

Đánh giá kết quả tính lũ thiết kế bằng các đường phân phối tần suất khác nhau

Lương Tuấn Anh^{1,*}, Nguyễn Thanh Sơn², Đặng Thu Hiền¹

¹Viện KHKTTV và BDKH, 62 Nguyễn Chí Thanh, Hà Nội, Việt Nam

²Trường Đại học Khoa học Tự nhiên, ĐHQGHN, 334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam

Nhận ngày 26 tháng 6 năm 2015

Chỉnh sửa ngày 28 tháng 7 năm 2015; Chấp nhận đăng ngày 6 tháng 8 năm 2015

Tóm tắt: Bài báo trình bày kết quả nghiên cứu áp dụng các đường phân phối Log-chuẩn, Gamma II, Pearson III, Log-Pearson III và Gumbel, đồng thời áp dụng các phương pháp xác định thông số phân phối khác nhau như phương pháp mô men, Khả năng hiện thực lớn nhất (Maximum Likelihood), phương pháp đồ thị để tính lưu lượng lũ thiết kế tại một số trạm trên Hồng có chuỗi quan trắc đủ dài. Tập hợp, so sánh các kết quả tính toán đã đưa ra một số đánh giá mức độ tin cậy của các kết quả tính toán tần suất lũ thiết kế.

Từ khóa: Lũ thiết kế, đường phân phối tần suất, thông số phân phối tần suất.

1. Đặt vấn đề

Một trong số các nhiệm vụ của thủy văn ứng dụng là xác định các đặc trưng thiết kế thủy văn phục vụ quy hoạch, quản lý tài nguyên nước nhằm mục đích phòng lũ, cấp, thoát nước, phục vụ cho các nhu cầu dân dụng, công nghiệp, nông nghiệp, giao thông và đô thị, Độ lớn của các biến số dùng trong thiết kế như lưu lượng lũ, mực nước lũ, lưu lượng dòng chảy cạn, ... được lựa chọn đều liên quan đến chi phí và an toàn công trình. Trị số tối ưu dùng cho thiết kế là trị số được lựa chọn sao cho sự hài hòa giữa hai mặt đối lập nhau là chi phí và độ an toàn của công trình được đảm bảo [1]. Tiêu chuẩn thiết kế của các loại công trình thủy được

xây dựng và được áp dụng cho mỗi nước dựa trên cơ sở quy mô, tầm quan trọng của công trình và thời kỳ lặp lại.

Độ lớn của các đặc trưng thiết kế thủy văn phụ thuộc vào việc lựa chọn đường cong phân phối tần suất, phương pháp xác định các thông số của đường phân phối tần suất thống kê.

Việc lựa chọn đường cong phân phối tần suất luôn là đề tài thảo luận của nhiều cuộc Hội thảo trong và ngoài nước [2] và mỗi nước đều có các đề xuất sử dụng một đường cong phân phối tần suất nào đó. Hội đồng Tài nguyên nước Hoa Kỳ đã đề nghị chọn đường cong Log-Pearson III như một phân phối cơ bản trong các nghiên cứu về tần suất lũ [1]. Ở Nga, phân phối Krixky-Menkel và Pearson III thường được áp dụng trong phân tích tần suất của các đặc trưng thủy văn. Đối với các nước Châu Âu, Nhật Bản,

* Tác giả liên hệ. ĐT.: 84-917357025
Email: lanh@imh.ac.vn

đường cong Gumbel hay được sử dụng để tính toán các đặc trưng mưa-lũ lớn nhất. Ở nước ta, các đặc trưng lũ thiết kế thông thường được tính theo đường cong phân phối xác suất Pearson III. Tuy vậy, chưa có một kiến nghị nào định hướng cho việc tính toán các đặc trưng thiết kế thủy văn và do đó các kết quả tính toán các đặc trưng thiết kế của các cơ sở nghiên cứu khác nhau thường thiếu sự thống nhất, mặc dù sử dụng cùng một chuỗi số liệu đầu vào. Do đó, việc đánh giá kết quả tính dòng chảy lũ thiết kế bằng các đường cong phân phối xác suất khác nhau là cần thiết, đặt vấn đề cho việc nghiên cứu định hướng, lựa chọn các đường cong phân phối xác suất, phương pháp xác định các thông số của phân phối.

2. Một số đường cong phân phối xác suất và phương pháp xác định thông số phân phối.

Giả thiết x là một đại lượng ngẫu nhiên tuân theo quy luật phân phối xác suất $p(x)$, nghĩa là xác suất xuất hiện trị số x nằm trong khoảng $(x, x+dx)$ là $p(x)dx$. Xác suất $P(x_0)$ sao cho các trị số x không vượt quá x_0 được xác định theo công thức:

$$P(x_0) = \int_0^{x_0} p(x)dx \quad (1)$$

Xác suất $P^1(x_0)$ sao cho trị số x vượt quá x_0 được xác định như sau:

$$P^1(x_0) = 1 - \int_0^{x_0} p(x)dx$$

2.1. Đường Log-chuẩn II:

$$P(x) = \frac{1}{x\delta_y\sqrt{2\pi}} \exp\left(-\frac{(\ln x - M_{oy}^1)^2}{2\delta_y^2}\right) \quad (2)$$

- $y = \ln(x)$
- Đường cong này thích hợp với trị số C_s của chuỗi y gần bằng 0

2.2. Đường Log-chuẩn III:

$$p(x) = \frac{1}{(x-A)\delta_y\sqrt{2\pi}} \exp\left(-\frac{(\ln(x-A) - M_{oy}^1)^2}{2\delta_y^2}\right) \quad (3)$$

- $y = \ln(x-A)$; A là một thông số cận dưới;
- Đường cong này thích hợp với trị số C_s của chuỗi gần bằng $\ln(y-A)$.

2.3. Đường Gamma II

$$p(x) = \frac{1}{\alpha\Gamma(\beta)} \left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left(-\frac{x}{\alpha}\right) \quad (4)$$

2.4. Đường Pearson III

$$p(x) = \frac{1}{\alpha\Gamma(\beta)} \left(\frac{x-\gamma}{\alpha}\right)^{\beta-1} \exp\left(-\frac{x-\gamma}{\alpha}\right) \quad (5)$$

Đường cong này thích hợp với điều kiện: $0 \leq C_s \leq 2$, [3];

Đường cong này trong một số trường hợp không có nghiệm hiện thực lớn nhất (maximum likelihood solution) [4].

2.5. Đường Log-Pearson III:

$$p(x) = \frac{1}{\alpha\Gamma(\beta)} \left(\frac{\ln x - \gamma}{\alpha}\right)^{\beta-1} \exp\left(-\frac{\ln x - \gamma}{\alpha}\right) \quad (6)$$

Tương tự đường phân phối xác suất Pearson III, đường cong Log-Pearson III thích hợp với điều kiện $0 \leq C_s \leq 2$;

2.6. Đường Gumbel:

$$p(x) = a \exp(-\alpha(x-\beta) - \exp(-\alpha(x-\beta))) \quad (7)$$

- Đường cong này có 2 thông số;
- Đường cong này thích hợp với trị số Cs trong khoảng 1.14 [3].

2.7. Các phương pháp xác định thông số phân phối:

(a) Phương pháp mô men:

Các thông số phân phối (2-5) có liên hệ với mô men xung quanh gốc tọa độ và mô men trung tâm bậc r:

$$M_0^r = \int_{-\infty}^{\infty} x^r p(x) dx \tag{8}$$

$$M_c^r = \int_{-\infty}^{\infty} (x - M_0^1)^r p(x) dx \tag{9}$$

(b) Phương pháp khả năng hiện thực lớn nhất

Phương pháp này dựa trên cơ sở kỹ thuật tối ưu hóa các thông số phân phối α, β, γ sao cho hàm số khả năng hiện thực LL (loga của tích các xác suất hoặc tổng loga của hàm mật độ $p(x_i, \alpha, \beta, \gamma)$ đạt cực trị:

$$LL = \ln \prod_{i=1}^n p(x_i, \alpha, \beta, \gamma) \tag{10}$$

hoặc:

$$LL = \sum_{i=1}^n \ln p(x_i, \alpha, \beta, \gamma) \tag{11}$$

Các tham số α, β, γ được xác định bằng việc tính đạo hàm riêng của hàm LL theo lần lượt các thông số α, β, γ và nghiệm cần tìm tại điểm các thông số này làm các đạo hàm riêng của hàm LL bằng 0.

(c) Phương pháp đồ thị

Phương pháp này dựa trên cơ sở so sánh đường phân phối tần suất thực nghiệm và

đường phân phối lý thuyết là phù hợp nhất trên cùng một hệ tọa độ.

3. Áp dụng các đường cong phân phối xác suất khác nhau để tính lũ thiết kế cho một số trạm đo lưu lượng trên sông Đà, sông Hồng

3.1. Công thức xác định độ lớn của đặc trưng thiết kế

Công thức tổng quát xác định độ lớn của các đặc trưng thiết kế bằng phương pháp mô men có dạng:

$$X_T = M_0^1 + K_T \sqrt{M_c^2} \tag{12}$$

Trong đó: - X_T = là độ lớn của đặc trưng thiết kế;

- K_T = hệ số phụ thuộc vào tần suất được xác định theo từng dạng phân phối khác nhau, đối với phân phối 3 tham số, hệ số này còn phụ thuộc vào mô men trung tâm bậc 3 (hệ số bất đối xứng);

- M_0^1 = mô men quanh gốc tọa độ bậc 1 (giá trị trung bình);

- M_c^2 = mô men trung tâm bậc 2 (phương sai).

- Đối với phương pháp khả năng hiện thực lớn nhất, trị số X_T được xác định trực tiếp thông qua các tham số phân bố.

3.2. Khoảng tin cậy của các đặc trưng tính toán

$$C.I. = X_T \pm t.S_T$$

Trong đó:

C.I. = khoảng tin cậy của các đặc trưng tính toán

t = Chỉ số Students phụ thuộc vào mức bảo đảm α cho trước;

S_T = chuẩn sai của đặc trưng tính toán với điều kiện bảo đảm T.

3.3. Đánh giá mức độ phù hợp của các đường cong phân phối:

Là mức độ phù hợp giữa kết quả tính đường tần suất lý thuyết và đường tần suất thực nghiệm, đường tần suất thực nghiệm lại phụ thuộc vào công thức thực nghiệm tính tần suất. Trong thực tế tính toán các đặc trưng thủy văn, công thức Weibull-Gumbel hoặc Chegodayev thường được lựa chọn.

3.4. Số liệu sử dụng để tính toán:

- Trạm thủy văn Hòa Bình, sông Đà: lưu lượng lũ lớn nhất thời kỳ quan trắc 1904-1983 (

năm 1983 là năm hồ Hòa Bình bắt đầu ngăn dòng).

- Trạm thủy văn Sơn Tây, sông Hồng: lưu lượng lũ lớn nhất thời kỳ quan trắc 1902-1983.

3.5. Các kết quả tính toán

Kết quả tính toán lưu lượng lũ thiết kế ứng với các tần suất lặp lại (số năm) sử dụng phân phối Log- chuẩn 2 thông số (Log-N-II), log- chuẩn 3 thông số (Log-N-III); Gama 2 thông số (Gama-II); Pearson-III (P-III) và log-Pearson III(L-P-III) được thể hiện ở bảng 1 và 2.

Bảng 1. Kết quả tính lưu lượng thiết kế (m^3/s) trạm thủy văn Hòa Bình

PR (Year)	Log-N-II		Log-N-III		Gamma II		P-III		L-P-III		Gumbel	
Method*	M	ML	M	ML	M	ML	M	ML	G	G	ML	G
2	9033	9033	8870	NS	9162	9213	8710	8987	8671	8804	8983	8971
5	11729	11729	11519	NS	11873	11685	11465	11594	11933	12076	11647	11362
10	13445	13445	13355	NS	13482	13137	13431	13271	14177	14371	13410	12946
50	17087	17087	17592	NS	16645	15968	17903	16791	19438	19496	17292	16431
100	18596	18596	19437	NS	17870	17075	19825	18223	21051	21205	18933	17904

Bảng 2. Kết quả tính lưu lượng thiết kế (m^3/s) trạm thủy văn Sơn Tây

PR (Year)	Log-N-II		Log-N-III		Gamma II		P-III		L-P-III		Gumbel	
Method*	M	ML	M	ML	M	ML	M	ML	G	G	ML	G
2	16082	16082	15594	NS	16269	16352	15239	NS	15161	15599	15933	15938
5	20276	20276	19538	NS	20461	20101	19385	NS	19768	19594	20076	19514
10	22887	22887	22533	NS	22914	22271	22628	NS	23201	23155	22818	21881
50	28309	28309	29803	NS	27680	26448	30500	NS	31955	31729	28855	27091
100	30516	30516	33203	NS	29509	28038	34029	NS	34673	34752	31406	29293

*Chú giải: M- ký hiệu tính theo phương pháp mô men; ML-phương pháp khả năng hiện thực lớn nhất; G-phương pháp đồ thị; NS-không có nghiệm khả năng hiện thực lớn nhất.

4. Đánh giá và Nhận xét các kết quả tính toán:

(i) Đối với các mức tần suất thấp, các kết quả tính toán lưu lượng lũ lớn nhất theo các đường cong khác nhau và phương pháp xác định các thông số khác nhau có sự khác biệt không đáng kể;

(ii) Đường phân phối xác suất Pearson III và Log-Pearson III cho kết quả thiên lớn so với các đường cong khác ở vùng có tần suất lặp lại hiếm;

(iii) Đối với chuỗi số liệu quan trắc dài như ở các trạm thủy văn Hòa Bình và Sơn Tây, phương pháp đồ thị có sự phù hợp tốt với

phương pháp mô men và phương pháp khả năng hiện thực lớn nhất.

(iv) Các đường cong Log-chuẩn II, Gamma II và đường Gumbel cho kết quả gần giống nhau;

(v) Các đường cong Log-chuẩn III, Pearson III và Log-Pearson III cho kết quả tính toán tương đương;

(vi) Các kết quả tính theo phương pháp khả năng hiện thực nhất là phương pháp có cơ sở lý thuyết chặt chẽ nhất, có kết quả thiên nhỏ so với các phương pháp xác định thông số khác của các dạng phân phối xác suất.

(vii) Như đã trình bày ở phần lý thuyết, việc lựa chọn đường cong khác nhau phân phối xác suất nào trước tiên ước tính hệ số bất đối xứng C_s của chuỗi số liệu quan trắc, sau đó xét các trường hợp sau:

- Nếu $0 \leq C_s \leq 2$ đường cong Pearson III hoặc Log-Pearson III có thể được lựa chọn để tính các đặc trưng thiết kế;

- Nếu C_s gần bằng 1.13 chọn đường phân phối Gumbel;

- Nếu C_s của chuỗi $\ln(x)$ gần bằng 0 chọn đường Log-chuẩn II.

Kết luận, kiến nghị:

Từ kết quả nghiên cứu có thể rút ra một số kết luận và kiến nghị sau:

a. Đối với các công trình có quy mô lớn, có trị số $C_s > 0$, đề nghị sử dụng đường phân phối Pearson III hoặc log-Pearson III để bảo đảm an toàn cho công trình;

b. Nên sử dụng một số dạng phân phối cùng nhau để so sánh và đánh giá độ tin cậy;

c. Đối với chuỗi quan trắc ngắn, các thông số phân phối cần được xác định theo phương pháp khả năng hiện thực lớn nhất hoặc phương pháp đồ thị.

Tài liệu tham khảo

- [1] Ven Te Chow, David R. Maidment, Larry W. Mays (1988): Applied Hydrology, Mc Graw Hill Book.
- [2] Proceeding of international Symposium on floods and Droughts. IHP-V/ Technical Documents in Hydrology/ No.4. UNESCO Jakarta Office, 1999.
- [3] W. Buydens (1987): HYFA a Fortran 77 program for hydrological frequency analysis. I.H.E, Delft.
- [4] G.W. Kite (1976): Frequency and Risk Analyses in Hydrology, Wat. Res. Br., Inland Water Directorate, Dept. Environment, Ottawa.

Assessing the Results of Designed Flood Estimation by Using Different Types of Frequency Distributions

Lương Tuấn Anh¹, Nguyễn Thanh Sơn², Đặng Thu Hiền¹

¹Institute of Meteorology, Hydrology and Climate Change, 62 Nguyễn Chí Thanh, Hanoi, Vietnam

²VNU University of Science, 334 Nguyễn Trãi, Thanh Xuân, Hanoi, Vietnam

Abstract: The article presents the results of applying Log-normal, Gamma II, Pearson III, Log-Pearson III and Gumbel Distributions, and the methods to estimate parameter of the distribution such as Moments method, Maximum Likelihood and Graphical method for estimating the designed flood for long-term observed data series of some stations on the Hồng river. Analysing the results of computation, the paper aims to assess the liability of computed results.

Keywords: Frequency distribution, parameters of distribution, designed flood.