บทที่ 6
ทฤษฎีที่ใช้วิเคราะห์ความไม่คงตัวของข้อมูลอุทกวิทยา

6.1 คำนำ

ข้อมูลอุทกวิทยาจัดเป็นข้อมูลที่มีลักษณะเป็นอนุกรมเวลา คือ ค่าของข้อมูลจะวัดติดต่อกันเป็นระยะ ๆ และเป็นกลุ่มของการสังเกตการณ์ที่ช่วงเวลาที่สม่ำเสมอ และมีคุณลักษณะส่วนใหญ่ คือ มีค่าการเปลี่ยนแปลงตามกาลเวลาสูง ซึ่งลักษณะที่ข้อมูลมีการเปลี่ยนแปลงตามกาลเวลา ก็คือ ข้อมูลมีความไม่คงตัว

ลักษณะความไม่คงตัวของข้อมูลจะมีรูปแบบต่าง ๆ ได้แก่ แนวโน้ม แนวโน้มลักษณะวงจร การเปลี่ยนแปลงตามฤดูกาล และการเปลี่ยนแปลงแบบสุ่ม ซึ่งวิธีการวิเคราะห์ลักษณะดังกล่าวของข้อมูลอุทกวิทยา มีด้วยกันหลายวิธี โดยสามารถแบ่งวิธีออกได้เป็น 2 แบบ คือ การวิเคราะห์แบบใช้ตัวพารามิเตอร์ และการวิเคราะห์แบบไม่ใช้ตัวพารามิเตอร์ ในการศึกษาได้เลือกใช้วิธีการวิเคราะห์ทั้งหมด 4 วิธี เป็นวิธีแบบใช้ตัวพารามิเตอร์ 3 วิธี ได้แก่ วิธี Moving Average การวิเคราะห์ด้วยกราฟค่าสะสมเฉลี่ย การตรวจสอบด้วยค่าสหสัมพันธ์ภายในตัวแปร และวิธีแบบไม่ใช้ตัวพารามิเตอร์ 1 วิธี ได้แก่ การตรวจสอบด้วยค่าสัมประสิทธิ์ความสัมพันธ์ตามตำแหน่งของ Spearman

6.2 ทฤษฎีที่ใช้ในการวิเคราะห์แบบใช้ตัวพารามิเตอร์

วิธีการวิเคราะห์ความไม่คงตัว ของข้อมูลอุทกวิทยา แบบใช้ตัวพารามิเตอร์ ที่เลือกใช้ในการวิเคราะห์ความไม่คงตัวของข้อมูลน้ำฝน จากสถานีตัวแทนในลุ่มน้ำปิงมี 3 วิธี ซึ่งเป็นวิธีที่นิยมใช้กันในการวิเคราะห์แนวโน้ม และวิธีการคำนวณทางสถิติที่แสดงลักษณะความไม่คงตัว และแนวโน้มของข้อมูลที่มีลักษณะเป็นอนุกรมเวลาได้ ดังต่อไปนี้

6.2.1 Moving Average

การวิเคราะห์ด้วยวิธี Moving Average เป็นวิธีที่นิยมใช้ในการวิเคราะห์แนวโน้ม มีขั้นตอนการวิเคราะห์ ดังต่อไปนี้

1. จัดเรียงข้อมูลที่ตรวจสอบแล้ว

2. หาค่าเฉลี่ย ($\bar{x}$) ของข้อมูลทั้งชุด

3. หาค่า 3-year Moving Average 5-year Moving Average และ 10-year Moving Average ของชุดข้อมูล ตามลำดับ ด้วยสมการดังนี้

$$\left[\frac{\sum_{i=1}^{m} x_i}{m}, \frac{\sum_{i=1}^{m} x_{i+1}}{m}, \frac{\sum_{i=1}^{m} x_{i+2}}{m}, \ldots, \frac{\sum_{i=1}^{m} x_{i+n}}{m}\right] \qquad (6.2.1)$$

x_i = ข้อมูลน้ำฝนรายปีหรือรายเดือน (มิลลิเมตร)

n = N - m

N = จำนวนข้อมูล (ปี หรือ เดือน)

m = จำนวนข้อมูลที่ใช้เฉลี่ย (ปี หรือ เดือน) และในการวิเคราะห์นี้ใช้ m = 3 , 5 และ 10 ตามลำดับ

4. วาดกราฟอนุกรมเวลาของชุดข้อมูล ระดับค่าเฉลี่ย ($\bar{x}$) และเส้นกราฟ 3-year Moving Average 5-year Moving Average และ 10-year Moving Average ในกราฟรูปเดียวกัน

5. การวิเคราะห์แนวโน้ม (Trend) และแนวโน้มลักษณะวงจร (Cycle Trend) จากเส้นกราฟทั้งสาม

6. ใช้การถดถอยเชิงเส้นตรง เพื่อเป็นตัวแทนเส้น Moving Average ที่วาดไว้ ด้วยการลากเส้นตรงแทนเส้น 3-year Moving Average 5-year Moving Average และ 10-year Moving Average ด้วยเส้นตรง โดยใช้วิธีกำลังสองน้อยที่สุด

สมการของเส้นถดถอย ได้แก่

$$y = \alpha_1 + \beta_1 x \qquad (a)$$

และ $$x = \alpha_2 + \beta_2 y \qquad (b)$$

7. ทำการทดสอบสมมติฐานว่าแนวโน้มจริงหรือไม่จากเส้นตรงที่ลากแทนเส้น Moving Average ทั้ง 3 แบบ โดยใช้ตัวทดสอบสถิติ Student - t

สมมติฐานเพื่อการทดสอบ

H_o : β_1 (Slope) ไม่แตกต่างจากศูนย์

สมมติฐานแย้ง

H_1 : β_1 แตกต่างจากศูนย์ และมีค่าได้ทั้งมากกว่าและน้อยกว่าค่าประชากร

จึงใช้การทดสอบด้วย Student - t แบบทดสอบสองด้าน (Two-Tail Test)

ถ้าหากความสหสัมพันธ์ แสดงว่า ตัวแปร x และ y มีความสัมพันธ์แบบเส้นตรง เส้นถดถอยแบบเส้นตรงสำหรับตัวแปร y ต่อตัวแปร x ก็คือ เส้นตรงที่ให้ค่าการประมาณที่ดีที่สุดของ y ต่อค่าของ x ที่กำหนด และในทำนองเดียวกัน คือ เส้นตรงที่ให้ค่าการประมาณที่ดีที่สุดของ x ต่อค่าของ y ที่กำหนด

$$y = a_1 + b_1 x \qquad (c)$$

และ $$x = a_2 + b_2 y \qquad (d)$$

b_1 , b_2 = สัมประสิทธิ์การถดถอยของ y ต่อ x และ x ต่อ y ตามลำดับ

a_1 , a_2 = Intercepts y ต่อ x และ x ต่อ y ตามลำดับ

$$b_1 = \frac{s_{xy}}{s_x^2} = \frac{\sum_{i=1}^{N} x_i y_i - N\overline{xy}}{\sum_{i=1}^{N} x_i^2 - N\overline{x}^2} \qquad (e)$$

$$a_1 = \overline{y} - b_1\overline{x} \qquad (f)$$

การแจกแจงแบบ Student - t ; NDF = N-2

$$t = b_1 \cdot \sqrt{\frac{(N-2)s_x^2}{s_{ey}^2}} = \sqrt{r^2 \frac{N-2}{1-r^2}} \qquad (g)$$

เปรียบเทียบค่า t กับ $t_{1-\alpha/2\,;\,N-2}$

ถ้า $|t| > t_{1-\alpha/2\,;\,N-2}$ ....ปฏิเสธสมมติฐาน

ขีดความเชื่อมั่น

$$(b_1)_1 = b_1 + \frac{b_1}{r} t_p \sqrt{\frac{1-r^2}{N-2}} \qquad (h)$$

$$(b_1)_2 = b_1 - \frac{b_1}{r} t_p \sqrt{\frac{1-r^2}{N-2}} \qquad (i)$$

โดย ตัวแปร x คือ ค่าตามลำดับของเส้น Moving Average ที่จะทดสอบ

ตัวแปร y คือ ค่าตามลำดับของเส้นตรงที่ลากแทนเส้น Moving Average ที่จะทดสอบ

$t_{1-\alpha/2\,;\,N-2}$ จากตาราง Cumulative Student - t Distribution โดยในการทดสอบใช้สมมติฐานที่ Two-Tail Test และมีขีดความเชื่อมั่น 95%

ดังนั้นเปิดตารางที่ $1-\alpha/2 = 1 - \frac{(1-0.95)}{2} = 0.975$

$n = N - 2$

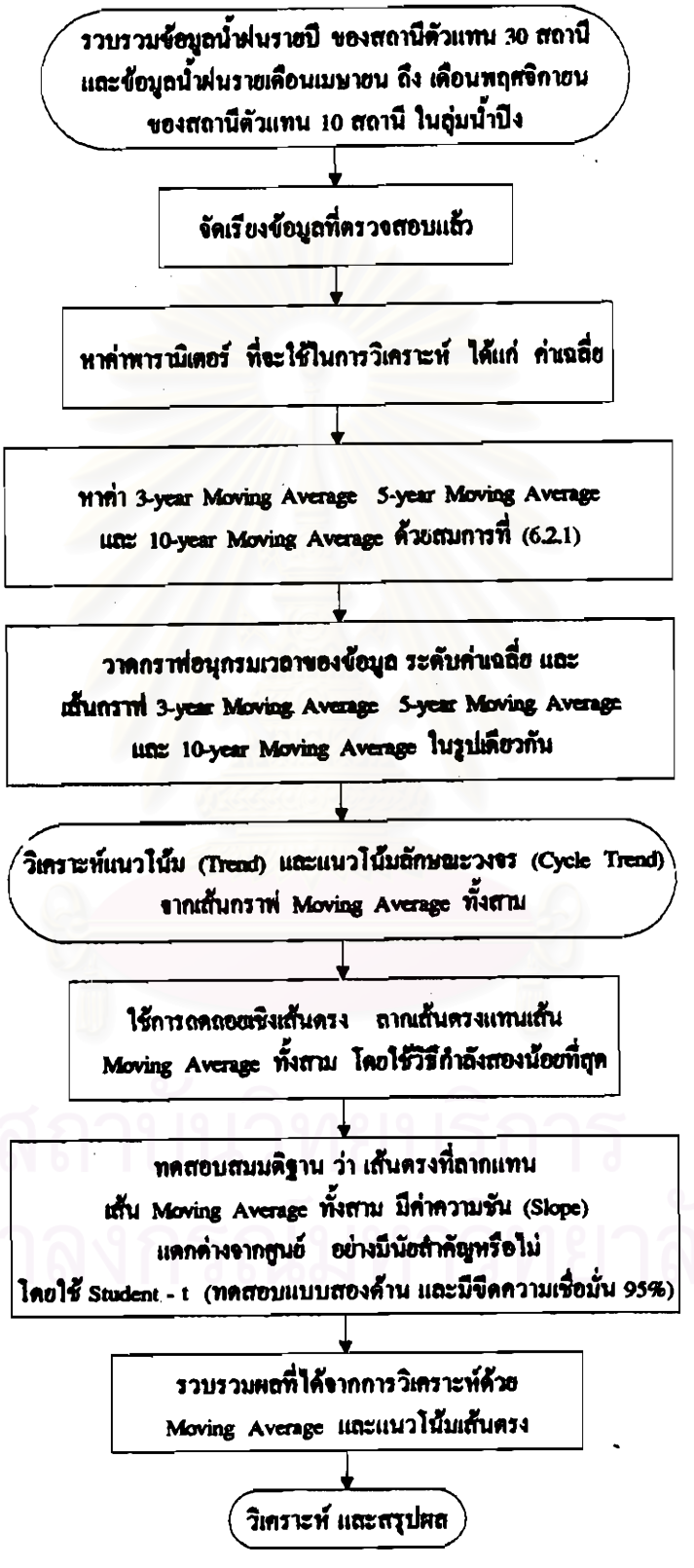
โดย N = จำนวนข้อมูล

8. รวบรวมผลที่ได้จากการวิเคราะห์ด้วย Moving Average และแนวโน้มเส้นตรง (Linear Trend)

9. สรุปผล

ซึ่งได้ใช้การวิเคราะห์ด้วย Moving Average กับข้อมูลปริมาณน้ำฝนรายปี ของสถานีตัวแทนทั้ง 31 สถานี และใช้ Moving Average ในการวิเคราะห์ข้อมูลน้ำฝนรายเดือน เดือนเมษายน ถึงเดือนพฤศจิกายน ของสถานีตัวแทน 10 สถานี ที่เลือกจากกลุ่มพื้นที่ส่วนต่าง ๆ ที่แบ่งไว้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 6.1 แผนภาพแสดงขั้นตอนการวิเคราะห์ด้วยวิธี Moving Average

6.2.2 วิเคราะห์ด้วยกราฟค่าสะสมเฉลี่ย

ด้วยสมมติฐานที่ว่า ถ้าข้อมูลไม่มีแนวโน้ม หรือมีความแปรปรวนไม่มาก ค่าของข้อมูลตามลำดับเวลาจะมีค่าใกล้เคียงกัน ดังนั้น ค่าเฉลี่ยของข้อมูล ซึ่งไม่ว่าจะเป็นการเฉลี่ยกี่ปีก็น่าที่จะไม่แตกต่างกัน

จึงทำการวิเคราะห์ด้วย การวาดกราฟค่าสะสมเฉลี่ย และลากเส้นแทนเส้นสะสมเฉลี่ยด้วยสมการเส้นตรงโดยวิธีกำลังสองน้อยที่สุด และวาดแผนภูมิแท่งจำนวนสถานีในแต่ละช่วงของค่าสัมบูรณ์ของความลาดชัน (| Slope |) และความถี่สถานีในแต่ละช่วง | Slope |

ขั้นตอนในการวิเคราะห์มีดังต่อไปนี้

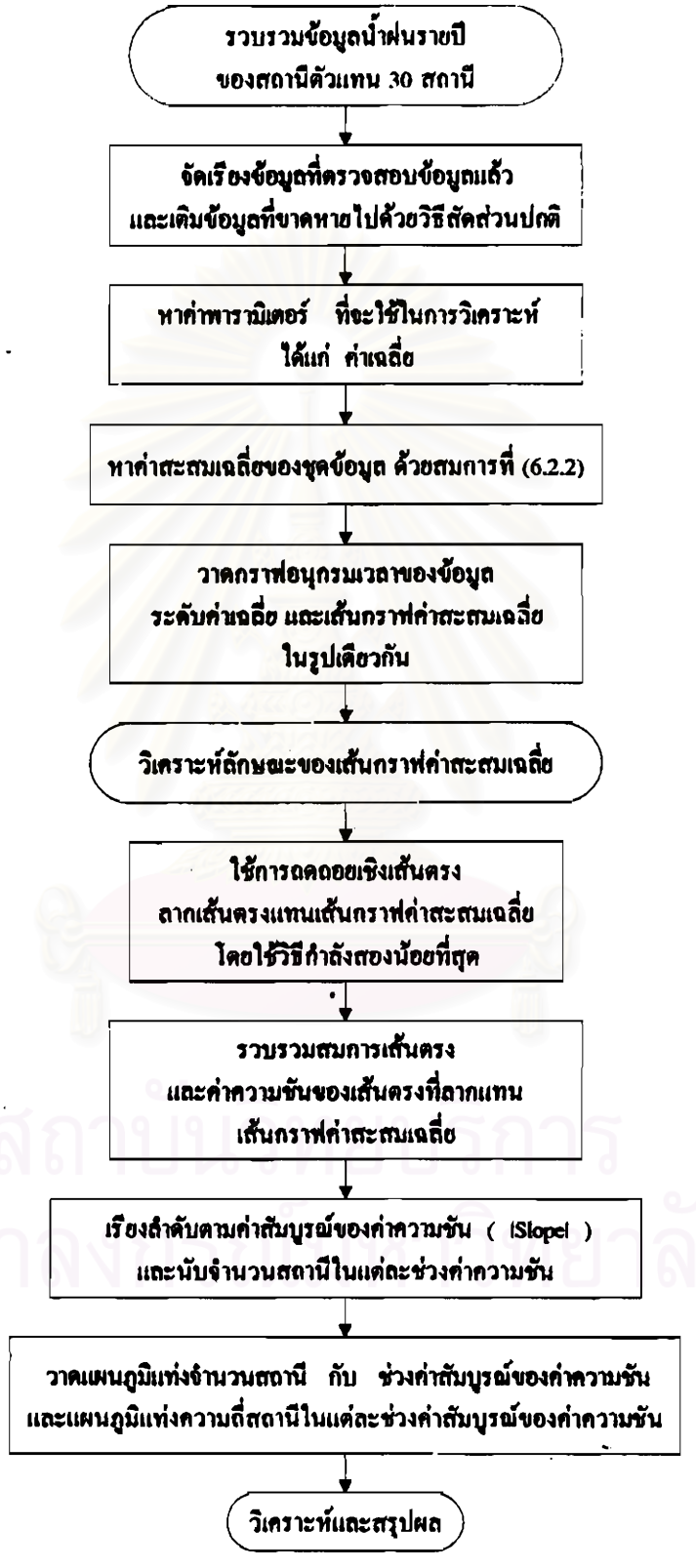
1. จัดเรียงข้อมูลตามลำดับเวลา
2. เนื่องจากต้องการใช้ข้อมูลที่ต่อเนื่องตามลำดับจึงเติมข้อมูลที่ขาดหายไปด้วยวิธีสัดส่วนปกติ
3. หาค่าสะสมเฉลี่ยของชุดข้อมูลดังสมการต่อไปนี้

$$\left[x_1, \frac{x_1+x_2}{2}, \frac{x_1+x_2+x_3}{3}, \ldots, \frac{x_1+x_2+\ldots+x_N}{N}\right] \qquad (6.2.2)$$

$x_1, x_2, x_3, \ldots, x_N$ = ข้อมูลปริมาณน้ำฝนตามลำดับเวลา (มิลลิเมตร)

N = จำนวนข้อมูล (ปี)

4. นำชุดข้อมูลค่าสะสมเฉลี่ยวาดกับกราฟอนุกรมเวลาของข้อมูล
5. วิเคราะห์ลักษณะของเส้นกราฟค่าสะสมเฉลี่ย
6. ลากเส้นแทนกราฟค่าสะสมเฉลี่ยด้วยสมการเส้นตรงโดยวิธีกำลังสองน้อยที่สุด เพื่อเป็นตัวแทนลักษณะการเปลี่ยนแปลงตามเวลาของเส้นกราฟ
7. รวบรวมสมการเส้นตรง และค่าความชันของเส้นตรงที่ลากแทนกราฟค่าสะสมเฉลี่ย
8. เรียงลำดับตามค่า | Slope | และนับจำนวนสถานีในแต่ละช่วงของค่า | Slope |
9. วาดแผนภูมิแท่ง จำนวนสถานีกับช่วง | Slope | และแผนภูมิแท่งความถี่สถานี ในแต่ละช่วง | Slope |
10. วิเคราะห์และสรุปผลที่ได้จากแผนภูมิแท่งทั้ง 2



รูปที่ 6.2 แผนภาพแสดงขั้นตอนการวิเคราะห์ด้วยวิธี กราฟค่าสะสมเฉลี่ย

6.2.3 การตรวจสอบด้วยค่าสหสัมพันธ์ภายในตัวแปร (Autocorrelation)

ความคงตัวของอนุกรมเวลา คือ ไม่มีการเปลี่ยนแปลงใดๆ ตามกาลเวลาในอนุกรม ถ้าข้อมูลอุทกวิทยามีความคงตัว หมายถึง ฟังก์ชั่นความหนาแน่นน่าจะเป็น (Probability Density Function) ของข้อมูลที่เวลาหนึ่ง จะเท่ากับ อีกที่เวลาหนึ่ง ในอนุกรมเดียวกัน ซึ่งสามารถแสดงได้ด้วยค่าความแปรปรวนร่วม (Covariance) ของชุดข้อมูลที่เวลาหนึ่งกับอีกเวลาหนึ่ง

ค่าความแปรปรวนร่วมเป็นค่าทางสถิติที่ใช้อธิบายถึง ความเป็นอิสระต่อกันของข้อมูล 2 ตัวแปร คือ ถ้าค่าความแปรปรวนร่วมของตัวแปรทั้ง 2 เท่ากับศูนย์ หมายถึง ตัวแปรทั้ง 2 ไม่มีความสหสัมพันธ์กันเชิงเส้นตรง แต่อาจจะแปรตามกันในรูปแบบอื่นก็ได้

ดังนั้น ค่าความแปรปรวนร่วม ของชุดข้อมูลที่เวลาหนึ่ง ($x(t)$) กับอีกเวลาหนึ่ง ($x(t+\tau)$) โดย τ คือ Lag Time จึงอาจอธิบายลักษณะความไม่คงตัวและแนวโน้มของข้อมูลอุทกวิทยาได้

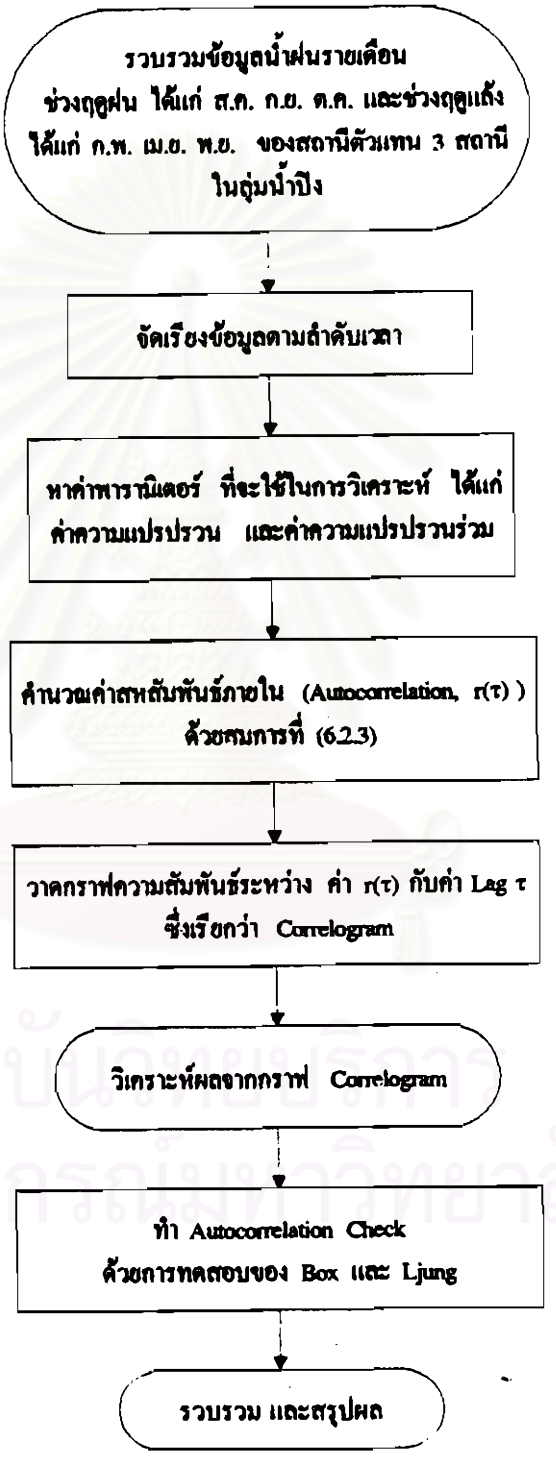
ขั้นตอนการวิเคราะห์มีดังต่อไปนี้

1. จัดเรียงข้อมูลตามลำดับเวลา
2. คำนวณหาค่าความแปรปรวน และความแปรปรวนร่วม
3. คำนวณค่า Autocorrelation ด้วยสมการดังนี้

$$r(\tau) = \frac{\mathrm{cov}(x(t), x(t+\tau))}{\left[\mathrm{var}(x(t)) \cdot \mathrm{var}(x(t+\tau))\right]^{0.5}} \qquad (6.2.3)$$

4. วาดกราฟความสัมพันธ์ระหว่างค่า $r(\tau)$ กับค่า Lag τ
5. วิเคราะห์ผลจากกราฟ และทำ Autocorrelation Check ด้วยการทดสอบของ Box และ Ljung

กราฟความสัมพันธ์ระหว่างค่า $r(\tau)$ กับค่า Lag τ เรียกว่า Correlogram ซึ่งสามารถใช้อธิบายความไม่คงตัวและลักษณะวงจรของข้อมูลอุทกวิทยาได้



รูปที่ 6.3 แผนภาพแสดงขั้นตอนการวิเคราะห์ด้วยวิธีการตรวจสอบด้วยค่าสหสัมพันธ์ภายในตัวแปร

6.3 ทฤษฎีที่ใช้ในการวิเคราะห์แบบไม่ใช้ตัวพารามิเตอร์

วิธีการวิเคราะห์ความไม่คงตัว และแนวโน้มของข้อมูลอุทกวิทยา แบบไม่ใช้ตัวแปรพารามิเตอร์ที่เลือกใช้ เป็นการวิเคราะห์ด้วยการเรียงลำดับค่าของข้อมูล และเป็นวิธีการใหม่ ที่ยังไม่เป็นที่นิยมใช้ในประเทศไทยซึ่งจากการศึกษาที่ผ่านมายังไม่พบผู้นำวิธีนี้มาใช้วิเคราะห์แนวโน้มในประเทศไทย แต่ผู้แต่ง (Dahmen, E. R. และ Hall, M. J., 1989) ได้ใช้ข้อมูลน้ำฝนรายปีของกรุงเทพมหานคร ในการทดสอบและนำเสนอวิธีการ Spearman's Rank - Correlation

6.3.1 การตรวจสอบด้วยค่าสัมประสิทธิ์ความสัมพันธ์ตามตำแหน่งของ Spearman

เป็นวิธีการทดสอบแนวโน้ม ซึ่งเป็นส่วนหนึ่งของวิธีการแยกแยะส่วนประกอบต่าง ๆ (Screen) ของข้อมูลอุทกวิทยาโดย Dahmen, E. R. และ Hall, M. J. ซึ่งเป็นการวิเคราะห์โดยการใช้สัมประสิทธิ์ความสัมพันธ์ตามตำแหน่งของ Spearman , r_{sp} (Spearman Rank-Correlation Coefficient)

โดย
$$r_{sp} = 1 - \frac{6 \cdot \sum_{i=1}^{N} \left(D_i^2\right)}{N \cdot \left(N^2 - 1\right)} \qquad (6.3.1)$$

N = จำนวนข้อมูล (ปี)

D_i = $Kx_i - Ky_i$ (m)

Kx_i = ตำแหน่งของตัวแปร x ซึ่งเป็นหมายเลขลำดับตามเวลาที่ได้จากการสังเกต

Ky_i = ตำแหน่งของอนุกรมที่ได้จากการสังเกต โดยเรียงลำดับค่าของตัวแปร x จากน้อยไปมาก

และมีการทดสอบสมมติฐาน ดังนี้

สมมติฐานเพื่อการทดสอบ , H_o ถ้า $\rho_{sp} = 0$ แสดงว่าไม่มีแนวโน้มในชุดข้อมูล

สมมติฐานแย้ง , H_1 ถ้า $\rho_{sp} <> 0$ แสดงว่า แนวโน้มในชุดข้อมูล

โดยการทดสอบด้วยวิธีการ Student - t และมีค่าองศาอิสระ , ν (Degree of Freedom) = N-2 ที่ Two-Tail Test และมีขีดความเชื่อมั่น 95 %

$$\text{โดย} \quad t = r_{sp}\left[\frac{N-2}{1-r_{sp}^2}\right]^{0.5} \qquad (n)$$

มีขั้นตอนการวิเคราะห์ดังต่อไปนี้

1. จัดเรียงข้อมูลตามลำดับเวลา

2. คำนวณค่า สัมประสิทธิ์ความสัมพันธ์ตามตำแหน่งของ Spearman , r_{sp}

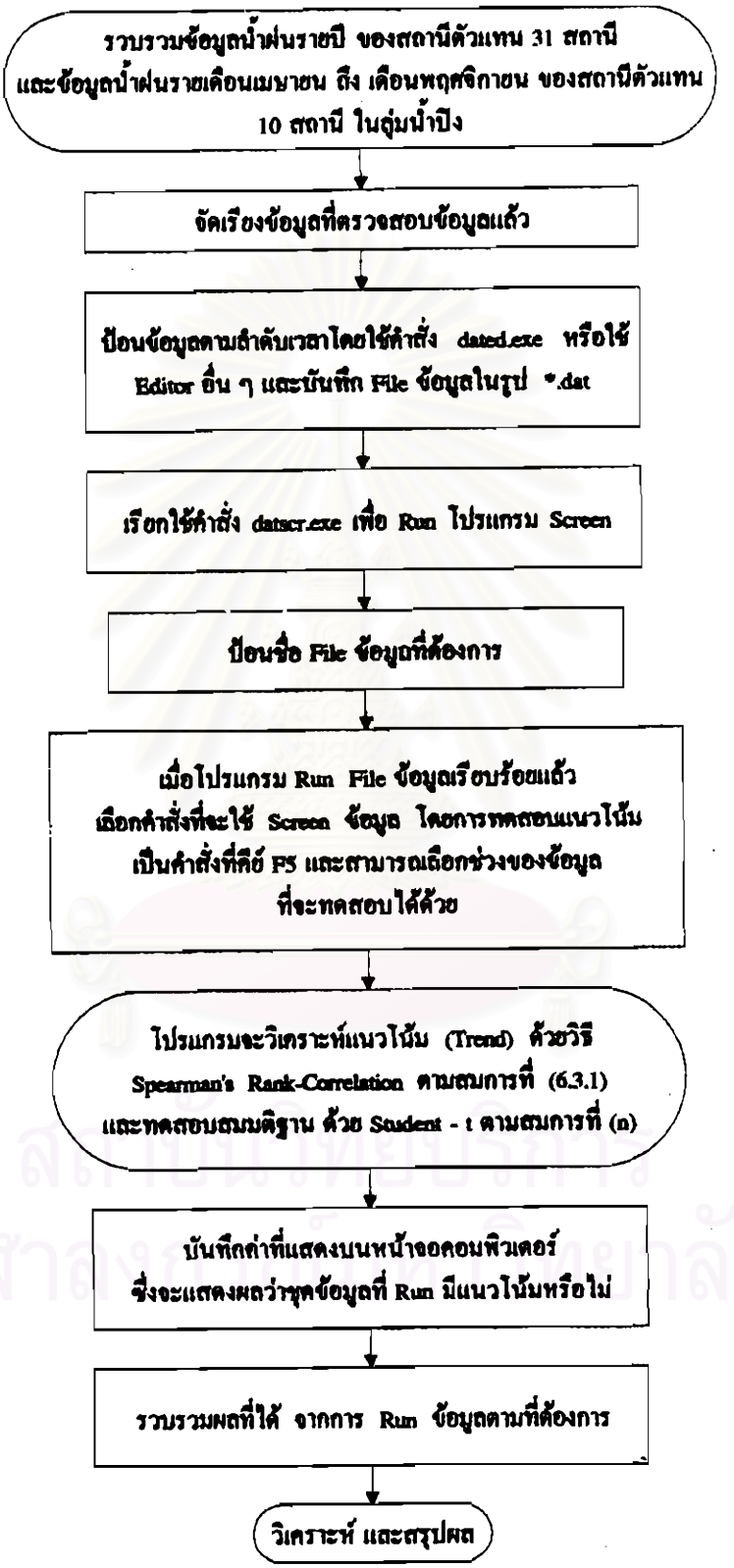
3. ทดสอบสมมติฐานว่าในชุดข้อมูลมีแนวโน้มหรือไม่ โดยการคำนวณค่า t และเปรียบเทียบค่าที่คำนวณได้ กับค่า t จากตาราง Student - t (ภาคผนวก ซ)

4. บันทึกผลการทดสอบแนวโน้ม

7. รวบรวมผลที่ได้จากการทดสอบ

8. สรุปและวิเคราะห์ผลที่ได้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 6.4 แผนภาพแสดงขั้นตอนการทำงานเพื่อวิเคราะห์ ด้วยวิธี
การตรวจสอบด้วยค่าสัมประสิทธิ์ความสัมพันธ์ของ Spearman