



ใบรับรองวิทยานิพนธ์
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตร์มหาบัณฑิต (สถิติ)

ปริญญา

สถิติ สาขา สถิติ ภาควิชา

เรื่อง การพยากรณ์ปริมาณน้ำฝนรายเดือนของจังหวัดลำปางโดยใช้เทคนิคการพยากรณ์เชิงสถิติ

Forecasting Monthly Rainfall of Lampang Province by Statistical Forecasting Techniques

นามผู้วิจัย นายปรีชา เรืองชัยศิวเวท

ได้พิจารณาเห็นชอบโดย

ประธานกรรมการ

(อาจารย์อำไพ ทองธีรภาพ, Ph.D.)

กรรมการ

(ผู้ช่วยศาสตราจารย์บุญอ้อม โฉมที, Ph.D.)

กรรมการ

(ผู้ช่วยศาสตราจารย์ทิพย์รัตน์ เล่าหิweiseียร, Ph.D.)

หัวหน้าภาควิชา

(อาจารย์อำไพ ทองธีรภาพ, Ph.D.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญญา วีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ เดือน พ.ศ.

สืบศิริ มหาวิทยาลัยเกษตรศาสตร์

วิทยานิพนธ์

เรื่อง

การพยากรณ์ปริมาณน้ำฝนรายเดือนของจังหวัดลำปางโดยใช้เทคนิคการพยากรณ์เชิงสถิติ

Forecasting Monthly Rainfall of Lampang Province by Statistical Forecasting Techniques

โดย

นายปรีชา เรืองชัยคิวเวท

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ)

พ.ศ. 2554

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ปริษา เรื่องชัชวาท 2554: การพยากรณ์ปริมาณน้ำฝนรายเดือนของจังหวัดลำปางโดยใช้เทคนิคการพยากรณ์เชิงสถิติ ปรินญาวิทยาศาสตร์มหาบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ ปรธานกรรมการที่ปรึกษา: อาจารย์อำไพ ทองธีรภาพ, Ph.D. 99 หน้า

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการพยากรณ์ 4 วิธี ได้แก่ วิธีแยกส่วนประกอบ วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์ วิธีบ็อกซ์และเจนกินส์ และวิธีการพยากรณ์ร่วม ในการพยากรณ์ปริมาณน้ำฝนรายเดือนของจังหวัดลำปาง เพื่อหารูปแบบที่เหมาะสมที่สุด โดยพิจารณาจากวิธีที่ให้ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (Mean Absolute Deviation : MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (Mean Absolute Percent Error : MAPE) ที่ต่ำที่สุด

ข้อมูลที่ใช้ในงานวิจัยนี้ เป็นข้อมูลน้ำฝนรายเดือนของจังหวัดลำปาง ซึ่งรวบรวมจากสำนักพัฒนาอุตุนิยมหาวิทยาลัย กรมอุตุนิยวิทยา กระทรวงเทคโนโลยีสารสนเทศ และการสื่อสาร โดยแบ่งข้อมูลออกเป็น 2 ชุด ชุดแรก ตั้งแต่เดือนมกราคม 2542 ถึงเดือนพฤษภาคม 2551 เพื่อพยากรณ์เปรียบเทียบกับข้อมูลในชุดที่ 2 ตั้งแต่เดือนมิถุนายน 2551 ถึงเดือนพฤษภาคม 2552

ผลการศึกษาพบว่า วิธีวินเตอร์จะเหมาะสมในการพยากรณ์ตัวแบบปริมาณน้ำฝนมากที่สุด เนื่องจากให้ค่า MAD, MSE และ MAPE ต่ำกว่าวิธีการพยากรณ์อื่นๆ สรุปสมการการพยากรณ์ของแต่ละวิธีได้ ดังนี้

วิธีแยกส่วนประกอบ สมการคือ $Y_t = (132.432 + 0.364579t)\hat{S}_t$

วิธีวินเตอร์ สมการคือ $\hat{Y}_{t+m} = (a_t + b_t(m))\hat{S}_{t-12+m}$,

$a_t = 0.001(Y_t / \hat{S}_{t-12}) + 0.999[a_{t-1} + b_{t-1}]$ และ $b_t = 0.01(a_t - a_{t-1}) + 0.99b_{t-1}$

วิธีบ็อกซ์และเจนกินส์ สมการคือ

$Y_t = Y_{t-12} - 0.5347Y_t + 0.5347Y_{t-24} - 0.2853Y_{t-24} + 0.2853Y_{t-36} + \varepsilon_t$

วิธีพยากรณ์ร่วม สมการคือ $\hat{Y}_{CF} = 0.057\hat{Y}^{(1)} + 0.54\hat{Y}^{(2)} + 0.403\hat{Y}^{(3)}$

ลายมือชื่อนิติ

ลายมือชื่อปรธานกรรมการ

Preecha Ruangchaisiwawet 2011: Forecasting Monthly Rainfall of Lampang Province by Statistical Forecasting Techniques. Master of Science (Statistics), Major Field: Statistics, Department of Statistics. Thesis Advisor: Mrs. Ampai Thongteeraparp, Ph.D. 99 pages.

The purpose of this research was the comparison of four methods namely: Decomposition Method, Triple exponential smoothing : Winter's method, Box – Jenkins Method and one Combination forecasting method to fit with forecasting of monthly rainfall of Lampang Province. Data which gives the lowest Mean Absolute Deviation, Mean Square Error and Mean Absolute Percent Error represents the most suitable method.

Data used in this research was monthly rainfall series of Lampang collected from Thai Meteorological Department Ministry of Information Technology and communication. Data was divided into two groups. first group consisted of data collected from January 1999 to May 2008, while data for June 2008 to May 2009 was kept to investigate the accuracy of forecast methods.

The results of the study of forecasting models shown that triple exponential smoothing : Winter's method was suitable for the four types of rainfall data when consider the lowest Mean Absolute Deviation, Mean Square Error and Mean Absolute Percent Error.

The predicted model of each method were

Model for Decomposition method was $Y_t = (132.432 + 0.364579t)\hat{S}_t$

Model for Winter method was $\hat{Y}_{t+m} = (a_t + b_t(m))\hat{S}_{t-12+m}$,

$$a_t = 0.001\left(Y_t / \hat{S}_{t-12}\right) + 0.999[a_{t-1} + b_{t-1}] \text{ and } b_t = 0.01(a_t - a_{t-1}) + 0.99b_{t-1}$$

Model for Box and Jenkins method was

$$Y_t = Y_{t-12} - 0.5347Y_t + 0.5347Y_{t-24} - 0.2853Y_{t-24} + 0.2853Y_{t-36} + \varepsilon_t$$

Model for Combination forecasting method was $\hat{Y}_{CF} = 0.057\hat{Y}^{(1)} + 0.54\hat{Y}^{(2)} + 0.403\hat{Y}^{(3)}$

Student's signature

Thesis Advisor's signature

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณ อาจารย์ ดร.อำไพ ทองธีรภาพ ประธานกรรมการที่ปรึกษา
วิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.บุญอ้อม โฉมทิ กรรมการที่ปรึกษาสาขาวิชาเอก และ
ผู้ช่วยศาสตราจารย์ ดร.ทิพย์รัตน์ เลาหวิเชียร กรรมการที่ปรึกษาสาขาวิชารอง ที่ให้คำปรึกษาในการ
เรียน การค้นคว้าวิจัย ตลอดจนการตรวจแก้ไขวิทยานิพนธ์จนกระทั่งเสร็จสมบูรณ์ และกราบ
ขอบพระคุณ อ.ดร.ชนิสวรา เลิศอมรพงษ์ผู้แทนบัณฑิตวิทยาลัย ที่ได้ให้ความกรุณาตรวจแก้ไข
วิทยานิพนธ์ให้สมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณคณาจารย์ภาควิชาสถิติทุกท่าน ที่ได้อบรมสั่งสอนและมอบความรู้อัน
เป็นประโยชน์อย่างยิ่งในการนำไปใช้ประโยชน์ต่อไป และขอขอบคุณ คุณพิสิษฐ์ ศิริเลิศชนกุล
ช่างคอมพิวเตอร์ และเจ้าหน้าที่ภาควิชาสถิติทุกท่าน ที่ได้ให้ความช่วยเหลือและให้คำแนะนำต่างๆ

ด้วยความดีหรือประโยชน์อันใดเนื่องจากวิทยานิพนธ์เล่มนี้ ขอมอบแต่คุณพ่อ คุณแม่ ที่ได้
อบรมและให้กำลังใจผู้วิจัยมาตลอดในทุกเรื่อง

ปรีชา เรืองชัยศิวเวท
มีนาคม 2554

สารบัญ

	หน้า
สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(4)
คำนำ	1
วัตถุประสงค์	3
การตรวจเอกสาร	5
อุปกรณ์และวิธีการ	54
อุปกรณ์	54
วิธีการ	54
ผลและวิจารณ์	57
ผล	57
วิจารณ์	60
สรุปและข้อเสนอแนะ	92
สรุป	92
ข้อเสนอแนะ	94
เอกสารและสิ่งอ้างอิง	95
ภาคผนวก	97
ประวัติการศึกษาและการทำงาน	99

สารบัญตาราง

ตารางที่		หน้า
1	ลักษณะของ ρ_k และ ρ_{kk} สำหรับรูปแบบ ARMA(p,q)	27
2	ความสัมพันธ์ของ ρ_k กับพารามิเตอร์ในรูปแบบต่างๆ	41
3	ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของปริมาณน้ำฝนจังหวัดลำปางในแต่ละปีตั้งแต่ปี 2542-2551	59
4	ดัชนีฤดูกาลของวิธีแยกส่วนประกอบ	62
5	ค่า MAD, MSE และ MAPE ของอนุกรมเวลาสำหรับค่าที่ได้จากการพยากรณ์เปรียบเทียบกับค่าจริงด้วยวิธีแยกส่วนประกอบ	63
6	ดัชนีฤดูกาลวินเตอร์	65
7	ค่า $r_k(e_t)$ และคอเรลโรแกรมของค่า $r_k(e_t)$ ของวิธีวินเตอร์	67
8	ค่า MAD, MSE และ MAPE ของอนุกรมเวลาสำหรับค่าที่ได้จากการพยากรณ์เปรียบเทียบกับค่าจริงด้วยวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์	68
9	ค่า r_k และคอเรลโรแกรมของค่า r_k ของวิธีวิธีบ็อกซ์และเจนกินส์	70
10	ค่า $r_k(z_t)$ และคอเรลโรแกรมของค่า $r_k(z_t)$ ของวิธีวิธีบ็อกซ์และเจนกินส์	74
11	ค่า $r_{kk}(z_t)$ และคอเรลโรแกรมของค่า $r_{kk}(z_t)$ ของวิธีวิธีบ็อกซ์และเจนกินส์ค่า	75
12	AIC, ค่า SBC และ SSE ของรูปแบบ	76
13	ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(1,1,0) ₁₂	76
14	ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(0,1,1) ₁₂ ค่าประมาณ	77
15	ของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(1,1,1) ₁₂	77
16	ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(2,1,0) ₁₂ ค่าประมาณ	77
17	ของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(2,1,1) ₁₂	77
18	ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(2,1,2) ₁₂	78
19	ค่า AIC, ค่า SBC และ SSE ของรูปแบบที่ตัดค่าคงที่ θ_0 ในรูปแบบออก	79
20	ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(2,1,2) ₁₂ ที่ตัดค่าคงที่ θ_0 ในรูปแบบออก	80

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
21	ค่า $r_k(e_t)$ และคอเรลโรแกรมของค่า $r_k(e_t)$ ของวิธีวิธีบ็อกซ์และเจนกินส์	82
22	ค่า MAD, MSE และ MAPE ของอนุกรมเวลาสำหรับค่าที่ได้จากการพยากรณ์ เปรียบเทียบกับค่าจริงด้วยวิธีบ็อกซ์และเจนกินส์	83
23	ค่า MAD, MSE และ MAPE ของวิธีการพยากรณ์เดี่ยวทั้งสามวิธี	84
24	การหาค่าความคลาดเคลื่อนของการพยากรณ์ด้วยวิธีแยกส่วนประกอบ วิธีวิน เตอร์ และวิธีบ็อกซ์และเจนกินส์	87
25	ค่า MAD, MSE และ MAPE ของอนุกรมเวลาสำหรับค่าที่ได้จากการพยากรณ์ เปรียบเทียบกับค่าจริงด้วยวิธีการพยากรณ์ร่วม	90
26	การเปรียบเทียบวิธีการพยากรณ์ปริมาณน้ำฝนรายเดือนของจังหวัดลำปาง ตั้งแต่ เดือนมิถุนายน 2551 ถึง เดือนพฤษภาคม 2552 ที่ได้จากการพยากรณ์ด้วยวิธี ต่างๆ	91
ตารางผนวกที่		
1	ข้อมูลปริมาณน้ำฝนของจังหวัดลำปาง ตั้งแต่ปี 2542 – 2552	98

สารบัญภาพ

ภาพที่		หน้า
1	ข้อมูลมีลักษณะ Horizontal	9
2	ข้อมูลมีลักษณะเป็นเส้นตรง	9
3	ข้อมูลมีลักษณะเป็นโพลีโนเมียลกำลังสอง หรือรูปพาราโบลา	10
4	ข้อมูลมีลักษณะเป็นโพลีโนเมียลกำลังสาม	10
5	ข้อมูลมีลักษณะเป็นเอ็กซ์โปเนนเชียล	11
6	ข้อมูลมีลักษณะเป็น Gompertz	12
7	ข้อมูลมีลักษณะเป็น Logistic	12
8	วัฏจักรทางธุรกิจ	15
9	ลักษณะการเคลื่อนไหวของอนุกรมเวลาที่มีความผันแปรของแนวโน้ม และฤดูกาลแบบบวกและแบบคูณ	18
10	กราฟแสดงค่า ρ_k และ ρ_{kk} สำหรับตัวแบบ ARMA(p,q) ต่างๆ	28
11	แผนภาพขั้นตอนในการวิจัย	56
12	การเคลื่อนไหวของปริมาณน้ำฝนจังหวัดลำปาง	58
13	ปริมาณน้ำฝนของจังหวัดลำปางเฉลี่ยเป็นรายปี ตั้งแต่ พ.ศ. 2542-2551	60
14	กราฟแสดงความคลาดเคลื่อนจากวิธีของวินเตอร์	66
15	ACF ของปริมาณน้ำฝนรายเดือนของจังหวัดลำปาง	69
16	ACF ของปริมาณน้ำฝนรายเดือนของจังหวัดลำปางที่แปลงข้อมูลด้วยการหาผลต่างฤดูกาล 12 เดือน	72
17	PACF ของปริมาณน้ำฝนรายเดือนของจังหวัดลำปางที่แปลงข้อมูลด้วยการหาผลต่างฤดูกาล 12 เดือน	73
18	กราฟแสดงความคลาดเคลื่อนจากวิธีของบ็อกซ์และเจนกินส์	81

การพยากรณ์ปริมาณน้ำฝนรายเดือนของจังหวัดลำปาง โดยใช้เทคนิคการพยากรณ์เชิงสถิติ

Forecasting Monthly Rainfall of Lampang Province by Statistical Forecasting Techniques

คำนำ

น้ำฝนนับว่าเป็นทรัพยากรที่มีความจำเป็นและสำคัญมากต่อทุกสิ่งมีชีวิตบนโลก นอกจากน้ำฝนมีประโยชน์ในด้านการอุปโภค บริโภค และด้านการเกษตรแล้ว น้ำฝนยังมีประโยชน์ในเชิงพลังงาน และในด้านการชลประทานอีกด้วย ดังนั้นการเปลี่ยนแปลงของปริมาณน้ำฝนที่สูงขึ้น หรือน้อยลงกว่าปกติ ล้วนส่งผลกระทบต่อและก่อให้เกิดความเสียหายต่อการดำเนินชีวิตในด้านเศรษฐกิจ สังคมความเป็นอยู่ และสิ่งแวดล้อมต่างๆ ถ้าปริมาณน้ำฝนลดลงมาก ก็จะทำให้เกิดความแห้งแล้งตามมา ซึ่งส่งผลกระทบต่อในด้านการอุปโภค บริโภค และในด้านการเกษตร พืชและสัตว์ที่ปลูกหรือเลี้ยงไว้เมื่อขาดน้ำก็จะล้มตาย ทำให้สูญเสียรายได้ แต่ถ้าปริมาณน้ำฝนเพิ่มขึ้นมาก ก็จะทำให้เกิดปัญหาทางด้านอุทกภัย น้ำป่าไหลหลาก เกิดปัญหาน้ำชะล้างหน้าดิน อาจเกิดอุบัติเหตุในการเดินทาง และเกิดโรคระบาด ทำให้มีสภาพความเป็นอยู่ที่ลำบาก เป็นต้น

ประเทศไทย มีการทำเกษตรกรรมเป็นหลัก น้ำฝนจัดเป็นปัจจัยสำคัญต่อการทำการเกษตร การมีน้ำใช้อย่างเพียงพอและสม่ำเสมอต่อการเพาะปลูก การเจริญเติบโตของพืช และการเลี้ยงสัตว์ ถือว่าเป็นสิ่งที่สำคัญที่สุดในความพยายามที่จะปรับปรุงผลผลิตทางการเกษตรให้มีจำนวนมาก และมีคุณภาพออกสู่ตลาด ทั้งนี้เพราะว่าการที่ฝนตกหรือฝนไม่ตก เป็นปัจจัยที่ไม่สามารถควบคุมได้ อีกทั้งปริมาณน้ำฝนมีความแปรปรวน ไม่แน่นอน และมีความแตกต่างกันในแต่ละท้องถิ่นอีกด้วย ดังนั้นจึงจำเป็นต้องใช้น้ำฝนที่มีให้เกิดประสิทธิภาพสูงสุด เพราะเราไม่สามารถทราบได้ว่าจะเกิดภัยแล้งขึ้นเมื่อไร แต่ถ้าสามารถพยากรณ์ปริมาณน้ำฝนล่วงหน้าได้ โดยการศึกษาข้อมูลในอดีต จะเป็นประโยชน์ในการวางแผน และเตรียมการป้องกันในด้านต่างๆ เช่น การเตรียมการป้องกันการเกิดอุทกภัย การเตรียมน้ำสำรองเพื่อป้องกันการเกิดภาวะขาดแคลนน้ำ เป็นต้น

จังหวัดลำปางมีการเพาะปลูกพืชเศรษฐกิจ ได้แก่ ข้าวนาปี อ้อย ถั่วเหลือง ถั่วลิสง ข้าวโพด ข้าวนาปรัง ยางพารา และมันสำปะหลัง เป็นต้น พืชเศรษฐกิจที่มีเนื้อที่เพาะปลูกมากที่สุดคือเนื้อที่ ข้าวนาปี เท่ากับ 416,095 ไร่ คิดเป็น ร้อยละ 73.42 ของเนื้อที่เพาะปลูกทั้งหมด รองลงมาคือเนื้อที่

เพาะปลูกอ้อย เท่ากับ 41,160 ไร่ คิดเป็น ร้อยละ 7.26 ของเนื้อที่เพาะปลูกทั้งหมด พืชที่มีเนื้อที่เพาะปลูกน้อยที่สุดได้แก่ มันสำปะหลัง เท่ากับ 598 ไร่ คิดเป็นร้อยละ 0.11 ของเนื้อที่เพาะปลูกทั้งหมด พื้นที่เพาะปลูกพืชเศรษฐกิจเหล่านี้มีแนวโน้มจะเพิ่มขึ้นทุกๆปี (www.thaienergydata.in.th) จากที่ได้กล่าวมาแล้วว่าการทำเกษตรกรรมต้องพึ่งพาน้ำในการเพาะปลูก ดังนั้นการจะทำเกษตรกรรมอย่างได้ผลนั้น จึงขึ้นอยู่กับปริมาณน้ำฝน

ผู้วิจัยจึงสนใจที่จะศึกษาการพยากรณ์ปริมาณน้ำฝนในจังหวัดลำปาง โดยศึกษาเทคนิคในการพยากรณ์น้ำฝน ทั้งวิธีการพยากรณ์เดี่ยว และ วิธีการพยากรณ์ร่วม วิธีการพยากรณ์เดี่ยวที่สนใจศึกษา ได้แก่ วิธีแยกส่วนประกอบ (Decomposition Method) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์ (Triple Exponential Smoothing : Winter's Method) วิธีบ็อกซ์และเจนกินส์ (Box – Jenkins Method) ซึ่งวิธีการพยากรณ์ทั้ง 3 นี้ เป็นวิธีที่เหมาะสมสำหรับการพยากรณ์ข้อมูลปริมาณน้ำฝน ซึ่งมีความผันแปรตามฤดูกาล อีกทั้งยังเหมาะกับการพยากรณ์ในระยะสั้น (ระยะเวลาประมาณ 1-3 เดือน) จนถึงการพยากรณ์ในระยะปานกลาง (ระยะเวลาประมาณ 3 เดือน - 2 ปี) และวิธีการพยากรณ์ร่วม ที่สนใจศึกษาได้แก่ วิธีเฉลี่ยถ่วงน้ำหนักซึ่งน้ำหนักผกผันกับค่าผลบวกกำลังสองของค่าความคลาดเคลื่อน

วัตถุประสงค์

1. เพื่อศึกษาเทคนิคการพยากรณ์ข้อมูลโดยวิธีแยกส่วนประกอบ (Decomposition Method) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์ (Triple Exponential Smoothing : Winter's Method) วิธีบ็อกซ์และเจนกินส์ (Box – Jenkins Method) และวิธีการพยากรณ์ร่วมใช้วิธีเฉลี่ยถ่วงน้ำหนักซึ่งน้ำหนักผูกพันกับค่าผลบวกกำลังสองของค่าความคลาดเคลื่อน
2. เพื่อสร้างสมการการพยากรณ์ปริมาณน้ำฝนในจังหวัดลำปาง
3. เพื่อเปรียบเทียบการพยากรณ์แต่ละวิธีว่าวิธีการพยากรณ์ใดที่เหมาะสมกับข้อมูลปริมาณน้ำฝนในจังหวัดลำปางมากที่สุด

ขอบเขตของการวิจัย

1. การศึกษาครั้งนี้ ศึกษาจากข้อมูลปริมาณน้ำฝนรายเดือนของจังหวัดลำปาง ซึ่งเก็บรวบรวมตั้งแต่เดือนมกราคม 2542 ถึงเดือนพฤษภาคม 2552 รวมระยะเวลา 125 เดือน ซึ่งแหล่งที่มาของข้อมูล คือ สำนักพัฒนาอคูนิยมวิทยา กรมอคูนิยมวิทยา กระทรวงเทคโนโลยีสารสนเทศ และการสื่อสาร
2. วิธีการพยากรณ์ในงานวิจัยนี้ใช้วิธีแยกส่วนประกอบ (Decomposition Method) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์ (Triple Exponential Smoothing : Winter's Method) วิธีบ็อกซ์และเจนกินส์ (Box – Jenkins Method) และวิธีการพยากรณ์ร่วมใช้วิธีเฉลี่ยถ่วงน้ำหนักซึ่งน้ำหนักผูกพันกับค่าผลบวกกำลังสองของค่าความคลาดเคลื่อน
3. เกณฑ์ที่ใช้ในการเปรียบเทียบความสามารถของวิธีการพยากรณ์ ได้แก่ ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (Mean Absolute Deviation : MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (Mean Absolute Percent Error : MAPE)
4. ระดับนัยสำคัญที่ใช้ในการทดสอบตลอดงานวิจัยนี้คือ $\alpha = 0.05$

ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถเลือกใช้วิธีการพยากรณ์ที่เหมาะสมในการพยากรณ์ปริมาณน้ำฝนในจังหวัดลำปาง
2. สมการพยากรณ์ปริมาณน้ำฝนของจังหวัดลำปาง
3. สามารถนำสมการพยากรณ์ปริมาณน้ำฝนไปใช้ในการวางแผน และเตรียมการในด้านต่างๆ เพื่อรองรับปัญหาเกี่ยวกับปริมาณน้ำฝนที่อาจเกิดขึ้น เช่น การเตรียมการป้องกันปัญหาภัยแล้ง หรือ การเตรียมการป้องกันการเกิดอุทกภัย น้ำป่าไหลหลาก เป็นต้น

การตรวจเอกสาร

การวิจัยนี้แบ่งออกเป็น 2 ส่วน ดังนี้

1. เทคนิคการพยากรณ์
2. ผลงานวิจัยที่เกี่ยวข้อง

ซึ่งมีรายละเอียดดังนี้

เทคนิคการพยากรณ์ ประกอบด้วย

1. วิธีการพยากรณ์เดี่ยว

1.1 วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์ (Triple Exponential Smoothing : Winters's Method)

เป็นวิธีหนึ่งของวิธีการทำให้เรียบแบบเอ็กซ์โปเนนเชียล ที่เหมาะสำหรับข้อมูลอนุกรมเวลาที่มีแนวโน้มและความผันแปรตามฤดูกาล (Trend-Season Data) และเหมาะสำหรับการพยากรณ์ในระยะสั้นจนถึงการพยากรณ์ในระยะปานกลางด้วย วิธีของวินเตอร์ขึ้นอยู่กับรูปแบบของข้อมูลอนุกรมว่าเป็นรูปแบบการบวกหรือการคูณ ซึ่งมีรายละเอียดดังนี้

ก. รูปแบบการคูณ (Multiplicative Seasonality Model)

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

เมื่อ Y_t คือ ข้อมูลหรือค่าสังเกต ณ เวลา t

β_0, β_1 คือ ค่าพารามิเตอร์ของรูปแบบ

ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t

S_t คือ ค่าความผันแปรตามฤดูกาล ณ เวลา t

สมการที่ใช้ในการพยากรณ์ คือ

$$\hat{Y}_{t+m} = \begin{cases} (a_t + b_t(m))\hat{S}_t & \text{เมื่อ } t \leq p \\ (a_t + b_t(m))\hat{S}_{t-p+m} & \text{เมื่อ } t > p \end{cases}$$

ค่า a_t คือระดับของข้อมูลหรือส่วนที่เป็นการทำให้เรียบ

$$a_t = \begin{cases} \frac{\alpha Y_t}{\hat{S}_t} + (1-\alpha)[a_{t-1} + b_{t-1}] & \text{เมื่อ } t \leq p \\ \frac{\alpha Y_t}{\hat{S}_{t-p}} + (1-\alpha)[a_{t-1} + b_{t-1}] & \text{เมื่อ } t > p \end{cases}$$

ค่า b_t คือส่วนที่เป็นแนวโน้ม

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

ค่า \hat{S}_t คือส่วนที่เป็นฤดูกาล

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

m คือ จำนวนช่วงเวลาที่ต้องการพยากรณ์ไปข้างหน้า

p คือ จำนวนฤดูกาล โดยการกำหนดค่า p ขึ้นอยู่กับลักษณะของข้อมูล ดังนี้

$p = 4$ สำหรับข้อมูลเป็นรายไตรมาส $p = 7$ สำหรับข้อมูลเป็นรายวัน และ $p = 12$ สำหรับข้อมูลรายเดือน

α คือ ค่าคงที่การทำให้เรียบระหว่างข้อมูลกับค่าพยากรณ์ โดยที่ $0 \leq \alpha \leq 1$

γ คือ ค่าคงที่การทำให้เรียบระหว่างแนวโน้มจริงกับค่าประมาณของแนวโน้ม โดยที่ $0 \leq \gamma \leq 1$

δ คือ ค่าคงที่การทำให้เรียบระหว่างค่าฤดูกาลจริงกับค่าประมาณฤดูกาล โดยที่ $0 \leq \delta \leq 1$

ข. รูปแบบการบวก (Additive Seasonality Model)

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

สมการที่ใช้ในการพยากรณ์ คือ $\hat{Y}_{t+m} = a_t + b_t(m) + \hat{S}_{t-p+m}$

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

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

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

การตรวจสอบความเหมาะสมของรูปแบบ

ในกรณีที่กราฟของอนุกรมเวลาแสดงลักษณะของการแปรผันตามฤดูกาลรูปแบบการบวก หรือรูปแบบการคูณไม่ชัดเจน การตรวจสอบว่ารูปแบบที่เลือกไว้มีความเหมาะสมกับอนุกรมเวลา หรือไม่ ทำได้โดยการพิจารณาค่าสหสัมพันธ์ในตัวเองของความคลาดเคลื่อนของค่าพยากรณ์หนึ่ง หน่วยเวลาล่วงหน้า ($r_k(e_t)$), $k = 1, 2, \dots, 16$ อยู่ในช่วง $(-2s[r_k(e_t)], 2s[r_k(e_t)])$ หรือไม่

และพิจารณาค่าสถิติ Box-Ljung Chi-square มีค่า p-value มากกว่า 0.05 หรือไม่ ถ้ามีค่ามากกว่า 0.05 นั่นคือยอมรับได้ว่าสหสัมพันธ์ในเองของ e_t ไม่ต่างจากศูนย์ แสดงว่าวิธีวินเตอร์รูปแบบนี้เหมาะสมแล้ว (มุกดา, 2549: 97)

1.2 วิธีแยกส่วนประกอบ (Decomposition Method)

วิธีแยกส่วนประกอบเป็นวิธีการศึกษาว่าอนุกรมเวลาที่สนใจได้รับอิทธิพลจากความผันแปรใดบ้าง จากนั้นทำการวัดความผันแปรเหล่านั้นเพื่อนำไปหาค่าพยากรณ์ในอนาคต ซึ่งอิทธิพลจากความผันแปรต่างๆ ที่เป็นส่วนประกอบของอนุกรมเวลามีดังนี้

ก. แนวโน้ม (Trend :T)

เป็นส่วนประกอบหนึ่งของอนุกรมเวลา การเปลี่ยนแปลงอนุกรมเวลาในระยะยาว แสดงให้เห็นการเคลื่อนไหวของอนุกรมเวลาในทางที่เพิ่มขึ้น (Upward Trend) หรือ ลดลง (Downward Trend) ซึ่งลักษณะของแนวโน้มเป็นไปได้ทั้งแบบเส้นตรงและแบบเส้นโค้ง

การวิเคราะห์ค่าแนวโน้มเป็นการศึกษาการเคลื่อนไหวของข้อมูลที่เกิดขึ้นในระยะยาว ว่าข้อมูลมีรูปแบบการเคลื่อนไหวเป็นแบบเส้นตรงหรือเส้นโค้ง เส้นที่ใช้แทนการเคลื่อนไหวของข้อมูล ต้องเป็นเส้นที่เรียบ ไม่มีมุมหัก และเรียกเส้นนี้ว่าเส้นแนวโน้ม เส้นแนวโน้มที่สร้างขึ้นมีประโยชน์สำหรับการพยากรณ์ข้อมูลในอนาคต ภายใต้อสมมติและสภาพแวดล้อมเดียวกับอดีต

ค่าที่พยากรณ์ได้จากเส้นแนวโน้มเรียกว่า ค่าแนวโน้ม เส้นแนวโน้มที่ดีจะต้องสามารถพยากรณ์ค่าได้ใกล้เคียงกับค่าจริงมากที่สุด การสร้างเส้นแนวโน้มโดยวิธีการคำนวณมีหลายวิธี แต่วิธีที่นิยมใช้ ได้แก่ วิธีกำลังสองน้อยที่สุด (Least square method) เป็นวิธีการสร้างเส้นแนวโน้มที่ใช้แนวคิดที่จะทำให้ผลรวมกำลังสองของผลต่างระหว่างค่าแนวโน้มกับค่าข้อมูลที่เก็บรวบรวมมาได้มีค่าน้อยที่สุด การสร้างเส้นแนวโน้มด้วยวิธีกำลังสองน้อยที่สุด เป็นวิธีที่ได้รับความนิยมเพราะเป็นวิธีที่ใช้คุณสมบัติทางคณิตศาสตร์ที่สำคัญ และน่าเชื่อถือในการหาเส้นที่เหมาะสมที่สุด ที่เป็นตัวแทนของข้อมูลที่เก็บรวบรวมมาได้และใช้สร้างเส้นแนวโน้มที่เป็นเส้นตรงหรือไม่เป็นเส้นตรงก็ได้ เมื่อพิจารณาแผนภาพการกระจายข้อมูลที่เก็บมาได้ อาจมีรูปการกระจายที่

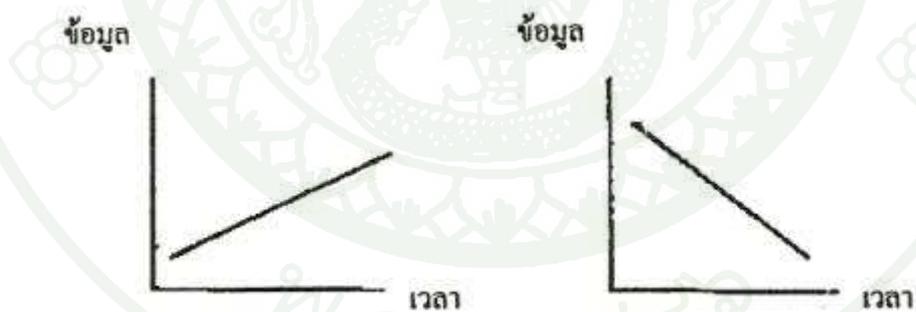
มีแนวโน้มเป็นเส้นตรงหรือเส้นโค้งแบบต่างๆ มากมาย รูปทั่วไปของสมการแนวโน้มที่ได้จากการคำนวณด้วยวิธีกำลังสองน้อยที่สุด จึงมีหลายรูปแบบ ดังต่อไปนี้ (แสดงดังภาพที่ 1-7)



ภาพที่ 1 ข้อมูลมีลักษณะ Horizontal

ที่มา: จันทรรักษ์ (2548)

จากภาพที่ 1 แสดงเส้นแนวโน้มที่เกิดจากข้อมูลที่มีลักษณะคงที่ ไม่มีแนวโน้มและฤดูกาล



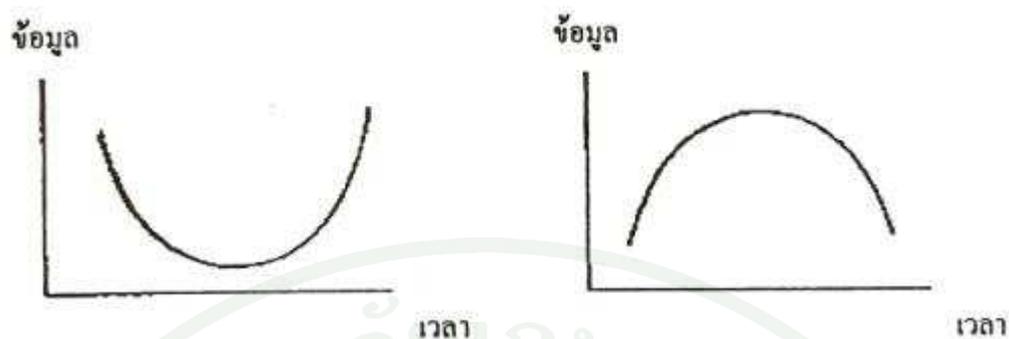
(ก) แสดงแนวโน้มเพิ่มขึ้น

(ข) แสดงแนวโน้มลดลง

ภาพที่ 2 ข้อมูลมีลักษณะเป็นเส้นตรง

ที่มา: ศิริลักษณ์ (2543)

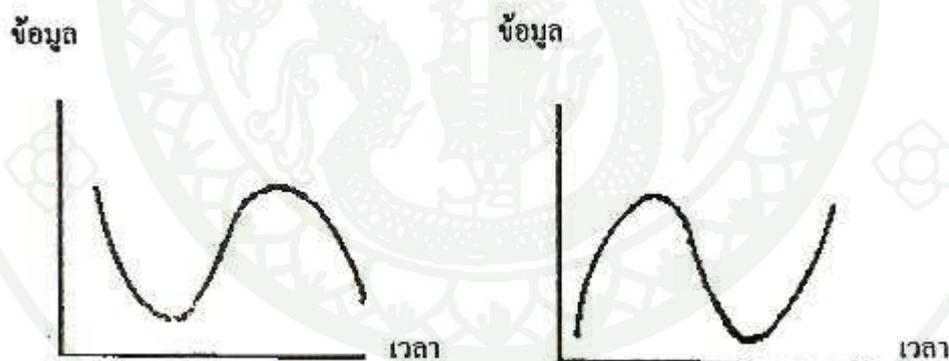
จากภาพที่ 2 แสดงเส้นแนวโน้มที่เกิดจากข้อมูลที่มีค่าเพิ่มขึ้น (ภาพ (ก)) หรือลดลงคงที่ (ภาพ (ข)) เมื่อเวลาเปลี่ยนไป



ภาพที่ 3 ข้อมูลมีลักษณะเป็นโพลีโนเมียลกำลังสอง หรือรูปพาราโบลา

ที่มา: ศิริลักษณ์ (2543)

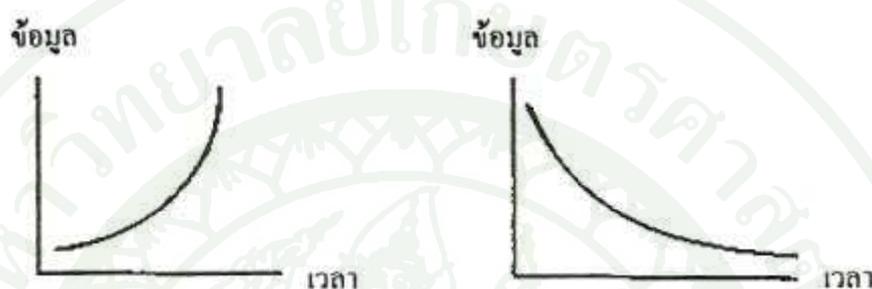
จากภาพที่ 3 แสดงเส้นแนวโน้มที่เกิดจากข้อมูลที่มีลักษณะการเพิ่มขึ้นหรือลดลงไม่คงที่ แต่ทิศทางการเพิ่มขึ้นหรือลดลงคงที่ จึงทำให้ลักษณะเส้นแนวโน้มเป็นกราฟหงายหรือกราฟคว่ำ



ภาพที่ 4 ข้อมูลมีลักษณะเป็นโพลีโนเมียลกำลังสาม

ที่มา: ศิริลักษณ์ (2543)

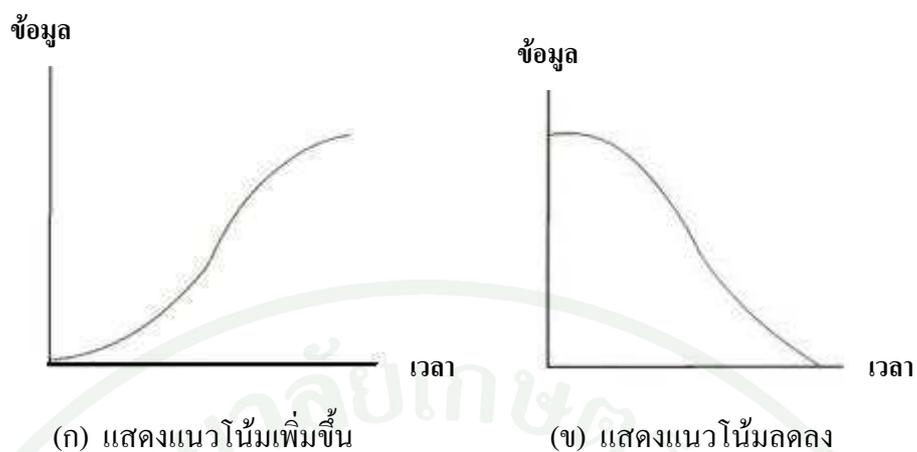
จากภาพที่ 4 แสดงเส้นแนวโน้มที่เกิดจากข้อมูลที่มีลักษณะการเพิ่มขึ้นหรือลดลงไม่คงที่ และทิศทางการเพิ่มขึ้นหรือลดลงไม่คงที่ ทำให้แผนภาพการกระจายมีลักษณะเป็นลูกคลื่น การวิเคราะห์แนวโน้มที่มีลักษณะนี้จะต้องระมัดระวังมาก เพราะแผนภาพการกระจายลักษณะเป็นลูกคลื่นที่เกิดขึ้นอาจเป็นผลเนื่องจากการผันแปรตามวัฏจักรก็ได้ และการวิเคราะห์แนวโน้มลักษณะนี้มีความยุ่งยากในการคำนวณด้วย



ภาพที่ 5 ข้อมูลมีลักษณะเป็นเอ็กซ์โปเนนเชียล

ที่มา: ศิริลักษณ์ (2543)

จากภาพที่ 5 แสดงเส้นแนวโน้มที่เกิดกับข้อมูลที่มีลักษณะการเพิ่มขึ้นหรือลดลงไม่เป็นเส้นตรง และค่าการเปลี่ยนแปลงไม่คงที่ แต่มีทิศทางการเพิ่มขึ้นหรือลดลงไปในทางเดียวกัน และอัตราการเพิ่มขึ้นหรือลดลงคงที่ แนวโน้มรูปเอ็กซ์โปเนนเชียลมักนิยมใช้กับอนุกรมเวลาของข้อมูลทางธุรกิจที่กำลังรุ่งเรือง เพราะสามารถพยากรณ์ค่าแนวโน้มได้ดีกว่ารูปเส้นตรงหรือโพลิโนเมียลกำลังสอง



(ก) แสดงแนวโน้มเพิ่มขึ้น

(ข) แสดงแนวโน้มลดลง

ภาพที่ 6 ข้อมูลมีลักษณะเป็น Gompertz

จากภาพที่ 6 แสดงเส้นแนวโน้มซึ่งมีส่วนเพิ่มหรือลดลงของลอการิทึมของข้อมูลเพิ่มขึ้นหรือลดลงด้วยเปอร์เซ็นต์คงที่



ภาพที่ 7 ข้อมูลมีลักษณะเป็น Logistic

ที่มา: จันทรธา (2548)

จากภาพที่ 7 แสดงเส้นแนวโน้มที่เป็นส่วนกลับของข้อมูลมีลักษณะเป็นเอ็กซ์โปเนนเชียล

ข. ฤดูกาล (Seasonal : S)

เป็นส่วนประกอบของอนุกรมเวลา การเปลี่ยนแปลงที่เกิดขึ้นซ้ำๆ กัน ในช่วงเวลาหนึ่งอาจเป็น 1 ปี 1 เดือน 1 สัปดาห์ หรือ 1 วัน การเปลี่ยนแปลงของฤดูกาลมี 2 ลักษณะคือ ฤดูกาลที่คงที่ (Constant Seasonal) เป็นการเปลี่ยนแปลงที่เกิดขึ้นทุกปี ด้วยรูปแบบที่คงที่ในช่วงเวลานั้น และฤดูกาลเปลี่ยนแปลง (Changing Seasonal) เป็นการเปลี่ยนแปลงของฤดูกาลที่มีรูปแบบการเปลี่ยนแปลงไปเรื่อยๆ อย่างช้าๆ ในช่วงเวลาหนึ่งๆ บางครั้งอาจจะมีรูปแบบของฤดูกาลคงที่อยู่เป็นเวลานานหลายปี แล้วค่อยๆ เปลี่ยนไป กรณีที่รูปแบบของอนุกรมเวลาเป็นแบบคูณจะวัดความผันแปรของฤดูกาลเป็นดัชนีฤดูกาล (Seasonal Index) และกรณีที่รูปแบบการเคลื่อนไหวของอนุกรมเวลาเป็นแบบบวกจะวัดความผันแปรของฤดูกาลเป็นค่าวัคคิทธิพลของฤดูกาล (Seasonal Factor)

การวิเคราะห์ความผันแปรตามฤดูกาล (Seasonal Variation) หมายถึงการแยกความผันแปรตามฤดูกาลออกจากข้อมูลอนุกรมเวลาที่มีความผันแปรตามฤดูกาลเป็นส่วนประกอบ การแยกความผันแปรตามฤดูกาลต้องคำนึงถึงรูปแบบอนุกรมเวลามีลักษณะรูปแบบการบวก หรือ รูปแบบการคูณ สำหรับการวัดความผันแปรตามฤดูกาลที่วัดออกมาในรูปดัชนีฤดูกาลมีหลายวิธี วิธีที่ใช้ในงานวิจัยนี้คือ

วิธีอัตราส่วนต่อค่าเฉลี่ยเคลื่อนที่ (Ratio to Moving Average)

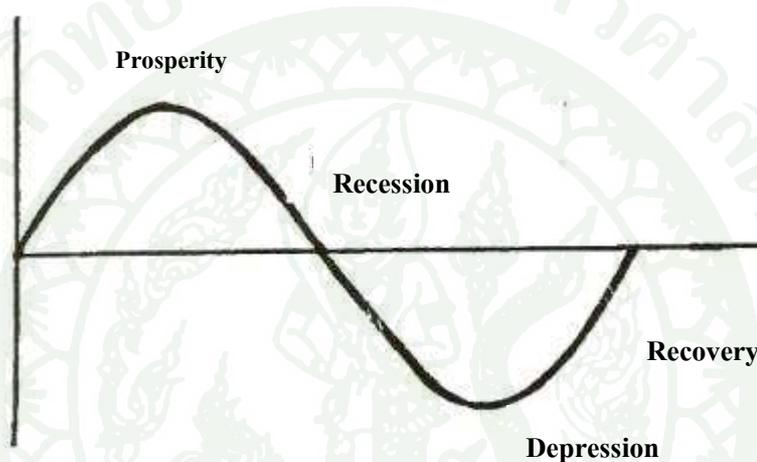
เป็นวิธีคำนวณค่าดัชนีฤดูกาลที่ค่อนข้างง่าย โดยใช้หลักการกำจัดส่วนประกอบอนุกรมเวลาอื่นๆ ให้หมดไป เพื่อคงเหลือแต่ความผันแปรตามฤดูกาลเพียงอย่างเดียว

ขั้นตอนการคำนวณค่าดัชนีฤดูกาลด้วยวิธีอัตราส่วนต่อค่าเฉลี่ยเคลื่อนที่ มีดังนี้

- 1) คำนวณค่าเฉลี่ยเคลื่อนที่จากข้อมูลอนุกรมเวลา ค่าเฉลี่ยที่คำนวณได้เป็นค่าแนวโน้มของข้อมูล และข้อมูลที่นำมาคำนวณค่าเฉลี่ยเคลื่อนที่เป็นข้อมูลในช่วงเวลาสั้นๆ เป็นฤดูกาลที่คาบเวลาไม่ถึง 1 ปี ข้อมูลลักษณะนี้จึงมีส่วนประกอบอนุกรมเวลาในส่วนที่มีความผันแปรตามวัฏจักรเพียงเล็กน้อย และเมื่อนำมาคำนวณหาค่าเฉลี่ยเคลื่อนที่ ส่วนประกอบที่เป็นความผันแปรตามวัฏจักรจะถูกแยกไปพร้อมกับค่าเฉลี่ยเคลื่อนที่ที่คำนวณได้
- 2) คำนวณหาอัตราส่วนค่าเฉลี่ยเคลื่อนที่ เป็นการกำจัดค่าแนวโน้มและความผันแปรตามวัฏจักรออกจากข้อมูล โดยการหารข้อมูลเดิมด้วยค่าเฉลี่ยเคลื่อนที่ที่คำนวณได้ ผลหารที่เกิดขึ้นจะเหลือแต่ส่วนประกอบความผันแปรตามฤดูกาล และความผันแปรเนื่องจากเหตุการณ์ที่ผิดปกติเท่านั้น
- 3) จัดรูปอัตราส่วนต่อค่าเฉลี่ยเคลื่อนที่ให้อยู่ในรูปร้อยละ โดยการนำ 100 มาคูณ คำนวณหาค่าเฉลี่ยของอัตราส่วนร้อยละต่อค่าเฉลี่ยเคลื่อนที่ เพื่อกำจัดความผันแปรเนื่องจากเหตุการณ์ที่ผิดปกติ ค่าเฉลี่ยที่ใช้อาจเป็นค่าเฉลี่ยเลขคณิตหรือค่ามัธยฐานก็ได้ ตามความเหมาะสมของข้อมูล
- 4) ค่าเฉลี่ยที่คำนวณได้ยังอยู่ในรูปร้อยละ ซึ่งเป็นค่าดัชนีฤดูกาลโดยประมาณจึงต้องปรับให้เป็นค่าดัชนีฤดูกาลที่แท้จริง

ค. วัฏจักร (Cycle : C)

เป็นการเปลี่ยนแปลงที่มีลักษณะคล้ายความผันแปรของฤดูกาล แต่มีช่วงเวลาที่นานกว่า วัฏจักรมักปรากฏในข้อมูลอนุกรมเวลาทางธุรกิจ โดยทั่วไปวัฏจักรหนึ่ง จะประกอบด้วยช่วงเวลา 4 ระยะ คือ ระยะเจริญรุ่งเรือง (Prosperity) ระยะที่ค่อยๆ ตกต่ำ (Recession) ระยะตกต่ำ (Depression) และระยะฟื้นตัว (Recovery) แสดงดังภาพที่ 8



ภาพที่ 8 วัฏจักรทางธุรกิจ

ที่มา: จันทรธา (2548)

ข้อมูลอนุกรมเวลาบางชุดอาจมีวัฏจักรไม่ครบทั้ง 4 ระยะ และวัฏจักรหนึ่งๆ อาจมีรูปร่างและคาบเวลาหรือความยาวของวัฏจักรไม่แน่นอน จึงยากที่จะพยากรณ์การเปลี่ยนแปลงการเคลื่อนที่ของวัฏจักร ทั้งนี้จะต้องอาศัยประสบการณ์ประกอบด้วย ดังนั้นการคำนวณหาความผันแปรตามวัฏจักรจึงแตกต่างจากการคำนวณหาความผันแปรตามฤดูกาล และกรณีที่วัฏจักรมีความยาวมากๆ จะไม่นิยมนำมาวิเคราะห์ห้อนุกรมเวลา เพราะข้อมูลที่เก็บรวบรวมมาไม่เพียงพอ โดยทั่วไปการวิเคราะห์ห้อนุกรมเวลาเพื่อการพยากรณ์ นิยมพยากรณ์ข้อมูลในช่วงเวลาสั้นๆ ซึ่งได้รับอิทธิพลจากแนวโน้มความผันแปรตามฤดูกาลหรือความผันแปรตามวัฏจักรที่มีระยะสั้นๆ เท่านั้น

การวิเคราะห์อนุกรมเวลาเพื่อประมาณค่าความผันแปรตามวัฏจักร สามารถทำได้โดยการกำจัดส่วนประกอบอนุกรมเวลา ได้แก่ แนวโน้ม ความผันแปรตามฤดูกาล และความผันแปรอื่นเนื่องมาจากเหตุการณ์ผิดปกติ ออกไปทีละส่วน หรือ พร้อมๆ กัน การประมาณค่าความผันแปรตามวัฏจักร โดยวิธีการกำจัดส่วนประกอบอื่นๆ ออกไป เรียกว่า วิธีแยกค้ำ การประมาณค่าความผันแปรด้วยวิธีแยกค้ำ จะขึ้นอยู่กับข้อมูลที่นำมาวิเคราะห์ว่าเป็นข้อมูลรายปี รายเดือน รายไตรมาส กรณีข้อมูลรายปีจะไม่มีส่วนประกอบของความผันแปรตามฤดูกาล การกำจัดส่วนประกอบอนุกรมเวลาจะกำจัดเพียงแนวโน้มและความผันแปรอื่นเนื่องมาจากความไม่สม่ำเสมอเท่านั้น แต่ถ้าข้อมูลเป็นรายไตรมาส หรือรายเดือน จะมีความผันแปรตามฤดูกาลเพิ่มเข้ามา การกำจัดส่วนประกอบอื่นๆ ออกจากส่วนประกอบอนุกรมเวลาจะต้องกำจัดแนวโน้ม ความผันแปรตามฤดูกาล และความผันแปรอื่นเนื่องมาจากความไม่สม่ำเสมอ เพื่อคงเหลือแต่ความผันแปรตามวัฏจักรเท่านั้น วิธีการกำจัดส่วนประกอบอื่นๆ ออกจากข้อมูล เพื่อคงเหลือแต่ความผันแปรตามวัฏจักร จะต้องคำนึงถึงรูปแบบของอนุกรมเวลาว่าเป็นรูปแบบการบวก หรือ รูปแบบการคูณ กรณีที่ตัวแบบของอนุกรมเวลาเป็นรูปแบบการบวก วิธีการกำจัดส่วนประกอบอื่นๆ ออกจากข้อมูล เพื่อคงเหลือแต่ความผันแปรตามวัฏจักร จะใช้วิธีการหักค่าส่วนประกอบอื่นๆ ในรูปแบบออกจากข้อมูล ถ้ารูปแบบอนุกรมเวลาเป็นรูปแบบการคูณ วิธีการกำจัดส่วนประกอบอื่นๆ ออกจากข้อมูล เพื่อคงเหลือแต่ความผันแปรตามวัฏจักร จะใช้วิธีการหารค่าส่วนประกอบอื่นๆ ในรูปแบบออกจากข้อมูล

ง. เหตุการณ์ผิดปกติ (Irregular : I)

เป็นการเปลี่ยนแปลงที่เกิดขึ้นเนื่องจากเหตุการณ์ที่ไม่คาดคิดมาก่อนว่าจะเกิดขึ้นในช่วงเวลาใด และเป็นส่วนประกอบที่ทำให้ข้อมูลมีค่ามากหรือน้อยจนรูปแบบการเคลื่อนที่ของข้อมูลผิดปกติ แต่ความผิดปกติจะเกิดขึ้นเพียงช่วงเวลาใดเวลาหนึ่งเท่านั้น ความผิดปกติอาจมีสาเหตุมาจากสภาพแวดล้อมหรือธรรมชาติ เช่น เกิดสงคราม การประกาศลดค่าเงิน เป็นต้น การพิจารณาว่าข้อมูลอนุกรมเวลาที่มีความผันแปรเนื่องจากเหตุการณ์ผิดปกติหรือไม่ สามารถพิจารณาได้จากแผนภาพของข้อมูลอนุกรมเวลา ว่ามีตำแหน่งใดสูงหรือต่ำกว่าปกติ และไม่เป็นไปตามลักษณะการเคลื่อนที่ตามปกติของข้อมูลหรือไม่ ถ้าข้อมูลมีลักษณะดังกล่าวเกิดขึ้น แสดงว่ามีความผันแปรเนื่องมาจากเหตุการณ์ผิดปกติในช่วงเวลานั้น

รูปแบบของอนุกรมเวลาที่มีสาเหตุของความผันแปรต่างๆมี 2 รูปแบบ ดังนี้

1. รูปแบบการบวก (Additive Model) เป็นรูปแบบที่มีความผันแปรแต่ละชนิดที่เป็นอิสระต่อกัน กล่าวคือส่วนประกอบใดส่วนประกอบหนึ่งมีค่าเปลี่ยนแปลงไม่ว่าจะเพิ่มขึ้นหรือลดลง จะไม่ส่งผลกระทบต่อค่าการเปลี่ยนแปลงของส่วนประกอบอื่นๆ ที่เหลือ มีรูปแบบดังนี้

$$Y = T + S + C + I$$

โดยที่ Y = อนุกรมเวลา

T = แนวโน้ม

S = ความผันแปรตามฤดูกาล

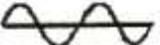
C = ความผันแปรตามวัฏจักร

I = ความผันแปรเนื่องจากเหตุการณ์ที่ผิดปกติ

2. รูปแบบการคูณ (Multiplicative Model) เป็นรูปแบบที่มีความผันแปรแต่ละชนิดที่มีความสัมพันธ์กัน หรือไม่เป็นอิสระต่อกัน กล่าวคือ ถ้าส่วนประกอบใดส่วนประกอบหนึ่งมีค่าเปลี่ยนแปลงไปไม่ว่าจะเพิ่มขึ้นหรือลดลง จะมีผลกระทบต่อค่าการเปลี่ยนแปลงของส่วนประกอบอื่นๆ ที่เหลือ มีรูปแบบดังนี้

$$Y = T \times S \times C \times I$$

การกำหนดว่ารูปแบบของอนุกรมเวลาเป็นรูปแบบการบวก หรือรูปแบบการคูณ และมีแนวโน้มเป็นแบบคงที่ เส้นตรง หรือเส้นโค้งนั้นขึ้นอยู่กับลักษณะของอนุกรมเวลา และเมื่อสร้างกราฟของอนุกรมเวลา ทำให้เห็นลักษณะของอนุกรมเวลาได้ ลักษณะของอนุกรมเวลารูปแบบการบวก และรูปแบบการคูณ เมื่อมีการเปลี่ยนแปลงเนื่องจากแนวโน้มและฤดูกาล มีลักษณะการเคลื่อนไหว ดังภาพที่ 9

รูปแบบ	อนุกรมเวลาที่มี แนวโน้ม	อนุกรมเวลา ที่มีแนวโน้มและ ความผันแปรของ ฤดูกาลแบบบวก	อนุกรมเวลา ที่มีแนวโน้มและ ความผันแปรของ ฤดูกาลแบบคูณ
แนวโน้มคงที่			
แนวโน้มเป็นเส้นตรง			
แนวโน้มเป็นเส้นโค้ง			
			

ภาพที่ 9 ลักษณะการเคลื่อนไหวของอนุกรมเวลาที่มีความผันแปรของแนวโน้มและฤดูกาลแบบบวกและแบบคูณ

ที่มา: จันทรธา (2548)

ในการพิจารณาว่าอนุกรมเวลาที่มีแนวโน้มคงที่หรือไม่สามารถพิจารณาได้จากค่าอนุกรมเวลาเมื่อมีค่าอยู่ระหว่างค่าคงที่ค่าหนึ่ง แนวโน้มเส้นตรงมีลักษณะการเคลื่อนไหวเพิ่มขึ้นหรือลดลงในอัตราที่คงที่ ส่วนแนวโน้มเส้นโค้งมีลักษณะการเคลื่อนไหวเพิ่มขึ้นหรือลดลงในอัตราที่ไม่คงที่ เช่น เพิ่มขึ้นแล้วค่อยๆ ลดลง และสามารถพิจารณาว่าอนุกรมเวลามีรูปแบบการบวก หรือรูปแบบการคูณได้จากลูกคลื่น ที่มีการเปลี่ยนแปลงขึ้นลง ถ้าลักษณะของลูกคลื่นมีการขึ้นลงอย่างสม่ำเสมอ แสดงว่ามีรูปแบบการบวก แต่ถ้าลักษณะของลูกคลื่นมีการเพิ่มขึ้นเรื่อยๆ เมื่อเวลาผ่านไป แสดงว่ามีรูปแบบการคูณ

สำหรับอนุกรมเวลาที่มีแนวโน้มและฤดูกาลจะมีขั้นตอนการหาค่าแนวโน้มและค่าวัดอิทธิพลของฤดูกาลจำแนกตามรูปแบบของอนุกรมเวลา ดังนี้

สำหรับรูปแบบการบวก

1. ปรับอนุกรมเวลาเดิม (Y_t) ด้วยค่าแนวโน้มที่วัดโดยการหาค่าเฉลี่ยเคลื่อนที่กลาง โดยการลบค่าเฉลี่ยเคลื่อนที่ (MA_t) ออกจากค่า Y_t ได้อนุกรมเวลาใหม่ ($Y_t - MA_t$)
2. หาค่าวัดอิทธิพลของฤดูกาลจากการเฉลี่ยอนุกรมเวลาที่ได้ในข้อที่ 1 แยกตามฤดูกาล ค่าวัดอิทธิพลของฤดูกาล (\hat{S}_i) มักมีผลรวมไม่เป็น 0 การปรับให้มีผลรวมเป็น 0 สามารถทำได้โดยนำค่าเฉลี่ยของค่าวัดอิทธิพลของฤดูกาลเดิม (\bar{S}) มาลบออกจาก \hat{S}_i จะได้ \hat{S}_i^* เป็นค่าวัดอิทธิพลของฤดูกาลใหม่ที่ i

ซึ่ง

$$\hat{S}_i^* = \hat{S}_i - \bar{S} \quad \text{ที่มี} \quad \sum \hat{S}_i^* = 0$$

3. หาอนุกรมเวลาปรับฤดูกาล โดยนำค่าวัดอิทธิพลของฤดูกาลที่ได้ในข้อที่ 2 ไปลบออกจากค่าสังเกตในอนุกรมเวลาเดิม นำอนุกรมเวลาปรับฤดูกาลนี้ไปสร้างสมการแนวโน้ม โดยวิธีกำลังสองน้อยที่สุด (Least Square Method)

กรณีแนวโน้มเส้นตรง สมการแนวโน้มคือ

$$\hat{T}_t = b_0 + b_1t$$

ได้สมการพยากรณ์เป็น

$$\hat{Y}_t = (b_0 + b_1t) + \hat{S}_t^*$$

กรณีแนวโน้มเส้นโค้ง สมการแนวโน้มคือ

$$\hat{T}_t = b_0 + b_1t + b_2t^2$$

ได้สมการพยากรณ์เป็น

$$\hat{Y}_t = (b_0 + b_1t + b_2t^2) + \hat{S}_t^*$$

สำหรับรูปแบบการคูณ

1. ปรับอนุกรมเวลาเดิม (Y_t) ด้วยค่าแนวโน้มที่วัดโดยการหาค่าเฉลี่ยเคลื่อนที่กลาง ด้วยค่าแนวโน้ม ได้อนุกรมเวลาใหม่ (Y_t / MA_t) ที่มีการเคลื่อนไหวเนื่องมาจากฤดูกาลและเหตุการณ์ที่ผิดปกติ

2. หาดัชนีฤดูกาลจากการเฉลี่ยอนุกรมเวลาที่ได้ในข้อที่ 1 แยกตามฤดูกาล ดัชนีฤดูกาล (\hat{S}_i) มักมีผลรวมไม่เท่ากับ L ปรับให้มีผลรวมเป็น L ทำได้โดยการนำค่าเฉลี่ยของดัชนีฤดูกาล (\bar{S}) มาหารนั่นคือ \hat{S}_i^* เป็นดัชนีฤดูกาลใหม่ที่ i

ซึ่ง

$$\hat{S}_i^* = \hat{S}_i / \bar{S} \quad \text{ที่มี} \quad \sum \hat{S}_i^* = L$$

3. หออนุกรมเวลาปรับฤดูกาลโดยนำดัชนีฤดูกาลที่ได้จากข้อที่ 2 ไปหารค่าสังเกตในอนุกรมเวลาเดิม จากนั้นนำอนุกรมเวลาปรับฤดูกาลนี้ไปสร้างสมการแนวโน้ม โดยวิธีกำลังสองน้อยที่สุด

กรณีแนวโน้มเส้นตรง สมการแนวโน้มคือ

$$\hat{T}_t = b_0 + b_1t$$

ได้สมการพยากรณ์เป็น

$$\hat{Y}_t = (b_0 + b_1t)\hat{S}_t^*$$

กรณีแนวโน้มเส้นโค้ง สมการแนวโน้มคือ

$$\hat{T}_t = b_0 + b_1t + b_2t^2$$

ได้สมการพยากรณ์เป็น

$$\hat{Y}_t = (b_0 + b_1t + b_2t^2)\hat{S}_t^*$$

กรณีแนวโน้มเอ็กซ์โปเนนเชียล สมการแนวโน้มคือ

$$\hat{T}_t = b_0b_1^t$$

ได้สมการพยากรณ์เป็น

$$\hat{Y}_t = (b_0b_1^t)\hat{S}_t^*$$

1.3 วิธีการบ็อกซ์และเจนกินส์ (Box – Jenkins Method)

วิธีการบ็อกซ์และเจนกินส์เป็นวิธีวิเคราะห์อนุกรมเวลา โดยข้อมูลอนุกรมเวลาที่นำมาวิเคราะห์ต้องเป็นอนุกรมเวลาที่มีคุณสมบัติสแตชันนารี (Stationary) กรณีที่อนุกรมเวลาไม่เป็นสแตชันนารี จะต้องทำการแปลงอนุกรมเวลาใหม่ให้มีคุณสมบัติสแตชันนารี

1.3.1 อนุกรมเวลาที่เป็นสแตชันนารี (Stationary Time Series)

เป็นอนุกรมเวลาที่ค่าสังเกต (Y_t) มีคุณสมบัติทางสถิติ คือ ค่าเฉลี่ย ($E(Y_t)$) ความแปรปรวน ($V(Y_t)$) และฟังก์ชันความน่าจะเป็นของค่าสังเกต ณ เวลาต่างๆ กันคงที่ กล่าวคือ ไม่เปลี่ยนแปลงไปตามเวลาที่เปลี่ยนไป นอกจากนี้อนุกรมเวลาที่เป็นสแตชันนารี จะต้องมีการสัมพันธ์สหสัมพันธ์ในตัวเองที่ lag (Autocorrelation Function) ขึ้นอยู่กับค่า k อย่างเดียว การพิจารณาว่าอนุกรมเวลาชุดใดชุดหนึ่งเป็นสแตชันนารีหรือไม่ พิจารณาได้ ดังนี้

1. ค่าเฉลี่ย ($E(Y_t)$) คงที่สำหรับทุกๆ ค่าของ t พิจารณาโดยแบ่งอนุกรมเวลาออกเป็นส่วนๆ แล้วหาค่าเฉลี่ยของอนุกรมเวลาแต่ละส่วน ถ้าค่าเฉลี่ยแต่ละส่วนย่อยไม่แตกต่างกันมากนัก จะสรุปได้ว่าค่าเฉลี่ยคงที่

2. ค่าความแปรปรวน ($V(Y_t)$) คงที่สำหรับทุกๆ ค่าของ t พิจารณาโดยแบ่งอนุกรมเวลาออกเป็นส่วนๆ แล้วหาค่าเบี่ยงเบนมาตรฐานของอนุกรมเวลาแต่ละส่วน ถ้าค่าเบี่ยงเบนมาตรฐานในแต่ละส่วนย่อยไม่แตกต่างกันมากนัก จะสรุปได้ว่าค่าความแปรปรวนคงที่

3. การพล็อตกราฟของอนุกรมเวลาเพื่อดูการเคลื่อนไหวของอนุกรมเวลา ถ้าการเคลื่อนไหวของอนุกรมเวลามีแนวโน้ม และ/หรือฤดูกาล แสดงว่าอนุกรมเวลาชุดนั้นขาดคุณสมบัติสเตชันนารี

4. พิจารณาออเรลโลแกรม (Correlogram) ของ r_k ซึ่งเป็นกราฟที่เกิดจากการพล็อตค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองที่ lag K (Coefficient autocorrelation at lag K : r_k) กับค่า K โดยที่ $K = L, 2L, 3L, \dots$ เมื่อ L เป็นจำนวนฤดูกาล ถ้าอนุกรมเวลาเป็นสเตชันนารี ค่า r_k จะมีลักษณะลดลงอย่างรวดเร็ว (Die down) เมื่อ K มีค่าเพิ่มขึ้น แต่ถ้าค่า r_k มีลักษณะลดลงค่อนข้างช้า เมื่อ K มีค่าเพิ่มขึ้น และ/หรือค่า r_k มีค่าค่อนข้างสูง แสดงว่าอนุกรมเวลาชุดนั้นขาดคุณสมบัติสเตชันนารี

สำหรับการกำหนดรูปแบบให้กับอนุกรมเวลาที่เป็นสเตชันนารี จะกำหนดรูปแบบในรูป ARMA(p,q) ซึ่งรูปแบบ ARMA(p,q) จะประกอบด้วยรูปแบบ AR(p) (Auto Regressive Order p) และรูปแบบ MA(q) (Moving Average Order q) สำหรับอนุกรมเวลาที่ไม่เป็น สเตชันนารี (Nonstationary Time Series) เป็นอนุกรมเวลาที่ค่าสังเกต มีคุณสมบัติทางสถิติไม่คงที่ โดยเปลี่ยนแปลงไปตามเวลาที่เปลี่ยนไป การกำหนดรูปแบบ ARMA(p,q) ได้ จะต้องแปลงอนุกรมเวลาให้เป็นอนุกรมเวลาใหม่ที่มีคุณสมบัติสเตชันนารีเสียก่อน

1.3.2 อนุกรมเวลาที่ไม่เป็นสเตชันนารี (Nonstationary Time Series)

เป็นอนุกรมเวลาที่ค่าสังเกต (Y_t) มีคุณสมบัติทางสถิติไม่คงที่ คือเปลี่ยนแปลงไปตามเวลาที่เปลี่ยนไป อนุกรมเวลาที่ไม่เป็นสเตชันนารี จะใช้รูปแบบ ARMA(p,q) ไม่ได้ ต้องแปลงอนุกรมเวลาดังกล่าว ให้เป็นอนุกรมเวลาใหม่ที่มีคุณสมบัติเป็นสเตชันนารีเสียก่อน แล้วจึงจะใช้รูปแบบ ARMA(p,q) ได้ การแปลงอนุกรมเวลาที่ไม่เป็นสเตชันนารีให้เป็นอนุกรมเวลาที่เป็นสเตชันนารี สามารถทำได้ด้วยวิธีการ ดังต่อไปนี้

ก. การหาผลต่างปกติ (Regular Differencing)

ถ้าอนุกรมเวลาเดิม (Y_t) เป็นอนุกรมเวลาที่มีแนวโน้ม สามารถแปลงให้เป็นอนุกรมเวลาที่ไม่มีแนวโน้ม (Z_t)

โดยที่ $Z_t = \nabla^d Y_t$ เมื่อ d เป็นลำดับของการหาผลต่าง เช่น

$$\text{เมื่อ } d=1, \quad Z_t = \nabla Y_t = Y_t - Y_{t-1}$$

$$\begin{aligned} \text{เมื่อ } d=2, \quad Z_t &= \nabla^2 Y_t = \nabla(Y_t - Y_{t-1}) = \nabla Y_t - \nabla Y_{t-1} \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2} \end{aligned}$$

ข. การหาผลต่างฤดูกาล (Seasonal Differencing)

ถ้าอนุกรมเวลาเดิม (Y_t) เป็นอนุกรมเวลาที่มีความผันแปรตามฤดูกาลเข้ามาเกี่ยวข้อง สามารถแปลงให้เป็นอนุกรมฤดูกาลที่ไม่มีความผันแปรตามฤดูกาล (Z_t) โดยที่ $Z_t = \nabla_L^D Y_t$ เมื่อ D เป็นลำดับของการหาผลต่างฤดูกาล และ L เป็นจำนวนฤดูกาลต่อปี เช่น $L = 4$ ถ้าอนุกรมเวลาเดิมเป็นอนุกรมเวลารายไตรมาส

$$\text{เมื่อ } D=1, \quad Z_t = \nabla_4 Y_t = Y_t - Y_{t-4}$$

$$\begin{aligned} \text{เมื่อ } D=2, \quad Z_t &= \nabla_4^2 Y_t = \nabla_4(Y_t - Y_{t-4}) = \nabla_4 Y_t - \nabla_4 Y_{t-4} \\ &= (Y_t - Y_{t-4}) - (Y_{t-4} - Y_{t-8}) = Y_t - 2Y_{t-4} + Y_{t-8} \end{aligned}$$

การหาผลต่างวิธีนี้จะกระทำกี่ครั้งก็ได้ จนกว่าจะได้อนุกรมเวลาที่เป็นสแตชันนารี

ค. การหาผลต่างปกติและผลต่างฤดูกาล

ถ้าอนุกรมเวลาเดิม (Y_t) เป็นอนุกรมเวลาที่มีทั้งแนวโน้ม และ ความผันแปรตามฤดูกาล การแปลงให้เป็นอนุกรมเวลาสแตชันนารี จะทำได้โดยหาผลต่างและผลต่างฤดูกาลควบคู่กันไป ค่า d และ D จะมีค่าเป็นเท่าไรนั้นขึ้นอยู่กับว่าอนุกรมเวลาที่แปลงแล้ว เป็นอนุกรมเวลาที่ สแตชันนารีแล้วหรือยัง เช่น อนุกรมเวลาเดิมเป็นอนุกรมเวลาที่มีทั้งแนวโน้มและมีอนุกรมเวลารายไตรมาส

จาก $Z_t = \nabla^d Y_t$ เมื่อ d เป็นลำดับของการหาผลต่าง

และ $Z_t = \nabla_L^D Y_t$ เมื่อ D เป็นลำดับของการหาผลต่างฤดูกาล และ L เป็นจำนวนฤดูกาลต่อปี

จะได้ $Z_t = \nabla^d \nabla_L^D Y_t$ เมื่อ d เป็นลำดับของการหาผลต่าง D เป็นลำดับของการหาผลต่างฤดูกาล และ L เป็นจำนวนฤดูกาลต่อปี

เมื่อ $d = 1, D = 1$ และ $L = 4$

$$\begin{aligned} Z_t &= \nabla^1 \nabla_4^1 Y_t = \nabla^1 (Y_t - Y_{t-4}) = \nabla^1 Y_t - \nabla^1 Y_{t-4} \\ &= (Y_t - Y_{t-1}) - (Y_{t-4} - Y_{t-5}) = Y_t - Y_{t-1} - Y_{t-4} + Y_{t-5} \end{aligned}$$

ง. การแปลงโดยใช้ลอการิทึม รากที่สอง หรือ กำลังสอง

การแปลงประเภทนี้มักใช้กับข้อมูลอนุกรมเวลาเดิม (Y_t) ที่มีความแปรปรวนไม่คงที่ นั่นคือ ความแปรปรวนไม่คงที่สำหรับค่า t ต่างๆ โดยที่สามารถแปลงอนุกรมเวลาได้รูปแบบต่างๆ เช่น $Z_t = \log(Y_t)$, $Z_t = \sqrt{Y_t}$ เป็นต้น

การพยากรณ์โดยใช้วิธีบ็อกซ์และเจนกินส์ ประกอบด้วย 4 ขั้นตอน ดังนี้

ขั้นที่ 1 การกำหนดรูปแบบ (Identification)

เป็นการหารูปแบบ ARMA(p,q) ที่เหมาะสมให้กับอนุกรมเวลา โดยเปรียบเทียบ
คอเรลโลแกรมของค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง และค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนใน
ตัวเองของตัวอย่าง (r_k และ r_{kk}) กับค่าคอเรลโลแกรมของค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง
และค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนในตัวเองของประชากร (ρ_k และ ρ_{kk}) ดังนั้นจะต้องมีการ
คำนวณหาค่า r_k และ r_{kk} สำหรับ k หลายค่า สำหรับอนุกรมเวลาฤดูกาลควรหาค่า r_k และ r_{kk}
สำหรับค่า k ที่เป็นผลคูณของจำนวนฤดูกาลต่อปี

r_k เป็นค่าประมาณของ ρ_k ซึ่ง

$$r_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad \text{สำหรับ } k = 1, 2, 3, \dots$$

โดยที่ Y_t คือ ข้อมูลหรือค่าสังเกต ณ เวลา t

k คือ จำนวนช่วงเวลาที่ข้อมูลอยู่ห่างกัน, $k = 1, 2, 3, \dots$

n คือ จำนวนข้อมูลทั้งหมด

\bar{Y} คือ ค่าเฉลี่ยของข้อมูลทั้งหมด

และความแปรปรวนของ r_k หาได้จาก

$$v(r_k) = \frac{1}{n} \left(1 + 2 \sum_{j=1}^{k-1} r_j^2 \right) \quad \text{สำหรับ } k = 1, 2, 3, \dots$$

r_k เป็นค่าวัดสหสัมพันธ์ระหว่างค่าสังเกตในอนุกรมเวลาที่อยู่ห่างกัน k ช่วงเวลา หรือค่าวัดสหสัมพันธ์ระหว่าง Y_t และ Y_{t+k}

คุณสมบัติของ r_k มีดังนี้

1. r_k มีค่าอยู่ระหว่าง -1 ถึง 1 เมื่อขนาดของ r_k ที่วัดด้วย $|r_k|$ มีค่าเข้าใกล้ 0 แสดงว่ามีความสัมพันธ์กันต่ำ ถ้า r_k มีค่าน้อยกว่า 0 แสดงว่ามีความสัมพันธ์กันในทิศทางตรงกันข้าม และถ้า r_k มีค่ามากกว่า 0 แสดงว่ามีความสัมพันธ์กันในทิศทางเดียวกัน

2. $r_k = r_{-k}$ คือการวัดความสัมพันธ์ระหว่าง Y_t และ Y_{t-k} หรือระหว่าง Y_t และ Y_{t+k} ซึ่งต่างก็เป็นค่าสังเกตที่อยู่ห่างกัน k ค่าไม่แตกต่างกัน

r_{kk} เป็นค่าประมาณของ ρ_{kk} ซึ่ง

$$r_{kk} = \begin{cases} r_1 & k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & k = 2, 3, \dots \end{cases}$$

เมื่อ $r_{k,j}$ คือ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนที่ช่วงเวลาห่างกัน k และ j โดยที่

$$r_{k,j} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \quad , \quad j = 1, 2, 3, \dots, k-1$$

และความแปรปรวนของ r_{kk} หาได้จาก

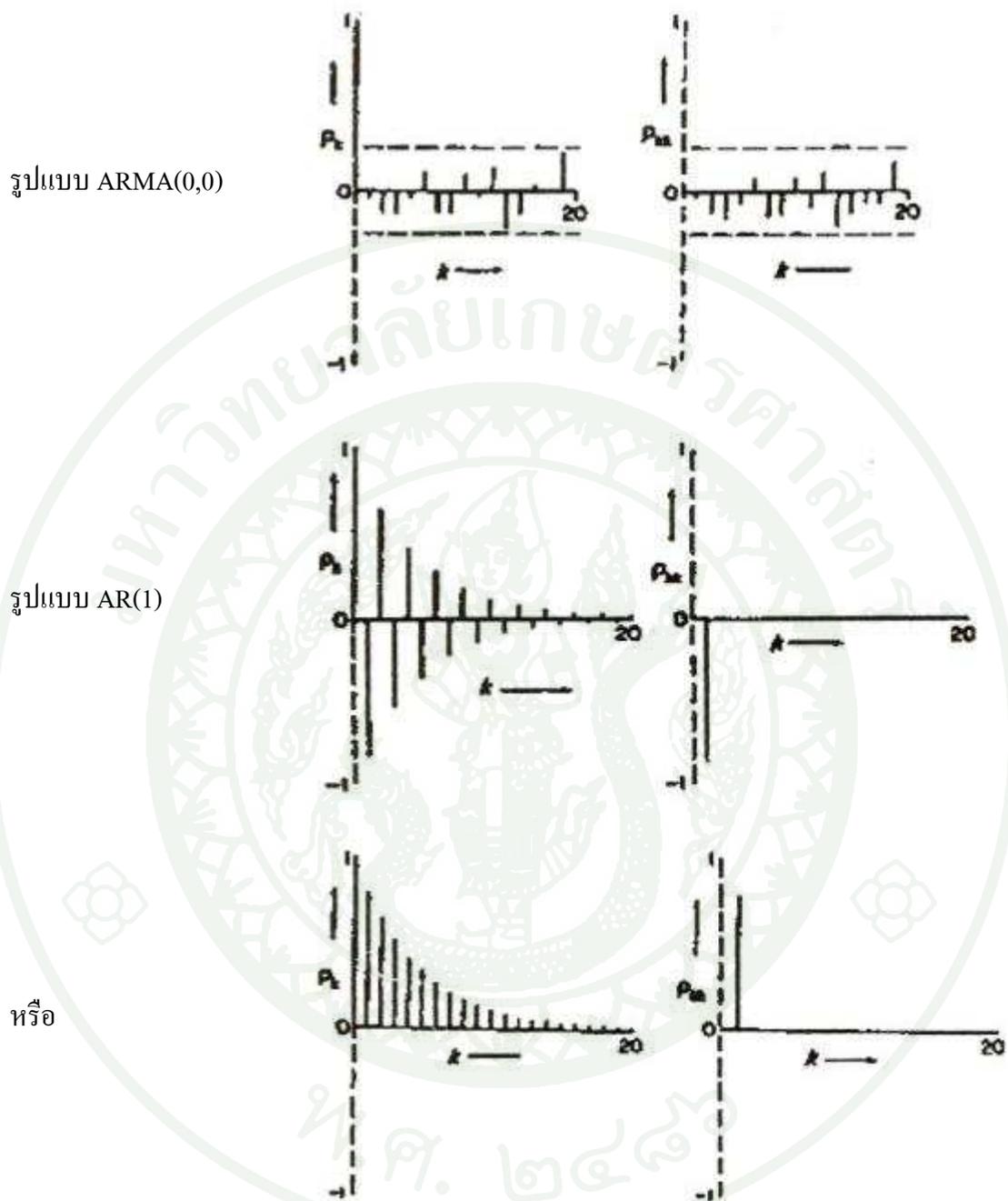
$$v(r_{kk}) = \frac{1}{n} \quad , \quad k = 1, 2, \dots$$

การหารูปแบบ ARMA(p,q) ให้กับข้อมูลอนุกรมเวลา โดยพิจารณาเปรียบเทียบคอเรลโลแกรมระหว่าง r_k กับ ρ_k และ คอเรลโลแกรมของ r_{kk} กับ ρ_{kk} สำหรับ ARMA(p,q) แต่ละรูปแบบจะมีลักษณะที่แตกต่างกัน ดังแสดงในตารางที่ 1 และภาพที่ 10

ตารางที่ 1 ลักษณะของ ρ_k และ ρ_{kk} สำหรับรูปแบบ ARMA(p,q)

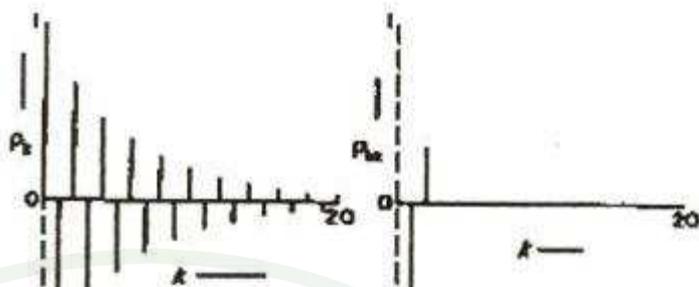
รูปแบบ	ρ_k	ρ_{kk}
ARMA(0,0)	$\rho_k = 0$ สำหรับ $k = 1, 2, \dots$	$\rho_{kk} = 0$ สำหรับ $k = 1, 2, \dots$
AR(1)	ρ_k มีค่าเข้าใกล้ 0 เมื่อ k มีค่ามากขึ้น	$\rho_{kk} \begin{cases} \neq 0 & \text{สำหรับ } k = 1 \\ = 0 & \text{สำหรับ } k \geq 2 \end{cases}$
AR(2)	ρ_k มีค่าเข้าใกล้ 0 เมื่อ k มีค่ามากขึ้น	$\rho_{kk} \begin{cases} \neq 0 & \text{สำหรับ } k = 1, 2 \\ = 0 & \text{สำหรับ } k \geq 3 \end{cases}$
MA(1)	$\rho_k \begin{cases} \neq 0 & \text{สำหรับ } k = 1 \\ = 0 & \text{สำหรับ } k \geq 2 \end{cases}$	ρ_{kk} มีค่าเข้าใกล้ 0 เมื่อ k มีค่ามากขึ้น
MA(2)	$\rho_k \begin{cases} \neq 0 & \text{สำหรับ } k = 1, 2 \\ = 0 & \text{สำหรับ } k \geq 3 \end{cases}$	ρ_{kk} มีค่าเข้าใกล้ 0 เมื่อ k มีค่ามากขึ้น
ARMA(1,1)	ρ_k มีค่าเข้าใกล้ 0 เมื่อ k มีค่ามากขึ้น	ρ_{kk} มีค่าเข้าใกล้ 0 เมื่อ k มีค่ามากขึ้น

ที่มา: สมเกียรติ (2548)



ภาพที่ 10 กราฟแสดงค่า ρ_k และ ρ_{kk} สำหรับตัวแบบ ARMA(p,q) ต่างๆ

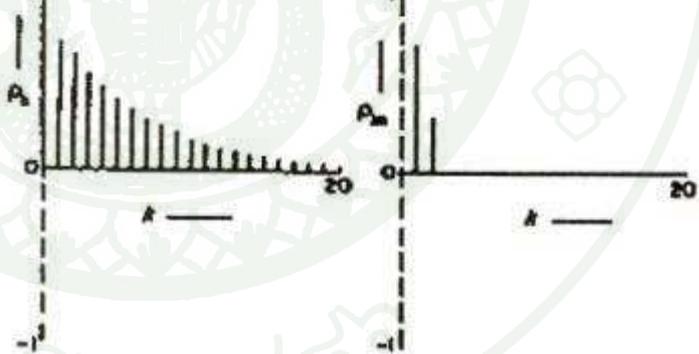
รูปแบบ AR(2)



หรือ



หรือ



ภาพที่ 10 (ต่อ)

รูปแบบ MA(1)

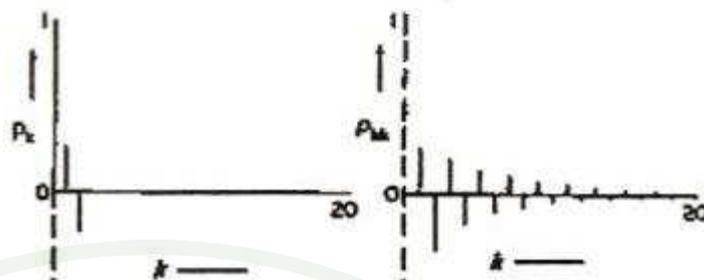
หรือ

หรือ



ภาพที่ 10 (ต่อ)

รูปแบบ MA(2)



หรือ

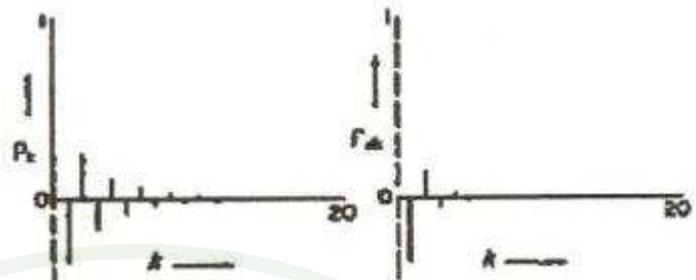


หรือ



ภาพที่ 10 (ต่อ)

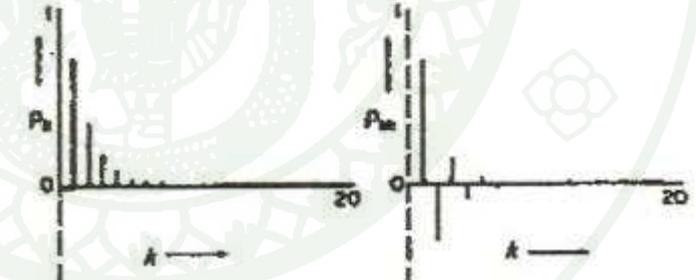
รูปแบบ ARMA(1,1)



หรือ



หรือ

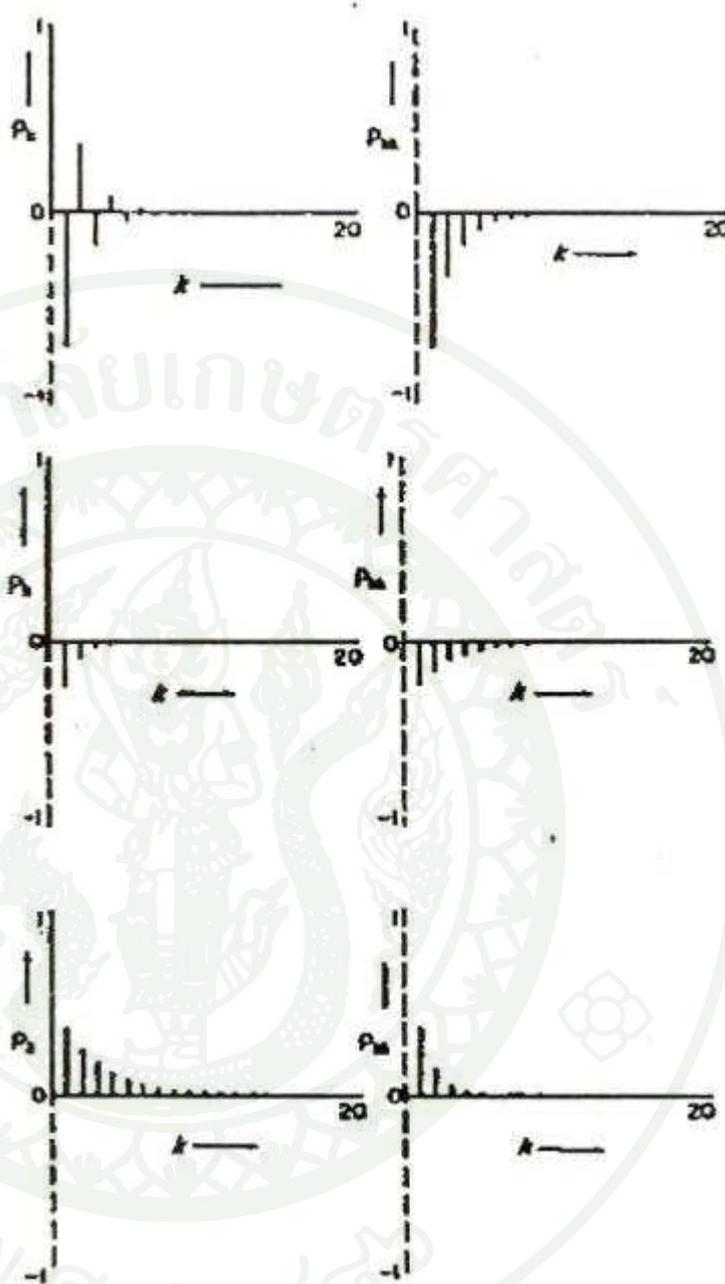


ภาพที่ 10 (ต่อ)

รูปแบบ ARMA(1,1)

หรือ

หรือ



ภาพที่ 10 (ต่อ)

ที่มา: จันทรรักษ์ (2548)

รูปแบบ ARMA(p,q) ของอนุกรมเวลาที่ใช้ในการพยากรณ์โดยวิธีบ็อกซ์และเจนกินส์
เมื่ออนุกรมเวลามีคุณสมบัติเป็นสเตชันนารี มีดังนี้

1. รูปแบบการถดถอยในตัวเองอันดับ p (Autoregressive Model of Order p : AR(p))
มีรูปแบบทั่วไปดังนี้

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

โดยที่ Y_t คือ ค่าสังเกตของอนุกรมเวลา ณ เวลา t

δ คือ ค่าคงที่

ϕ_i คือ ค่าพารามิเตอร์ของการถดถอยในตัวเองที่ i

ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t มีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ย
เท่ากับ 0 และความแปรปรวนเท่ากับ δ^2 และมีการแจกแจงที่เป็นอิสระกัน

จาก $Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$

หรือ $Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \dots - \phi_p Y_{t-p} = \delta + \varepsilon_t$

สามารถเขียนในรูปแบบของ Backward Operator ได้เป็น

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = \delta + \varepsilon_t$$

B เรียกว่าตัวถอยหลัง (Backward Shift หรือ Backshift) โดยที่

$$B^p Y_t = Y_{t-p}$$

รูปแบบการถดถอยในตัวเองอันดับ p (AR(p)) ที่นิยมใช้กันเมื่อ $p = 1$ และ $p = 2$ ซึ่งมีรูปแบบดังนี้

1.1 รูปแบบถดถอยในตัวเองอันดับที่ 1 (AR(1)) มีรูปแบบดังนี้

$$Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t$$

หรือ

$$Y_t - \phi_1 Y_{t-1} = \delta + \varepsilon_t$$

หรือ

$$(1 - \phi_1 B) Y_t = \delta + \varepsilon_t$$

มี $|\phi_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลามีคุณสมบัติเป็นสเตชันนารี

1.2 รูปแบบถดถอยในตัวเองอันดับที่ 2 มีรูปแบบดังนี้

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

หรือ

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} = \delta + \varepsilon_t$$

หรือ

$$(1 - \phi_1 B - \phi_2 B^2) Y_t = \delta + \varepsilon_t$$

มี $\phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1, |\phi_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลามีคุณสมบัติเป็นสเตชันนารี

2. รูปแบบค่าเฉลี่ยเคลื่อนที่อันดับ q (Moving Average Model of Order q : MA(q))
มีรูปแบบทั่วไปดังนี้

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

โดยที่ Y_t คือ ค่าสังเกตของอนุกรมเวลา ณ เวลา t

μ คือ ค่าคงที่

θ_i คือ ค่าพารามิเตอร์ของค่าเฉลี่ยเคลื่อนที่ที่ i

ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t ซึ่ง ε_t มีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ σ^2 และมีการแจกแจงที่เป็นอิสระกัน

จาก
$$Y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}$$

สามารถเขียนในรูปแบบของ Backward Operator ได้เป็น

$$Y_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t$$

เมื่อ

$$B^q \varepsilon_t = \varepsilon_{t-q}$$

รูปแบบค่าเฉลี่ยเคลื่อนที่อันดับ q MA(q) ที่นิยมใช้กันเมื่อ $q = 1, q = 2$ ซึ่งมีรูปแบบดังนี้

2.1 รูปแบบค่าเฉลี่ยเคลื่อนที่อันดับ 1 (MA(1)) มีรูปแบบ ดังนี้

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

หรือ
$$Y_t = \mu + (1 - \theta_1 B) \varepsilon_t$$

มี $|\theta_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลามีคุณสมบัติเป็นอินเวอร์ติเบิล (Invertible) ซึ่งคุณสมบัตินี้หมายถึง การที่สามารถเขียน ε_t ในเทอมของ $Y_t, Y_{t-1}, Y_{t-2}, \dots$ ได้ และหาค่าประมาณของค่าความคลาดเคลื่อน ε_t ได้

2.2 รูปแบบค่าเฉลี่ยเคลื่อนที่อันดับ 2 (MA(2)) มีรูปแบบ ดังนี้

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

$$\text{หรือ } Y_t = \mu + (1 - \theta_1 B - \theta_2 B^2) \varepsilon_t$$

มี $\theta_1 + \theta_2 < 1$, $\theta_2 - \theta_1 < 1$ และ $|\theta_2| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลา มีคุณสมบัติเป็นอินเวอร์ติเบิล

3. รูปแบบผสมการถดถอยในตัวเองและค่าเฉลี่ยเคลื่อนที่อันดับ p และ q (Mixed Autoregressive and Moving Average Model of Order p and q : ARMA(p, q))

มีรูปแบบทั่วไป ดังนี้

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

ARMA(p, q) ที่มีค่า p และ q เท่ากับ 1 หรือ ARMA(1,1) มีรูปแบบ ดังนี้

$$Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

มี $|\phi_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลา มีคุณสมบัติเป็นสเตชันนารี และ $|\theta_1| < 1$ เป็นเงื่อนไขที่ระบุให้อนุกรมเวลา มีคุณสมบัติเป็นอินเวอร์ติเบิล

4. รูปแบบอินทิเกรตค่าเฉลี่ยเคลื่อนที่และออโตรีเกรสซีฟ (Autoregressive Integrated Moving Average): ARIMA(p, d, q)

มีรูปแบบทั่วไป ดังนี้

$$W_t = \delta + \phi_1 W_{t-1} + \phi_2 W_{t-2} + \dots + \phi_p W_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

โดยที่ W_t คือ ค่าสังเกตของอนุกรมเวลา ณ เวลา t ซึ่ง $W_t = \nabla^d Y_t$

δ คือ ค่าคงที่

ϕ_i คือ ค่าพารามิเตอร์ของการถดถอยในตัวเองที่ i

θ_i คือ พารามิเตอร์ของค่าเฉลี่ยเคลื่อนที่ที่ i

ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t

p เป็น อันดับที่ของการถดถอยในตัวเอง

d คือ อันดับที่ของผลต่าง

q คือ อันดับที่ของค่าเฉลี่ยเคลื่อนที่

ตัวอย่างของรูปแบบ ARIMA(p,q) เช่น

ARIMA(0,1,1) หรือ IMA(1,1) มีรูปแบบ

$$Y_t - Y_{t-1} = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1}, \quad \text{เมื่อ } |\theta_1| < 1$$

ARIMA(1,1,0) หรือ ARI(1,1) มีรูปแบบ

$$(Y_t - Y_{t-1}) - \phi_1(Y_{t-1} - Y_{t-2}) = \delta + \varepsilon_t, \quad \text{เมื่อ } |\phi_1| < 1$$

ARIMA(1,1,1) มีรูปแบบ

$$(Y_t - Y_{t-1}) - \phi_1(Y_{t-1} - Y_{t-2}) = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1}, \quad \text{เมื่อ } |\phi_1| < 1 \text{ และ } |\theta_1| < 1$$

ARIMA(0,1,0) หรือ รูปแบบ Random Walk มีรูปแบบ

$$Y_t - Y_{t-1} = \varepsilon_t$$

5. รูปแบบอินทิเกรตค่าเฉลี่ยเคลื่อนที่และการถดถอยในตัวเอง เมื่อมีอิทธิพลฤดูกาล (Seasonal Autoregressive Integrated Moving Average): SARIMA (P,D,Q)_L โดยที่ D เป็นอันดับที่ของผลต่าง และ L เป็นช่วงการเกิดฤดูกาล

ถ้าอนุกรมเวลาที่พิจารณามีทั้งแนวโน้ม และอิทธิพลฤดูกาลเข้ามาเกี่ยวข้อง หาโดยรูปแบบ ARIMA(p,d,q) (P,D,Q)_L ซึ่งแสดงความสัมพันธ์ของอนุกรมเวลาที่อยู่ในช่วงเวลาที่ห่างกัน L ช่วงเวลา อาจเป็น เดือน หรือ ปี เช่น อนุกรมเวลา มีรูปแบบ ARIMA (1,1,0) (0,1,2)₁₂ เขียนได้ดังนี้

$$W_t = \phi_1 W_{t-1} + \varepsilon_t - \theta_{12}\varepsilon_{t-12} - \theta_{24}\varepsilon_{t-24}$$

โดยที่ W_t เป็นผลต่างของค่าสังเกตที่อยู่ห่างกัน 12 เดือน ซึ่ง

$$W_t = \Delta_{12} Y_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$$

เกณฑ์ที่ใช้ในการพิจารณาเลือกรูปแบบเมื่อมีรูปแบบที่เหมาะสมมากกว่าหนึ่งรูปแบบ

โดยทั่วไปจะมีความเป็นไปได้ที่ว่าอนุกรมเวลาหนึ่งชุดใดๆ อาจมีรูปแบบที่เหมาะสมมากกว่าหนึ่งรูปแบบ การพิจารณาการเลือกใช้เพียงหนึ่งตัวแบบจากที่เหมาะสมทั้งหมด ในกรณีนี้จะเลือกรูปแบบที่ให้ค่า AIC (เงื่อนไขของ Akaike), ค่า SBC (เงื่อนไขของ Schwartz) (มุกดา, 2549: 237) และ ค่า SSE มีค่าน้อยที่สุด

เงื่อนไขของ Schwartz (SBC)

เงื่อนไขของ Schwartz (SBC) กำหนดในรูป

$$SBC(M) = n \ln \sigma_a^2 + M \ln(n)$$

โดยที่ σ_a^2 เป็นตัวประมาณโดยวิธีแมกซิมัมไลค์ลิตู๊ดของ σ_a^2

M เป็นจำนวนพารามิเตอร์ในตัวแบบที่ทำให้ $AIC(M)$ มีค่าน้อยที่สุด

n เป็นจำนวนรีซิดวอลทั้งหมดที่ได้จากรูปแบบที่พิจารณา

เงื่อนไขของ Akaike (AIC)

ภายใต้ข้อสมมติว่าได้กำหนดรูปแบบที่มีพารามิเตอร์เป็นจำนวน M ค่า AIC (Akaike's Information Criterion) กำหนดในรูป

$$AIC(M) = -2\ln(\text{maximum likelihood}) + 2M$$

ในกรณีที่รูปแบบเป็น ARMA(p,q) จะได้

$$AIC(M) = n \ln \sigma_a^2 + M \ln(n)$$

โดยที่ σ_a^2 เป็นตัวประมาณโดยวิธีแมกซิมัมไลค์ลิฮูดของ σ_a^2

M เป็นจำนวนพารามิเตอร์ในตัวแบบที่ทำให้ $AIC(M)$ มีค่าน้อยที่สุด

ขั้นที่ 2 การประมาณค่าพารามิเตอร์ (Parameter Estimation)

การประมาณค่าพารามิเตอร์ในรูปแบบโดยวิธีกำลังสองน้อยที่สุด ซึ่งจะใช้เทคนิคการวิเคราะห์ตัวเลข โดยมีการกำหนดค่าประมาณเบื้องต้น (Initial Estimates) เพื่อหาค่าประมาณสุดท้าย (Final Estimates) ค่าประมาณสุดท้ายจะได้อาจจากการทำซ้ำ (Iteration) หลายๆ รอบ โดยเป็นค่าประมาณที่ทำให้ผลรวมกำลังสองของความคลาดเคลื่อนมีค่าต่ำสุด ค่าประมาณเบื้องต้นจะได้อาจจากการแก้สมการที่แสดงความสัมพันธ์ของ p_k ที่แทนด้วย r_k และพารามิเตอร์ที่แทนด้วยค่าประมาณพารามิเตอร์ ซึ่งแสดงในตารางที่ 2 โดยจำนวนสมการจะเท่ากับจำนวนพารามิเตอร์

ตารางที่ 2 ความสัมพันธ์ของ ρ_k กับพารามิเตอร์ในรูปแบบต่างๆ

รูปแบบ	พารามิเตอร์	ความสัมพันธ์	ขอบเขตของพารามิเตอร์
AR(1)	ϕ_1	$\rho_1 = \phi_1$	$ \phi_1 < 1$
AR(2)	ϕ_1, ϕ_2	$\rho_1 = \frac{\phi_1}{1 - \phi_2}$ $\rho_2 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2$	$\phi_1 + \phi_2 < 1$ $\phi_1 - \phi_2 < 1, \phi_1 < 1$
MA(1)	θ_1	$\rho_1 = \frac{-\theta_1}{1 - \theta_1^2}$	$ \theta_1 < 1$
MA(2)	θ_1, θ_2	$\rho_1 = \frac{-\theta_1(1 - \theta_1)}{1 + \theta_1^2 + \theta_2^2}$ $\rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}$	$\theta_2 - \theta_1 < 1$ $\theta_1 + \theta_2 < 1$ $ \theta_2 < 1$
ARMA(1,1)	θ_1, ϕ_1	$\rho_1 = \frac{(1 - \theta_1\phi_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1}$ $\rho_2 = \phi_1\rho_1$	$ \phi_1 < 1$ $ \theta_1 < 1$

ที่มา: สมเกียรติ (2548)

ขั้นที่ 3 การตรวจสอบความเหมาะสมของรูปแบบ (Diagnostic Checking)

การตรวจสอบว่ารูปแบบที่เลือกไว้มีความเหมาะสมกับอนุกรมเวลาหรือไม่ มีหลายวิธี วิธีการตรวจสอบส่วนใหญ่จะอาศัยค่าความคลาดเคลื่อนของการพยากรณ์ (e_t) ซึ่งเป็นค่าประมาณของ ε_t ในรูปแบบ โดยที่ $e_t = Y_t - \hat{Y}_t$ วิธีการตรวจสอบที่ใช้ในงานวิจัยนี้จะใช้การตรวจสอบ 3 วิธี ดังนี้

1. การทดสอบว่าพารามิเตอร์ (θ) ในรูปแบบมีค่าเท่ากับ 0 หรือไม่ โดยกำหนดสมมติฐานหลัก และ สมมติฐานทางเลือก ดังนี้

$$H_0 : \theta = 0$$

$$H_1 : \theta \neq 0$$

สำหรับสถิติทดสอบที่ใช้คือ t โดยที่

$$t = \frac{\hat{\theta}}{S_{\hat{\theta}}}$$

เมื่อ $\hat{\theta}$ คือ ค่าประมาณพารามิเตอร์ θ

$S_{\hat{\theta}}$ คือ ค่าความคลาดเคลื่อนมาตรฐานของ $\hat{\theta}$

2. การตรวจสอบข้อจำกัดของรูปแบบที่กำหนดว่า ε_t จะต้องมีการแจกแจงที่เป็นอิสระกัน โดยพิจารณาว่าค่าความคลาดเคลื่อนมีลักษณะการเคลื่อนไหวที่เป็นอิสระกันหรือไม่ หากพบว่าอนุกรมเวลาของค่าความคลาดเคลื่อนเป็นอิสระกัน จะสรุปได้ว่ารูปแบบที่กำหนดให้กับอนุกรมเวลานั้นมีความเหมาะสม การทดสอบดังกล่าวกำหนดสมมติฐานหลัก และ สมมติฐานทางเลือก ดังนี้

$$H_0 : \rho_k = 0$$

$$H_1 : \rho_k \neq 0 \quad \text{สำหรับ } k = 1, 2, \dots$$

การทดสอบจะใช้ r_k ของอนุกรมเวลาค่าความคลาดเคลื่อน

$$r_k = \frac{\sum_{t=1}^{n-k} (e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2}$$

โดยที่

$$\bar{e} = \frac{\sum_{t=1}^n e_t}{n}$$

เมื่อ \bar{e} เป็นค่าเฉลี่ยของ e_t และ n เป็นขนาดของอนุกรมเวลา จะยอมรับ H_0 เมื่อค่า r_k อยู่ในช่วง ± 2 เท่าของส่วนเบี่ยงเบนมาตรฐานของ r_k กำหนดระดับนัยสำคัญ 0.05 แสดงว่ารูปแบบที่กำหนดให้กับอนุกรมเวลามีความเหมาะสม

3. การตรวจสอบว่าความคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน 1, 2, 3, ..., m ช่วงเวลาเป็นอิสระกันหรือไม่ พิจารณา $\rho_k = 0$ สำหรับ $k=1, 2, 3, \dots, m$ หรือไม่ กำหนดสมมติฐานหลักและสมมติฐานทางเลือก ดังนี้

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_m = 0 \quad \text{สำหรับ } k = 1, 2, 3, \dots, m$$

$$H_1 : \rho_k \text{ อย่างน้อยหนึ่งค่าไม่เท่ากับ } 0 \quad \text{สำหรับ } k = 1, 2, 3, \dots, m$$

สถิติทดสอบที่ใช้ คือ สถิติทดสอบของ Box และ Ljung มีรูปแบบ ดังนี้

$$Q_m = n(n+2) \sum_{k=1}^m \left\{ \frac{r_k^2}{n-k} \right\}$$

โดยที่ n คือ ขนาดของอนุกรมเวลา e_t

m คือ ช่วงเวลาที่ห่างกันมากที่สุดของ e_t ที่นำมาพิจารณา

ค่าวิกฤติของการทดสอบคือ $\chi_{\alpha, m-a}^2$ (เมื่อ a คือ จำนวนพารามิเตอร์ทั้งหมดรวมค่าคงที่ (μ, δ)) ด้วย จะสรุปการทดสอบว่า ปฏิเสธ H_0 เมื่อ Q_m ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าวิกฤติเป็นต้นไป และสรุปว่า ρ_k อย่างน้อยหนึ่งค่าไม่เท่ากับ 0 สำหรับ $k = 1, 2, 3, \dots, m$ นั่นคือค่าความคลาดเคลื่อนไม่เป็นอิสระต่อกัน หรือรูปแบบที่กำหนดให้กับอนุกรมเวลายังไม่เหมาะสม

ขั้นที่ 4 การพยากรณ์ (Forecasting)

คือขั้นตอนการนำสมการที่สร้างจากรูปแบบไปใช้ในการพยากรณ์ วิธีของบ็อกซ์และ เจนกินส์ จะให้สมการพยากรณ์ที่จะใช้ในการหาค่าพยากรณ์ล่วงหน้าได้หลายช่วงเวลา สมการพยากรณ์จะสร้างจากรูปแบบที่ได้ทดสอบแล้วว่าเหมาะสมกับการพยากรณ์ โดยส่วนใหญ่ค่าพยากรณ์ที่ได้จากสมการพยากรณ์ หากเป็นการพยากรณ์ในระยะยาว ค่าพยากรณ์นั้นจะใช้สาระจากข้อมูลจริงน้อยลง และความแม่นยำที่ได้จากการพยากรณ์ก็จะมีค่าลดลง สมการพยากรณ์ล่วงหน้า m ช่วงเวลาที่ได้จากรูปแบบ ARMA(p,q) ต่างๆ มีดังนี้

1. รูปแบบ ARMA(0,0)

$$\hat{Y}_{t+m} = \hat{\delta} = \bar{Y} \quad \text{เมื่อ } m \geq 1$$

2. รูปแบบ AR(1)

$$\hat{Y}_{t+m} = \begin{cases} \hat{\delta} + \hat{\phi}_1 Y_t & \text{เมื่อ } m = 1 \\ \hat{\delta} + \hat{\phi}_1 Y_{t+m-1} & \text{เมื่อ } m \geq 2 \end{cases}$$

3. รูปแบบ AR(2)

$$\hat{Y}_{t+m} = \begin{cases} \hat{\delta} + \hat{\phi}_1 Y_t + \hat{\phi}_2 Y_{t-1} & \text{เมื่อ } m = 1 \\ \hat{\delta} + \hat{\phi}_1 Y_{t+1} + \hat{\phi}_2 Y_t & \text{เมื่อ } m = 2 \\ \hat{\delta} + \hat{\phi}_1 Y_{t+m-1} + \hat{\phi}_2 Y_{t+m-2} & \text{เมื่อ } m \geq 3 \end{cases}$$

4. รูปแบบ MA(1)

$$\hat{Y}_{t+m} = \begin{cases} \hat{\mu} - \hat{\theta}_1 e_t & \text{เมื่อ } m = 1 \\ \hat{\mu} & \text{เมื่อ } m \geq 2 \end{cases}$$

5. รูปแบบ MA(2)

$$\hat{Y}_{t+m} = \begin{cases} \hat{\mu} - \hat{\theta}_1 e_t - \hat{\theta}_2 e_{t-1} & \text{เมื่อ } m = 1 \\ \hat{\mu} - \hat{\theta}_2 e_t & \text{เมื่อ } m = 2 \\ \hat{\mu} & \text{เมื่อ } m \geq 3 \end{cases}$$

6. รูปแบบ ARMA(1,1)

$$\hat{Y}_{t+m} = \begin{cases} \hat{\delta} + \hat{\phi}_1 Y_t - \hat{\theta}_1 e_t & \text{เมื่อ } m = 1 \\ \hat{\delta} + \hat{\phi}_1 Y_{t+m-1} & \text{เมื่อ } m \geq 2 \end{cases}$$

7. รูปแบบ ARIMA(0,1,0)

$$\hat{Y}_{t+m} = \begin{cases} \hat{\delta} + Y_t & \text{เมื่อ } m = 1 \\ \hat{\delta} + Y_{t+m-1} & \text{เมื่อ } m \geq 2 \end{cases}$$

2. วิธีการพยากรณ์ร่วม (Combination Forecasting Method)

วิธีการพยากรณ์ร่วมเป็นวิธีการพยากรณ์ที่ใช้ค่าพยากรณ์จากวิธีการพยากรณ์เดี่ยวหลายวิธี โดยการนำค่าพยากรณ์ที่ได้จากแต่ละวิธีการพยากรณ์เดี่ยวมาหาค่าพยากรณ์ใหม่ โดยข้อดีของวิธีการพยากรณ์ร่วมคือให้ค่าคลาดเคลื่อนต่ำกว่าวิธีการพยากรณ์เดี่ยวทุกวิธี

สำหรับงานวิจัยนี้เลือกวิธีเฉลี่ยถ่วงน้ำหนักซึ่งน้ำหนักผกผันกับค่าผลบวกกำลังสองของค่าความคลาดเคลื่อน

วิธีนี้เสนอโดย Delurgio (1998) มีหลักการว่าความถูกต้องของการพยากรณ์จะสูง ถ้าความคลาดเคลื่อนของการพยากรณ์ต่ำ และในทางกลับกันความถูกต้องของการพยากรณ์จะต่ำ ถ้าความคลาดเคลื่อนของการพยากรณ์สูง วิธีเฉลี่ยถ่วงน้ำหนักจึงใช้ค่าน้ำหนักของวิธีที่ i ซึ่งแปรผกผันกับกับค่า SSE_i นั่นคือความถูกต้องของวิธีที่ i จะแปรผกผันกับค่า SSE_i และเพื่อให้ w_i มี $\sum w_i = 1$ ดังนั้น

$$w_i = \frac{1/SSE_i}{1/SSE_1 + 1/SSE_2 + \dots + 1/SSE_p}$$

ซึ่ง

$$SSE_i = \sum_{t=1}^n e_{it}^2, \quad i = 1, 2, \dots, p$$

ในการคำนวณ SSE ต้องกำหนดก่อนว่าจะใช้จำนวนค่าความคลาดเคลื่อนเท่าใด ถ้าหากใช้ค่าสังเกตเป็นจำนวนมากๆ ความแม่นยำในการหาค่าพยากรณ์จะลดน้อยลง เนื่องจากค่าพยากรณ์สร้างจากอนุกรมเวลาที่มีการเก็บรวบรวมมานานแล้ว

รูปแบบที่ใช้ในการพยากรณ์ร่วม (Combined Forecast: CF) มีรูปแบบคือ

$$\hat{Y}_{CF} = w_1 \hat{Y}^{(1)} + w_2 \hat{Y}^{(2)} + \dots + w_p \hat{Y}^{(p)}$$

โดยที่ w_i คือ ตัวถ่วงน้ำหนักของวิธีพยากรณ์; $i=1, 2, \dots, p$

\hat{Y}_{CF} คือ ค่าพยากรณ์

p คือ วิธีการพยากรณ์ที่ p

t คือ คาบเวลา; $t=1, 2, 3, \dots, n$

เกณฑ์ที่ใช้ในการเปรียบเทียบความสามารถของวิธีการพยากรณ์

การตัดสินใจว่าวิธีการพยากรณ์ใดเหมาะสมและมีความถูกต้องในการพยากรณ์มากที่สุด สามารถพิจารณาได้จากค่าความคลาดเคลื่อนระหว่างค่าจริงกับค่าพยากรณ์ที่ได้ ณ เวลา t ซึ่งค่าวัดความถูกต้องของการพยากรณ์ที่นิยมใช้มีอยู่หลายวิธี ได้แก่ ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (Mean Absolute Deviation : MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (Mean Absolute Percent Error : MAPE)

1. ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (Absolute Deviation : MAD) เป็นการวัดความแม่นยำที่วัดจากค่าความคลาดเคลื่อนของการพยากรณ์ โดยไม่คำนึงถึงทิศทางของความคลาดเคลื่อน MAD จะมีหน่วยวัดเหมือนกับค่าสังเกต (สมเกียรติ, 2548: 8)

$$MAD = \frac{\sum_{i=1}^n |e_i|}{n}$$

โดยที่ e_i คือ ค่าความคลาดเคลื่อนของการพยากรณ์ชุดที่ i

n คือ จำนวนข้อมูลทั้งหมด

ในการนำค่า MAD มาเป็นเกณฑ์ในการเปรียบเทียบหาวิธีการพยากรณ์ที่เหมาะสม สำหรับอนุกรมเวลาแต่ละชุดนั้น จะสรุปว่าแต่ละขนาดของอนุกรมเวลาและแต่ละช่วงการพยากรณ์ วิธีการที่เหมาะสมจะเป็นวิธีการพยากรณ์ที่ให้ค่า MAD ที่ต่ำที่สุด

2. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) เป็นการวัดความแม่นยำจากค่าเฉลี่ยผลรวมกำลังสองของค่าความคลาดเคลื่อนของการพยากรณ์ (Sum Square Error : SSE) (สมเกียรติ, 2548: 8)

โดยที่
$$SSE = \sum_{i=1}^n e_i^2$$

และ
$$MSE = \frac{\sum_{i=1}^n e_i^2}{n}$$

โดยที่ e_i คือ ค่าความคลาดเคลื่อนของการพยากรณ์ชุดที่ i

n คือ จำนวนข้อมูลทั้งหมด

ในการนำค่า MSE มาเป็นเกณฑ์ในการเปรียบเทียบหาวิธีการพยากรณ์ที่เหมาะสมสำหรับอนุกรมเวลาแต่ละชุดนั้น จะสรุปว่าแต่ละขนาดของอนุกรมเวลาและแต่ละช่วงการพยากรณ์วิธีการที่เหมาะสมจะเป็นวิธีการพยากรณ์ที่ให้ค่า MSE ที่ต่ำที่สุด

3. ค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (Mean Absolute Percent Error : MAPE) เนื่องจากเป็นค่าวัดความแม่นยำที่ไม่มีหน่วย จึงเหมาะสำหรับการเปรียบเทียบอนุกรมเวลาหลายชุดเมื่อใช้วิธีการพยากรณ์เดียวกัน หรือเปรียบเทียบวิธีการพยากรณ์หลายวิธีเมื่อใช้อนุกรมเวลาชุดเดียวกัน (สมเกียรติ, 2548: 9)

$$MAPE = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_i| / Y_i \times 100}{n}$$

ในการนำค่า MAPE มาเป็นเกณฑ์ในการเปรียบเทียบหาวิธีการพยากรณ์ที่เหมาะสมสำหรับอนุกรมเวลาแต่ละชุดนั้น จะสรุปว่าแต่ละขนาดของอนุกรมเวลาและแต่ละช่วงการพยากรณ์วิธีการที่เหมาะสมจะเป็นวิธีการพยากรณ์ที่ให้ค่า MAPE ที่ต่ำที่สุด

งานวิจัยที่เกี่ยวข้อง

นงนุช (2534) ศึกษาเปรียบเทียบเทคนิคการพยากรณ์ระหว่างวิธีแยกส่วนประกอบ กับ วิธีบ็อกซ์และเจนกินส์ ที่เหมาะสมสำหรับราคาข้าว มันสำปะหลัง และถั่วเขียว ที่เกษตรกรขายได้ที่ไร่นา ผลการศึกษาสรุปว่า การวิเคราะห์ด้วยวิธีแยกส่วนประกอบ มีอนุกรมเวลาที่นำมาศึกษามีแนวโน้มเป็นเส้นโค้งโพลีโนเมียลอันดับที่ 3 วิธีนี้เหมาะสำหรับการพยากรณ์ราคาข้าวเปลือกเหนียวเมล็ดยาว และราคามันสำปะหลังสดคละ ส่วนวิธีบ็อกซ์และเจนกินส์ พบว่าอนุกรมเวลาที่นำมาศึกษาทำให้เป็น สเตชันนารี ได้จากการหาค่าผลต่างเพียงครั้งเดียว และวิธีนี้เหมาะสำหรับการพยากรณ์ข้าวเปลือกเจ้า 5 % และราคาถั่วเขียวผิวมันเมล็ดใหญ่ชนิดคละ

รัชนิวรรณ (2539) ศึกษาเปรียบเทียบวิธีการพยากรณ์ด้วยค่าวัดความถูกต้อง 3 ค่า ซึ่งเป็นฟังก์ชันของความคลาดเคลื่อน ได้แก่ MAPE (Mean Absolute Percentage Error) MAD (Mean Absolute Deviation) และ MSE (Mean Square Error) โดยใช้ข้อมูลอนุกรมเวลารายเดือนที่มีการเคลื่อนไหวที่แตกต่างกัน 4 ลักษณะ ลักษณะละ 15 อนุกรมเวลา ผลการศึกษาพบว่าโดยส่วนใหญ่ค่าวัดความถูกต้องทั้ง 3 ค่าให้ผลสรุปสอดคล้องกัน คือ วิธีนี้พหุหนึ่ง และวิธีบ็อกซ์และเจนกินส์ วิธีการพยากรณ์ที่เหมาะสม สำหรับอนุกรมเวลาที่มีแนวโน้มแต่ไม่มีฤดูกาล คือ วิธีบ็อกซ์และเจนกินส์ วิธีการพยากรณ์ที่เหมาะสมสำหรับอนุกรมเวลาที่ไม่มีความโน้มแต่มีฤดูกาล คือ วิธีโฮลท์และวินเตอร์ และวิธีของบ็อกซ์และเจนกินส์ ส่วนอนุกรมเวลาที่มีอิทธิพลของแนวโน้มและฤดูกาล วิธีการพยากรณ์ที่เหมาะสมคือ วิธีโฮลท์และวินเตอร์

มนฤดี (2542) ศึกษาการพยากรณ์พื้นที่เพาะปลูก ผลผลิต และราคาสินค้าเกษตร ในกรณีข้าว (ข้าวนาปีและข้าวนาปรัง) ข้าวโพดเลี้ยงสัตว์ ถั่วเขียว และถั่วเหลือง โดยมีวัตถุประสงค์เพื่อหาวิธีการพยากรณ์ที่เหมาะสมกับพื้นที่เพาะปลูก ผลผลิต และราคาสินค้าเกษตรทั้ง 4 ชนิด โดยใช้เทคนิคการพยากรณ์ ได้แก่ วิธีการวิเคราะห์การถดถอย วิธีบ็อกซ์และเจนกินส์ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล วิธีตัดถดถอย วิธีการวิเคราะห์อนุกรมเวลาแบบคลาสสิก และเปรียบเทียบกับวิธีของศูนย์สารสนเทศเกษตร โดยใช้ค่าเฉลี่ยของเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เป็นเกณฑ์ในการเปรียบเทียบวิธีการทั้ง 6 วิธี

ผลการศึกษพบว่ารูปแบบการถดถอยจะเหมาะกับวิธีการพยากรณ์พื้นที่เพาะปลูกข้าวนาปี ข้าวนาปรัง ถั่วเขียว ถั่วเหลือง ข้าวโพดเลี้ยงสัตว์ความชื้นไม่เกิน 14 % รูปแบบบ็อกซ์และ เจนกินส์เหมาะสำหรับราคาข้าวเปลือกเจ้านาปี 5 % และราคาถั่วเหลืองชนิดคละ รูปแบบอัตโนมัติเหมาะสำหรับการพยากรณ์พื้นที่เพาะปลูกข้าวโพดเลี้ยงสัตว์ ถั่วเหลือง ปริมาณ ผลผลิตถั่วเขียว และราคาข้าวเปลือกเจ้านาปรังความชื้น 14-15 % สำหรับรูปแบบอนุกรมเวลาแบบ คลาสสิก เหมาะสำหรับการพยากรณ์ราคาถั่วเขียวผิวมันเมล็ดใหญ่คละชนิด

รัศมี (2542) ศึกษาการพยากรณ์ปริมาณน้ำฝนในภาคตะวันออกเฉียงเหนือ เพื่อวางแผนการ เพาะปลูกพืช โดยมีวัตถุประสงค์เพื่อเปรียบเทียบรูปแบบพยากรณ์จากการใช้เทคนิคการพยากรณ์ สำหรับอนุกรมเวลาที่มีระดับค่าเฉลี่ยคงที่ วิธีการพยากรณ์ที่ใช้มี 3 วิธี ได้แก่ วิธีแยกส่วนประกอบ วิธีการพยากรณ์ของวินเตอร์ และเทคนิคการวิเคราะห์อนุกรมเวลาบ็อกซ์และเจนกินส์ การเปรียบเทียบวิธีการพยากรณ์ที่เหมาะสมจะพิจารณาจากค่าวัดความถูกต้องของการพยากรณ์ MAPE (Mean Absolute Percentage Error) ผลการศึกษสรุปได้ว่า วิธีแยกส่วนประกอบเป็นวิธีการ ที่เหมาะสมกับลักษณะข้อมูลปริมาณน้ำฝนมากที่สุด สำหรับทุกจังหวัดในภาคตะวันออกเฉียงเหนือ และนอกจากนี้ทำการวิเคราะห์แนวโน้มพื้นที่ผิว เปรียบเทียบค่าพยากรณ์จากการวิเคราะห์แนวโน้ม พื้นที่ผิว กับการพยากรณ์ด้วยค่าเฉลี่ยคงที่ พบว่าการวิเคราะห์แนวโน้มพื้นที่ผิว ให้ค่าพยากรณ์ปริมาณ น้ำฝนใกล้เคียงกับปริมาณน้ำฝนจริงมากที่สุด

นงลักษณ์ (2544) ศึกษาประสิทธิภาพของวิธีการพยากรณ์รวม โดยมีวัตถุประสงค์เพื่อ พิจารณาหาวิธีการพยากรณ์เดี่ยวและวิธีการพยากรณ์รวมที่เหมาะสมกับขนาดของอนุกรมเวลาและ ช่วงการพยากรณ์ การพิจารณาเปรียบเทียบวิธีการพยากรณ์ที่เหมาะสมใช้ค่าวัดความถูกต้องของการ พยากรณ์ ได้แก่ ค่า MAPE (Mean Absolute Percentage Error) วิธีการพยากรณ์เดี่ยวที่นำมาพิจารณา สำหรับอนุกรมเวลาที่มีการเคลื่อนไหวของแนวโน้มมี 4 วิธี ได้แก่ วิธีเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง วิธีการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง วิธีการวิเคราะห์การถดถอยและ วิธีของ บ็อกซ์และเจนกินส์ สำหรับอนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากอิทธิพลของแนวโน้มและ ฤดูกาลมี 4 วิธี ได้แก่ วิธีวิเคราะห์การถดถอย วิธีแยกส่วนประกอบ วิธีไฮลท์และวินเตอร์ และวิธี ของบ็อกซ์และเจนกินส์ วิธีการพยากรณ์รวมมี 4 วิธี ได้แก่ วิธีเฉลี่ยอย่างง่าย วิธีเฉลี่ยถ่วงน้ำหนัก วิธีเฉลี่ยถ่วงน้ำหนักซึ่งน้ำหนักผกผันกับค่าผลบวกกำลังสองของค่าความคลาดเคลื่อน และวิธีเฉลี่ย ถ่วงน้ำหนักซึ่งน้ำหนักได้จากวิธีการวิเคราะห์การถดถอย ผลการศึกษสรุปได้ว่าวิธีการพยากรณ์ รวมให้วิธีที่คิดว่าวิธีการพยากรณ์เดี่ยว โดยวิธีเฉลี่ยถ่วงน้ำหนักได้จากวิธีการวิเคราะห์การถดถอย

เหมาะสมกับอนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากอิทธิพลของแนวโน้ม ขนาด 4, 5 และ 6 ปี ในช่วงเวลาการพยากรณ์ 6, 12, 18 และ 24 เดือน และเหมาะสมกับอนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากอิทธิพลของแนวโน้มและฤดูกาล ขนาด 4, 5 และ 6 ปี ในช่วงเวลาการพยากรณ์ 6, 12, 18 และ 24 เดือน

ทักดาว (2545) ศึกษาเปรียบเทียบการพยากรณ์ด้วยวิธีโครงข่ายประสาทเทียม และวิธีของบ็อกซ์และเจนกินส์ กรณีศึกษา อัตราแลกเปลี่ยนเงินตราต่างประเทศ โดยมีวัตถุประสงค์เพื่อศึกษารูปแบบการพยากรณ์ที่เหมาะสมในการพยากรณ์อัตราแลกเปลี่ยนเงินตราต่างประเทศ 6 สกุล ด้วยวิธีโครงข่ายประสาทเทียม โดยทำการศึกษาหาขนาดอนุกรมเวลาที่เหมาะสมสำหรับสร้างรูปแบบการพยากรณ์ที่มีความถูกต้องที่สุด จากอนุกรมเวลา 3 ขนาด โดยกำหนดโครงสร้างของโครงข่ายประสาทเทียม 5 ลักษณะ เพื่อศึกษาปัจจัยที่สำคัญ 3 ปัจจัย อนุกรมเวลาที่นำมาศึกษาเป็นอนุกรมเวลารายวันของอัตราแลกเปลี่ยนเงินดอลลาร์สหรัฐฯ เงินคอยซ์มาร์กเยอรมัน เงินเยนญี่ปุ่น เงินปอนด์สเตอร์ลิงอังกฤษ เงินฟรังก์สวิสเซอร์แลนด์ และเงินยูโรกับเงินไทย การพิจารณาเปรียบเทียบวิธีการพยากรณ์ที่เหมาะสมจะใช้ค่าวัดความถูกต้องของการพยากรณ์ 3 ค่า ได้แก่ ME (Mean Error), MSE (Mean Square Error) และ MAPE (Mean Absolute Percentage Error) ผลการศึกษาสรุปได้ว่า สำหรับอนุกรมเวลาอัตราแลกเปลี่ยนเงินดอลลาร์สหรัฐฯ และเงินคอยซ์มาร์กเยอรมัน ขนาดอนุกรมเวลาที่เหมาะสมคือ 30 สำหรับเงินเยนญี่ปุ่น และเงินปอนด์สเตอร์ลิงอังกฤษ ขนาดอนุกรมเวลาที่เหมาะสมคือ 60 สำหรับเงินฟรังก์สวิสเซอร์แลนด์ ขนาดอนุกรมเวลาที่เหมาะสมคือ 30 และ 60 ส่วนเงินยูโรขนาดอนุกรมเวลาที่เหมาะสมคือ 150 ส่วนการเปรียบเทียบความถูกต้องของการพยากรณ์ด้วยวิธีวิธีโครงข่ายประสาทเทียมและวิธีของบ็อกซ์และเจนกินส์ ยกเว้น เงินเยนญี่ปุ่น วิธีของบ็อกซ์และเจนกินส์ ให้ค่าพยากรณ์ที่มีความถูกต้องสูงกว่าวิธีโครงข่ายประสาทเทียม

จันทร์ธา (2548) ศึกษาการพยากรณ์ปริมาณความต้องการใช้ไฟฟ้า กรณีศึกษาการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย งานวิจัยนี้ได้ศึกษาหาวิธีการพยากรณ์ ขนาดอนุกรมเวลา ช่วงการพยากรณ์ที่เหมาะสม และศึกษาตัวแปรชี้้นำ โดยใช้คุณทฤษฎีเฉลี่ย ข้อมูลการพยากรณ์คือปริมาณความต้องการใช้ไฟฟ้า 4 ชุดของการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย วิธีการพยากรณ์ที่ใช้ในการวิจัยนี้มี 5 วิธี ได้แก่ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล วิธีแยกส่วนประกอบ วิธีวิเคราะห์การถดถอยที่ใช้ตัวแปรคัมมี วิธีของบ็อกซ์และเจนกินส์ และวิธีการพยากรณ์ร่วม การเลือกวิธีการพยากรณ์ที่เหมาะสมได้เปรียบเทียบโดยใช้ค่าวัดความคลาดเคลื่อน 3 วิธี ได้แก่ ค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) ค่าเฉลี่ยส่วนเบี่ยงเบนสัมบูรณ์ (MAD) และ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (MSE) ผลการศึกษาเปรียบเทียบวิธีการพยากรณ์เดี่ยวทั้ง 4 วิธี พบว่าวิธีของบ็อกซ์และเจนกินส์เหมาะสมกว่าวิธีอื่นๆ และวิธีการพยากรณ์ร่วมให้ค่าคลาดเคลื่อนต่ำกว่าวิธีเดี่ยวทุกวิธี ในการเลือกขนาดอนุกรมเวลาพยากรณ์พบว่า ขนาดอนุกรมเวลา 10 ปี เหมาะสมสำหรับอนุกรมเวลาชุดที่ 1 ขนาดอนุกรมเวลา 4 ปี เหมาะสมสำหรับอนุกรมเวลาชุดที่ 2 3 และ 4 ในการพยากรณ์ล่วงหน้าไป 2 และ 6 เดือน วิธีที่เหมาะสม ได้แก่ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล และวิธีของบ็อกซ์และเจนกินส์ ส่วนการพยากรณ์ล่วงหน้าไป 12 เดือนวิธีที่เหมาะสม ได้แก่ วิธีของบ็อกซ์และเจนกินส์ และวิธีแยกส่วนประกอบ ในการศึกษาตัวแปรชี้้นำ พบว่าคุณทฤษฎีเฉลี่ยสามารถใช้เป็นตัวแปรชี้้นำของตัวแปรทั้ง 4 ชุด

Boluwade (2008) ศึกษาตัวแบบของการไหลของแม่น้ำไนเจอร์ด้วยวิธีบ็อกซ์และเจนกินส์ สำหรับอนุกรมเวลาที่มีความผันแปรตามฤดูกาล ใช้ข้อมูลปริมาณน้ำรายเดือนของแม่น้ำไนเจอร์ ที่เก็บรวบรวมจาก NEPA Hydrogenerating Station เมือง Jebba ประเทศไนจีเรีย โดยมีวัตถุประสงค์เพื่อพยากรณ์ปรากฏการณ์ปริมาณน้ำในแม่น้ำไนเจอร์ เพื่อเตรียมการวางแผนรับมือ โดยจะมีอยู่สองปรากฏการณ์ คือปรากฏการณ์แรกเรียกว่า Black flood เป็นปรากฏการณ์ที่เกิดขึ้นในช่วงเดือนพฤศจิกายนถึงเดือนกรกฎาคม ซึ่งมีปริมาณน้ำอยู่ที่ 900 ถึง 1000 ลูกบาศก์เมตรต่อวินาที ซึ่งน้ำในช่วงเดือนนี้จะมีปริมาณน้อย ทำให้เกิดความแห้งแล้ง ปรากฏการณ์ที่ 2 เรียกว่า White flood เป็นปรากฏการณ์ที่เกิดขึ้นในช่วงเดือนสิงหาคมถึงเดือนตุลาคม ซึ่งมีปริมาณน้ำอยู่ที่ 1500 ถึง 1600 ลูกบาศก์เมตรต่อวินาที ซึ่งน้ำในช่วงเดือนนี้จะมีปริมาณมาก ทำให้เกิดอุทกภัยได้

นำข้อมูลปริมาณน้ำรายเดือนของแม่น้ำไนเจอร์มาวิเคราะห์ โดยแบ่งออกเป็นสามชุด ชุดแรกเป็นข้อมูลตั้งแต่ปีค.ศ.1984 ถึง 1995 นำมากำหนดรูปแบบ, ประมาณค่าพารามิเตอร์ และตรวจสอบความถูกต้องให้กับรูปแบบให้กับอนุกรมเวลา ชุดที่ 2 เป็นข้อมูลตั้งแต่ปี ค.ศ.1996 ถึง 1997 ใช้ตรวจสอบตัวแบบ และชุดที่ 3 เป็นข้อมูลตั้งแต่ปีค.ศ.1984 ถึง 1998 ใช้สำหรับการพยากรณ์ ค่าในอนาคต และใช้ค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (APE) ในการตรวจสอบความเหมาะสมของค่าพยากรณ์ที่ได้

ผลที่ได้จากงานวิจัยคือ ได้รูปแบบที่เหมาะสมกับข้อมูลปริมาณน้ำรายเดือนของแม่น้ำไนเจอร์ ซึ่งทำให้สามารถพยากรณ์ปริมาณน้ำในแม่น้ำไนเจอร์ของปี 1998 – 2000 ทำให้มีการจัดระบบการกักเก็บน้ำ เช่น สร้างอ่างเก็บน้ำขึ้นมา เมื่อปริมาณน้ำมาก เพื่อใช้ในการชลประทาน เมื่อปริมาณน้ำน้อย เป็นต้น

จากการศึกษางานวิจัยที่เกี่ยวข้อง สรุปได้ว่าการพยากรณ์มีประโยชน์ในการวิจัยและวางแผนในการดำเนินงานในสาขาวิชาต่างๆมากมาย ทั้งในด้านการเกษตร ด้านเศรษฐกิจ ด้านการดำรงชีวิต และอื่นๆ ผู้วิจัยจึงมีความสนใจที่จะศึกษาเปรียบเทียบเทคนิคการพยากรณ์ในแต่ละวิธี เพื่อเลือกวิธีการพยากรณ์ที่เหมาะสมกับอนุกรมเวลาปริมาณน้ำฝนของจังหวัดลำปาง โดยเลือกวิธีที่เหมาะสมที่สุด โดยใช้ค่าวัดความถูกต้องของการพยากรณ์ 3 ค่า ได้แก่ MAD, MSE และ MAPE

อุปกรณ์และวิธีการ

อุปกรณ์

อุปกรณ์ที่ใช้ในงานวิจัยประกอบด้วย

1. ข้อมูลปริมาณน้ำฝนรายเดือนของสถานีอุตุนิยมวิทยาลัยลำปาง จังหวัดลำปาง ซึ่งเก็บรวบรวมตั้งแต่เดือนมกราคม 2542 ถึงเดือนพฤษภาคม 2552 ซึ่งแหล่งที่มาของข้อมูล คือ สำนักพัฒนาอุตุนิยมวิทยา กรมอุตุนิยมวิทยา กระทรวงเทคโนโลยีสารสนเทศ และการสื่อสาร
2. โปรแกรมสำเร็จรูปทางสถิติ SPSS (Statistical Package for the Social Sciences) เวอร์ชัน 11.5, Minitab for Windows เวอร์ชัน 14 และ Microsoft Office Excel 2003
3. เครื่องไมโครคอมพิวเตอร์ของภาควิชาสถิติ มหาวิทยาลัยเกษตรศาสตร์

วิธีการ

1. รวบรวมข้อมูลปริมาณน้ำฝนรายเดือนของสถานีอุตุนิยมวิทยาลัยลำปาง จังหวัดลำปาง ตั้งแต่เดือนมกราคม 2542 ถึงเดือนพฤษภาคม 2552
2. นำข้อมูลอนุกรมเวลารายเดือนตั้งแต่เดือนมกราคม 2542 ถึง เดือนพฤษภาคม 2552 แบ่งออกเป็นสองชุด คือ

ชุดที่ 1 ข้อมูลปริมาณน้ำฝนรายเดือน ตั้งแต่เดือนมกราคม 2542 ถึงเดือนพฤษภาคม 2551 รวมระยะเวลา 113 เดือน ใช้สำหรับการวิเคราะห์เพื่อสร้างรูปแบบพยากรณ์

ชุดที่ 2 ข้อมูลปริมาณน้ำฝนรายเดือน ตั้งแต่เดือนมิถุนายน 2551 ถึงเดือนพฤษภาคม 2552 ใช้สำหรับการตรวจสอบผลของการพยากรณ์ของรูปแบบพยากรณ์แต่ละรูปแบบที่สร้างขึ้น ด้วยข้อมูลชุดที่ 1

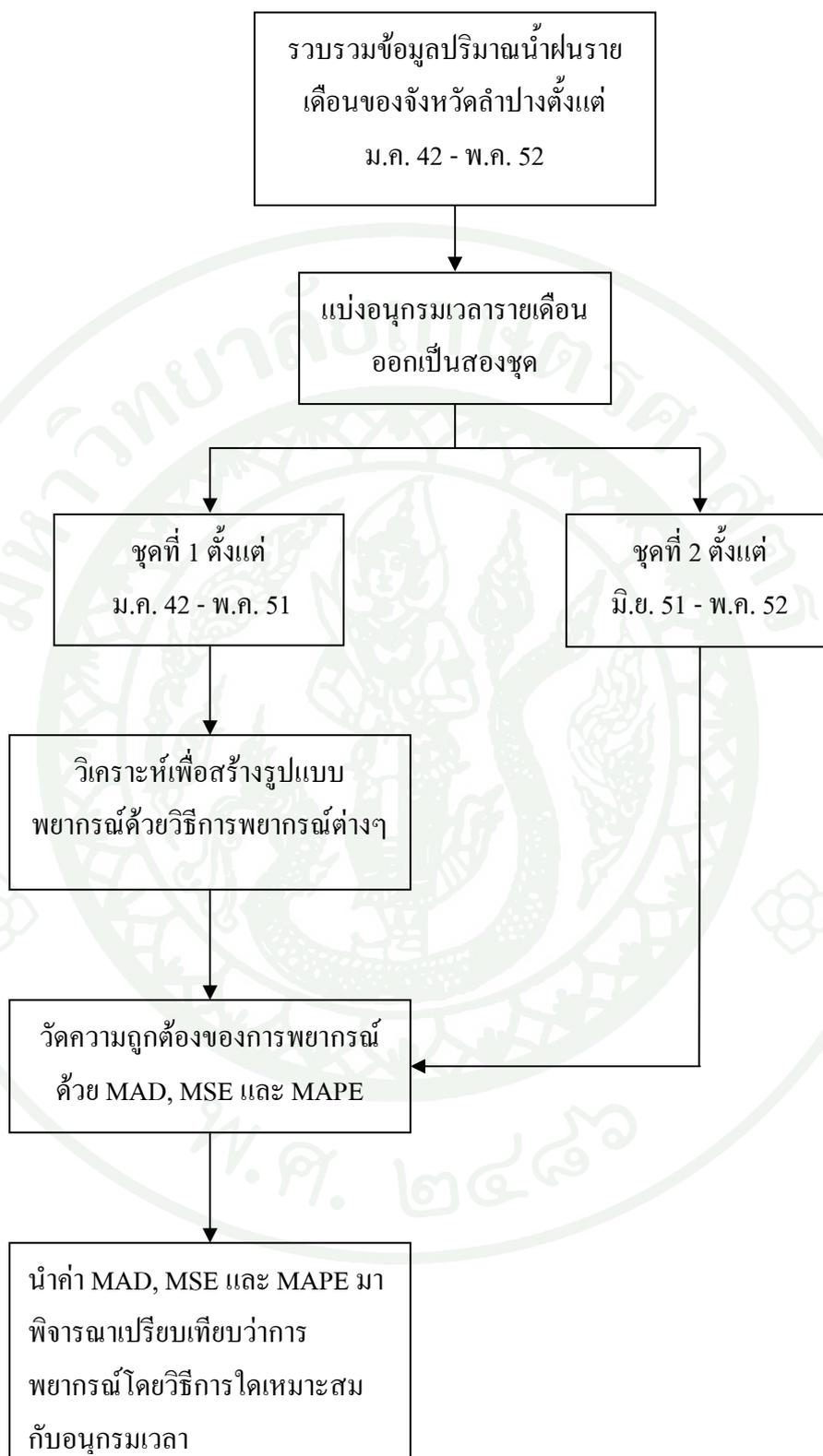
3. นำอนุกรมเวลาชุดที่ 1 มาสร้างสมการพยากรณ์ด้วยวิธีการพยากรณ์เดี่ยว 3 วิธี และวิธีการพยากรณ์ร่วม 1 วิธี

4. นำค่าพยากรณ์ที่ได้จากข้อ 3 มาคำนวณค่าวัดความถูกต้องของการพยากรณ์โดยเปรียบเทียบกับข้อมูลจริง ตั้งแต่เดือนเดือนมิถุนายน 2551 ถึงเดือนพฤษภาคม 2552 และวัดความถูกต้องโดยค่า MAD (Mean Absolute Deviation), MSE (Mean Square Error) และ MAPE (Mean Absolute Percent Error)

5. นำค่า MAD, MSE และ MAPE ที่ได้จากอนุกรมเวลา มาพิจารณาเปรียบเทียบว่าการพยากรณ์ โดยวิธีการใดเหมาะสมกับอนุกรมเวลา โดยแยกเปรียบเทียบวิธีการพยากรณ์เดี่ยว จากนั้นทำการเปรียบเทียบวิธีการพยากรณ์เดี่ยวที่เหมาะสมที่สุดกับวิธีการพยากรณ์ร่วม

สถานที่และระยะเวลาในการทำการวิจัย

สถานที่ทำการวิจัย คือ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ และระยะเวลาที่ทำการวิจัย เริ่มตั้งแต่เดือนเมษายน 2552 ถึง เดือนเมษายน 2553



ภาพที่ 11 แผนภาพขั้นตอนในการวิจัย

ผลและวิจารณ์

ผล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบเทคนิคในการพยากรณ์น้ำฝนโดยศึกษาจากข้อมูลปริมาณน้ำฝนรายเดือนของจังหวัดลำปาง ซึ่งเก็บรวบรวมตั้งแต่เดือนมกราคม 2542 ถึงเดือนพฤษภาคม 2552 รวมระยะเวลา 125 เดือน วิเคราะห์ข้อมูลดังกล่าวด้วยเทคนิคการพยากรณ์ทั้งวิธีการพยากรณ์เดี่ยวและการพยากรณ์ร่วม วิธีการพยากรณ์เดี่ยวที่ใช้ในงานวิจัยนี้ ได้แก่ วิธีแยกส่วนประกอบ (Decomposition Method) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์ (Triple Exponential Smoothing : Winter's Method) วิธีบ็อกซ์และเจนกินส์ (Box – Jenkins Method) ส่วนวิธีการพยากรณ์ร่วม ได้แก่ วิธีเฉลี่ยถ่วงน้ำหนักซึ่งน้ำหนักผกผันกับค่าผลบวกกำลังสองของค่าความคลาดเคลื่อน ผลการศึกษาแบ่งออกเป็น 3 ส่วนดังนี้

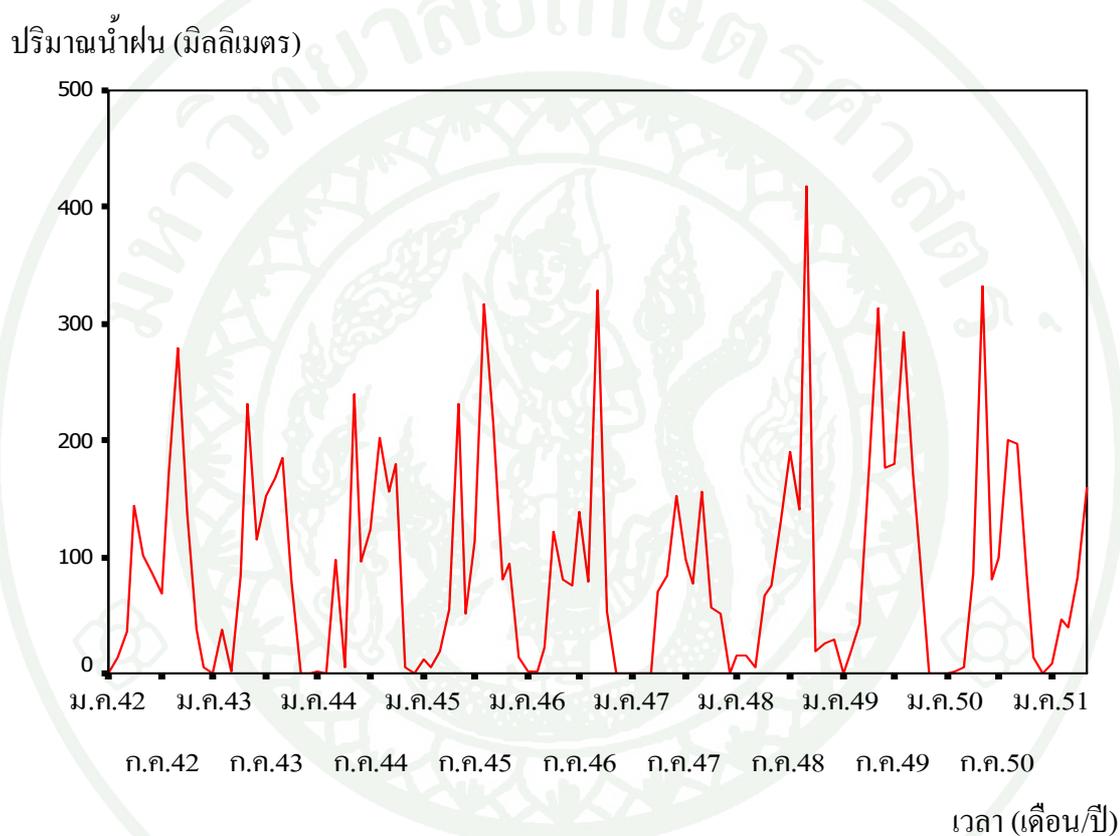
ส่วนที่ 1 ผลการวิเคราะห์ข้อมูลปริมาณน้ำฝนเบื้องต้น

ส่วนที่ 2 ผลการวิเคราะห์ข้อมูลปริมาณน้ำฝนด้วย วิธีการพยากรณ์เดี่ยว

ส่วนที่ 3 ผลวิเคราะห์ข้อมูลปริมาณน้ำฝนด้วย วิธีการพยากรณ์ร่วม

1. ผลการวิเคราะห์ข้อมูลปริมาณน้ำฝนเบื้องต้น

จากการนำข้อมูลปริมาณน้ำฝนของจังหวัดลำปางที่เก็บรวบรวมรายเดือนตั้งแต่ เดือน มกราคม 2542 ถึงเดือนพฤษภาคม 2551 มาสร้างกราฟการเคลื่อนไหวอนุกรมเวลา เพื่อพิจารณาองค์ประกอบของอนุกรมเวลา ดังแสดงในภาพที่ 11



ภาพที่ 12 การเคลื่อนไหวอนุกรมเวลาของปริมาณน้ำฝนจังหวัดลำปาง

จากภาพที่ 12 พบว่าแนวโน้มการเคลื่อนไหวของปริมาณน้ำฝนในแต่ละปีไม่แตกต่างกันมากนัก นั่นคือในช่วง 3 เดือนแรก และช่วง 3 เดือนสุดท้ายของแต่ละปี มีค่าน้อยหรือเข้าใกล้ศูนย์ จากนั้นนำข้อมูลที่เก็บรวบรวมได้มาคำนวณหาค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานเป็นรายปี รายละเอียดแสดงดังตารางที่ 3 และนำค่าเฉลี่ยของปริมาณน้ำฝนรายปีมาสร้างกราฟ ดังภาพที่ 13

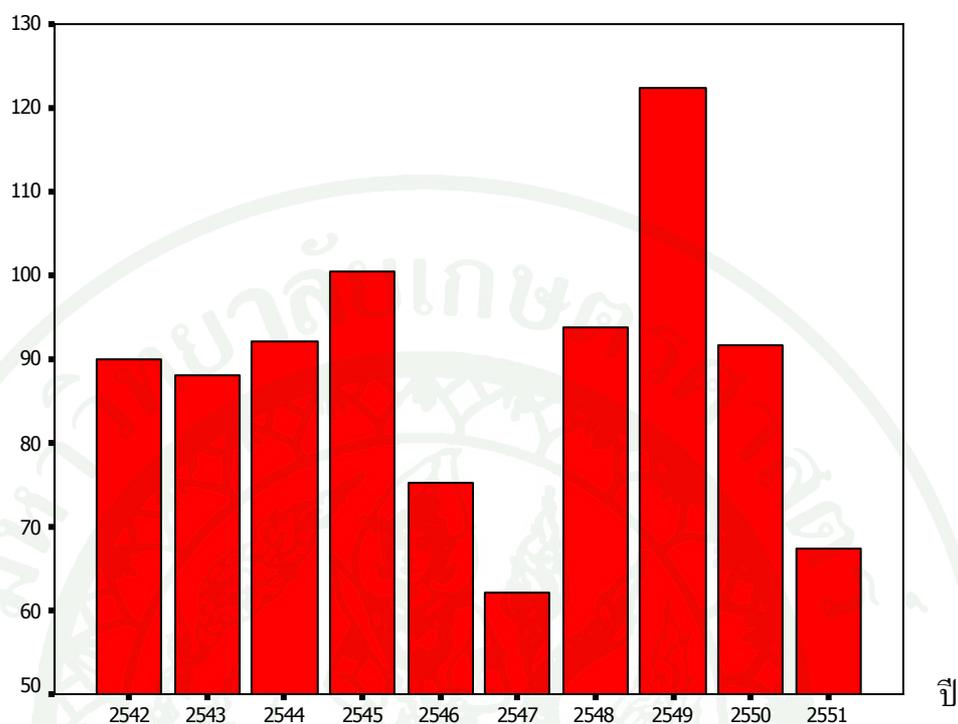
ตารางที่ 3 ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของปริมาณน้ำฝนจังหวัดลำปางในแต่ละปี ตั้งแต่ปี 2542-2551

ปี (พ.ศ.)	ค่าเฉลี่ย (\bar{x})	ส่วนเบี่ยงเบนมาตรฐาน (S.D.)
2542	90.09	82.32
2543	88.00	82.40
2544	92.08	88.88
2545	100.43	101.26
2546	75.13	93.54
2547	62.25	56.12
2548	93.90	117.92
2549	122.32	112.75
2550	91.77	104.37
2551	67.34	57.32

หมายเหตุ ค่าเฉลี่ยปริมาณน้ำฝนของปี 2551 ได้จากการคำนวณเพียง 6 เดือน (ม.ค. – พ.ค.)

จากตารางที่ 3 และภาพที่ 13 พบว่าปริมาณน้ำฝนเฉลี่ยรายปีของจังหวัดลำปาง ตั้งแต่ปี 2542-2551 ปีที่มีปริมาณน้ำฝนเฉลี่ยมากที่สุด 3 ลำดับแรก คือ ปี พ.ศ. 2549 มีปริมาณน้ำฝนมากที่สุดคือ 122.32 มิลลิเมตร รองลงมาคือ ปี พ.ศ. 2545 มีปริมาณน้ำฝนเฉลี่ยเท่ากับ 100.43 มิลลิเมตร และปี พ.ศ. 2548 มีน้ำฝนเฉลี่ยเท่ากับ 93.90 มิลลิเมตร

ปริมาณน้ำฝน (มิลลิเมตร)



ภาพที่ 13 ปริมาณน้ำฝนของจังหวัดลำปางเฉลี่ยเป็นรายปี ตั้งแต่ พ.ศ. 2542-2551

วิจารณ์

2. ผลการวิเคราะห์ข้อมูลด้วยวิธีการพยากรณ์เดี่ยว ดังนี้

2.1 วิธีแยกส่วนประกอบ (Decomposition Method)

วิธีแยกส่วนประกอบเป็นวิธีการศึกษาว่าอนุกรมเวลาที่สนใจได้รับอิทธิพลจากความผันแปรใดบ้าง จากนั้นก็ทำการวัดความผันแปรเหล่านั้นเพื่อนำไปหาค่าพยากรณ์ในอนาคต ในการวิเคราะห์ใช้รูปแบบการคูณ มีค่าพยากรณ์ Y หน่วยเวลาล่วงหน้า n เวลา t

สำหรับงานวิจัยนี้ เลือกใช้รูปแบบการคูณเนื่องจาก รูปแบบการคูณเป็นรูปแบบที่มีส่วนประกอบของอนุกรมเวลาที่ส่วนมีความสัมพันธ์กัน กล่าวคือ ถ้าส่วนประกอบใดส่วนประกอบหนึ่งมีค่าเปลี่ยนแปลงไปไม่ว่าจะเพิ่มขึ้นหรือลดลง จะมีผลกระทบต่อการ

เปลี่ยนแปลงของส่วนประกอบอื่นๆที่เหลือ รูปแบบชนิดนี้เป็นที่นิยมใช้กันมากกว่ารูปแบบอื่น เพราะเป็นรูปแบบพื้นฐานที่สามารถทำการแยกส่วนประกอบแต่ละชนิดของอนุกรมเวลาได้สะดวก และมีหลักการที่สมเหตุสมผลว่าส่วนประกอบแต่ละชนิดของอนุกรมเวลาจะมีส่วนเกี่ยวข้องซึ่งกันและกัน (ศิริลักษณ์, 2543: 19)

รูปแบบการคูณ (Multiplicative Model) เป็นรูปแบบที่มีความผันแปรแต่ละชนิดที่มีความสัมพันธ์กัน มีรูปแบบดังนี้

$$Y_t = (b_0 + b_1t)\hat{S}_t$$

โดยที่

$$Y_t = \text{อนุกรมเวลา}$$

$$b_0 = \text{ค่าเฉลี่ยของข้อมูล}$$

$$b_1 = \text{ความชัน}$$

$$\hat{S}_t = \text{ความผันแปรตามฤดูกาล}$$

ผลจากการวิเคราะห์ข้อมูลด้วยโปรแกรมสำเร็จรูปวิธีแยกส่วนประกอบ ได้ผลการของวิธีการแยกส่วนประกอบ และดัชนีฤดูกาล แสดงดังตารางที่ 4

$$Y_t = (132.432 + 0.364579t)\hat{S}_t$$

โดยที่

$$b_0 = 132.432$$

และ

$$b_1 = 0.364579$$

ตารางที่ 4 คำนวณฤดูกาลของวิธีแยกส่วนประกอบ

เดือนที่	1	2	3	4	5	6
\hat{S}_t	0.01143	0.04562	0.13314	0.96924	2.60561	1.24236

เดือนที่	7	8	9	10	11	12
\hat{S}_t	1.56011	2.06145	2.28563	0.92679	0.15861	0.00000

ตัวอย่างการคำนวณหาค่าพยากรณ์

พยากรณ์ปริมาณน้ำฝนเดือนมิถุนายน พ.ศ.2551 (ข้อมูลที่ 114)

จากสมการ $Y_t = (132.432 + 0.364579t)\hat{S}_t$

จะได้ว่า $Y_{114} = (132.432 + 0.364579(114))(1.24236)$

$$Y_{114} = 216.16$$

นั่นคือ ปริมาณน้ำฝนเดือนมิถุนายน พ.ศ.2551 ที่ได้จากการพยากรณ์ มีค่าเท่ากับ 216.16 มิลลิเมตร

พยากรณ์ปริมาณน้ำฝนเดือนกุมภาพันธ์ พ.ศ.2552 (ข้อมูลที่ 122)

จากสมการ $Y_t = (132.432 + 0.364579t)\hat{S}_t$

จะได้ว่า $Y_{122} = (132.432 + 0.364579(122))(0.04562)$

$$Y_{122} = 8.07$$

นั่นคือ ปริมาณน้ำฝนเดือนกุมภาพันธ์ พ.ศ.2552 ที่ได้จากการพยากรณ์ มีค่าเท่ากับ 8.07 มิลลิเมตร

คำนวณค่าพยากรณ์ของปริมาณน้ำฝน คำนวณค่า MAD, MSE และ MAPE ผลการวิเคราะห์ข้อมูล แสดงได้ดังตารางที่ 5

ตารางที่ 5 ค่า MAD, MSE และ MAPE ของอนุกรมเวลาสำหรับค่าที่ได้จากการพยากรณ์
เปรียบเทียบกับค่าจริงด้วยวิธีแยกส่วนประกอบ

เดือน/ปี	ปริมาณน้ำฝนจริง (มิลลิเมตร)	ปริมาณน้ำฝนจากการพยากรณ์ (มิลลิเมตร)	MAD	MSE	MAPE
มิ.ย. 51	74.10	216.16	142.06	20,181.04	191.71
ก.ค. 51	81.70	272.02	190.32	36,221.70	232.95
ส.ค. 51	100.00	360.18	260.18	67,693.63	260.18
ก.ย. 51	191.60	400.19	208.59	43,509.79	108.88
ต.ค. 51	160.40	162.61	2.21	4.88	1.38
พ.ย. 51	37.70	27.89	9.81	96.24	26.02
ธ.ค. 51	12.30	0.00	12.30	151.29	100.00
ม.ค. 52	0.00	2.02	2.02	4.08	-
ก.พ. 52	0.00	8.07	8.07	65.12	-
มี.ค. 52	26.90	23.60	3.30	10.89	12.27
เม.ย. 52	55.20	172.18	116.98	13,684.32	211.92
พ.ค. 52	237.70	463.81	226.11	31,125.73	95.12
ค่าเฉลี่ย			98.50	17,729.06	103.37

จากตารางที่ 5 แสดงปริมาณน้ำฝนจริงและปริมาณน้ำฝนจากการพยากรณ์ด้วยวิธีแยกส่วนประกอบ ได้ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) มีค่าเท่ากับ 98.50 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าเท่ากับ 17,729.06 และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) มีค่าเท่ากับ 103.37

2.2 วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์ (Triple Exponential Smoothing : Winters's Method)

เป็นวิธีหนึ่งของวิธีการทำให้เรียบแบบเอ็กซ์โปเนนเชียล ที่ใช้ในการพยากรณ์สำหรับอนุกรมเวลาที่มีแนวโน้มและความผันแปรตามฤดูกาล และเหมาะสำหรับการพยากรณ์ในระยะสั้นจนถึงการพยากรณ์ในระยะปานกลาง ซึ่งมีรายละเอียดดังนี้

รูปแบบการคูณ (Multiplicative Seasonality Model) มีรูปแบบดังนี้

$$\hat{Y}_{t+m} = (a_t + b_t(m))\hat{S}_{t-p+m}$$

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

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

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

เมื่อ \hat{Y}_{t+m} คือ ข้อมูลหรือค่าสังเกต ณ เวลา $t+m$

m คือ จำนวนช่วงเวลาที่ต้องการพยากรณ์ไปข้างหน้า

p คือ จำนวนฤดูกาล

a_t คือ ระดับของข้อมูลหรือส่วนที่เป็นการทำให้เรียบ

b_t คือ ส่วนที่เป็นแนวโน้ม

\hat{S}_t คือ ส่วนที่เป็นฤดูกาล

วิธีของวินเตอร์วิเคราะห์โดยใช้รูปแบบการคูณและกำหนดค่า α , γ และ δ ระหว่าง 0.001 – 1.0 โดยการทดลองเปลี่ยนค่า α , γ และ δ ให้เพิ่มขึ้นทีละ 0.01 จากนั้นเลือกค่า α , γ และ δ ที่ให้ค่า MSE ต่ำที่สุด ซึ่งค่า α , γ และ δ ของวิธีของวินเตอร์คือ $\alpha = 0.001$, $\gamma = 0.01$ และ $\delta = 0.001$

มีค่าพยากรณ์ m หน่วยเวลาล่วงหน้า ที่พยากรณ์ ณ เวลา t และข้อมูลเป็นรายเดือน คือ

$$\hat{Y}_{t+m} = (a_t + b_t(m))\hat{S}_{t-12+m}$$

จากการวิเคราะห์ จะได้

$$a_0 = 89.98698$$

$$b_0 = 0.01745$$

จากการวิเคราะห์ข้อมูล ได้ดัชนีฤดูกาลวินเตอร์ \hat{S}_t เมื่อ $t = -11, -10, \dots, 0$ แสดงดังตารางที่ 6

ตารางที่ 6 ดัชนีฤดูกาลวินเตอร์

t	-11	-10	-9	-8	-7	-6
\hat{S}_t	0.0315	0.0778	0.1602	0.9787	2.1656	1.2136
t	-5	-4	-3	-2	-1	0
\hat{S}_t	1.5106	1.9685	2.6513	0.9631	0.2489	0.0302

$$\text{และ } a_t = 0.001 \left(\frac{Y_t}{\hat{S}_{t-12}} \right) + 0.999[a_{t-1} + b_{t-1}]$$

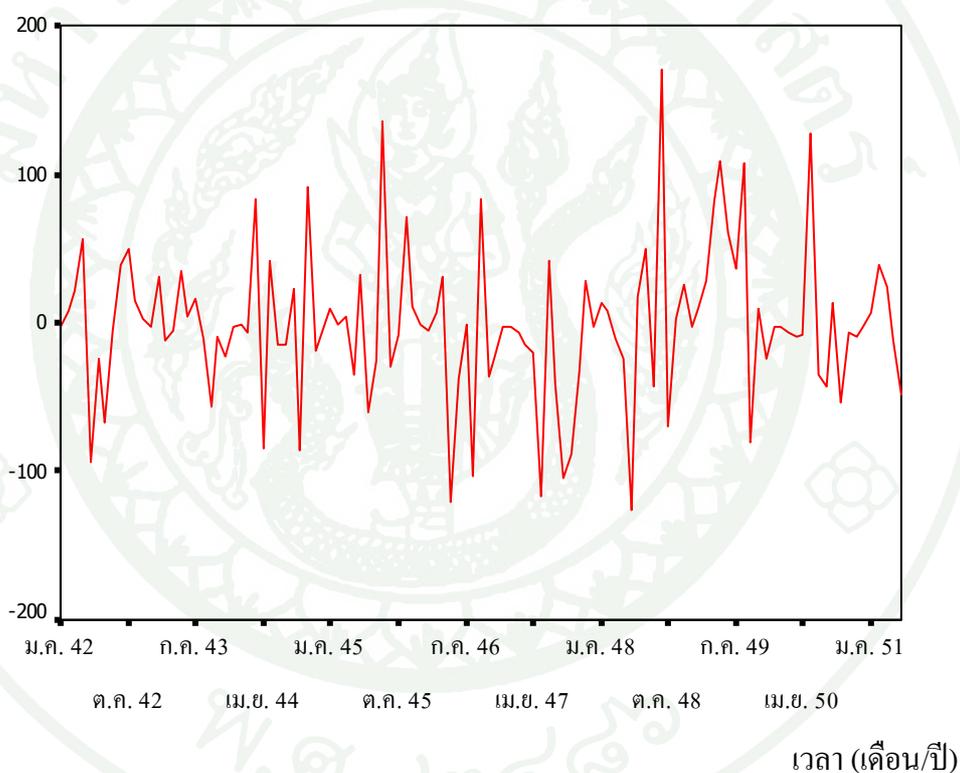
$$b_t = 0.01(a_t - a_{t-1}) + 0.99b_{t-1}$$

$$\hat{S}_t = 0.001 \left(\frac{Y_t}{a_t} \right) + 0.999 \hat{S}_{t-12}$$

ผลการวิเคราะห์ข้อมูล สรุปได้ดังตารางที่ 8

การตรวจสอบความเหมาะสมของรูปแบบ เมื่อทดสอบค่าความคลาดเคลื่อนกับเวลา ดังภาพที่ 13 พบว่าค่าความคลาดเคลื่อนกระจายอยู่รอบค่าศูนย์ และความแปรปรวนคงที่ (วราฤทธิ์, 2550: 144)

ความคลาดเคลื่อนจากวิธีวินเตอร์



ภาพที่ 14 กราฟแสดงความคลาดเคลื่อนจากวิธีของวินเตอร์

ตารางที่ 7 ค่า $r_k(e_t)$ และคอเรลโลแกรมของค่า $r_k(e_t)$ ของวิธีวินเตอร์

Lag	Corr. $r_k(e_t)$	Err.	-.1 -.75 -.5 -.25 0 .25 .5 .75 1	Box-Ljung	p-value
1	-.117	.093	. ** .	1.575	.210
2	.098	.092	. **	2.701	.259
3	.164	.092	. ***	5.862	.119
4	-.188	.092	**** .	10.096	.039
5	.099	.091	. **	11.269	.046
6	-.125	.091	. ** .	13.164	.040
7	.108	.090	. **	14.593	.042
8	.040	.090	. *	14.795	.063
9	.082	.089	. **	15.632	.075
10	.130	.089	. ****	17.772	.059
11	-.035	.089	. *	17.926	.083
12	-.058	.088	. *	18.352	.105
13	.025	.088	. *	18.430	.142
14	-.040	.087	. *	18.644	.179
15	.017	.087	. *	18.682	.229
16	-.073	.086	. *	19.400	.248

จากตารางที่ 7 พิจารณาความเหมาะสมของรูปแบบ โดยพิจารณาสหสัมพันธ์ในตัวเองของ e_t จะเห็นว่า $r_k(e_t)$, $k = 1, 2, \dots, 16$ อยู่ในช่วง $(-2s[r_k(e_t)], 2s[r_k(e_t)])$ และค่าสถิติ Box-Ljung Chi-square มีค่า p-value เท่ากับ 0.248 ซึ่งมากกว่า 0.05 นั่นคือยอมรับได้ว่า สหสัมพันธ์ในตัวเองของ e_t ไม่ต่างจากศูนย์ แสดงว่าวิธีวินเตอร์รูปแบบการคูณนี้เหมาะสมแล้ว

ตารางที่ 8 ค่า MAD, MSE และ MAPE ของอนุกรมเวลาสำหรับค่าที่ได้จากการพยากรณ์
เปรียบเทียบกับค่าจริงด้วยวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์

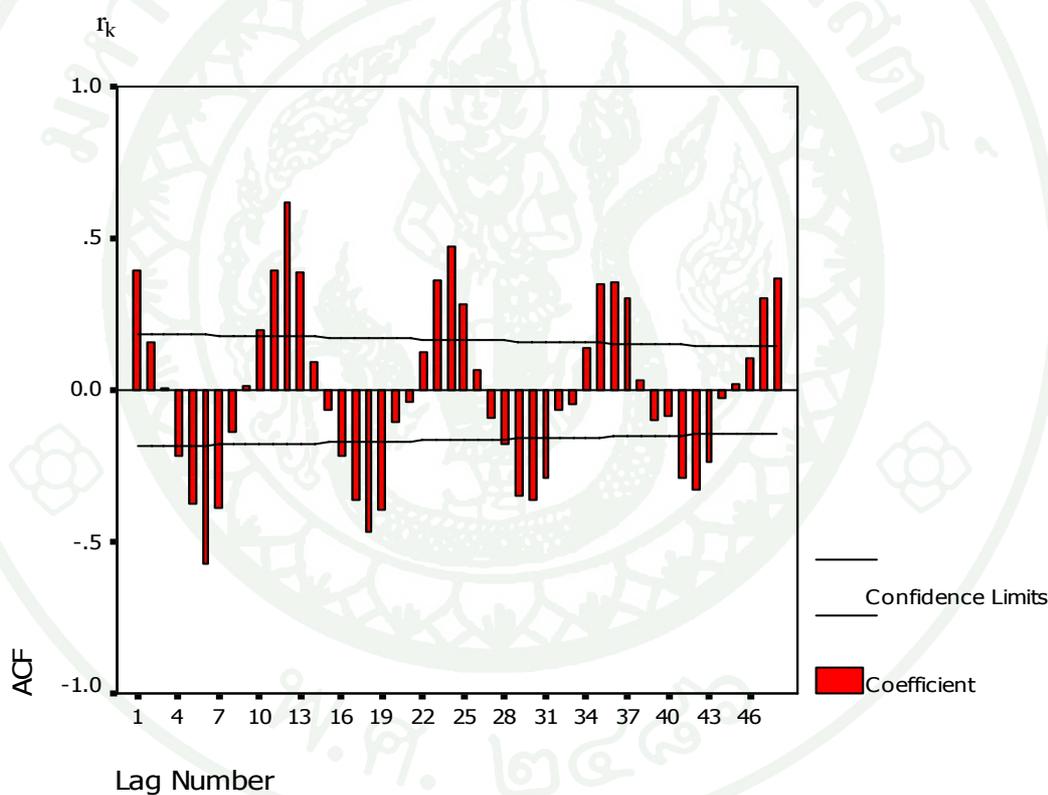
เดือน/ปี	ปริมาณน้ำฝนจริง (มิลลิเมตร)	ปริมาณน้ำฝนจากการพยากรณ์ (มิลลิเมตร)	MAD	MSE	MAPE
มิ.ย. 51	74.10	116.30	42.20	1,780.84	56.95
ก.ค. 51	81.70	144.79	63.09	3,980.34	77.22
ส.ค. 51	100.00	188.91	88.91	7,904.99	88.91
ก.ย. 51	191.60	254.44	62.84	3,948.87	32.80
ต.ค. 51	160.40	92.50	67.90	4,610.41	42.33
พ.ย. 51	37.70	23.94	13.76	189.34	36.50
ธ.ค. 51	12.30	2.93	9.37	87.80	76.18
ม.ค. 52	0.00	3.04	3.04	9.24	-
ก.พ. 52	0.00	7.56	7.56	57.15	-
มี.ค. 52	26.90	15.55	11.35	128.82	42.19
เม.ย. 52	55.20	94.23	39.03	1,523.34	70.71
พ.ค. 52	237.70	208.49	29.21	853.22	12.29
ค่าเฉลี่ย			36.52	2,089.53	53.61

จากตารางที่ 8 แสดงปริมาณน้ำฝนจริงและปริมาณน้ำฝนจากการพยากรณ์ด้วยวิธีของวินเตอร์ ได้ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) มีค่าเท่ากับ 36.52 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าเท่ากับ 2,089.53 และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) มีค่าเท่ากับ 53.61

2.3 วิธีบ็อกซ์และเจนกินส์ (Box – Jenkins Method)

ขั้นตอนการใช้วิธีบ็อกซ์และเจนกินส์ มีดังนี้

1) พิจารณาว่าอนุกรมเวลาเป็นสเตชันนารีหรือไม่ โดยพิจารณาจากคอเรลโลแกรม (Correlogram) จากกราฟ ACF (Autocorrelation Function: r_k) คือ การพิจารณาความสัมพันธ์ในตัวเอง ในลักษณะที่เป็นฟังก์ชันของช่วงเวลาที่ห่างกัน k หน่วย (มุกดา, 2549: 157) แสดงดังภาพที่ 15



ภาพที่ 15 ACF ของปริมาณน้ำฝนรายเดือนของจังหวัดลำปาง

ตารางที่ 9 ค่า r_k และคอเรลโลแกรมของค่า r_k ของวิธีวิธีบ็อกซ์และเจนกินส์

Autocorrelations: TOTALRAI TOTALRAIN

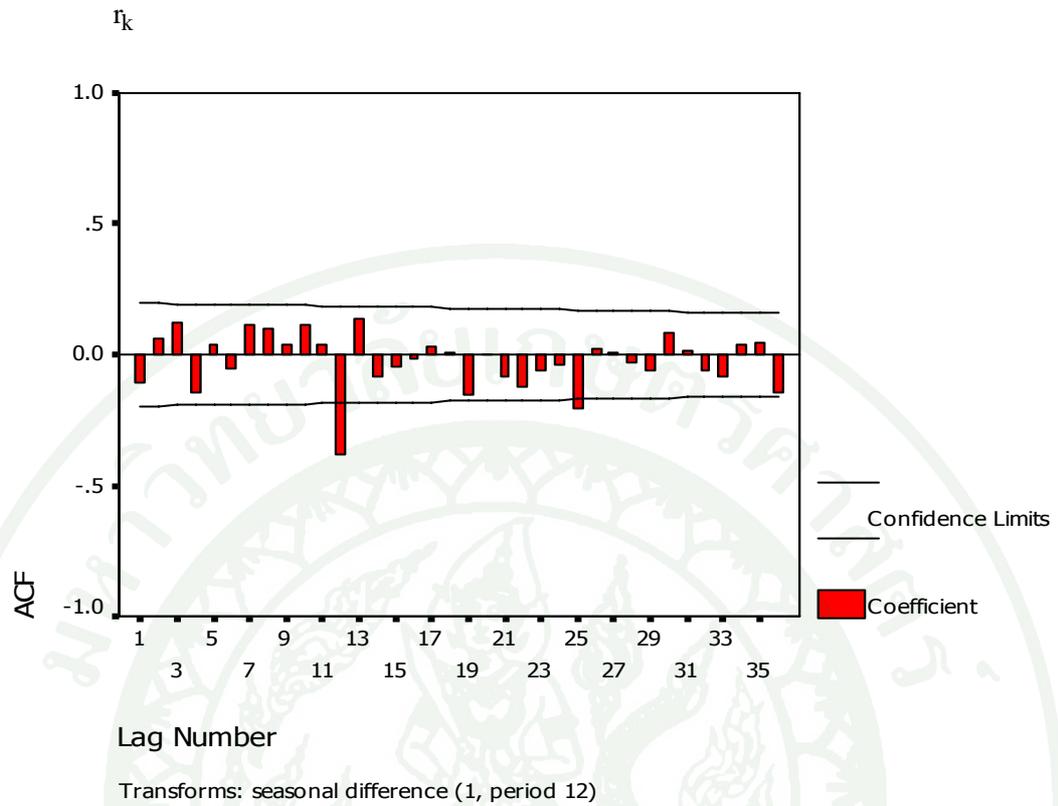
Lag	Corr. (r_k)	Err.		Box-Ljung	Prob.
1	.392	.093	17.806	.000
2	.160	.092	20.803	.000
3	.004	.092	20.805	.000
4	-.218	.092	26.462	.000
5	-.375	.091	43.360	.000
6	-.572	.091	83.163	.000
7	-.388	.090	101.613	.000
8	-.136	.090	103.897	.000
9	.014	.089	103.923	.000
10	.197	.089	108.843	.000
11	.396	.089	128.822	.000
12	.615	.088	177.560	.000
13	.387	.088	196.974	.000
14	.095	.087	198.152	.000
15	-.064	.086	198.688	.000
16	-.217	.086	204.975	.000
17	-.365	.086	222.967	.000
18	-.468	.086	252.870	.000
19	-.394	.085	274.313	.000
20	-.106	.085	275.870	.000
21	-.042	.084	276.119	.000
22	.122	.084	278.236	.000
23	.361	.083	297.072	.000
24	.471	.083	329.434	.000
25	.282	.082	341.142	.000
26	.066	.082	341.787	.000
27	-.092	.081	343.062	.000
28	-.180	.081	347.991	.000
29	-.350	.080	366.917	.000
30	-.363	.080	387.491	.000
31	-.292	.079	401.036	.000
32	-.065	.079	401.707	.000
33	-.044	.078	402.016	.000
34	.140	.078	405.239	.000
35	.347	.077	425.302	.000
36	.353	.077	446.378	.000
37	.300	.076	461.724	.000
38	.035	.076	461.934	.000
39	-.096	.075	463.541	.000
40	-.084	.075	464.785	.000
41	-.287	.074	479.663	.000

ตารางที่ 9 (ต่อ)

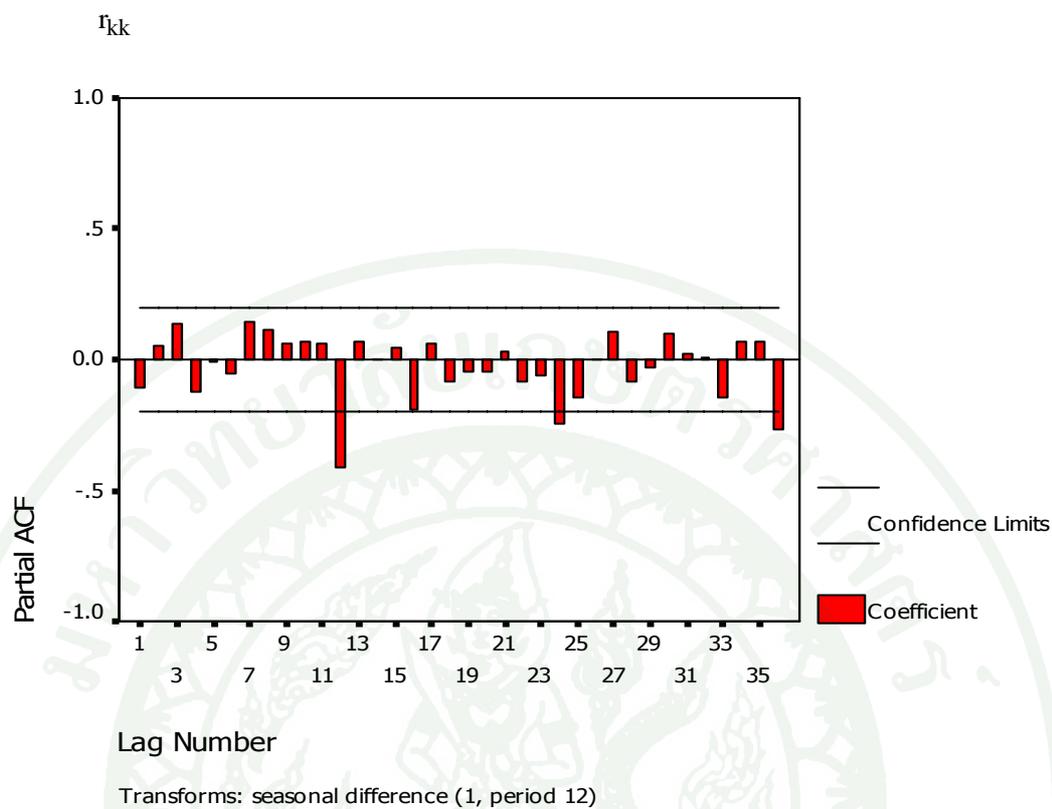
Autocorrelations: TOTALRAI TOTALRAIN

Lag	Corr. (r_k)	Err.	.-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
42	-.327	.074			****	**		.				499.197	.000
43	-.237	.073			**	**		.				509.661	.000
44	-.025	.073			.	*		.				509.779	.000
45	.023	.072			.	*		.				509.878	.000
46	.105	.072			.	**		.				512.017	.000
47	.305	.071			.	**	****					530.344	.000
48	.365	.071			.	**	****					557.023	.000

จากภาพที่ 12 เมื่อพิจารณาการเคลื่อนไหวของข้อมูลจะเห็นว่าข้อมูลอนุกรมเวลาปริมาณน้ำฝนรายเดือนไม่มีลักษณะเป็นสเตชันนารี เพราะว่าการเคลื่อนไหวของข้อมูลมีความผันแปรตามฤดูกาลเข้ามาเกี่ยวข้อง และพิจารณาค่า r_k และคอเรลโลแกรมของ r_k ในภาพที่ 15 และตารางที่ 9 จะเห็นว่าค่าดังกล่าวมีการเคลื่อนไหวเป็นลูกคลื่น และจะมีลักษณะซ้ำกันในทุกๆ 12 ค่าของค่า k ดังนั้น เมื่ออนุกรมเวลาชุดนี้ไม่เป็นสเตชันนารี เราต้องแปลงอนุกรมชุดนี้ให้เป็นอนุกรมเวลาชุดใหม่ที่ไม่มีอิทธิพลของความผันแปรตามฤดูกาลเข้ามาเกี่ยวข้อง โดยการหาค่าผลต่างของฤดูกาลแสดงได้ดังภาพที่ 16 และ 17



ภาพที่ 16 ACF ของปริมาณน้ำฝนรายเดือนของจังหวัดลำปางที่แปลงข้อมูลด้วยการ
หาผลต่างฤดูกาล 12 เดือน



ภาพที่ 17 PACF ของปริมาณน้ำฝนรายเดือนของจังหวัดลำปางที่แปลงข้อมูลด้วยการหาผลต่างฤดูกาล 12 เดือน

ตารางที่ 10 ค่า $r_k(z_t)$ และคอเรลโลแกรมของค่า $r_k(z_t)$ ของวิธีวิธีบ็อกซ์และเจนกินส์

Autocorrelations: TOTALRAI TOTALRAIN

Transformations: seasonal difference (1 at 12)

Auto- Stand.										Box-Ljung	Prob.		
Lag	Corr. $r_k(z_t)$	Err.	-.1	-.75	-.5	-.25	0	.25	.5	.75	1		
1	-.107	.098				**		.				1.181	.277
2	.065	.098				.		*	.			1.621	.445
3	.126	.097				.		***	.			3.300	.348
4	-.145	.097				***		.				5.551	.235
5	.036	.096				.		*	.			5.688	.338
6	-.054	.096				.		*	.			6.012	.422
7	.112	.095				.		**	.			7.394	.389
8	.100	.095				.		**	.			8.510	.385
9	.037	.094				.		*	.			8.661	.469
10	.118	.094				.		**	.			10.255	.418
11	.040	.093				.		*	.			10.438	.491
12	-.384	.092	****	****		.		.				27.694	.006
13	.135	.092				.		***	.			29.839	.005
14	-.080	.091				.		**	.			30.608	.006
15	-.045	.091				.		*	.			30.856	.009
16	-.015	.090				.		*	.			30.885	.014
17	.028	.090				.		*	.			30.982	.020
18	.010	.089				.		*	.			30.994	.029
19	-.149	.089				***		.				33.819	.019
20	.002	.088				.		*	.			33.820	.027
21	-.083	.088				.		**	.			34.722	.030
22	-.124	.087				.		**	.			36.742	.025
23	-.063	.087				.		*	.			37.272	.030
24	-.038	.086				.		*	.			37.465	.039
25	-.208	.085				**		**	.			43.366	.013
26	.025	.085				.		*	.			43.453	.017
27	.006	.084				.		*	.			43.457	.024
28	-.030	.084				.		*	.			43.585	.031
29	-.064	.083				.		*	.			44.175	.035
30	.086	.083				.		**	.			45.252	.037
31	.017	.082				.		*	.			45.294	.047
32	-.063	.081				.		*	.			45.889	.053
33	-.086	.081				**		.				47.009	.054
34	.037	.080				.		*	.			47.220	.065
35	.045	.080				.		*	.			47.539	.077
36	-.146	.079				***		.				50.970	.050

ตารางที่ 11 ค่า $r_{kk}(z_t)$ และคอเรลโลแกรมของค่า $r_{kk}(z_t)$ ของวิธีวิธีบ็อกซ์และเจนกินส์

Partial Autocorrelations: TOTALRAI TOTALRAIN

Transformations: seasonal difference (1 at 12)

Pr-Aut- Stand.														
Lag	Corr. $r_{kk}(z_t)$	Err.	-.1 -.75 -.5 -.25 0 .25 .5 .75 1											
1	-.107	.100	.	**
2	.054	.100	.	.	*
3	.140	.100
4	-.125	.100	.	.	**
5	-.009	.100	.	.	.	*
6	-.051	.100	.	.	.	*
7	.142	.100
8	.116	.100
9	.058	.100	.	.	.	*
10	.070	.100	.	.	.	*
11	.063	.100	.	.	.	*
12	-.412	.100
13	.065	.100
14	.001	.100
15	.045	.100
16	-.193	.100
17	.058	.100
18	-.085	.100
19	-.047	.100	.	.	.	*
20	-.047	.100	.	.	.	*
21	.030	.100	.	.	.	*
22	-.081	.100
23	-.061	.100	.	.	.	*
24	-.246	.100	.	.	.	*
25	-.142	.100
26	-.001	.100
27	.104	.100
28	-.084	.100
29	-.031	.100	.	.	.	*
30	.100	.100
31	.027	.100
32	.004	.100
33	-.141	.100
34	.067	.100
35	.066	.100
36	-.270	.100

2) การกำหนดรูปแบบ จากกราฟ ACF และ PACF ที่ได้จากข้อ 1) และค่า $r_k(z_t)$ และ $r_{kk}(z_t)$ จากการแปลงข้อมูลด้วยผลต่างฤดูกาล 12 เดือน ในตารางที่ 10 และ 11 จะเห็นว่า มีอนุกรมเวลา z_t มีลักษณะเป็นสเตรชันนารี และค่า $r_k(z_t)$ และ $r_{kk}(z_t)$ ส่วนใหญ่จะมีค่าเข้าใกล้ 0 เมื่อค่า k มีค่ามากขึ้น และค่า r_1 และ r_{11} ไม่เท่ากับ 0 ดังนั้นรูปแบบ ARIMA(p,d,q)(P,D,Q)_s ที่คาดว่าน่าจะเหมาะสมกับอนุกรมเวลาชุดนี้คือ ARIMA(0,0,0)(1,1,0)₁₂, ARIMA(0,0,0)(0,1,1)₁₂, ARIMA(0,0,0)(1,1,1)₁₂, ARIMA(0,0,0)(2,1,0)₁₂, ARIMA(0,0,0)(2,1,1)₁₂ และ ARIMA(0,0,0)(2,1,2)₁₂

จากนั้นพิจารณาค่า AIC (เงื่อนไขของ Akaike), ค่า SBC (เงื่อนไขของ Schwartz) และ SSE จากรูปแบบทั้งหมดที่ปรากฏในผลการวิเคราะห์ของโปรแกรมสำเร็จรูป ได้ดังตารางที่ 13

ตารางที่ 12 ค่า AIC, ค่า SBC และ SSE ของรูปแบบ

รูปแบบ	AIC	SBC	SSE
ARIMA(0,0,0)(1,1,0) ₁₂	1,149.2453	1,154.4755	496,979.81
ARIMA(0,0,0)(0,1,1) ₁₂	1,124.704	1,129.9342	377,325.37
ARIMA(0,0,0)(1,1,1) ₁₂	1,125.5898	1,133.5342	379,843.94
ARIMA(0,0,0)(2,1,0) ₁₂	1,144.229	1,152.0743	463,587.09
ARIMA(0,0,0)(2,1,1) ₁₂	1,127.2806	1,137.7411	375,052.33
ARIMA(0,0,0)(2,1,2) ₁₂	1,129.5129	1,142.5885	376,581.09

ตารางที่ 13 ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(1,1,0)₁₂

พารามิเตอร์	ค่าประมาณ	ค่าความคลาดเคลื่อน	ค่าสถิติ t	P-value
AR1+Season	-0.40784753	0.0945312	-4.3144204	0.00003793
CONSTANT	0.37691232	5.1336856	0.0734194	0.94162049

ตารางที่ 14 ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(0,1,1)₁₂

พารามิเตอร์	ค่าประมาณ	ค่าความคลาดเคลื่อน	ค่าสถิติ t	P-value
MA1+Season	0.9981394	12.905297	0.07734339	0.93850647
CONSTANT	1.1910468	1.872982	0.63590940	0.52630282

ตารางที่ 15 ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(1,1,1)₁₂

พารามิเตอร์	ค่าประมาณ	ค่าความคลาดเคลื่อน	ค่าสถิติ t	P-value
AR1+Season	0.1333482	0.1633308	0.8164304	0.41623419
MA1+Season	0.9563946	0.9194282	1.0402058	0.30080383
CONSTANT	1.2736411	2.1233369	0.4821859	0.63074906

ตารางที่ 16 ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(2,1,0)₁₂

พารามิเตอร์	ค่าประมาณ	ค่าความคลาดเคลื่อน	ค่าสถิติ t	P-value
AR1+Season	-0.5373905	0.1023604	-5.2499817	0.00000088
AR2+Season	-0.2885848	0.1115293	-2.5875256	0.01113456
CONSTANT	1.2736411	3.9665124	0.3210985	0.74881954

ตารางที่ 17 ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(2,1,1)₁₂

พารามิเตอร์	ค่าประมาณ	ค่าความคลาดเคลื่อน	ค่าสถิติ t	P-value
AR1+Season	0.1350614	0.212971	0.63417904	0.52745676
AR2+Season	-0.0848091	0.153402	-0.55285447	0.58163493
MA1+Season	0.9968314	19.099460	0.05219160	0.95848335
CONSTANT	1.0421609	2.016599	0.51679125	0.60647820

ตารางที่ 18 ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(2,1,2)₁₂

พารามิเตอร์	ค่าประมาณ	ค่าความคลาดเคลื่อน	ค่าสถิติ t	P-value
AR1+Season	-0.8935732	3.963122	-0.22547202	0.82209138
AR2+Season	0.0432750	0.178192	0.24285604	0.80863506
MA1+Season	0.0691857	23.516323	0.00294203	0.99765871
MA2+Season	0.9268687	18.731044	0.04948302	0.96063717
CONSTANT	1.1269227	1.960368	0.57485268	0.56673571

พิจารณาตารางที่ 13 รูปแบบ ARIMA(0,0,0)(1,1,0)₁₂ จะได้ค่า P-value = 0.00003793 (สำหรับ ϕ_1) ซึ่งน้อยกว่าค่าระดับนัยสำคัญ 0.05 และ P-value = 0.94162049 (สำหรับ θ_0) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 นั่นคือ เฉพาะค่า P-value ของค่าคงที่ θ_0 เท่านั้นที่ยอมรับสมมติฐาน $H_0 : \theta_0 = 0$ ฉะนั้นต้องกำหนดรูปแบบนี้ใหม่โดยไม่ให้มีค่าคงที่ θ_0

จากนั้นพิจารณาตารางที่ 14 รูปแบบ ARIMA(0,0,0)(0,1,1)₁₂ จะได้ค่า P-value = 0.93850647 (สำหรับ θ_1) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 และ P-value = 0.52630282 (สำหรับ θ_0) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 นั่นคือ รูปแบบนี้ไม่เหมาะสม

จากนั้นพิจารณาตารางที่ 15 รูปแบบ ARIMA(0,0,0)(1,1,1)₁₂ จะได้ค่า P-value = 0.41623419 (สำหรับ ϕ_1) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 P-value = 0.30080383 (สำหรับ θ_1) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 และ P-value = 0.63074906 (สำหรับ θ_0) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 นั่นคือ รูปแบบนี้ไม่เหมาะสม

จากนั้นพิจารณาตารางที่ 16 รูปแบบ ARIMA(0,0,0)(2,1,0)₁₂ จะได้ค่า P-value = 0.00000088 (สำหรับ ϕ_1) ซึ่งน้อยกว่าค่าระดับนัยสำคัญ 0.05 P-value = 0.01113456 (สำหรับ ϕ_2) ซึ่งน้อยกว่าค่าระดับนัยสำคัญ 0.05 และ P-value = 0.74881954 (สำหรับ θ_0) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 นั่นคือ เฉพาะค่า P-value ของค่าคงที่ θ_0 เท่านั้นที่ยอมรับสมมติฐาน $H_0 : \theta_0 = 0$ ฉะนั้นต้องกำหนดรูปแบบนี้ใหม่โดยไม่ให้มีค่าคงที่ θ_0

จากนั้นพิจารณาตารางที่ 17 รูปแบบ ARIMA(0,0,0)(2,1,1)₁₂ จะได้ว่า P-value = 0.52745676 (สำหรับ ϕ_1) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 P-value = 0.58163493 (สำหรับ ϕ_2) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 P-value = 0.95848335 (สำหรับ θ_1) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 และ P-value = 0.60647820 (สำหรับ θ_0) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 นั่นคือ รูปแบบนี้ไม่เหมาะสม

จากนั้นพิจารณาตารางที่ 18 รูปแบบ ARIMA(0,0,0)(2,1,2)₁₂ จะได้ว่า P-value = 0.82209138 (สำหรับ ϕ_1) ซึ่งน้อยกว่าค่าระดับนัยสำคัญ 0.05 จะได้ว่า P-value = 0.80863506 (สำหรับ ϕ_2) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 P-value = 0.99765871 (สำหรับ θ_1) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 P-value = 0.96063717 (สำหรับ θ_2) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 และ P-value = 0.56673571 (สำหรับ θ_0) ซึ่งมากกว่าค่าระดับนัยสำคัญ 0.05 นั่นคือ รูปแบบนี้ไม่เหมาะสม

จากนั้นพิจารณารูปแบบใหม่ โดยตัดค่าคงที่ θ_0 ในรูปแบบออก และจะได้ค่า AIC, SBC กับค่า SSE ดังตารางที่ 19 โดยจะเลือกรูปแบบที่ให้ค่า AIC, ค่า SBC กับค่า SSE ต่ำสุด

ตารางที่ 19 ค่า AIC, ค่า SBC และ SSE ของรูปแบบที่ตัดค่าคงที่ θ_0 ในรูปแบบออก

รูปแบบ	AIC	SBC	SSE
ARIMA(0,0,0)(1,1,0) ₁₂	1,147.2575	1,149.8726	4,97006.81
ARIMA(0,0,0)(2,1,0) ₁₂	1,142.3514	1,147.5816	464,068.79

พิจารณาค่าในตารางที่ 19 จะเลือกรูปแบบที่ให้ค่า AIC, ค่า SBC กับค่า SSE ต่ำสุด ดังนั้นจะได้รูปแบบ ARIMA(0,0,0)(2,1,0)₁₂ เป็นรูปแบบที่ดีที่สุด และจากการวิเคราะห์ด้วยโปรแกรม SPSS จะได้ว่าพารามิเตอร์ของรูปแบบ ดังตารางที่ 20

ตารางที่ 20 ค่าประมาณของสัมประสิทธิ์ของรูปแบบ ARIMA(0,0,0)(2,1,0)₁₂ ที่ตัดค่าคงที่ θ_0 ในรูปแบบออก

พารามิเตอร์	ค่าประมาณ	ค่าความคลาดเคลื่อน	ค่าสถิติ t	P-value
AR1+Season	-.53470359	0.10159822	-5.2629229	0.00000082
AR2+Season	-0.28534493	0.11055130	-2.5811088	0.01131455

3) ประมาณค่าพารามิเตอร์ จากรูปแบบที่เหมาะสมในข้อ 2) ด้วยวิธีกำลังสองน้อยที่สุด แสดงดังตารางที่ 20

จะได้รูปแบบ ARIMA(0,0,0)(2,1,0)₁₂ คือ

$$(1 - B^{12})^1 (1 - \phi_1 B^{12} - \phi_2 B^{24}) Y_t = \varepsilon_t$$

$$Y_t - B^{12} Y_t - \phi_1 B^{12} Y_t + \phi_1 B^{24} Y_t - \phi_2 B^{24} Y_t + \phi_2 B^{36} Y_t = \varepsilon_t$$

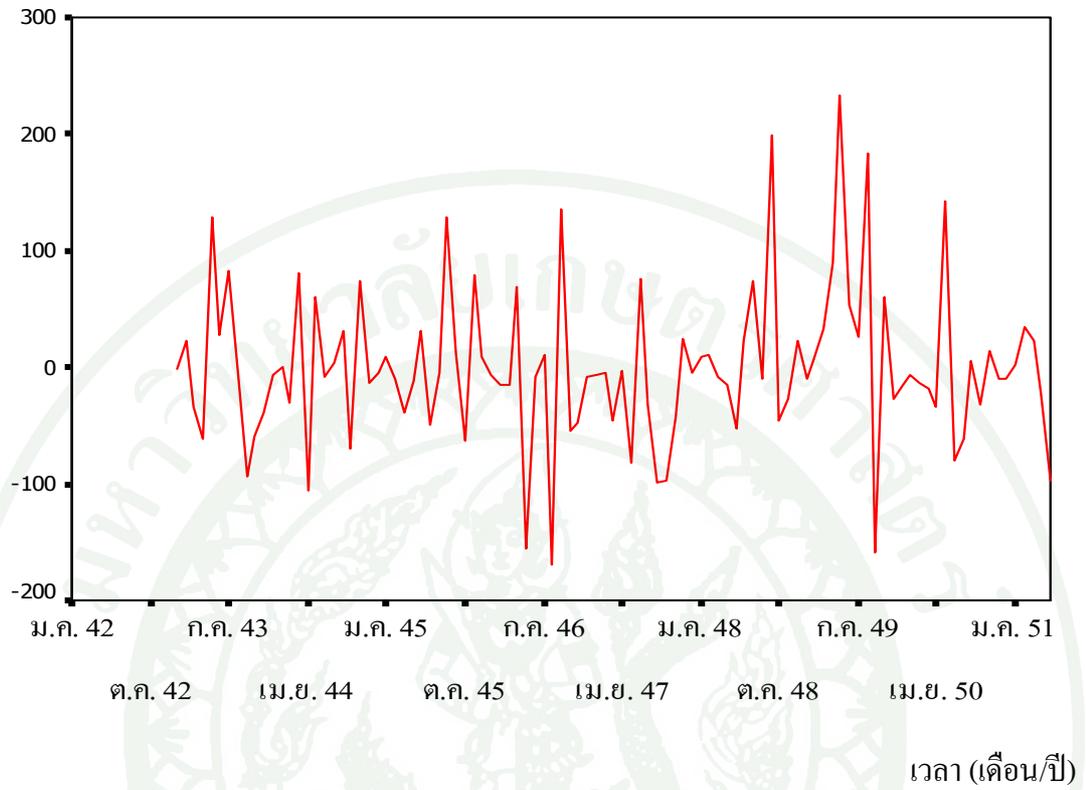
$$Y_t - Y_{t-12} - \phi_1 Y_{t-12} + \phi_1 Y_{t-24} - \phi_2 Y_{t-24} + \phi_2 Y_{t-36} = \varepsilon_t$$

$$Y_t - Y_{t-12} + 0.5347 Y_t - 0.5347 Y_{t-24} + 0.2853 Y_{t-24} - 0.2853 Y_{t-36} = \varepsilon_t$$

$$Y_t = Y_{t-12} - 0.5347 Y_t + 0.5347 Y_{t-24} - 0.2853 Y_{t-24} + 0.2853 Y_{t-36} + \varepsilon_t$$

4) การตรวจสอบความเหมาะสมของรูปแบบ เมื่อทดสอบค่าความคลาดเคลื่อน โดยพิจารณาความคลาดเคลื่อนกับเวลา ดังภาพที่ 18 พบว่าค่าความคลาดเคลื่อนกระจายอยู่รอบค่าศูนย์ แสดงว่าความคลาดเคลื่อนมีค่าเฉลี่ยเท่ากับศูนย์ และความแปรปรวนคงที่ (วารุทธิ์, 2550: 147)

ความคลาดเคลื่อนจากวิธีบ็อกซ์และเงินสิน



ภาพที่ 18 กราฟแสดงความคลาดเคลื่อนจากวิธีของบ็อกซ์และเงินสิน

ตารางที่ 21 ค่า $r_k(e_t)$ และคอเรลโลแกรมของค่า $r_k(e_t)$ ของวิธีวิธีบ็อกซ์และเจนกินส์

Autocorrelations: ERR_1 Error for TOTALRAI from ARIMA, MOD_1 CON

Auto- Stand.

Lag	Corr. $r_k(e_t)$	Err.	-.1 -.75 -.5 -.25 0 .25 .5 .75 1							Box-Ljung	Prob.
1	-.154	.098			***		.			2.470	.116
2	.087	.098			.		**			3.271	.195
3	.179	.097			.		****			6.663	.083
4	-.193	.097			****		.			10.673	.030
5	.108	.096			.		**			11.948	.036
6	-.093	.096			.		**			12.888	.045
7	.117	.095			.		**			14.393	.045
8	.020	.095			.		*			14.438	.071
9	.054	.094			.		*			14.764	.098
10	.127	.094			.		****			16.621	.083
11	-.041	.093			.		*			16.812	.114
12	-.087	.092			.		**			17.696	.125
13	.025	.092			.		*			17.771	.166
14	-.029	.091			.		*			17.874	.213
15	.018	.091			.		*			17.913	.267
16	-.073	.090			.		*			18.566	.292
17	.058	.090			.		*			18.990	.329
18	-.012	.089			.		*			19.009	.391
19	-.124	.089			.		**			20.975	.338
20	.033	.088			.		*			21.114	.390
21	-.147	.088			.		****			23.940	.296
22	-.082	.087			.		**			24.824	.306
23	-.025	.087			.		*			24.905	.355
24	-.162	.086			.		***			28.463	.241

จากนั้นทำการทดสอบความเหมาะสมของรูปแบบ ARIMA(0,0,0)(2,1,0)₁₂ กับอนุกรมเวลาชุดนี้ โดยการทดสอบสมมติฐาน ดังนี้

1. ทดสอบว่าพารามิเตอร์ (ϕ) ในรูปแบบมีค่าเท่ากับ 0 หรือไม่

1.1 กำหนด $H_0 : \phi_1 = 0$ vs $H_1 : \phi_1 \neq 0$

ผลการทดสอบได้ค่า P-value = 0.00000088 (แสดงดังตารางที่ 20)

P-value มีค่าน้อยกว่าระดับนัยสำคัญคือ 0.05 ปฏิเสธสมมติฐานหลัก

1.2 กำหนด $H_0 : \phi_2 = 0$ vs $H_1 : \phi_2 \neq 0$

ผลการทดสอบได้ค่า P-value = 0.0113456 (แสดงดังตารางที่ 20)

P-value มีค่าน้อยกว่าระดับนัยสำคัญคือ 0.05 ปฏิเสธสมมติฐานหลัก

สรุปว่ารูปแบบ ARIMA(0,0,0)(2,1,0)₁₂ เหมาะสมแล้ว

2. ทดสอบว่าค่าความคลาดเคลื่อนของการพยากรณ์ e_t มีการแจกแจงที่เป็นอิสระกันหรือไม่ โดยที่หากจาก e_t ซึ่งเป็นค่าประมาณของ ε_t คำนวณได้จาก $e_t = Y_t - \hat{Y}_t$ โดยกำหนดสมมติฐาน ดังนี้

$$H_0 : \rho_1(e_t) = \rho_2(e_t) = \dots = \rho_{24}(e_t) = 0$$

$$H_1 : \rho_i(e_t) \text{ อย่างน้อยหนึ่งค่าที่ไม่เท่ากับ } 0$$

จากตารางที่ 21 เมื่อ $k = 24$ ได้ P-value = 0.241 มีค่ามากกว่าระดับนัยสำคัญที่กำหนดคือ 0.05 หรือพิจารณาตัวสถิติ $Q_{24} = 28.463$ ซึ่งน้อยกว่าค่าวิกฤตคือ $\chi^2_{0.05,21} = 32.6705$ สรุปว่ายอมรับ H_0 นั่นคือ ค่า e_t เป็นอิสระต่อกัน หรือรูปแบบ ARIMA(0,0,0)(2,1,0)₁₂ เหมาะสมแล้ว

ผลจากการวิเคราะห์ข้อมูลแสดงดังตารางที่ 22

ตารางที่ 22 ค่า MAD, MSE และ MAPE ของอนุกรมเวลาสำหรับค่าที่ได้จากการพยากรณ์ เปรียบเทียบกับค่าจริงด้วยวิธีบ็อกซ์และเจนกินส์

เดือน/ปี	ปริมาณน้ำฝนจริง (มิลลิเมตร)	ปริมาณน้ำฝนจากการพยากรณ์ (มิลลิเมตร)	MAD	MSE	MAPE
มิ.ย. 51	74.10	118.23	44.13	1,947.46	59.55
ก.ค. 51	81.70	145.48	63.78	4,067.89	78.07
ส.ค. 51	100.00	206.28	106.28	11,295.44	106.28

ตารางที่ 22 (ต่อ)

เดือน/ปี	ปริมาณน้ำฝนจริง (มิลลิเมตร)	ปริมาณน้ำฝนจากการพยากรณ์ (มิลลิเมตร)	MAD	MSE	MAPE
ก.ย. 51	191.60	253.35	61.75	3,813.06	32.23
ต.ค. 51	160.40	70.18	90.22	8,139.65	56.25
พ.ย. 51	37.70	13.85	23.85	568.82	63.26
ธ.ค. 51	12.30	8.10	4.20	17.64	34.15
ม.ค. 52	0.00	4.37	4.37	19.10	-
ก.พ. 52	0.00	26.80	26.80	718.24	-
มี.ค. 52	26.90	32.32	5.42	29.38	20.15
เม.ย. 52	55.20	108.88	53.68	2,881.54	97.25
พ.ค. 52	237.70	246.47	8.77	76.91	3.69
ค่าเฉลี่ย			41.10	2,797.93	55.09

จากตารางที่ 22 แสดงปริมาณน้ำฝนจริงและปริมาณน้ำฝนจากการพยากรณ์ด้วยวิธีบ็อกซ์และเจนกินส์ (Box – Jenkins Method) ได้ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) มีค่าเท่ากับ 41.10 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าเท่ากับ 2,797.93 และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) มีค่าเท่ากับ 55.09

ตารางที่ 23 ค่า MAD, MSE และ MAPE ของวิธีการพยากรณ์เดี่ยวทั้งสามวิธี

วิธีการ	MAD	MSE	MAPE
วิธีแยกส่วนประกอบ	98.50	17,729.06	103.37
วิธีวินเตอร์	36.52	2,089.53	53.61
วิธีบ็อกซ์และเจนกินส์	41.10	2,797.93	55.09

จากตารางที่ 23 ทำการเปรียบเทียบประสิทธิภาพของความถูกต้องของการพยากรณ์สำหรับอนุกรมเวลาโดยใช้ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) ความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) พบว่าวิธีการพยากรณ์ที่ให้ค่า MAD, MSE และ MAPE ต่ำที่สุด คือ วิธีของวินเตอร์

3. ผลการวิเคราะห์ข้อมูลด้วยวิธีการพยากรณ์ร่วม

งานวิจัยนี้เลือกวิธีเฉลี่ยถ่วงน้ำหนักซึ่งน้ำหนักผูกพันกับค่าผลบวกกำลังสองของความคลาดเคลื่อน

Delurgio (1998) อาศัยจากหลักการที่ว่าความถูกต้องของการพยากรณ์จะสูง ถ้าความคลาดเคลื่อนของการพยากรณ์ต่ำ และในทางกลับกันความถูกต้องของการพยากรณ์จะต่ำ ถ้าความคลาดเคลื่อนของการพยากรณ์สูง วิธีเฉลี่ยถ่วงน้ำหนักจึงใช้ค่าน้ำหนักของวิธีที่ i ซึ่งแปรผกผันกับค่า SSE_i นั่นคือความถูกต้องของวิธีที่ i จะแปรผกผันกับค่า SSE_i และเพื่อให้ w_i มี $\sum w_i = 1$

ดังนั้น

$$w_i = \frac{1/SSE_i}{1/SSE_1 + 1/SSE_2 + \dots + 1/SSE_p} \quad (1)$$

ซึ่ง

$$SSE_i = \sum_{t=1}^n e_{it}^2, \quad i = 1, 2, \dots, p$$

รูปแบบที่ใช้ในการพยากรณ์ร่วม (Combined Forecast: CF) มีรูปแบบคือ

$$\hat{Y}_{CF} = w_1 \hat{Y}^{(1)} + w_2 \hat{Y}^{(2)} + \dots + w_p \hat{Y}^{(p)}$$

โดยที่ w_i คือ ตัวถ่วงน้ำหนักของวิธีพยากรณ์, $i=1, 2, \dots, p$

\hat{Y}_{CF} คือ ค่าพยากรณ์

p คือ วิธีการพยากรณ์ที่ p

t คือ คาบเวลา ; $t=1, 2, 3, \dots, n$

ตารางที่ 24 การหาค่าความคลาดเคลื่อนของการพยากรณ์ด้วยวิธีแยกส่วนประกอบ วิธีวินเตอร์ และวิธีบ็อกซ์และเจนกินส์

เดือน/ปี	ปริมาณ น้ำฝน จริง(มิลลิเมตร)	ปริมาณน้ำฝนที่ได้จากการพยากรณ์ด้วยวิธี			e_1	e_1^2	e_2	e_2^2	e_3	e_3^2
		แยก ส่วนประกอบ	วินเตอร์	บ็อกซ์และเจนกินส์						
มิ.ย. 51	74.10	216.16	116.30	118.23	-142.06	20,181.04	-42.20	1,780.84	-44.13	1,947.46
ก.ค. 51	81.70	272.02	144.79	145.48	-190.32	36,221.70	-63.09	3,980.35	-63.78	4,067.89
ส.ค. 51	100.00	360.18	188.91	206.28	-260.18	67,693.63	-88.91	7,904.99	-106.28	11,295.44
ก.ย. 51	191.60	400.19	254.44	253.35	-208.59	43,509.79	-62.84	3,948.87	-61.75	3,813.06
ต.ค. 51	160.40	162.61	92.50	70.18	-2.21	4.88	67.90	4,610.41	90.22	8,139.65
พ.ย. 51	37.70	27.89	23.94	13.85	9.81	96.24	13.76	189.34	23.85	568.82
ธ.ค. 51	12.30	0.00	2.93	8.10	12.30	151.29	9.37	87.80	4.20	17.64
ม.ค. 52	0.00	2.02	3.04	4.37	-2.02	4.08	-3.04	9.24	-4.37	19.10
ก.พ. 52	0.00	8.07	7.56	26.80	-8.07	65.12	-7.56	57.15	-26.80	718.24
มี.ค. 52	26.90	23.60	15.55	32.32	3.30	10.89	11.35	128.82	-5.42	29.38
เม.ย. 52	55.20	172.18	94.23	108.88	-116.98	13,684.32	-39.03	1,523.34	-53.68	2,881.54
พ.ค. 52	237.70	463.81	208.49	246.47	-226.11	51,125.73	29.21	853.22	-8.77	76.91

หมายเหตุ: e_1 , e_2 และ e_3 คือ ค่าความคลาดเคลื่อนของการพยากรณ์ด้วยวิธีแยกส่วนประกอบ วินเตอร์ และบ็อกซ์และเจนกินส์ ตามลำดับ

จากตารางที่ 24 จะได้ว่า

$$SSE_1 = \sum_{t=1}^{12} e_{1t}^2 = 20,181.04 + 36,221.70 + \dots + 51,125.73 = 232,748.72$$

$$SSE_2 = \sum_{t=1}^{12} e_{2t}^2 = 1,780.84 + 3,980.35 + \dots + 853.22 = 25,074.37$$

$$SSE_3 = \sum_{t=1}^{12} e_{3t}^2 = 2,167.83 + 4,403.65 + \dots + 132.25 = 33,575.13$$

$$\begin{aligned} w_1 &= \frac{1/SSE_1}{1/SSE_1 + 1/SSE_2 + 1/SSE_3} \\ &= \frac{1/232,748.72}{1/232,748.72 + 1/25,074.37 + 1/33,575.13} \\ &= 0.0570 \end{aligned}$$

$$\begin{aligned} w_2 &= \frac{1/SSE_2}{1/SSE_1 + 1/SSE_2 + 1/SSE_3} \\ &= \frac{1/25,074.37}{1/232,748.72 + 1/25,074.37 + 1/33,575.13} \\ &= 0.5400 \end{aligned}$$

$$\begin{aligned} w_3 &= \frac{1/SSE_3}{1/SSE_1 + 1/SSE_2 + 1/SSE_3} \\ &= \frac{1/33,575.13}{1/232,748.72 + 1/25,074.37 + 1/33,575.13} \\ &= 0.4030 \end{aligned}$$

สำหรับงานวิจัยนี้คำนวณค่าตัวถ่วงน้ำหนักจากสมการที่ 1 สรุปได้ ดังนี้

$$w_1 = 0.0570, w_2 = 0.5400 \text{ และ } w_3 = 0.4030$$

โดยที่ w_1 คือ ตัวถ่วงน้ำหนักของวิธีแยกส่วนประกอบ

w_2 คือ ตัวถ่วงน้ำหนักของวิธีวินเตอร์

w_3 คือ ตัวถ่วงน้ำหนักของวิธีบ็อกซ์และเจนกินส์

นำมาสร้างสมการพยากรณ์คือ

$$\hat{Y}_{CF} = 0.0570\hat{Y}^{(1)} + 0.5400\hat{Y}^{(2)} + 0.4030\hat{Y}^{(3)}$$

โดยที่ $\hat{Y}^{(1)}$ คือ ค่าพยากรณ์ของวิธีแยกส่วนประกอบ

$\hat{Y}^{(2)}$ คือ ค่าพยากรณ์ของวิธีวินเตอร์

$\hat{Y}^{(3)}$ คือ ค่าพยากรณ์ของวิธีบ็อกซ์และเจนกินส์

ผลจากการวิเคราะห์ข้อมูล แสดงดังตารางที่ 24

ตารางที่ 25 ค่า MAD, MSE และ MAPE ของอนุกรมเวลาสำหรับค่าที่ได้จากการพยากรณ์
เปรียบเทียบกับค่าจริงด้วยวิธีการพยากรณ์ร่วม

เดือน/ปี	ปริมาณน้ำฝนจาก		MAD	MSE	MAPE
	ปริมาณน้ำฝนจริง (มิลลิเมตร)	การพยากรณ์ (มิลลิเมตร)			
มิ.ย. 51	74.10	122.77	48.67	2,368.77	65.68
ก.ค. 51	81.70	152.32	70.62	4,987.18	86.44
ส.ค. 51	100.00	205.67	105.67	11,166.15	105.67
ก.ย. 51	191.60	262.31	70.71	4,999.90	36.91
ต.ค. 51	160.40	87.50	72.90	5,314.41	45.45
พ.ย. 51	37.70	20.10	17.60	309.76	46.68
ธ.ค. 51	12.30	4.85	7.45	55.50	60.57
ม.ค. 52	0.00	3.52	3.52	12.39	-
ก.พ. 52	0.00	15.34	15.34	235.32	-
มี.ค. 52	26.90	22.77	4.13	17.06	15.35
เม.ย. 52	55.20	104.58	49.38	2,438.38	89.46
พ.ค. 52	237.70	238.35	0.65	0.42	0.27
			38.89	2,660.26	55.25

จากตารางที่ 25 แสดงปริมาณน้ำฝนจริงและปริมาณน้ำฝนจากการพยากรณ์ด้วยวิธีการ
พยากรณ์ร่วม ได้ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) มีค่าเท่ากับ 38.89 ค่าความคลาดเคลื่อนกำลังสอง
เฉลี่ย (MSE) มีค่าเท่ากับ 2,660.26 และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน
(MAPE) มีค่าเท่ากับ 55.25

ตารางที่ 26 การเปรียบเทียบวิธีการพยากรณ์ปริมาณน้ำฝนรายเดือนของจังหวัดลำปาง ตั้งแต่เดือนมิถุนายน 2551 ถึง เดือนพฤษภาคม 2552 ที่ได้จากการพยากรณ์ด้วยวิธีต่างๆ

วิธีการ	MAD	MSE	MAPE
วิธีแยกส่วนประกอบ	98.50	17,729.06	103.37
วิธีวินเตอร์	36.52	2,089.53	53.61
วิธีบ็อกซ์และเจนกินส์	42.40	2,916.03	54.84
วิธีการพยากรณ์ร่วม	38.89	2,660.26	55.25

จากการเปรียบเทียบค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) ของวิธีพยากรณ์ทั้งการพยากรณ์เดี่ยว และ พยากรณ์ร่วม จากตารางที่ 26 พบว่าวิธีวินเตอร์เหมาะสมในการพยากรณ์รูปแบบปริมาณน้ำฝนมากที่สุด เนื่องจากให้ค่า MAD, MSE และ MAPE ต่ำกว่าวิธีการพยากรณ์ทุกวิธี

สรุปและข้อเสนอแนะ

สรุป

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อสมการการพยากรณ์และเปรียบเทียบเทคนิคการพยากรณ์สำหรับพยากรณ์ปริมาณน้ำฝนในจังหวัดลำปาง โดยวิธีที่นำมาศึกษาเปรียบเทียบ ได้แก่ วิธีแยกส่วนประกอบ (Decomposition Method) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของ วินเตอร์ (Triple Exponential Smoothing : Winter's Method) วิธีบ็อกซ์และเจนกินส์ (Box – Jenkins Method) และวิธีการพยากรณ์ร่วม ใช้วิธีเฉลี่ยถ่วงน้ำหนักซึ่งน้ำหนักผกผันกับค่าผลบวกกำลังสองของค่าความคลาดเคลื่อน

ข้อมูลที่นำมาศึกษาคือ ข้อมูลปริมาณน้ำฝนจากจังหวัดลำปาง ตั้งแต่เดือนมกราคม 2542 ถึง เดือนพฤษภาคม 2552 แบ่งออกเป็น 2 ชุด ดังนี้

ชุดที่ 1 เป็นข้อมูลปริมาณน้ำฝนรายเดือน ตั้งแต่เดือนมกราคม 2542 ถึงเดือนพฤษภาคม 2551 รวมระยะเวลา 113 เดือน ใช้สำหรับการวิเคราะห์เพื่อสร้างตัวแบบพยากรณ์

ชุดที่ 2 เป็นข้อมูลปริมาณน้ำฝนของเดือนมิถุนายน 2551 ถึงเดือนพฤษภาคม 2552 ใช้สำหรับการตรวจสอบผลของการพยากรณ์ของตัวแบบพยากรณ์ที่สร้างขึ้นด้วยข้อมูลชุดที่ 1 โดยวิธีการพยากรณ์ต่างๆ

เปรียบเทียบปริมาณน้ำฝนที่เกิดขึ้นจริงและปริมาณน้ำฝนที่ได้จากการพยากรณ์ด้วยวิธีต่างๆ โดยใช้ค่า MAD, MSE และ MAPE สรุปผลได้ดังนี้

1. วิธีการพยากรณ์เดี่ยว

ผลการวิเคราะห์ข้อมูลปริมาณน้ำฝนของจังหวัดลำปางด้วยวิธีแยกส่วนประกอบ (Decomposition Method) ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) มีค่าเท่ากับ 98.50 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าเท่ากับ 17,729.06 และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) มีค่าเท่ากับ 103.37

$$\text{ได้สมการการพยากรณ์คือ } Y_t = (132.432 + 0.364579t)\hat{S}_t$$

ผลการวิเคราะห์ข้อมูลปริมาณน้ำฝนของจังหวัดลำปางด้วยวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลด้วยวิธีของวินเตอร์ ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) มีค่าเท่ากับ 36.52 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าเท่ากับ 2,089.53 และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) มีค่าเท่ากับ 53.61

$$\text{ได้สมการการพยากรณ์คือ } \hat{Y}_{t+m} = (a_t + b_t(m))\hat{S}_{t-12+m}$$

$$\text{โดยที่ } a_t = 0.001 \left(\frac{Y_t}{\hat{S}_{t-12}} \right) + 0.999[a_{t-1} + b_{t-1}]$$

$$\text{และ } b_t = 0.01(a_t - a_{t-1}) + 0.99b_{t-1}$$

ผลการวิเคราะห์ข้อมูลปริมาณน้ำฝนของจังหวัดลำปางด้วยวิธีบ็อกซ์และเจนกินส์ (Box – Jenkins Method) ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) มีค่าเท่ากับ 42.40 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าเท่ากับ 2,916.03 และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) มีค่าเท่ากับ 54.84

$$\text{ได้รูปแบบคือ } Y_t - Y_{t-12} + 0.5347Y_t - 0.5347Y_{t-24} + 0.2853Y_{t-24} - 0.2853Y_{t-36} = \varepsilon_t$$

$$\text{หรือ } Y_t = Y_{t-12} - 0.5347Y_t + 0.5347Y_{t-24} - 0.2853Y_{t-24} + 0.2853Y_{t-36} + \varepsilon_t$$

2. วิธีการพยากรณ์ร่วม

ผลการวิเคราะห์ข้อมูลปริมาณน้ำฝนของจังหวัดลำปางด้วยพยากรณ์ร่วม ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) มีค่าเท่ากับ 38.89 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าเท่ากับ 2,660.26 และค่าเฉลี่ยของค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) มีค่าเท่ากับ 55.25

$$\text{ได้สมการการพยากรณ์คือ } \hat{Y}_{CF} = 0.0570\hat{Y}^{(1)} + 0.5400\hat{Y}^{(2)} + 0.4030\hat{Y}^{(3)}$$

3. สรุปผลการพยากรณ์

สรุปผลการศึกษาเปรียบเทียบวิธีการพยากรณ์ที่ได้จากงานวิจัยนี้ โดยพิจารณาจากค่า MAD, MSE และ MAPE ที่ต่ำที่สุด พบว่าวิธีวินเตอร์จะเหมาะสมในการพยากรณ์ตัวแบบปริมาณน้ำฝนมากที่สุด เนื่องจากให้ค่า MAD, MSE และ MAPE ต่ำกว่าวิธีการพยากรณ์ทุกวิธี

ข้อเสนอแนะ

1. สำหรับงานวิจัยนี้

ค่าที่ได้จากการพยากรณ์ในงานวิจัยนี้ เป็นเพียงกรณีศึกษาปริมาณน้ำฝนในจังหวัดลำปาง ยังไม่เป็นภาพรวมของทั้งประเทศ ดังนั้นในการวิจัยครั้งต่อไปอาจนำข้อมูลอนุกรมเวลาจากหลายๆ จังหวัด หลายๆภูมิภาค หรือทั้งประเทศ เพื่อเปรียบเทียบว่าวิธีการพยากรณ์ของข้อมูลแต่ละวิธี สอดคล้องหรือแตกต่างจากงานวิจัยนี้

2. สำหรับงานวิจัยครั้งต่อไป

ควรใช้วิธีการพยากรณ์ที่ต่างจากงานวิจัยครั้งนี้ เช่น การศึกษาการพยากรณ์แบบการวิเคราะห์การถดถอย และอาจศึกษาปริมาณน้ำฝนในหลายจังหวัด หรือในแต่ละภาค เพื่อศึกษาเปรียบเทียบหารูปแบบการพยากรณ์ที่เหมาะสม ซึ่งอาจแตกต่างจากข้อมูลปริมาณน้ำฝนของจังหวัดลำปาง

เอกสารและสิ่งอ้างอิง

จันทร์ธา วงษ์อุทอง. 2548. การพยากรณ์ปริมาณความต้องการใช้ไฟฟ้า : กรณีศึกษาการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

ศักดิ์ดาว แนนเนียน. 2545. การเปรียบเทียบการพยากรณ์ด้วยวิธีโครงข่ายประสาทเทียม และวิธีของบ็อกซ์และเจนกินส์ : กรณีศึกษา อัตราแลกเปลี่ยนเงินตราต่างประเทศ. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

นนุช ดีแท้. 2534. การศึกษาเพื่อเปรียบเทียบเทคนิคการพยากรณ์ราคาข้าว มันสำปะหลัง และถั่วเขียวที่เกษตรกรได้ขายที่ไร่นา. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

นงลักษณ์ กลีบบัว. 2544. การศึกษาประสิทธิภาพของการพยากรณ์ร่วม โดยพิจารณาหาวิธีการพยากรณ์เดี่ยวและวิธีการพยากรณ์ร่วมที่เหมาะสมกับขนาดของอนุกรมเวลาและช่วงการพยากรณ์. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

มุกดา แม่นมินทร์. 2549. อนุกรมเวลาและการพยากรณ์. สำนักพิมพ์ประกายพริก, กรุงเทพฯ.

มนฤดี เกิดสมบุญ. 2542. การพยากรณ์พื้นที่เพาะปลูก ผลผลิต และราคาสินค้าเกษตร ในกรณีข้าว (ข้าวนาปีและข้าวนาปรัง) ข้าวโพดเลี้ยงสัตว์ ถั่วเขียว และถั่วเหลือง. วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย.

รัชนิวรรณ อนุตระกูลชัย. 2539. การศึกษาเปรียบเทียบวิธีการพยากรณ์ด้วยค่าวัดความถูกต้อง 3 ค่า. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

รัศมี หนานสาขอ. 2542. การพยากรณ์ปริมาณน้ำฝนในภาคตะวันออกเฉียงเหนือ เพื่อวางแผนการเพาะปลูกพืช. วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย.

วราฤทธิ์ พานิชกิจโกศลกุล. 2550. การพยากรณ์ปริมาณน้ำฝนรายเดือนของจังหวัดหนองคาย โดยใช้เทคนิคการพยากรณ์เชิงสถิติ. วารสารวิชาการ มหาวิทยาลัยหอการค้าไทย 27: 138-150.

ศิริลักษณ์ สุวรรณวงศ์. 2543. เทคนิคการพยากรณ์ การวิเคราะห์หอนุกรมเวลา. สุวีริยาสาส์น, กรุงเทพฯ.

สมเกียรติ เกตุเอี่ยม. 2548. เทคนิคการพยากรณ์. พิมพ์ครั้งที่ 2 ภารกิจเอกสารและตำรา กลุ่มงาน บริการการศึกษา มหาวิทยาลัยทักษิณ, สงขลา.

Boluwade, A. 2008. Box-Jenkins Seasonal Modeling of Niger River Flow, **African Journal of Science and Technology (AJST)**. 9: 110-117.

Delurgio, A. 1998. **Forecasting Principle and Applications**. McGraw-Hall International Editions, New York.



ตารางผนวกที่ 1 ข้อมูลปริมาณน้ำฝนของจังหวัดลำปาง ตั้งแต่ปี 2542 – 2552 (หน่วย : มิลลิเมตร)

เดือน / ปี	2542	2543	2544	2545	2546	2547	2548	2549	2550	2551	2552
มกราคม	0.7	0	1.7	12.1	2	0.1	16	0	0	9.4	0
กุมภาพันธ์	14.4	38.3	0	5.9	1.8	0.4	14.8	16.8	1.3	46.6	0
มีนาคม	35.4	2.1	98.4	18.7	21.6	0	4.7	43	5.4	40.2	26.9
เมษายน	144.3	84	4.6	55.3	121.4	70.3	66.2	174.4	85.2	81.4	55.2
พฤษภาคม	101.5	231.9	239.4	231.1	79.8	83.8	74.5	313.6	332.7	159.1	237.7
มิถุนายน	84.8	114.5	96.1	51.1	74.5	153.1	129.7	175.7	80.4	74.1	-
กรกฎาคม	68.6	153	123.1	112.7	139.1	98.4	189.5	179.8	100.1	81.7	-
สิงหาคม	172.3	168.1	202	316.6	79.1	77	139.9	293.4	200.3	100	-
กันยายน	278.5	185.5	155.9	214	329.3	155.8	417.3	170.5	197.2	191.6	-
ตุลาคม	136.9	78.3	179.2	80.8	53	57	19.6	100.6	84.9	160.4	-
พฤศจิกายน	37.8	0.3	4.6	93.8	0	51.1	26.2	0	13.7	37.7	-
ธันวาคม	5.9	0	0	13.1	0	0	28.4	0	0	12.3	-

หมายเหตุ: ปี 2552 มีข้อมูล 5 เดือน

ที่มา: สำนักพัฒนาอุดมศึกษา กรมอุดมศึกษา กระทรวงเทคโนโลยีสารสนเทศ และการสื่อสาร

ประวัติการศึกษา และการทำงาน

ชื่อ – นามสกุล	นายปรีชา เรืองชัยทิเวท
วัน เดือน ปี ที่เกิด	31 พฤษภาคม 2524
สถานที่เกิด	อำเภอเมือง จังหวัดลำปาง
ประวัติการศึกษา	วท.บ. (คณิตศาสตร์) มหาวิทยาลัยศรีนครินทรวิโรฒ
ตำแหน่งหน้าที่การงานปัจจุบัน	-
สถานที่ทำงานปัจจุบัน	-
ผลงานดีเด่นและรางวัลทางวิชาการ	-
ทุนการศึกษาที่ได้รับ	-