

## PHÂN TÍCH TẦN SUẤT MƯA CỰC HẠN CHO TỈNH GIA LAI DỰA TRÊN CÁCH TIẾP CẬN VÙNG

Nguyễn Chí Công<sup>1</sup>

**Tóm tắt:** Hiện nay ở Việt Nam, việc phân tích tần suất mưa phục vụ thiết kế công trình thủy đang sử dụng cách tiếp cận địa phương. Cách tiếp cận này dựa trên số liệu thực đo hạn chế của một trạm và sử dụng suy luận tần suất để ước tính các giá trị mưa cực hạn ứng với thời gian lặp lại rất lớn (từ 100 đến 1000 năm). Tuy nhiên, bài báo này giới thiệu cách tiếp cận khác, trong đó sử dụng cách tiếp cận vùng và phương pháp suy luận Bayesian để làm lớn kích thước mẫu dữ liệu thống kê và ước tính được độ tin cậy của suy luận. Cơ sở dữ liệu sử dụng là tài liệu mưa ngày của 26 trạm đo mưa trên địa bàn tỉnh Gia Lai và lân cận. Các thời đoạn mưa tính toán thường dùng trong thiết kế công trình thủy là 1, 3, 5 và 7 ngày lớn nhất được xác định. Các bước thực hiện bao gồm: kiểm tra tính đồng nhất của dữ liệu, phân vùng đồng nhất, lựa chọn hàm phân phối xác suất phù hợp và phân tích tần suất vùng.

**Từ khóa:** cách tiếp cận vùng, suy luận Bayesian, độ tin cậy, cực hạn, tỉnh Gia Lai.

### 1. ĐẶT VẤN ĐỀ

Trong tính toán thủy văn, các kỹ sư thường dựa vào dữ liệu hạn chế của một trạm đo mưa để phân tích tần suất và ước tính mưa thiết kế (cách tiếp cận địa phương). Tuy nhiên, thời gian lặp lại ứng với tần suất thiết kế thường là rất lớn (từ 100 đến 1000 năm) nên việc ước tính mưa thiết kế là không chắc chắn. Để khắc phục hạn chế này, các nghiên cứu trên thế giới đã áp dụng cách tiếp cận vùng trong phân tích tần suất để làm lớn kích thước mẫu dữ liệu và giảm sự không chắc chắn của suy luận (Hosking et al 1997; Ngogondo et al 2011; Nguyen Chi Cong et al 2014). Nếu một vùng được xem là đồng nhất thì có thể nhóm dữ liệu của các trạm đo trong vùng, thông qua đó kích thước mẫu dữ liệu vùng sẽ lớn hơn rất nhiều và sau khi phân tích tần suất vùng sẽ phân phối lại cho các trạm đo thông qua chỉ số mưa vùng. Cách tiếp cận vùng sử dụng thuật toán Bayesian Markov chain Monte Carlo (MCMC) để ước tính độ tin cậy và cho phép thêm các

thông tin nhằm giảm sự không chắc chắn của giá trị ước tính tần suất (Gaume et al 2010; Nguyen Chi Cong et al 2014).

Khu vực Tây Nguyên nói chung và tỉnh Gia Lai nói riêng có chế độ mưa khá phức tạp. Cụ thể là phần diện tích phía Tây Trường Sơn chịu tác động của gió Tây Nam hoạt động mạnh vào tháng 5 đến tháng 10. Ngược lại, phần diện tích phía Đông Trường lại chịu tác động của gió Đông Bắc hoặc áp thấp nhiệt đới từ biển vào gây mưa lớn vào tháng 9 đến tháng 12. Với đặc thù này, khi áp dụng cách tiếp cận mưa vùng cho tỉnh Gia Lai sẽ xuất hiện các vấn đề đặt ra là (i) dữ liệu tất cả các trạm đo trong vùng có đồng nhất hay không, (ii) Nếu không đồng nhất thì việc phân chia vùng sẽ như thế nào, (iii) phân phối xác suất nào là phù hợp nhất cho các vùng và cho các thời đoạn mưa tính toán là 1; 2; 3; 5 và 7 ngày lớn nhất.

### 2. GIỚI THIỆU VÙNG NGHIÊN CỨU VÀ DỮ LIỆU

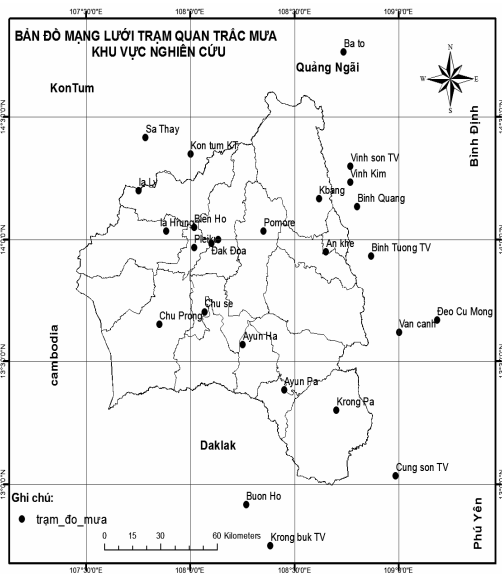
#### 2.1. Giới thiệu vùng nghiên cứu

Tỉnh Gia Lai thuộc khu vực Tây Nguyên với tổng diện tích trên 15.536 km<sup>2</sup>. Địa hình vùng

---

<sup>1</sup> Khoa Xây dựng Thủy lợi - Thủy điện, Trường Đại học Bách Khoa – Đại học Đà Nẵng.

ngiên cứu đa phần là núi cao và bị chia cắt bởi dải Trường Sơn, phía Bắc giáp với tỉnh Kon Tum và Quảng Ngãi; phía Nam giáp với tỉnh Đắk Lắk và Phú Yên; phía Tây giáp với Campuchia và phía Đông giáp với tỉnh Bình Định. Khí hậu của vùng nghiên cứu có đặc trưng của khí hậu Tây và Đông Trường Sơn. Do đó, hình thái mây mưa lớn trong vùng Đông Trường Sơn thường do ảnh hưởng của hoạt động gió mùa Đông Bắc và áp thấp nhiệt đới từ biển Đông (từ tháng 10 đến tháng 12). Đối với vùng Tây Trường Sơn, mưa lớn thường do hoạt động mạnh của gió Tây Nam (từ tháng 5 đến tháng 9). Các hoạt động khí hậu này kết hợp với địa hình núi cao tạo sườn đón gió nên thường tạo ra lượng mưa rất lớn trong vùng.



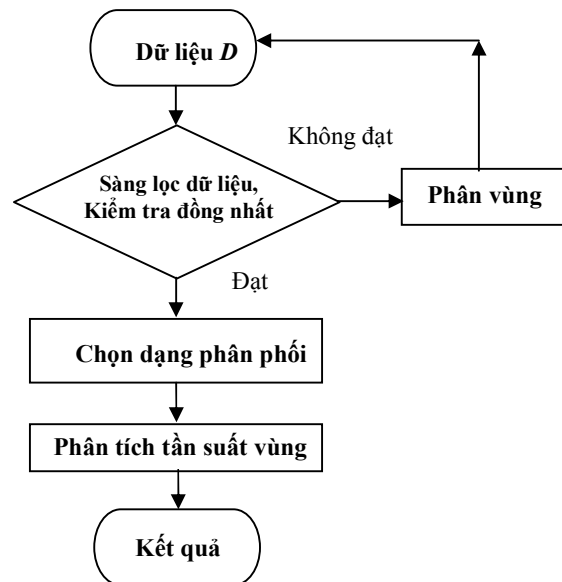
Hình 1. Mạng lưới trạm đo mưa (hình trái) và quy trình thực hiện (hình phải)

Bảng 1. Thông tin số năm và thời gian đo mưa của 26 trạm

TT	Trạm	Năm	Thời gian	TT	Trạm	Năm	Thời gian
1	Sa Thầy	27	1988-2014	14	Vĩnh Sơn	20	1995-2014
2	Kon Tum	39	1976-2014	15	Vĩnh Kim	32	1983-2014
3	Ia Ly	15	2000-2014	16	Bình Quang	17	1998-2014
4	Ia Hrun	15	2000-2014	17	Kbang	26	1989-2014
5	Biển Hồ	22	1993-2014	18	An Khê	38	1977-2014
6	Đắk Đoa	35	1980-2014	19	Bình Tường	38	1977-2014
7	Thôn 4	22	1993-2014	20	Cù Mông	38	1977-2014
8	Pleiku	59	1956-2014	21	Vân Canh	24	1991-2014
9	Pomore	37	1978-2014	22	Ayun Hạ	16	1999-2014
10	Chư Sê	21	1994-2014	23	Krông Pa	35	1980-2014
11	Chư Prông	37	1978-2014	24	Củng Sơn	36	1979-2014
12	Ayun Hạ	16	1999-2014	25	Buôn Hồ	33	1982-2014
13	Ba Tơ	38	1977-2014	26	Krông Buk	37	1977-2013

## 2.2. Dữ liệu

Trong cách tiếp cận vùng, để làm lớn kích thước mẫu dữ liệu vùng bằng cách nhóm dữ liệu của các trạm đo mưa trong vùng nhưng với điều kiện vùng đó phải đồng nhất về dữ liệu. Do đó nghiên cứu này sử dụng 14 trạm đo mưa thuộc tỉnh Gia Lai và 12 trạm đo mưa thuộc các tỉnh lân cận. Trong đó trạm có số năm quan sát dài nhất là trạm Pleiku với 59 năm (từ 1956 đến 2014) và trạm có số năm quan sát ngắn nhất là trạm IaLy và IaHrun với 15 năm (từ 2000 đến 2014). Mật độ các trạm đo phân bố không đồng đều về không gian. Chủ yếu tập trung ở những nơi có địa hình thuận lợi. Vùng núi cao hiểm trở hoặc tiếp giáp với Campuchia không có trạm đo (hình 1 bên trái và bảng 1).



### 3. PHƯƠNG PHÁP

Giả sử có một mẫu dữ liệu mưa vùng  $D$ , với  $s$  là số lượng các trạm đo mưa và  $n_i$  là số năm quan sát của trạm đo thứ  $i$ th. Khi đó, một phần tử trong mẫu dữ liệu  $D$  là  $x_{ij}$  với  $i = 1, \dots, s$  và  $j = 1, \dots, n_i$ . Quy trình thực hiện nghiên cứu này có thể tóm lược theo sơ đồ hình 1 bên phải.

#### 3.1. Sàng lọc dữ liệu

Nghiên cứu này sử dụng phương pháp kiểm tra được đề xuất bởi Hosking và Wallis (1997). Phương pháp này kiểm tra dữ liệu đo của một trạm nào đó có quy luật phân phối khác với quy luật phân phối của những trạm còn lại trong vùng nghiên cứu. Việc sàng lọc này thông qua đánh giá tính không phù hợp  $D_i$  dựa trên L-moment của từng trạm đo (Hosking và Wallis, 1997). Nếu dữ liệu của một trạm bị lỗi, thì  $D_i \geq 3$ . Giá trị phân tán  $D_i$  cho một trạm phụ thuộc vào chính dữ liệu của trạm đó và được Hosking và Wallis (1997) đề xuất theo công thức sau:

$$D_i = \frac{1}{3} N(u_i - \bar{u})^T S^{-1} (u_i - \bar{u}) \quad (1)$$

Trong đó  $u_i = [\tau^{(i)}, \tau_3^{(i)}, \tau_4^{(i)}]^T$  là vector chứa các giá trị  $\tau$ ,  $\tau_3$ , and  $\tau_4$  tương ứng với giá trị trung bình của hệ số biến đổi (L-CV), hệ số thiên lệch L-skewness (L-CA) và hệ số nhọn L-Kurtosis của trạm thứ  $i$  trong vùng, số mũ  $T$  biểu thị chuyển vị của vector hoặc ma trận,  $\bar{u}$  là trung bình trọng số của tỉ số L-moment và  $S$  được xác định theo hai công thức dưới đây:

$$s = \sum_{i=1}^N (u_i - \bar{u})(u_i - \bar{u})^T \quad (2)$$

#### 3.2. Kiểm tra đồng nhất

Theo Hosking và Wallis (1997) lượng mưa trong vùng sẽ được mô phỏng ( $N_{sim} = 500$  lần lặp) với các mẫu lấy từ 4 tham số của phân phối Kappa (Hosking và Wallis, 1997, pp. 202-204), 4 tham số của phân phối Kappa đặc trưng cho các giá trị trung bình L-moment:  $l_1^R, \tau^R, \tau_3^R$  and  $\tau_4^R$ . Trong mỗi lần mô phỏng  $V$ ,  $V_2$  và  $V_3$  sẽ được tính toán. Với  $\mu_V, \mu_{V_2}$  và  $\mu_{V_3}$  là giá trị trung bình,  $\sigma_V, \sigma_{V_2}$ , và  $\sigma_{V_3}$  là độ lệch chuẩn ứng với  $N$  lần mô phỏng của  $V$ ,  $V_2$  and  $V_3$ . Các giá trị thống kê này được ước tính theo ba chỉ số đồng nhất sau:

$$\begin{cases} H_1 = \frac{(V - \mu_V)}{\sigma_V} \\ H_2 = \frac{(V_2 - \mu_{V_2})}{\sigma_{V_2}} \\ H_3 = \frac{(V_3 - \mu_{V_3})}{\sigma_{V_3}} \end{cases} \quad (3)$$

Hosking và Wallis (1997) đưa ra các điều kiện để một vùng có thể xem: “chấp nhận đồng nhất” nếu  $H_{1,2,3} < 1$ , “có thể đồng nhất” nếu  $1 < H_{1,2,3} < 2$ , và “không đồng nhất” nếu  $H_{1,2,3} > 2$ . Khi dữ liệu vùng có  $H_{1,2,3} > 2$  thì cần tiến hành phân chia vùng. Nếu một trong ba giá trị của  $H > 2$  thì vùng được xem là không đồng nhất.

#### 3.3. Phân vùng đồng nhất

Theo Hosking và Wallis 1997, vùng đồng nhất là vùng mà trong đó các trạm quan trắc khác nhau có cùng tỷ lệ phân bố xác suất. Số lượng vùng đồng nhất và số lượng trạm trong mỗi vùng bước đầu được xác định thông qua phương pháp K-means. K-means là phương pháp tự động phân nhóm dữ liệu trong thống kê và được áp dụng rộng rãi trong nhiều lĩnh vực. Trong thủy văn, phương pháp này được Satyanarayana và Srinivas giới thiệu vào năm 2008. Dữ liệu của mỗi trạm đo được xem như một vector. Để loại bỏ các tác động gây ra bởi sự khác biệt giữa phương sai và trung bình của các chuỗi số liệu khi so sánh với nhau, do đó các vector được điều chỉnh lại tỷ lệ theo công thức sau.

$$y_{ij} = \frac{(x_{ij} - \bar{x}_j)}{\sigma_j} \text{ for } 1 \leq j \leq n \quad (4)$$

Trong đó  $y_{ij}$  là giá trị điều chỉnh tỷ lệ của  $x_{ij}$ ,  $\sigma_j$  là độ lệch chuẩn thứ  $j$ , và  $\bar{x}_j$  là giá trị trung bình thứ  $j$ . Thông qua một thủ tục lặp, thuật toán K-means dịch chuyển các vector từ nhóm này sang nhóm khác để làm nhỏ hàm mục tiêu  $F$ , và  $F$  được xác định như sau:

$$F = \sum_{k=1}^K \sum_{j=1}^n \sum_{i=1}^{n_k} d^2(y_{ij}^k - y_j^k) \quad (5)$$

Trong đó số nhóm  $K$  được giả định ngay ban đầu,  $n_k$  là số các vector trong một nhóm  $k$ ,  $d$  là khoảng cách của mỗi vector đến trung tâm của nhóm,  $y_{ij}^k$  là giá trị điều chỉnh tỷ lệ của thuộc tính trong vector  $i$ ,  $y_j^k$  là giá trị trung bình thứ  $j$  cho nhóm  $k$  và được tính bằng:

$$y_j^k = \frac{\sum_{i=1}^{n_k} y_{ij}^k}{n_k} \quad (6)$$

### 3.4. Chọn phân phối xác suất

Trong phân tích tần suất (PTTS) vùng, hàm phân phối xác suất ( $F$ ) được chọn dựa trên tỷ lệ L-moment và giá trị  $Z^{Dist}$  (goodness-of-fit). Với mỗi dạng phân phối,  $Z^{Dist}$  được tính toán như sau:

$$Z^{Dist} = \frac{t_4^R - \tau_4^{Dist}}{\sigma_4} \quad (7)$$

Trong đó  $\tau_4^R$  là giá trị trung bình L-kurtosis tính từ dữ liệu vùng,  $\tau_4^{Dist}$  là giá trị L-kurtosis lý thuyết tính từ mô phỏng cho một dạng phân phối, và  $\sigma_4$  là giá trị độ lệch chuẩn của L-kurtosis nhận được từ mô phỏng dữ liệu. Theo Hosking và Wallis (1997) đề xuất giá trị chấp nhận là  $|Z^{Dist}| \leq 1.64$  và phân phối nào có giá trị  $|Z^{Dist}|$  thỏa mãn điều kiện này sẽ được lựa chọn (Hosking và Wallis, 1997, pp. 80-83).

### 3.5. Phân tích tần suất mưa vùng

Trong phân tích tần suất mưa vùng, dữ liệu vùng được tính bằng tỷ lệ dữ liệu thực đo của mỗi trạm chia cho chỉ số mưa của trạm đó.

$$X' = \sum_{j=1}^{n_i} \frac{X_{ij}}{\mu_i} \quad (8)$$

Trong đó  $\mu_i$  là chỉ số mưa vùng được đề xuất bởi Hosking và Wallis (1997). Chỉ số mưa vùng được xác định như sau:

$$\mu_i = \frac{1}{n_j} \sum_{j=1}^{n_i} X_{ij} \quad (9)$$

Thông qua thuật toán Bayesian MCMC hiện đang được sử dụng khá rộng rãi cho các ứng dụng thủy văn (Gaume et al 2010; Nguyen Chi Cong et al 2014) để ước tính lượng mưa vùng  $\hat{X}(F)$  ứng với các tần suất dựa trên một quy luật phân bố  $F$  đã lựa chọn. Kết quả ước tính lượng mưa của vùng sẽ được phân phối lại cho các trạm thông qua chỉ số mưa vùng  $\mu_i$ . Khi đó giá trị ước tính lượng mưa ứng với các tần suất tại các trạm được xác định như sau:

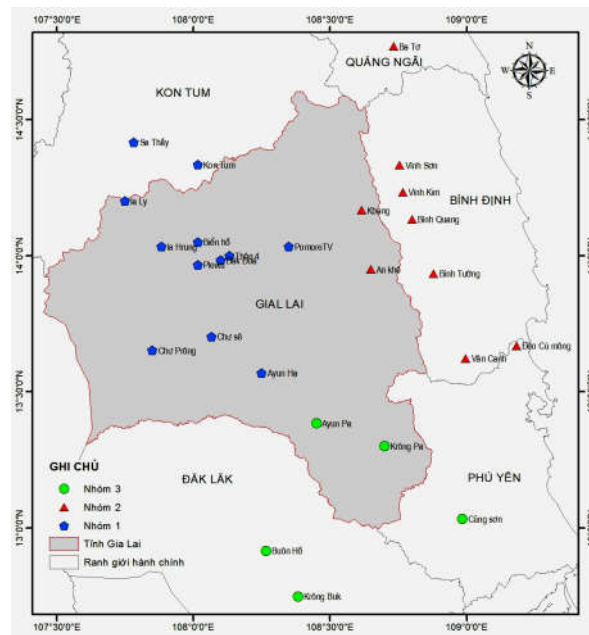
$$\hat{X}(F) = \mu_j \hat{X}(F) \quad (10)$$

Thông qua đó, kích thước mẫu dữ liệu tại các

trạm đo được làm lớn lên rất nhiều và bằng tổng kích thước mẫu dữ liệu của các trạm đo trong vùng. Điều này đã làm tăng độ tin cậy các giá trị suy luận, đặc biệt trong vùng ngoại suy (phần đuôi của đường cong tần suất). Quy trình thực hiện của nghiên cứu này được trình bày trong hình 1 bên phải.

## 4. KẾT QUẢ VÀ BÀN LUẬN

Từ dữ liệu mưa ngày của 26 trạm, tiến hành xây dựng các tổ hợp mẫu dữ liệu (**D1**, **D3**, **D5** và **D7**) tương ứng với các thời đoạn mưa tính toán (1, 3, 5 và 7 ngày lớn nhất) cho mỗi trạm. Cho phân nhóm  $K = 1, 2$  và 3 nhóm vùng và tiến hành kiểm tra đồng nhất dữ liệu của từng vùng. Kết quả cho thấy khi  $K=1$  và 2 thì dữ liệu vùng không thỏa mãn điều kiện đồng nhất ( $H_n > 2$ ). Khi  $K=3$  có nghĩa là phân chia 3 vùng, với số lượng trạm đo mưa mỗi vùng như sau: vùng I có 12 trạm, vùng II có 9 trạm và vùng III có 5 trạm (hình 2). Kết quả kiểm tra đồng nhất và sàng lọc dữ liệu của 3 vùng ứng với các mô hình mưa bất lợi được thể hiện trong bảng 2. Kết quả bảng 2 cho thấy dữ liệu mưa các mô hình của 3 vùng đều thỏa mãn điều kiện đồng nhất  $H_n < 2$  và giá trị phân tán của mỗi trạm  $D_i < 3$ .



Hình 2. Kết quả phân vùng đồng nhất

**Bảng 2. Kết quả sàng lọc và kiểm tra đồng nhất dữ liệu vùng**

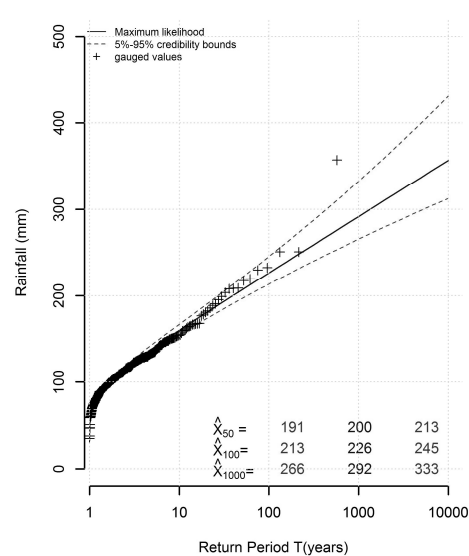
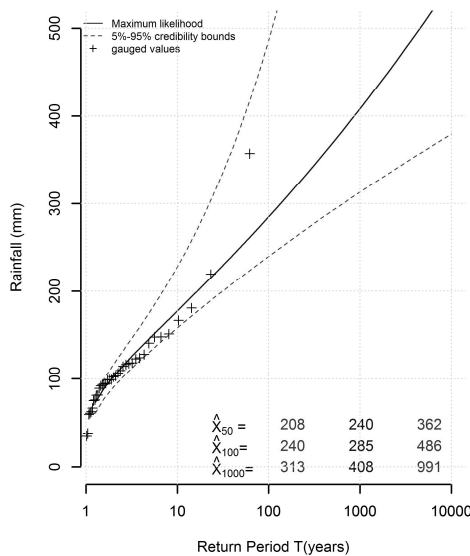
Thời đoạn mưa	Dữ liệu vùng	Tên trạm và giá trị phân tán ( $D_i$ ) mỗi trạm	Giá trị $H_n$		
			$H_1$	$H_2$	$H_3$
1 ngày	<b>D1-I</b>	Sa Thầy (0.27), Kun Tum (0.74), Ia Ly (2.96), Ia Hrunng (0.11), Biền Hồ (0.46), Đăk Đoa (0.87), Thôn 4 (1.69), Pleiku (0.59), Pomore (0.32), Chư Sê (0.37), Chư Prông (2.77), Ayun Ha (0.80)	0.34	0.76	0.77
	<b>D1-II</b>	Ba Tơ (0.97), Vĩnh Sơn (1.02), Vĩnh Kim (0.44), Bình Quang (0.59), Kbang (1.05), An Khê (0.23), Bình Tường (1.20), Cù Mông (1.46), Vân Canh (2.00)	-1.85	0.50	0.72
	<b>D1-III</b>	Ayun Pa (0.23), Krong Pa (1.32), Củng Sơn (1.30), Buôn Hồ (1.25), Krông Buk (0.87)	<b>1.56</b>	-0.42	-0.90
3 ngày	<b>D3-I</b>	Sa Thầy (0.73), Kun Tum (0.14), Ia Ly (2.53), Ia Hrunng (2.75), Biền Hồ (0.88), Đăk Đoa (0.43), Thôn 4 (1.79), Pleiku (0.21), Pomore (0.33), Chư Sê (0.49), Chư Prông (1.01), Ayun Ha (0.66)	-0.67	-0.01	0.26
	<b>D3-II</b>	Ba Tơ (1.19), Vĩnh Sơn (0.37), Vĩnh Kim (0.49), Bình Quang (1.24), Kbang (0.91), An Khê (1.56), Bình Tường (1.54), Cù Mông (0.70), Vân Canh (0.94)	-1.04	<b>1.67</b>	<b>1.95</b>
	<b>D3-III</b>	Ayun Pa (1.33), Krong Pa (1.26), Củng Sơn (1.33), Buôn Hồ (0.74), Krông Buk (0.32)	0.04	0.04	-0.47
5 ngày	<b>D5-I</b>	Sa Thầy (1.15), Kun Tum (0.29), Ia Ly (2.76), Ia Hrunng (2.24), Biền Hồ (0.88), Đăk Đoa (0.22), Thôn 4 (0.60), Pleiku (0.24), Pomore (0.10), Chư Sê (0.63), Chư Prông (1.15), Ayun Ha (1.67)	-1.15	-0.87	-0.46
	<b>D5-II</b>	Ba Tơ (1.24), Vĩnh Sơn (0.16), Vĩnh Kim (2.25), Bình Quang (1.14), Kbang (1.07), An Khê (1.15), Bình Tường (0.84), Cù Mông (0.20), Vân Canh (0.90)	-0.58	0.79	0.47
	<b>D5-III</b>	Ayun Pa (1.25), Krong Pa (0.94), Củng Sơn (1.31), Buôn Hồ (0.90), Krông Buk (0.58)	0.75	-0.33	-0.92
7 ngày	<b>D7-I</b>	Sa Thầy (1.02), Kun Tum (0.62), Ia Ly (1.70), Ia Hrunng (0.52), Biền Hồ (0.84), Đăk Đoa (1.00), Thôn 4 (0.71), Pleiku (0.30), Pomore (0.60), Chư Sê (1.73), Chư Prông (1.25), Ayun Ha (1.66)	-1.47	-1.16	-0.58
	<b>D7-II</b>	Ba Tơ (1.44), Vĩnh Sơn (1.30), Vĩnh Kim (0.54), Bình Quang (1.59), Kbang (1.40), An Khê (0.83), Bình Tường (0.70), Cù Mông (0.15), Vân Canh (1.01)	-1.23	-1.17	-1.61
	<b>D7-III</b>	Ayun Pa (1.11), Krong Pa (0.95), Củng Sơn (1.27), Buôn Hồ (0.76), Krông Buk (0.88)	0.79	-0.78	-0.92

Bảng 3 trình bày kết quả lựa chọn phân phối xác suất phù hợp với từng mẫu dữ liệu vùng. Các giá trị được lựa chọn phải thỏa mãn  $|Z^{Dist}| < 1.64$ . Trong đó GEV (Generalized extreme value), GLO (Generalized logistic), GNO (Generalized Normal) và PE3 (Pearson type III). Các dạng phân phối này đều thuộc họ 3 tham số.

Kết quả cho thấy mỗi vùng dữ liệu đều có ít nhất 2 dạng phân phối phù hợp (vùng I có GEV và GLO; vùng II có GEV, GNO và PE3; vùng III có GEV, GLO và GNO). Tuy nhiên, phân phối GEV là phổ biến nhất cho 3 vùng. Do đó tác giả sử dụng phân phối GEV để tiến hành PTTS vùng.

**Bảng 3. Kết quả lựa chọn hàm phân phối xác suất phù hợp**

Vùng	Dữ liệu mưa	Giá trị $ Z^{Dist}  < 1.64$			
		GEV	GLO	GNO	PE3
I (12 trạm)	<b>D1-I</b>	1.55	0.16	---	---
	<b>D3-I</b>	0.97	0.86	1.12	---
	<b>D5-I</b>	1.62	0.07	---	---
	<b>D7-I</b>	1.38	0.30	1.58	---
II (9 trạm)	<b>D1-II</b>	0.21	---	0.19	0.21
	<b>D3-II</b>	0.28	---	0.35	0.01
	<b>D5-II</b>	0.35	---	0.57	0.33
	<b>D7-II</b>	0.29	---	0.06	0.09
III (05 trạm)	<b>D1-III</b>	0.11	0.94	0.33	1.12
	<b>D3-III</b>	1.27	0.74	---	---
	<b>D5-III</b>	0.59	0.03	1.05	---
	<b>D7-III</b>	0.22	0.95	0.27	1.15



Hình 3. So sánh kết quả PTTS giữa cách tiếp cận địa phương (hình trái) và cách tiếp cận vùng (hình phải) sau khi phân phối cho trạm Chu Prông thông qua chỉ số mưa vùng  $\mu_i$ .

Để thấy được sự không chắc chắn trong suy luận giữa hai cách tiếp cận, tác giả lựa chọn đại diện 1 trạm đo mưa trong vùng mà trạm đó có chứa giá trị đo lớn nhất trong vùng. Trạm

được chọn là trạm Chu Prông với 37 năm đo mưa (vùng I), trong đó có giá trị mưa 1 ngày lớn nhất vào năm 1979 với lượng mưa đo được là 357.2 (mm). Hình 3 thể hiện kết quả

PTTS thời đoạn tính toán 1 ngày lớn nhất cho trạm Chư Prông dựa trên suy luận Bayesian MCMC bằng cách tiếp cận địa phương (hình trái) và tiếp cận vùng (hình phải), đường nét liền thể hiện giá trị suy luận Maximum Likelihood, hai đường nét đứt thể hiện giá trị suy luận tương ứng với độ tin cậy 90%, các điểm chấm là giá trị đo.

Hình 3 (trái) cho thấy do hạn chế về kích thước mẫu thống kê và giá trị đặc biệt của năm 1997 nên phần đuôi của đường tần suất bị điều

chỉnh lên rất cao, đồng thời sự không chắc chắn của suy luận ở phần đuôi là rất lớn (khoảng cách giữa 2 đường đứt nét là rất rộng).

Hình 3 (Phải) cho thấy cách tiếp cận vùng giúp làm lớn kích thước mẫu số liệu (795 năm) và tăng sự chắc chắn trong suy luận (khoảng cách giữa 2 đường nét đứt được thu hẹp).

Bảng 4 trình bày kết quả PTTS vùng cho các trạm đo mưa thuộc tỉnh Gia Lai ứng với các thời đoạn mưa tính toán và thời gian lặp lại là T=100 và 1000 năm.

**Bảng 4. Kết quả ước tính lượng mưa ứng với T=100 và 1000 năm (đơn vị: mm)**

Vùng	Trạm	1-NLN		3-NLN		5-NLN		7-NLN	
		100	1000	100	1000	100	1000	100	1000
I	Ia Ly	201	259	300	384	382	487	451	565
	Ia Hrung	243	313	406	522	545	695	648	812
	Biển Hồ	230	297	385	494	508	647	587	735
	Đăk Đoa	217	280	329	423	424	541	483	605
	Thôn 4	218	280	353	453	461	587	530	664
	Pleiku	222	286	372	477	476	606	551	691
	Pomore	236	304	355	456	427	545	494	619
	Chư Sê	228	294	343	440	448	570	500	626
	Chư Prông	226	292	414	531	552	704	644	806
	Ayun Hạ	148	191	193	248	237	302	281	352
II	Kbang	291	372	465	605	525	662	588	726
	An Khê	294	376	481	626	536	675	596	737
III	Ayun Pa	293	437	397	602	433	647	483	706
	Krông Pa	343	511	494	749	553	827	588	859

## 5. KẾT LUẬN

Nghiên cứu này đã áp dụng cách tiếp cận vùng trong PTTS mưa cực hạn cho tỉnh Gia Lai. Dựa trên dữ liệu mưa ngày thu thập được, tác giả đã xây dựng 4 thời đoạn mưa tính toán thường dùng trong thiết kế công trình thủy. Kết quả nghiên cứu cho thấy tổ hợp dữ liệu 26 trạm là không đồng nhất và đã tìm ra 3 vùng đồng nhất với số lượng trạm đo trong mỗi vùng là: vùng I (12 trạm), vùng II (9 trạm) và vùng III (5 trạm). Hàm phân phối xác suất phù hợp và phổ biến cho dữ liệu 3 vùng là phân phối GEV. Kết quả PTTS vùng cho thấy

đã khắc phục được những hạn chế của cách tiếp cận địa phương như: mẫu dữ liệu ngắn, sự không chắc chắn trung sự luận cao. Thông qua cách tiếp cận vùng và thuật toán Bayesian MCMC đã ước tính được lượng mưa cực hạn (T=100 và 1000 năm) cho các mô hình mưa bất lợi tại các trạm đo trên địa bàn tỉnh Gia Lai.

## LỜI CẢM ƠN

Tác giả xin chân thành cảm ơn quỹ Khoa học công nghệ, Đại học Đà Nẵng đã hỗ trợ tác giả thu thập số liệu đo mưa của 26 trạm đo trong vùng nghiên cứu.

## TÀI LIỆU THAM KHẢO

- Nguyễn Chí Công, Gaume E, Payrastre O (2014). *Regional flood frequency analyses involving extraordinary flood events at ungauged sites : further developments and validation*. Journal of Hydrology, 508, 385-396.
- Gaume E, Gaal L, Viglione A, Szolgay J, Kohnova S, Bloschl G (2010). *Bayesian MCMC approach to regional flood frequency analyses involving extraordinary flood events at ungauged sites*. Journal of Hydrology 394, 101-117.
- Halbert K, Nguyễn Chí Công, Payrastre O, Gaume E (2016). *Reducing uncertainty in flood frequency analyses : A comparison of local and regional approaches involving information on extreme historical*. Journal of Hydrology, 541, 90-98.
- Hosking, J. and J. Wallis (1997), *Regional frequency analysis: An approach Based on L-Moments*, Cambridge University Press, London, UK.
- Ngogondo CS, C-Y. Xu, L.M.Tallaksen, B. Alemaw and T. Chirwa (2011). *Regional frequency analysis of rainfall extremes in Southern Malawi using the index rainfall and L-moments approaches*. Stoch. Env. Res. Risk A. 25, 939-955.
- Satyanarayana P, SrinivasVV (2008). *Regional frequency analysis of precipitation using large-scale atmospheric variables*. Journal Geophys Res 113 :D24110. Doi :10.1029/2008JD01412.

### Abstract:

#### REGIONAL FREQUENCY ANALYSIS OF RAINFALL EXTREMES IN GIA LAI PROVINCE

*Currently in Vietnam, the rainfall frequency analysis for designing hydraulic works is mainly based on local approach. This approach use limited measured data from a gauge station and use frequency inference to estimate extreme rainfall which have time repeat from 100 to 1000 years. In this paper, the authors introduce a different approach which uses a regional approach and Bayesian inference methods to extend data sample and estimate the reliability of inference. The daily rainfall data of 26 rain gauge station in the Gia Lai province and its neighboring are used. The rainfall disadvantage framework commonly used in designing works are 1, 3, 5 and 7 maximum daily rainfall, will be determined. This approach includes 4 steps: Regional homogeneous test, homogeneous clusters, Distribution selection and regional frequency analysis.*

**Keywords:** regional frequency analysis, Bayesian inference, credibility, rainfall extremes, Gia Lai province.

---

*BBT nhận bài: 14/2/2017*

*Phản biện xong: 21/3/2017*