

การย้ายข้อมูลเพื่อเพิ่มความเชื่อถือได้ในการประเมินความถี่ของน้ำท่วม ตอน 2
(Transfer of Information to Improve Estimates
of Flood Frequencies, Part II)

วราวุธ วุฒินิชัย Ph.D.*

1. บทนำ

ในกรณีที่มีข้อมูลน้ำท่วมจากหลายสถานีที่อยู่ในลำน้ำหรือลุ่มน้ำเดียวกัน และแต่ละสถานีมีข้อมูลยาวนาน อาจสามารถย้ายข้อมูลจากสถานีวัดน้ำต่าง ๆ เหล่านั้นมาที่สถานีที่กำลังพิจารณา ซึ่งมีข้อมูลสั้นกว่าได้ โดยใช้วิธีการหาค่าเส้นตรงแบบหลายตัวแปร (Multiple Linear Regression) ในทำนองเดียวกันกับการใช้สมการหาค่าเส้นตรงแบบง่าย (Simple Linear Regression) ย้ายข้อมูลจากสถานีหนึ่งไปยังอีกสถานีหนึ่งตามที่ได้กล่าวมาแล้วในตอนที่ 1 ซึ่งใช้สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร (Cross Correlation Coefficient) เป็นเกณฑ์ในการตัดสินใจพาราเมเตอร์ เช่น ค่าเฉลี่ย แนวโน้ม ฯลฯ ของข้อมูลที่ถูกขยายให้ยาวขึ้นมีความเชื่อถือได้มากขึ้นหรือไม่

2. วิธีการหาค่าเส้นตรงแบบหลายตัวแปร

สมมติให้ Y คือข้อมูลของสถานีที่มีข้อมูลสั้น ซึ่งมีจำนวนข้อมูลเท่ากับ N_1 และเวกเตอร์ X คือข้อมูลของสถานีที่มีข้อมูลยาว p สถานี ซึ่งแต่ละสถานีมีจำนวนข้อมูลเท่ากับ (N_1+N_2) อนุกรมข้อมูล Y และ X คือ (Gilroy, E.J., 1970)

* ผู้ช่วยศาสตราจารย์ ภาควิชาวิศวกรรมชลประทาน คณะวิศวกรรมศาสตร์ มหาวิทยาลัย-เกษตรศาสตร์ กำแพงแสน

$$\begin{aligned}
& y_1, y_2, \dots, y_{N_1}, \\
& x_1^1, x_2^1, \dots, x_{N_1}^1, x_{N_1+1}^1, \dots, x_{N_1+N_2}^1 \\
& x_1^2, x_2^2, \dots, x_{N_1}^2, x_{N_1+1}^2, \dots, x_{N_1+N_2}^2 \\
& \cdot \\
& \cdot \\
& x_1^p, x_2^p, \dots, x_{N_1}^p, x_{N_1+1}^p, \dots, x_{N_1+N_2}^p
\end{aligned}$$

เมื่อ Subscript คือ ลำดับที่ของข้อมูลในอนุกรม

และ Superscript คือ หมายเลขของสถานี

สมมุติว่าตัวแปร y, x^i มีการแจกแจงความน่าจะเป็นแบบปกติ หรือสามารถแปลง (Transform) ให้มีการแจกแจงเป็นแบบปกติได้ โดยมี $\mu_y, \mu_x^i, \sigma_y^2, \sigma_x^{2i}$ และ R เป็นพารามิเตอร์ เมื่อ $i = 1, \dots, p$

μ_y และ σ_y^2 คือ ค่าเฉลี่ยและแวกเรียนซ์ของปอบปูลเลชันของข้อมูล y

μ_x^i และ σ_x^{2i} คือ ค่าเฉลี่ยและแวกเรียนซ์ของปอบปูลเลชันของข้อมูล x^i เมื่อ $i = 1, \dots, p$

และ R คือ สัมประสิทธิ์สหสัมพันธ์แบบหลายตัวแปรของปอบปูลเลชัน (Multiple Correlation Coefficient)

สมการรีเกรชันเส้นตรงระหว่างตัวแปร y และ x_t^i คือ

$$y_t = a + \sum_{i=1}^p b^i x_t^i + \alpha \sqrt{1-R^2} \sigma_y \epsilon_t \quad \dots \dots \dots \textcircled{1}$$

เมื่อ a และ b^i ตัวพารามิเตอร์ของสมการ ($i = 1, \dots, p$)

ϵ_t คือ ตัวแปรแรนดอมปกติหรือน้อยส์ (Noise) ซึ่งมีค่าเฉลี่ยเท่ากับศูนย์ และแวกเรียนซ์เท่ากับ 1

θ มีค่า 0 หรือ 1 ขึ้นอยู่กับว่าต้องการลดน้อยลงในสมการที่ 1 หรือไม่
ถ้าไม่คิด ให้ $\theta = 0$ ถ้าคิดให้ $\theta = 1$

α คือ ค่าที่ช่วยจัดการลำเอียงในการ ประมาณแนวรีชันซ์ของ Y

3. การประมาณค่าพารามิเตอร์

จากการประมาณค่าพารามิเตอร์โดยวิธีลีสสแควร์ (Least Square) ค่าประมาณ
ของ $\hat{a}$ และ $\hat{b}^i$ ในสมการที่ 1 คือ

$$\hat{a} = \bar{Y}_1 - \sum_{i=1}^p \hat{b}^i \bar{X}_1^i \dots\dots\dots 2$$

และ

$$\hat{b}^i = \sum_{j=1}^p d_1^{ij} c_1^j \quad (i = 1, \dots\dots, p) \dots\dots 3$$

เมื่อ

d_1^{ij} คือ elements ของ Inverse ของโคแวลเรียนซ์แมทริกซ์
(Covariance Matrix) ซึ่ง elements ของโคแวลเรียนซ์แมทริกซ์
(g_1^{ij}) จะคำนวณได้จากสมการ

$$g_1^{ij} = \sum_{t=1}^{N_1} (X_t^i - \bar{X}_1^i)(X_t^j - \bar{X}_1^j) \quad (i, j = 1, \dots\dots, p) \dots\dots 4$$

และ

$$c_1^j = \sum_{t=1}^{N_1} (X_t^j - \bar{X}_1^j)(Y_t - \bar{Y}_1) \quad (j = 1, \dots\dots, p) \dots\dots 5$$

ค่าเฉลี่ย $\bar{X}_1^i$ และ $\bar{Y}_1$ คำนวณได้จากสมการ

$$\bar{X}_1^i = \frac{1}{N_1} \sum_{t=1}^{N_1} X_t^i \quad (i = 1, \dots, p) \quad \dots\dots\dots 6$$

$$\bar{Y}_1 = \frac{1}{N_1} \sum_{t=1}^{N_1} Y_t \quad \dots\dots\dots 7$$

ค่าประมาณของส่วนเบี่ยงเบนมาตรฐาน $S_1(Y)$ และ α ของสมการที่ 1 คำนวณได้จากสมการ

$$S_1(Y) = \sqrt{\frac{1}{N_1-1} \sum_{t=1}^{N_1} (Y_t - \bar{Y}_1)^2} \quad \dots\dots\dots 8$$

$$\alpha = \sqrt{\frac{N_2(N_1-2p-2)(N_1-1)}{(N_2-1)(N_1-p-2)(N_1-p-1)}} \quad \dots\dots\dots 9$$

ส่วนค่า $\hat{R}$ ประเมินจากสมการ (Yevjevich, V., 1972)

$$\hat{R} = \frac{\sum_{i=1}^p \hat{b}_i C_1^i}{\sqrt{\sum_{t=1}^{N_1} (Y_t - \bar{Y}_1)^2}} \quad \dots\dots\dots 10$$

และโดยทั่ว ๆ ไป แนะนำว่า อในสมการที่ 1 มีค่าเท่ากับศูนย์ หรือไม่ต้องคิดน้อยส์ ϵ_t

สมการที่ 1 สามารถนำไปใช้ขยายข้อมูล Y ให้มีจำนวนเท่ากับความยาวของข้อมูล X^i

ซึ่งจากข้อมูล $x_{N_1+1}^i, \dots, x_{N_1+N_2}^i$ จะสามารถคำนวณหา $\hat{y}_{N_1+1}, \dots, \hat{y}_{N_1+N_2}$ ได้ อนุกรม Y ใหม่ที่ถูกขยายให้ยาวขึ้นคือ

$$y_1, y_2, \dots, y_{N_1}, \hat{y}_{N_1+1}, \dots, \hat{y}_{N_1+N_2}$$

ซึ่งสามารถนำไปคำนวณหาค่าเฉลี่ยและแนวเรียนชี้ใหม่ได้ดังนี้ (Gilroy, E.J., 1970)

$$\bar{y} = \bar{y}_1 + \frac{N_2}{(N_1+N_2)} \sum_{i=1}^P \hat{b}^i (\bar{x}_2^i - x_1^i) \dots\dots\dots (11)$$

และ

$$\begin{aligned} S^2(y) &= \frac{1}{(N_1+N_2-1)} \{ (N_1-1) S_1^2(y) + \sum_{i=1}^P \sum_{j=1}^P \hat{b}^i \hat{b}^j g_2^{ij} \\ &\quad + \frac{N_1 N_2}{(N_1+N_2)} [\sum_{i=1}^P \hat{b}^i (\bar{x}_2^i - \bar{x}_1^i)]^2 \\ &\quad + (N_2-1) \alpha^2 \theta^2 (1-\hat{R}^2) S_1^2(y) \} \dots\dots\dots (12) \end{aligned}$$

เมื่อ

$$g_2^{ij} = \sum_{t=N_1+1}^{N_1+N_2} (x_t^i - \bar{x}_2^i) (x_t^j - \bar{x}_2^j) \quad (i, j = 1, \dots, p) \dots\dots\dots (13)$$

และ

$$\bar{x}_2^i = \frac{1}{N_2} \sum_{t=N_1+1}^{N_1+N_2} x_t^i \quad (i = 1, \dots, p) \dots\dots\dots (14)$$

4. เกณฑ์ในการตัดสินว่าการย้ายข้อมูลช่วยเพิ่มความเชื่อถือได้ ในการประเมินพารามิเตอร์หรือไม่

หลังจากที่ใช้สมการรีเกรซชันเส้นตรงแบบหลายตัวแปร ขยายข้อมูลอนุกรม Y จาก N_1 จำนวน เป็น $(N_1 + N_2)$ จำนวนแล้ว จะต้องมีการทดสอบว่าพารามิเตอร์ เช่น ค่าเฉลี่ย แนวเรชัน ฯลฯ ที่ประเมินจากข้อมูลที่ถูกขยายให้ยาวขึ้น มีความเชื่อถือได้มากขึ้นหรือไม่ ซึ่งโดยทั่วไป จะใช้แนวเรชันของพารามิเตอร์เป็นเกณฑ์ในการทดสอบ ถ้าแนวเรชันของพารามิเตอร์ที่ประเมินได้จากข้อมูลที่ขยายให้ยาวขึ้น $\text{Var}(\hat{Y})$ มีค่าน้อยกว่าแนวเรชันที่ประเมินจากข้อมูลเดิม $\text{Var}(\hat{Y}_1)$ แสดงว่า การย้ายข้อมูลจาก p สถานี โดยวิธีเกรซชันเส้นตรงแบบหลายตัวแปร จะช่วยเพิ่มความเชื่อถือได้ในการประเมินค่าพารามิเตอร์ แต่ถ้า $\text{Var}(\hat{Y})$ มากกว่า $\text{Var}(\hat{Y}_1)$ ข้อสรุปจะเป็นตรงกันข้าม คือ การย้ายข้อมูลจาก p สถานีจะลดความเชื่อถือได้ในการประเมินค่าพารามิเตอร์ให้น้อยลง หรืออาจใช้เกณฑ์ข้อมูลสัมพัทธ์ I ตามที่กล่าวไว้ในตอน 1 (วราวุธ 2530)

5. เกณฑ์การทดสอบว่าการย้ายข้อมูลเพิ่มความเชื่อถือได้ในการประเมินค่าเฉลี่ย

แนวเรชันของค่า $\bar{Y}$ ในสมการที่ 11 คือ (Gilroy, E.J., 1970)

$$\text{Var}(\bar{Y}) = \frac{\sigma_Y^2}{N_1} \left\{ 1 + \frac{N_2}{(N_1 + N_2)} \frac{[p - (N_1 - 2)R^2]}{(N_1 - p - 2)} \right\} \quad \dots \quad (15)$$

เนื่องจาก $\text{Var}(\bar{Y}_1) = \frac{\sigma_Y^2}{N_1}$
 $\bar{Y}$ จะเป็นตัวแทน μ_Y ได้ดีกว่า $\bar{Y}_1$ ก็ต่อเมื่อ

$$\text{Var}(\bar{Y}) < \text{Var}(\bar{Y}_1)$$

ซึ่งกรณีนี้จะเกิดก็ต่อเมื่อ

$$R > \sqrt{\frac{p}{N_1 - 2}}$$

ค่าทางด้านขวามือของสมการที่ 16 คือค่าสัมประสิทธิ์สัมพันธ์หลายตัวแปรต่ำสุดวิกฤต (Critical Minimum Multiple Correlation Coefficient) ซึ่งใช้เป็นเกณฑ์ในการทดสอบว่าการย้ายข้อมูลช่วยเพิ่มความเชื่อถือได้ในการประเมินค่าเฉลี่ยหรือไม่ ถ้าค่า R ที่คำนวณได้มากกว่าค่าสัมประสิทธิ์วิกฤต แสดงว่าการย้ายข้อมูลโดยวิธีรีเกรชันเส้นตรงแบบหลายตัวแปรช่วยเพิ่มความเชื่อถือได้ในการประเมินค่าเฉลี่ย μ_Y

6. เกณฑ์การทดสอบว่าการย้ายข้อมูลเพิ่มความเชื่อถือได้ในการประเมินค่าเฉลี่ย

แวเรียนซ์ของค่า $S^2(Y)$ ในสมการที่ 12 คือ (Gilroy, E.J., 1970)

$$\text{Var} [S^2(Y)] = \frac{2\sigma_Y^4}{N_1-1} \left[1 + \frac{N_2(N_1-1)}{2(N_1+N_2-1)^2} (AR^4 + BR^2 + C) \right] \dots \quad (17)$$

เมื่อ A , B และ C เป็นฟังก์ชันของ N_1, N_2, θ และ p และเนื่องจากเป็นฟังก์ชันที่ยุ่งยากไม่เหมาะแก่การจำจึงไม่กล่าวถึงในที่นี้ ผู้สนใจหาดูได้จากเอกสารอ้างอิง

$$\text{เนื่องจาก } \text{Var} [S^2(Y_1)] = \frac{2\sigma_Y^4}{(N_1-1)}$$

$S^2(Y)$ จะเป็นตัวแทน σ_Y^2 ได้ดีกว่า $S_1^2(Y)$ ก็ต่อเมื่อ

$$\text{Var}[S^2(Y)] < \text{Var}[S_1^2(Y)]$$

ซึ่งจากสมการที่ 17 กรณีนี้จะเกิดก็ต่อเมื่อ

$$R > \left[\frac{-B + \sqrt{B^2 - 4AC}}{2A} \right]^{\frac{1}{2}} \dots \dots \dots \quad (18)$$

ค่าทางด้านขวามือของสมการที่ 18 คือ ค่าสัมประสิทธิ์สหสัมพันธ์หลายตัวแปรต่ำสุดวิกฤต ซึ่งใช้เป็นเกณฑ์ในการตัดสินใจการย้ายข้อมูลเพิ่มความเชื่อถือได้ในการประเมินค่าแวนรียนซ์หรือไม่ ค่าสัมประสิทธิ์วิกฤตจะแปรผันไปตามค่า N_1 , N_2 , θ และ p ดังแสดงในตารางที่ 1 สำหรับ $\theta = 0$ และตารางที่ 2 สำหรับ $\theta = 1$

7. สิ่งที่จะต้องพิจารณาในการย้ายข้อมูล

การย้ายข้อมูลน้ำท่วมจากสถานีต่าง ๆ มายังจุดที่ต้องการ จะต้องพิจารณาถึงสิ่งต่างๆ ดังต่อไปนี้ (US. Water Resources Councils, 1976; Natural Environment Research Council, 1975)

1. สถานีวัดน้ำต้องอยู่บนลำน้ำหรือลุ่มน้ำเดียวกัน และควรอยู่ในรัศมี 50 ไมล์ ของจุดที่ต้องการ ถ้ามีหลายสถานีให้เลือก การเลือกสถานีจะขึ้นอยู่กับระยะทาง ลองดูจุด ภูมิประเทศ ลักษณะทางธรณีวิทยา ความสอดคล้องกันของข้อมูล ฯลฯ ถ้ามีการวิเคราะห์ทางสถิติ จะพิจารณาจากสัมประสิทธิ์สหสัมพันธ์ระหว่างสถานีที่จะเลือกและจุดที่ต้องการที่มีค่าสูงสุด
2. ควรพิจารณาใช้ข้อมูลทางอุทกวิทยาอื่น ๆ ที่เกี่ยวข้องกับน้ำท่วม เช่น ฝนประจำวัน และความชื้นในดินที่ขาดหายไป
3. ควรเลือกช่วงเวลาของข้อมูลที่เหมาะสม เช่น ในการขยายข้อมูลน้ำท่วมสูงสุดประจำปี ควรต้องใช้ข้อมูลน้ำท่วมสูงสุดประจำปีของสถานีต่างๆ แต่บ่อยครั้งจะพบว่าข้อมูลน้ำท่วมสูงสุดประจำปีของแต่ละสถานีมีสหสัมพันธ์กันน้อยมาก กรณีเช่นนี้อาจต้องเลือกใช้ข้อมูลน้ำท่วมประจำเดือนมาวิเคราะห์แทน ดังตัวอย่างในตอนต้นที่ 1
4. ควรมีการวิเคราะห์ความถูกต้อง (Consistency analysis) ของข้อมูลจากสถานีต่างๆ ที่เลือกไว้สำหรับการย้ายข้อมูล

8. ตัวอย่าง

ข้อมูลน้ำท่วมประจำเดือนของ Trinchera Creek ระหว่างปี 1961-1970 อยู่ในตารางที่ 4 ของตอนที่ 1 สถานีวัดน้ำใกล้เคียงคือ Ute Creek และ Cucharas River ซึ่งมีข้อมูลน้ำท่วมประจำเดือนระหว่างปี 1956-1975 ข้อมูลของ Ute Creek อยู่ในตารางที่ 4 ของตอนที่ 1 ส่วนข้อมูลของ Cucharas River อยู่ในตารางที่ 3 ต้องการทดสอบว่าการประเมินค่าเฉลี่ยและแวนรียนซ์ของน้ำท่วมประจำปีของ Trinchera Creek

ตารางที่ 1 ค่า R วิกฤตที่สุด ในการทดสอบความเชื่อถือได้ของนักเรียน สำหรับ $\theta = 0$

$p = 2$

N_1 N_2	9	10	12	14	16	18	20	25	30	35	40	45	50	55	60
6.	.77	.60	.45	.42	.44	.47	.49	.52	.54	.55	.56	.57	.58	.58	.58
8.	.79	.67	.61	.61	.61	.62	.63	.65	.66	.66	.67	.67	.68	.68	.68
10.	.80	.70	.67	.67	.67	.68	.69	.70	.71	.71	.72	.72	.73	.73	.73
12.	.80	.72	.70	.70	.70	.71	.72	.73	.74	.74	.75	.75	.76	.76	.76
14.	.81	.74	.72	.72	.73	.73	.74	.75	.76	.76	.77	.77	.78	.78	.78
16.	.81	.75	.73	.73	.74	.75	.75	.76	.77	.78	.78	.79	.79	.79	.79
18.	.82	.75	.74	.74	.75	.76	.76	.77	.78	.79	.80	.80	.81	.81	.81
20.	.82	.76	.75	.75	.76	.77	.77	.79	.80	.80	.81	.81	.81	.81	.82
25.	.82	.77	.76	.77	.78	.78	.79	.80	.81	.82	.82	.83	.83	.83	.83
30.	.83	.78	.77	.78	.79	.79	.80	.81	.82	.83	.83	.84	.84	.84	.84
35.	.83	.78	.78	.79	.79	.80	.81	.82	.83	.84	.85	.85	.85	.85	.85
40.	.83	.79	.78	.79	.80	.81	.81	.83	.84	.85	.85	.86	.86	.86	.87
45.	.83	.79	.79	.80	.81	.82	.82	.84	.85	.85	.86	.86	.86	.87	.87
50.	.83	.79	.79	.80	.81	.82	.82	.84	.85	.85	.86	.86	.87	.87	.87
55.	.83	.80	.79	.80	.81	.82	.83	.84	.85	.85	.86	.86	.87	.87	.87
60.	.84	.80	.80	.80	.81	.82	.83	.84	.85	.86	.86	.87	.87	.87	.88

 $p = 3$

6.	.87	.69	.51	.31	.26	.28	.31	.37	.41	.43	.45	.46	.47	.48	.48
8.	.88	.72	.51	.50	.49	.50	.51	.54	.56	.58	.59	.60	.60	.61	.61
10.	.89	.75	.58	.58	.57	.58	.60	.62	.64	.65	.66	.66	.67	.67	.67
12.	.89	.76	.63	.63	.63	.63	.64	.67	.68	.69	.70	.70	.71	.71	.71
14.	.89	.77	.66	.66	.66	.67	.68	.70	.71	.72	.72	.73	.74	.74	.74
16.	.90	.78	.68	.68	.68	.69	.70	.72	.73	.74	.75	.75	.76	.76	.76
18.	.90	.79	.68	.68	.68	.69	.70	.72	.73	.74	.75	.76	.77	.77	.77
20.	.90	.79	.68	.68	.68	.69	.70	.72	.73	.74	.75	.76	.77	.77	.77
25.	.90	.80	.74	.71	.72	.71	.72	.75	.76	.77	.77	.78	.78	.79	.79
30.	.90	.81	.74	.75	.74	.75	.75	.77	.78	.79	.80	.80	.81	.81	.81
35.	.91	.82	.76	.76	.76	.76	.77	.79	.80	.81	.82	.82	.83	.83	.83
40.	.91	.82	.77	.77	.78	.77	.78	.80	.81	.82	.83	.83	.84	.84	.85
45.	.91	.82	.78	.78	.78	.78	.79	.81	.82	.83	.84	.84	.85	.85	.85
50.	.91	.83	.78	.78	.79	.79	.80	.82	.83	.84	.84	.85	.85	.86	.86
55.	.91	.83	.79	.79	.79	.80	.80	.82	.83	.84	.84	.85	.85	.86	.86
60.	.91	.83	.79	.79	.80	.80	.81	.82	.83	.84	.85	.85	.86	.86	.87

ตารางที่ 1 (ต่อ)

P = 4

$\frac{N_1}{N_2}$	10	12	14	16	18	20	25	30	35	40	45	50	55	60
6.	.80	.59	.39	.25	.19	.19	.25	.30	.35	.36	.37	.39	.40	.41
8.	.82	.64	.48	.40	.39	.40	.44	.47	.49	.51	.52	.53	.54	.54
10.	.83	.67	.54	.48	.48	.49	.53	.56	.58	.59	.60	.61	.61	.62
12.	.84	.70	.58	.54	.54	.55	.59	.61	.63	.64	.65	.65	.66	.66
14.	.85	.71	.61	.58	.58	.60	.63	.65	.66	.67	.68	.69	.69	.70
16.	.85	.73	.64	.61	.62	.63	.66	.68	.69	.70	.71	.71	.72	.72
18.	.86	.74	.66	.64	.64	.65	.68	.70	.71	.72	.73	.73	.74	.74
20.	.86	.75	.68	.66	.66	.67	.70	.72	.73	.74	.74	.75	.75	.76
25.	.87	.77	.71	.69	.70	.71	.73	.75	.76	.77	.77	.78	.78	.79
30.	.87	.78	.73	.72	.72	.73	.75	.77	.78	.79	.79	.80	.80	.81
35.	.88	.79	.74	.73	.74	.75	.77	.79	.80	.81	.81	.82	.82	.82
40.	.88	.80	.76	.75	.75	.76	.78	.79	.80	.81	.82	.82	.83	.83
45.	.88	.81	.76	.76	.76	.77	.78	.80	.81	.82	.83	.83	.84	.84
50.	.88	.81	.77	.76	.77	.78	.80	.81	.82	.83	.83	.84	.84	.85
55.	.89	.82	.78	.77	.78	.79	.81	.82	.83	.84	.84	.85	.85	.85
60.	.89	.82	.78	.77	.78	.79	.81	.82	.83	.84	.84	.85	.85	.86

P = 5

6.	.89	.71	.51	.34	.23	.17	.17	.22	.26	.28	.31	.32	.33	.34
8.	.90	.74	.55	.41	.34	.32	.35	.39	.42	.44	.45	.47	.48	.48
10.	.91	.76	.59	.46	.41	.40	.44	.48	.50	.52	.54	.55	.56	.56
12.	.91	.78	.62	.51	.47	.46	.50	.54	.56	.58	.59	.60	.61	.62
14.	.92	.79	.64	.54	.51	.51	.55	.58	.60	.62	.63	.64	.65	.65
16.	.92	.80	.66	.57	.54	.55	.59	.62	.64	.65	.66	.67	.68	.68
18.	.92	.81	.68	.60	.57	.58	.61	.64	.66	.68	.69	.70	.71	.71
20.	.93	.82	.70	.62	.59	.60	.64	.66	.68	.70	.71	.72	.72	.72
25.	.93	.83	.73	.66	.64	.65	.68	.70	.72	.73	.74	.75	.75	.76
30.	.93	.84	.75	.69	.67	.68	.71	.73	.75	.76	.77	.77	.78	.78
35.	.93	.85	.76	.71	.69	.70	.73	.75	.77	.78	.79	.79	.80	.80
40.	.94	.85	.77	.72	.71	.72	.75	.77	.78	.79	.80	.81	.81	.81
45.	.94	.86	.78	.74	.73	.73	.76	.78	.79	.80	.81	.82	.82	.83
50.	.94	.86	.79	.75	.74	.74	.77	.79	.80	.81	.82	.82	.83	.83
55.	.94	.87	.80	.76	.75	.75	.78	.79	.81	.82	.82	.83	.83	.84
60.	.94	.87	.80	.76	.75	.76	.78	.79	.81	.82	.83	.83	.84	.84

ตารางที่ 2 ค่า Rวิกฤตค่าสุด ในการทดสอบความซื่อสัตย์ของนักเรียน สำหรับ $\theta = 1$

$p = 2$

N_1	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
6	.78	.66	.54	.44	.36	.30	.26	.18	.14	.11	.09	.08	.07	.06	.06
8	.79	.67	.56	.47	.39	.32	.27	.19	.14	.11	.10	.08	.07	.06	.06
10	.80	.69	.58	.49	.40	.34	.28	.20	.15	.12	.10	.08	.07	.06	.06
12	.80	.69	.59	.50	.42	.35	.30	.21	.15	.12	.10	.09	.08	.07	.06
14	.81	.70	.61	.52	.44	.37	.31	.21	.16	.13	.10	.09	.08	.07	.06
16	.81	.71	.61	.53	.45	.38	.32	.22	.16	.13	.11	.09	.08	.07	.06
18	.81	.71	.62	.54	.46	.39	.33	.23	.17	.13	.11	.09	.08	.07	.06
20	.82	.72	.63	.55	.47	.40	.34	.24	.17	.14	.11	.09	.08	.07	.06
25	.82	.72	.64	.56	.49	.43	.37	.26	.19	.15	.12	.10	.09	.08	.07
30	.82	.73	.65	.58	.51	.45	.39	.27	.20	.15	.12	.10	.09	.08	.07
35	.82	.73	.66	.59	.52	.46	.41	.29	.21	.16	.13	.11	.09	.08	.07
40	.83	.74	.66	.60	.53	.47	.42	.30	.22	.17	.14	.11	.10	.09	.08
45	.83	.74	.67	.61	.55	.49	.44	.33	.23	.18	.14	.12	.10	.09	.08
50	.83	.74	.67	.61	.55	.49	.44	.33	.24	.19	.15	.12	.10	.09	.08
55	.83	.74	.67	.61	.56	.50	.45	.34	.25	.20	.16	.13	.11	.09	.08
60	.83	.75	.68	.62	.56	.51	.46	.35	.26	.20	.16	.13	.11	.09	.08

$p = 5$

N_1	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
6	.87	.72	.57	.44	.35	.29	.24	.17	.13	.11	.09	.08	.07	.06	.05
8	.88	.76	.60	.47	.38	.31	.26	.18	.14	.11	.09	.08	.07	.06	.06
10	.89	.75	.62	.50	.40	.33	.27	.19	.14	.11	.09	.08	.07	.06	.06
12	.89	.76	.64	.52	.42	.34	.28	.19	.15	.12	.10	.08	.07	.06	.06
14	.89	.77	.65	.54	.44	.36	.30	.20	.15	.12	.10	.08	.07	.06	.06
16	.90	.78	.67	.55	.45	.37	.31	.21	.16	.13	.10	.09	.07	.06	.06
18	.90	.79	.68	.57	.47	.39	.32	.22	.16	.13	.10	.09	.08	.07	.06
20	.90	.79	.68	.58	.48	.40	.33	.22	.17	.13	.11	.09	.08	.07	.06
25	.90	.80	.70	.60	.51	.43	.36	.24	.18	.14	.11	.09	.08	.07	.06
30	.90	.81	.71	.62	.54	.45	.38	.26	.19	.14	.12	.10	.09	.08	.07
35	.91	.81	.72	.64	.55	.48	.40	.27	.20	.15	.12	.10	.09	.08	.07
40	.91	.82	.73	.65	.57	.49	.42	.29	.21	.16	.13	.11	.09	.08	.07
45	.91	.82	.74	.66	.58	.51	.44	.30	.22	.17	.14	.11	.10	.09	.08
50	.91	.82	.74	.67	.59	.52	.45	.32	.23	.18	.14	.12	.10	.09	.08
55	.91	.82	.75	.67	.60	.53	.47	.33	.24	.19	.15	.12	.10	.09	.08
60	.91	.82	.75	.68	.61	.54	.48	.34	.25	.19	.15	.12	.10	.09	.08

ตารางที่ 2 (ต่อ)

P = 4

N ₁	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
6.	.80	.63	.48	.38	.31	.26	.18	.14	.11	.09	.08	.07	.06	.06	.06
8.	.82	.66	.52	.41	.33	.27	.19	.14	.11	.10	.08	.07	.06	.06	.06
10.	.83	.68	.54	.43	.35	.29	.20	.15	.12	.10	.08	.07	.06	.06	.06
12.	.84	.70	.57	.45	.37	.30	.21	.15	.12	.10	.09	.07	.06	.06	.06
14.	.85	.72	.59	.47	.38	.32	.21	.16	.13	.10	.09	.08	.07	.06	.06
16.	.85	.73	.60	.49	.40	.33	.22	.16	.13	.11	.09	.08	.07	.06	.06
18.	.86	.74	.62	.51	.41	.34	.23	.17	.13	.11	.09	.08	.07	.06	.06
20.	.86	.75	.63	.52	.43	.35	.24	.17	.14	.11	.09	.08	.07	.06	.06
25.	.87	.77	.66	.55	.46	.38	.25	.18	.14	.12	.10	.09	.08	.07	.06
30.	.87	.78	.68	.58	.49	.41	.27	.20	.15	.12	.10	.09	.08	.07	.06
35.	.88	.79	.70	.60	.51	.43	.29	.21	.16	.13	.11	.09	.08	.07	.06
40.	.88	.80	.71	.62	.53	.45	.30	.22	.17	.13	.11	.09	.08	.07	.06
45.	.88	.80	.72	.63	.55	.47	.32	.23	.17	.14	.11	.09	.08	.07	.06
50.	.88	.81	.73	.64	.56	.49	.33	.24	.18	.14	.12	.10	.09	.08	.08
55.	.89	.81	.75	.66	.58	.50	.35	.25	.19	.15	.13	.11	.10	.09	.08
60.	.89	.81	.74	.66	.59	.52	.36	.26	.19	.15	.13	.11	.10	.09	.08

P = 5

N ₁	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
6.	.89	.71	.54	.42	.34	.28	.19	.15	.12	.10	.08	.07	.07	.07	.06
8.	.90	.74	.58	.45	.36	.30	.20	.15	.12	.10	.09	.08	.07	.07	.06
10.	.91	.76	.60	.48	.38	.31	.21	.16	.13	.10	.09	.08	.07	.07	.06
12.	.91	.78	.63	.50	.40	.33	.22	.16	.13	.11	.09	.08	.07	.06	.06
14.	.92	.79	.65	.52	.42	.35	.23	.17	.13	.11	.09	.08	.07	.06	.06
16.	.92	.80	.67	.54	.44	.36	.24	.18	.14	.11	.10	.08	.07	.06	.06
18.	.92	.81	.68	.56	.45	.37	.25	.18	.14	.11	.10	.08	.07	.06	.06
20.	.93	.82	.70	.57	.47	.39	.26	.19	.14	.12	.10	.09	.07	.07	.06
25.	.93	.83	.72	.61	.50	.42	.28	.20	.15	.12	.10	.09	.08	.07	.06
30.	.93	.84	.74	.63	.53	.44	.29	.21	.16	.13	.11	.09	.08	.07	.06
35.	.93	.85	.76	.66	.56	.47	.31	.22	.17	.14	.11	.10	.08	.07	.06
40.	.94	.85	.77	.67	.58	.49	.33	.23	.18	.14	.12	.10	.09	.08	.07
45.	.94	.86	.78	.69	.60	.51	.34	.24	.19	.15	.12	.10	.09	.08	.08
50.	.94	.86	.78	.70	.61	.53	.36	.26	.19	.15	.13	.11	.10	.09	.08
55.	.94	.87	.79	.71	.63	.54	.37	.27	.20	.16	.13	.11	.10	.09	.08
60.	.94	.87	.79	.72	.64	.56	.39	.28	.21	.16	.13	.11	.10	.09	.08

ที่มีความเชื่อถือได้มากขึ้นหรือไม่ โดยการย้ายข้อมูลจาก Ute Creek และ Cucharas River ซึ่งมีข้อมูลยาวนานกว่า โดยวิธีรีเกรสชันเส้นตรงแบบหลายตัวแปร

การคำนวณมีขั้นตอนดังนี้

ขั้นที่ 1 ให้ Y_t' คือปริมาณน้ำท่วมประจำเดือนของ Trinchera Creek ระหว่าง 1961-1970 ซึ่งมีจำนวนข้อมูล $N_1 = 12(10) = 120$ เดือน

ให้ $X_t^{1'}$ คือปริมาณน้ำท่วมประจำเดือนของ Ute Creek ระหว่าง 1956-1975 ซึ่งมีจำนวนข้อมูล $N_2 = 12(10) = 120$ เดือน

$X_t^{2'}$ คือปริมาณน้ำท่วมประจำเดือนของ Cucharas River ระหว่าง 1956-1975 ซึ่งมีจำนวนข้อมูล $N_2 = 12(10) = 120$ เดือน

ข้อมูลน้ำท่วมประจำเดือนของลำน้ำทั้งสามระหว่าง 1961-1970 จะถูกนำมาใช้หาพารามิเตอร์ของสมการรีเกรสชันเส้นตรงแบบหลายตัวแปร

ขั้นที่ 2 ใช้ฟังก์ชัน Log แปลงข้อมูลให้มีการแจกแจงเป็นแบบปกติโดยประมาณดังนี้ .-

$$Y_t = \log Y_t' \text{ (Trinchera Creek)}$$

$$X_t^1 = \log X_t^{1'} \text{ (Ute Creek)}$$

$$X_t^2 = \log X_t^{2'} \text{ (Cucharas River)}$$

ขั้นที่ 3 คำนวณหา $\hat{R}$ จากสมการที่ 10

$$\hat{R} = 0.954$$

ขั้นที่ 4 จากสมการที่ 16 เมื่อ $N_1 = 120$ และ $p = 2$

$$R \text{ วิฤต} = \sqrt{\frac{2}{120-2}} = 0.130$$

ตารางที่ 3 ข้อมูลพื้นที่ประมงเคื่อนของ Cucharas River ช่วงปี 1956-1975
ระหว่าง 1956-1975

	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUNE	JULY	AUG	SEPT
1956	11,000	8,500	8,500	7,000	9,000	16,000	72,000	70,000	89,000	42,000	18,000	12,000
1957	12,000	11,000	9,000	6,400	8,000	9,000	58,000	97,000	70,000	39,000	19,000	10,000
1958	26,000	28,000	15,000	9,000	8,500	13,000	39,000	186,000	128,000	29,000	16,000	11,000
1959	7,200	9,000	4,800	5,000	6,000	11,000	50,000	171,000	320,000	176,000	64,000	27,000
1960	14,000	12,000	8,800	11,000	8,000	12,000	17,000	45,000	37,000	14,000	11,000	5,500
1961	10,000	12,000	9,000	8,000	8,000	18,000	105,000	144,000	143,000	55,000	78,000	33,000
1962	23,000	13,000	12,000	12,000	16,000	12,000	125,000	160,000	66,000	36,000	18,000	10,000
1963	11,000	11,000	8,000	6,500	7,000	16,000	14,000	11,000	10,000	5,500	7,400	11,000
1964	6,400	7,400	6,000	5,800	5,000	12,000	39,000	72,000	55,000	20,000	15,000	18,000
1965	7,400	9,400	8,300	8,800	7,500	14,000	26,000	64,000	107,000	28,000	58,000	10,000
1966	14,000	11,000	11,000	8,500	7,500	8,800	14,000	29,000	22,000	13,000	15,000	13,000
1967	8,800	8,000	6,400	6,200	6,800	8,000	9,400	15,000	117,000	53,000	38,000	19,000
1968	9,000	6,000	5,000	5,500	5,500	13,000	42,000	119,000	90,000	37,000	24,000	15,000
1969	13,000	11,000	9,500	9,500	8,000	15,000	25,000	57,000	64,000	46,000	20,000	15,000
1970	14,000	11,000	9,500	10,000	9,500	23,000	42,000	113,000	90,000	39,000	23,000	13,000
1971	12,000	12,000	14,000	10,000	7,000	13,000	13,000	40,000	39,000	23,000	15,000	7,700
1972	15,000	9,100	8,000	6,500	7,500	9,100	10,000	16,000	19,000	12,000	8,000	14,000
1973	6,200	7,800	7,500	6,500	7,500	7,000	98,000	265,000	225,000	110,000	27,000	8,800
1974	11,000	8,500	8,000	6,000	7,500	19,000	23,000	40,000	32,000	19,000	17,000	16,000
1975	10,000	7,400	6,500	7,500	6,500	10,000	40,000	57,000	100,000	62,000	28,000	16,000

$\hat{R}$ (= 0.954) มากกว่า R (= 0.130) แสดงว่าการย้ายข้อมูลจาก Ute

Creek และ Cucharas River มาใช้ที่ Trinchera Creek ช่วยเพิ่มความเชื่อถือได้ในการประเมินค่าเฉลี่ยของลือกของปริมาณน้ำท่วมประจำปี

จากตารางที่ 2 สำหรับ $\theta = 1$ เมื่อ $p = 2$, $\tilde{N}_1 = 120$ และ $\tilde{N}_2 = 120$ ได้ R น้อยกว่า 0.08 ซึ่งน้อยกว่า $\hat{R}$ มาก ซึ่งแสดงว่าการย้ายข้อมูลช่วยเพิ่มความเชื่อถือได้ในการประเมินค่าเฉลี่ยของลือกของปริมาณน้ำท่วมประจำปี

จากข้อสรุปดังกล่าวจึงคาดว่า การประเมินค่าเฉลี่ยและแนวโน้มของปริมาณน้ำท่วมประจำปี จะมีความเชื่อถือได้มากขึ้นด้วย

ขั้นที่ 5 ใช้สมการที่ (1) ขยายข้อมูลน้ำท่วมประจำปีของ Trinchera Creek ในช่วง 1956-1960 และ 1971-1975 โดยอาศัยข้อมูลของ Ute Creek และ Cucharas River โดยมีขั้นตอนการคำนวณดังต่อไปนี้ .-

- หาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของลือกของข้อมูลน้ำท่วมประจำปีของ Trinchera Creek ในช่วง 1961 - 1970

$$\bar{Y}_1 = 1.253 \text{ และ } S_1(Y) = 0.353$$

- หาค่าเฉลี่ยของลือกของข้อมูลน้ำท่วมประจำปีของ Ute Creek และ Cucharas River สำหรับช่วง 1961-1970

$$\bar{X}_1^1 = 1.274 \text{ และ } \bar{X}_1^2 = 1.241$$

- หาสัมประสิทธิ์รีเกรสชัน $\hat{b}^1$ และ $\hat{b}^2$ จากสมการที่ 3

$$\hat{b}^1 = 0.2297 \text{ และ } \hat{b}^2 = 0.6180$$

- หาค่าพารามิเตอร์ $\hat{a}$ จากสมการที่ 2

$$\begin{aligned} a &= 1.253 - [0.2297(1.274) + 0.6180(1.241)] \\ &= 0.193 \end{aligned}$$

- เลือกใช้ $\theta = 1$ และคำนวณหา α จากสมการที่ 9

$$= \sqrt{\frac{120(120-2 \times 2-2)(120-1)}{(120-1)(120-2-2)(120-2-1)}} = 1.004$$

$$\hat{R} = 0.954$$

- แทนค่าสัมประสิทธิ์ต่างๆ ลงในสมการที่ 1 จะได้สมการรีเกรซชัน
เส้นตรงแบบตัวแปร

$$\begin{aligned} Y_t &= 0.193 + 0.2297 X_t^1 + 0.618 X_t^2 \\ &\quad + 1.004(1) \sqrt{1-0.954^2} (0.353) \varepsilon_t \\ &= 0.193 + 0.2297 X_t^1 + 0.618 X_t^2 + 0.106 \varepsilon_t \end{aligned}$$

ตัวอย่าง ปริมาณน้ำท่วมในเดือนมิถุนายน 1956 ของ Ute Creek จากตารางที่ 4 ของตอนที่ 1 คือ $X_t^1 = 124$ ลบ.ฟุต/วินาที และของ Cucharas River จากตารางที่ 3 คือ $X_t^2 = 89$ ลบ.ฟุต/วินาที

$$X_t^1 = \log 124 = 2.0934$$

$$X_t^2 = \log 89 = 1.9494$$

สมมุติให้เลขแรนด้อมปกติ (0.1) ที่สร้างขึ้นคือ = 0.461

$$Y_t = 0.193 + 0.2297(2.0934) + 0.618(1.9494) + 0.106(0.461)$$

$$Y_t = 1.9274$$

$$\text{หรือ } Y'_t = \log^{-1}(1.9274) = 84.6 \text{ ลบ.ฟุต/วินาที}$$

ดังนั้นปริมาณน้ำท่วมของ Trinchera Creek ในเดือนมิถุนายน 1956 จะมีค่าเท่ากับ 84.6 ลบ.ฟุต/วินาที

ในทำนองเดียวกันจะสามารถคำนวณปริมาณน้ำท่วมของ Trinchera Creek ในเดือนต่าง ๆ ระหว่าง 1956-1960 และ 1971-1975 ได้ ดังแสดงในตารางที่ 4

ตารางที่ 4 ข้อมูลปริมาณน้ำท่วมประจำเดือนของ Trinchera Creek ที่ขยายเพิ่มขึ้น โดยใช้สมการที่ 1

	ตค.	พย.	ธค.	มก.	กพ.	มีค.	เมย.	พค.	มิย.	กค.	สค.	กย.
1956	4.661	8.437	11.127	7.343	9.692	15.655	38.910	66.144	84.600	54.149	18.471	11.905
1957	15.891	7.494	6.651	7.083	9.110	5.082	37.032	48.006	57.661	36.034	24.855	15.626
1958	28.338	22.813	9.263	10.357	8.708	15.338	52.024	147.752	82.131	17.273	26.601	14.038
1959	7.035	9.405	6.268	4.334	5.156	16.644	37.662	118.683	191.760	88.885	63.801	37.581
1960	10.908	10.912	10.838	8.600	5.572	9.395	14.292	28.291	30.984	15.362	15.438	2.836
	ตค.	พย.	ธค.	มก.	กพ.	มีค.	เมย.	พค.	มิย.	กค.	สค.	กย.
1971	20.285	13.749	15.008	10.026	15.476	13.686	12.191	22.401	48.693	16.994	14.740	14.111
1972	13.997	14.472	9.761	9.116	12.912	7.107	14.355	20.637	24.400	8.843	5.473	11.751
1973	12.744	10.228	9.699	8.508	18.300	11.221	48.867	145.781	131.494	81.844	30.476	19.048
1974	11.524	11.815	7.040	7.377	8.538	17.027	21.702	28.085	21.077	18.146	22.949	5.466
1975	10.519	8.491	10.730	8.879	7.261	13.900	38.870	72.263	62.090	46.116	29.424	14.155

ข้อ 6 จากตารางที่ 4 จะสามารถหาปริมาณน้ำท่วมสูงสุดประจำปีของ Trinchera Creek ที่ขยายเพิ่มขึ้นได้ ส่วนค่าเฉลี่ยและค่าแวนเรียนซ์ใหม่จะสามารถคำนวณได้จากสมการที่ 11 และ 12 ในทำนองเดียวกับที่ได้กล่าวถึงในตอนที่ 1

9. สรุป

ในการย้ายข้อมูลจากสถานีหนึ่งหรือหลายสถานีเพื่อการวิเคราะห์น้ำท่วมตรงจุดที่ต้องการนั้น ข้อมูลทางอุทกวิทยาที่สำคัญในการย้ายข้อมูลคือ ข้อมูลน้ำท่วม และข้อมูลน้ำฝน ปกติแล้วถ้าจำนวนข้อมูลเท่ากัน ข้อมูลน้ำท่วมจะมีประโยชน์กว้างข้อมูลน้ำฝน แต่ผลจากการศึกษาเปรียบเทียบในประเทศอังกฤษ (Natural Environment Research Council, 1975) แนะนำว่าการย้ายทั้งข้อมูลน้ำท่วมและข้อมูลน้ำฝนพร้อมกันจะให้ผลลัพธ์ที่ดีกว่า

เอกสารอ้างอิง

1. วราวุธ วุฒินิชย์. 2530. การย้ายข้อมูลเพื่อเพิ่มความเชื่อถือได้ในการประเมินความถี่ของน้ำท่วม ตอน 1. วารสารวิศวกรรมโยธาและการก่อสร้าง มก.ปีที่ 2 ฉบับที่ 3 เดือนธันวาคม, หน้า 281-297
 2. Gilroy, E.J., 1970, Reliability of A Variance Estimate Obtained from A Sample Agumented by Multivariate Regression, Water Resources Research, Vol. 6, No. 6, pp. 1955-1600.
 3. Natural Environment Research Council, 1975, Flood Studies Report, Vol. 1 : Hydrologic Studies, Whitefriars Press Ltd, London.
 4. US. Water Resources Councils, 1976, Flood Flow Frequency, Bulletin No. 17 of the Hydrology Committee, Washington, D.C.
 5. Yevjevich, V., 1972, Probability and Statistics in Hydrology, Water Resources Publications, Fort Collins, Colorado.
-