บทที่ 2

เอกสารและทฤษฎีที่เกี่ยวข้อง

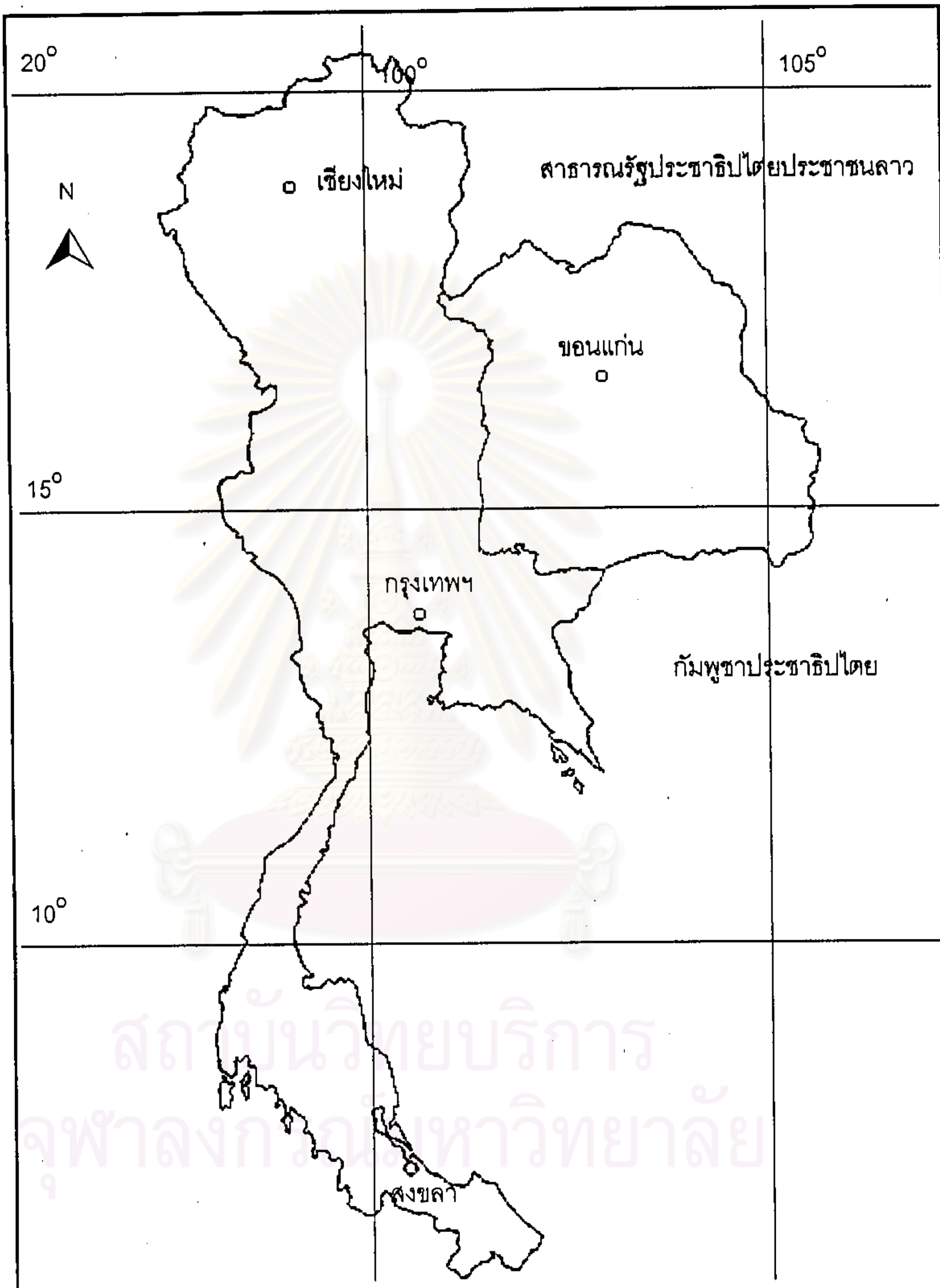
2.1 สภาพภูมิประเทศ และภูมิอากาศในภาคตะวันออกเฉียงเหนือ

ภาคตะวันออกเฉียงเหนือของประเทศไทยตั้งอยู่ระหว่างเส้นรุ้งที่ 14-18 องศาของซีกโลกเหนือ และเส้นแวงที่ 100-106 องศาตะวันออก มีเนื้อที่ 106,250,000 ไร่ โดยมีเนื้อที่หนึ่งในสามของประเทศ ดังภาพที่ 2.1 ประกอบด้วยจังหวัดต่าง ๆ 19 จังหวัดคือ จังหวัดเลย อุดรธานี หนองคาย หนองบัวลำภู สกลนคร นครพนม ชัยภูมิ ขอนแก่น กาฬสินธุ์ มุกดาหาร มหาสารคาม ร้อยเอ็ด ยโสธร อำนาจเจริญ อุบลราชธานี ศรีสะเกษ สุรินทร์ บุรีรัมย์ และนครราชสีมา ดังภาพที่ 2.2

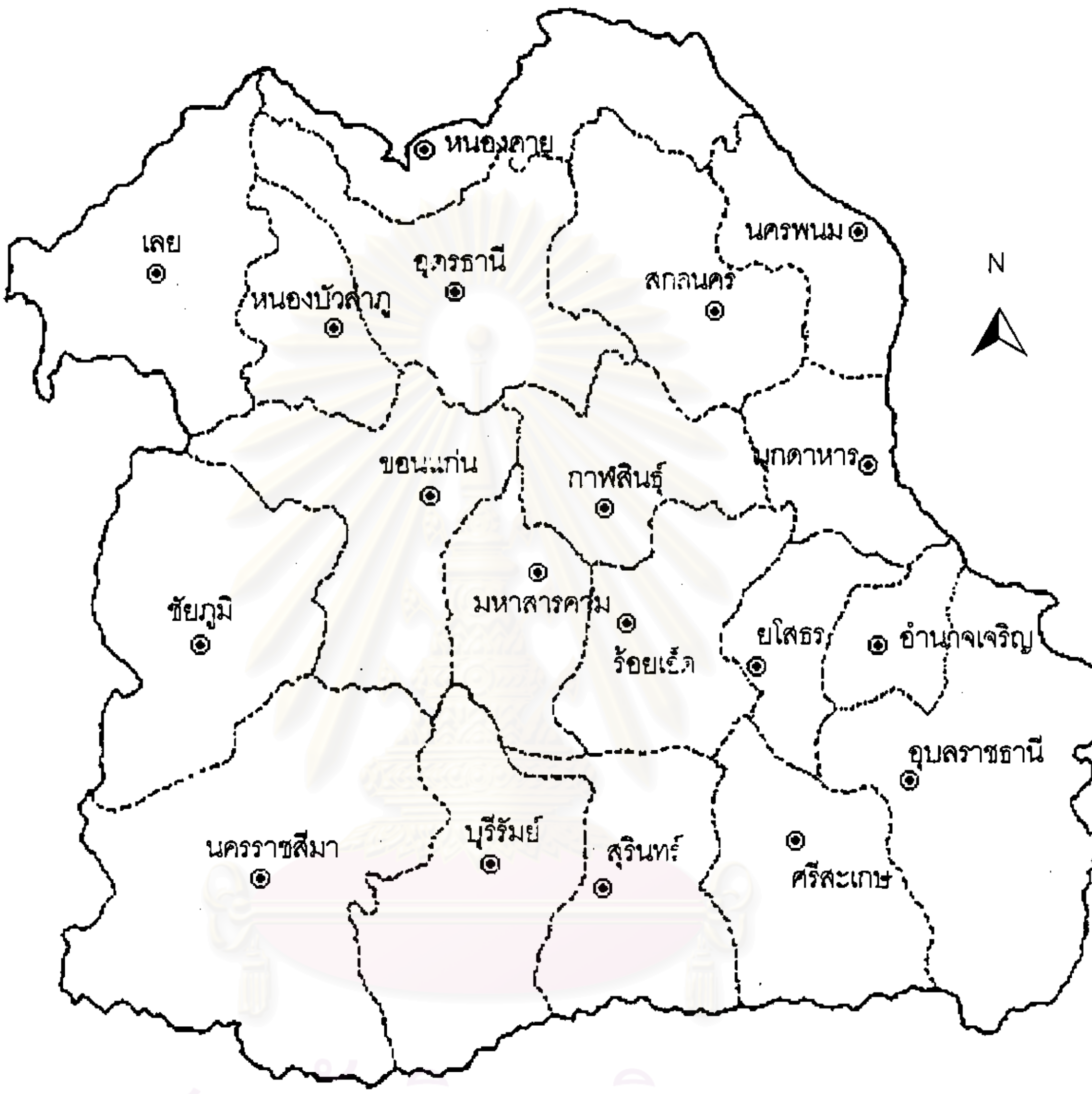
ลักษณะทางภูมิอากาศของภาคตะวันออกเฉียงเหนือ มี 3 ฤดูกาล คือ ฤดูฝน ฤดูร้อน และฤดูหนาว ฤดูฝนจะเริ่มประมาณกลางเดือนพฤษภาคมถึงเดือนตุลาคม แต่เนื่องจากภาคตะวันออกเฉียงเหนือมีพื้นที่กว้างใหญ่ปริมาณน้ำฝนในภาคนี้จึงแตกต่างกัน และสามารถแยกออกได้เป็น 3 เขตด้วยกัน ดังภาพที่ 2.3 ดังนี้

1. เขตที่มีน้ำฝนเฉลี่ยต่อปี 1400 มม. หรือมากกว่า ถือว่าเป็นเขตที่มีน้ำฝนมาก
2. เขตที่มีน้ำฝนเฉลี่ยต่อปี อยู่ระหว่าง 1200-1400 มม.
3. เขตที่มีน้ำฝนเฉลี่ยต่อปีน้อยกว่า 1200 มม. ถือว่าเป็นเขตที่มีน้ำฝนน้อย

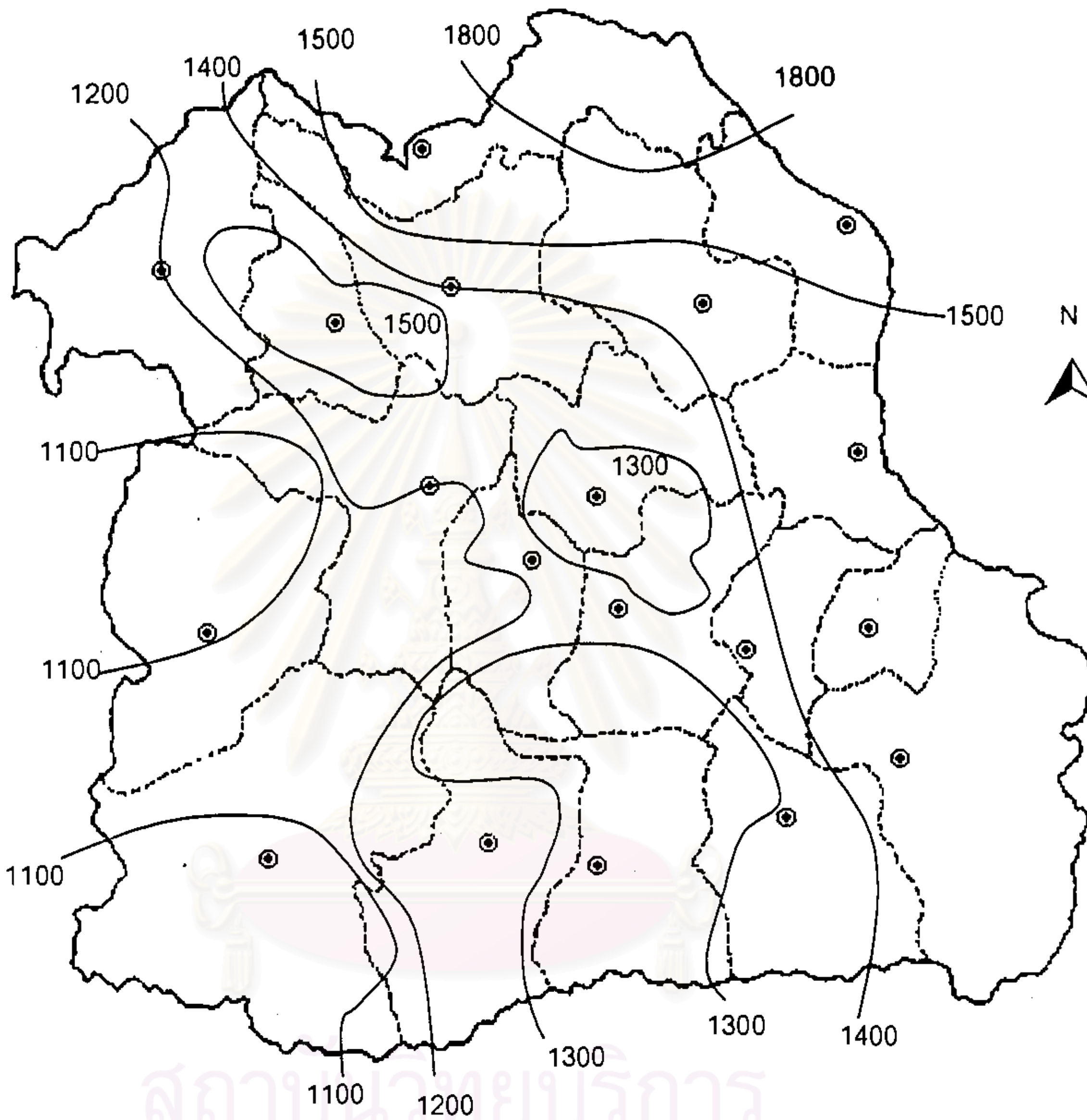
ภาคตะวันออกเฉียงเหนือมีประชากรหนึ่งในสามของประเทศ แต่ขณะเดียวกันรายได้ของประชากรต่อหัวกลับต่ำที่สุด เนื่องจากประชากรส่วนใหญ่ประกอบอาชีพหลักด้านเกษตรกรรม การผลิตพืชโดยส่วนใหญ่จะปลูกข้าวซึ่งมีเนื้อที่ปลูกถึง 33,750,000 ไร่ แหล่งปลูกข้าวที่สำคัญของภาคจะอยู่ทางราบลุ่มทิศตะวันออกของภาค พื้นที่ริมฝั่งแม่น้ำโขง ตอนกลางของภาค และบริเวณที่ราบชายฝั่งแม่น้ำชี และแม่น้ำมูล คือจังหวัดร้อยเอ็ด ยโสธร สุรินทร์ และศรีสะเกษ ซึ่งเป็นเขตที่ได้รับน้ำฝนมาก (มากกว่า 1400 มม.ต่อปี) ส่วนบริเวณที่มีน้ำฝนต่ำกว่า 1200 มม.ต่อปีจะมีโอกาสเกิดภาวะฝนแล้งในระหว่างฤดูฝนได้



ภาพที่ 2.1 แผนที่ประเทศไทย และภาคตะวันออกเฉียงเหนือ



ภาพที่ 2.2 แสดงเขตจังหวัดต่าง ๆ ในภาคตะวันออกเฉียงเหนือ



ภาพที่ 2.3 แสดงเขตน้ำฝนเฉลี่ย (มม./ปี) ในภาคตะวันออกเฉียงเหนือ

ในการผลิตพืชที่ต้องอาศัยน้ำฝนเพียงอย่างเดียวนี้ จำเป็นต้องมีการจัดระบบการปลูกพืชที่ดี ต้องมีความรู้เกี่ยวกับความแปรปรวนของฝน ลักษณะการกระจายของฝน การเริ่มฤดูที่แน่นอน ฯลฯ จะสามารถทำให้สามารถวางแผนที่ดีได้ เพื่อลดความเสี่ยงต่อความเสียหายที่อาจมีต่อการเพาะปลูกพืช (Vorasoot, et al., 1985)

มีการศึกษาที่เกี่ยวกับการวิเคราะห์และการพยากรณ์ปริมาณน้ำฝนหลายงานวิจัยด้วยกัน โดยมีเหตุผลเพื่อจะนำไปใช้ประโยชน์ในการวางแผนการเพาะปลูกพืช และจัดการระบบพืช ซึ่งมีวิธีการและหลักการแตกต่างกันออกไป แต่มีวัตถุประสงค์รวมกันคือเพื่อช่วยกำหนดระบบการปลูกและการจัดการที่เหมาะสมกับสภาพท้องถิ่น

พูนศักดิ์ (2537) ศึกษาถึงสาระอุตุนิยมวิทยา 8 ชนิด คือ อุณหภูมิเฉลี่ยต่อเดือน ปริมาณความชื้นสัมพัทธ์ต่อเดือน อุณหภูมิจุดน้ำค้างเฉลี่ยต่อเดือน ปริมาณน้ำระเหยเฉลี่ยต่อเดือน จำนวนเมฆเฉลี่ยต่อเดือน จำนวนชั่วโมงเฉลี่ยต่อเดือนของระยะเวลาที่มีแสงแดด ระยะทางเฉลี่ยต่อเดือนของทัศนวิสัยและความเร็วเฉลี่ยต่อเดือน เพื่อดูผลกระทบต่อการพยากรณ์ปริมาณน้ำฝนเฉลี่ย และสร้างสมการพยากรณ์ปริมาณน้ำฝนเฉลี่ยรายปีในอำเภอเมือง จังหวัดเชียงใหม่ ปรากฏว่าอิทธิพลทางตรงต่อปริมาณน้ำฝนเฉลี่ยคือ อุณหภูมิจุดน้ำค้างเฉลี่ยต่อเดือน ปริมาณความชื้นสัมพัทธ์ อุณหภูมิและจำนวนเมฆ ในสมการการพยากรณ์ปริมาณน้ำฝนเฉลี่ยต่อเดือนยังให้ค่าปริมาณน้ำฝนเฉลี่ยล่วงหน้าของอำเภอเมืองด้วย

สำนักงานเศรษฐกิจการเกษตร (2528) ได้ศึกษาวิเคราะห์รูปแบบน้ำฝนเพื่อนำไปใช้เป็นข้อมูลพื้นฐานในการพิจารณาแนวทางการพัฒนาการเกษตรของเกษตรกรในเขตน้ำฝน โดยการกำหนดแผนการเพาะปลูกพืชให้มีความสัมพันธ์กับรูปแบบน้ำฝนที่เป็นอยู่ในแต่ละเขตการเพาะปลูก ซึ่งจะทำให้เกษตรกรและผู้ที่จะนำรูปแบบน้ำฝนตลอดจนข้อมูลที่ได้ทำการศึกษาวิเคราะห์นี้ไปใช้ในการพัฒนาการปลูกพืช และปรับปรุงระบบการปลูกพืชให้มีประสิทธิภาพในทางการผลิตต่อไป

วิยะดาและคณะ (2530) ได้ศึกษาข้อมูลปริมาณน้ำฝนแยกตามลักษณะภูมิประเทศที่แตกต่างกันจากสถานีต่าง ๆ ในเชียงใหม่ ลำพูน โดยศึกษาปริมาณน้ำฝนต่ำสุดและสูงสุดในแต่ละสัปดาห์และแต่ละปี ลักษณะโค้งการแจกแจงน้ำฝนที่

เปลี่ยนแปลงไปตามลักษณะภูมิประเทศ 2 กลุ่มคือ สถานีตรวจวัดน้ำฝนที่อยู่บนที่ราบคือ อยู่ไม่สูงเกิน 500 เมตรจากระดับน้ำทะเล และสถานีตรวจวัดน้ำฝนที่อยู่บนที่สูงคืออยู่สูงเกิน 500 เมตรจากระดับน้ำทะเล โดยวิเคราะห์ความสัมพันธ์ระหว่างปริมาณน้ำฝนรายเดือน รายฤดูกาล และรายปี กับความสูงของพื้นที่จากระดับน้ำทะเล พบว่าปริมาณน้ำฝนจะมากขึ้นตามความสูงของพื้นที่ รวมทั้งช่วงฤดูฝนจะยาวขึ้นทำให้โอกาสที่จะเพิ่มประสิทธิภาพการใช้พื้นที่เพาะปลูกมีมากขึ้น เมื่อนำผลการวิเคราะห์ไปประเมินปริมาณฝนในพื้นที่ที่ไม่มีสถานีแต่มีลักษณะภูมิประเทศคล้ายคลึงกัน ก็จะสามารถเลือกชนิดของพืชที่เหมาะสมในการเพาะปลูกในพื้นที่ และสามารถเตรียมการป้องกันน้ำท่วมหรือไฟป่าได้อีกด้วย

จุไรพร (2541) ได้ศึกษาเพื่อพัฒนาฐานข้อมูลเชิงพื้นที่ของปริมาณน้ำฝนในจังหวัดเชียงใหม่ และจังหวัดใกล้เคียงจาก 201 สถานีตรวจวัดน้ำฝน จากการบันทึกข้อมูลเป็นระยะ 6-46 ปี โดยเก็บข้อมูลในระบบสารสนเทศภูมิศาสตร์ (Geographic Information System ; GIS) เพื่อแก้ปัญหาข้อจำกัดของการนำข้อมูลปริมาณน้ำฝนและภูมิอากาศไปใช้ในแบบจำลองพืชเมื่อพื้นที่เป้าหมายอยู่ระหว่างสถานีอุตุนิยมวิทยาหลายสถานี ซึ่งจะทำให้ตัดสินใจไม่ได้ว่าจะใช้ข้อมูลจากสถานีใด ผลคือทำให้ข้อมูลภูมิอากาศคลาดเคลื่อนไปจากความเป็นจริง จากฐานข้อมูลน้ำฝนเฉลี่ยรายเดือน และจำนวนวันที่ฝนตกเฉลี่ยต่อเดือนที่พัฒนาขึ้น โดยนำไปประมาณค่าต่อเนื่องเชิงพื้นที่จาก 3 วิธีการ แล้วเปรียบเทียบวิธีการที่สามารถประมาณค่าได้ใกล้เคียงข้อมูลจากการจดบันทึกมากที่สุด

2.2 **ทฤษฎีและสถิติที่ใช้ในการวิจัย** ในการศึกษาวิจัยในครั้งนี้ เป็นการพยากรณ์ปริมาณน้ำฝนในภาคตะวันออกเฉียงเหนือ เพื่อการวางแผนการเพาะปลูกพืช

การพยากรณ์ หมายถึง การคาดคะเนหรือทำนายการเกิดของเหตุการณ์ หรือสภาพต่าง ๆ ในอนาคตโดยศึกษารูปแบบการเกิดของเหตุการณ์จากข้อมูลและหรือความรู้ในอดีต การพยากรณ์มีความจำเป็นอย่างยิ่งต่อในการประกอบการวางแผน และตัดสินใจเกี่ยวกับการดำเนินงานในทุกสาขาอาชีพ หากนักวางแผนหรือผู้ตัดสินใจในองค์กรทราบว่าเหตุการณ์ใดจะเกิดขึ้นหรืออาจจะเกิดขึ้นในอนาคตด้วยความเชื่อมั่นระดับหนึ่งจะทำให้การวางแผนหรือการตัดสินใจในการดำเนินงานเป็นไปได้อย่างถูกต้อง

อนุกรมเวลา หมายถึง กลุ่มของค่าสังเกตที่เก็บรวบรวมมาตามเวลาอย่างต่อเนื่อง ใช้สัญญลักษณ์ $\{y_t\}$ แทนอนุกรมเวลา $y_1 ....y_t$, t = 1,2......,T ดังนั้นการวิเคราะห์อนุกรมเวลา จึงเป็นการศึกษาแผนแบบการเคลื่อนไหวของอนุกรมเวลา จากรูปแบบที่ได้จะนำไปใช้พยากรณ์ ซึ่งการเคลื่อนไหวของอนุกรมเวลา จะขึ้นอยู่กับส่วนประกอบหลัก ได้แก่ แนวโน้ม (trend) ฤดูกาล (seasonal) วัฏจักร (cyclical effect) และเหตุการณ์ผิดปกติ (irregular)

เทคนิคการพยากรณ์ที่นำมาใช้ในการศึกษามีดังนี้

2.2.1 การพยากรณ์สำหรับอนุกรมเวลาที่มีระดับค่าเฉลี่ยคงที่ ลักษณะตัวแบบ คือ

$$Y_t = \beta + \varepsilon_t , \quad t= 1,2,...,T \tag{1}$$

ซึ่งพารามิเตอร์ β เป็นระดับค่าเฉลี่ยคงที่

ตัวแปรสุ่ม ε_t , t= 1,2,....,T แทนความคลาดเคลื่อนที่มีค่าเฉลี่ยเป็นศูนย์ ไม่มีสหสัมพันธ์และความแปรปรวนคงที่

ตัวแบบพยากรณ์ค่าในอนาคตแบบจุดสำหรับ Y_{T+1} ที่เวลา T+1 จากเวลาปัจจุบัน T คือ

$$\hat{Y}_T(l) = \bar{Y} = \frac{1}{T}\sum_{t=1}^{T} Y_t , \qquad l = 1,2,...$$

2.2.2 เทคนิคการปรับให้เรียบครั้งเดียวแบบเอกซ์โปเนลเชียล เป็นเทคนิคการพยากรณ์เมื่อระดับค่าเฉลี่ยไม่คงที่ตลอดเวลา T แต่จะแปรเปลี่ยนอย่างช้า ๆ ตามเวลา จึงควรให้น้ำหนักหรือความสำคัญกับข้อมูลไม่เท่ากัน กล่าวคือควรให้น้ำหนักกับข้อมูลในช่วงเวลาปัจจุบันมากกว่าข้อมูลในอดีตที่ห่างออกไป โดยมีตัวแบบอนุกรมเวลาดังตัวแบบใน (1) และ ตัวแบบพยากรณ์ค่าในอนาคตแบบจุดสำหรับ $Y_{T+\ell}$ $(l = 1,2,...)$ จากเวลาปัจจุบัน T คือ

$$\hat{Y}_T(l) = S_t = \alpha Y_t + (1-\alpha)S_{t-1}, \quad l = 1,2,...$$

ซึ่งพารามิเตอร์ α เป็นค่าคงที่หรือสัมประสิทธิ์ปรับให้เรียบ (smoothing constant) มีค่าอยู่ระหว่าง 0 และ 1 โดยทั่ว ๆ ไปจะเลือก α อยู่ระหว่าง 0.05 และ 0.3 ถ้าระดับค่าเฉลี่ยเปลี่ยนแปลงอย่างช้า ๆ ควรเลือก α มีค่าเล็ก หรือ ทดลองแปรค่า α เช่นเริ่มจาก 0.01, 0.02 ไปเรื่อย ๆ และแต่ละค่า α คำนวณค่า S_t และหาค่าความคลาดเคลื่อนกำลังสอง (MSE) และเลือกค่า α ที่ให้ค่า MSE ต่ำที่สุด

สำหรับ S_t คือ ตัวสถิติปรับให้เรียบ (smoothing statistics) นักพยากรณ์จะต้องกำหนดค่าเริ่มต้น S_0 ถ้าระดับค่าเฉลี่ยเปลี่ยนแปลงอย่างช้า ๆ โดยทั่วไปจะเลือกกำหนด S_0 เป็นค่าเฉลี่ยของข้อมูลอนุกรมเวลา คือ $S_0 = \bar{Y}$ แต่ถ้าข้อมูลมีการแกว่งขึ้นลงอย่างมาก จะเลือกกำหนด S_0 โดยใช้ค่าข้อมูลตัวแรกคือ $S_0 = Y_1$ เป็นต้น

2.2.3 **เทคนิคการปรับให้เรียบสองครั้งแบบเอกซ์โปเนลเชียล** เมื่ออนุกรมเวลามีแนวโน้ม เป็นเชิงเส้นคงที่ ซึ่งตัวแบบเชิงเส้นตรงดังนี้

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

ซึ่ง Y_t คือตัวแปรอนุกรมเวลาที่เวลา t

β_0 เป็นระดับค่าเฉลี่ยของ Y_t ที่เวลา t=0

β_1 อัตราการเปลี่ยนแปลงของ Y_t ต่อหนึ่งหน่วยเวลา

ε_t, t= 1,2,...,T แทนความคลาดเคลื่อน ที่มีค่าเฉลี่ยเป็นศูนย์ ไม่มีสหสัมพันธ์ และความแปรปรวนคงที่

มีสูตรพยากรณ์ค่าจริงสำหรับ $Y_{T+\ell}$ ที่เวลา $T+\ell$ จากเวลาปัจจุบัน T คือ

$$\hat{Y}_T(l) = \left(2+\frac{\alpha l}{1-\alpha}\right)S_T^{[1]} - \left(1+\frac{\alpha l}{1-\alpha}\right)S_T^{[2]},$$

$$S_T^{[1]} = \alpha Y_T + (1-\alpha)S_{T-1}^{[1]} \quad \text{และ} \quad S_T^{[2]} = \alpha S_T^{[1]} + (1-\alpha)S_{T-1}^{[2]}$$

ซึ่งค่าเริ่มต้น $S_0^{[1]}$ และ $S_0^{[2]}$ ประมาณได้ดังนี้

$$S_0^{[1]} = \hat{\beta}_0 - \left(\frac{1-\alpha}{\alpha}\right)\hat{\beta}_1, \qquad S_0^{[2]} = \hat{\beta}_0 - 2\left(\frac{1-\alpha}{\alpha}\right)\hat{\beta}_1$$

ซึ่ง $\hat{\beta}_0 = \bar{Y} - \frac{T+1}{2}\hat{\beta}_1$, $\hat{\beta}_1 = 12\sum_{t=1}^{T}\frac{(t-(T+1)/2)Y_t}{T^3 - T}$

2.2.4 **วิธีแยกส่วนประกอบ** (Decompositon) การวิเคราะห์ข้อมูลอนุกรมเวลาวิธีนี้เป็นที่นิยมกันมาก เนื่องจากสามารถวิเคราะห์โดยแยกอนุกรมเวลาออกเป็นส่วนประกอบต่าง ๆ คือแนวโน้ม ฤดูกาล วัฏจักร และเหตุการณ์ผิดปกติ ซึ่งตัวแบบที่ใช้ในการศึกษาครั้งนี้อยู่ในรูปตัวแบบผลคูณ คือ

$$Y_T(l) = (\mu_t + \beta_t t) I_t C_t \varepsilon_t$$

ซึ่งพารามิเตอร์ μ_t, β_t, C_t และ I_t เป็นระดับของข้อมูล ความชัน ส่วนประกอบของวัฏจักร และส่วนประกอบฤดูกาล ตามลำดับ
ตัวแปรสุ่ม ε_t , t= 1,2,...,T แทนความคลาดเคลื่อน ที่มีค่าเฉลี่ยเป็นศูนย์ ไม่มีสหสัมพันธ์ และความแปรปรวนคงที่

สูตรพยากรณ์คือ

$$\hat{Y}_T(l) = (\hat{\mu}_t + l\hat{\beta}_t)\hat{C}_t\hat{I}_t$$

2.2.5 **วิธีการพยากรณ์ของวินเตอร์** เมื่ออนุกรมเวลามีองค์ประกอบฤดูกาล ควรใช้วิธีการพยากรณ์ที่พิจารณาองค์ประกอบฤดูกาลร่วมด้วย ตัวแบบผลคูณของวินเตอร์ คือ

$$Y_T(l) = (\mu_t + \beta_t t) I_t + \varepsilon_t$$

ซึ่งพารามิเตอร์ μ_t, β_t และ I_t เป็นระดับของข้อมูล ความชัน และฤดูกาล ตามลำดับ
ตัวแปรสุ่ม ε_t , t= 1,2,...,T แทนความคลาดเคลื่อน ที่มีค่าเฉลี่ยเป็นศูนย์ ไม่มีสหสัมพันธ์ และความแปรปรวนคงที่

สูตรพยากรณ์คือ

$$\hat{Y}_T(l) = (\hat{\mu}_t + l\hat{\beta}_t) I_{t+1-m} \quad , \quad t=m,m+1,...$$

ซึ่ง

$$\hat{\mu}_t = \alpha_1\left(\frac{Y_t}{\hat{I}_{t-m}}\right) + (1-\alpha_1)(\hat{\mu}_{t-1} + \hat{\beta}_{t-1})$$

$$\hat{\beta}_t = \alpha_2(\hat{\mu}_t - \hat{\mu}_{t-1}) + (1-\alpha_2)\hat{\beta}_{t-1}$$

$$\hat{I}_t = \alpha_3\left(\frac{Y_t}{\mu_t}\right) + (1-\alpha_3)\hat{I}_{t-m}$$

m คือความยาวของคาบฤดูกาล เช่น m=12 (สำหรับอนุกรมเวลารายเดือน) หรือ m=4 (สำหรับอนุกรมเวลารายสามเดือน) หรือใช้การกำหนดค่าเริ่มต้นให้

$$\hat{\mu}_m = \frac{(Y_1 + Y_2 + \ldots + Y_m)}{m} \quad , \hat{I}_t = \frac{Y_t}{\hat{\mu}_m}, \quad t = 1,2,\ldots,m \text{ และ } \beta_m = 0$$

2.2.6 เทคนิคการวิเคราะห์อนุกรมเวลาบอกซ์-เจนกินส์ วิธีการพยากรณ์ข้อมูลอนุกรมเวลาบอกซ์-เจนกินส์ จะหาตัวแบบอนุกรมเวลาโดยพิจารณาสหสัมพันธ์ระหว่าง Y ที่ตำแหน่งเวลาหรือคาบเวลา $t(Y_t)$ และ Y_t ที่ตำแหน่งเวลาหรือคาบเวลาต่าง ๆ ที่ผ่านมา $(Y_{t-2}, Y_{t-1}, \ldots)$ เมื่อได้ตัวแบบที่แสดงความสัมพันธ์ระหว่าง Y_t กับ $Y_{t-2}, Y_{t-1}, \ldots$ จะใช้ตัวแบบนี้ในการพยากรณ์ $Y_{t+1}, Y_{t+2}, \ldots$ ในอนาคต ซึ่งมีลักษณะตัวแบบคือ

$$Y_t = \mu + a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \ldots \qquad (3)$$

Y_t เกิดจากผลบวกเชิงเส้นของอนุกรมเวลา ค่าความคลาดเคลื่อนสุ่ม a_t กับ $a_{t-1}, a_{t-2}, \ldots$ ซึ่งมีข้อสมมติว่าแต่ละตัวมีค่าเฉลี่ยศูนย์ $E(a_t) = 0$ ไม่มีสหสัมพันธ์กัน $E(a_t, a_{t-k})=0$ และความแปรปรวนคงที่ $V(a_t) = E(a_t^2) = \sigma^2$ และ $Cov(a_t,a_{t-k}) = 0$, ทุกค่า $k \neq 0$ และมีการแจกแจงปกติ

2.2.6.1 ตัวแบบภายใต้ภาวะคงที่ (Stationary Model)

2.2.6.1.1 ตัวแบบอัตตถดถอย (Autoregressive Models ; AR) ตัวแบบอัตตถดถอยอันดับ p แทนด้วย AR(p) มีรูปแบบทั่วไปคือ

$$Y_t - \mu = \phi_1(Y_{t-1}-\mu) + \phi_2(Y_{t-2}-\mu) + \ldots + \phi_p(Y_{t-p}-\mu) + a_t$$

หรือ

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + a_t$$

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + a_t$$

โดยให้

$Z_t = Y_t - \mu$, $Z_{t-1} = Y_{t-1} - \mu$,.... และ $c = \mu(1-\phi_1-\phi_2-...-\phi_p)$

หรือเขียนในรูปสั้น ๆ คือ

$$\phi_p(B)Z_t = a_t$$

โดยที่ $\phi_p(B) = (1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p)$

และ $BZ_t = Z_{t-1}$, $B^2 Z_t = Z_{t-2}$, ... , $B^p Z_t = Z_{t-p}$

และ $\mu, \phi_1, \phi_2 ...$ เป็นพารามิเตอร์ซึ่งไม่ทราบค่าต้องทำการประมาณค่า และ B คือ ตัวถอยหลังเวลา

โดยทั่วไปอันดับ p จะไม่สูงมาก ถ้า p=1,2 สามารถเขียนตัวแบบ AR(1) ได้ดังนี้

$$Y_t = c + \phi_1 Y_{t-1} + a_t \quad , -1<\phi_1<1$$

และตัวแบบ AR(2) คือ

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t \quad , \quad \begin{array}{l} \phi_1+\phi_2<1 \\ \phi_1-\phi_2<1 \\ -1<\phi_2<1 \end{array}$$

2.2.6.1.2 ตัวแบบค่าเฉลี่ยเคลื่อนที่ (Moving Average Model ; MA) ตัวแบบค่าเฉลี่ยเคลื่อนที่อันดับ q แทนด้วย MA(q) มีตัวแบบคือ

$$Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - ... - \theta_q a_{t-q}$$

อันดับ q จะไม่สูงมาก ถ้า q=1,2 สามารถเขียนตัวแบบ MA(1) ได้ดังนี้

$$Y_t = \mu + a_t - \theta_1 a_{t-1} \quad , \quad -1<\theta_1<1$$

และ MA(2) คือ

$$Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad , \quad \begin{array}{l} -1<\theta_2<1 \\ \theta_1+\theta_2<1 \\ \theta_1-\theta_2<1 \end{array}$$

2.2.6.1.3 ตัวแบบอัตตถดถอย-ค่าเฉลี่ยเคลื่อนที่ (Autoregressive-Moving average Model ; ARMA) เป็นตัวแบบผสมระหว่างตัวแบบ AR และ MA จะเป็นตัวแบบที่ประหยัดแทนการใช้ตัวแบบ AR และ MA อันดับสูง ๆ แทน

ด้วย ARMA(p,q) หมายถึงตัวแบบผสมอัตตถดถอยอันดับ p และค่าเฉลี่ยเคลื่อนที่อันดับ q มีรูปแบบดังนี้

$$Y_t = \mu + \phi_1(Y_{t-1}-\mu) + \phi_2(Y_{t-2}-\mu) + ... + \phi_p(Y_{t-p}-\mu) + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - ... - \theta_q a_{t-q}$$

อันดับ p และ q จะไม่สูงมากนัก ถ้า p=1 และ q=1 จะได้ตัวแบบ ARMA(1,1) คือ

$$Y_t = \mu + \phi_1(Y_{t-1}-\mu) + a_t - \theta_1 a_{t-1}$$

หรือ $$Y_t = c + \phi_1 Y_{t-1} + a_t - \theta_1 a_{t-1} \quad , -1<\phi_1<1 \text{ และ } -1<\theta_1<1$$

โดยมีเงื่อนไขในภาวะคงที่ และผกผันได้

2.2.6.2 ตัวแบบภายใต้ภาวะไม่คงที่ (Nonstationary Models) และตัวแบบ ARIMA (Autoregressive Integrated Moving Average Model) ถ้าข้อมูลอนุกรมเวลาไม่อยู่ในภาวะคงที่ในค่าเฉลี่ย และ/หรือความแปรปรวน จะต้องทำการแปลงข้อมูลนั้นให้อยู่ในภาวะคงที่ก่อนพิจารณากำหนดตัวแบบ การแปลงข้อมูลอนุกรมเวลาเมื่อค่าเฉลี่ยไม่คงที่ โดยวิธีการทำผลต่าง แทนด้วยสัญญลักษณ์ d ซึ่งอาจทำมากกว่าหนึ่งครั้งหรือมากกว่าก็ได้ จึงจะทำให้อนุกรมเวลามีค่าเฉลี่ยคงที่ แต่การทำผลต่างหลายครั้งอาจทำให้ค่าพยากรณ์มีความคลาดเคลื่อนสูง ตัวแบบผสม ARIMA ด้วยอันดับ p,d,q ; ARIMA(p,d,q) มีรูปแบบทั่วไปดังนี้

$$\phi_p(B)(1-B)^d Y_t = \delta + \theta_q(B)a_t$$

หรือ

$$\phi_p(B)W_t = \delta + \theta_q(B)a_t$$

ซึ่งให้ $W_t = (1-B)^d Y_t$ และ δ (อาจมีค่าเป็นศูนย์) เป็นพารามิเตอร์แสดงระดับค่าเฉลี่ยคงที่ของอนุกรม W_t เป็นอนุกรมที่มีภาวะคงที่ในค่าเฉลี่ย เช่นตัวแบบ ARIMA (1,1,1) คือ

$$(1-\phi_1 B)(1-B)Y_t = \delta + (1-\theta_1 B)a_t$$

หรือ

$$Y_t = \delta + (1+\phi_1)Y_{t-1} - \phi_1 Y_{t-2} + a_t - \theta_1 a_{t-1}$$

หรือ

$$W_t = \delta + \phi_1 W_{t-1} + a_t - \theta_1 a_{t-1} \quad , W_t = Y_t - Y_{t-1}$$

และตัวแบบ ARIMA (2,1,0) คือ

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

หรือ

$$W_t = \delta + \phi_1 W_{t-1} + \phi_2 W_{t-2} + a_t \quad , W_t = Y_t - Y_{t-1}$$

และตัวแบบ ARIMA (1,2,1) คือ

$$(1-\phi_1 B)(1-B^2)Y_t = \delta + (1-\theta_1)a_t$$

หรือ

$$W_t = \delta + \phi_1 W_{t-1} + a_t - \theta_1 a_{t-1} \quad , W_t = (1-B)^2 Y_t$$

2.2.6.3 ตัวแบบ ARIMA ที่มีองค์ประกอบฤดูกาล เมื่อนำองค์ประกอบในส่วนที่ไม่ใช่ฤดูกาล และส่วนที่มีฤดูกาลมาผนวกเข้าด้วยกันจะได้ตัวแบบ ARIMA ซึ่งแทนด้วย ARIMA $(p,d,q)(P,D,Q)_s$ ที่ตัวแบบดังนี้

$$\phi_p(B)\Phi_p(B^s)(1-B)^d(1-B)^D Y_t = \delta + \theta_q(B)\ \Theta_q(B^s)a_t$$

โดยที่ $\Phi_p(B^s) = (1-\Phi_s B^s - \Phi_{2s} B^{2s} - ... - \Phi_{ps} B^{ps})$

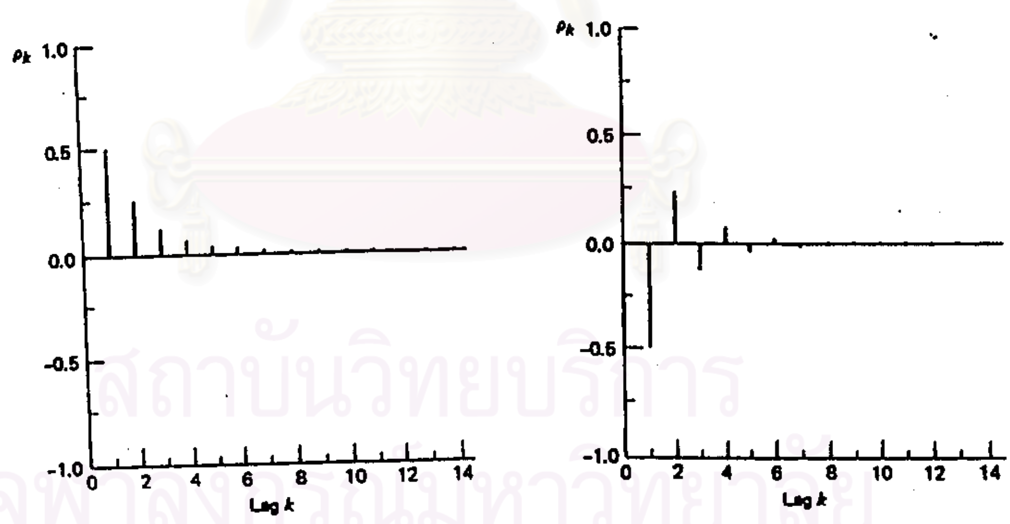
$$\Theta_q(B^s) = (1-\Theta_s B^s - \Theta_{2s} B^{2s} - ... - \Theta_{qs} B^{qs})$$

ซึ่ง P คือ อันดับในส่วนของกระบวนการ AR , Q คือ อันดับในส่วนของกระบวนการ MA และ D คือจำนวนครั้งของการทำผลต่าง อนุกรมเวลาห่างกัน s คาบเวลา

ซึ่งขั้นตอนการพยากรณ์โดยวิธีของบอกซ์-เจนกินส์มี 4 ขั้นตอนคือ

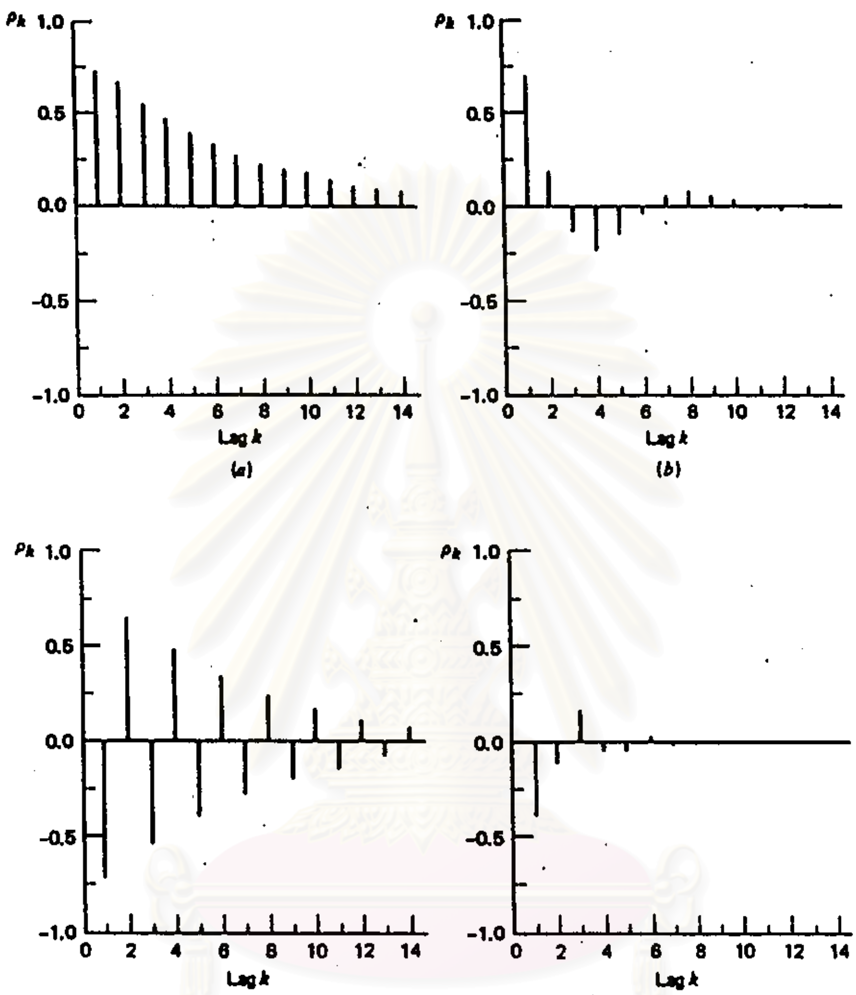
ขั้นตอนที่ 1 *การกำหนดรูปแบบ (Model identification)* เป็นการหารูปแบบที่คาดว่าจะเหมาะสมให้กับอนุกรมเวลา โดยการตรวจสอบกราฟฟังก์ชันอัตตสหสัมพันธ์ (Autocorrelation Function; ACF) ของอนุกรมเวลา $\{y_t\}$ ซึ่งแทนด้วย ρ_k เป็นค่าสัมประสิทธิ์อัตตสหสัมพันธ์ที่ lag k ซึ่ง k หมายถึงคาบเวลาระยะห่างระหว่างอนุกรม เรียกคาบนี้ว่าแลคเค lag k เป็นค่าที่แสดงว่าค่าสังเกตในอนุกรมเวลาที่อยู่ห่างกัน k

ตำแหน่ง มีความสัมพันธ์อย่างไร โดยพิจารณาร่วมกับกราฟฟังก์ชันอัตตสหสัมพันธ์บางส่วน (Partial Autocorrelation Function; PACF) แทนด้วย ϕ_{kk} ซึ่งการพิจารณาค่าสัมประสิทธิ์อัตตสหสัมพันธ์และค่าสัมประสิทธิ์อัตตสหสัมพันธ์บางส่วนนั้นจะพิจารณาจากค่าสัมประสิทธิ์อัตตสหสัมพันธ์ของตัวอย่างและค่าสัมประสิทธิ์อัตตสหสัมพันธ์บางส่วนของตัวอย่าง (Sample Autocorrelation Function; SACF และ Sample Partial Autocorrelation Function; SPACF) จากกราฟ ACF, PACF ที่ได้จึงนำมากำหนดรูปแบบ โดยอนุกรมเวลาที่จะนำมากำหนดรูปแบบนั้นจะต้องอยู่ภายใต้ภาวะคงที่เท่านั้น หากไม่อยู่ภายใต้ภาวะคงที่ต้องแปลงข้อมูลให้อยู่ภายใต้ภาวะคงที่เสียก่อน ถ้าฟังก์ชันอัตตสหสัมพันธ์มีค่าน้อยไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญอาจกล่าวได้ว่าข้อมูลมีแต่อิทธิพลจากเหตุการณ์ผิดปกติเท่านั้น ซึ่งลักษณะของ ACF และ PACF ของกระบวนการภายใต้ภาวะคงที่ แสดงดังภาพที่ 2.4



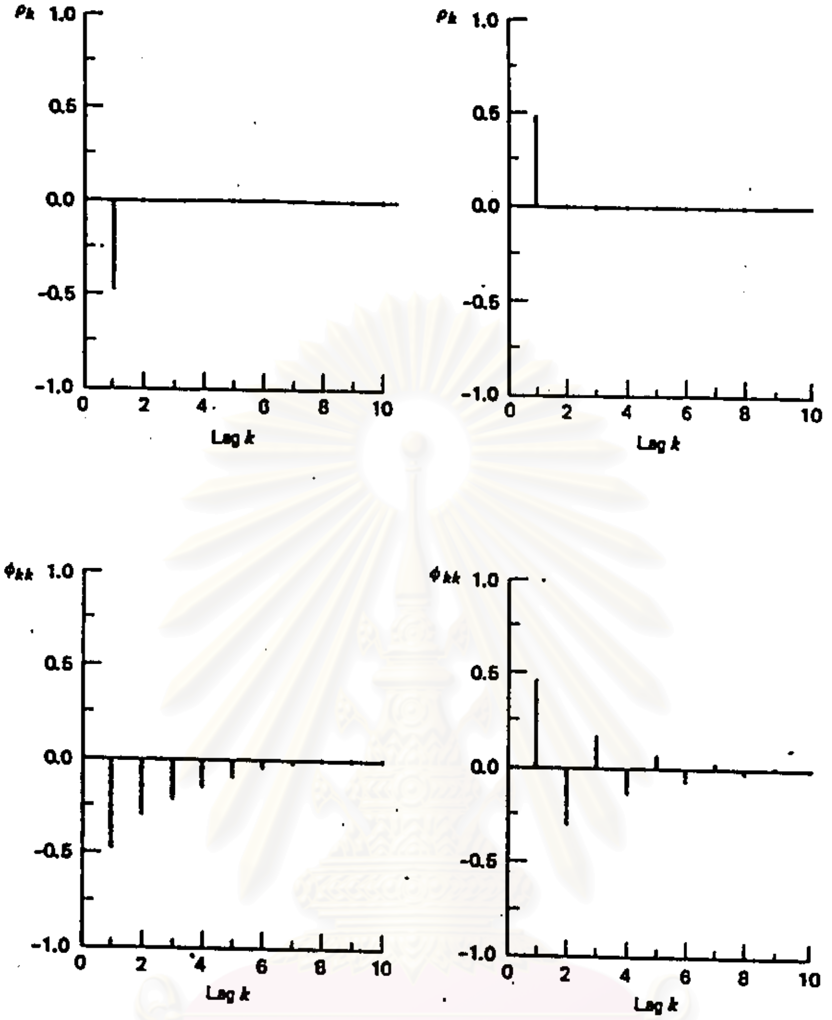
AR(1)

ภาพที่ 2.4 ลักษณะของ ACF และ PACF ของกระบวนการภายใต้ภาวะคงที่



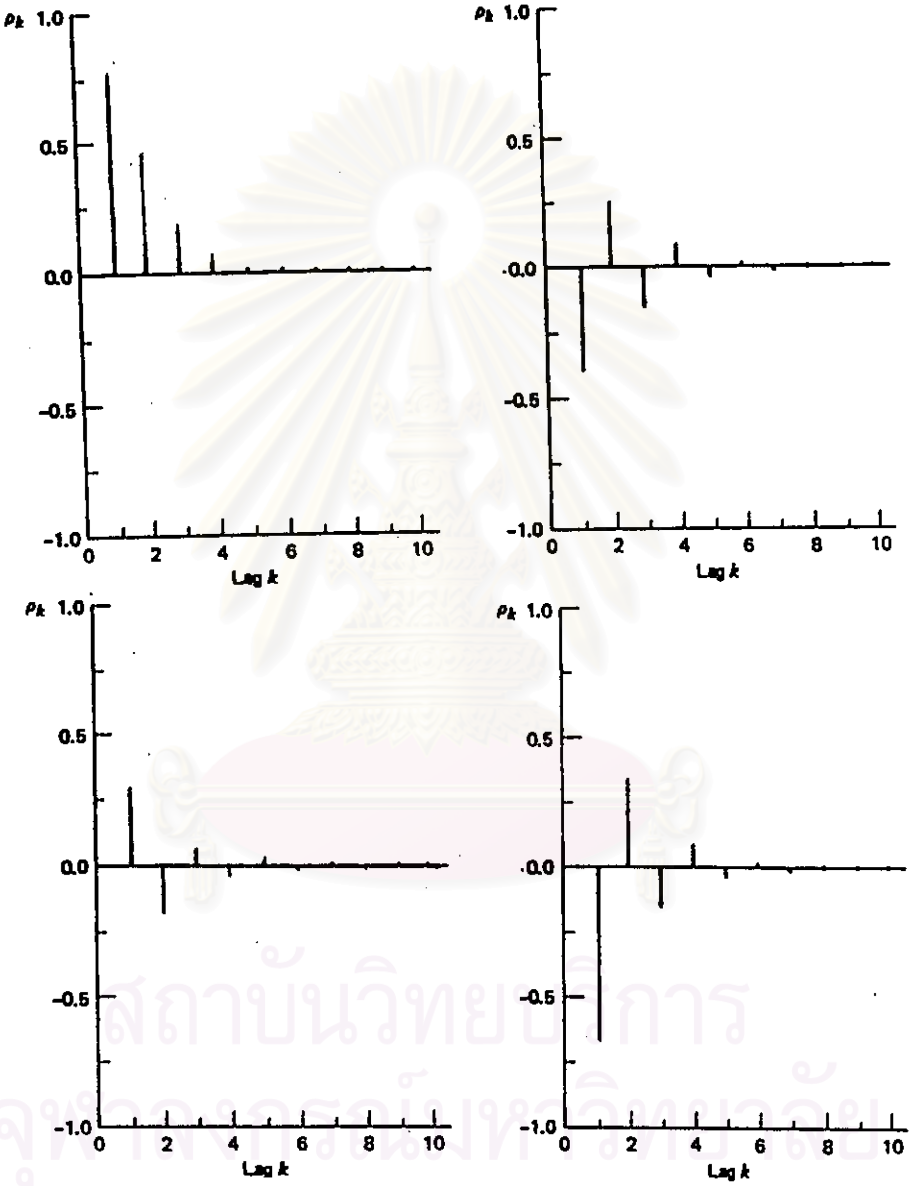
AR(2)

ภาพที่ 2.4 (ต่อ)



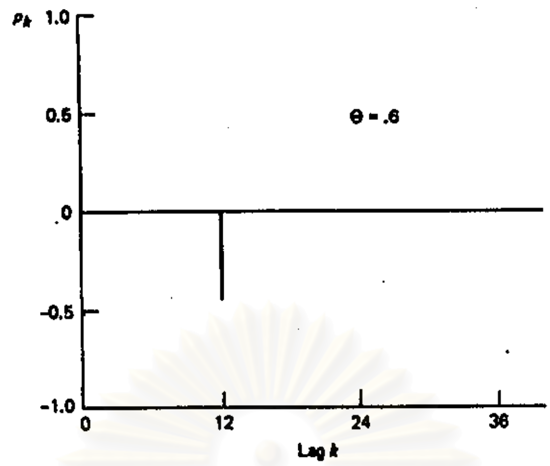
MA(1)

ภาพที่ 2.4 (ต่อ)



ARMA(1,1)

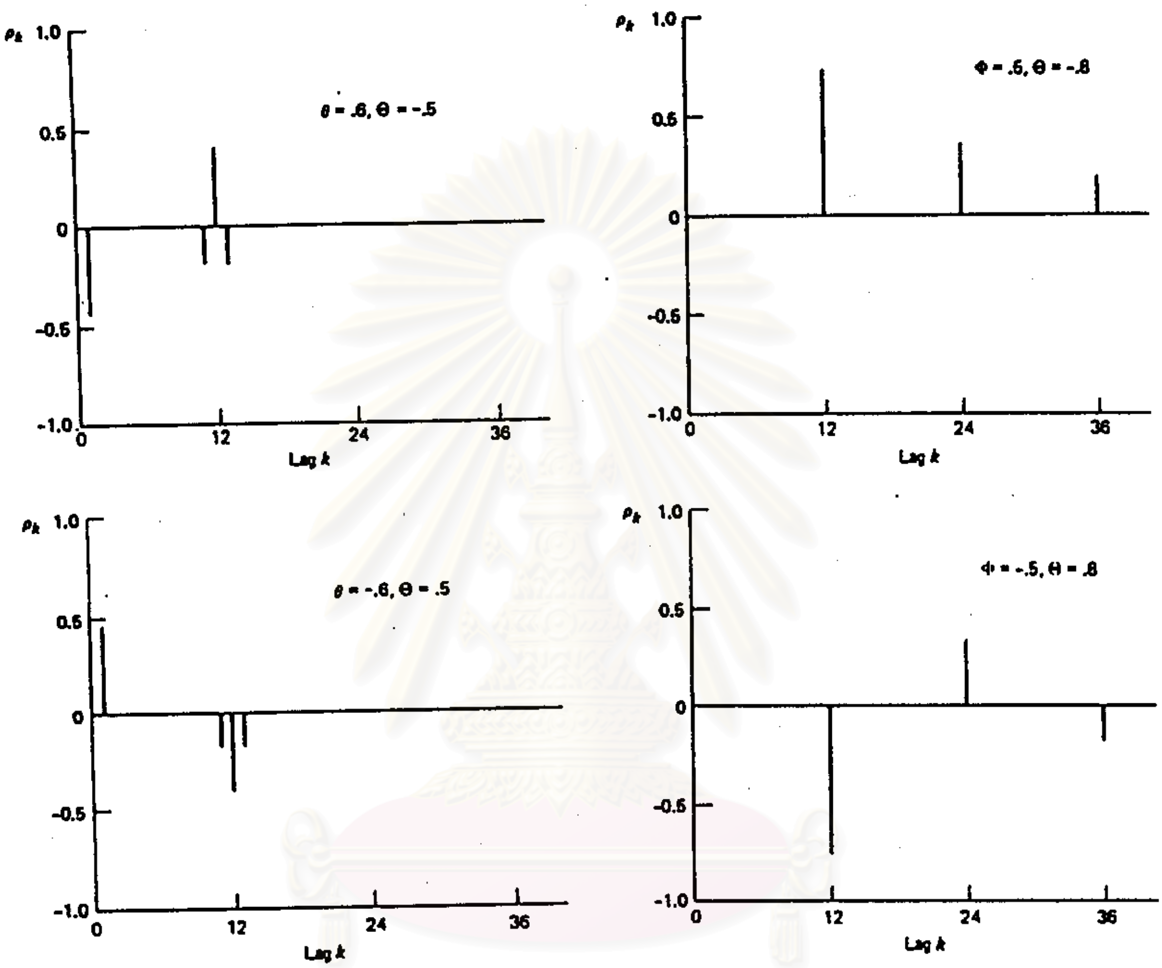
ภาพที่ 2.4 (ต่อ)



ตัวแบบ $w_t = (1-\Theta B^{12})a_t$

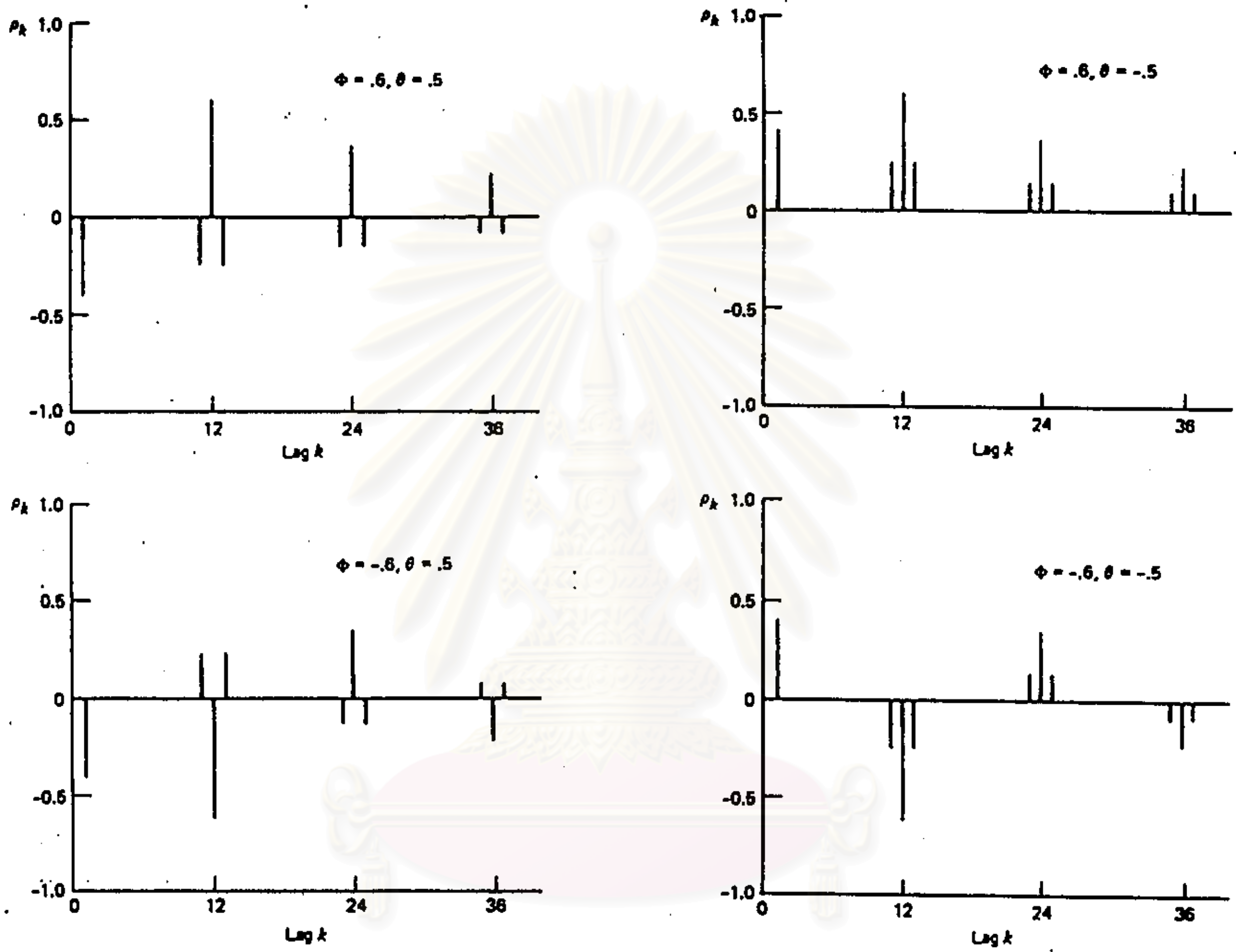
ตัวแบบ $(1-\Phi B^{12})w_t = a_t$

ภาพที่ 2.4 (ต่อ)



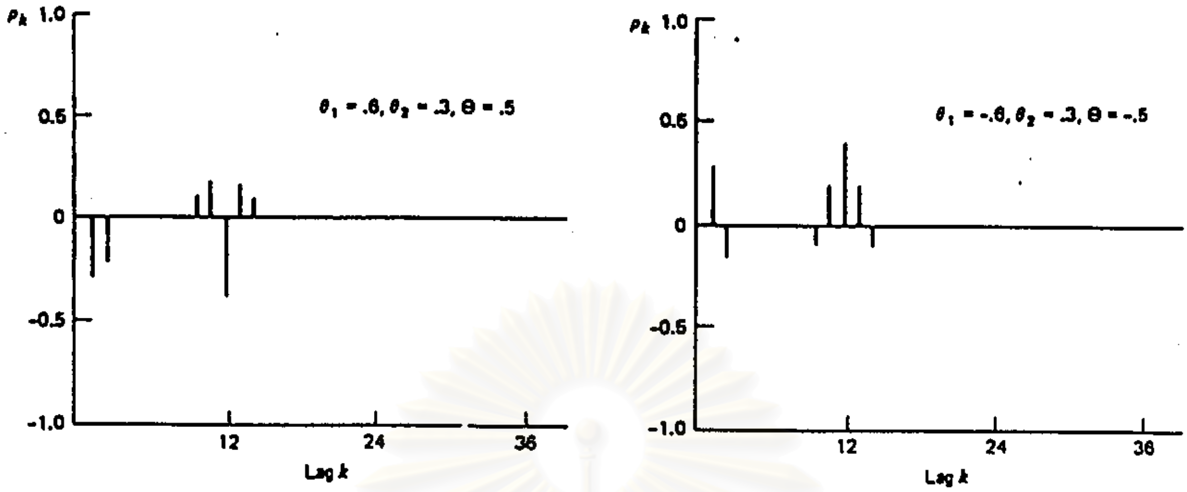
ตัวแบบ $w_t = (1-\theta B)(1-\Theta B^{12})a_t$ **ตัวแบบ** $(1-\Phi B^{12})w_t = (1-\Theta B^{12})a_t$

ภาพที่ 2.4 (ต่อ)

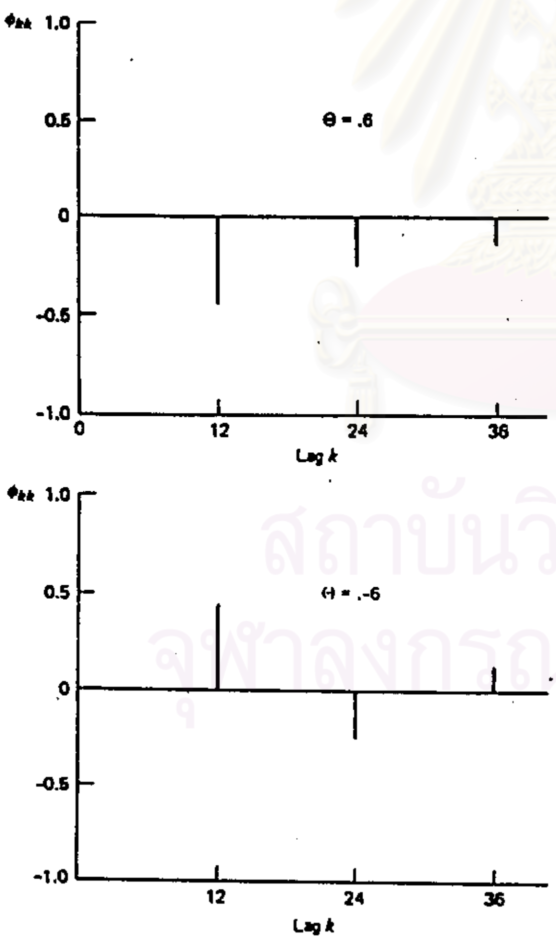


ตัวแบบ $(1-\Phi B^{12})w_t = (1-\theta B)a_t$

ภาพที่ 2.4 (ต่อ)



ตัวแบบ $w_t = (1-\phi_1 B-\phi_2 B)(1-\Theta B^{12})a_t$



ตัวแบบ $w_t = (1-\Theta B^{12})a_t$

ตัวแบบ $(1-\Phi B^{12})w_t = a_t$

ภาพที่ 2.4 (ต่อ)

ขั้นตอนที่ 2 *การประมาณค่าพารามิเตอร์ในรูปแบบ (Model estimation)* โดยการสร้างสมการที่มาจากความสัมพันธ์ระหว่าง p_k และพารามิเตอร์ สมการที่สร้างขึ้นจะมีจำนวนเท่ากับจำนวนพารามิเตอร์ที่ต้องการประมาณส่วนค่าประมาณที่ได้จากการวิเคราะห์ตัวเลขจะได้จากการแก้สมการที่สร้างขึ้นจากวิธีกำลังสองน้อยที่สุด

ขั้นตอนที่ 3 *การตรวจสอบรูปแบบ (diagnostic checking)* เมื่อกำหนดรูปแบบและการประมาณค่าพารามิเตอร์ในรูปแบบแล้ว รูปแบบที่กำหนดอาจจะไม่ใช่รูปแบบที่เหมาะสมดังนั้นจึงจะต้องตรวจสอบทุกครั้งว่ารูปแบบที่กำหนดนั้น มีความเหมาะสมจริงหรือไม่ ซึ่งทำได้หลายวิธีได้แก่

3.1 การพิจารณากราฟความคลาดเคลื่อนของการพยากรณ์ $(\hat{e}_t)$ กับเวลา ถ้าอนุกรมเวลา $\{\hat{e}_t\}$ มีลักษณะการเคลื่อนไหวเป็นอิสระกันจริงโดยค่า $\hat{e}_t$ เป็นค่าประมาณของความคลาดเคลื่อนในประชากร ε_t, t= 1,2,....,T แทนความคลาดเคลื่อนที่มีค่าเฉลี่ยเป็นศูนย์ ไม่มีสหสัมพันธ์ และความแปรปรวนคงที่ แสดงว่าตัวแบบนั้นเหมาะสมกับลักษณะของอนุกรมเวลาแล้ว หรือการพิจารณากราฟ ACF, PACF ของค่าคลาดเคลื่อนของการพยากรณ์ $(\hat{e}_t)$ ซึ่งความถูกต้องของการพยากรณ์ ขึ้นอยู่กับค่าความคลาดเคลื่อนของการพยากรณ์ $(\hat{e}_t)$ ซึ่งเป็นผลต่างของค่าจริงและค่าพยากรณ์ $(\hat{e}_t = y_t - \hat{y}_t)$ ถ้าความคลาดเคลื่อนมากนั่นคือค่าจริงห่างจากค่าพยากรณ์มาก และเพื่อความมั่นใจยิ่งขึ้นจึงต้องทดสอบสมมติฐานค่าความคลาดเคลื่อนโดยใช้ตัวสถิติ Q โดยเปรียบเทียบกับค่า $\chi^2_{k-p-q-P-Q}$ เมื่อ

$$Q = (n)(n+2)\sum_{k=1}^{K}\frac{r_{\hat{a}}^2(k)}{n-k}$$

ถ้า Q< χ^2 สรุปได้ว่ารูปแบบนั้นมีความเหมาะสมแล้ว

3.2 การพิจารณาค่าความแปรปรวนของค่าคลาดเคลื่อน s^2 รูปแบบที่เหมาะสมจะให้ s^2 ต่ำ

3.3 การพิจารณาว่าค่าพารามิเตอร์ในรูปแบบมีค่ามีค่าแตกต่างจากศูนย์หรือไม่ นั่นคือเมื่อ $\phi, \theta, \Phi, \Theta, \delta$ และ σ^2 เป็นค่าพารามิเตอร์ที่ต้องการประมาณในการทดสอบสมมุติฐานจะใช้ตัวสถิติทดสอบ t

ขั้นตอนที่ 4 การพยากรณ์ เป็นการนำสมการพยากรณ์ที่สร้างจากรูปแบบดังกล่าวไปใช้ประโยชน์ โดยการพยากรณ์จะใช้สมการพยากรณ์ที่สร้างจากรูปแบบการพยากรณ์ที่กำหนดและผ่านการตรวจสอบในขั้นตอนที่ผ่านมาแล้ว

ในการเปรียบเทียบตัวแบบพยากรณ์ ด้วยวิธีวัดค่าความถูกต้องของค่าพยากรณ์ โดยใช้ค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percent Error; MAPE)

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

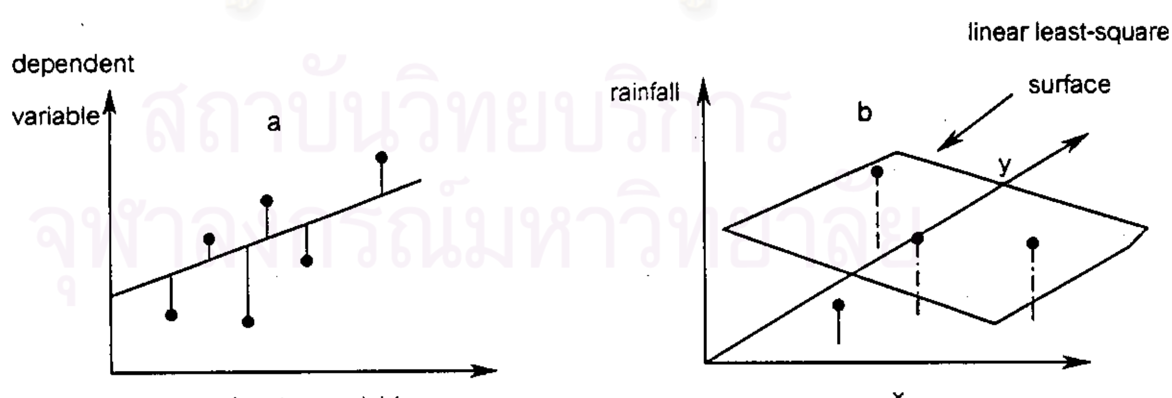
2.2.7 การวิเคราะห์แนวโน้มพื้นผิว โดยใช้ตัวแบบการถดถอยเชิงเส้นพหุ (Multiple Linear Regression) โดยมีรูปแบบคือ

$$Y = f(x;\beta) + \varepsilon_t$$

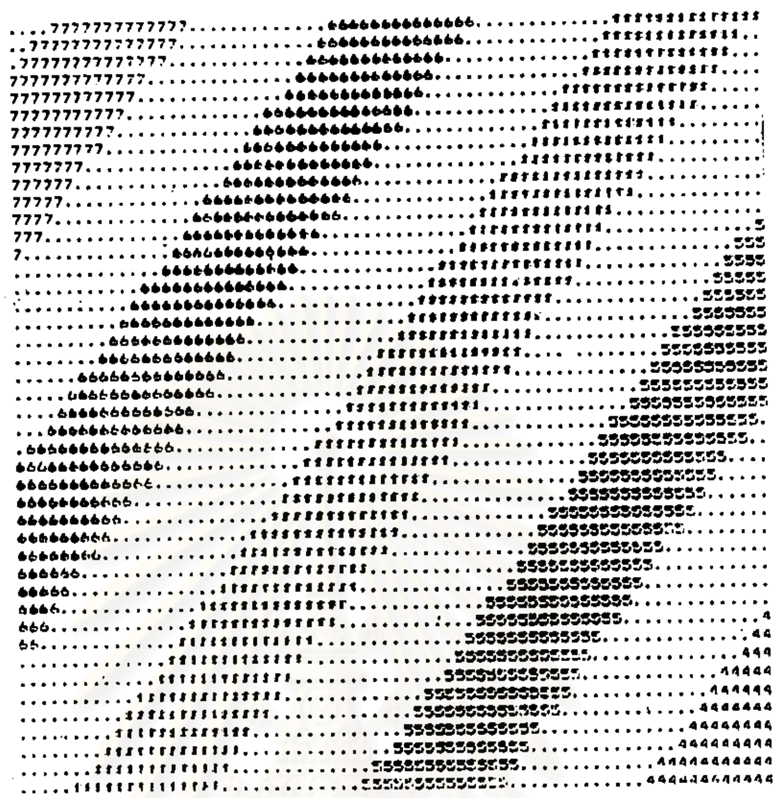
โดยมีข้อสมมติว่า $E(\varepsilon_t) = 0$ และ $V(\varepsilon_t) = \sigma^2$ ความคลาดเคลื่อน ε_t ไม่มีสหสัมพันธ์กัน $Cov(\varepsilon_t, \varepsilon_{t-k}) = E(\varepsilon_t\varepsilon_{t-k}) = 0$, ทุกค่า $k \neq 0$ และสมการพยากรณ์คือ

$$\hat{y} = rainfall = a + b_1 LAT - b_2 LONG$$

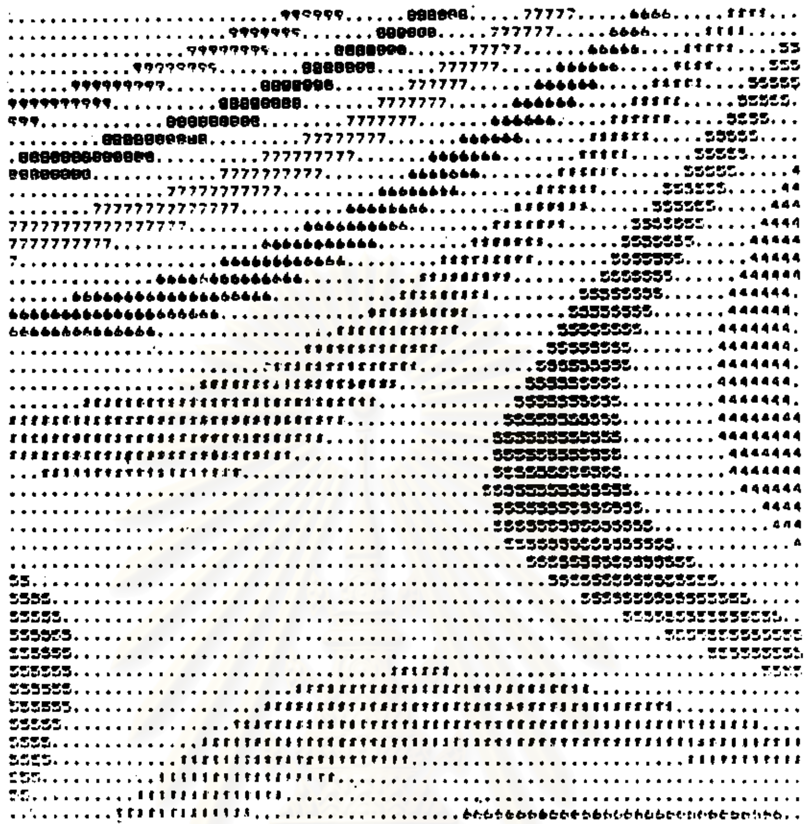
เมื่อ LAT, LONG คือตำแหน่งที่ตั้งเส้นรุ้ง และเส้นแวงของสถานีในจังหวัดต่าง ๆ ตามลำดับ



ภาพที่ 2.5 (a) แสดงสมการการถดถอยเชิงเส้นอย่างง่าย Y = a+bX. ด้วยวิธีกำลังสองน้อยที่สุด
(b) แสดงพื้นผิวแนวโน้มระหว่าง x และ y



ภาพที่ 2.6 แสดงพื้นผิวแนวโน้มในรูปเชิงเส้น



ภาพที่ 2.7 แสดงพื้นผิวแนวโน้มในรูปโพลีโนเมียล