วารสารวิชาการและวิจัย มทร.พระนคร ฉบับพิเศษ
การประชุมวิชาการมหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 5

การเปรียบเทียบวิธีการพยากรณ์ดัชนีราคาเซ็ท 100 ในเชิงสถิติ

Comparison of Statistical Techniques for Forecasting SET 100 Index

ยุพาพิน อติกานต์กุล[1*]

[1]อาจารย์ สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร กรุงเทพฯ 10300

บทคัดย่อ

การวิจัยครั้งนี้เป็นการเปรียบเทียบวิธีการพยากรณ์ดัชนีราคาเซ็ท 100 ด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย วิธีการปรับให้เรียบของโฮลท์ และวิธีบ็อกซ์และเจนกินส์ โดยรวบรวมข้อมูล ตั้งแต่วันที่ 2 มกราคม พ.ศ. 2551 ถึงวันที่ 30 กันยายน พ.ศ. 2554 ในการสร้างตัวแบบ และใช้ข้อมูลตั้งแต่วันที่ 3 ตุลาคม พ.ศ. 2554 ถึงวันที่ 14 ตุลาคม พ.ศ. 2554 ในการทดสอบตัวแบบ พิจารณาวิธีการพยากรณ์ที่เหมาะสมจาก ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (MAE) ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซนต์ของคลาดเคลื่อน (MAPE) ที่มีค่าน้อยที่สุด และความเหมาะสมของตัวแบบ ผลการวิจัยพบว่า วิธีการปรับให้เรียบของโฮลท์ ให้ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (MAE) ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซนต์ของคลาดเคลื่อน (MAPE) น้อยที่สุด รองลงมาได้แก่ วิธีบ็อกซ์และเจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย ตามลำดับ แต่วิธีบ็อกซ์และเจนกินส์เป็นวิธีการพยากรณ์เพียงวิธีเดียวที่ให้ตัวแบบซึ่งความคลาดเคลื่อนไม่มีสหสัมพันธ์ในตัวเอง ดังนั้นตัวแบบที่เหมาะสมสำหรับการพยากรณ์ดัชนีราคาเซ็ท 100 คือ ARIMA (2,1,3)

Abstract

This research is the comparative study of three forecasting methods on SET 100 index. These methods are Simple Exponential Smoothing method, Holt's method and Box-Jenkins method. Two sets of data were used in this study. The first one used as a training set was collected from January 2, 2008 to September 30, 2011. The other one used as a testing set was collected from October 3, 2011 to October 14, 2011. We propose that the most appropriate forecasting method is the one which give the lowest MAE, MSE, MAPE and diagnostic checking. The results show that Holt's method displays the lowest MAE, MSE, and MAPE. The next one is Box-Jenkins method while the Simple Exponential Smoothing method exhibits the highest value of all three indicators. However, the model from Box-Jenkins method is the only one way that reveals no residual autocorrelation. This leads to the conclusion that the appropriate models for forecasting of SET 100 Index is ARIMA (2,1,3).

คำสำคัญ : การพยากรณ์ ดัชนีราคาเซ็ท 100

Keywords : FORECASTING, SET100

[*]*ผู้นิพนธ์ประสานงานไปรษณีย์อิเล็กทรอนิกส์ yupapin.a@rmutp.ac.th โทร. 0 2913 2424 ต่อ 197*

1. บทนำ

การลงทุนในตลาดหลักทรัพย์มีส่วนสำคัญในการพัฒนาเศรษฐกิจชาติ เนื่องจากเป็นแหล่งในการระดมเงินทุนสำหรับธุรกิจที่ต้องการเงินทุนในการลงทุน นอกจากนี้แล้วประชาชนทั่วไปสามารถมีส่วนร่วมในการเป็นเจ้าของกิจการ ซึ่งหากบริษัทมีกำไร นักลงทุนมีสิทธิ์ในการได้รับเงินปันผล หรืออาจได้รับกำไรจากผลต่างระหว่างราคาซื้อและราคาขาย แต่ความเสี่ยงจากการลงทุนในตลาดหลักทรัพย์มีความเสี่ยงสูง นักลงทุนจึงต้องศึกษาหาวิธีการลงทุนที่เหมาะสมเพื่อลดความเสี่ยง

จากการสำรวจงานวิจัยที่เกี่ยวข้อง บุญหญิง สมร่าง (2552 : 19-32) ได้ศึกษาพบว่าวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเหมาะสำหรับพยากรณ์กลุ่มอสังหาริมทรัพย์ ส่วนวิธีการวิเคราะห์ความถดถอยเหมาะสำหรับพยากรณ์กลุ่มธนาคาร แต่ศราวุฒิ จินตนาสุนทรศิริ (2548 : บทคัดย่อ) ได้ศึกษาพบว่าวิธีของบ็อกซ์และเจนกินส์เป็นวิธีการพยากรณ์ที่เหมาะสำหรับการพยากรณ์ราคาปิดรายวันของรายหลักทรัพย์

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการพยากรณ์ 3 วิธี ได้แก่ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย วิธีการปรับให้เรียบของโฮลท์ และวิธีบ็อกซ์และเจนกินส์ ในการพยากรณ์ดัชนีราคาเซ็ท 100 ที่คัดเลือกจากหุ้นสามัญ 100 ตัวแรกที่มีการซื้อขายสูงอย่างสม่ำเสมอในรอบ 6 เดือน และจดทะเบียนในตลาดหลักทรัพย์แห่งประเทศไทย (ตลาดหลักทรัพย์แห่งประเทศไทย, 2555) ซึ่งวิธีการพยากรณ์ที่เหมาะสมนอกจากจะสามารถพยากรณ์ดัชนีราคาเซ็ท 100 แล้ว ยังอาจใช้ในการพยากรณ์ราคารายหลักทรัพย์ส่วนใหญ่ได้อีกด้วย

2. วิธีการทดลอง

2.1 การรวบรวมข้อมูล ใช้ข้อมูลทุติยภูมิของดัชนีราคาเซ็ท 100 เป็นระยะเวลา 923 วัน ตั้งแต่วันที่ 2 มกราคม พ.ศ. 2551 ถึงวันที่ 14 ตุลาคม พ.ศ. 2554 โดยใช้ข้อมูลตั้งแต่วันที่ 2 มกราคม พ.ศ. 2551 ถึงวันที่ 30 กันยายน พ.ศ. 2554 ในการสร้างตัวแบบพยากรณ์ และใช้ข้อมูลตั้งแต่วันที่ 3 ตุลาคม พ.ศ. 2554 ถึงวันที่ 14 ตุลาคม พ.ศ. 2554 ในการทดสอบตัวแบบพยากรณ์

2.2 การวิเคราะห์ข้อมูล โดยใช้โปรแกรมสำเร็จรูป SPSS for Windows

2.2.1 การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย (Simple Exponential Smoothing)

เป็นวิธีการพยากรณ์สำหรับข้อมูลที่ไม่มีแนวโน้ม และไม่มีความผันแปรเนื่องมาจากฤดูกาล ซึ่งจะให้ความสำคัญมากกับข้อมูลที่ใกล้กับเวลาปัจจุบัน และให้ความสำคัญลดลงกับข้อมูลในอดีตที่ห่างออกไปจากปัจจุบัน

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

โดยที่ $\hat{Y}_t(1)$ แทนค่าพยากรณ์ 1 หน่วยเวลาล่วงหน้า เมื่อพยากรณ์ ณ เวลา t

α แทนค่าคงที่ในการปรับให้เรียบ ซึ่ง $0 < \alpha < 1$

Y_t แทนค่าสังเกต ณ เวลา t

เลือก α ที่ให้ค่าผลรวมกำลังสองของความคลาดเคลื่อน (SSE) ที่มีค่าน้อยที่สุด

2.2.2 การปรับให้เรียบของโฮลท์ (Holt's Method)

วิธีนี้เหมาะสำหรับการพยากรณ์ข้อมูลที่มีแนวโน้มเชิงเส้น ซึ่งมีค่าคงที่ในการปรับให้เรียบ 2 ตัวคือ α และ γ

$$\hat{Y}_t(l) = S_t + b_t l$$

โดยที่ $S_t = \alpha Y_t + (1-\alpha)(S_{t-1} + b_{t-1})$

$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1}$

$\hat{Y}_t(l)$ แทนค่าพยากรณ์ l หน่วยเวลาล่วงหน้า เมื่อพยากรณ์ ณ เวลา t

S_t แทนระดับของข้อมูล

b_t แทนความชันของข้อมูล

เลือก α และ γ ที่ให้ค่าผลรวมกำลังสองของความคลาดเคลื่อน (SSE) ที่มีค่าน้อยที่สุด

2.2.3 บ็อกซ์และเจนกินส์ (Box-Jenkins Method)

เป็นการวิเคราะห์ข้อมูลอนุกรมเวลาที่อนุกรมเวลาต้องมีคุณสมบัติสเตชันนารี หากอนุกรมเวลาไม่มีคุณสมบัติสเตชันนารี ต้องทำการแปลงข้อมูลเสียก่อน และจะกำหนดตัวแบบโดยพิจารณาจากสหสัมพันธ์ในตัวเอง (Autocorrelation) และสหสัมพันธ์ในตัวเองบางส่วน (Partial autocorrelation) สำหรับตัวแบบอนุกรมเวลา ARIMA (p,d,q) คือ

$$Y_t = Y_{t-1} + \phi_1(Y_{t-1}-Y_{t-2}) + \phi_2(Y_{t-2}-Y_{t-3}) + \ldots + \phi_p(Y_{t-p}-Y_{t-p-1}) + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \ldots - \theta_q\varepsilon_{t-q}$$

โดยที่ Y_t แทนค่าสังเกต ณ เวลา t

$\phi_1, \ldots, \phi_p$ แทนพารามิเตอร์ของกระบวนการ AR

$\theta_1, \ldots, \theta_q$ แทนพารามิเตอร์ของกระบวนการ MA

ε_t แทนตัวแปรสุ่มที่เป็นอิสระกัน มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนคงที่ซึ่งเท่ากับ σ_ε^2

เมื่อคัดเลือกตัวแบบจากสหสัมพันธ์ในตัวเอง (Autocorrelation) และสหสัมพันธ์ในตัวเองบางส่วน (Partial autocorrelation) แล้ว หลังจากนั้นจะประมาณค่าพารามิเตอร์ โดยใช้สถิติที

2.3 ทดสอบว่าความคลาดเคลื่อนมีสหสัมพันธ์ในตัวเองหรือไม่ โดยใช้สถิติ Q หรือ เรียกว่าสถิติ Box-Ljung

$$Q = n(n+2)\sum_{j=1}^{k}\frac{r_j^2(\hat{\varepsilon}_t)}{n-j}$$

โดยที่ n แทนขนาดของอนุกรมเวลาของความคลาดเคลื่อน

k แทนช่วงเวลาห่างกันสูงสุดของความคลาดเคลื่อนที่พิจารณา

$\hat{\varepsilon}_t$ แทนความคลาดเคลื่อนของการพยากรณ์ ณ เวลา t

$r_j(\hat{\varepsilon}_t)$ แทนสหสัมพันธ์ในตัวเองของตัวอย่างของ $\hat{\varepsilon}_t$

ถ้า Q มีค่าน้อยกว่า χ^2 ที่ระดับนัยสำคัญ α องศาเสรีเท่ากับ k ลบด้วยจำนวนพารามิเตอร์ที่ประมาณค่าในตัวแบบ แสดงว่าความคลาดเคลื่อนไม่มีสหสัมพันธ์ในตัวเอง ตัวแบบที่ได้เป็นตัวแบบที่เหมาะสม

2.4 คำนวณหาค่าความคลาดเคลื่อน จากผลต่างระหว่างค่าจริงและค่าพยากรณ์จากตัวแบบที่ได้ โดยใช้ข้อมูลตั้งแต่วันที่ 3 ตุลาคม พ.ศ. 2554 ถึงวันที่ 14 ตุลาคม พ.ศ. 2554

$$e_t = y_t - \hat{y}_t = \hat{\varepsilon}_t$$

โดยที่ e_t แทนความคลาดเคลื่อนของการพยากรณ์ ณ เวลา t

2.5 เปรียบเทียบวิธีการพยากรณ์ โดยพิจารณาจาก ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (MAE) ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซนต์ของคลาดเคลื่อน (MAPE) ที่มีค่าน้อยที่สุด

2.5.1 ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (Mean Absolute Error)

$$MAE = \frac{\sum|e_t|}{n}$$

2.5.2 ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (Mean Square Error)

$$MSE = \frac{\sum e_t^2}{n}$$

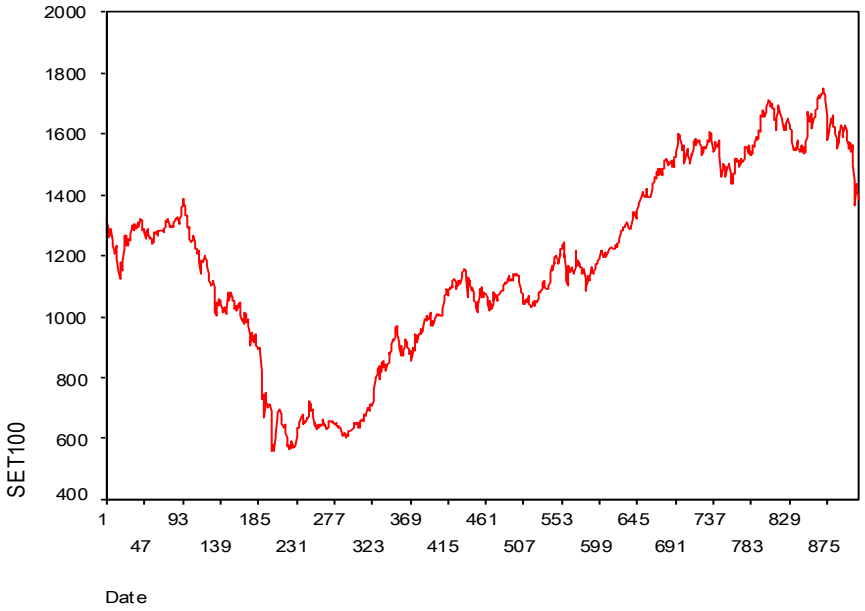
2.5.3 ค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซนต์ของคลาดเคลื่อน (Mean Absolute Percent Error)

$$MAPE = \frac{\sum\left|\frac{e_t}{y_t}\right| \times 100}{n}$$

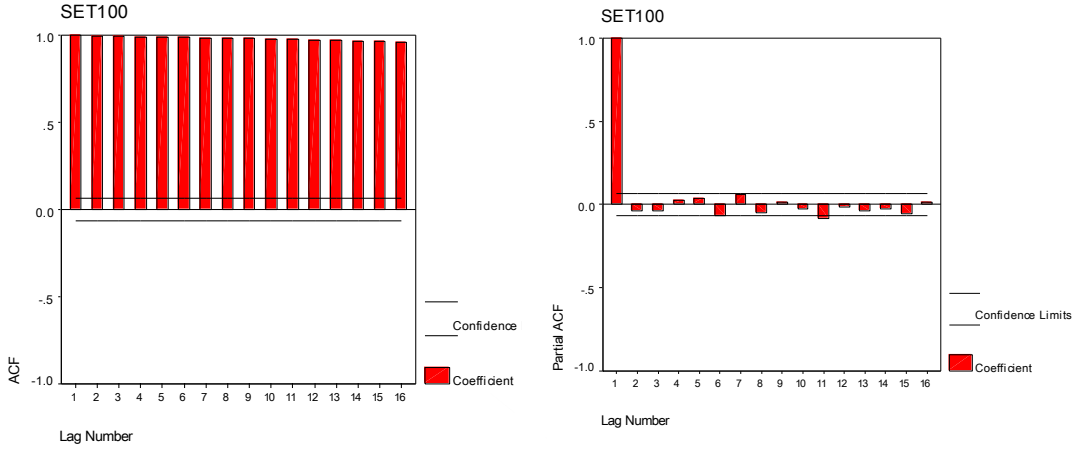
3. ผลการทดลองและวิจารณ์ผล

3.1 ผลการทดลอง

จากการรวบรวมข้อมูลดัชนีราคาเซ็ท 100 ตั้งแต่วันที่ 2 มกราคม พ.ศ. 2551 ถึงวันที่ 30 กันยายน พ.ศ. 2554 สามารถวาดกราฟได้ดังนี้



รูปที่ 1 ดัชนีราคาเซ็ท 100 ตั้งแต่วันที่ 2 มกราคม พ.ศ. 2551 ถึงวันที่ 30 กันยายน พ.ศ. 2554



รูปที่ 2 กราฟฟังก์ชันสหสัมพันธ์ในตัวเอง (ACF) และฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (PACF) ของดัชนีราคาเซ็ท 100

จากรูปที่ 2 พิจารณาจากกราฟฟังก์ชันสหสัมพันธ์ในตัวเอง (ACF) และฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (PACF) ที่ได้ ตัวแบบที่คาดว่าน่าจะเหมาะสมคือ ARIMA (2,1,3)

ตารางที่ 1 การประมาณค่าพารามิเตอร์ของตัวแบบ ARIMA (2,1,3)

	B	SEB	T-RATIO	APPROX.PROB
AR1	-1.5331095	0.10518185	-14.575798	0.00000000
AR2	-0.8127424	0.10303459	-7.888054	0.00000000
MA1	-1.5845535	0.10857880	-14.593581	0.00000000
MA2	-0.9578925	0.10563553	-9.067901	0.00000000
MA3	-0.1103242	0.03544026	-3.112962	0.00191027

จากตารางที่ 1 สรุปได้ว่าพารามิเตอร์ทุกตัวมีค่าแตกต่างจาก 0 อย่างมีนัยสำคัญทางสถิติ ที่ระดับ 0.05

ตารางที่ 2 การเปรียบเทียบสมการพยากรณ์ดัชนีราคาเซ็ท 100

วิธีการพยากรณ์	MAE	MSE	MAPE
การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย $\alpha = 1$	33.87	1776.61	2.46
การปรับให้เรียบของโฮลท์ $\alpha = 1, \gamma = 0$	28.38	1248.72	2.04
บ็อกซ์และเจนกินส์ ARIMA (2,1,3)	29.57	1524.78	2.15

จากตารางที่ 2 พบว่าวิธีการปรับให้เรียบของโฮลท์ ให้ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (MAE) ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซนต์ของคลาดเคลื่อน (MAPE) น้อยที่สุด รองลงมาได้แก่วิธีบ็อกซ์และเจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย ตามลำดับ

แต่เมื่อพิจารณาจากความคลาดเคลื่อนไม่มีสหสัมพันธ์ในตัวเองด้วยสถิติ Box-Ljung พบว่าวิธีบ็อกซ์และเจนกินส์ เป็นวิธีการพยากรณ์เพียงวิธีเดียวที่เหมาะสม

3.2 การอภิปรายผล

การเปรียบเทียบวิธีการพยากรณ์ดัชนีราคาเซ็ท 100 พบว่าวิธีบ็อกซ์และเจนกินส์เป็นวิธีการที่เหมาะสมสำหรับการพยากรณ์ โดยพิจารณาจากค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (MAE) ค่าเฉลี่ยกำลังสองของ

ความคลาดเคลื่อน (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซนต์ของคลาดเคลื่อน (MAPE) ที่มีค่าค่อนข้างน้อย และจากการตรวจสอบความคลาดเคลื่อนไม่มีสหสัมพันธ์ในตัวเองด้วยสถิติ Box-Ljung ซึ่งขัดแย้งกับบุญหญิง สมร่าง (2552 : 19-32) ที่ได้ศึกษาพบว่าวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเหมาะสำหรับพยากรณ์กลุ่มอสังหาริมทรัพย์ ส่วนวิธีการวิเคราะห์ความถดถอยเหมาะสำหรับพยากรณ์กลุ่มธนาคาร แต่เป็นไปในแนวทางเดียวกับ ศราวุฒิ จินตนาสุนทรศิริ (2548 : บทคัดย่อ) ที่พบว่าวิธีของบ็อกซ์และเจนกินส์เป็นวิธีการพยากรณ์ที่เหมาะสำหรับการพยากรณ์ราคาปิดรายวันของรายหลักทรัพย์ ซึ่งในการศึกษาของบุญหญิง สมร่าง และศราวุฒิ จินตนาสุนทรศิริ เป็นการศึกษารายหลักทรัพย์ แต่ในการศึกษาครั้งนี้เป็นการศึกษาดัชนีราคาเซ็ท 100

4. สรุป

การเปรียบเทียบวิธีการพยากรณ์ดัชนีราคาเซ็ท 100 โดยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย วิธีการปรับให้เรียบของโฮลท์ และวิธีบ็อกซ์และเจนกินส์ ซึ่งรวบรวมข้อมูล ตั้งแต่วันที่ 2 มกราคม พ.ศ. 2551 ถึงวันที่ 30 กันยายน พ.ศ. 2554 ในการสร้างตัวแบบ พิจารณาวิธีการพยากรณ์จากค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (MAE) ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซนต์ของคลาดเคลื่อน (MAPE) ที่มีค่าน้อยที่สุด พบว่าวิธีการปรับให้เรียบของโฮลท์ให้ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (MAE) ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซนต์ของคลาดเคลื่อน (MAPE) น้อยที่สุด รองลงมาได้แก่ วิธีบ็อกซ์และเจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย ตามลำดับ แต่เมื่อพิจารณาประกอบกับความคลาดเคลื่อนไม่มีสหสัมพันธ์ในตัวเองด้วยสถิติ Box-Ljung พบว่าวิธีบ็อกซ์และเจนกินส์เป็นวิธีที่เหมาะสมสำหรับการพยากรณ์ ซึ่งตัวแบบที่เหมาะสมคือ ARIMA (2,1,3)

5. กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณมหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร ที่ให้การสนับสนุนการวิจัยในครั้งนี้

6. เอกสารอ้างอิง

ตลาดหลักทรัพย์แห่งประเทศไทย. 2555. **หลักเกณฑ์การคัดเลือกหลักทรัพย์ใน SET 50 Index และ 100 Index** [Online]. Available from : http://www.set.or.th/th/products/index/files/201207_SET50100_Selection_Criteria_TH.pdf

บุญหญิง สมร่าง. 2552. **การเปรียบเทียบวิธีพยากรณ์ราคาหลักทรัพย์เชิงสถิติ**. วารสารวิทยาศาสตร์ประยุกต์. ปีที่ 5 (ฉบับที่ 2) : 19-32.

มุกดา แม้นมินทร์. 2549. **อนุกรมเวลาและการพยากรณ์**. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร. สำนักพิมพ์ประกายพรึก.

ศราวุฒิ จินตนาสุนทรศิริ. 2548. **การใช้วิธีทางสถิติพยากรณ์ราคาปิดรายวันของหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย**.วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์.