บทที่ 3

ทฤษฎีที่นำมาใช้ในการพิจารณาเพื่อการศึกษา

3.1 สถิติของข้อมูลอุทกวิทยา (Basic Statistical of Hydrological data)

ในข่ายงานทางอุทกวิทยา ข้อมูลส่วนใหญ่ที่ได้มา จะเป็นรูปแบบของอนุกรมเวลา (Time Series) ดังนั้นข้อมูลที่ได้เหล่านี้จะนำมาเป็นตัวแทนของข้อมูลได้ จะต้องมีการเก็บข้อมูลเป็นเวลาต่อเนื่องไม่ขาดตอน และควรจะมีช่วงเวลาในการเก็บข้อมูลยาวนานพอสมควร ทางด้านอุทกวิทยา ค่าของตัวแปรแต่ละตัว (Variate) จะเป็นตัวแปรอิสระ (Independent Variable) และตัวแปร(Variable) ในสมการถูกเรียกว่า ตัวแปรสุ่ม (Random Variable) ดังปรากฎการณ์ที่ถูกกำหนดลักษณะให้มีค่าที่แน่นอน โดยให้แปรเปลี่ยนตามกาลเวลา และช่องว่างแห่งปรากฎการณ์ อุทกวิทยาลักษณะนี้คือตัวแปร (Variable) และค่าเฉพาะของตัวแปร คือ Variate เช่น ปริมาณน้ำเป็น Variable และปริมาณน้ำหลาก (Flood Discharge) จะเป็น Variate

การสุ่มคุณลักษณะของตัวแปรทางอุทกวิทยา มีความสำคัญที่สุดเพื่อใช้ในกรรมวิธีทางสถิติ อนุกรมเวลาทั้งหมด อาจจะถูกแสดงออกมาให้เห็นโดยพารามิเตอร์ทางสถิติ (Statistical Parameter)

3.2 พารามิเตอร์หลักของสถิติอนุกรม

โดยทั่ว ๆ ไป จุดประสงค์ในการศึกษาทางด้านสถิติ เพื่อรวบรวมข่าวสารที่ได้จากการสังเกตุเข้าไว้ด้วยกันเป็นอนุกรม ผลจากการคำนวณได้จัดแบ่งเป็น 2 กลุ่มด้วยกัน คือ

ก. ค่ากลาง, ค่าเด่น หรือค่าแนวโน้มของอนุกรม
(The central or dominant or trend value of the series)

ข. การกระจาย หรือการแปรเปลี่ยนของอนุกรมรอบค่ากลาง
(The dispersion or fluctuation of the various terms of the series round the above central value)

3.2.1 การวัดค่ากลาง (Measurement of Central Value)

จุดมุ่งหมายของพารามิเตอร์นี้ เพื่อแสดงลักษณะโดยค่า (typical value) ซึ่งเป็นเลขจำนวนเดียว (Single Number) ซึ่งเป็นการประเมินอย่างหยาบ ๆ เท่านั้น แต่จะเป็นการประเมินอย่างดีสำหรับการเปรียบเทียบระหว่างสองอนุกรมที่ต่างกัน

ก. การหาค่าเฉลี่ยเลขคณิต (Arithmetic Mean)

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{i=N} X_i \quad \ldots\ldots\ldots\ldots\ldots \; 3.1$$

โดยที่ X_i = ตัวแปร (Variate)

N = จำนวนตัวแปรทั้งหมดในอนุกรม

i = แสดงลำดับที่ของตัวแปรมีค่าเป็นเลขจริงจำนวนเต็ม โดยมีค่าตั้งแต่ 1 ถึง N

ในทางอุทกวิทยา ค่า $\bar{X}$ โดยทั่วไปจะเป็นค่าเฉลี่ย (Mean) ของตัวอย่างอนุกรมที่ไม่แน่นอน จำนวน N ค่าของการสังเกตุ ซึ่งจะขยายขอบเขตจำนวนปีได้ไม่สิ้นสุด เพราะฉะนั้น $\bar{X}$ เป็นเพียงการประเมินหาค่าเฉลี่ยเลขคณิต (μ) ของอุดมคติอนุกรม (Ideal Series) เท่านั้น

ข. การหาค่ามัธยฐาน (Median = M)

มัธยฐาน M จะเป็นค่าตัวแปร X ที่อยู่ตรงกลางของอนุกรม ซึ่งถูกจัดเรียงจากน้อยไปมาก (Ascending) หรือจากมากไปน้อย (Descending) ตามลำดับของขนาด (Magnitude)

ถ้าจำนวนตัวอย่าง มีจำนวนคี่ (N = 2p + 1) :

$$\text{ค่า } M = \text{ตัวที่ } (p + 1)$$

ถ้าจำนวนตัวอย่างเป็นจำนวนคู่ (N = 2p) :

$$\text{ค่า } M = \frac{(\text{ตัวที่ } p) + (\text{ตัวที่ } p+1)}{2}$$

ค. การหาค่าฐานนิยม (Mode = M_o)

ฐานนิยม คือค่าเฉพาะของตัวแปรที่มีความถี่มากที่สุด ถ้าโค้งคณิตศาสตร์ (Mathematical Curve) สำหรับตัวแปรต่อเนื่อง ถูกทำให้เหมาะสมในแผนภาพ (Histogram) ฐานนิยม คือ ค่าของตัวแปร ซึ่งมีความหนาแน่นที่เป็นไปได้ของฟังชั่นมากที่สุด

โดยให้ p = ความหนาแน่นที่เป็นไปได้ของฟังชั่น

(probability density function)

แล้ว $\frac{dp}{dx} = 0$ & $\frac{d^2p}{dx^2} = 0$

3.2.2 การวัดการกระจาย (Measurement of the dispersion or fluctuation)

การวัดการกระจายของจำนวนต่าง ๆ ในทางคณิตศาสตร์รอบค่ากลาง คุณลักษณะนี้เป็นปัจจัยของวิศวกร สามารถถูกแสดงจำนวนโดย parameter หลักตามลำดับ ซึ่งจะพิจารณาตามลำดับความยุ่งยากที่เพิ่มขึ้น

ก. พิสัย (Rangs - W) เป็นค่าความแตกต่างระหว่างค่าไกลสุดของอนุกรม (extreme values of the series) หรือตัวอย่าง (Sample)

$$W = X\ max - X\ min \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ 3.2$$

โดยที่ X max = ค่ามากที่สุดของตัวอย่าง (Extreme Maximum)

X min = ค่าน้อยที่สุดของตัวอย่าง (Extreme Minimum)

W = ค่าพิสัยของอนุกรม (Range)

ข. ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation - σ)ของสถิติประชากร คือ กรณีที่สองของค่าเฉลี่ย ผลรวมกำลังสองของส่วนเบี่ยงเบน (Mean Square Deviation) ของแต่ละจำนวนในอนุกรมรอบค่าเฉลี่ย μ ของอนุกรมนี้ ในกรณีอนุกรมยาวมาก (Long Series)

$$\sigma = \sqrt{\frac{\sum_{i=1}^{i=N} (X_i - \mu)^2}{N}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ 3.3$$

สำหรับส่วนเบียงเบนมาตรฐาน ซึ่งมีจำนวน N ตัวอย่างของอนุกรม ในกรณีอนุกรมสั้น (Short Series) เราหาค่านี้โดยใช้สูตร

$$S = \sqrt{\frac{\sum_{i=1}^{i=N} (X_i - \bar{X})^2}{N-1}} \quad \ldots\ldots\ldots\ldots\ldots 3.4$$

ในกรณีอนุกรมยาวมาก เมื่อใช้ค่า N-1 ไปหาร ก็จะไม่ทำให้ค่าส่วนเบียงเบนมาตรฐานผิดไปมากเท่าใด ดังนั้นในการคำนวณสำหรับการศึกษานี้ จะยอมรับส่วนเบียงเบนมาตรฐานที่คำนวณจากสูตรในสมการ 3.4

ค. ความแปรปรวน (Variance-V) วัดเป็นกำลังสองของส่วนเบียงเบนมาตรฐาน

$$V = \delta^2 \text{ หรือ } S^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots 3.5$$

ง. สัมประสิทธิ์ความแปรปรวน (Coefficient of Variation-C_v) เป็นค่าที่ไม่มีหน่วย (dimensionless) แสดงค่าด้วยอัตราส่วนของส่วนเบียงเบนมาตรฐานกับค่าเฉลี่ย

$$C_v = \frac{\delta}{\mu} \text{ หรือ } \frac{S}{\bar{X}} \quad \ldots\ldots\ldots\ldots\ldots 3.6$$

3.3 วิธีการวิเคราะห์ความถี่ของข้อมูลอุทกวิทยา

(Frequency Analysis of Hydrological Data)

การวิเคราะห์ความถี่เป็นทฤษฎีทางสถิติ ซึ่งถูกนำมาใช้อย่างแพร่หลายในทางอุทกวิทยา เช่น การประเมินค่าน้ำหลากสูงสุดในรอบปีต่าง ๆ, การประเมินค่าน้ำฝนสูงสุด เช่น น้ำฝน ๑ วันสูงสุด, น้ำฝน ๒ วันสูงสุด, น้ำฝน ๓ วันสูงสุด เป็นต้น ในรอบปีต่าง ๆ เพื่อนำค่าที่ได้ไปประกอบการพิจารณา และเพื่อออกแบบในเชิงอุทกวิทยา ซึ่งการวิเคราะห์ความถี่นี้มีอยู่ด้วยการหลายวิธี เช่น Plotting-Position Formula, Pearson Type 3 Distribution, Hazen Formula, Gumbel Formula เป็นต้น

สำหรับการศึกษานี้ เราศึกษาการวิเคราะห์ความถี่ของข้อมูลอุทกวิทยา โดยสูตรของกัมเบล(Gumbel's Formula)Gumbel's Law ถูกใช้อย่างกว้างขวางในการศึกษา

น้ำหลากสูงสุด ซึ่งเป็นวิธีหาความถี่สำหรับค่าสูงสุดขีดทางสถิติ

แนวความคิดของคำว่า "RETURN PERIOD T" นี้ไม่ได้บ่งชี้ให้เห็นว่า ขนาดของน้ำหลากที่คำนวณได้จะเกิดขึ้นอีกครั้งในช่วงสม่ำเสมอ (Regular interval)หรือเกือบจะสม่ำเสมอในเวลา T ปี แต่หมายถึงว่าความเป็นไปได้ของช่วงเฉลี่ยระหว่าง ๒ เหตุการณ์ของขนาดที่พิจารณา

$$T = \frac{1}{p} \qquad \text{..................} \ 3.7$$

ซึ่ง p = ความเป็นไปได้สูงสุด (Exceedence Probability)

ถ้าให้ $x_1, x_2, x_3, \ldots\ldots, x_n$ เป็นค่าสูงสุดของค่าที่ได้จากการสังเกต n ตัวอย่าง มีขนาด N และ x มีค่าได้ไม่จำกัด เพราะว่า n และ N มีค่าเข้าใกล้อนันต์ (infinity) แล้ว ค่าเป็นไปได้สะสม (Cumulative Probability=P)ของ n ตัวอย่าง จะแสดงนิพจน์ได้เป็น

$$P = e^{-e^{-b}} \qquad \text{..................} \ 3.8$$

ซึ่ง e = ฐานของ Napierian ลอกการิทึม มีค่า 2.71828

b = ค่า Reduce Variate และถูกกำหนดให้เป็น

$$b = a\,(X - X_f) \qquad \text{..................} \ 3.9$$

a = Dispersion Parameter

X_f = Mode of the distribution

สำหรับตัวอย่างที่ใหญ่มาก (infinitely Large Sample) จากทฤษฎีนี้จะแสดงค่า X_f และ a เป็นฟังชันของค่าเฉลี่ยเลขคณิต ($\bar{X}$) และส่วนเบี่ยงเบนมาตรฐาน(S_x)

$$X_f = \bar{X} - 0.45005\, S_x \qquad \text{..................} \ 3.10$$

$$a = 1.28255/S_x \qquad \text{..................} \ 3.11$$

โดยที่ $\bar{X}$ และ S_x คำนวณได้จากสมการที่ 3.1, 3.3 หรือ 3.4

ในสมการที่ 3.8 เป็นนิพจน์ที่แสดงถึงโอกาสที่จะไม่เกิดขึ้น(Probability of

Non-occurence) ดังนั้นนิพจน์ที่แสดงรอบปีที่จะกลับมาเกิดอีกครั้ง (Return Period) สามารถคำนวณจากสูตร

$$T = \frac{1}{1-P} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.12$$

รวมสมการ 3.9, 3.10 และ 3.11 จะได้

$$b = \frac{1.28255}{S_x}(X_T - \bar{X} + 0.45005\, S_x) \quad \ldots\ldots\ldots\ldots 3.13$$

ผลของสมการ 3.8 และ 3.12 จะได้

$$b = -\log\left(-\log\left(1-\frac{1}{T}\right)\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots 3.14$$

จับเทียบสมการ 3.13 และ 3.14 จะได้

$$X_T = \bar{X} - 0.45005\, S_x - \frac{S_x}{1.28255}\log\left(-\log\left(1-\frac{1}{T}\right)\right) \quad \ldots\ldots\ldots 3.15$$

จากการวิเคราะห์ความถี่ของน้ำหลากสูงสุดในที่นี้จะมี $X_1, X_2, X_3, \ldots\ldots\ldots, X_n$ เป็นค่าน้ำหลากสูงสุดรายปี (maximum flow for each water year, N=1) และ X_T เป็นน้ำหลากสูงสุดในรอบปีที่ T (maximum flood peak for return period T years)

จากขั้นตอนของทฤษฎีดังกล่าว จึงได้เขียน program สำหรับการวิเคราะห์ความถี่ โดยใช้ COMPUTER ขึ้น ดังแสดง program ไว้ในภาคผนวก ค. หน้า 117

3.4 สหสัมพันธ์และเส้นถดถอยเชิงซ้อน

(Multiple Correlation and Regression)

3.4.1 รายละเอียดทั่วไปของสหสัมพันธ์และเส้นถดถอยเชิงซ้อน

ตามปกติแล้ว ปรากฏการณ์ที่เกิดขึ้นในทางอุทกวิทยาและวิทยาศาสตร์ของแหล่งน้ำ เป็นเรื่องที่ซับซ้อนมากจนขาดการเข้าไปศึกษาอย่างใกล้ชิดเพื่อทำให้มันง่ายขึ้น ตัวอย่างของปรากฏการณ์ที่เราสนใจ เช่น การประเมินน้ำท่า (Runoff) จากน้ำฝน (Rainfall), การประเมินจำนวนตะกอน (Sediment Yield) และ Evapotranspiration จากลุ่มน้ำ เป็นต้น ดังนั้น การวางแผนเพื่อพัฒนาลุ่มน้ำ จึงได้นำการวิเคราะห์สหสัมพันธ์และเส้นถดถอยมาใช้ในการรวมข้อมูลที่ได้มาจากการสังเกตของตัวแปรไว้ในโจทย์

ในงานการศึกษาการวิจัยทางอุทกวิทยา มีปรากฏอยู่บ่อยครั้งที่การวิเคราะห์ผลการทดลองโดยวิธีเส้นถดถอยอย่างง่าย (Simple Linear Regression) นั้นไม่เพียงพอ เนื่องจากมีค่าตัวแปรอิสระ (Independent Variable) อยู่หลายตัว ดังนั้น การวิเคราะห์ผลการทดลองเมื่อมีตัวแปรอิสระตั้งแต่สองตัวขึ้นไป สามารถกระทำโดยวิธีเส้นถดถอยเชิงซ้อน (Multiple Regression) ซึ่งประโยชน์ที่สำคัญของการวิเคราะห์นี้มีดังนี้

ก. ใช้ในการคำนวณเพื่อตั้งสมการทำนาย เมื่อตัวแปรอิสระมีมากกว่า 1 ตัวขึ้นไป

ข. ใช้ในการสำรวจเพื่อหาตัวแปรอิสระ (Independent Variable) พวกที่สำคัญที่สุด เพื่อใช้สำหรับคำนวณสมการทำนาย ในกรณีที่มีจำนวนตัวแปรอิสระอยู่เป็นจำนวนมาก

ค. ในการทดลองบางอย่าง อาจมุ่งศึกษาว่า มีตัวแปรอิสระอะไรบ้างที่เกี่ยวข้องกับตัวแปรตาม (Dependent Variable) และมีความสัมพันธ์กันอย่างไร แค่ไหน

ง. เพื่อ Interpolating ระหว่างค่าของฟังชั่น

จ. เพื่อกำหนดเงื่อนไขสูงสุดในการทำงาน

ฉ. เพื่อประเมินหาค่าของสัมประสิทธิ์เส้นถดถอย

ในการศึกษาครั้งนี้ ได้ใช้การวิเคราะห์เส้นถดถอยเชิงซ้อน โดยวิธี Stepwise เพราะเป็นวิธีที่ใช้กันมากในการทดสอบของค่าตัวแปรต่าง ๆ ในงานทางด้านอุทกวิทยา และงานวิศวกรรมสาขาอื่น ๆ ทั้งนี้ เนื่องจากตามศูนย์คอมพิวเตอร์ทั่ว ๆ ไป มีโปรแกรมมาตรฐานของคอมพิวเตอร์นั้น ๆ ซึ่งสะดวกแก่การใช้สำหรับผู้ทำการศึกษา (User) ได้ไม่ยุ่งยากนัก

3.4.2 สมการทั่วไปของสหสัมพันธ์และเส้นถดถอยเชิงซ้อน

เส้นถดถอยเชิงซ้อน ถูกนำมาใช้อย่างกว้างขวางในทางอุทกวิทยาเพื่อตรวจสอบความสัมพันธ์ระหว่างตัวแปรตาม (Dependent Variable) และกลุ่มตัวแปรอิสระ (Group of Independent Variable) สมการทั่วไปของเส้นถดถอยเชิงซ้อนซึ่งประกอบด้วย p ตัวแปร โดยมีตัวแปรตาม (Dependent Variable) 1 ตัว และตัวแปรอิสระ (Independent Variable) k = p - 1 ตัว คือ

$$X_1 = b_1+b_2X_2+b_3X_3+ \text{---} + b_pX_p \text{ ------ } 3.16$$

โดยที่ b_1 = จุดที่เส้นตรงตัดแกน X_1 (intercept)

b_i ซึ่ง i = 2, 3, ----, p = สัมประสิทธิ์เส้นถดถอยเชิงซ้อนของตัวแปรอิสระ

ในการประเมินสมการทำนาย (3.16) ได้ใช้วิธีการของกำลังสองน้อยที่สุด (Least Square Method) ซึ่งทำให้ผลรวมกำลังสองของส่วนเบี่ยงเบน (ผลต่างของ

ค่าสังเกตและค่า X_1 ซึ่งได้จากการประเมินจากสมการ 3.16) มีค่าน้อยที่สุด

สัมประสิทธิ์สหสัมพันธ์เชิงซ้อน (Multiple Correlation Coefficient $= R_1$) ตามปกติแล้วถูกใช้เป็นมาตราในการวัดสหสัมพันธ์เส้นตรง หาได้โดยสูตร

$$R_1 = \frac{S_e}{S} \qquad \text{----------} \quad 3.17$$

โดยที่ S_e = ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตาม ซึ่งประเมินได้จากสูตรสมการเส้นถดถอยเชิงซ้อน

S = ส่วนเบี่ยงเบนมาตรฐานของค่าสังเกตของตัวแปรตามขนาดของ R_1 มีค่าแปรเปลี่ยนระหว่าง 0 ถึง 1 ซึ่งเทียบได้จากไม่มีความสัมพันธ์กันเลย จนถึงมีความสัมพันธ์กันอย่างสมบูรณ์ ดังนี้ (หนังสืออ้างอิง 26)

$R_1 = 1$	มีความสัมพันธ์กันโดยตรงมากที่สุด
$0.6 \leq R_1 < 1$	มีความสัมพันธ์กันโดยตรงในขั้นดี
$0 < R_1 < 0.6$	มีความสัมพันธ์กันโดยตรงยังไม่พอเพียง
$R_1 = 0$	ไม่มีความสัมพันธ์กันเลย
$-0.6 < R_1 < 0$	มีความสัมพันธ์ผกผันยังไม่พอเพียง
$-1 < R_1 \leq -0.6$	มีความสัมพันธ์ผกผันในขั้นดี
$R_1 = -1$	มีความสัมพันธ์ผกผันมากที่สุด

3.4.3 การวิเคราะห์เส้นถดถอยเชิงซ้อนโดยกรรมวิธีของ Stepwise

ปัญหาส่วนใหญ่ในทางปฏิบัติเนื่องจากมีตัวแปรมาก ซึ่งดูเหมือนว่าเป็นสาเหตุของการแปรเปลี่ยนในตัวแปรตาม ถ้าตัวแปรอิสระทั้งหมดเหล่านี้ถูกนำมาใช้ในสมการ 3.16

ความเชื่อถือได้ของสมการเส้นถดถอยจะลดลงเนื่องจากการสูญเสียองศาแห่งเสรีภาพ (The Loss of degree of freedom) โดยเฉพาะอย่างยิ่งเมื่อขนาดตัวอย่างของข้อมูลมีจำนวนน้อย ดังนั้นกลวิธีของ Stepwise ถูกใช้อย่างมาก เพื่อจุดมุ่งหมายในการเลือก M ตัวแปรอิสระ จากจำนวนทั้งหมด p-1 หรือ k ตัวแปรอิสระที่เป็นไปได้ ($M \leqslant k$) ซึ่งมีนัยสำคัญ สัมพันธ์กับตัวแปรตาม

กรรมวิธีของ Stepwise เป็นกลวิธีที่ช่วยในการตรวจสอบเพื่อหลีกเลี่ยงปัญหาในการคำนวณ เริ่มต้นด้วยการเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามสูงที่สุด เข้ามาสัมพันธ์กัน แล้วจะเพิ่มตัวแปรอิสระอันดับที่สอง ซึ่งมีสัมประสิทธิ์สหสัมพันธ์รองลงมาจากตัวแรก นำเข้าพิจารณานัยสำคัญว่ามีความสัมพันธ์กับตัวแปรตามประการใด กรรมวิธีในตอนนี้จะเป็นการตรวจสอบความช่วยเหลือของตัวแปรอิสระตัวแรกที่ถูกนำเข้าไปรวมในสมการ ถ้าตัวแปรอิสระตัวแรกนั้นถูกนำเข้าไปรวมในสมการหลังจากตัวแปรตัวที่สอง ผลปรากฏว่า การที่ตัวแปรอิสระตัวแรกถูกนำเข้าไปช่วยเหลือนั้น ไม่มีนัยสำคัญกับตัวแปรตาม ตัวแปรซึ่งถูกรวมเข้าไปครั้งแรกนั้นจะถูกนำออกจากสมการ หรือมิฉะนั้นมันจะถูกเก็บในสมการ แล้วตัวแปรอิสระตัวต่อไปจะถูกเลือกเข้ามาโดยกรรมวิธีเดียวกัน กรรมวิธีเหล่านี้จะถูกดำเนินต่อไปจนกระทั่งไม่มีตัวแปรอิสระที่มีนัยสำคัญต่อตัวแปรตามอีก หรือจำนวนตัวแปรอิสระที่ต้องการถูกนำเข้าไปรวมในสมการจนครบ

คุณสมบัติที่สำคัญของกรรมวิธีของ Stepwise คือการคำนวณแต่ละขั้นตอนของตัวแปรอิสระแต่ละตัวที่ถูกนำเข้าไปในสมการเส้นถดถอยทุกครั้ง จะต้องถูกทดสอบนัยสำคัญก่อน ถ้าการช่วยเหลือของตัวแปรถูกพบว่าไม่มีนัยสำคัญ ตัวแปรนั้นจะถูกนำออกจากสมการก่อนที่จะนำตัวแปรตัวถัดไปเพิ่มเข้าไปในสมการ จะเห็นได้ว่ากรรมวิธีนี้ไม่เพียงแต่จะเลือกตัวแปรเข้าในสมการเท่านั้น ยังจะขจัดตัวแปรที่เห็นว่าไม่มีความสำคัญออกด้วย แต่กรรมวิธีของ Stepwise นี้ไม่ได้ประกันว่า ชุดของตัวแปรที่ดีที่สุดจะถูกรวมอยู่ในสมการขั้นสุดท้าย อย่างไรก็ตาม ก็ยังเป็นวิธีที่มีประสิทธิภาพในการค้นหาสมการเส้นถดถอยที่ดี และให้ผลประโยชน์ที่สมบูรณ์แบบมากวิธีหนึ่ง สำหรับรายละเอียดเกี่ยวกับกรรมวิธีนี้สามารถหาอ่านได้

จากหนังสืออ้างอิงที่ 15

กรรมวิธีของ Stepwise ได้นำ Experimental Data Generating Technique มาใช้เพื่อทำให้เกิดจำนวนสุ่มตัวอย่าง p อนุกรมขึ้นมาอย่างอิสระ โดยแต่ละอนุกรมประกอบด้วยขนาดของตัวอย่าง (Sample Size) = N และมีการกระจายอย่างปกติ (Normal Distribution) ด้วยค่าเฉลี่ยเป็นศูนย์ (Mean Zero) และความแปรปรวนเท่ากับ (Variance Unity) ในการวิเคราะห์เส้นถดถอยเชิงซ้อนโดยของ Stepwise นี้ได้ใช้จำนวน p อนุกรม ซึ่งจะเป็นตัวแปรอิสระ (Independent Variable) จำนวน k = p-1 อนุกรม และเป็นตัวแปรอิสระที่ถูกคัดเลือกแล้วจำนวน M อนุกรม ($M \leq k$) เป็นที่ปรากฏชัดว่า ความถูกต้องของการกระจายของตัวอย่างขึ้นอยู่กับจำนวนของตัวอย่าง นั่นคือ สำหรับค่า Percentile ของ R_s สูงนั้น ต้องการใช้จำนวนของตัวอย่างมากกว่าจำนวนของตัวอย่างสำหรับค่า Percentile ของ R_s ต่ำ ในที่นี้ R_s หมายถึงสัมประสิทธิ์สหสัมพันธ์เชิงซ้อนของการกระจายของตัวอย่าง

3.4.4 ขั้นตอนในการคำนวณหาค่าสหสัมพันธ์และเส้นถดถอยเชิงซ้อน

รูปแบบทั่วไปของเส้นถดถอยเชิงซ้อน (General Model of Multiple Regression)

$$Y = B_0+B_1X_1+B_2X_2+ \cdots +B_kX_k+E \quad \text{-----} \quad 3.18$$

โดยที่ B = Unknown Parameter

E = Random Error

X = Independent Variable

Y = Dependent Variable

ในการประเมินสมการทำนาย (Estimating the Prediction Equation)

$$\hat{Y} = b_o+b_1X_1+b_2X_2+ \cdots +b_kK_k \quad \text{-----} \quad 3.19$$

โดยที่ $\widehat{Y}$ = Predicted Value of Y for the given values of

$X_1, X_2, ----, X_k$

โดยการใช้จากหลักความจริงที่ว่า

$$b_o = \overline{Y} - b_1\overline{X}_1 - b_2\overline{X}_2 - --- -b_k\overline{X}_k$$

โดยให้ $x_i = X_i - \overline{X}_i$

$$\therefore \quad \widehat{Y} = \overline{Y} + b_1x_1 + b_2x_2 + --- + b_kx_k \text{ ---- } 3.20$$

เพื่อให้สูตรง่ายขึ้น Observations สามารถแสดงในรูปของ Deviation จากค่าเฉลี่ยของตัวอย่าง เช่น m th observation

$$x_{im} = X_{im} - \overline{X}_i$$

และ $y_m = Y_m - \overline{Y}$

แล้ว Least Squares Estimates ของ k+1 Parameter ของ Multivariable Linear Model (3.18) จะได้สมการ k+1 ชุดของสมการเส้นตรง

$$\left.\begin{array}{l} b_1\Sigma x_{1m}^2 + b_2\Sigma x_{1m}x_{2m} + --- b_k\Sigma x_{1m}x_{km} = \Sigma x_{1m}y_m \\ b_1\Sigma x_{1m}x_{2m} + b_2\Sigma x_{2m}^2 + --- + b_k\Sigma x_{2m}x_{km} = \Sigma x_{2m}y_m \\ ------------------------------------ \\ b_1\Sigma x_{1m}x_{km} + b_2\Sigma x_{2m}x_{km} + --- + b_k\Sigma x_{km}^2 = \Sigma x_{km}y_m \end{array}\right\} \text{--- } 3.21$$

สมการ 3.21 สามารถเขียนในรูปแบบของ Matrix ได้เป็น

$$\begin{bmatrix} \Sigma x_{1m}^2 & \Sigma x_{1m}x_{2m} & --- & \Sigma x_{1m}x_{km} \\ \Sigma x_{2m}x_{1m} & \Sigma x_{2m}^2 & --- & \Sigma x_{2m}x_{km} \\ --- & --- & --- & --- \\ \Sigma x_{km}x_{1m} & \Sigma x_{km}x_{2m} & --- & \Sigma x_{km}^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ -- \\ b_m \end{bmatrix} = \begin{bmatrix} \Sigma x_{1m}y_m \\ \Sigma x_{2m}y_m \\ --- \\ \Sigma x_{km}y_m \end{bmatrix} \quad ---- 3.22$$

ให้ S แทน Matrix k x k

และ a แทน Matrix k x 1 ทางด้านขวามือในสมการ 3.22

จะได้ $Sb = a$ ---- 3.23

ในการคำนวณ S_{ij} และ a_i ทำได้ดังนี้

$$\left.\begin{aligned} S_{ij} &= \Sigma x_{im}x_{jm} \\ &= \Sigma X_{im}X_{jm} - \frac{(\Sigma X_{im})(\Sigma X_{jm})}{N} \\ \text{และ } a_i &= \Sigma x_{im}y_m \\ &= \Sigma X_{im}Y_m - \frac{(\Sigma X_{im})(\Sigma Y_m)}{N} \end{aligned}\right\} \quad ----3.24$$

จากชุดของสมการเส้นตรงนี้ จะสังเกตเห็นว่า ค่า b เป็นตัวไม่รู้ค่า เราสามารถหาค่าได้โดย

$$b = C.a \quad ---- 3.25$$

ในกรณี 3.25 ค่า C คือ Inverse Matrix ของ S สำหรับ

สำหรับการตรวจสอบสมการทำนาย

หลังจากหาค่า C, b แล้วสร้างตาราง ANOVA (Analysis of Variance)

เนื่องจากว่าใน Model มี k variable, sum of squares due to regression มี k degree of freedom, residual sum of square มี N-(k+1) = N-k-1 degree of freedom

$$F\text{ statistic} = \frac{\text{Mean Square of Regression}}{\text{Mean Square of Residual}} \quad \text{-----} \quad 3.26$$

$$\text{Mean Square} = \frac{\text{Sum of Square}}{\text{Degree of Freedom}} \quad \text{-----} \quad 3.27$$

และ s = Standard error of estimate

$$s = \sqrt{\frac{\text{sum square of residual}}{N-k-1}} \quad \text{-----} \quad 3.28$$

<u>ตาราง ANOVA</u>

Source	Sum of Square (SS)	Degree of Freedom (df)	Mean Square (MS)
1. Due to regression (REG)	b'a	k	$MS(REG)=\frac{SS(REG)}{k}$
2. Due to residual (RES)	$\Sigma(Ym-\hat{Y}m)^2$	N-k-1	$MS(RES)=\frac{SS(RES)}{N-k-1}$
3. Total corrected forth mean	$\Sigma(Ym-\vec{Y})^2$	N-1	

หมายเหตุ: b' = Transpose of the b vector

$$b'a = b_1 \Sigma x_{1m}y_m + b_2 \Sigma x_{2m}y_m + \text{---} + b_k \Sigma x_{km}y_m \quad \text{---} \quad 3.29$$

ในกรณีนี้ F statistic ถูกใช้เพื่อตรวจสอบสมมติฐานทางสถิติ Proportion of Variation อธิบายด้วย R^2 ซึ่งเชื่อตามตาราง ANOVA สังเกตว่า R^2 ไม่ขึ้นอยู่กับจำนวนตัวแปรในสมการ แต่ s^2 ใช่ จากข้อเท็จจริงนี้ ถ้าตัวแปรตัวใหม่ถูกเพิ่มเข้าใน Model ค่า R^2 จะไม่ลดลง แต่ s^2 อาจเพิ่มหรือลด เนื่องจาก s^2 ขึ้นอยู่กับ Residual Degree of Freedom ซึ่งจะลดลงเมื่อตัวแปรตัวใหม่ถูกเพิ่มเข้าไปในสมการ

การคำนวณ R^2

$$\because \text{Total Corrected for the Mean} = \Sigma(Y_m-\bar{Y})^2$$

$$\therefore \Sigma(Y_m-\bar{Y})^2 = SS(REG) + SS(RES) \quad ---- \quad 3.30$$

โดยที่ R^2 = Square of sample multiple correlation coefficient

$$\therefore \quad R^2 = \frac{SS(REG)}{\Sigma(Y_m-\bar{Y})^2} \quad ------- \quad 3.31$$

F statistic สามารถคำนวณหาได้จากค่า Mean Square ซึ่งMean Square (MS) หาได้โดยการหาร Sum of Square (SS) ด้วย Degree of Freedom (df) จากสูตร 3.27

F_{CALC} = F statistical ที่ได้จากการคำนวณดังนี้

$$F_{CALC} = \frac{MS(REG)}{MS(RES)} \quad ------ \quad 3.32$$

โดยการเปรียบเทียบ F_{CALC} กับ F_{TAB} (จากตารางสถิติ F) ที่ระดับนัยสำคัญที่เหมาะสม

ถ้า $F_{CALC} > F_{TAB}$ สรุปได้ว่า เส้นถดถอยมีนัยสำคัญและให้ค่าทำนาย Y จากสมการทำนายดีกว่า $\bar{Y}$

3.5 การคำนวณหาค่าฝนเฉลี่ยทั้งปี

3.5.1 การหาค่าฝนเฉลี่ยของฝนโดยวิธีเส้นชั้นระดับ (Isohyetal Method)

ในการทำการศึกษานี้ การคำนวณหาค่าฝนเฉลี่ยทั้งปีว่าจะมีปริมาณฝนที่คาดว่าจะตกภายในเนื้อที่หนึ่ง ๆ นั้น ว่าเป็นจำนวนเท่าใด ได้นำวิธีเส้นชั้นระดับ (Isohyetal Method) มาใช้ เนื่องจากว่าวิธีนี้นิยมกันว่าเป็นวิธีหาฝนเฉลี่ยได้สมเหตุสมผลดี และได้ผลถูกต้องที่สุด ความผิดพลาดของวิธีนี้อาจเกิดได้ในกรณีที่ผู้เขียนเส้นชั้นระดับไม่อาจเขียนได้ถูกต้องตรงกับความเป็นจริงที่ฝนตกประการหนึ่ง กับการที่มีสถานีวัดน้ำฝนภายในเนื้อที่ไม่พอที่จะตรวจสอบการแผ่กระจายของฝนที่ตกจริงได้ทั่วเนื้อที่ เป็นสาเหตุให้เขียนเส้นชั้นระดับคลาดเคลื่อนได้อีกประการหนึ่ง

วิธีทำก็คือ กำหนดจุดของสถานีวัดน้ำฝนลงในแผนที่ พร้อมกับเขียนค่าน้ำฝนลงไปด้วย แล้วลากเส้นที่มีปริมาณน้ำฝนตกเท่า ๆ กัน (Isohyetal Line) โดยอาศัยค่าปริมาณน้ำฝนของแต่ละสถานีที่เขียนไว้นั้นเป็นหลัก รูปแบบที่ได้เรียกว่า แผนที่แสดงชั้นระดับ (Isohyetal Map)

$$\overline{X} = \frac{\Sigma AX}{\Sigma A} \qquad \text{------} \qquad 3.33$$

โดยที่ $\overline{X}$ = ค่าฝนเฉลี่ยโดยน้ำหนัก

X = ค่าฝนเฉลี่ยระหว่างเส้นชั้นระดับที่ติดกัน

A = เนื้อที่ระหว่างเส้นชั้นระดับที่พิจารณา

3.5.2 หลักในการเขียนเส้นชั้นระดับ

การเขียนเส้นชั้นระดับ คือการเขียนเส้นแสดงบริเวณที่ฝนตกเท่า ๆ กัน มีกฎเกณฑ์ที่ควรทราบดังต่อไปนี้

ก. ถ้าเป็นการหาค่าฝนเฉลี่ยโดยใช้จำนวนฝนรายเดือนหรือรายปีแล้ว ควรเลือก

แต่เฉพาะสถานีที่ทำการตรวจวัดปริมาณน้ำฝนไว้ติดต่อกันตลอดไม่ขาดตอน นอกจากจะมีน้อยแห่ง จึงอาจใช้สถิติที่ไม่สมบูรณ์นั้นได้ แต่จะต้องพิจารณาให้ถี่ถ้วนเสียก่อนว่า สถิติช่วงที่ขาดสำคัญหรือไม่เพียงใด เช่น ช่วงที่ขาดมีฝนตกมากน้อยเพียงใด

ข. สถิติของแต่ละสถานีให้ถือว่าถูกต้อง คือ ถือว่าเป็นความลึกจริงของฝนที่ตกที่สถานีนั้น แม้ฝนของสองสถานีนั้นใกล้เคียงกัน แต่วัดฝนได้แตกต่างกันมาก ก็จะต้องใช้ทั้งสองสถานีนั้น โดยไม่พิจารณาตัดทอนหรือยกเลิกสถานีหนึ่งสถานีใดเสีย

ค. ความถี่ห่างของเส้นชั้นระดับ ให้กะประมาณเอาด้วยตาเปล่า ไม่ควรแบ่งด้วยอัตราส่วนหรือเทียบอัตราอย่างใดทั้งสิ้น

ง. การลากเส้นชั้นระดับ ควรลากให้โค้งเป็น Smooth Curve คือ ไม่เป็นเส้นหัก และลากด้วยมือเปล่า

จ. ถ้ามีค่าของฝนบางสถานีทำให้เขียนได้หลายทาง ก็ควรเลือกเขียนอย่างที่ง่ายกว่าที่จะเป็นเส้นโค้งได้เหมาะสมและมองดูไม่ขัดตา

ฉ. การเลือกระดับชั้นไอโซไฮต์ เลือกให้เหมาะสมกับปริมาณฝนมากน้อย เพื่อมิให้มีเส้นหนาแน่นเกินไปหรือโปร่งเกินไป และควรให้เลขลงตัว

ช. เส้นชั้นระดับบริเวณเทือกเขา อาจคล้ายคลึงกับเส้นแสดงระดับดิน (Ground Contour) สำหรับฝนตลอดฤดูกาล เช่น ปริมาณฝนรายปี (Annual Rainfall)

ซ. การที่จะเขียนเส้นชั้นระดับให้ถูกต้องใกล้เคียงกับความเป็นจริงเพียงใดนั้นขึ้นกับความชำนิชำนาญของผู้เขียนแต่ละคน และจำนวนสถานีวัดน้ำฝนภายในเนื้อที่นั้นอีกด้วย

การหาค่าปริมาณฝนเฉลี่ยโดยวิธีเส้นชั้นระดับนี้ใช้ได้กับเนื้อที่ทุกขนาด และจำเป็นอย่างยิ่งที่จะต้องมีสถานีวัดน้ำฝนภายในเนื้อที่นั้น ๆ ให้พอเพียงที่จะเขียนเส้นชั้นระดับได้

จากการหาค่าเฉลี่ยของปริมาณน้ำฝนรายปีภายในลุ่มน้ำยม โดยวิธีเส้นชั้นระดับนี้ ได้แสดงแผนที่เส้นชั้นระดับ (Isohyetal Map) ดังในรูปที่ 3 หน้า 124 และแสดงค่าของปริมาณฝนเฉลี่ยรายปีของลุ่มน้ำย่อย ๆ ของลุ่มน้ำยม ดังในตารางที่ ก-1 หน้า 75 ซึ่งคำนวณได้โดยใช้สูตรในสมการ 3.33 โดยที่เนื้อที่ระหว่างเส้นชั้นระดับวัดโดย Polar Planimeter จำนวน 3 ครั้ง แล้วนำค่าที่ได้มาหาเฉลี่ยโดยสูตรในสมการที่ 3.1

3.6 คุณลักษณะของลุ่มน้ำ (Basin Characteristics)

กรรมวิธีทางกายภาค (Physical Process) ที่เกิดขึ้นเนื่องจากฝนที่ตกในลุ่มน้ำ แล้วกลายเป็นน้ำท่าไหลออกมา เป็นเรื่องที่ซับซ้อนมาก และเป็นเรื่องยุ่งยากในการที่จะให้คำอธิบายอย่างละเอียด ในส่วนของวัฏจักรทางอุทกวิทยา (Hydrogic Cycle) ดังนั้น การวิเคราะห์ทางอุทกวิทยาส่วนใหญ่ได้ทำโดยการใช้ 2-3 คุณลักษณะของลุ่มน้ำที่มีอิทธิพลต่อกรรมวิธี เป็นที่ปรากฏชัดว่า คุณลักษณะของลุ่มน้ำมีความสำคัญสำหรับการวิเคราะห์ใด ๆ จะขึ้นอยู่กับจุดประสงค์และวัตถุประสงค์ของการวิเคราะห์ ตามรายงานการวิเคราะห์นี้ จะเป็นการพิจารณาโดยตรงต่อคุณลักษณะซึ่งอาจจะมีผลกระทบต่อน้ำหลากที่จะเกิดในลุ่มน้ำยม ได้จัดแบ่งออกเป็นกลุ่ม ๆ ซึ่งอาจจะพบได้ในตำราต่าง ๆ ดังนี้คือ

ก. ลักษณะภูมิประเทศ (Topography)

ข. ลักษณะทางพืชพรรณ (Vegetation)

ค. ลักษณะทางปฐพีวิทยา (Soil)

ง. ลักษณะทางธรณี (Geology)

จ. ลักษณะดินฟ้าอากาศ (Climate)

ในการศึกษานี้ จะนำคุณลักษณะของลุ่มน้ำบางอย่างที่กล่าวนี้มาทำการศึกษา ดังจะกล่าวในหัวข้อต่อไป

3.6.1 พื้นที่ลุ่มน้ำ (Drainage Area - DA)

พื้นที่ลุ่มน้ำของจุดที่กำหนดให้บนลำน้ำ คือ เงาฉายของพื้นที่จริงที่เกิดขึ้นบนพื้นราบ ซึ่งจะรวมถึงการไหลของน้ำท่าที่ไหลตามพื้นผิวดิน มารวมกันที่จุดที่กำหนดนั้น

$$A_T = \sum_{i=1}^{i=N} A_i \qquad ------ \qquad 3.34$$

โดยที่ A_T = พื้นที่ลุ่มน้ำทั้งหมดซึ่งมีพื้นที่ลุ่มน้ำย่อย ๆ ทั้งหมด i อันดับ

A_i = พื้นที่ลุ่มน้ำย่อย ๆ อันดับที่ i

ซึ่ง i เป็นค่าจริงจำนวนเต็ม ตั้งแต่ 1 ขึ้นไป

เราสามารถวัดค่าพื้นที่ลุ่มน้ำต่าง ๆ ได้จากแผนที่ 1:50000 หรือ 1:250000 โดยใช้ Polar Planimeter ซึ่งได้แสดงอาณาเขตของพื้นที่ด้วยเส้นขอบเขต (B oundary Line) ด้วยการลากเส้นเชื่อมต่อระหว่างสันเขาถึงจุดที่กำหนดให้ เป็นเส้นแสดงขอบเขตของพื้นที่ที่จะทำการวัด

3.6.2 ตัวเลขแสดงถึงรูปร่างลุ่มน้ำ (Shape Number - SN)

รูปร่างของลุ่มน้ำ สามารถแสดงเป็นนิพจน์ในเทอมอย่างง่าย ๆ คือ

$$SN = \frac{A}{L^2} \qquad ---------- \qquad 3.35$$

โดยที่ SN = ตัวเลขแสดงถึงรูปร่างลุ่มน้ำ

A = พื้นที่ลุ่มน้ำที่พิจารณา

L = ความยาวของลำน้ำหลัก

นั่นคือ การแปรเปลี่ยนของ Shape Number จะเปลี่ยนแปลงตามขนาดของพื้นที่ลุ่มน้ำและความยาวของลำน้ำหลัก เพื่อพิจารณารูปร่างของลุ่มน้ำจาก Shape Number เมื่อมีลักษณะอื่น ๆ เหมือนกัน เช่น เมื่อลุ่มน้ำยิ่งมี SN มากขึ้น ลุ่มน้ำก็จะมีความกว้างมากขึ้นด้วย และปริมาณน้ำหลากย่อมมีค่าสูงขึ้นด้วย ซึ่งสามารถแบ่งเป็นลักษณะที่สำคัญของ

ลุ่มน้ำต่าง ๆ ได้ดังนี้

ก. ลุ่มน้ำรูปขนนก (Featherlike Basin) คือพื้นที่ขนาดเล็ก ซึ่งมีลำน้ำสาขาไหลลงสู่ลำน้ำสายใหญ่ทั้ง 2 ฝั่ง ปริมาณน้ำหลากจากลุ่มน้ำชนิดนี้จะมีค่าค่อนข้างต่ำ เพราะว่าระยะเวลาที่ปริมาณน้ำท่วมของแควสาขาต่าง ๆ ไหลมาถึงไม่เท่ากัน แต่น้ำจะท่วมอยู่เป็นเวลานาน

ข. ลุ่มน้ำรูปวงกลม (Radial Basin) คือพื้นที่ลุ่มน้ำที่มีสันปันน้ำเป็นรูปพัดหรือวงกลม และมีลำน้ำสาขาไหลลงสู่ลำน้ำสายใหญ่ที่จุดใดจุดหนึ่ง ดุจเป็นรัศมีของวงกลม ในพื้นที่ลุ่มน้ำที่มีลักษณะเช่นนี้ ปริมาณน้ำหลากจะมารวมกันที่จุด ๆ เดียว ทำให้เกิดน้ำท่วมใหญ่ใกล้กับจุดบรรจบหรือสบแม่น้ำ

ค. ลุ่มน้ำรูปขนาน (Parallel Basin) นี้มีพื้นที่ลุ่มน้ำ 2 ส่วน ซึ่งจะมาบรรจบกันในตอนล่าง หรือด้านท้ายน้ำของลุ่มน้ำ และน่าจะเกิดน้ำท่วมบริเวณตอนใต้ของสบแม่น้ำ

ง. ลุ่มน้ำรูปผสม (Complex Basin) ตามปกติแล้ว ลุ่มน้ำมีเพียงไม่กี่แห่งที่มีลักษณะอย่างใดอย่างหนึ่งดังกล่าว แต่ส่วนใหญ่แล้วจะมีลักษณะหลาย ๆ ชนิดรวมอยู่ด้วยกัน ซึ่งเรียกลุ่มน้ำลักษณะนี้ว่า ลุ่มน้ำรูปผสม

3.6.3 ความลาดชันของลำน้ำหลัก (Slope of Main Stream - S)

ความลาดชันของลำน้ำหลัก เป็นตัวแปรที่มีนัยสำคัญที่สุดของคุณลักษณะลุ่มน้ำรองลงมาจากพื้นที่ลุ่มน้ำ (DA) หาได้โดยนำค่าความสูงเฉลี่ยของลำน้ำตามยาว (Stream Profile) หารด้วยความยาวของลำน้ำหลัก ซึ่งความสูงเฉลี่ยของลำน้ำตามยาวหาได้จากผลรวม (Integration) ของ Elevation-distance curve ซึ่งหาได้โดยใช้ Polar Planimeter วัดพื้นที่ใต้เส้นโค้งนี้ ความลาดชันถูกลากโดยเส้นซึ่งลากผ่านจุดเริ่มต้นของ Stream profile curve นี้ โดยให้พื้นที่ใต้เส้นนี้เท่ากับพื้นที่ใต้

Stream profile curve ดังดูตัวอย่างได้จากภาพที่แสดง รูปที่ 4 หน้า 125

รูปตัดตามยาวของแม่น้ำ (Stream profile) คือ เส้นแสดงความสัมพันธ์ระหว่างระยะทางกับระดับท้องแม่น้ำ นับตั้งแต่ปากแม่น้ำขึ้นไปหาต้นน้ำ รูปตัดตามยาวนี้ค่อย ๆ เปลี่ยนไปจนเกือบจะคงที่หลังจากมีการกัดพาและการตกตะกอนในแม่น้ำ อันเนื่องมาจากอิทธิพลของปริมาณน้ำที่ไหลมา รูปตัดตามยาวที่คงที่แล้วนี้เรียกว่า รูปตัดตามยาวอยู่ในสมดุลย์ (Balanced profile) อย่างไรก็ตาม รูปตัดตามยาวที่อยู่ในสมดุลย์นี้ จะไม่มีเพียงรูปเดียวที่เหมือนกัน ทั้งนี้ เพราะว่าทั้งการกัดพาและการตกตะกอน เปลี่ยนแปลงไปตามจำนวนปริมาณน้ำที่ไหลในแม่น้ำนั้น

ความลาดชันของลำน้ำ เป็นองค์ประกอบหนึ่งซึ่งมีผลกระทบต่อรูปร่างของ Hydrograph เป็นอย่างมาก เช่น ถ้าความลาดชันของลำน้ำมีค่ามาก จะทำให้ความเร็วของน้ำที่ไหลในลำน้ำมีความเร็วสูง ซึ่งเป็นผลให้เกิด Peak ของ Hydrograph เกิดขึ้นเร็วและพร้อมกันนั้น อัตราการลดลงของปริมาณน้ำจะเร็วด้วย ซึ่งถ้าเปรียบเทียบลักษณะของ 2 ลุ่มน้ำเหมือนกันทุกอย่าง ยกเว้นความลาดชันของลำน้ำ จะทำให้สรุปผลได้ว่า ความลาดชันที่มีค่ามากจะให้ Peak เกิดเร็วกว่า, มีค่ามากกว่า และลักษณะรูปร่างของ Hydrograph จะผอมกว่าของความลาดชันที่มีค่าน้อย

3.6.4 ความยาวของลำน้ำ (Length of Stream)

ในการศึกษานี้ จะพิจารณาความยาวของลำน้ำ 2 ชนิด คือ

ก. L = ความยาวของลำน้ำหลัก โดยเริ่มพิจารณาจากจุดต้นกำเนิดบนยอดเขาจนถึงจุดทางออกของลำน้ำ

ข. L_c = ความยาวของลำน้ำ จากจุดบนลำน้ำที่ใกล้จุด Centroid มากที่สุด จนถึงจุดทางออกของลำน้ำ

3.6.4.1 ความยาวของลำน้ำหลัก (Length of Main Stream - L)

ความยาวของลำน้ำหลัก จะเป็นความยาวที่ยาวที่สุด ซึ่งวัดตามลำน้ำหลัก โดยเริ่มจากจุดที่พิจารณาจนถึงจุดบนเส้นแบ่งขอบเขตลุ่มน้ำ ความยาว L นี้ตามปกติวัดจากแผนที่แสดงภูมิประเทศ (Topographical Map) ซึ่งได้แสดงเงาฉายของลำน้ำหลัก ถึงเส้นแบ่งอาณาเขตลุ่มน้ำ ในการวัดความยาวตามลำน้ำหลักนี้ อาจมีการคลาดเคลื่อนบ้าง เนื่องจากเครื่องมือแต่ละเครื่อง, ความคดเคี้ยวมากน้อยของลำน้ำและการเปลี่ยนแปลงระดับของลำน้ำ ทำให้การลากเครื่องวัดคลาดเคลื่อนได้ ถึงอย่างไรก็ตาม ความคลาดเคลื่อนนี้ โดยทั่วไปแล้ว มันมีผลกระทบต่อการคำนวณทางอุทกวิทยาไม่มากนัก เราทำการแก้ไขข้อบกพร่องนี้โดยการหาค่าเฉลี่ยจากการวัดหลาย ๆ ครั้ง เราทำการวัด 3 ครั้ง สำหรับการศึกษานี้ เครื่องมือที่ใช้ในการวัดความยาวของลำน้ำนี้เรียกว่า Curvimeter

3.6.4.2 ความยาวของลำน้ำหลักจากจุดใกล้ศูนย์ถ่วงถึงจุดออกของลำน้ำ - L_c

L_c ถูกกำหนดให้เป็นความยาวที่ถูกวัดจากลำน้ำหลัก ที่จุดออกของลำน้ำที่กำหนดให้ ถึงจุดศูนย์ถ่วงของพื้นที่ลุ่มน้ำ หรือ L_c ก็คือ ระยะทางจากจุดออกที่กำหนดให้ของลำน้ำหลัก ถึงจุดซึ่งอยู่ใกล้กับ Centroid ของลุ่มน้ำมากที่สุด

จากรูปที่ 5 หน้า 125 แสดงถึงพื้นที่ลุ่มน้ำ ซึ่งมีแกน X-X ลากตั้งฉากกับ Major Axis ของพื้นที่ และผ่านฐานของพื้นที่ที่จุดออกของลำน้ำนั้น Area-centroid จริง ๆ แล้ว ถูกกำหนดด้วยระยะทาง $\overline{Y}$ เหนือแกน X-X นี้ ซึ่ง

$$Y = \frac{\Sigma Y_1 \cdot \Delta A}{A} \qquad \text{--------} \qquad 3.36$$

โดยที่ A = พื้นที่ลุ่มน้ำที่พิจารณา

Y_1 = ระยะจากจุด ΔA ถึงจุดออกบนแกน X-X

$\triangle A$ = พื้นที่เล็ก ๆ ใด ๆ ในพื้นที่ลุ่มน้ำที่พิจารณา

เทอม L_c สามารถแสดงเป็นสูตรที่คล้ายคลึงกัน คือ

$$L_c = \frac{\Sigma Y_2 \cdot \triangle A}{A} \qquad \text{-------} \qquad 3.37$$

โดยที่ Y_2 = ระยะทางซึ่งถูกวัดตามลำน้ำของจุดบนลำน้ำ ถึงจุดกำหนดให้ ซึ่งมาจากจุดบนลำน้ำซึ่งน้ำท่าที่ไหลบนผิวดินที่เกิดมาจากฝนตก บนเนื้อที่ $\triangle A$ ครั้งแรกที่มาถึง

3.6.5 องค์ประกอบรูปร่างลุ่มน้ำ (Basin Shape Factor - LCS)

ในการศึกษาคุณลักษณะของลุ่มน้ำนี้ ได้มุ่งที่จะหาความสัมพันธ์ระหว่างปริมาณน้ำหลากในรอบปีต่าง ๆ กับพื้นที่ลุ่มน้ำ, Shape Number, ปริมาณฝนเฉลี่ยรายปี, ความยาวลำน้ำหลักของ L และ L_c และความลาดชันของลำน้ำหลัก จากสถานีสำรวจปริมาณน้ำในลุ่มน้ำยม จำนวน 8 สถานี เนื่องจากตัวแปรอิสระมีจำนวนมากเกินไป เมื่อเทียบกับข้อมูล ทำให้ Loss degree of freedom ดังนั้น จึงรวมกลุ่มระหว่าง L, L_c และ S ในรูปแบบของ $\frac{L \cdot L_c}{\sqrt{S}}$ ซึ่งเรียกว่า Basin shape factor ในการรวมกลุ่มรูปแบบ Basin shape factor นี้ ได้ความคิดมาจากนิพจน์การหาค่า Lag time ของ Linsley, Kohler และ Paulhus ซึ่งได้แสดงค่าปริมาณน้ำหลากสูงสุดของหน่วยไฮโดรกราฟต่อหน่วยพื้นที่ลุ่มน้ำ แปรผกผันกับ Lag time และอีกกรณีหนึ่ง เพื่อหาความสัมพันธ์ระหว่างพื้นที่ลุ่มน้ำกับ Basin shape factor ว่าสัมพันธ์กันอย่างไร