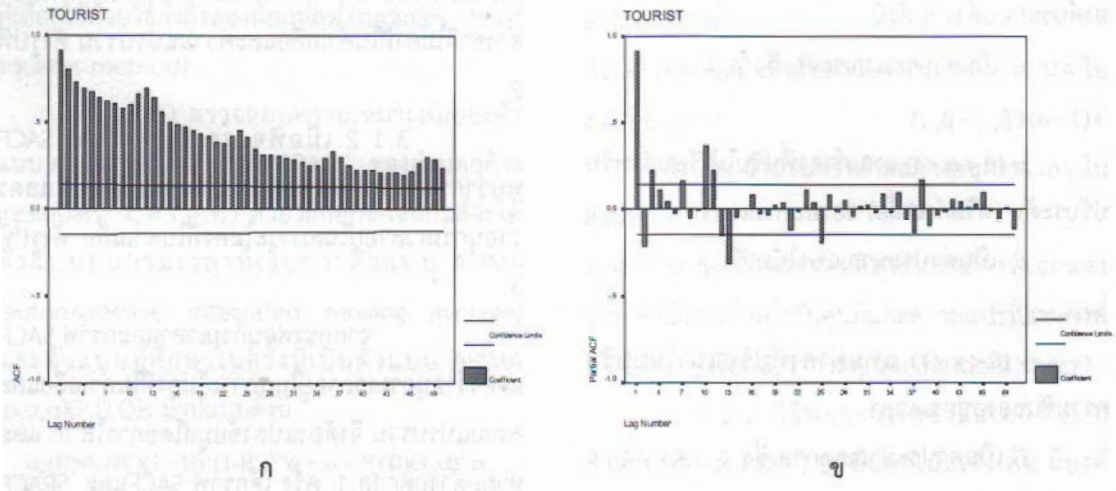
จากกราฟ (t,Y_t) ซึ่งแสดงอนุกรมเวลาจำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวประเทศไทยรายเดือน ตั้งแต่เดือนมกราคม 2540 ถึงเดือนกันยายน 2555 จำนวน 189 เดือน พบว่าการเคลื่อนไหวของอนุกรมเวลาของจำนวนนักท่องเที่ยวต่างชาติไม่คงที่ในค่าเฉลี่ยและความแปรปรวน ดังรูปที่ 2

3.1.2 เมื่อพิจารณาจากกราฟ SACF พบว่าการเคลื่อนไหวของ r_k มีลักษณะลดลงช้า แสดงว่าอนุกรมเวลาอยู่ในสภาวะไม่คงที่ในค่าเฉลี่ย ดังรูปที่ 3

จากกราฟอนุกรมเวลาและกราฟ SACF แสดงว่าอนุกรมเวลาอยู่ในสภาวะไม่คงที่ในค่าเฉลี่ยและความแปรปรวน จึงต้องแปลงข้อมูลโดยการใส่ ln และหาผลต่างฤดูกาล 1 ครั้ง ได้กราฟ SACFและ SPACF ของอนุกรมเวลาชุดใหม่ดังรูปที่ 4

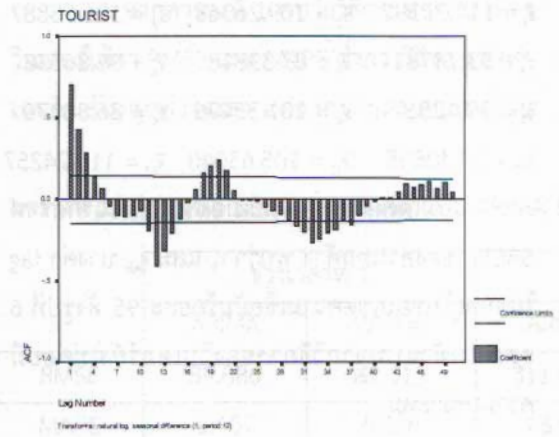


รูปที่ 2 การเคลื่อนไหวของอนุกรมเวลาจำนวนนักท่องเที่ยวต่างชาติ

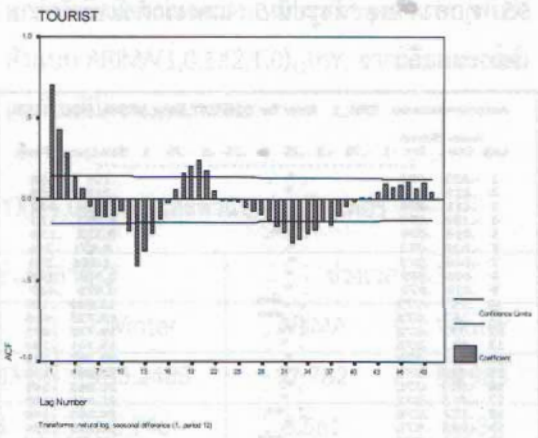


ก ข

รูปที่ 3 กราฟ SACF (ก) และ SPACF (ข) ของอนุกรมเวลาจำนวนนักท่องเที่ยวต่างชาติ



ก



ข

รูปที่ 4 กราฟ SACF (ก) และ SPACF (ข) ของอนุกรมเวลาจำนวนนักท่องเที่ยวต่างชาติที่ใส่ ln และหาผลต่าง 1 ครั้ง

3.1.3 กำหนดตัวแบบที่เหมาะสม โดยพิจารณาจากกราฟ SACF และ SPACF ของอนุกรมเวลาชุดใหม่ พบว่าอนุกรมเวลามีอิทธิพลจากฤดูกาล ดังนั้นตัวแบบที่เหมาะสมคือ ARIMA ของอนุกรมเวลา โดยที่ คือ $(1-\varphi B)(1-\Phi_{12}B^{12}-\Phi_{24}B^{24})(1-B^{12})Z_t = \delta + (1-\theta B)a_t$ หรือ $(1-\varphi B)(1-\Phi_{12}B^{12}-\Phi_{24}B^{24})W_t = \delta + (1-\theta B)a_t$, $W_t = (1-B^{12})Z_t$ โดยที่ $W_t = (1-B^{12})Z_t$ หรือ $W_t = \delta + \varphi W_{t-1} + \Phi_{12}W_{t-12} - \varphi\Phi_{12}W_{t-13} + \Phi_{24}W_{t-24} - \varphi\Phi_{24}W_{t-25} + a_t - \theta a_{t-1}$

3.1.4 จากตัวแบบที่กำหนด ARIMA ของอนุกรมเวลา โดยประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดได้ค่าพารามิเตอร์ดังผลลัพธ์ ดังต่อไปนี้

Standard error	.08876209
Log likelihood	175.27664
AIC	-340.55328
SBC	-324.67253

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	172	1.4296558	.00787871

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
AR1	.66473453	.06963140	9.546477	.00000000
MA1	-.32955975	.08678124	-3.797592	.00020237
SAR1	-.75242215	.07232541	-10.403289	.00000000
SAR2	-.40719619	.07295478	-5.581487	.00000009
CONSTANT	.06973760	.01272573	5.480047	.00000015

และตัวแบบพยากรณ์ของอนุกรม W_t สำหรับเดือนที่ $t = n+l,\ l = 1,2,\ldots$ คือ

$\hat{W}_n(l) = \hat{\delta} + \hat{\varphi}\,\hat{W}_n(l-1) + \hat{\Phi}_{12}\hat{W}_n(l-12) - \hat{\varphi}\,\hat{\Phi}_{12}\hat{W}_n(l-13) + \hat{\Phi}_{24}\hat{W}_n(l-24) - \hat{\varphi}\,\hat{\Phi}_{24}\hat{W}_n(l-25) - \hat{\theta}e_{n+l-2}(1)$ โดยที่ $\hat{\delta} = \hat{\mu}(1-\hat{\varphi}-\hat{\Phi}_{12}-\hat{\Phi}_{24}+\hat{\varphi}\,\hat{\Phi}_{12}+\hat{\varphi}\,\hat{\Phi}_{24})$ และให้ $\hat{W}_n(l-k) = \ln y_{n+l-k} - \ln y_{n+l-k-12}$ สำหรับ $l \le k$ และ $e_{n+l-2}(1) = 0$ สำหรับ $l \ge 2$ จากตัวแบบพยากรณ์ของ W_t ได้ตัวแบบพยากรณ์ของอนุกรม Y_t สำหรับเดือนที่ $t = n+l, l = 1,2,\ldots$ คือ $\hat{Y}_n(l) = e^{\hat{W}_n(l)} \times \hat{Y}_n(l-12)$ โดยที่ $\hat{Y}_n(l-12) = y_{n+l-12}$ สำหรับ $l \le 12$

จากการประมาณค่าพารามิเตอร์ต่าง ๆ ได้ดังนี้

$\hat{\mu}$ = 0.06973760

$\hat{\varphi}$ = 0.66473453

$\hat{\theta}$ = -0.32955975

$\hat{\Phi}_{12}$ = -0.75242215

$\hat{\Phi}_{24}$ = -0.40719619

3.15 ตรวจสอบความเหมาะสมของตัวแบบ โดยตรวจสอบจากค่าสถิติ Box-Ljung ของส่วนตกค้าง พบว่า r_k และ r_{kk} ตกอยู่ในขอบเขตความเชื่อมั่นร้อยละ

95 ทุกค่า lag ดังรูปที่ 5 แสดงว่าตัวแบบมีความเหมาะสมแล้ว

Autocorrelations: ERR_1 Error for TOURIST from ARIMA, MOD_12 LN

Lag	Auto-Corr.	Stand. Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
1	-.033	.075										.199	.656
2	.011	.074										.219	.896
3	.111	.074										2.464	.482
4	-.152	.074										6.702	.153
5	.094	.074										8.322	.139
6	-.028	.073										8.471	.206
7	-.049	.073										8.924	.258
8	.006	.073										8.931	.348
9	.019	.073										8.999	.437
10	.157	.073										13.648	.190
11	.163	.072										18.733	.066
12	-.072	.072										19.739	.072
13	-.003	.072										19.741	.102
14	-.054	.072										20.309	.121
15	-.038	.072										20.589	.150
16	-.063	.071										21.363	.165
17	-.015	.071										21.406	.209
18	.122	.071										24.360	.144
19	-.065	.071										25.215	.154
20	.164	.070										30.613	.061
21	.082	.070										31.985	.059
22	-.015	.070										32.034	.077
23	-.006	.070										32.042	.099
24	-.122	.069										35.118	.067
25	-.063	.069										35.946	.072
26	-.103	.069										38.170	.058
27	-.028	.069										38.331	.073
28	-.035	.069										38.587	.088
29	-.011	.068										38.615	.109
30	-.009	.068										38.632	.134
31	-.045	.068										39.073	.151
32	-.031	.068										39.280	.176
33	-.034	.067										39.527	.201
34	-.041	.067										39.897	.224
35	-.072	.067										41.068	.222
36	-.129	.067										44.824	.149

รูปที่ 5 ค่าสถิติ Box-Ljung ของส่วนตกค้าง

3.2 วิธีพยากรณ์ของวินเตอร์

วิธีการพยากรณ์จะเลือกค่า α, γ และ δ อยู่ระหว่าง 0 – 1 โดย α เพิ่มค่าขึ้นครั้งละ 0.1 ส่วน γ และ φ เพิ่มค่าขึ้นครั้งละ 0.2 ผลการวิเคราะห์ค่า α และ γ และ δ ที่ให้ค่า MSE ต่ำสุด คือ α = 0.9 γ = 0 และ δ = 0 ดังนั้นตัวแบบพยากรณ์ค่าจริง Y_{t+l} แบบจุดที่เวลา t+l จากเวลาปัจจุบัน t (t = 1,2,...,l) เป็นดังนี้

$$\hat{Y}_t(l) = (\hat{\mu}_t + l\hat{\beta}_t)\hat{s}_{t+l-m}, \; l = 1,2,...,m$$

$$\hat{Y}_t(l) = (\hat{\mu}_t + l\hat{\beta}_t)\hat{s}_{t+l-km}, \; l = (k-1)m+1,(k-1)m+2,...,km; \; k = 2,3,...$$

โดยที่

$$\hat{\mu}_t = 0.9(Y_t / \hat{s}_{t-m}) + 0.1(\hat{\mu}_{t-1} + \hat{\beta}_{t-1})$$

$$\hat{\beta}_t = \hat{\beta}_{t-l}$$

$$\hat{s}_t = \hat{s}_{t-m}$$

กำหนด $\hat{\mu}_0$ = 566.03771 $\hat{\beta}_0$ = 5.95691 และค่าเริ่มต้นของดัชนีฤดูกาล คือ

$\bar{s}_1$ = 114.22392 $\bar{s}_2$ = 109.26368 $\bar{s}_3$ = 106.58887

$\bar{s}_4$ = 93.78781 $\bar{s}_5$ = 83.39818 $\bar{s}_6$ = 86.86538

$\bar{s}_7$ = 99.42517 $\bar{s}_8$ = 101.38496 $\bar{s}_9$ = 86.88070

$\bar{s}_{10}$ = 93.30835 $\bar{s}_{11}$ = 105.63040 $\bar{s}_{12}$ = 119.24257

ผลการพยากรณ์เมื่อตรวจสอบจากกราฟ SACF ของส่วนตกค้าง พบว่า r_k และ r_{kk} บางค่า lag ไม่ตกอยู่ในขอบเขตความเชื่อมั่นร้อยละ 95 ดังรูปที่ 6 แสดงว่าตัวแบบจากวิธีการของวินเตอร์ยังไม่ค่อยมีความเหมาะสม

Autocorrelations: ERR_1 Error for TOURIST from EXSMOOTH, MOD_2 L

Lag	Auto-Corr.	Stand. Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
1	.044	.072										.374	.541
2	-.121	.072										3.209	.201
3	-.010	.072										3.227	.358
4	-.197	.072										10.782	.029
5	-.011	.071										10.807	.055
6	-.090	.071										12.414	.053
7	-.110	.071										14.806	.039
8	-.009	.071										14.823	.063
9	-.065	.071										15.657	.074
10	.102	.070										17.760	.059
11	.191	.070										25.185	.009
12	-.058	.070										25.860	.011
13	-.065	.070										26.728	.014
14	-.050	.070										27.250	.018
15	-.045	.069										27.670	.024
16	-.057	.069										28.347	.029
17	.026	.069										28.485	.040
18	.182	.069										35.463	.008
19	.019	.069										35.537	.012
20	.097	.068										37.551	.010
21	.050	.068										38.087	.013
22	-.046	.068										38.540	.016
23	-.040	.068										38.880	.020
24	-.063	.068										39.748	.023
25	-.032	.067										39.979	.029
26	-.025	.067										40.120	.038
27	-.013	.067										40.158	.050
28	.050	.067										40.724	.057
29	-.002	.067										40.725	.073
30	-.011	.066										40.751	.091
31	.005	.066										40.757	.113
32	-.040	.066										41.127	.129
33	.023	.066										41.244	.154
34	-.043	.066										41.674	.172
35	.036	.065										41.978	.194
36	.086	.065										43.703	.177

รูปที่ 6 กราฟ SACF ของส่วนตกค้าง

3.3 การพยากรณ์

การวิจัยครั้งนี้ได้จำแนกการพยากรณ์จำนวนนักท่องเที่ยวต่างชาติออกเป็น 3 ช่วงเวลา คือ ช่วงเวลาที่ 1 ตั้งแต่เดือนมกราคม 2541 ถึงเดือนกันยายน 2555 (177 คาบเวลา) ช่วงเวลาที่ 2 ตั้งแต่เดือนตุลาคม 2555 ถึงเดือนธันวาคม 2555 (3 คาบเวลาล่วงหน้า) ช่วงเวลาที่ 3 ตั้งแต่เดือนมกราคม 2541 ถึงเดือนธันวาคม 2555 (180 คาบเวลา)

จากตารางที่ 1 พบว่าวิธีการพยากรณ์ของวินเตอร์ให้ค่า RMSE และ MAPE ต่ำกว่าวิธีของบอกซ์-เจนกินส์ในทุกช่วงเวลา แต่จากการตรวจสอบส่วนตกค้างเพื่อพิจารณาความเหมาะสมของตัวแบบ พบว่าตัวแบบ ARIMA$(1,0,1)(2,1,0)_{12}\ln Y_t$ จากวิธีของบอกซ์-เจนกินส์มีความเหมาะสม

ตารางที่ 1 ค่า RMSE และ MAPE จากตัวแบบ ARIMA$(1,0,1)(2,1,0)_{12}\ln Y_t$ และตัวแบบของวินเตอร์

	ช่วงเวลาที่ 1		ช่วงเวลาที่ 2		ช่วงเวลาที่ 3	
	ARIMA	Winter	ARIMA	Winter	ARIMA	Winter
RMSE	89.886	82.313	311.4303	183.2488	97.782	84.983
MAPE	6.407	5.536	13.548	5.746	6.561	5.537

4. สรุปและข้อเสนอแนะ

ในการเปรียบเทียบค่าพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวในประเทศไทยทั้ง 3 ช่วงเวลา คือ ช่วงเวลาที่ 1 ตั้งแต่เดือนมกราคม 2541 ถึงเดือนกันยายน 2555 (177 คาบเวลา) ช่วงเวลาที่ 2 ตั้งแต่เดือนตุลาคม 2555 ถึงเดือนธันวาคม 2555 (3 คาบเวลาล่วงหน้า) ช่วงเวลาที่ 3 ตั้งแต่เดือนมกราคม 2541 ถึงเดือนธันวาคม 2555 (เป็นการรวมช่วงเวลาที่ 1 และ 2 เข้าด้วยกัน เป็น 180 คาบเวลา) จะพบว่าค่าพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวในประเทศไทยจากวิธีการพยากรณ์ของวินเตอร์ให้ค่า RMSE และ MAPE ต่ำกว่าวิธีของบอกซ์-เจนกินส์ในทุกช่วงเวลา แต่เมื่อตรวจสอบส่วนตกค้างพบว่าตัวแบบจากวิธีการพยากรณ์ของวินเตอร์ไม่เหมาะสม ในขณะที่ตัวแบบ ARIMA(1,0,1)(2,1,0)12lnYt จากวิธีของบอกซ์-เจนกินส์มีความเหมาะสม ดังนั้นผู้สนใจศึกษาอาจพิจารณาวิธีพยากรณ์วิธีอื่นต่อไป

5. กิตติกรรมประกาศ

งานวิจัยครั้งนี้ได้รับคำปรึกษาและข้อเสนอแนะเป็นอย่างดียิ่งจาก รองศาสตราจารย์มานพ วราภักดิ์ คณะผู้วิจัยจึงขอกราบขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้

6. เอกสารอ้างอิง

[1] สำนักงานสถิติแห่งชาติ, เที่ยวเมืองไทย ใครใครก็อยากมา, แหล่งที่มา: http://service.nso.go.th/nso/nsopublish/citizen/news/news_travel.jsp, 9 มิถุนายน 2556.

[2] คมชัดลึก, ปี 56 ยอดนักท่องเที่ยงต่างชาติทะลุ 24 ล้านคน, แหล่งที่มา: http://www.Komchadluek.net/detail/20130413/156003/ปี56ยอดนทท.ต่างชาติทะลุ24ล้านคน.html#.UdZyZtQ5Prc, 7 มิถุนายน 2556.

[3] ไทยรัฐ, กทม. คว้าอันดับ 1 'สุดยอดเมืองท่องเที่ยวโลก', แหล่งที่มา: http://www.thairath.co.th/content/region/347702, 9 มิถุนายน 2556.

[4] อัครพงศ์ อั้นทอง และปวีณา คำพุกกะ, 2552, การพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวในประเทศไทยโดยใช้แบบจำลอง Sarima intervention, ว.วิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยอุบลราชธานี 11: 196-218.

[5] ไอริณ ถาวรนันท์, การพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่มาท่องเที่ยวในประเทศไทย, โครงงานระดับปริญญาตรี, ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล, กรุงเทพฯ.

[6] Chaitip, P., Chaiboonsri, C. and Mukhjang, R., 2008, Time series models for forecasting international visitor arrivals to Thailand, International Conference on Applied Economics – ICOAE 2008: 159-163.

[7] เมตตา โกสินานนท์, 2550, ระบบการพยากรณ์จำนวนนักท่องเที่ยวโดยใช้โครงข่ายใยประสาทเทียม : กรณีศึกษาการท่องเที่ยวแห่งประเทศไทย, สารนิพนธ์, สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, กรุงเทพฯ.

[8] ธนาคารแห่งประเทศไทย, ข้อมูลจำนวนนักท่องเที่ยว, แหล่งที่มา : http://www2.bot.or.th/statistics/ReportPage.aspx?reportID=624&language=th, 9 มิถุนายน 2556.

[9] Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., 1994, Time Series Analysis: Forecasting and Control, 3rd Ed., Prentice-Hall, Englewood Cliffs, New Jersey.

[10] มานพ วราภักดิ์, 2552, การวิจัยดำเนินการ, สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.