

Hàm tự tương quan và hàm mật độ phổ của chuỗi dòng chảy sông Hồng tại Sơn Tây

PTS. TRẦN THANH XUÂN

KS. DẶNG LAN HƯƠNG

Viện KTTV

Hàm tự tương quan và hàm phổ được sử dụng rộng rãi trong nghiên cứu tính biến đổi trong quá trình nhiều năm của các yếu tố KTTV nhằm phục vụ cho công tác tính toán và dự báo KTTV. Trong bài này chúng tôi xin giới thiệu một số kết quả nghiên cứu về hệ số tương quan và hàm mật độ phổ của chuỗi dòng chảy sông Hồng tại Sơn Tây [1].

1- Hàm tự tương quan

Hệ số tự tương quan giữa các thành phần của chuỗi số liệu được tính theo công thức sau[5]:

$$r_{\tau} = \frac{\sum_{i=1}^{n-\tau} (X_i - \bar{X})(X_{i+\tau} - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (1)$$

trong đó X_i - giá trị của thành phần thứ i , từ X_1 đến $X_{n-\tau}$,

$X_{i+\tau}$ - giá trị của thành phần thứ $i + \tau$, từ $X_{1+\tau}$ đến X_n ,

n - số thành phần (số năm quan trắc) của chuỗi,

τ - bước trượt (dịch chuyển, trễ) từ 1 đến M ,

M là bước trượt lớn nhất, lấy bằng số nguyên lớn nhất, bằng hoặc nhỏ hơn $(n-1)/6$.

r_{τ} - hệ số tự tương quan tương ứng với bước trượt thứ τ .

Sai số chuẩn của hệ số tự tương quan ($SE_{(r_{\tau})}$) được tính theo công thức dưới đây của Box và Jenkins (1970):

$$SE_{(r_{\tau})} = \left[\left(1 + 2 \sum_{i=1}^q r_i^2 \right) / n \right]^{1/2} \quad (2)$$

trong đó lấy $q = 1$ khi $\tau = 1$ và $q = 2$ khi $\tau \geq 2$.

Chuỗi số liệu dòng chảy của sông Hồng tại trạm thủy văn Sơn Tây dùng để tính toán dài 87 năm, từ năm 1902 đến năm 1988. Số liệu dòng chảy thực đo của trạm Sơn Tây trong các năm 1972-1988 đã được hiệu chỉnh để loại trừ ảnh hưởng điều tiết của hai hồ chứa Thác Bà trên sông Cháy (từ năm 1972) và Hòa Bình trên sông Đà (từ năm 1983).

Căn cứ vào chuỗi số liệu dòng chảy nói trên của trạm Sơn Tây đã tiến hành tính toán hệ số tự tương quan theo công thức (1) cho các chuỗi dòng chảy tháng và năm dưới đây:

- Chuỗi dòng chảy năm,
- Chuỗi dòng chảy tháng của từng tháng trong năm,
- Chuỗi dòng chảy tháng lớn nhất,
- Chuỗi dòng chảy tháng nhỏ nhất,
- Chuỗi dòng chảy tháng do tất cả các tháng của toàn chuỗi tạo thành.

Trong bảng 1 đưa ra giá trị của hệ số tự tương quan tại bước trượt đầu tiên (r_1) và sai số chuẩn của nó.

Trong hình 1 đưa ra đồ thị hàm tự tương quan ($r_T \sim \tau$) của chuỗi dòng chảy năm và một số tháng đại biểu: tháng IV - tháng giữa mùa cạn, tháng V - tháng chuyển tiếp từ mùa cạn sang mùa lũ.

Bảng 1 - Hệ số tự tương quan tại bước trượt đầu tiên và sai số chuẩn tương ứng của các chuỗi dòng chảy sông Hồng tại Sơn Tây.

Tháng Đặc trưng	Tháng												Tháng lớn nhất	Tháng nhỏ nhất	Năm
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII			
r_1	0,036	0,229	0,176	0,047	-0,217	-0,103	-0,08	0,049	0,066	0,024	0,026	0,011	-0,017	0,378	0,006
SE (r_1)	0,107	0,113	0,110	0,107	0,112	0,108	0,108	0,107	0,108	0,107	0,107	0,107	0,107	0,112	0,107

Từ kết quả tính toán có thể nhận xét về hệ số tự tương quan của chuỗi dòng chảy sông Hồng tại Sơn Tây như sau:

a) Đối với chuỗi dòng chảy năm

Giá trị r_T tương đối nhỏ và dao động trong phạm vi khá rộng, từ - 0,024 (tương ứng với $\tau = 14$) đến 0,173 ($\tau = 5$); giá trị tuyệt đối của r_T dao động từ 0,006 đến 0,173. Phân tích quan hệ $r_T \sim \tau$ (hình 1) có thể thấy hệ số tự tương quan tại bước trượt đầu tiên rất nhỏ ($r_1 = 0,006$), tăng dần và đạt giá trị lớn nhất tại bước trượt thứ 5 ($r_5 = 0,173$); sau đó giảm dần và dao động trong phạm vi hẹp (- 0,011 ~ 0,080), từ bước trượt thứ 8 đến thứ 14 giá trị r_T thường nhỏ hơn 0.

b) Đối với chuỗi dòng chảy tháng

Cũng như chuỗi dòng chảy năm, giá trị r_T của các chuỗi dòng chảy tháng cũng đều tương đối nhỏ, dao động trong phạm vi từ -0,283 đến 0,378; giá trị tuyệt đối của r_T dao động trong phạm vi từ 0,002 đến 0,378. Giá trị r_1 đều tương đối nhỏ, chỉ có các tháng II, III, V, VI và dòng chảy tháng nhỏ nhất lớn hơn 0,10; trong đó chuỗi dòng chảy tháng nhỏ nhất có r_1 lớn nhất và r_1 của chuỗi dòng chảy các tháng V, VI và VII nhỏ hơn 0. Nói chung, r_T có thể nhỏ hơn 0 ở bất kỳ bước trượt nào, nhưng các chuỗi dòng chảy tháng trong mùa lũ thường có $r_1 < 0$.

Đồ thị hàm tự tương quan của các chuỗi dòng chảy tháng trong mùa cạn khác với đồ thị hàm tự tương quan của các tháng trong mùa lũ. Đối với các chuỗi dòng chảy tháng trong mùa cạn, giá trị

r_τ tương đối lớn ngay từ những bước trượt đầu tiên, sau đó giảm dần rồi dao động trong phạm vi hẹp. Đối với các chuỗi dòng chảy tháng trong mùa lũ, giá trị r_1 đều tương đối nhỏ và có khi nhỏ hơn 0; trên đồ thị $r_\tau \sim \tau$ có một số đỉnh. Đồ thị hàm tự tương quan của chuỗi dòng chảy tháng V tương đối đặc biệt so với đồ thị hàm tự tương quan của các chuỗi dòng chảy tháng khác. Đối với tháng này, tuy giá trị r_1 nhỏ hơn 0, nhưng có giá trị tuyệt đối tương đối lớn ($r_1 = -0,217$), sau đó tăng nhanh và đạt lớn nhất tại bước trượt thứ 3 ($r_3 = 0,378$), tiếp sau đó r_τ giảm dần và từ $\tau = 6$ đều có $r_\tau < 0$.

Đồ thị hàm tự tương quan của chuỗi dòng chảy tháng lớn nhất tương tự như đồ thị hàm tự tương quan của các tháng trong mùa lũ.

Đồ thị hàm tự tương quan của chuỗi dòng chảy tháng nhỏ nhất cũng tương tự như các chuỗi dòng chảy tháng trong mùa cạn; hệ số tự tương quan đạt giá trị lớn nhất ngay từ bước trượt đầu tiên ($r_1 = 0,376$), sau đó giảm dần khi τ tăng.

Đồ thị hàm tự tương quan của chuỗi dòng chảy tháng của tất cả các tháng trong năm có dạng hình sin với chu kỳ 12 tháng (Hình 2). Các đỉnh xuất hiện tại các bước trượt $\tau = 12, 24, 36, \dots$ và giá trị r_τ tại các đỉnh dao động trong phạm vi từ 0,715 ($\tau = 168$) đến 0,826 ($\tau = 24$), có xu thế giảm dần khi τ tăng; các chân xuất hiện tại các bước trượt $r = 6, 18, 30, \dots$ và cũng giảm dần khi τ tăng:

$$r_{18} = -0,617, r_{162} = -0,542.$$

Tóm lại, hệ số tự tương quan của chuỗi dòng chảy sông Hồng tại Sơn Tây tương đối nhỏ và dao động trong phạm vi khá rộng. Điều đó chứng tỏ mối liên quan nội tại của chuỗi không lớn và có thể coi chuỗi dòng chảy như là một quá trình ngẫu nhiên. Tính ngẫu nhiên của các chuỗi dòng chảy tháng trong mùa lũ cao hơn so với các chuỗi dòng chảy tháng trong mùa cạn. Nguyên nhân có thể là do nguồn cung cấp nước sông trong mùa lũ chủ yếu do mưa cung cấp, còn trong mùa cạn thì chủ yếu do lượng nước trữ trong lưu vực và nước ngầm cung cấp. Rõ ràng, tính ổn định của nước ngầm cao hơn so với lượng mưa. Do vậy, mối quan hệ nội tại trong chuỗi dòng chảy cạn cao hơn so với chuỗi dòng chảy lũ. Tuy vậy, cũng nhận thấy rằng mô hình chuỗi Markov đơn với quy luật biến đổi $r_\tau = r_1^\tau$ không phù hợp với đồ thị hàm tự tương quan của chuỗi dòng chảy sông Hồng.

Cũng cần chỉ ra rằng, tuy chuỗi số liệu dài tới 87 năm nhưng sai số chuẩn của r_τ khá lớn, nhiều khi bằng hoặc lớn hơn bản thân giá trị r_τ . Từ số liệu trong bảng 1 có thể thấy chỉ có giá trị r_1 của các chuỗi dòng chảy tháng II, III, V và chuỗi dòng chảy tháng nhỏ nhất lớn hơn giá trị sai số chuẩn, còn các chuỗi dòng chảy khác thì $SE(r_1) \geq r_1$.

2- Hàm mật độ phổ

Hàm mật độ phổ chuẩn hóa được hình thành theo công thức sau đây [5]:

$$S_{p,k} = (1/M) \left[1 + 2 \sum_{\tau=1}^{M-1} r_{\tau k} \cos(\pi p \tau / M + r_{MK} \cos(\pi p)) \right] \quad (3)$$

trong đó $P = 0, 1, 2, \dots, M$ và phép làm tròn như sau:

$$S_{0,k} = 0,5 (S_{0,k} + S_{1,k}) \quad (4)$$

$$S_{p,k} = 0,25 S_{p-1,k} + 0,5 S_{p,k} + 0,25 S_{p+1,k} \quad (5)$$

$$S_{M,k} = 0,5 (S_{M-1,k} + S_{M,k}) \quad (6)$$

tần số tương ứng với P được xác định như sau:

$$f = P/2M, \quad \text{vòng/năm.} \quad (7)$$

trong các công thức trên k là chỉ số thời đoạn tính ($k = 1, 2, \dots, m$); m là số thời đoạn tính toán; các ký hiệu τ và M như trong công thức (1).

Khoảng tin cậy tương ứng với mức bảo đảm 95% được tính như sau:

$$[\vartheta S_{pk}/\chi^2_{0,025}(\vartheta), \vartheta S_{pk}/\chi^2_{0,475}(\vartheta)]. \quad (8)$$

trong đó bậc tự do của phân bố χ^2 được tính như sau: $\vartheta = 2,667n/M$; n là số năm của mỗi thời đoạn tính toán.

Phân tích kết quả tính hàm mật độ phổ của chuỗi dòng chảy của sông Hồng tại Sơn Tây cho thấy:

- Hàm mật độ phổ của chuỗi dòng chảy năm (hình 2) tại bước trượt tương ứng với tần số $f = 0$ đạt giá trị lớn nhất (0,136), giảm nhanh đến bước trượt thứ 3 (0,0328), sau đó tăng dần tới bước trượt thứ 6 (0,076) rồi biến đổi trong phạm vi hẹp (0,068 ~ 0,081). Đỉnh phổ xuất hiện tại bước trượt thứ 6 (chu kỳ 4, 7 năm) không được rõ ràng. Như vậy, có thể coi hàm mật độ phổ của chuỗi dòng chảy năm như là hàm mật độ phổ của quá trình "nhiều đờ".

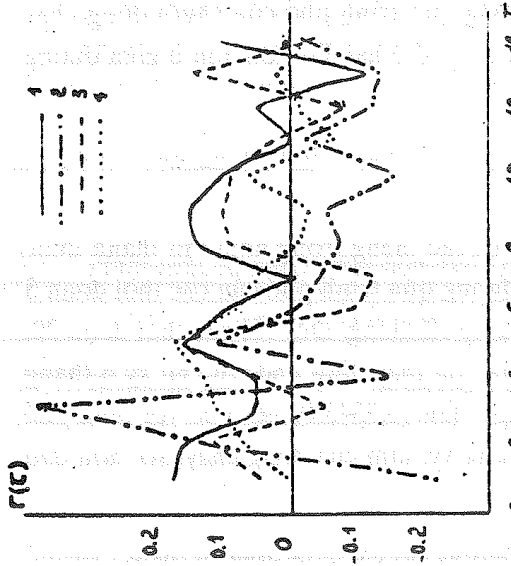
- Đồ thị hàm mật độ phổ của các chuỗi dòng chảy tháng trong mùa cạn có sự khác biệt so với hàm mật độ của chuỗi dòng chảy của các tháng trong mùa lũ (Hình 2).

Đối với chuỗi dòng chảy của các tháng trong mùa cạn, ngay từ bước trượt đầu tiên ($\tau = 0$) xuất hiện giá trị mật độ phổ lớn nhất, sau đó giảm dần khi τ tăng lên và dao động trong phạm vi hẹp. Đường quá trình phổ của các chuỗi dòng chảy tháng trong mùa lũ cũng không hoàn toàn giống nhau, nhưng thường nhận thấy có một vài đỉnh phổ với chu kỳ 2,5 - 5, 6 năm, cá biệt tới 14 năm. Đặc biệt, trên đồ thị hàm mật độ phổ của chuỗi dòng chảy tháng V có 1 đỉnh phổ cao nhất xuất hiện tại bước trượt thứ 10 ($S_{10} = 0,1533$) với chu kỳ 2, 8 năm. Đây là đỉnh phổ có giá trị lớn nhất so với các đỉnh phổ khác của các chuỗi dòng chảy tháng trong mùa cạn và trong mùa lũ. Cũng như đồ thị hàm tự tương quan, đồ thị hàm mật độ phổ của chuỗi dòng chảy tháng V tương đối đặc biệt và khác với đồ thị hàm mật độ phổ của các chuỗi dòng chảy khác. Đường quá trình phổ của chuỗi dòng chảy tháng lớn nhất không thể hiện rõ một đỉnh phổ nào và mật độ phổ ở hai đầu lớn hơn ở giữa đường quá trình.

Đường quá trình phổ của chuỗi dòng chảy tháng nhỏ nhất tương tự như đường quá trình phổ của chuỗi dòng chảy tháng trong mùa cạn.

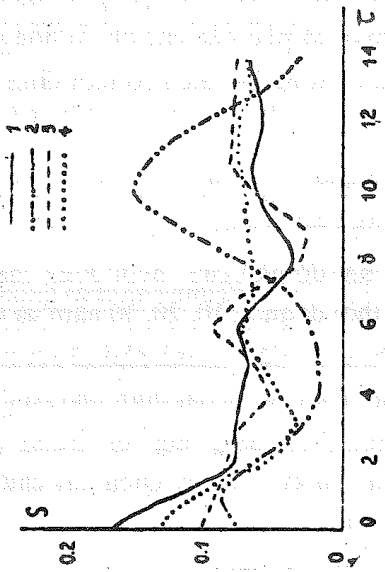
Hàm mật độ phổ của chuỗi dòng chảy tháng cho tất cả các tháng trong năm tạo thành được tính cho các thời đoạn 5, 10, 20, 30 năm và cả chuỗi. Trên đường quá trình phổ của các thời đoạn 5 và 10 năm chỉ xuất hiện 1 đỉnh phổ rõ rệt tương ứng với chu kỳ 12 tháng. Đối với các thời đoạn 20, 30 năm và cả chuỗi thì ngoài đỉnh phổ có chu kỳ 12 tháng ra còn thấy đỉnh phổ với chu kỳ 6 tháng (Hình 4). Đỉnh phổ tương ứng với chu kỳ 12 tháng lớn gấp 7 lần (0,3345/0,0467) so với đỉnh phổ tương ứng với chu kỳ 6 tháng. Điều này chứng tỏ dao động chu kỳ năm của dòng chảy trội hơn dao động chu kỳ mùa.

Tóm lại, chuỗi dòng chảy sông Hồng tại Sơn Tây dao động theo chu kỳ mùa (6 tháng) và chu kỳ năm (12 tháng). Đường quá trình phổ của phần lớn chuỗi dòng chảy tháng và chuỗi dòng chảy năm thường có 1 hay vài đỉnh phổ với chu kỳ 2 - 5 năm, có khi hơn 10 năm nhưng chỉ có đỉnh phổ của



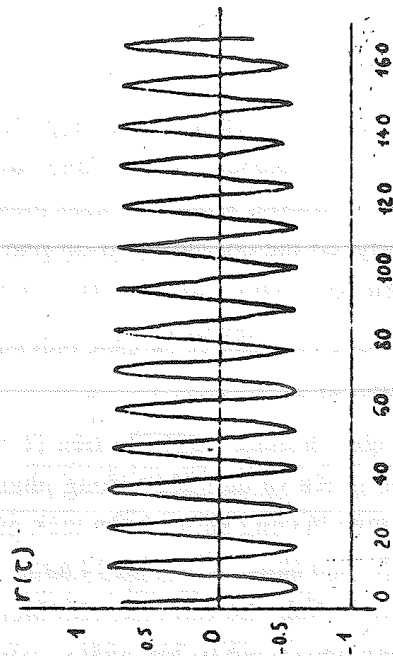
Hình 1 - Sơ đồ hàm tự tương quan dòng chảy sông Hồng tại Sơn Tây

1- Tháng III, 2- Tháng V, 3- Tháng VIII, 4- Năm

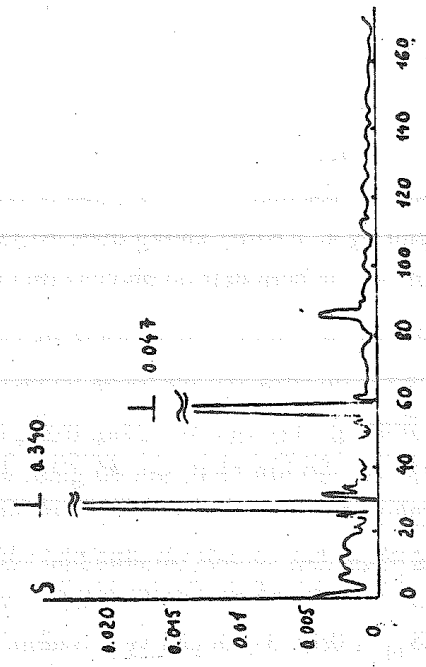


Hình 3 - Sơ đồ hàm mật độ phổ dòng chảy sông Hồng tại Sơn Tây

1- Tháng III, 2- Tháng V, 3- Tháng VIII, 4- Năm



Hình 2 - Đồ thị hàm tự tương quan dòng chảy tháng sông Hồng tại Sơn Tây



Hình 4 - Đồ thị hàm mật độ phổ dòng chảy, tháng, sông Hồng tại Sơn Tây

chuỗi dòng chảy tháng V là đáng tin cậy. Chuỗi dòng chảy năm cũng như chuỗi dòng chảy tháng trong mùa cạn có thể coi là quá trình "nhiều đỏ" với dao động phân bố chủ yếu ở những tần số thấp, còn chuỗi dòng chảy tháng trong mùa lũ thì có thể coi là quá trình "nhiều trắng" với các dao động phân bố đều ở các tần số.

TÀI LIỆU THAM KHẢO

1- Trần Thanh Xuân, Đặng Lan Hương, v.v. "Phân tích chuỗi số liệu thủy văn nhiều năm và các chỉ số thủy văn liên quan đến sự biến đổi của khí hậu". Báo cáo tổng kết đề tài nghiên cứu, 1990.

2 - КАЙСЛ. Ч. Анализ временных рядов гидрологических данных
Л., Гидрометеиздат, 1972.

3 - РАТКОВИЧ. Д. Я. Многолетние колебания речного стока.
Л., Гидрометеиздат, 1976.

4 - РОЖДЕСТВЕНСКИЙ. А. В., ЧЕБОТАРЕВ. А. И. Статистические методы в гидрологии. Л., Гидрометеиздат, 1974.

5 - РОЖДЕСТВЕНСКИЙ. А. В. Пространственно - временные колебания стока рек СССР. Л., Гидрометеиздат, 1988.

6 - WMO. Analyzing long time series of hydrological data with respect to climate variability., 1988.