บทความ ทางวิชาการ

เทคนิคไร้พารามิเตอร์เพื่อการวิเคราะห์ความถี่น้ำหลาก

สำนักวิจัยและพัฒนา

ดร.สมเกียรติ อภิพัฒนศว์
นายวชิร สามวัง
กลุ่มงานชลศาสตร์ ส่วนวิจัยและพัฒนาด้านวิศวกรรม
สำนักวิจัยและพัฒนา กรมชลประทาน

มิถุนายน 2553

เทคนิคไร้พารามิเตอร์เพื่อการวิเคราะห์ความถี่น้ำหลาก

(Nonparametric Technique for Analyzing Flood Frequency)

ดร.สมเกียรติ อภิพัฒนวิศว์ [1]

นายวชิร สามวัง [2]

[1] วิศวกรโยธาชำนาญการ กลุ่มวิจัยและพัฒนาด้านวิศวกรรม สำนักวิจัยและพัฒนา กรมชลประทาน email:skavis@yahoo.com

[2] วิศวกรโยธาชำนาญการพิเศษ กลุ่มวิจัยและพัฒนาด้านวิศวกรรม สำนักวิจัยและพัฒนา กรมชลประทาน

email:wachira34_river@yahoo.com

บทคัดย่อ

สมมุติฐานของความถี่น้ำหลากที่นิยมใช้กันแพร่หลายที่ว่า น้ำหลากเกิดอย่างอิสระและเกิดขึ้นอย่างมีรูปแบบตามการแจกแจงความถี่ (independent and identical distribution) หรือ น้ำหลากประชากรเดี่ยว (Single Population Flood) ทำให้ต้องวิเคราะห์หาการแจกแจงความถี่ของน้ำหลากก่อน แล้วจึงวิเคราะห์ความถี่น้ำหลากได้ ต่อมาเหล่านักวิจัย พบว่าลุ่มน้ำหลายแห่งในโลก มีน้ำหลากที่มีคุณลักษณะแตกต่างจากสมมุติฐานข้างต้น เนื่องจากลุ่มน้ำเหล่านั้นมีขบวนการทางธรรมชาติที่ก่อให้เกิดน้ำหลากหลายขบวนการ เช่น น้ำหลากจากหิมะละลาย น้ำหลากจากมรสุม เป็นต้น ก่อให้เกิดเป็นน้ำหลากประชากรผสมขึ้น (Mixed population flood) ดังนั้นจึงพัฒนาวิธีวิเคราะห์ความถี่สำหรับน้ำหลากประชากรผสมขึ้น ซึ่งยุ่งยาก ซับซ้อนมากขึ้น เนื่องจากต้องศึกษาวิเคราะห์ขบวนการที่ก่อให้เกิดน้ำหลากในลุ่มน้ำ และแบ่งแยกข้อมูลน้ำหลากตามธรรมชาติที่เกิด แล้วจึงวิเคราะห์ความถี่น้ำหลากได้

การศึกษานี้เสนอ การใช้เทคนิคไร้พารามิเตอร์ (Nonparametric approach) แบบโลคัลโพลิโนเมียล (Local Polynomial Technique) เพื่อวิเคราะห์ความถี่น้ำหลาก ซึ่งสามารถใช้วิเคราะห์ความถี่ของน้ำหลากได้ทั้งสองประเภทคือ น้ำหลากประชากรเดียว และ น้ำหลากประชากรผสม เนื่องจากเทคนิคนี้เป็นแบบไร้พารามิเตอร์ ทำให้ไม่ต้องวิเคราะห์การแจกแจงความถี่ของน้ำหลาก และไม่ต้องแยกข้อมูลน้ำหลากออกตามขบวนการธรรมชาติก่อนทำการวิเคราะห์ ซึ่งช่วยลดขั้นตอนการวิเคราะห์ความถี่น้ำหลากได้ ช่วยให้ทำงานได้รวดเร็วขึ้น ประกอบกับคุณสมบัติของเทคนิคนี้ที่ทำงานตามการขับเคลื่อนของข้อมูล (data-driven) ทำให้สามารถประยุกต์ใช้เทคนิคนี้ได้กับทุกลุ่มน้ำทั่วโลก (portable across sites)

การศึกษานี้มีการทดสอบประสิทธิภาพของเทคนิคนี้ และเปรียบเทียบผลกับประสิทธิภาพของวิธีแบบพาราเมตริก (Parametric approach) ที่ใช้กันมานาน โดยทดสอบกับข้อมูลน้ำหลากสังเคราะห์ (Synthetic data) แบบต่างๆเช่น ข้อมูลเอ็กโปเนนเชียล ข้อมูลแกมมา ข้อมูลเกาเซียน และข้อมูลแบบผสม นอกจากนี้ยังมีการทดสอบกับข้อมูลบันทึกน้ำหลากทั้งในไทยและสหรัฐเอมริกาอีกด้วย พบว่าเทคนิคที่นำเสนอนี้มีประสิทธิภาพสูงกว่าให้ผลที่น่าเชื่อถือกว่า โดยเฉพาะอย่างยิ่งกับน้ำหลากแบบประชากรผสม และที่คาบการเกิดซ้ำสูงมากกว่า 100 ปี

Keywords: Flood Frequency Analysis; Local Polynomial Regression; Mixed Population Flood

คำนำ

การวิเคราะห์ความถี่น้ำหลาก คือการวิเคราะห์หาความสัมพันธ์ระหว่างขนาดของปริมาณน้ำหลากและโอกาสการเกิดของน้ำหลากขนาดนั้นๆ ณ สถานที่ศึกษา โดยทั่วไปจะสนใจน้ำหลากขนาดใหญ่เช่น น้ำหลาก 100 ปี หรือ 500 ปี เป็นต้น เพื่อประโยชน์ในการออกแบบอาคารชลศาสตร์ต่างๆ เช่น เขื่อน อาคารระบายน้ำ สะพาน เป็นต้น การวิเคราะห์ความถี่น้ำหลากมีความสำคัญต่อการวางแผนและการบริหารจัดการทรัพยากรน้ำมาก เช่น การป้องกันน้ำท่วม การปรับปรุงทางน้ำ การออกแบบระบบระบายน้ำ เป็นต้น

การวิเคราะห์ความถี่น้ำหลากแบบดั้งเดิม ใช้สมมุติฐานที่ว่า น้ำหลากเกิดขึ้นโดยอิสระ และการกระจายตัวของปริมาณน้ำหลาก มีรูปแบบการแจกแจงความถี่เฉพาะ (independent and identical distribution) ซึ่งหมายถึงว่าข้อมูลน้ำหลาก เป็นประชากรของการแจงแจกความถี่หนึ่ง (Kite, 1977, Chow et al., 1988) ดังนั้น การวิเคราะห์จึงเริ่มด้วยการทดสอบก่อนว่าข้อมูลน้ำหลากนั้น มีการแจกแจงความถี่แบบใด (เช่น Log Normal, Log Pearson type III, Extreme Value type I เป็นต้น) เมื่อได้การแจกแจงความถี่แล้ว จึงสร้างความสัมพันธ์ระหว่างคาบการเกิดซ้ำ กับ ปริมาณน้ำหลากตามการแจกแจงความถี่นั้น อย่างไรก็ตาม ปัญหาของการวิเคราะห์แบบนี้ คือ เราไม่สามารถแยกแยะได้อย่างชัดเจนว่าข้อมูลน้ำหลาก มีการแจกแจงความถี่แบบใดได้อย่างชัดเจน เนื่องจากว่าข้อมูลน้ำหลากจะสั้น ส่งผลต่อความเชื่อมั่นทางหลักสถิติ

ในปัจจุบัน นักวิจัยจำนวนมากพบอีกว่า ข้อมูลน้ำหลากในหลายลุ่มน้ำทั่วโลกหลายแห่ง ไม่ได้มีลักษณะตามสมมุติฐานเดิมๆ ที่ใช้กันอยู่ โดยพบว่าข้อมูลน้ำหลากมีประชากรน้ำหลากมาจากหลายประเภทการแจกแจงความถี่ อันเนื่องมาจาก ลุ่มน้ำเหล่านั้นมีขบวนการทางธรรมชาติหลายขบวนการที่ก่อให้เกิดฝน ไม่ได้มีขบวนการเดียว (Webb and Betancourt, 1992; Singh et al.,2005) ตัวอย่างเช่น ลุ่มน้ำหนึ่งมีพายุฤดูร้อนก่อให้เกิดฝนในฤดูร้อน และ ฝนจากอิทธิพลของลมมรสุมในฤดูฝน เป็นต้น ทำให้เกิดน้ำหลากที่มีประชากรแบบผสม ทำให้การวิเคราะห์ความถี่ของน้ำหลากแบบนี้มีความซับซ้อนและยุ่งยากกว่าแบบน้ำหลากเดี่ยวมาก โดยต้องเริ่มทำการวิเคราะห์หาจำนวนขบวนการธรรมชาติที่ทำให้เกิดฝนก่อน จากนั้นจึงแยกข้อมูลน้ำหลากแบบผสมนี้ ออกเป็นข้อมูลย่อย ตามขบวนการเกิดฝนของลุ่มน้ำนั้น จากนั้นจึงวิเคราะห์ว่าข้อมูลย่อยเหล่านั้น มีการแจกแจงความถี่อย่างไร ท้ายสุดจึงจะสามารถวิเคราะห์ความถี่น้ำหลากของลุ่มน้ำได้

การศึกษานี้นำเสนอการวิเคราะห์ความถี่น้ำหลากแบบไร้พารามิเตอร์ โดยใช้เทคนิคโลคัลโพลิโนเมียล (local polynomial technique) ด้วยคุณสมบัติแบบไร้พารามิเตอร์ทำให้ไม่ต้องคำนึงถึงการแจกแจงความถี่ของประชากรน้ำหลาก ส่งผลให้การวิเคราะห์ด้วยเทคนิคนี้ใช้ได้ทั้งกับประชากรน้ำหลากแบบเดี่ยวและประชากรแบบผสม (Apipattanavis et al., 2010; Apipattanavis and Sarmwung, 2010) ในบทต่อไปจะกล่าวถึงทฤษฎีการวิเคราะห์ จากนั้นก็เป็นการทดสอบประสิทธิภาพของเทคนิคนี้ กับข้อมูลน้ำหลากสังเคราะห์ และข้อมูลน้ำหลากที่บันทึกได้จากสถานีวัดน้ำในประเทศไทย นอกจากนี้ยังมีการเปรียบเทียบประสิทธิภาพของเทคนิคที่นำเสนอ กับวิธีแบบดั้งเดิมที่ใช้วิเคราะห์กันอย่างแพร่หลายอีกด้วย

การวิเคราะห์ด้วยเทคนิคโลคัลโพลิโนเมียล

จากข้อมูลน้ำหลากสูงสุดที่บันทึกได้ ณ สถานีใดๆ จำนวน n ปี เรานิยาม empirical quantile function ตามลำดับคู่ข้อมูลได้เป็น (X_i, Y_i), $i=1,2,3,...,n$ โดย $X_i = (i)/(n+1)$ และ X_i คือ ค่าน้ำหลากสูงสุดรายปี (เรียงจากค่าน้อยไปหามาก) ทั้งนี้ X_i เรียกชื่อว่า plotting positions จากนั้น เราจึงสร้างแบบจำลองรูปแบบทั่วไปได้เป็น

$$Y_i = \mu(X_i) + \varepsilon_i \quad (1)$$

โดย $\mu(.)$ คือ ฟังก์ชันเส้นไม่ตรง (nonlinear function), ε_i คือ ค่าผิดพลาดซึ่งมีการกระจายตัวตามการแจกแจงความถี่เฉพาะ (identical distributed errors) ที่มีค่าเฉลี่ยเป็นศูนย์ และค่าความแปรปรวนจำกัด $\varepsilon_i \in [0,1]$ เช่นเราสนใจคำนวณหา น้ำหลาก T ปี นั่นคือเราจะหาค่า $\mu(X_T)$ โดยที่ $X_T = 1-1/T$ ในการวิเคราะห์ด้วยเทคนิคนี้ เราประมาณค่า $\mu(X_T)$ ด้วยฟังก์ชันโพลีโนเมียลกำลัง p และใช้ข้อมูลน้ำหลากที่มีค่าใกล้เคียงกับค่าที่พิจารณานี้เพียงจำนวนหนึ่งเท่านั้น เช่น $k = 0.1*n$ หมายความใช้ข้อมูลใกล้เคียงที่สุดเพียงร้อยละ 10 หากมีข้อมูล 80 ปี ก็จะใช้ข้อมูลใกล้เคียงที่สุดเพียง 8 ปีมาทำการวิเคราะห์ (แต่หากเราใช้ค่า $k = 0.1$ การวิเคราะห์นี้จะใช้ทั้งหมด ซึ่งกลายเป็นการวิเคราะห์แบบโพลีโนเมียลทั่วไปนั่นเอง) สำหรับรายละเอียดการวิเคราะห์ด้วยเทคนิคโลคัลโพลิโนเมียล สามารถศึกษาเพิ่มเติมได้จาก Loader (1999)

การวิเคราะห์ความถี่น้ำหลากด้วยเทคนิคโลคัลโพลิโนเมียล มีขั้นตอนคร่าวๆ ดังนี้

1) ที่ค่า X_T ใดๆ เราจะเลือกจำนวนค่า k ใกล้เคียงที่สุด (nearest neighbors) ที่จะใช้คำนวณ ซึ่ง $k = \alpha n$ โดย α มีค่าระหว่าง 0 ถึง 1 สำหรับค่าช่วงกว้าง (bandwidth, $h(X)$ คือหน้าต่างของข้อมูล k neighbors ที่อยู่ไกลที่สุดไปยังจุดที่ต้องการคำนวณ
2) ทุกคู่ของข้อมูลใกล้เคียง k จะได้รับตัวคูณค่าน้ำหนัก โดยค่าที่ใกล้มาก จะได้ค่าตัวคูณน้ำหนักมาก ในทางตรงข้าม ค่าที่ใกล้เคียงน้อยกว่า จะได้รับค่าตัวคูณน้ำหนักน้อยกว่า ซึ่งนิยมใช้ค่าน้ำหนักเป็นฟังก์ชัน (เช่น Bisquare, Tricute เป็นต้น) ในการศึกษานี้ให้น้ำหนักโดยใช้ Bisquare function คือ $W(u_i) = (15/16)*(1-u_i^2)^2$ โดย $u_i = (X_i - X_T)/h(X_T)$ และ $|u| \le 1$
3) ในช่วงข้อมูลที่พิจารณา (คือ k neighbors), ค่า X_T ประมาณได้จากสมการโพลิโนเมียลกำลัง p ตัวอย่างเช่น

$$\mu(X) = a_0 + a_1(X) + a_2(X)^2 \quad (2)$$

ค่าสัมประสิทธิ์ a_0, a_1 และ a_2 คำนวณได้จาก ค่าน้อยสุดของฟังก์ชัน

$$\sum_{i=1}^{k} W_i(X_T)(Y_i - \mu(X_i))^2 \quad (3)$$

ขั้นตอน 1 ถึง 3 กระทำซ้ำสำหรับทุกจุดข้อมูลที่ต้องการวิเคราะห์ความถี่

ในการวิเคราะห์ด้วยเทคนิคนี้ ค่าพารามิเตอร์หลักสองตัวที่สำคัญต่อการคำนวณมาก คือ ค่าที่เหมาะสมของ k และกำลังของโพลิโนเมียล p เพื่อให้ได้ค่าที่เหมาะสมที่สุด เราใช้ฟังก์ชัน Generalized Cross Validation (GCV)

$$GCV(a,p) = n\sum_{i=1}^{n}(Y_i - \hat{\mu}(X_i))^2 \Big/ \left(1 - \sum_{i=1}^{n} h_{ii}\right) \quad (4)$$

ในการวิเคราะห์ หากเราใช้ค่า $h(X)$ ที่มีขนาดเล็กเกินไป ข้อมูลที่ตกในช่วงพิจารณาจะไม่เพียงพอ ทำให้ผลที่ได้มีความแปรปรวนสูงมาก (high variance) ในทางตรงข้าม หากใช้ค่า $h(X)$ ที่มีขนาดใหญ่เกินไป จะก่อให้เกิดเบี่ยงเบนสูง

วิธีการศึกษา

การศึกษานี้มีการทดสอบประสิทธิภาพของเทคนิคที่นำเสนอนี้ โดยทดสอบกับข้อมูลน้ำหลากสังเคราะห์และข้อมูลน้ำหลากบันทึกจริงจากสถานีวัดน้ำ แล้วเปรียบเทียบกับประสิทธิภาพที่ได้จากวิธีที่นิยมใช้กันทั่วไปที่ทดสอบด้วยข้อมูลเดียวกัน

สำหรับข้อมูลน้ำหลากสังเคราะห์นั้น สังเคราะห์ประชากรเดี่ยวจาก Log-Normal distribution (LN), Log-Pearson type 3 distribution (LPIII), และ Extreme Value I distribution (EVI) และสังเคราะห์ข้อมูลประชากรผสมจากการผสมการแจกแจงข้างต้น โดยในแต่ละการสังเคราะห์นั้น เราสังเคราะห์ข้อมูลน้ำหลากจำนวน 500 ชุดข้อมูล แต่ละชุดข้อมูลมีความยาว 80 ปี โดยมีรายละเอียดดังนี้

Log-Normal:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right], \; x > 0 \;\; \text{โดย} \;\; y = \log x, \; \mu_y = \bar{y}, \; \sigma_y = s_y$$

Log-Pearson type 3:

$$f(x) = \frac{\lambda^{\beta}(x-\varepsilon)^{\beta-1} e^{-\lambda(x-\varepsilon)}}{x\Gamma(\beta)}, \; \log x \ge \varepsilon \;\; \text{โดย} \;\; y = \log x, \lambda = s_y/\sqrt{\beta}, \beta = [2/C_s(y)],$$

$$\varepsilon = \bar{y} - s_y\sqrt{\beta} \text{ (assuming is positive)}$$

Extreme Value Type I:

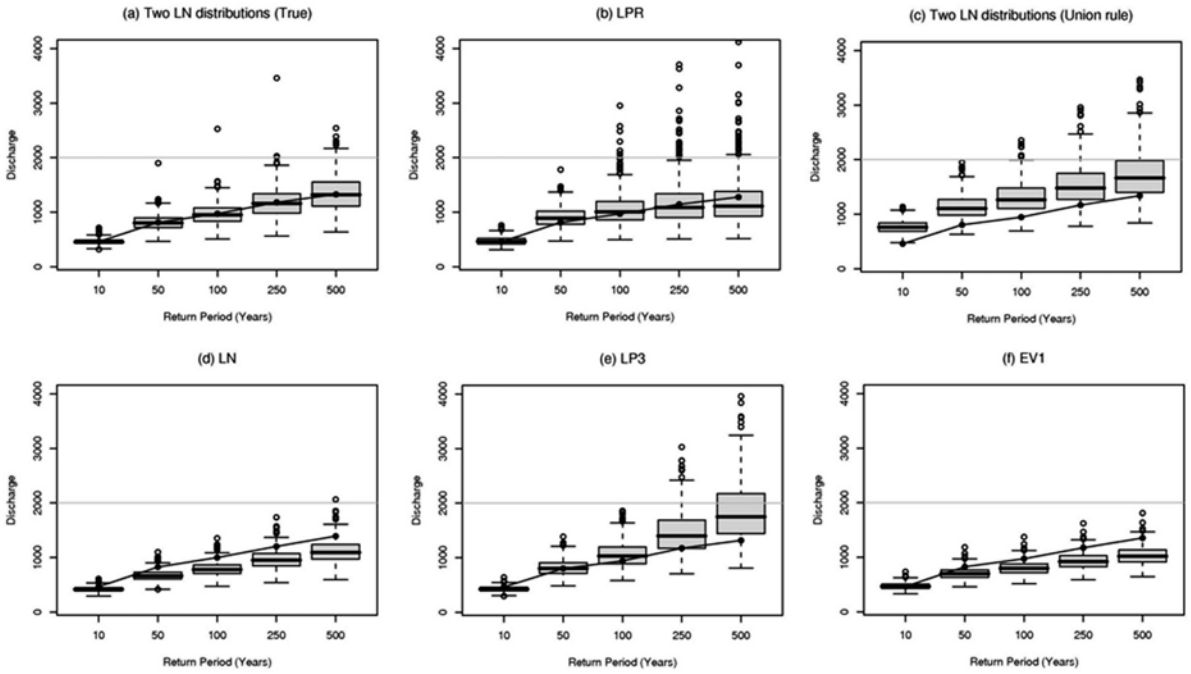
$$f(x)=\frac{1}{\alpha}\exp\left[-\frac{x-u}{\alpha}-\exp\left(-\frac{x-u}{\alpha}\right)\right],-\infty<x<\infty \text{ where } \alpha=\sqrt{6}s_x/\pi,\ u=x-0.5772\alpha$$

ในส่วนของข้อมูลน้ำหลากบันทึกนั้น ใช้ข้อมูลน้ำหลากจากแม่น้ำสองสายคือ แม่น้ำน่าน และแม่น้ำปิง

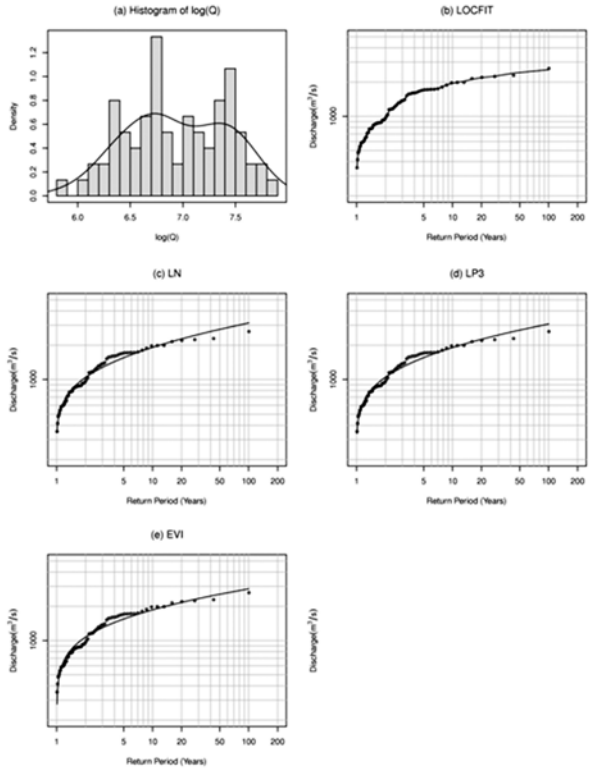
ผลการศึกษา

จากการทดสอบ พบว่า ในกรณีน้ำหลากประชากรเดี่ยว เทคนิคโลลัคโพลิโนเมียลมีประสิทธิภาพใกล้เคียงกับ LPIII อย่างไรก็ตาม เทคนิคนี้ก็มีข้อดีกว่าที่ว่าเราไม่ต้องหาการแจกแจงความถี่ของข้อมูลก่อนทำการวิเคราะห์ความถี่น้ำหลาก สำหรับกรณีน้ำหลากประชากรผสมนั้น เทคนิคนี้ให้ผลลัพธ์ที่ดีกว่าอย่างชัดเจน เนื่องจากวิธีดั้งเดิมที่ใช้วิเคราะห์กันอยู่ทั่วไป (traditional methods) ใช้สมมุติฐานที่ไม่ตรงกับคุณลักษณะของลุ่มน้ำ ซึ่งส่งผลต่อน้ำหลากอย่างมากต่อน้ำหลากที่มีคาบการเกิดซ้ำสูงมากกว่า 100 ปีขึ้นไป ด้วยข้อจำกัดในเรื่องพื้นที่ ผู้ศึกษาสามารถแสดงผลการทดสอบเพียงบางส่วนเท่านั้น ผู้อ่านสามารถดูผลการทดสอบประสิทธิภาพของเทคนิคโพลิโนเมียลนี้ได้จาก Apipattanavis and Sarmwung (2010) ในกรณีน้ำหลากประชากรเดี่ยว และ Apipattanavis et al. (2010) ในกรณีน้ำหลากแบบประชากรผสม

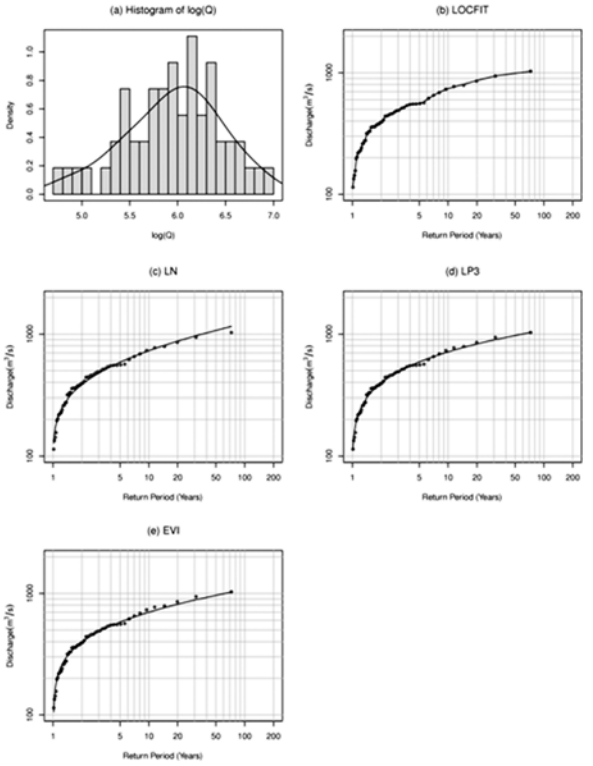
ประสิทธิภาพของวิธีวิเคราะห์แบบต่างๆ บนข้อมูลน้ำหลากสังเคราะห์ประชากรผสม ที่เกิดจาก การแจกแจงผสมของ 2 LN (แสดงในรูปที่ 1) แสดงให้เห็นถึงประสิทธิภาพของเทคนิคโพลิโนเมียล (LPR) ที่เหนือกว่าวิธีการอื่นๆ (LN, LP3, EVI) ข้อสังเกตว่า รูปที่ 1(a) ให้ผลที่ดีกว่า รูปที่ 1(b) นั้นเนื่องจาก รูปที่ 1(a) เป็นการวิเคราะห์โดยใช้ค่าจริงของการแจกแจงน้ำหลากสังเคราะห์ จึงไม่มี bias มีแต่เพียง variance เนื่องจาก sample size



รูปที่ 1 ประสิทธิภาพการวิเคราะห์ความถี่บนข้อมูลสังเคราะห์แบบประชากรผสม (a) True distribution (b) Local Polynomial Regression, LPR (c) Union rule (d) LN (e) LP3 and (f) EVI โดยที่จุดดำต่อเชื่อมด้วยเส้นทึบแสดงค่าจริงของ distribution และ boxplot แสดงค่าจากการวิเคราะห์ที่ได้จากข้อมูลจำนวน 500 ชุด หากจุดดำที่อยู่ใน boxplot แสดงว่ามีประสิทธิภาพดีในเชิงสถิติ ความสูงของ boxplot แสดงถึงความแปรปรวน(variance) ของการวิเคราะห์ และ จุดดำที่เคลื่อนออกจาก box แสดงถึง ความเบี่ยงเบน (bias)



รูปที่ 2 น้ำหลากของแม่น้ำน่าน (a) histogram น้ำหลาก แสดงให้เห็นลักษณะของประชากรผสม, เส้นโค้งความถี่ของ (b) LPR, (c) LN, (d) LP3, และ (e) EVI



รูปที่ 3 เช่นเดียวกับรูปที่ 2 แต่เป็นของแม่น้ำปิง

การทดสอบด้วยข้อมูลบันทึกจากสถานีวัดน้ำสองแห่งในแม่น้ำน่าน และแม่น้ำปิง พบว่า วิธีวิเคราะห์ต่างๆให้ผลที่แตกต่างกันออกไป อย่างไรก็ตาม LPR วิเคราะห์ข้อมูลได้ดีกว่า ในขณะที่วิธีอื่นๆในผลที่สูงเกินจริง (overestimate) หรือต่ำกว่า (underestimate)

สรุปผลการศึกษา

เทคนิคไร้พารามิเตอร์ (Nonparametric approach) แบบโลคัลโพลิโนเมียล (Local Polynomial Technique) เป็นทางเลือกที่ดีอีกทางเลือกหนึ่งในการวิเคราะห์ความถี่น้ำหลากของลุ่มน้ำที่มีข้อมูลไม่ยาวมากพอที่จะวิเคราะห์ได้อย่างชัดเจนในเรื่องการแจกแจงความถี่ของข้อมูล ตลอดจนแยกแยะข้อมูลย่อยตามลักษณะการเกิดของฝนในลุ่มน้ำ เนื่องจากเทคนิคนี้ไม่ต้องคำนึงถึงการแจกแจงความถี่ของข้อมูลน้ำหลากก่อนการวิเคราะห์
โลคัลโพลิโนเมียลสามารถใช้วิเคราะห์ความถี่ของน้ำหลากได้ทั้งสองประเภทคือ น้ำหลากประชากรเดียว และ น้ำหลากประชากรผสม เนื่องจากเทคนิคนี้เป็นแบบไร้พารามิเตอร์ ทำให้ไม่ต้องวิเคราะห์การแจกแจงความถี่ของน้ำหลาก และไม่ต้องแยกข้อมูลน้ำหลากออกตามขบวนการธรรมชาติก่อนทำการวิเคราะห์ ซึ่งช่วยลดขั้นตอนการวิเคราะห์ความถี่น้ำหลากได้ ช่วยให้ทำงานได้รวดเร็วขึ้น ประกอบกับคุณสมบัติของเทคนิคนี้ที่ทำงานตามการขับเคลื่อนของข้อมูล (data-driven) ทำให้สามารถประยุกต์ใช้เทคนิคนี้ได้กับทุกลุ่มน้ำทั่วโลก (portable across sites)

กิตติกรรมประกาศ

คณะผู้ศึกษาขอขอบคุณ ท่านศุภชัย รุ่งศรี ผู้อำนวยการสำนักวิจัยและพัฒนา ที่ให้การสนับสนุนและคำแนะนำที่เป็นประโยชน์อย่างยิ่งต่อการศึกษานี้

เอกสารอ้างอิง

Apipattanavis S., B. Rajagopalan, and U. Lall (2010) Local Polynomial Based Flood Frequency Estimator for Mixed Population, Journal of Hydrologic Engineering (in press).

Apipattanavis S. and W. Sarmwung (2010). Nonparametric Method for Flood Frequency Analysis, Proceeding of 15th National Convention on Civil Engineering, Ubon Ratchthani, Thailand.

Chow, V. T., D. R. Maidment and L. W. Mays, (1988). Applied Hydrology. 572 p.p., McGraw Hill, New York.

Kite, G. W. (1977). Frequency and Risk Analyses in Hydrology, Water Resources Publications. Fort Collins, Colorado.

Loader, C., (1999). Local Regression and Likelihood, 290 p.p., Springer, New York.

Webb R. H. and J. L. Betancourt, (1992). Climatic variability and flood frequency of the Santa Cruz River, Pima County, Arizona, U.S. Geological Survey, Water Supply Paper; 2379.

Singh, V. P., S. X. Wang, and L. Zhang, (2005). Frequency analysis of nonidentically distributed hydrologic flood data, Journal of Hydrology, 307, 175-195.