

# **Khả năng nâng cao chất lượng phương pháp thống kê khách quan trong dự báo đinh lũ năm**

PTS. BÙI VĂN ĐỨC, KS. ĐỖ QUANG HUYỀN

Cục Dự báo KTTV

## **MỞ ĐẦU**

Thống kê khách quan (TKKQ) hiện đang là một trong những phương pháp thông dụng trong dự báo đinh lũ năm trên các sông ngòi Việt Nam. Cơ sở vật lý của phương pháp dựa trên giả thiết : "Thời tiết mùa đông và mùa xuân chứa đựng nhiều thông tin về dòng chảy của mùa lũ tiếp theo". Đinh lũ năm được dự báo tổng hợp từ các "bản đồ thống kê" thành phần [2,3]. Độ tin cậy của phương pháp phụ thuộc vào nhiều điều kiện khác nhau, như chọn trạm khí tượng đại biểu, khoảng thời gian thu thập số liệu dự báo, phương pháp lọc tuyển và tổ hợp các nhân tố dự báo v.v.

Nhiều năm qua, trong dự báo thủy văn hạn dài, TKKQ được xử lý bằng phương pháp đồ giải, còn mang nhiều hạn chế, không tiện cập nhật số liệu hàng năm và không có điều kiện xác định cấu trúc tối ưu cho sơ đồ dự báo (với suất đảm bảo cao nhất và "gọn nhẹ nhất").

Trong điều kiện hiện nay, với ưu thế về tốc độ tính nhanh của máy vi tính chúng ta có thể nhận được sơ đồ dự báo có cấu trúc tối ưu một cách dễ dàng.

Kết quả sơ tuyển các nhân tố dự báo trong một số nghiên cứu gần đây [2,3], bài viết tập trung vào nội dung chính là mô phỏng các tần đồ thống kê khách quan trên máy vi tính và phương pháp tìm kiếm cấu trúc tối ưu cho sơ đồ dự báo. Kết quả nghiên cứu được kiểm nghiệm và minh họa trên tập số liệu sông Cà - Trạm Nam Đàm.

### **1. Nhân tố dự báo**

Từ những nhân tố đặc trưng khí tượng của một hay nhiều trạm khí tượng đại biểu, trong khoảng thời gian đã chọn của mùa đông, như mưa, nhiệt, ẩm, áp v.v. (các giá trị trung bình và cực trị), qua sơ lọc thông tin theo các chỉ tiêu thống kê [2], đã chọn được tập nhân tố ảnh hưởng ( $X_1, X_2, \dots, X_m$ ) để dự báo yếu tố ( $Y$ ), biểu diễn dưới dạng ma trận sau:

$$\begin{matrix} X_{1,1} & X_{1,2} & \dots & X_{1,m} & Y_1 \\ X_{2,1} & X_{2,2} & \dots & X_{2,m} & Y_2 \\ \vdots & \vdots & & \vdots & \vdots \\ X_{n,1} & X_{n,2} & \dots & X_{n,m} & Y_n \end{matrix} \quad (1)$$

Trong đó:

$m$  - số nhân tố,

$n$  - độ dài dãy quan trắc

$Y$  - yếu tố dự báo ( $H_{max}$ , hoặc  $Q_{max}$ ).

Yếu tố dự báo không mang giá trị thực mà được gán các giá trị cấp lũ (5 cấp) theo quy phạm hiện hành của Cục Dự báo KTTV [1].

- |               |                                    |
|---------------|------------------------------------|
| 1- Lũ rất nhò | $H_{max} \leq H_1 = 0,7.H_o$       |
| 2- Lũ nhò     | $H_1 < H_{max} \leq H_2 = 0,9.H_o$ |
| 3- Lũ vừa     | $H_2 < H_{max} \leq H_3 = 1,1.H_o$ |
| 4- Lũ lớn     | $H_3 < H_{max} \leq H_4 = 1,3.H_o$ |
| 5- Lũ rất lớn | $H_4 < H_{max}$                    |

Trong đó  $H_o$  là trị số trung bình của chuỗi  $H_{max}$ .

## 2. Nội dung phương pháp

Từ tập số liệu (1) tiến hành tổ hợp chéo các cặp nhân tố  $X_i, X_j$ . Mỗi cặp nhân tố trên tạo ra một không gian hai chiều, được xác định bởi hai trục số  $X_i$  và  $X_j$  cắt nhau vuông góc tại giá trị trung bình của chúng, chia không gian hai chiều thành 4 miền. Mỗi miền của không gian được mô phỏng bằng một hệ bất phương trình tương ứng:

		Xi				
		I	II			
$X_j$	$\begin{cases} X_{it} < X_i \\ X_{jt} < X_j \end{cases}$	$\begin{cases} X_{it} < X_i \\ X_{jt} \geq X_j \end{cases}$	$\begin{cases} X_{it} \leq X_i \\ X_{jt} \geq X_j \end{cases}$			
		$\begin{cases} X_{it} \leq X_i \\ X_{jt} < X_j \end{cases}$	$\begin{cases} X_{it} \geq X_i \\ X_{jt} < X_j \end{cases}$			
		IV	III			
$X_i$						

Với m nhân tố ta có M không gian 2 chiều ( $M = \frac{1}{2} (m^2 + m)$ ).

Số lần xuất hiện các cấp lũ trong mỗi miền của mỗi không gian được lưu trữ trong ma trận ba chiều U. Ma trận U mang các đặc tính sau:

- Đối với một tán đồ k bất kỳ,

$$\sum_{l=1}^4 \sum_{h=1}^5 U_{l,h} = n \quad (3)$$

- Đối với tổ hợp chập  $M'$  tán đồ thành phần,

$$\sum_{l=1}^4 \sum_{h=1}^5 \sum_{k=1}^M U_{l,h,k} = n.M' \quad (4)$$

Trong đó l - chỉ số miền ( $l = 1 \div 4$ ),

h - chỉ số cấp lũ ( $h = 1 \div 5$ ),

k - chỉ số không gian ("tán đồ thống kê khách quan") thành phần,,

$M'$  - số không gian 2 chiều của mỗi tổ hợp chập bất kỳ.

Phương pháp thống kê khán quan coi tổng số lần xuất hiện của từng cấp lũ trên từng miền của các không gian 2 chiều, là khả năng xuất hiện cấp lũ tương ứng trong tương lai. Khả năng trung bình ( $\bar{Z}_h$ ) xuất hiện cấp lũ ( $h$ ) của  $M'$  không gian trong  $n$  năm được xác định bởi biểu thức (5). Khả năng xuất hiện cấp lũ ( $h$ ) của năm bất kỳ của  $M'$  không gian được xác định bằng độ lệch  $\Delta Z_h$ , t, biểu thức (6).

$$\bar{Z}_h = \frac{1}{n} \sum_{k=1}^{M'} \sum_{t=1}^n U L_t, k, h, \quad (5)$$

$$\Delta Z_h = \text{Max} [ Z_h, t - \bar{Z}_h ] \quad (6)$$

Dự báo được coi là đúng khi cấp lũ dự báo trùng với cấp lũ thực tế xảy ra. Mức bảo đảm ( $P\%$ ) của số đồ dự báo là tỷ số phần trăm số lần dự báo đúng ( $n'$ ) trên tổng số lần dự báo ( $n$ ).

$$P = \frac{n'}{n} 100\%. \quad (7)$$

Trên cơ sở các biểu thức (5), (6) và (7), tiến hành tổ hợp chập các không gian hai chiều để xác định chuỗi suất đảm bảo lớn nhất tương ứng [ $P^{M'} \text{ max}]_M$ .

Số đồ dự báo có cấu trúc tối ưu là số đồ có suất đảm bảo cao nhất và có số không gian ("tần đồ dự báo") ít nhất.

$$X(X_1, X_2, \dots, X_m) \sim \begin{cases} \text{Max} [ P^{M'} \text{ max}]_M \\ \text{Min} [ M' ] \end{cases} \quad (8)$$

## 2. Kết quả ứng dụng cho sông Cà - Nam Đàm

### 2.1. Chuẩn bị số liệu

Yếu tố dự báo là mực nước lớn nhất năm ( $H_{max}$ ) trại trạm Nam Đàm - S. Cà. Theo quy phạm hiện hành về dự báo thủy văn,  $H_{max}$  được chia thành 5 cấp. Sau khi gộp 2 cấp dưới thành một cấp nhỏ ta có các cấp lũ như sau:

Lũ nhỏ	$H_{max} \leq$	6,1m
Lũ vừa	$6,1 m <$	$H_{max} \leq$
Lũ lớn	$7,5 m <$	$H_{max} \leq$
Lũ rất lớn	$8,8 m <$	$H_{max}$

Đối với s. Cà - Nam Đàm, trạm khí tượng đại biểu được chọn là Vinh - Nghệ Tĩnh, khoảng thời gian dùng thu thập thông tin dự báo là các tháng XII năm trước, tháng I và II năm dự báo. Các nhân tố được chọn là:

- $X_1$  - tổng lượng mưa, mm
- $X_2$  - tổng số ngày mưa, ngày
- $X_3$  - tổng lượng mưa bốc hơi, mm
- $X_4$  - độ ẩm tương đối bình quân, %
- $X_5$  - nhiệt độ không khí bình quân,  $^{\circ}\text{C}$

$X_6$  - nhiệt độ không khí tối thấp tuyệt đối,  $^{\circ}\text{C}$

$X_7$  - nhiệt độ không khí tối cao tuyệt đối,  $^{\circ}\text{C}$

$X_8$  - khí áp bình quân, mb

$X_9$  - khí áp tối thấp tuyệt đối, mb

$X_{10}$  - khí áp tối cao tuyệt đối, mb

Các chuỗi số liệu khí tượng thủy văn đồng bộ và liên tục từ năm 1957 đến 1989 với độ dài 32 năm. Từ tập nhân tố này đã tiến hành tổ hợp chéo và xác định suất đảm bảo của từng tán đồ thống kê thành phần. Kết quả tính toán được ghi trong biểu 1.

Bảng 1- Suất bảo đảm của các tán đồ thành phần, %

$X_i \setminus X_j$	1	2	3	4	5	6	7	8	9	10
1	39									
2	39	42								
3	55	35	48							
4	42	42	42	55						
5	42	32	36	55	32					
6	39	45	42	55	45	26				
7	48	52	55	45	48	42	52			
8	48	39	45	35	42	42	52	48		
9	42	45	45	52	58	45	52	55	52	
10	42	35	52	32	45	48	42	58*	52	29

\* - giá trị lớn nhất.

Tổng số tán đồ thành phần được xác định theo công thức sau:

$$M = \frac{1}{2} (m^2 + m). \quad (9)$$

Đối với s. Cà - Nam Đàm, với 10 nhân tố được sơ tuyển đã tổ hợp chéo được 55 tán đồ thành phần khác nhau.

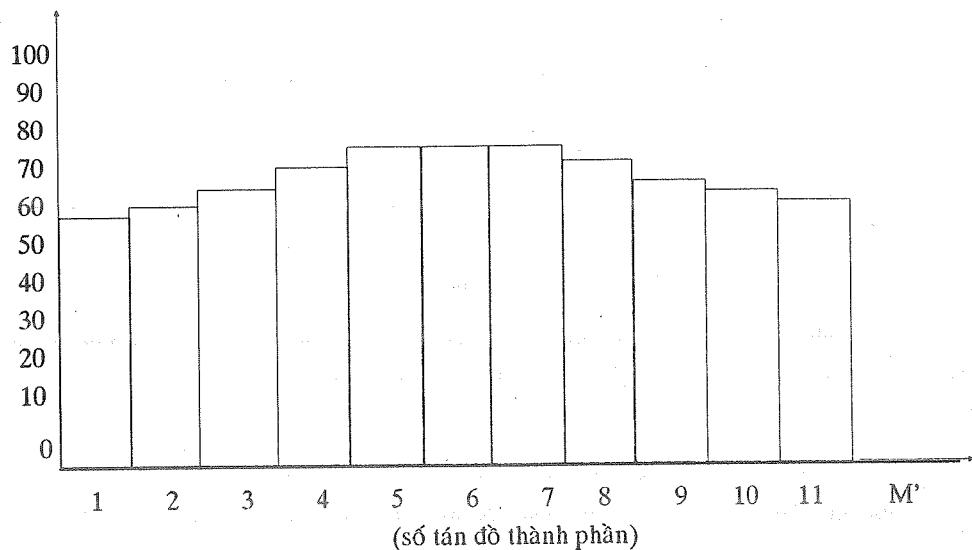
## 2.2. Tìm cấu trúc tối ưu

Từ 55 tán đồ khác nhau nói trên tiến hành tính toán thử dần cho mỗi tổ hợp chập có thể xảy ra và xác định suất đảm bảo của chúng. Từ mỗi loại tổ hợp chập bậc  $M'$  cho phép chọn và giữ lại một tổ hợp có suất đảm bảo lớn nhất. Ví dụ, tổ hợp chéo ở bảng 1 được coi là tổ hợp chập một, và tán đồ được tạo ra từ nhân tố thứ 8 và 10 có suất đảm bảo lớn nhất ( $[P_{\max}]_{55} = 58\%$ ). Tiếp tục thực hiện tổ hợp chập hai, chập ba... và chọn ra chuỗi suất đảm bảo cao nhất tương ứng  $[P^{M'}_{\max}]_{55}$ . Để thuận lợi cho xác định cấu trúc tối ưu, đã tiến hành xây dựng biểu đồ biến thiên  $[P^{M'}_{\max}]$  theo  $M'$  (hình 1). Nhìn vào biến trình đó dễ dàng nhận thấy rằng cấu trúc tối ưu nhất chứa 5 tán đồ thành phần và có suất đảm bảo bằng 74%. Các cặp nhân tố thành phần của sơ đồ tối ưu được ghi trong bảng 2.

Bảng 2- Thành phần, cấu trúc sơ đồ dự báo tối ưu

Tán đồ thành phần	I	II	III	IV	V
i	2	2	1	2	10
j	5	9	7	7	5

$P^{M'}_{\max}$ , %



Hình 1- Biến trình suất dầm bảo lớn nhất ( $P^{M'}_{\max}$  - đã được chọn trong từng loại tổ hợp chập bậc  $M'$  từ 55 tán đồ thành phần)

### KẾT LUẬN

Việc mô phỏng các tán đồ thống kê khách quan trên máy vi tính mở ra một khả năng mới trong ứng dụng phương pháp TKKQ và dự báo đỉnh lũ năm cho các sông ngòi ở Việt Nam.

Sơ đồ dự báo đỉnh lũ năm cho sông Cà Lồ tại Nam Đàm mới chỉ tối ưu trên tập số nhân tố đã được lựa chọn. Muốn nâng tiếp chất lượng sơ đồ dự báo cần nghiên cứu đưa thêm các nhân tố mới, như các đặc trưng hoàn lưu khí quyển trong mùa đông vào tập nhân tố.

Sơ đồ mô phỏng và tìm kiếm cấu trúc tối ưu trên có thể ứng dụng cho các sông khác, góp phần nâng cao chất lượng dự báo đỉnh lũ năm trên các sông ngòi nước ta, một vấn đề đang được nhiều người quan tâm.

## TÀI LIỆU THAM KHẢO

1. Cục Dự báo KTTV. Quy phạm hiện hành về dự báo thủy văn.
  2. Nguyễn Lan Châu. Đề tài dự báo hạn đài định lũ năm trên lưu vực sông Hồng. Tổng cục KTTV, 1985
  3. Đỗ Quang Huyên. Dự báo định lũ năm trên lưu vực sông Cà Nam Dàn bằng phương pháp thống kê khách quan. Đề tài nghiên cứu Cục Dự báo KTTV, 1988.
- 

## Phân loại chất gây ô nhiễm nước

Chất gây ô nhiễm nước có thể chia làm 8 loại:

1. Chất thải cần ô-xi: là hợp chất hữu cơ bị giảm cấp về sinh học, chứa trong nước thải gia đình hoặc trong dòng chảy công nghiệp. Khi những hợp chất đó bị phân hủy bởi vi khuẩn, ô-xi bị loại khỏi nước. Nếu mức ô-xi bị giảm quá thấp thì cá chết.
2. Tác nhân gây bệnh: là những vi sinh vật gây bệnh khác nhau, thường xâm nhập vào nước cùng với nước thải của người. Tiếp xúc với những vi khuẩn đó do uống nước hoặc các hoạt động tiếp xúc với nước.
3. Hợp chất hữu cơ tổng hợp: bao gồm chất tẩy rửa, thuốc trừ sâu và các hóa chất công nghiệp tổng hợp khác. Nhiều hợp chất đó là độc đối với thủy sinh và có thể có hại đối với con người.
4. Chất dinh dưỡng cho cây: như ni-tơ và phốt-pho trôi từ đất bón phân cũng như dòng chảy từ nhà máy xử lý nước thải, thúc đẩy sự phát triển của tảo và cỏ nước.
5. Hóa chất vô cơ và chất khoáng: bao gồm các a-xít được hình thành khi tiêu nước từ các mỏ cũng như kim loại nặng như thủy ngân và cat-mi-um.
6. Trầm tích: là những hạt đất, cát và khoáng vật bị rửa trôi từ đất. Chúng có thể bóp chết đối sống ở đáy cũng như bồi lấp hồ chứa và bến cảng. Quản lý đất không đúng sẽ dẫn đến xói mòn và là nguyên nhân chính gây ra trầm tích.
7. Chất phóng xạ: có thể xâm nhập vào nước do khai mỏ và xử lý quặng phóng xạ, từ hoạt động năng lượng hạt nhân, từ thiết bị y tế và từ thử vũ khí hạt nhân. Tính phóng xạ tập trung trong chuỗi thực phẩm cũng như DDT.
8. Sự tiết nhiệt từ các nhà máy nhiệt điện làm tăng nhiệt độ của nước  $20^{\circ}\text{F}$  (khoảng  $11^{\circ}\text{C}$ ), gây ra những biến đổi khác nhau trong hệ sinh thái địa phương.

Ô nhiễm nước thường do tổ hợp của 8 nguồn kể trên và có thể gây ra vấn đề nghiêm trọng.

Phân loại chất gây ô nhiễm nước nói trên được sử dụng trong Cơ quan bảo vệ môi trường Hoa Kỳ.

Theo "Introduction to Environmental Science and Technology"