MÔ HÌNH TOÁN THỦY VĂN

Nguyễn Hữu Khải -Nguyễn Thanh Sơn

NXB Đại học Quốc gia Hà Nội 2003

Từ khoá: Tần suất, Chuẩn dòng chảy năm, Dòng chảy lũ, mặt dệm, dao động dòng chảy năm, phân phối dòng chảy năm, dòng chảy lũ, cường độ tới hạn, vi phân, dòng chảy kiệt, tài nguyên nước, môi trường, mô hình, tất định, ngẫu nhiên, phương pháp, Monte -Carlo

---

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

NGUYỄN HỮU KHẢI
NGUYỄN THANH SƠN

MÔ HÌNH TOÁN THUỶ VĂN

NHÀ XUẤT BẢN ĐẠI HỌC QUỐC GIA HÀ NỘI

MỤC LỤC

LỜI NÓI ĐẦU

Mô hình toán trong thuỷ văn đang ngày càng phát triển, được ứng dụng rộng rãi trong thực tế và bắt đầu được đưa vào chương trình giảng dạy và học tập ở bậc đại học. Tuy nhiên hiện nay chưa có giáo trình chính thức và đầy đủ về vấn đề này. Để đáp ứng yêu cầu nghiên cứu và học tập của sinh viên ngành thuỷ văn và tài nguyên nước, giáo trình đã được khẩn trương biên soạn. Các tác giả đã cố gắng tập hợp và hệ thống hoá những nghiên cứu gần đây về vấn đề này.

Tài liệu này rất cần thiết cho sinh viên và học viên cao học ở ngành thuỷ văn, Khoa Khí tượng-Thuỷ văn và Hải dương học, đồng thời là tài liệu tham khảo rất bổ ích cho sinh viên cũng như các học viên cao học ở các ngành có liên quan. Cuốn sách được các giảng viên đã giảng dạy và nghiên cứu nhiều về lĩnh vực mô hình toán thuỷ văn biên soạn.

Các tác giả chân thành cảm ơn các bạn đồng nghiệp về những đóng góp quý báu cho nội dung của cuốn sách. Cảm ơn Khoa Khí tương-Thuỷ văn và Hải dương học, Trường Đại học Khoa học tự nhiên, Đai học Quố gia Hà nội đã tạo mọi điều kiện thuận lợi cho việc xuất bản tài liệu này.

Đây là giáo trình được biên soạn lần đầu tiên, nên chắc rằng còn có những khiếm khuyết và thiếu sót, rất mong được sự đóng góp của bạn đọc.

Các tác giả

Chương 1
PHÂN TÍCH HỆ THỐNG VÀ MÔ HÌNH TOÁN THUỶ VĂN

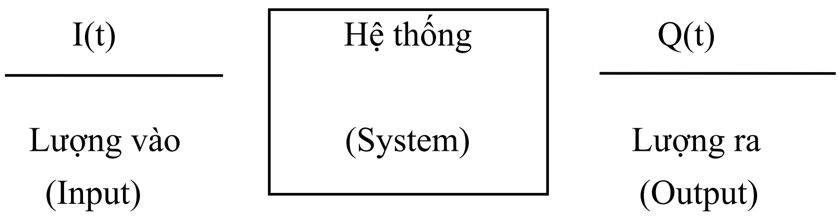
1.1. KHÁI NIỆM VỀ PHÂN TÍCH HỆ THỐNG VÀ MÔ HÌNH TOÁN THỦY VĂN

Ngày nay sự hiểu biết của con người về các quá trình thuỷ văn đã tiến được những bước dài. Con người đã hiểu biết khá sâu sắc về các quá trình hình thành dòng chảy, các cơ chế tác động và từ đó thiết lập các mô hình mô phỏng chúng. Tuy nhiên trong thực tế các hiện tượng thuỷ văn vô cùng phức tạp , chúng ta chỉ hiểu được một phần không đầy đủ về chúng và thiếu những lý thuyết hoàn chỉnh để mô tả tất cả các quá trình xẩy ra trong tự nhiên. Vì lẽ đó trong thuỷ văn vẫn sử dụng khái niệm hệ thống,cho phép mô tả các hiện tượng thuỷ văn một cách đơn giản hơn.

1.1.1 Khái niệm về phân tích hệ thống (Systematical analysis)

1.1.1.1. Hệ thống(System)

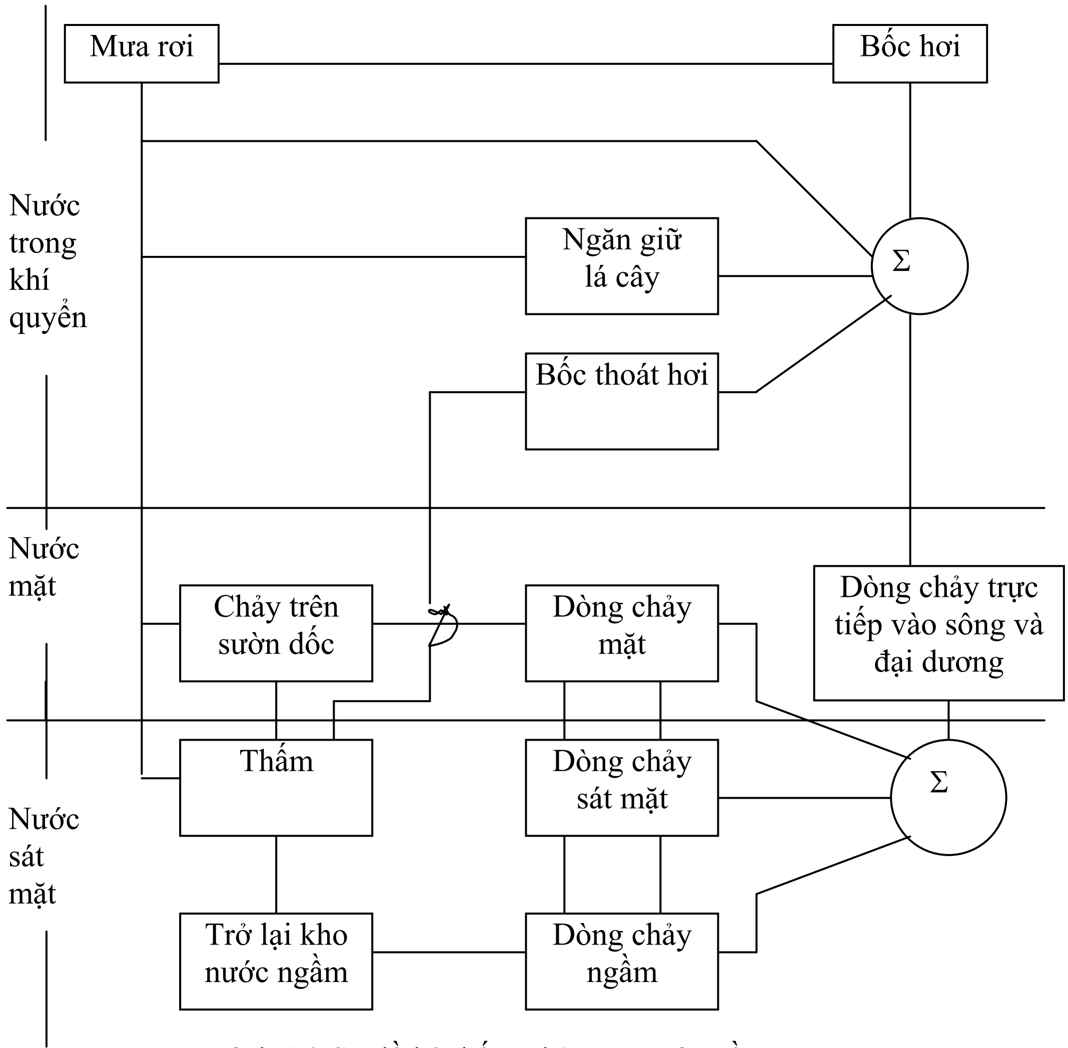
Hệ thống được hiểu là một tập hợp các thành phần có quan hệ liên thông với nhau để tạo thành một tổng thể. Theo Dooge (1964) hệ thống là bất kỳ một cấu trúc, thiết bị hoặc sơ đồ, trình tự nào đó, thực hay trừu tượng, được gắn với bước thời gian nhất định, liên hệ giữa lượng vào(nguyên nhân, năng lượng, thông tin) với lượmg ra(hệ quả, phản ứng, năng lượng) như hình 1.1.



Hình 1.1. Sơ đồ hệ thống

Hệ thống thuỷ văn (Hydrologic system) là các quá trình thuỷ văn (chu trình thuỷ văn) trên một vùng không gian nhất định và đó là các hệ thống thực. Ta có thể coi tuần hoàn thuỷ văn như một hệ thống với các thành phần là nước, bốc hơi, dòng chảy và các pha khác nhau của chu trình. Các thành phần này lại có thể tập hợp thành các hệ thống con của chu trình lớn. Để phân tích hệ thống toàn cục ta tiến hành xử lý, phân tích riêng rẽ các hệ thống con đơn giản hơn và tổng hợp các kết quả dựa trên mối quan hệ qua lại giữa chúng.

Trong hình 1.2 tuần hoàn thuỷ văn toàn cầu được miêu tả như một hệ thống. Các đường đứt quãng chia hệ thống thành 3 hệ thống con: Hệ thống nước khí quyển bao gồm các quá trình mưa rơi , bốc hơi ngăn giữ bởi cây cối và bốc thoát hơi sinh vật, hệ thống nước trên mặt đất với các quá trình chảy trên sườn dốc, dòng chảy mặt, quá trình chảy dòng sát mặt, dòng ngầm và các quá trình chảy trong sông và đổ ra biển, hệ thống nước dưới đất bao gồm các quá trình thấm, bổ sung nước ngầm, các dòng sát mặt và dòng ngầm. Các quá trình thuỷ văn, cũng theo định nghĩa của Dooge không chỉ bó hẹp trong số lượng dòng chảy mà là tập hợp các quá trình vật lý, hoá học cũng như sinh học của dòng chảy sông ngòi. Các quá trình này có thể do một hay nhiều biến vào, phản ứng của hệ thống có thể tạo ra nhiều quá trình ra.



Hình 1.2 Sơ đồ hệ thống thủy văn toàn cầu

Trong hầu hết các bài toán thực hành chúng ta chỉ xét một số ít quá trình trong tuần hoàn thủy văn tại một thời gian và một phạm vi không gian nhỏ bé nào đó của trái đất. Để nghiên cứu các bài toán này, người ta dùng một khái niệm hẹp hơn, thích hợp hơn đó là khái niệm " thể tích kiểm tra ". Đó là khái niệm được dùng trong cơ học chất lỏng biểu thị một không gian ba chiều, có chất lỏng chảy qua và các nguyên lý cơ bản về khối lượng, năng lượng và động lượng được áp dụng cho nó. Thể tích kiểm tra cung cấp cho chúng ta một cái khung để áp dụng các định luật về bảo toàn khối lượng, năng lượng và định luật II Niutơn, từ đó rút ra các phương trình động lực dùng trong thực hành. Trong quá trình suy diễn đó ta không cần biết mô hình chính xác của các dòng chất lỏng bên trong thể tích kiểm tra, mà chỉ cần biết tính chất của chất lỏng trên mặt kiểm tra, tức là biên giới của thể tích kiểm tra đang xét. Chất lỏng bên trong thể tích kiểm tra được coi như một khối mà khi xét đến tác dụng của các lực ngoài, ví dụ trọng lực, ta coi khối chất lỏng này như một điểm trong không gian tại đó tập trung khối lượng của chất lỏng .

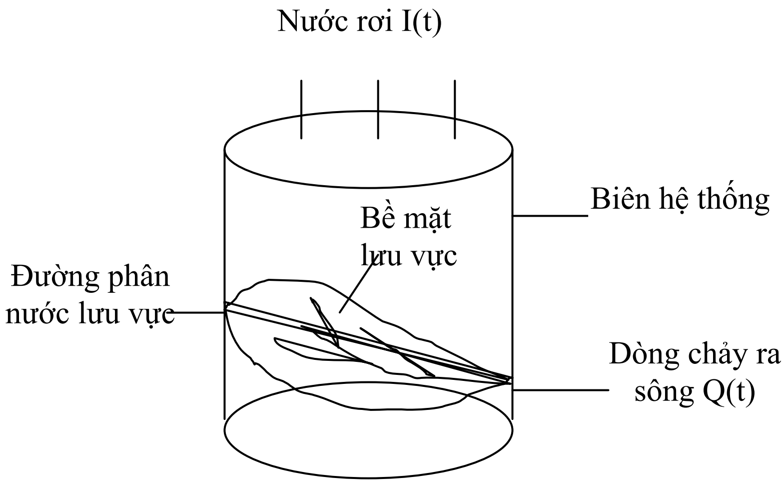
Tương tự, hệ thống thủy văn được định nghĩa như một cấu trúc hay một thể tích không gian bao quanh bởi một mặt biên. Cấu trúc này tiếp nhận các yếu tố đầu vào (Input) qua mặt biên như mưa theo phương thẳng đứng, dòng chảy theo phương ngang, thao tác phân tích các yếu tố đó ở bên trong và biến đổi chúng thành các yếu tố đầu ra (Output) ở mặt biên bên kia. Có thể hiểu cấu trúc của hệ thống (hay thể tích không gian) là toàn bộ các đường đi, các phương thức khác nhau để qua đó nước xuyên suốt qua hệ thống từ điểm đi vào cho đến điểm đi ra. Biên của hệ thống là một mặt liên tục, xác định trong không gian 3 chiều bao quanh cấu trúc hay thể tích đang xét. Một đối tượng nghiên cứu nào đó đi vào hệ thống như một yếu tố đầu vào, tác động qua lại với cấu trúc và các yếu tố khác, rồi rời khỏi hệ thống thành yếu tố đầu ra. Nhiều quá trình vật lý, hoá học và sinh học khác nhau ở bên trong cấu trúc đã tác động lên đối tượng.

1.1.1.2. Phân tích hệ thống

Phân tích hệ thống là tìm hiểu cấu trúc và sự vận hành của hệ thống, xác lập các mô hình mô tả chúng .

Người ta tiến hành thiết lập các phương trình và mô hình của các hiện tượng thủy văn theo các bước tương tự như cơ học chất lỏng. Tuy nhiên, việc áp dụng các định luật vật lý mang tính xấp xỉ gần đúng nhiều hơn bởi vì hệ thống nhiều hơn, phức tạp hơn, có thể bao hàm nhiều yếu tố cần xét. Mặt khác phần lớn các hệ thống thủy văn mang tính ngẫu nhiên bởi vì yếu tố đi vào hệ thống là mưa, một hiện tượng có tính biến động lớn và tính ngẫu nhiên cao. Cũng chính vì vậy, phân tích thống kê giữ một vai trò quan trọng trong này.

Ví dụ ta có thể biểu thị quá trình mưa rào dòng chảy trên một lưu vực như là một hệ thống thủy văn (hình 1.3). Lượng mưa là yếu tố đầu vào được phân bố trong không gian trên mặt phẳng phía trên. Lưu vực là diện tích tập trung nước của một con sông. Biên của hệ thống được dựng xung quanh lưu vực bằng cách chiếu thẳng đứng đường phân nước tới hai mặt nằm ngang tại đỉnh và đáy. Yếu tố đầu ra là dòng nước tập trung trong không gian tại cửa ra của lưu vực. Lượng bốc hơi và dòng sát mặt cũng có thể coi là yếu tố đầu ra nhưng thường rất nhỏ so với dòng chảy sinh ra trong một trận mưa nên có thể bỏ qua.



Hình 1.3 : Minh hoạ lưu vực như một hệ thống thủy văn .

Cấu trúc của hệ thống là tập hợp các đường đi của dòng chảy trên mặt hoặc trong đất bao gồm cả các dòng nhánh, những dòng này cuối cùng sẽ hoà nhập thành dòng chảy tại mặt cắt cửa ra. Cấu trúc của hệ thống chịu ảnh hưởng của các đặc tính lưu vực như địa hình, địa chất, thổ nhưỡng, các đặc trưng hình thái lưu vực và sông

Nếu khảo sát thật chi tiết bề mặt và các tầng đất của lưu vực ta thấy số lượng các đường di chuyển của dòng chảy có thể vô cùng lớn. Dọc theo một đường đi bất kỳ, hình dạng, độ nhám, độ dốc bề mặt có thể thay đổi liên tục từ vị trí này sang vị trí khác, đồng thời thay đổi theo thời gian. Mặt khác mưa cũng biến đổi ngẫu nhiên theo không gian và thời gian. Do sự phức tạp như vậy ta không thể mô tả một số quá trình thủy văn bằng những định luật vật lý chính xác. Sử dụng khái niệm hệ thống người ta tập trung xây dựng một mô hình liên hệ các yếu tố đầu vào và sản phẩm đầu ra hơn là miêu tả một cách chính xác các chi tiết của hệ thống.

Sự miêu tả chính xác như vậy có thể không mang ý nghĩa thực tiễn hoặc không thực hiện được vì nó vượt quá khả năng hiểu biết của chúng ta. Tuy nhiên sự hiểu biết về hệ thống vật lý sẽ giúp ích rất nhiều trong việc thiết lập mô hình một cách đúng đắn và kiểm chứng độ chính xác của nó .

1.1.2. Khái niệm mô hình toán thủy văn

1.1.2.1 Mô hình toán học hệ thủy văn.

Mục tiêu của phân tích hệ thống là nghiên cứu sự vận hành của hệ thống và dự toán kết quả đầu ra. Mô hình hệ thống thủy văn là phản ánh gần đúng của một hệ thống thủy văn có thật. Các yếu tố đầu vào và sản phẩm đầu ra là các biến lượng thủy văn đo được .

Mô hình hệ thống thủy văn có thể là mô hình vật lý, tương tự hay toán học. Mô hình vật lý bao gồm các mô hình tỉ lệ tức là các mô hình biểu thị hệ thống thật dưới dạng thu nhỏ như mô hình thủy lực của đập tràn. Mô hình tương tự là một mô hình vật lý khác có tính chất tương tự như mô hình nguyên thể, chẳng hạn một số mô hình điện trong thủy lực .

Mô hình toán học miêu tả hệ thống dưới dạng toán học. Mô hình toán học là tập hợp các phương trình toán học, các mệnh đề logic thể hiện các quan hệ giữa các biến và các thông số của mô hình để mô phỏng hệ thống tự nhiên (Reepgaard) hay nói cách khác mô hình toán học là một hệ thống biến đổi đầu vào (hình dạng, điều kiện biên, lực v.v...) thành đầu ra (tốc dộ chảy, mực nước, áp suất v.v...) (Novak).

Chúng ta biểu thị đầu vào và đầu ra của hệ thống là các hàm của thời gian, thứ tự là I(t) và Q(t) , trong đó t là biến thời gian trong khoảng thời gian T đang xét. Hệ thống thực hiện một phép biến đổi, biến yếu tố đầu vào I(t) thành đầu ra Q(t) theo phương trình :

$$Q = \Omega I(t) \tag{1.1}$$

Phương trình này được gọi là phương trình biến đổi của hệ thống .

Ω là một hàm truyền (Propogation function) giữa các yếu tố đầu vào và đầu ra. Đôi khi người ta còn gọi là hàm ảnh hưởng hay hàm phản ứng. Nếu mối liên hệ này có thể biểu thị bằng một phương trình đại số thì Ω là một toán tử đại số. Ví dụ nếu có :

$$Q(t)=C.I(t) \tag{1.2}$$

trong đó C là một hằng số thì hàm truyền sẽ là một toán tử:

$$\Omega = \frac{Q(t)}{I(t)} \tag{1.3}$$

Nếu phép biến đổi được mô tả bởi một phương trình vi phân thì hàm truyền là một toán tử vi phân. Ví dụ trong một kho nước tuyến tính lượng trữ S liên hệ với lưu lượng ra Q qua phương trình :

$$S = KQ \tag{1.4}$$

trong đó K là một hằng số. Từ tính liên tục của dòng chảy ta có lượng biến thiên của lượng trữ trong một đơn vị thời gian dS/dt bằng hiệu giữa lượng vào I(t) và lượng ra Q(t) :

$$\frac{dS}{dt} = I\ (t) - Q\ \ (t) \qquad (1.5)$$

Thay S từ (1.4) vào (1.5) ta có :

$$K\ .\frac{dQ}{dt} + \ Q\ (t) = I\ (t) \qquad (1.6)$$

Do đó:

$$\Omega = \frac{Q(t)}{I(t)} = \frac{Q(t)}{K.\frac{dQ}{dt} + Q(t)} = \frac{Q}{Q + KD} \qquad (1.7)$$

trong đó D là một toán tử vi phân d/dt .

Nếu phương trình biến đổi hệ thống (1.7) đã được xác định và có thể giải được thì nó cho ta kết quả đầu ra như là hàm của yếu tố đầu vào.

Cũng có thể viết mô hình toán học của hệ thống theo dạng sau :

$$f\left[I(t), Q(t), \frac{\partial I}{\partial t}, \frac{\partial Q}{\partial t}, ..., \frac{\partial^2 I}{\partial t^2}, \frac{\partial^2 Q}{\partial t^2}, ..., \theta_1, \theta_2, ... \right] = 0 \qquad (1.8)$$

trong đó f [...] là một hàm số có dạng xác định. Còn θ_1, θ_2,... là các thông số có thể trực tiếp đo đạc trên bản đồ hoặc xác định theo tài liệu thực đo .

Trong thực tế các biến I(t), Q(t) không thể đo liên tục mà đo rời rạc theo ccác thời đoạn bằng nhau. Do vậy để thuận tiện ta viết I(t)=Q(t) biểu thị các giá trị của các biến I(t) , Q(t) tại thời điểm t , và thay các đạo hàm riêng $\left[\frac{\partial I}{\partial t}, \frac{\partial Q}{\partial t}, \frac{\partial^2 I}{\partial t^2}, \frac{\partial^2 Q}{\partial t^2} \right]$ bằng các sai phân thì phương trình (1.8) có thể viết lại như sau :

$$f\left[I_t, Q_t, I_{t-1}, Q_{t-1}, I_{t-2}, Q_{t-2}, ..., \theta_1, \theta_2 \right] = 0 \qquad (1.9)$$

Nói chung hệ thống thực rất phức tạp khi mô hình hóa thường dùng một hàm tương đối đơn giản f*[...] trong phương trình 1.9 khi đó sẽ mắc một sai số. Ta có thể viết lại (1.9) có tính đến sai số này như sau :

$$f\left[I_t, Q_t, I_{t-1}, Q_{t-1}, I_{t-2}, Q_{t-2}, ..., \theta_1, \theta_2 \right] + \varepsilon_t = 0 \qquad (1.10)$$

Hay f= $f\left[I_t, Q_t, I_{t-1}, Q_{t-1}, I_{t-2}, Q_{t-2}, ..., \theta_1, \theta_2 \right] + \varepsilon_t = 0$ (1.11)

Phương trình (1.11) biểu thị một mô hình toán học với hàm số f* là hàm số mô phỏng mô hình. Việc chọn dạng f* để mô tả hệ thống thực là một vấn đề chủ yếu khi xây dựng mô hình .

1.1.2.2 Thông số mô hình (Parametter of model).

Thông số là đặc trưng số lượng của hệ thống thủy văn. Ví dụ diện tích lưu vực là một thông số biểu thị độ lớn của lưu vực. Nói chung thông số của hệ thống không thay đổi theo thời gian trong điều kiện các nhân tố ảnh hưởng đến hệ thống ổn định. Đặc tính của hệ thống có thể biểu thị qua nhiều thông số khác nhau .

Hiệu quả của mô hình phụ thuộc trước hết vào độ chính xác xác định thông số. Nếu thông tin ban đầu không đầy đủ thì khi tăng số thông số, mặc dù cho phép mô tả đầy đủ hơn và chính xác hơn quá trình, nhưng có thể đưa đến những kết quả kém hơn bởi vì các thông số được lựa chọn sẽ có sai số lớn hơn. Vì vậy phải lựa chọn một cấu trúc mô hình tối ưu nào đó, bao gồm một số lượng tối ưu các thông số, có thể mô tả tốt các quá trình cơ bản trong hệ thống thông tin đã có, đồng thời phải đưa ra các phương pháp xác định chính xác các thông số. Thực tế cho thấy khả năng thay đổi cấu trúc mô hình luôn lớn hơn khả năng thay đổi các phương pháp xác định thông số .

1.1.2.3 Cấu trúc mô hình (Structure of model).

Cấu trúc mô hình phản ánh thứ tự các khối tính toán và mô tả từ hàm vào đến hàm ra. Có 3 khuynh hướng lựa chọn cấu trúc mô hình :

- Thứ nhất là chọn một cấu trúc chung nhất bao hàm tất cả các hiện tượng và tập hợp các nhân tố tác động.

- Thứ hai là chọn cấu trúc mô tả tốt nhất các hiện tượng và đối tượng thủy văn riêng biệt cho từng bài toán cụ thể .

- Thứ ba là lựa chọn một cấu trúc nào đó đã được nghiên cứu và chỉnh lý tốt để áp dụng cho các hiện tượng thủy văn. Trong thực tế có nhiều mô hình có thể áp dụng cho một lớp rộng rãi các bài toán. Tuy nhiên khi đó đã sử dụng tính tương tự giả tạo và không tính được các đặc điểm riêng biệt quan trọng của quá trình thủy văn .

Lựa chọn khuynh hướng này hay khuynh hướng khác phụ thuộc vào ý chí chủ quan của những người thiết lập mô hình. Nhưng dù sao cấu trúc mô hình phải tận dụng đến mức tối đa các thông tin đã có và độ chính xác xác định các thông số. Trong khi xác lập cấu trúc mô hình cần chú ý đến lý thuyết chung về loại hiện tượng cũng như các quan hệ đặc thù vốn có của một hiện tượng riêng biệt. Cấu trúc mô hình thường biểu hiện cho các thông tin cơ bản về một loại quá trình, còn các thông số của nó đặc trưng cho mỗi hiện tượng, khu vực cụ thể .

Thông tin nhận được nhờ tính toán theo mô hình không thể nhiều hơn những thông tin vốn có của chính mô hình. Cấu trúc mô hình càng tổng hợp thì những thông tin có trong nó phản ánh cho các hiện tượng riêng biệt càng ít. Việc lựa chọn cấu trúc mô hình liên hệ chặt chẽ với vấn đề đưa vào nó các thông tin chứa trong các quan trắc cụ thể và các thông số.

Để lựa chọn cấu trúc mô hình tối ưu có thể sử dụng nguyên tắc phức tạp dần mô hình. Thực chất của nó là việc tối ưu hóa được tiến hành theo từng giai đoạn. Trong các thông số mô hình, tỷ trọng của từng thông số không đồng đều nhau, tính chất của các thông số không giống nhau. Do vậy không thể đồng thời đưa tất cả tối ưu vào cùng một lúc. Việc phức tạp hóa dần cấu trúc mô hình được bắt đầu bằng việc thể nghiệm một mô hình đơn giản nhất, với một số thông số tối thiểu. Sau khi đã tối ưu được các thông số đó, mô hình được chính xác hoá dần nhờ việc đưa thêm dần các thông số mới, mô tả chính xác thêm hiện tượng. Ở từng giai đoạn, các thông số được tối ưu một cách độc lập trên cơ sở các thông số của giai đoạn trước, tức là lấy giá trị ban đầu bằng các giá trị đã được tối ưu.

1.1.2.4. Xây dựng và ứng dụng mô hình toán thuỷ văn.

Để xây dựng mô hình toán cần thực hiện các bước sau:

- Xác định bài toán: Định nghĩa hệ thống, xác định hàm vào, hàm ra, các điều kiện mô phỏng hệ thống .
- Xây dựng cấu trúc mô hình toán .
- Mô phỏng toán học các thành phần trong mô hình và các quan hệ giữa chúng.
- Xây dựng các chương trình trên máy tính cho các nội dung của mô hình toán

Khi giải quyết các bài toán về mô hình có hai loại bài toán sau :

- Bài toán thuận: Cho đầu vào I(t) và cấu trúc mô hình, yêu cầu xác định được đầu ra. Nếu mô hình là các phương trình vi phân thì bài toán này là giải các phương trình vi phân đó với điều kiện ban đầu và điều kiện biên đã cho .
- Bài toán ngược: Các đại lượng ra đã biết, cần xác định dạng cấu trúc mô hình cùng các thông số của nó hoặc hàm đầu vào (điều kiện ban đầu và điều kiện biên), trong đó quan trọng nhất là xác định cấu trúc và thông số của mô hình .

Để ứng dụng mô hình toán cần tiến hành theo các bước:

- Chọn mô hình tuỳ theo điều kiện của bài toán, tuỳ theo tình hình tài liệu và đặc điểm khu vực ứng dụng .
- Thu thập chỉnh lý các tài liệu Khí tượng- thủy văn (hàm vào, hàm ra), tính toán các thông số biểu thị đặc tính của hệ thống, lưu vực.
- Hiệu chỉnh xác định thông số mô hình theo số liệu quan trắc đồng bộ của hàm vào và hàm ra.
- Kiểm tra mô hình theo tài liệu độc lập.

Nếu các tiêu chuẩn đánh giá mô hình được đảm bảo thì các mô hình với các thông số ở trên có thể sử dụng trong tính toán và dự báo tiếp theo. Ở đây cần thừa nhận các thông số mô hình là không thay đổi cho đến thời gian dự báo hoặc tính toán.

Với các mô hình có cấu trúc phức tạp, khối lượng tính toán thực hiện rất lớn. Vì vậy hầu hết các nội dung tính toán phải thực hiện trên các máy tính điện tử. Ngày nay cùng với sự phát triển của tin học các mô hình toán thủy văn ngày càng phát triển.

1.2. PHÂN LOẠI MÔ HÌNH TOÁN THỦY VĂN

Có nhiều cách phân loại mô hình tùy theo quan điểm và ý tưởng của người phân loại. Một trong các cách phân loại là dựa trên cơ sở xem xét sự phân bố của các biến vào và ra hệ thống trong trường không gian, thời gian.

Mô hình toán thủy văn là mô hình miêu tả hệ thống dưới dạng toán học. Sự vận hành của hệ thống được mô tả bằng một hệ phương trình liên kết giữa các biến vào, ra của hệ thống. Các biến này có thể là hàm của thời gian và không gian và cũng có thể là các biến ngẫu nhiên, không lấy giá trị xác định tại một điểm riêng biệt trong không gian, thời gian mà được mô tả bằng các phân bố xác suất. Biểu thị tổng quát cho các biến như vậy là một trường ngẫu nhiên, một vùng của không-thời gian, trong đó các biến tại những điểm khác nhau trong trường được xác định bởi một phân bố xác suất.

Xây dựng mô hình với các biến ngẫu nhiên phụ thuộc cả vào thời gian và không gian 3 chiều, đòi hỏi một khối lượng công việc khổng lồ. Vì thế trong thực hành người ta xây dựng các mô hình giản hoá bằng cách bỏ qua một số nguồn biến đổi. Các mô hình thủy văn có thể phân loại theo các đường lối giản hoá được áp dụng. Đối với một mô hình, người ta xem xét 3 quyết định cơ bản sau:

- Các biến trong mô hình có là ngẫu nhiên không?
- Chúng biến đổi theo không gian như thế nào?
- Chúng biến đổi theo thời gian ra sao?

Tùy thuộc sự lựa chọn các quyết định trên, các mô hình có thể phân loại theo "cây phân loại" như hình 1.4 .

Ở mức tổng quát nhất có thể chia ra thành mô hình tất định và mô hình ngẫu nhiên. Trong mô hình tất định không xét đến tính ngẫu nhiên còn trong mô hình ngẫu nhiên, sản phẩm đầu ra ít nhiều mang đặc tính ngẫu nhiên.

Tại mức thứ hai của cây phân loại 1.4 chúng ta nghiên cứu phân loại theo tính biến thiên theo không gian của hiện tượng. Nói chung các hiện tượng thủy văn đều biến thiên theo một không gian 3 chiều. Nhưng sự xem xét đầy đủ tất cả các biến đổi sẽ làm cho bài toán cồng kềnh. phân biệt mô hình tất định với thông số tập trung và mô hình tất định với thông số phân bố. Trong mô hình tất định với thông số tập trung hệ thống được trung bình hoá trong không gian hoặc có thể coi hệ thống như một điểm đơn độc trong không gian. Trong mô hình tất định với thông số phân bố người ta xem xét diễn biến của các quá trình thủy văn tại các vị trí khác nhau trong không gian.

Mô hình ngẫu nhiên tại mức trung gian này được chia ra thành mô hình không gian độc lập và không gian tương quan tuỳ theo mức độ ảnh hưởng lẫn nhau của các biến ngẫu nhiên tại các vị trí khác nhau trong không gian.

Tại mức thứ ba của cây phân loại chúng ta xem xét tính biến thiên theo thời gian của hiện tượng. Ở đây dòng chất lỏng trong mô hình tất định được phân ra thành dòng ổn định(có tốc độ dòng chảy không thay đổi theo thời gian) và dòng không ổn định. Còn trong mô hình ngẫu nhiên có thể phân ra thành mô hình ngẫu nhiên thời gian độc lập hay thời gian tương quan. Mô hình thời gian độc lập miêu tả một dãy các sự kiện thủy văn không ảnh hưởng lẫn nhau, trong khi đó mô hình ngẫu nhiên thời gian tương quan mô phỏng một dãy trong đó sự kiện tiếp theo bị ảnh hưởng một phần bởi sự kiện hiện tại hoặc một số sự kiện khác trong dãy. Sau đây chúng ta phân tích chi tiết hơn từng loại mô hình.

1.2.1. Mô hình tất định (Deterministic model)

Trong mô hình này người ta không xét đến tính ngẫu nhiên, các biến vào ra không mang tính ngẫu nhiên, không mang một phân bố xác suất nào cả. Các đầu vào như nhau đi qua hệ thống sẽ cho ta cùng một sản phẩm đầu ra. VenteChow(1964) có nêu định nghĩa “Nếu các cơ hội xảy ra của các biến của quá trình thủy văn được bỏ qua trong mô hình toán, mô hình coi như tuân theo qui luật tất định và có thể gọi là mô hình tất định”. Mặc dù các hiện tượng thủy văn đều ít nhiều mang tính ngẫu nhiên, nhưng đôi khi mức độ biến đổi ngẫu nhiên của đầu ra có thể rất nhỏ bé so với sự biến đổi gây ra bởi các nhân tố đã biết. Trong trường hợp đó sử dụng mô hình tất định là thích hợp.

Về ý nghĩa khái niệm tất định như trên biểu thị mối quan hệ nhân quả của mô hình toán thủy văn. Việc mô tả hệ thống thủy văn thực theo mô hình tất định gọi là mô phỏng tất định (deterministic simulation) hệ thủy văn. Thông qua mô phỏng các thành phần chủ yếu hoặc toàn bộ quá trình thủy văn theo các phương trình toán học, các mô hình toán thuỷ văn có khả năng dần dần thể hiện và tiếp cận hệ thống, biểu đạt gần đúng qui luật của hệ thống. Trong mô hình, hệ thống thủy văn luôn được coi là hệ thống kín, các biến vào ra thực tế là các quá trình biến đổi theo thời gian và có thể đo đạc được. Sử dụng mô hình có thể tính toán các quá trình ra (biến ra) theo các giá trị đo đạc được của quá trình vào (biến vào).

Những mô hình toán thủy văn tất định trong thực tế thường dùng để mô phỏng quá trình hình thành dòng chảy trên lưu vực, quá trình vận động nước trong sông. Nó cho khả năng xem xét, đánh giá được những phản ứng của hệ thống khi cấu trúc bên trong thay đổi. Thí dụ như khi xây dựng các hồ chứa điều tiết hay trồng rừng, phá rừng thượng nguồn.

1.2.1.1. Mô hình tất định với thông số tập trung (Lumped parametter model)

Trong mô hình này hệ thống được trung bình hoá trong không gian và có thể coi hệ thống như một điểm đơn độc trong không gian. Các thông số coi như không thay đổi theo không gian mà chỉ nhận một giá trị đặc trưng cho cả hệ thống. Trong mô hình tất định với thông số tập trung, các quan hệ toán học thường biểu đạt bằng các phương treình vi phân thường với các quá trình lượng vào và lượng ra hệ thống chỉ phụ thuộc vào thời gian. Chẳng hạn mô hình mưa dòng chảy nêu trong hình (1.3) đã coi lượng mưa phân bố đều trên lưu vực và bỏ qua sự biến đổi theo không gian của dòng chảy. Mô hình tất định với thông số tập trung còn được gọi là mô hình diễn toán thủy văn.

- *Mô hình tất định với thông số tập trung ổn định (Steady lumped parametter model).* Trong mô hình này dòng chuyển động là dòng ổn định, không thay đổi theo thời gian và không gian nghĩa là dòng vào và dòng ra bằng nhau, lượng biến đổi lượng trữ bên trong hệ thống bằng không, mối quan hệ giữa lượng nhất và lượng ra là đơn nhất.

- *Mô hình tất định với thông số tập trung không ổn định(Unsteady lumped parametter model).* Trong mô hình này dòng vào và dòng ra đều biến đổi theo thời gian và không bằng nhau. Từ đó dẫn đến sự thay đổi lượng trữ bên trong hệ thống. Quan hệ giữa lượng trữ và dòng ra có dạng vòng dây. Các mô hình toán thuỷ văn hiện nay hầu hết thuộc loại này.

1.2.1.2. Mô hình tất định với thông số phân bố(Distributed parametter model).

Trong mô hình này xem xét sự diễn biến của quá trình thủy văn tại các điểm khác nhau trong không gian và định nghĩa các biến trong mô hình như là hàm toạ độ. Các thông số được xem xét theo sự biến đổi không gian của hệ thống. Các phương trình biểu đạt các quan hệ là các phương trình đạo hàm riêng, chứa cả biến không gian và thời gian. Để diễn tả hệ thống theo mô hình này thường chia hệ thống ra các ô lưới, mỗi ô lưới diễn tả đặc tính riêng của hệ thống toạ độ cùng với các thông số của chúng.

Mô hình tất định với thông số phân bố cho phép mô tả sự biến đổi không gian của hiện tượng thủy văn. Nhưng khi đó bài toán xác định các thông số trở nên phức tạp hơn. Khi sử dụng nó cần phải thay đổi về chất các phương pháp xác định thông số cũng như phương pháp đo đạc các đặc trưng của hệ thống. Điều này cho đến nay chúng ta chưa làm được bao nhiêu. Mô hình điển hình trong loại này hiện nay là hệ phương trình SaintVenant, đó là mô hình lâu đời nhất và được nghiên cứu tốt nhất. Hệ phương trình này được sử dụng để tính toán chuyển động không ổn định trong sông và trong kênh, nhưng cũng có thể dùng để mô tả các quá trình xảy ra trên lưu vực. Mô hình tất định với thông số phân bố còn được gọi là mô hình diễn toán thủy lực. Mô hình này lại được chia ra:

- *Mô hình tất định với thông số phân bố ổn định (Steady distributed parametter model).* Trong mô hình xem xét các dòng vào, dòng ra thay đổi theo không gian nhưng lại không thay đổi theo thời gian. Có thể coi dòng ổn định trong kênh phi lăng trụ với độ dốc đáy khác nhau thuộc loại mô hình này, ở đó các thông số thay đổi theo dòng chảy nhưng không thay đổi theo thời gian.

- *Mô hình tất định với thông số phân bố không ổn định(Unsteady distributed parametter model)* Đây là mô hình tổng quát nhất trong mô hình tất định. Dòng ra, dòng vào, các thông số đều thay đổi theo thời gian và không gian. Các giá trị của mô hình được xác định trên một mạng lưới của các điểm trên mặt phẳng không-thời gian. Loại mô hình này được dựa trên các phương trình đạo hàm riêng mô tả một chiều thời gian và ba chiều không gian. Mô hình hệ thống Saint Venant đầy đủ thuộc loại này.

Việc giải mô hình đầy đủ là rất phức tạp. Do đó người ta thường đơn giản hoá một số điều kiện để việc giải dễ dàng hơn và từ đó ta có các mô hình khác nhau.

1.2.2. Mô hình ngẫu nhiên(Stochastic model)

Trong mô hình ngẫu nhiên các kết quả đầu ra luôn mang tính ngẫu nhiên tức là luôn tuân theo một quy luật xác suất nào đấy. Ta có thể nói mô hình tất định thực hiện một "dự báo"(forecast), còn mô hình ngẫu nhiên thực hiện một "dự đoán"(Prediction). Nếu tính biến đổi ngẫu nhiên của đầu ra là lớn thì kết quả đầu ra có thể rất khác biệt với giá trị đơn nhất tính toán theo mô hình tất định. Ví dụ ta có thể xây dựng các mô hình tất định với chất lượng tốt tại một điểm cho trước bằng các số liệu về cung cấp năng lượng và vận chuyển hơi nước, nhưng cũng với số liệu này ta không thể xây dựng được mô hình tin cậy về lượng mưa ngày rất lớn. Vì vậy hầu hết các mô hình mưa ngày đều là ngẫu nhiên.

Thực sự các quá trình thuỷ văn, trong đó có dòng chảy là một hiện tượng ngẫu nhiên dưới tác động của nhiều nhân tố. Từng nhân tố đến lượt mình lại là hàm của rất nhiều nhân tố khác mà qui luật của nó, con người chưa thể nào mà tả đầy đủ được. Cuối cùng các quá trình thủy văn lại là sự tổ hợp của vô vàn các mối quan hệ phức tạp, biểu hiện là một hiện tượng ngẫu nhiên và được mô tả bằng một mô hình ngẫu nhiên. Với quan điiểm cho rằng dòng chảy là một quá trình ngẫu nhiên, trong cấu trúc mô hình ngẫu nhiên không hề có các nhân tố hình thành dòng chảy và nguyên liệu để xây dựng mô hình chính là bản thân số liệu chuỗi dòng chảy trong quá khứ. Vì vậy chuỗi số liệu phải đủ dài để bộc lộ hết đặc tính của nó. Lớp này không quan tâm đến các nhân tố tác động đến quá trình thủy văn mà chỉ xem xét khả năng diễn biến của bản thân quá trình đó, và chủ yếu là sản sinh ra những thể hiện mới đầy đủ hơn của một quá trình ngẫu nhiên. Ngày nay lĩnh vực này tách ra thành một chuyên ngành riêng dưới tên gọi là "Thủy văn ngẫu nhiên".

Trong thời gian gần đây người ta xem xét đưa vào các mô hình tất định các thành phần ngẫu nhiên và hình thành lớp mô hình tất định-ngẫu nhiên. Việc đưa tình ngẫu nhiên vào mô hình tất định diễn ra theo 3 hướng sau:

- Xét sai số tính toán như một quá trình ngẫu nhiên và trở thành một thành phần trong mô hình.

- Sử dụng các mô tả xác suất cho các hàm vào.

- Xét qui luật phân bố không gian của các tác động Khí tượng-Thủy văn dưới dạng hàm phân bố xác suất.

Vì tình phức tạp của vấn đề, lớp mô hình này chỉ ở giai đoạn đầu nghiên cứu.

1.2.2.1. Mô hình ngẫu nhiên độc lập không gian (Spatial independent Stochactic model)

Trong mô hình này coi các biến và các thông số có phân bố xác suất như nhau tại mọi điểm không gian và độc lập đối với nhau, hay nói cách khác chúng không có tương quan với nhau, giá trị tại một vị trí này không làm ảnh hưởng tới vị trí khác. Dạng mô hình này được dùng nhiều trong thống kê thủy văn.

- *Mô hình ngẫu nhiên độc lập không-thời gian (Spatial-temporal indeperdent Stochactic model)* Trong mô hình này hàm phân bố xác suất là duy nhất và chỉ là hàm một chiều. Các đại lượng xuất hiện tại các thời điểm khác nhau không làm ảnh hưởng lẫn nhau và giá trị tại vị trí này không liên quan đến vị trí khác. Các mô hình xác suất thống kê thủy văn hiện nay hầu hết thuộc loại này.

- *Mô hình ngẫu nhiên độc lập không gian nhưng tương quan thời gian (Spatial indeperdent and temporal correlational Stochactic model*) Trong mô hình này coi khả năng(xác suất) xuất hiện của các biến trong không gian không làm ảnh hưởng lẫn nhau. Nhưng giá trị của biến tại một thời điểm chịu ảnh hưởng của các giá trị tại một số thời điểm trước, nói cách khác giá trị của các biến theo thời gian có tương quan với nhau. Hàm phân bố xác suất là hàm phân bố nhiều chiều. Mô hình này mô tả một quá trình ngẫu nhiên tại một vị trí hay tuyến riêng biệt. Xích Markov là một mô hình thuộc loại này, được sử dụng nhiều trong việc mô tả dao động của dòng chảy tháng và năm.

1.2.2.2 Mô hình ngẫu nhiên tương quan không gian (Spatial correlational Stochectic model).

Trong mô hình này cho rằng các biến và các thông số có phân bố xác suất và có tương quan với nhau theo không gian. Hàm phân bố xác suất tại vị trí này có ảnh hưởng đến hàm phân bố tại vị trí khác. Ví dụ trong hệ thống bể chứa nối tiếp, giá trị xác định theo hàm phân bố của bể chứa trên có ảnh hưởng đến hàm phân bố của bể chứa phía dưới. Đây là bài toán có ý nghĩa thực tiễn lớn, tuy nhiên vì tính phức tạp của vấn đề nên các mô hình dạng này chưa nhiều.

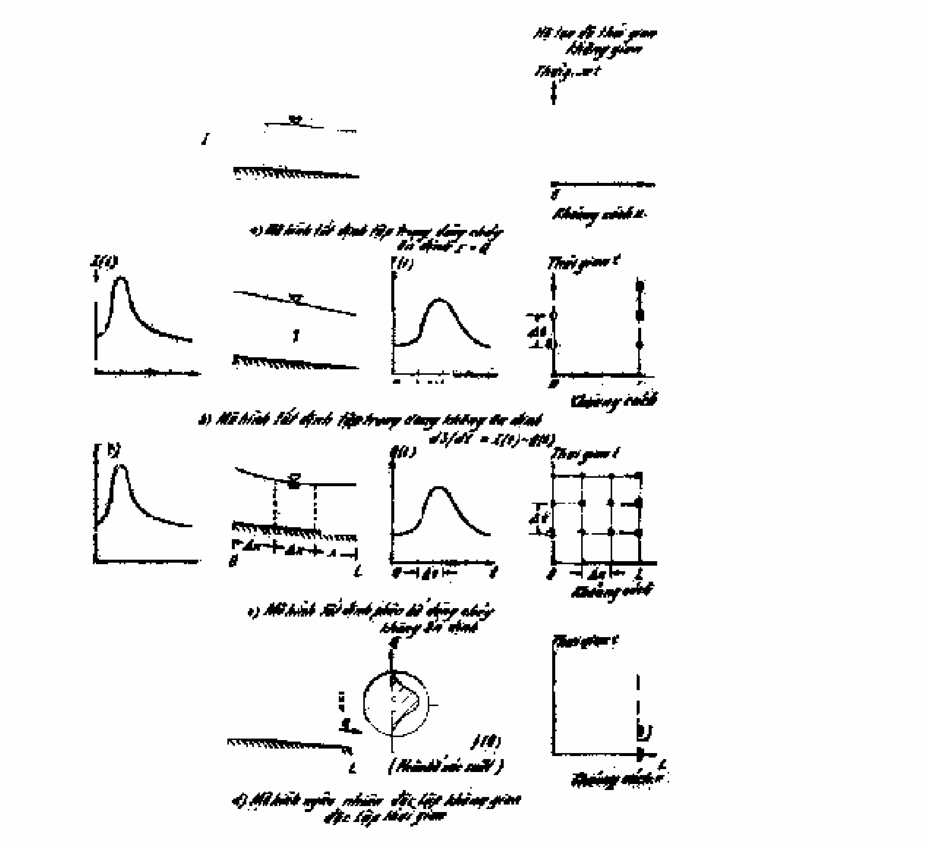
- *Mô hình ngẫu nhiên tương quan không gian nhưng độc lập thời gian (Spatial correlational and tempora independerent Stochactic model*) Trong mô hình xem xét tác động tương hỗ giữa xác suất xuất hiện các biến tại các vị trí khác nhau, nhưng theo thời gian không bị ảnh hưởng. Mô hình mô tả một trường ngẫu nhiên các quá trình thủy văn. Dạng mô hình này được xem xét trong các bài toán tổ hợp xác suất, tuy nhiên còn ở trong những dạng đơn giản.

- *Mô hình ngẫu nhiên tương quan không-thời gian (Spactial-Temporal correlational Stochactic model)* Đây là mô hình tổng quát nhất trong lớp mô hình ngẫu nhiên. Trong mô hình xem xét xác suất xuất hiện của các biến phụ thuộc lẫn nhau cả theo không gian, cả theo thời gian. Loại này đang được đầu tư nghiên cứu vì ý nghĩa thực tiễn của nó. Tuy nhiên kết quả còn hạn chế vì bài toán trở nên rất phức tạp. Một số phiên bản của mô hình Markov cho chuỗi dòng chảy có quan hệ tương hỗ là một thử nghiệm của mô hình này.

Mọi mô hình thủy văn chỉ là một mẫu gần đúng của thực tế, do đó sản phẩm của hệ thống thật không bao giờ dự báo được một cách chắc chắn. Các hiện tượng thủy văn thường biến đổi theo thời gian và trong không gian 3 chiều, nhưng việc xem xét đồng thời tất cả 5 nguồn biến động(ngẫu nhiên, theo thời gian và theo không gian 3 chiều) cũng chỉ có thể thực hiện trong một số ít trường hợp lý tưởng. Mô hình thực tế thường chỉ có thể đề cập đến 1 hay 2 nguồn biến động mà thôi.

Có thể minh hoạ cho một số mô hình của cây phân loại 1.4 bằng cách sử dụng mặt cắt của một khúc sông như hình 1.5. Phần bên phải của hình 1.5 mô tả một vùng không-thời gian sử dụng cho các trường hợp nghiên cứu, trong đó trục hoành biểu thị toạ độ không gian, hay khoảng cách dọc sông còn trục tung biểu thị thời gian.

Trường hợp đơn giản nhất (a) là mô hình tất định với thông số tập trung và ổn định. Trong mô hình này dòng vào và dòng ra bằng nhau và không thay đổi theo thời gian và được minh hoạ bởi các chấm cùng kích thước trên các đường thẳng góc tại x=0 và x=L. Trong trường hợp thứ hai (b) là mô hình tất định với thông số tập trung không ổn định. Dòng vào I(t) và dòng ra Q(t) biến đổi theo thời gian và được mô tả bằng các chấm có kích thước khác nhau trên các đường thẳng góc tại x=o và x=L. Trong mô hình với thông số tập trung, không xem xét sự biến thiên theo không gian giữa hai đầu đoạn sông, do đó ta không vẽ các chấm trong vùng này.



Hình 1.5.

Trường hợp thứ ba (c) là mô hình tất định với thông số phân bố không ổn định. Trong mô hình xem xét sự biến thiên của dòng chảy theo không-thời gian và được mô tả bằng các chấm không đều nhau trong mạng lưới các điểm trên mặt không thời gian. Nếu là dòng phân bố ổn định thì các điểm là kích thước như nhau.

Trường hợp thứ tư (d) là mô hình ngẫu nhiên độc lập không-thời gian. Ở đây kết quả ra của hệ thống được biểu thị không phải bằng một chấm đơn lẻ mà bằng một phân bố, trong đó mỗi giá trị có thể nhận của biến được gán một xác suất tương ứng. Các hàm phân bố như nhau theo thời gian.

Trường hợp cuối cùng là mô hình ngẫu nhiên độc lập không gian nhưng tương quan thời gian. Hàm phân bố xác suất thay đổi theo thời gian, phụ thuộc vào giá trị có thể nhận được ở đầu ra.

Thực tế các mô hình rất đa dạng, vì vậy có một cách phân loại khác không mang tính tổng quát như cây phân loại 1.4, nhưng trong từng phạm vi hẹp hơn nó lại tỏ ra khái quát phù hợp với các mô hình cụ thể. Sự phân loại này khác nhau từ mức cây trung gian thứ hai.

Trong mô hình tất định được chia thành mô hình hộp đen (Black box model) và mô hình nhận thức (Conceptual model).

- Mô hình hộp đen là mô hình mà cấu trúc bên trong nó chưa biết hoặc không được mô tả. Hàm truyền(hay hàm ảnh hưởng) của hệ thống được xác định từ dòng ra và dòng vào. Sự khác nhau giữa các mô hình hộp đen là do cách xác định hàm truyền. Các hàm truyền của mô hình hộp đen phản ánh tác động của lưu vực dưới dạng ẩn, do vậy nó không đánh giá được đặc tính riêng biệt của hệ thống đến các quá trình dòng chảy.

- Mô hình nhận thức ra đời sau mô hình hộp đen, nhưng phát triển nhanh và ứng dụng rộng rãi. Mô hình nhận thức xuất phát từ sự tìm hiểu và nhận thức từng thành phần của hệ thống và tiếp cận hệ thống bằng phương pháp mô phỏng từng thành phần (ví dụ quá trình tổn thất, quá trình tập trung nước ...). Cấu trúc của mô hình nhận thức phần nhiều chứa đựng các mô hình thành phần được rút ra từ lí thuyết cơ học chất lỏng và các mô hình thành phần này được xây dựng trên cơ sở phân tích hộp đen. Các mô hình nhận thức có nhiều tham số phản ánh đặc tính vật lí của hệ thống. Sự khác nhau giữa các mô hình nhận thức được thể hiện qua cách thức mô phỏng các qui luật vật lí, những mối quan hệ giữa các nhân tố trong hệ thống và đặc tính của thông số trong mô hình. Các mô hình tất định phổ biến hiện nay phần lớn là các mô hình tất định nhận thức. Do mô tả cấu trúc bên trong của hệ thống thông qua phân tích và nhận thức hệ thống nên các mô hình tất định nhận thức còn gọi là mô hình hộp xám (gray-box model).

Từ mô hình tất định nhận thức người ta lại chia ra thành mô hình với thông số tập chung và mô hình với thông số phân bố .

Trong mô hình tập trung lại có các mô hình tuyến tính và phi tuyến. Mô hình hệ thống tuyến tính là mô hình trong đó hàm lượng trữ là một phương trình tuyến tính có các hệ số là hằng số. Ngược lại mô hình hệ thống phi tuyến là mô hình mà hàm lượng trữ là một hàm phi tuyến .

Trong mô hình tất định với thông số phân bố lại có thể chia ra các mô hình một chiều (1D), hai chiều(2D) hay ba chiều(3D). Mô hình một chiều xem xét sự thay đổi của đặc trưng thuỷ văn chỉ trong không gian một chiều, dọc theo chiều dòng chảy. Mô hình hai chiều tính tới cả sự thay đổi theo phương ngang, phương vuông góc với chiều dòng chảy. Còn mô hình ba chiều xét đến cả sự thay đổi theo chiều sâu tức là xét đến sự biến đổi theo cả không gian ba chiều .

Từ mô hình một chiều lại có thể tách ra thành mô hình sóng động học, sóng khuếch tán hay sóng động lực tuỳ thuộc vào số thành phần (hay số hạng) được xem xét trong phương trình động lượng của hệ thống phương trình vi phân không ổn định của dòng chảy.

Mô hình ngẫu nhiên có thể được chia thành mô hình ngẫu nhiên dừng và mô hình ngẫu nhiên không dừng. Mô hình ngẫu nhiên dừng mô tả quá trình thuỷ văn có các đặc trưng thống kê (hay phân phối xác suất) không thay đổi theo thời gian. Đa số các mô hình ngẫu nhiên thuỷ văn thừa nhận tính dừng để mô phỏng. Còn đối với mô hình ngẫu nhiên không dừng thì hàm phân phối xác suất thay đổi theo thời gian. Có thể coi mô hình dòng chảy tháng theo xích Markov phức là một mô hình ngẫu nhiên không dừng.

Còn có thể có những cách phân loại khác. Tuy nhiên hợp lí nhất đối với đa số các bài toán thuỷ văn là sử dụng mô hình động lực thống kê (vật lí thống kê) căn cứ trên qui luật tất định, nhưng có thông số và hàm vào mang tính ngẫu nhiên, có ý nghĩa xác suất .

1.3. SƠ LƯỢC QUÁ TRÌNH PHÁT TRIỂN MÔ HÌNH TOÁN THỦY VĂN

Vấn đề xây dựng mô hình toán học thủy văn không phải là hoàn toàn mới. Ngay từ khi bắt đầu phát triển của thủy văn học đã có sự liên hệ chặt chẽ với cơ sở toán-lý trong sự tạo thành những mô hình toán cơ bản của hàng loạt các quá trình thủy văn. Có thể coi mô hình về dòng thấm của Green-Amp(1911), đường đơn vị Sherman(1932) và phương pháp tương quan hợp trục của Linsley(1949) là sự những bước đi đầu tiên trong mô hình hoá. Ngày nay các mô hình tất định và ngẫu nhiên đã thu được rất nhiều thành tựu. Các mô hình này đã góp phần đáng kể trong các bài toán tính toán và dự báo thủy văn. Tuy nhiên do sự phức tạp của các quá trình thủy văn, do thiêu những tài liệu thực nghiệm và các khái niệm vật lý chuẩn xác cùng với sự phát triển chưa đầy đủ của các công cụ toán học và phương pháp tính nên nhiều bài toán thủy văn thiếu cơ sở vật lý-toán. Một hướng khác để mô phỏng các quá trình thủy văn là mô hình hoá hệ thống đã ra đời cho phép mô hình hoá nó mà không cần biết chi tiết các quá trình vật lý xảy ra bên trong hệ thống.

Đa số các nghiên cứu thủy văn không nhằm nghiên cứu các quá trình thủy văn nói chung, mà nhằm giải quyết các bài toán công trình riêng biệt. Trong khi đó mỗi một quá trình thủy văn đều khác nhau và việc tổng hợp các kết quả này rất khó khăn và không phải lúc nào cũng có thể làm được.

Việc ra đời của máy tính và phương pháp tính làm tăng mối quan tâm đến việc xây dựng các mô hình toán thủy văn và đưa nó vào sản xuất. Trong những năm gần đây nó đã tạo một hướng nghiên cứu độc lập, có các bài toán và phương pháp riêng của mình. Những bài toán trước đây như giải hệ phương trình vi phân chuyển động không ổn định (hệ phương trình Saint Venant) phải đơn giản hoá thì ngày nay có thể giải đầy đủ bằng các mô hình 1chiều, 2 chiều, 3 chiều. Việc giải hệ thống Saint Venant đã thu hút cả các nhà toán học, những người quan tâm đến ứng dụng thực tế của phương pháp giải bằng số các phương trình vi phân cũng như các nhà thủy văn học, những người muốn đưa các kỹ thuật và phương pháp tính hiện tại vào các tính toán thủy văn.

Lý thuyết hệ thống được Dooge(1964), Nash(1959) và sau đó là Rockwood(1956), Sugawara(1960) cùng với những người khác phát triển. Ở Liên Xô(cũ) được Kalinin-Miliucov nghiên cứu, trong đó đã hình thành những tư tưởng cơ bản của các mô hình tuyến tính với các thông số tập trung. Phương pháp lý thuyết hệ thống rất gần về mặt tư tưởng với các phương pháp truyền thống của thủy văn công trình, nhanh chóng được áp dụng trong thực tế và nhanh chóng có đội ngũ riêng của mình. Với sự phát triển của quan điểm này, hàng loạt mô hình ra đời song song với các mô hình căn cứ trên quan điểm vật lý-toán. Năm 1965 đã hình thành nhóm thủy văn thông số, thống nhất các thuật ngữ và các phương pháp chủ yếu của thủy văn hệ thống.

Với quan điểm coi các số liệu thủy văn là các đại lượng ngẫu nhiên độc lập có phân bố đồng nhất và các hệ thống thủy văn sản sinh ra chúng cũng là một hệ thống ngẫu nhiên độc lập, một loạt các mô hình xác suất ra đời, bắt đầu từ phương pháp tính tần suất của Hazen(1914)và được phát triển bởi Pearson, Kritski-Mekel, Gumbel(1941), Frehet(1927), Chow(1953) và Weibull(1929)...

Sau này với sự phát triển của nghiên cứu thủy văn người ta thấy rằng các số hạng của chuỗi thủy văn không hoàn toàn độc lập mà có tương quan với nhau. Quá trình thủy văn được coi là một quá trình ngẫu nhiên và từ đó hình thành các mô hình mô phỏng quá trình ngẫu nhiên. Ứng dụng mô hình Markov cho các quá trình thủy văn được đưa ra trong các tác phẩm của Kritxki-Menkel(1946), sau đó được phát triển trong một loạt các tác phẩm của Xvanhiđde(1977), Ratkovich(1975)... Những mô hình này khi xác lập đều quan tâm đến bản chất vật lý của các mối liên hệ nội tại của qúa trình thuỷ văn và các thông số được xác định từ chúng. Song song với nó là một loạt các mô hình thông số theo quan điểm hệ thống. Đó là các mô hình ARIMA của Box-Jenkin(1970), mô hình với bước nhảy ngẫu nhiên của Klemes(1974).Các mô hình Thormat-Fiering(1970),Winter(1960). Từ đó đã hình thành một nhóm nghiên cứu riêng lẻ thủy văn ngẫu nhiên.

Năm 1967 đã hình thành nhóm thứ ba trong Uỷ ban mô hình toán thủy văn quốc tế, nhóm thủy văn ngẫu nhiên. Những năm gân đây hình thành các mô hình liên kết giữa tính tất định và ngẫu nhiên, mô tả đầy đủ hơn bức tranh phức tạp về các quá trình thủy văn.

Mô hình toán thủy văn ngày nay được phát triển rộng rãi và ứng dụng trong tất cả các lĩnh vực liên quan đến thủy văn học. Ở Việt Nam, mo hình toán đựoc đưa vào từ cuối những năm 1950 với các mô hình SSAAR(1956), Delta(1970) cho đồng bằng sông Cửu Long.Sau đó là việc sử dụng các mô hình Muskingum(1938), Kalinin-Miliucov(1964), Tank(1968) trong những năm 1960-1980. Trong những năm gần đây rất nhiều mô hình thủy lực-thủy văn tất định, ngẫu nhiên, 1 chiều đến 3 chiều được sử dụng cho các bài toán dự báo, tính toán thủy văn, tính toán thủy lợi, bảo vệ môi trường và thu được những kết quả tốt đẹp

Chương 2
MÔ HÌNH TẤT ĐỊNH

Mặc dù bản chất của dòng chảy là ngẫu nhiên, cũng thừa nhận những giai đoạn hình thành dòng chảy, trong đó những thành phần tất định đóng vai trò chủ yếu. Quá trình hình thành một trận lũ do mưa rào là một thí dụ minh họa. Như vậy, nếu những mô hình ngẫu nhiên là mô hình tạo chuỗi dòng chảy thì mô hình tất định hình thành dòng chảy.

2.1 NGUYÊN TẮC CẤU TRÚC MÔ HÌNH TẤT ĐỊNH

2.1.1 Nguyên tắc mô phỏng

Trong việc mô hình hoá sự hình thành dòng chảy có hai cách tiếp cận:

2.1.1.1. Cách tiếp cận vật lý - toán

Bài toán biến đổi mưa thành dòng chảy có thể được giải cho các khu vực nghiên cứu theo cách sau. Trên cơ sở phân tích tài liệu quan trắc mưa và dòng chảy cho nhiều lưu vực thuộc vùng địa lý - khí hậu khác nhau, tiến hành nghiên cứu chi tiết các hiện tượng vật lý tạo nên quá trình hình thành dòng chảy và xây dựng những quy luật tương ứng, được biểu diễn dưới dạng phương trình, các công thức toán v.v.. Nói chung, các phương trình, các công thức đều chỉ là các cách để biểu diễn ba quy luật chung nhất của vật chất trong trường hợp riêng cụ thể:

a) Bảo toàn vật chất (phương trình liên tục hoặc cần bằng nước),

b) Bảo toàn năng lượng (phương trình cân bằng động lực hay phương trình chuyển động thể hiên nguyên lý Dalambera),

c) Bảo toàn động lượng (phương trình động lượng).

Sau đó, có các đặc trưng địa hình- thuỷ văn mạo lưu vực, độ ẩm ban đầu, quá trình mưa cùng các đặc trưng khí tượng, có thể trực tiếp biến đổi ngay quá trình mưa thành quá trình dòng chảy ở mặt cắt cửa ra lưu vực theo các phương trình và các công thức đã được thiết lập. Trong trường hợp tổng quát, những công thức được biểu diễn dưới dạng các phương trình vi phân đạo hàm riêng thì: Đặc trưng địa hình - thủy địa mạo lưu vực đóng vai trò các thông số phương trình (các hằng số hoặc trong trường hợp chung sẽ biến đổi theo thời gian) quá trình mưa cho chúng ta điều kiện biên, còn trạng thái lưu vực ban đầu. Hệ Saint - Venant cùng với những phương pháp số cụ thể giải nó cho ta một minh hoạ về cách tiếp cận này trong việc mô hình hoá giai đoạn cuối cùng trong sự hình thành dòng chảy- giai đoạn chảy trên bề mặt lưu vực và trong mạng lưới sông.

Lĩnh vực này của mô hình hoá dòng chảy có những đặc thù và phương pháp nghiên cứu riêng biệt không thể thiếu được những tài liệu nghiên cứu cơ bản cùng với những tài liệu nghiên cứu rất chi tiết và tốn kém về địa hình , về các đặc trưng thuỷ địa mạo khu vực, về các đặc trưng diễn biến của mưa theo không gian...

Khước từ sử dụng bộ tài liệu chi tiết về địa hình - địa mạo cùng các đặc trưng khác về lưu vực, chúng ta chỉ có một cách coi lưu vực như là một hệ động lực. Và trong việc mô hình hoá sự hình thành dòng chảy sử dụng cách tiếp cận thông số hoá.

2.1.12. Cách tiếp cận thông số hoá l

Đây là cách tiếp cận thị trường dựa trên việc sử dụng tài liệu quan trắc đồng bộ giữa mưa và dòng chảy. Điều này cho phép lựa chọn các thông số của các biểu tức toán học theo tài liệu đo đạc.

Trong đó, từ những ý niệm vật lý (căn nguyên) sẽ xây dựng cấu trúc chung mô hình, chứa hàng loạt các thông số cùng các giá trị ban đầu của chúng cố gắng xuất phát từ những ý nghĩa vật lý. Sau đó theo tài liệu quan trắc mưa - dòng chảy của nhiều trận lũ trên một lưu vực cụ thể, tiến hành xác định bộ thông số.

Khi mô hình hoá, lưu vực sông hoạt động như một toán tử biến đổi hàm vào q(t) - mô tả lượng nước đến bề mặt lưu vực thành hàm ra Q(t) - mô tả quá trình dòng chảy hình thành. Hai cách tiếp cận trên dẫn đến 2 dạng toán tử lưu vực L1 và L2:

$$Q = L1(Q, q, x, y, z)\{q(x,y,z)\} \quad (2.1)$$

$$z = f(x,y)$$

$$Q = L2(Q,q,t)\{q(t)\} \quad (2..2)$$

Toán tử L2 - cách tiếp cận thông số hoá mô tả sự chuyển đổi hàm vào thành hàm ra không phụ thuộc và từng điểm cụ thể của lưu vực, có nghĩa là loại bỏ sự thay đổi theo không gian các đặc trưng lưu vực. Trong trường hợp này có thể coi các thông số tạp trung tại một điểm. Do đó nhưng mô hình được xây dựng theo cách thông số hoá được gọi là mô hình các thông số tập trung.

Toán tử L1 mô tả sự chuyển đổi có xét sự phân bố không đều theo không gian không nhưng của các đặc trưng lưu vực mà còn cả hàm vào và hàm ra. Đó là những mô hình có thông số rải (phân bố) hay được gọi là những mô hình vật lý - toán.

Các toán tử lưu vực không phụ thuộc hàm vào và hàm ra:

$L(Q, q, t) \Leftrightarrow L(t)$

từ đây có thể rút ra nguyên lý xếp chồng:

$L\{q1(t) + q2(t\} = L\{q1(t)\} + L\{q2(t)\}.$

$L\{ cq(t)\} = cL\{q)t\}$

với những mô hình dừng, toán tử lưu vực không phụ thuộc vào thời gian:

$L(Q,q,t) \Leftrightarrow L(Q,q)$

Nếu mô hình tuyến tính dừng:

$L(Q,q,t) \Leftrightarrow L$

Đây là mô hình đơn giản nhất, được sử dụng trong trường hợp có thông tin gì về các đặc trưng lưu vực.

2.1.2 Cấu trúc mô hình tất định

Những mô hình có thông số tập trung (toán tử lưu vực dạng L2) đến lượt mình lại được chia làm hai loại: Mô hình "hộp đen" và mô hình " quan niệm".

2.1.2.1. Mô hình " hộp đen" .

"Hộp đen" thuật ngữ dùng trong điều khiển học để chỉ những hệ thống mà cấu tạo và các thông số của nó hoàn toàn không rõ ràng, chỉ có thể được xác định trên cơ sở những thông tin vào - ra. Trong thực tế sản xuất, đôi khi xuất hiện tình huống khi cần xây dựng những quan hệ mưa - dòng chảy cũng chỉ có những quan trắc ở đầu vào (mưa) đầu ra (dòng chảy) hệ thống. Những trường hợp này buộc phải coi lưu vực là một "hộp đen" . Tình trạng thiếu thông tin về lưu vực chỉ cho phép xây dựng những mô hình thô sơ nhất, và khi xây dựng chúng người ta cũng hoàn toàn không có thông tin gì về lưu vực ngoài việc coi nó là một hệ thống tuyến tính và dừng. Do vậy, trong thuỷ văn: mô hình "hộp đen" đồng nghĩa với mô hình tuyến tính - dừng.

Lớp mô hình " hộp đen " xuất hiện khá sớm vào thời kỳ đầu của sự phát triển mô hình thuỷ văn tất định. Ngày nay lớp mô hình này chỉ còn tồn tại với tư cách mô tả một giai đoạn cuối trong sự hình thành dòng chảy - giai đoạn chảy: giai đoạn biến đổi lớp cấp nước trên lưu vực thành dòng chảy ở cửa ra.

2.1.2.2. Mô hình quan niệm

Quá trình biến đổi mưa thành dòng chảy - một quá trình phi tuyến phức tạp gồm nhiều giai đoạn. Cùng với sự phát triển của lý thuyết hình thành dòng chảy, mô hình quan niệm ra đời. Có thể định nghĩa mô hình quan niệm là loại mô hình được mô tả bởi một tập hợp các quan hệ toán học, từng quan hệ biểu diễn từng mặt riêng của quá trình, nhưng kết hợp lại chúng mô hình hoá cả quá trình trọn vẹn. Với sự xuất hiện của máy tính điện tử vào giữa những năm 50, lớp mô hình "hộp đen" hoàn toàn lùi bước trước những mô hình "quan niệm" cho phép mô tả đầy đủ hơn, chính xác hơn quá trình " mưa -dòng chảy" được hình thành từ hàng loạt các quá trình thành phần mưa, bốc hơi, điền trũng, thảm thực vật, nước thấm, chảy mặt, sát mặt, ngầm ... Ngày nay, có thể thấy hàng loạt các mô hình quan niệm rất phát triển như mô hình SSARR (Mỹ), TANK (Nhật), STANFORD - 4 (Mỹ), CLS (Ý), GMC (Liên Xô), SMART (Bắc Ailen), GIRARD - 1(Pháp).v.v...

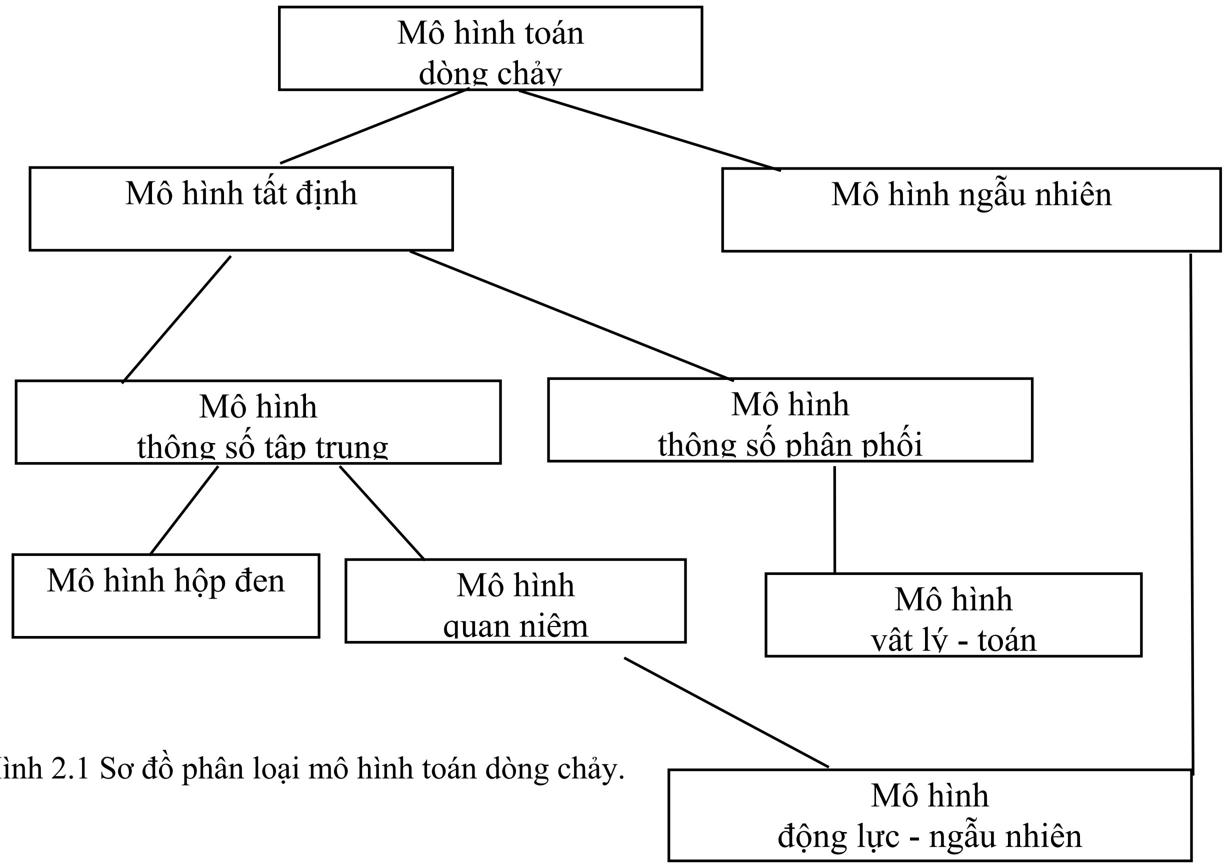
Trong những năm gần đây đã xuất hiện những xu hướng liên kết cách tiếp cận tất định và ngẫu nhiên vào việc mô tả các hiện tượng thuỷ văn. Việc xét tính ngẫu nhiên của các quá trình trong mô hình tất định diễn ra theo 3 phương hướng:

1. Xét sai số tính toán như một quá trình ngẫu nhiên và trở thành một thành phần trong các mô hình tất định.

2. Sử dụng các mô tả xác suất - thống kê (luật phân bố) của các tác động khí tượng - thuỷ văn với tư cách là hàm vào của mô hình tất định.

3. Xét các quy luật phân bố xác suất theo không gian của tác động khí tượng - thuỷ văn vào lưu vực.

Với những tư tưởng này đã hình thành những mô hình động lực - ngẫu nhiên. Do sự phức tạp của vấn đề, lớp mô hình này mới chỉ ở giai đoạn đầu của sự khai sinh. Sự phân loại mô hình nêu trên được trình bày như trên hình 2.1



Hình 2.1 Sơ đồ phân loại mô hình toán dòng chảy.

2.2 NHỮNG NGUYÊN LÝ CHUNG TRONG VIỆC XÂY DỰNG MÔ HÌNH " HỘP ĐEN

Khi xây dựng mô hình "hộp đen" chúng ta hoàn toàn không có thông tin gì về các đặc trưng lưu vực cùng với những quá trình xảy ra trên nó ngoài giả thiết : lưu vực là hệ thống tuyến tính - dừng. Cần làm sáng tỏ, trong những điều kiện nào có thể coi lưu vực hoặc đoạn sông là hệ tuyến tính - dừng?

1. Như phần trên đã nêu để đảm bảo nguyên lý "xếp chồng", cấu tạo hệ thống cùng những đặc trưng của nó không được phụ thuộc vào hàm vào(tác động) và hàm ra (phản ứng). Điều này còn nghĩa rằng: Các đặc trưng thuỷ địa mạo lưu vực và đoạn sông(độ dốc mặt nước, hệ số nhám, tốc độ truyền lũ và thời gian chảy truyền)

không được phụ thuộc vào lưu lượng nước. Như vậy hệ thủy văn không phải là tuyến tính, nhưng giả thuyết về tính tuyến tính của nó trong nhiều trường hợp tỏ ra rất hữu ích với tư cách là sự xấp xỉ ban đầu.

2. Nếu như thời gian của quá trình hình thành dòng chảy nhỏ hơn nhiều so với khoảng thời gian trong đó những đặc trưng của lưu vực hay đoạn sông có những thay đổi đáng kể thì có thể coi lưu vực (đoạn sông) là một hệ dừng (với nghĩa là không thay đổi theo thời gian).

Trong trường hợp tổng quát, hoạt động của một hệ động lực tuyến tính - dừng được mô tả bởi những phương trình vi phân thường, liên hệ phản ứng hệ thống Q(t) với tác động q(t).

$$\alpha_n \frac{d^n Q}{dt^n} + \cdots + \alpha_1 \frac{dQ}{dt} + \alpha_0 Q(t) = \beta_n \frac{d^n q}{dt^n} + \cdots + \beta_1 \frac{dq}{dt} + \beta_0 Q(t) \quad (2.3)$$

Các hệ số αi, βi các hằng số mô tả đặc trưng của lưu vực (đoạn sông).

Như vậy, công cụ toán học để mô tả và phân tích những mô hình hộp đen và lý thuyết phương trình vi phân thường tuyến tính. Trong khi xây dựng các mô hình "hộp đen" về dòng chảy, các tác giả thường kết hợp sự mô tả toán học với sự tương tự vật lý thông qua các nguyên tố vật lý. Hai nguyên tố vật lý cơ bản nhất, có mặt hầu hết trong các mô hình "hộp đen" khác nhau là: Bể chứa tuyến tính Ai và kênh tuyến tính.

1. Bể chứa tuyến tính Ai, đó là bể chứa tượng trưng có lưu lượng chảy ra tỷ lệ thuận với thể tích nước trong đó:

$$Q_i = C_i W_i \quad (2.4)$$

Như sẽ thấy rõ sau này, hoạt động của bể chứa tuyến tính luôn luôn có được sự mô tả bởi tính luôn luôn có thể được mô tả bởi toán tử cơ bản có dạng :

$$A_i = a_i \frac{d}{dt} + b_i \quad (2.5)$$

Trong đó, ai và bi là các đặc trưng của bể chứa. Một bể chứa tuyến tính có thể coi một hoặc vài cửa vào, một hoặc vài cửa ra. Các mô hình dòng chảy khác nhau cũng một phần do sự do sự kết hợp khác nhau của bể chứa tuyến tính.

Mô hình dòng chảy vùng núi do nhóm nghiên cứu I.M. Đenhixốp đề xuất hai bể chứa thẳng đứng. Trong mô hình TANK, M.Sugawara sử dụng nhiều bể mắc nối tiếp - song song . Mô hình Kalinhin -Miliukốp - Nash gồm nhiều bể chứa tuyến tính mắc nối tiếp.

2. Kênh tuyến tính: đó là kênh tượng trưng có chiều dài x với thời gian chảy truyền τ không đổi với mọi cấp lưu lượng Q. Như vậy, khi lan truyền trên kênh tuyến tính, hình dáng đường quá trình lưu lượng không bị biến dạng. Có nghĩa, nếu hàm vào q = f(t), thì hàm ra:

$$Q=f(t-\tau)$$

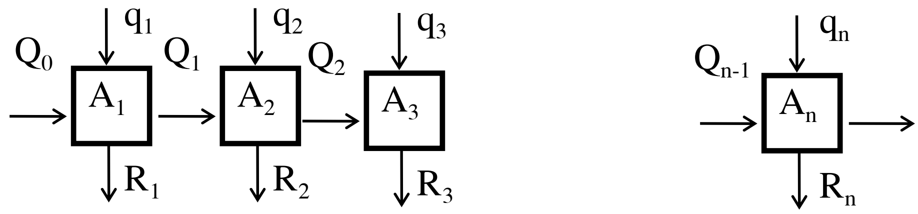
Bể tuyến tính có tác dụng làm biến dạng (bẹt) sóng lũ, kênh tuyến tính có tác dụng dịch chuyển sóng lũ. Đó là hai nguyên tố cơ bản nhất tạo nên mô hình khác nhau. Trong mô hình của Dooge J.C.I. Các bể tuyến tính và các kênh tuyến tính được mắc xen kẽ xen từng đôi một.

Diện tích lưu vực được chia thành n phần bởi các đường đẳng thời. Từng diện tích bộ phận được coi là một cặp kênh tuyến tính và bể tuyến tính. Như vậy, lượng nước đến bể thứ i gồm 2 bộ phận : dòng chảy từ bể (i-1) qua kênh tuyến tính vào bể i và lượng mưa rơi rực tiếp xuống bể i. Mô hình của Dooge trực tiếp hoàn thiện mô hình của Nash.

Khi xây dựng mô hình tuỳ thuộc vào khả năng điều tiết của lưu vực cùng sự cảm nhận tinh tế của người xây dựng, để quyết định số bể chứa, kiểu kết hợp giữa chúng và với các kênh tuyến tính. Nên lưu ý lựu chọn cấu trúc đơn giản nhất mà vẫn đảm bảo độ chính xác. Sự phức tạp hoá mô hình đôi khi tỏ ra thừa và dẫn đến luỹ tích sai số tính toán. Trong việc xác định bộ thông số. Mô hình phức tạp, nhiều thông số, sẽ thường gặp phải hiệu ứng "rà quá kỹ" khi xây dựng mô hình, hoàn toàn có thể sử dụng các loại bể chứa phi tuyến và kênh phi tuyến. Trong mục này chỉ trình bày những kỹ thuật cơ bản nhất của việc xây dựng lớp mô hình tuyến tính - dừng.

2.2.1. Một số cấu trúc mô hình tuyến tính cơ bản

1. Để mô phỏng tác dụng điều tiết của lòng sông trên đoạn sông có lượng nhập khu giữa, người ta sử dụng kỹ thuật mặc nối tiếp các bể tuyến tính.



Hình 2..2 Sơ đồ mắc nối tiếp các bể tuyến tính

Hoạt động của bể tuyến tính này được mô tả bởi phương trình vi phân dạng:

$$\frac{dW_i}{dt} = Q_{i-1} + q_i - Q_i - R_i \tag{2.6}$$

Các lưu lượng ra khỏi bể tỷ lệ thuận với lượng nước trong bể

$$Q_i = C_i W_i \tag{2.7}$$

$$R_i = \gamma_i W_i \tag{2.8}$$

từ (2.7) và (2.8) ta có

$$\frac{dW_i}{dt} = \frac{1}{c_i}\frac{dQ_i}{dt} \tag{2.9}$$

$$R_i = \frac{\gamma_i}{c_i} Q_i \tag{2.10}$$

Thay (2.9), (2.10) vào (2.6)

$$a_i \frac{dQ_1}{dt} + b_i Q_i = Q_{i-1} + q_i \qquad i=1,2,...,n \tag{2.11}$$

với $a_i = \dfrac{1}{c_i}, \qquad bi = 1 + \dfrac{\gamma_i}{c_i}$

Quá trình truyền lũ trên đoạn sông được mô tả bởi hệ n phương trình vi phân :

$$a_1 \frac{dQ_1}{dt} + b_1 Q_1 = Q_0 + q_1$$

$$a_2 \frac{dQ_2}{dt} + b_2 Q_2 = Q_1 + q_2$$

....

$$a_n \frac{dQ_n}{dt} + b_n Q_n = Q_{n-1} + q_n \quad (2.12)$$

Hệ (2.12) tương đương với một phương trình vi phân bậc n. Để đạt được điều đó tiến hành như sau: Giải phương trình thứ hai trong hệ đối với Q1, lấy đạo hàm của nó, thay $\overline{Q}$ và $\frac{dQ_1}{dt}$ tìm được vào phương trình 1 sẽ có:

$$a_1 a_2 \frac{d^2 Q_2}{dt^2} + (a_1 b_2 + a_2 b_1)\frac{dQ_2}{dt} + b_1 b_2 Q_2 =$$
$$= Q_0 + q_1 + \ldots + a_1 \frac{dq_2}{dt} + b_1 q_2 \quad (2.13)$$

hoặc:

$$\left(a_1 \frac{d}{dt} + b_1\right)\left(a_2 \frac{d}{dt} + b_2\right) Q_2 = Q_0 + q_1 + \left(a_1 \frac{d}{dt} + b_1\right) q_2$$

Tương tự giải phương trình thứ ba trong (2.12) đối với Q, lấy đạo hàm bậc 1, bậc 2 đối với Q2 và thế vào (2.13) . Tiếp tục thuật toán này đối với Qn và cuối cùng ta được:

$$\left[\prod_{i=1}^{n} (a_i \frac{d}{dt} + b_i\right] Q = Q_0 + q_1 + \sum_{k=1}^{n-1} \left[\prod_{i=1}^{k} (a_i \frac{d}{dt} + b_i)\right] q_{k+1} \quad (2.14)$$

Như vậy vế trái của phương trình dạng (2.3) luôn có thể đưa về dạng tích của các toán tử Ai dạng (2.4) như trong (2.14)

Trong trường hợp các bể tuyến tính Ai đều như nhau ai=a và bi=b đối với mọi i:

$$(a\frac{d}{dt}+b)^n Q = Q_0 + \sum_{k=0}^{n-1}(a\frac{d}{dt}+b)q_{k+1} \qquad (2.15)$$

Kết hợp với điều kiện lượng nhập khu giữa phân bố đều trên đoạn sông qk=q với mọi k

AnQ=Q0 + q(1+ A + A2 + ... + An-1) (2.16)

với A là toán tử từ (11.4)

Trong trường hợp không có lượng nhập khu giữa qi = 0.

$$\left[\prod_{i=1}^{n}(a_i\frac{d}{dt}+b_i)\right]Q = Q_0 \qquad (2.17)$$

và nếu như các bể tuyến tính như nhau:

$$\left(a\frac{d}{dt}+b\right)^n Q = Q_0 \qquad (2.18)$$

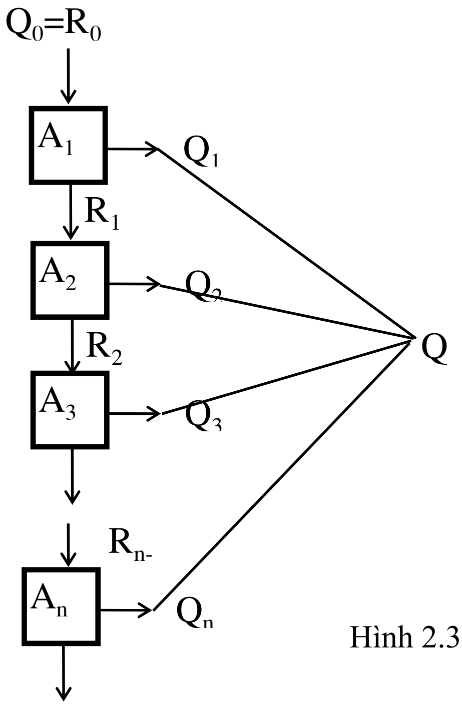
2. Để mô tả tác dụng điều tiết lưu vực thường sử dụng kỹ thuật mắc nối tiếp - song song n bể tuyến tính, tượng trưng cho các tầng đất dẫn nước khác nhau:

Q0 = R0 - lượng cấp nước trên bề mặt lưu vực.

$Q = \sum_{1}^{n} Q_i$ - lưu lượng nước tại mặt cắt cửa ra lưu vực.

Ri lưu lượng ra tại bể Ai nhưng vào bể Ai+1 tượng trưng cho sự thấm.

Qi lưu lượng ra khỏi bể Ai tượng trưng cho dòng chảy mặt.



Hình 2.3 Sơ đồ mắc nối tiếp - song song các bể

Hoạt động của từng bể Ai được mô tả bởi phương trình:

$$\frac{dW_i}{dt} = R_{i-1} - Q_i - R_i \qquad (2.19)$$

$$Q_i = C_i W_i$$
$$R_i = \gamma_i W_i \qquad (2.20)$$

Quá trình điều tiết trên toàn lưu vực được mô tả bởi hệ phương trình tuyến tính

: $$a_i \frac{dQ_i}{dt} + b_i Q_i = Q_{i-1} \qquad \text{i= 1,2,3,..., n} \qquad (2.21)$$

$$a_1 = \frac{1}{c_1}, \qquad b_1 = \frac{c_1 + \gamma_1}{c_1},$$

với $$a_i = \frac{c_{i-1}}{c_i \gamma_{i-1}}, \qquad b_i = \frac{c_{i-1}(c_i + \gamma_i)}{c_i \gamma_{i-1}} \qquad (2.22)$$

Như vậy tương tự thuật toán đã trình bày ở trên có thể viết:

$$\left.\begin{array}{c}
(a_1 \frac{d}{dt} + b_1)Q_1 = Q_0 \\
\left[(a_1 \frac{d}{dt} + b_1)(a_2 \frac{d}{dt} + b_2)\right]Q_2 = Q_0 \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
\left[\prod_{k=1}^{n}(a_k \frac{d}{dt} + b_k)\right]Q_i = Q_0 \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
\left[\prod_{k=1}^{n}(a_k \frac{d}{dt} + b_k)\right]Q_n = Q_0
\end{array}\right\} \qquad (2.23)$$

Nhân hai vế của (n-1) phương trình đầu của (2.23) với toán tử dạng:

$$\prod_{k=i+1}^{n}(a_k \frac{d}{dt} + b_k)$$

rồi tiến hành cộng tất cả các phương trình (2.23) sẽ có dạng:

$$\prod_{k=1}^{n}(a_k \frac{d}{dt} + b_k)(Q_1 + Q_2 + \ldots + Q_n) =$$
$$\left[\prod_{k=2}^{n}(a_k \frac{d}{dt} + b_k) + \prod_{k=3}^{n}(a_k \frac{d}{dt} + b_k) + \ldots + (a_n \frac{d}{dt} + b_n) + 1\right]Q_0 \qquad (2.24)$$

Nhưng vì:

$$Q = \sum_{1}^{n} Q_i$$

có:

$$\left[\prod_{k=1}^{n}(a_k \frac{d}{dt} + b_k)\right]Q = \left[\sum_{j=1}^{n-1}\prod_{k=+1}^{n}(a_k \frac{d}{dt} + b_k + 1)\right]Q_0$$

Trong việc mô phỏng sự điều tiết của lưu vực do mối quan hệ (2.22), các bể chỉ có thể tương tự nhau từ bể thứ hai trở đi:

ai=a; bi=b i=2,3,...,n

Trong trường hợp này:

$$\left[(a_1\frac{d}{dt}+b_1)(a\frac{d}{dt}+b)^n-1\right]Q=\left[\sum_{j=1}^{n}(a\frac{d}{dt}+b)^n-j\right]Q \qquad (2.25)$$

2.2.2 Hàm ảnh hưởng. Biểu thức toán học lớp mô hình tuyến tính

Từ lý thuyết phương trình vi phân tuyến tính tính đạo hàm thường thấy rằng nghiệm của phương trình (2.3) thoả mãn những điều kiện ban đầu: Q(t0)=Q0,Q'(t0)=....=Q0(n-1) có thể biểu diễn dưới dạng:

$$Q(t)=\widetilde{Q}(t)+Q^{\bullet}(t) \qquad (2.26)$$

trong đó:

$\widetilde{Q}(t)$ - nghiệm của phương trình thuần nhất

$Q^{\bullet}(t)$ - nghiệm riêng của phương trình không thuần nhất thoả mãn điều kiện ban đầu bằng 0.

Q(t0) ≡ Q'(t0) ≡ ≡ Q(n-1)(t0) ≡ 0,

Do tính chất tuyến tính $\widetilde{Q}(t)$ có thể biểu diễn dưới dạng một tổ hợp tuyến tính của n nghiệm riêng của phương trình thuần nhất.

$$\widetilde{Q}(t)=\sum_{k=1}^{n}C_k Q_k(t) \qquad (2.27)$$

trong đó Ck - các hằng số được xác định bởi điều kiện ban đầu qua việc giải hệ phương trình đại số tuyến tính sau:

$$\left.\begin{array}{l} C_1Q_1(t_0) + C_2Q_2(t_0)+\ldots+C_nQ_n(t_0) = Q_0 \\ C_1Q'_1(t_0) + C_2Q'_2(t_0)+\ldots+C_nQ'_n(t_0) = Q'_0 \\ \ldots \quad \ldots \quad \ldots\ldots \quad \ldots \quad \ldots \quad \ldots \\ C_1Q^{(n-1)}(t_0) + C_2Q^{(n-1)}(t_0)+\ldots+C_nQ^{(n-1)}(t_0) = Q_0^{(n-1)} \end{array}\right\} \quad (2.28)$$

Định thức ma trận hệ số vế trái là định thức Vronski tại t0:

$$\Delta = \begin{vmatrix} Q_1(t_0) & Q_2(t_0)\ldots & Q_n(t_0) \\ Q'_1(t_0) & Q'_2(t_0)\ldots & Q'_n(t_0) \\ Q_1^{(n-1)}(t_0) & Q_2^{(n-1)}(t_0) & Q_n^{(n-1)}(t_0) \end{vmatrix} \quad (2.29)$$

Do các nghiệm $Q_i(t)$ (i=1,2,...,n)độc lập tuyến tính nên định thức Vronski luôn luôn tồn tại một nghiệm duy nhất có thể xác định theo công thức Crame:

$$C_k = \frac{\Delta_k}{\Delta},$$

trong đó Δk là định thức nhận được từ định thức Vronski sau khi thay cột thứ k trong (2.29) bằng cột các điều kiện ban đầu:

$$\begin{pmatrix} Q_0 \\ Q'_0 \\ \ldots \\ Q_0^{(n-1)} \end{pmatrix}$$

Trong toán học đã chứng minh, với điều kiện ban đầu bằng 0, phương trình phụ trợ của (2.3) có dạng:

$$Q(P) = \frac{L_\beta(P)}{L_\alpha(P)} q(P) \quad (2.30)$$

trong đó: P=a+ib (a>0) - một số phức;

Lα(P)=αnPn+αn-1Pn-1+...+α1P+α0

Lβ(P)=βnPn+βn-1Pn-1+...+β1P+β0

$$Q(P) \Rightarrow Q(t) \text{ và } q(P) \Rightarrow q(t)$$

có nghĩa là Q(P) và q(P) là các tạo hình của Q(t) và q(t) nhận được bằng biến đổi Laplace.

$$Q(P) = \int_0^\infty e^{-P.t} Q(t)dt$$

$$q(P) = \int_0^\infty e^{-P.t} q(t)dt$$

Hàm $P(P) = \dfrac{L_\beta(P)}{L_\alpha(P)}$ được gọi là hàm truyền, và (2.30) được viết dưới dạng:

Q(P)=P(P).q(P) (2.31)

Từ (2.31) suy ra:

$$Q(P) \to \int_0^t P(t-\tau)q(\tau)d\tau$$

và theo định lý về nguyên bản duy nhất ta có:

$$Q(t) = \int_0^t P(t-\tau)q(\tau)d\tau \qquad (2.32)$$

Biểu thức(2.32) được gọi là tích phân Duhamel và đó cũng chính là nghiệm riêng của phương trình vi phân tuyến tính không thuần nhất với các điều kiện ban đầu bằng 0.

$$Q^\bullet(t) = \int_{t_0}^t P(t-\tau)q(\tau)d\tau \qquad (2.33)$$

Hàm P(t-τ) trong (2.32) được gọi là hàm ảnh hưởng và là nguyên bản của hàm truyền P(P).

$$P(t) \leftarrow \frac{L_\beta(P)}{L_\alpha(P)} = P(P)$$

Trong quá trình xây dựng mô hình hàm truyền P(P) luôn luôn có thể xác định được dễ dàng và sau đó sử dụng bảng tra tạo hình - nguyên bản của phép biến đổi Laplace để xác định hàm ảnh hưởng P(t).

Mô hình hàm tuyến tính đều có dạng chung là:

$$Q(t) = \sum_{1}^{n} \frac{\Delta_i}{\Delta} Q(t) + \int_{t_0}^{t} P(t-\tau) q(\tau) d\tau \qquad (2.34)$$

Biểu thức (2.34) là dạng tổng quát của tất cả mô hình "hộp đen". Các mô hình "hộp đen" được phân biệt với nhau bởi:

1. Dạng giải tích hàm ảnh hưởng P(t-τ),

2. Cách xác định hàm ảnh hưởng

3. Cách xét Qi(t).

Với chức năng của mình mô hình "hộp đen" mô tả quá trình chảy điều tiết của lòng dẫn học lưu vực với những tầng đất khác nhau. do vậy ngày nay mô hình "hộp đen" là bộ phận không thể thiếu được trong các mô hình "quan niệm' sự hình thành dòng chảy.

2.3. NGUYÊN LÝ XÂY DỰNG MÔ HÌNH "QUAN NIỆM" DÒNG CHẢY.

Cách tiếp cận trong việc xây dựng mô hình "quan niệm' là cách tiếp cận thông số hoá:

1. Cho dãy các số liệu quan trắc về mưa X(t) và dòng chảy ở mặt cắt cửa ra lưu vực Q(t).

2. Cần tìm toán tử chuyển đổi tốt nhất từ mưa ra dòng chảy.

Cấu trúc của toán tử cùng các thông số của nó, nói chung là không có sẵn.

Tuy nhiên, trong học thuyết dòng chảy đã có những cơ sở lý thuyết và thực nghiệm về sự hình thành dòng chảy nói chung và trên 1 số lưu vực cụ thể. Điều đó dẫn đến hình thành 1 số thông tin về các lớp toán tử cần thiết cùng phạm vi biến đổi các thông số của chúng (lý thuyết thấm, tích đọng, ảnh hưởng của rừng, dòng chảy sườn dốc, chảy ngầm v.v...)

Xây dựng mô hình gồm 2 giai đoạn:

- Thiết lập cấu trúc mô hình

- Xác đinh thông số mô hình

2.3.1. Xây dựng cấu trúc mô hình

Đây là khâu xác định những quan hệ toán học mô tả diễn biến hiện tượng.

Trong công việc này, nhà mô hình phải rất am hiểu hiện tượng, hiểu rõ những tác động chính đến diễn biến hiện tượng và có trí tưởng tượng phong phú để khái quát hoá hiện tượng. Khi thiết lập cấu trúc mô hình hình thành dòng chảy, cần phác thảo sơ đồ khối về từng quá trình thành phần cùng sự tác động tương hỗ giữa chúng.

Trong mô hình STANFORD-4, nước có thể được trao đổi theo hai chiều: đi xuống và đi lên. Với một số mô hình khác, nước chỉ có một chiều đi xuống (mô hình SSARR). Nét chung của các mô hình quan niệm là đều sử dụng các bể chứa để mô tả các dạng tổn thất và điều tiết khác nhau, do vậy, phương trình tính toán chủ đạo trong mô hình là phương trình cân bằng nước. Việc đưa ra bể chứa ngầm vào mô hình cho phép mô hình mô tả được cả phần dòng chảy mùa kiệt.

Nói chung, sự hình thành dòng chảy trên các lưu vực cụ thể rất khác nhau, do vậy không có một mô hình vạn năng nào dùng cho tất cả mọi trường hợp.Nhà thiết kế mô hình phải lắm vững hiện tượng cụ thể để có sự cải biến cần thiết.

Nói chung, khi thiết lập mô hình hình thành dòng chảy cần đề cập và giải quyết những vấn đề sau:

1. Vấn đề mưa trên lưu vực (hàm vào): có cần hiệu chỉnh số liệu mưa tại các điểm đó (bằng thùng hoặc máy tự ghi)? Nếu cần, cách hiệu chỉnh. Có cần hiệu sự phân phối không đều của mưa theo không gian? Nếu cần, cách hiệu chỉnh?

2. Vấn đề tổn thất do thảm thực vật, do tích đọng trên mặt lưu vực, do thấm, cách xét tác động của độ ẩm ban đầu. Những giả thiết nào về diễn biến quá trình thấm, có xét đến đặc tính của tầng thổ nhưỡng? Nếu có, như thế nào?

3. Có xét đến tổn thất do bốc hơi? nếu có, cách xét (với độ chi tiết nào xét đến các yếu tố khí tượng: tốc độ gió, nhiệt độ không khí, độ thiếu hụt bão hoà v.v...).

4. Cách tách quá trình dòng chảy ngầm ra khỏi dòng chảy tổng cộng tại mặt cắt cửa ra lưu vực?

5. Có xét dòng chảy sát mặt(nếu có, cách xét). Có xét lượng nước hồi quy từ tầng thổ nhưỡng vào sông?

6. Có xét tình huống dòng chảy không phải được hình thành lên toàn bộ diện tích lưu vực (có những chỗ trũng khép kín)nếu có, bằng cách tính diện tích hiệu quả?

7. Cách xét chuyển động sóng lũ trong mạng sông-sự giao thoa của sóng lũ trên dòng chính với các sông nhánh, sự bẹt sóng lũ v.v..

8. Bằng cách nào xét được một bộ phận trên đường quá trình lưu lượng được gây ra bởi lượng nước tồn lại của trận lũ trước v.v..

Giải quyết những vấn đề nêu trên, thiết lập những công thức mô tả quá trình, đồng thơi luôn luôn phải suy xét: Những đại lượng nào trong các công thức cho dưới dạng những giá trị số xác định, những đại lượng nào có thể được tính theo những công thức vật lý và những đại lượng nào đóng vai trò thông số cần phải xác định nhờ những tài liệu quan trắc vào - ra. Chỉ sau khi giải quyết những vấn đề nêu trên mới có thể thiết lập một cấu trúc nào đó của mô hình. Cần chú ý rằng mô hình toán dòng chảy là một chỉnh thể thống nhất, các quá trình thành phần liên quan với nhau một cách mật thiết và hữu cơ, do vậy xét sự ảnh hưởng của một quá trình nào đó đến dòng chảy chỉ có thể làm được sau khi đã xây dựng trọn vẹn mô hình. Ngoài ra các nhân tố hình thành dòng chảy rất biến động theo không gian, có cơ chế hoạt động và số liệu quan trắc của một quá trình nào đó tại một điểm, không khi nào có thể chuyên rập khuôn cho toàn khu vực. Vai trò của từng quá trình thành phần biến đổi từ điểm này sang điểm khác, từ lưu vực này sang lưu vực khác. Điều này dẫn đến việc lựa chọn cấu trúc mô hình quan niệm mang tính mò mẫm-cảm nhận. Điều này cũng cắt nghĩa vì sao việc lắp ghép những kết quả nghiên cứu hiện đại về từng quá trình thành phần (mưa, thấm, bốc hơi, điểm trũng, dòng mặt, sát mặt, ngầm v.v...)của nhiều tác giả khác nhau để hòng được 1 mô hình tốt đã thất bại. Điều này cũng cho thấy vì sao các mô hình quan niệm khác xa nhau cả về cấu trúc lẫn số liệu ban đầu sử dụng.

Việc xây dựng mô hình mang đầy tính sáng tạo cùng với việc am hiểu tường tận hiện tượng trên từng lưu vực cụ thể.

2.3.2 Xác định thông số mô hình

Các mô hình thông số tập trung đều chứa đựng nhiều thông số. Cần xác định cách này trên cơ sở những tài liệu quan trắc vào-ra của hệ thống. Về mặt toán học, có hai phương trình thiết lập thông số mô hình: phương pháp tối ưu hoá và phương pháp giải bài toán ngược. Phương pháp thường dùng trong thực tế hiện nay là khử-sai được coi là phương án thô sơ nhất của phương pháp tối ưu hoá

2.3.1.1. Phương pháp tối ưu hoá.

Đây là bài toán thuận, cho biết thông số vào và bộ thông số mô hình, cần xác định hàm ra của hệ thống. Thực chất tối ưu hoá là bài toán điều khiển hệ thống. Mục tiêu điều khiển là hàm ra phải đúng với tín hiệu đo đạc, còn biến điều khiển là chính véc tơ thông số mô hình.

Cần phải xác định biểu thức toán học của mục tiêu:

$$K = \sum_{i=1}^{n} \int_{0}^{T} \left[Q(t) - \tilde{Q}(t,a) \right]^2 fQ(t)dt \rightarrow \min \qquad (2.35\ )$$

Trong đó: n - Tổng số trận lũ, T - thời gian một trận lũ,

$Q(t), \tilde{Q}(t,a)$ - các quá trình đo đạc và tính toán

a=(a1, a2, am) - véc tơ thông số mô hình.

Hàm f(Q(t) được đưa vào nhằm tăng tỷ trọng những tung lộ lớn (đỉnhlũ). Cần xác định véc tơ a để hàm mục tiêu K đạt cực tiểu. Ngày nay đã có nhiều thuật toán tối ưu đủ mạnh để tìm cực trị của những phiếm hàm mục tiêu phức tạp. Một trong những thuật toán thường dùng là thuật toán Rosenbroc . Nhưng ở đây, bản thân những phương pháp toán học không giải quyết sự chính xác của những thông số cũng như sự thành công của quá trình tối ưu hoá. Một lần nữa, chúng ta thấy nổi lên vai trò cùng những kinh nghiệm và sự hiểu biết hiện tượng vật lý của người thiết lập mô hình.

Sau đây trình bày những kinh nghiệm có tính nguyên tắc trong việc điều hành quá trình tối ưu.

a, Nguyên tắc lựa chọn số liệu. Trong quá trình tối ưu, một số thông số tỏ ra không ảnh hưởng gì tới hàm mục tiêu. Nguyên nhân chính của hiện tượng này là trong những số liệu dùng để tối ưu, chưa có những số liệu mà vai trò của thông số này hay thông số khác tỏ ra rõ rệt. Để khắc phục tình hình này, những số liệu dùng trong quá trình tối ưu phải bao gồm những trận lũ có điều kiện hình thành hết sức khác nhau: đủ lớn, đủ nhỏ, đủ dạng.

Độ chính xác việc xác định thông số phụ thuộc nhiều vào độ chính xác, mức đại biểu và khối lượng của những tài liệu ban đầu. Những trận lũ không đủ tin cậy sẽ gây ra những sai lệch đáng kể cho từng thông số riêng biệt. Do vậy, để tối ưu phải chọn những trận lũ có độ tin cậy cao nhất.

b. Nguyên tắc tiến hành: có hai cách tiến hành quá trình tối ưu:

Cách 1: Tối ưu riêng rẽ từng trận lũ, được các bộ thông số khác nhau, sau đó lấy bộ thông số trung bình cho tất cả các trận.

Cách 2: Tiến hành tối ưu đồng thời cho nhiều trận lũ, được một bộ thông số chung cho tất cả các trận lũ. Kinh nghiệm cho thấy hai cách tối ưu này cho kết quả rất khác nhau. Với từng trận lũ, luôn luôn tìm được một thông số thích hợp. Do đặc thù riêng của từng trận lũ, một số thông số có thể bị sai lệch. Điều này dẫn đến các bộ thông số của các trận lũ rất khác nhau.

Để đảm bảo ý nghĩa của các thông số, đảm bảo độ bền vững, ổn định của chúng, để tối ưu phải sử dụng nhiều trận lũ. Kinh nghiệm cho thấy số liệu dùng để tối ưu không ít hơn 5 quá trình dòng chảy khác nhau.

c. Nguyên tắc phức tạp hoá dần mô hình, do giáo sư Kuchmen đề ra. Thực chất của nó là việc tối ưu hoá được tiến hành theo từng giai đoạn. Trong bộ thông số mô hình, trọng lượng của từng thông số không đồng đều nhau, tính chất của các thông số cũng không giống nhau, có thông số ảnh hưởng tới đỉnh, cóp thông số chỉ ảng hưởng đến tổng lượng, có thông số ảnh hưởng tới nhánh lên, có thông số ảnh hưởng tới nhánh xuống. Thật sai lầm nếu đưa tất cả những thông số đó vào tối ưu cùng một lúc.

Việc phức tạp hoá dần cấu trúc mô hình được bắt đầu bằng việc thử nghiệm mô hình đơn giản nhất, bao gồm các thông số tối thiểu. Trên cơ sở đã tối ưu được các thông số đó, mô hình sẽ được chính xác hoá nhờ việc đưa dần thêm các thông số mới, mô tả chính xác thêm hiện tượng. Ở từng giai đoạn,các thông số được tối ưu một cách độc lập trên cơ sở các thông số của giai đoạn trước nhận những trị số ban đầu bằng các trị số đã được tối ưu.

2.3.1.2. Phương pháp giải bài toán ngược.

Đây là bài toán biết các thông tin vào - ra của hệ thống, cần xác định bộ thông số mô hình. Tính chất của bài toán này là phi chỉnh, có nghĩa là những sai số không lớn lắm của số liệu ban đầu (dùng để giải bài toán ngược) sẽ dẫn đến những sai số rất lớn của những đại lượng cần xác định. Thí dụ khi giải bài toán thuận, những đặc trưng của lưu vực (độ dốc, sườn dốc, khả năng thấm của đất, thảm thực vật, địa hình bề mặt lưu vực v.v) rất biến động theo không gian và chúng cần phải được trung bình hoá theo một cách nào đó và cách trung bình hoá này dù sao cũng ít ảnh hưởng tới kết quả tính toán - dòng chảy ở mặt cắt cửa ra lưu vực. Khi giải bài toán ngược, những thay đổi nhỏ trong số liệu ban đầu (quá trình dòng chảy) có thể tương ứng với những thay đổi rất lớn của các đặc trưng lưu vực, do vậy cũng ảnh hưởng rất lớn đến các thông số mô hình.

Trong những năm 70, những nhà toán học Xô viết Tikhônốp, Lavrenchev, Ivanov đã xây dựng lí thuyết bài toán phi chỉnh. Nhưng công trình toán học này mới chỉ dừng ở việc giải phương trình Volte bậc một. Giáo sư Kuchmen đã vận dụng lí thuyết này trong việc xác định các thông số của hàm ảnh hưởng Kalinhin-Miulikốp-Nash.

Như vậy, lý thuyết toán phi chỉnh mới chỉ áp dụng được trong mô hình tuyến tính đơn giản nhất, vận dụng những mô hình đơn giản quan niệm, những thành tựu trên mới nhất của lý thuyết này chưa đáp ứng được.

2.4. CÁC PHƯƠNG PHÁP XÁC ĐỊNH THÔNG SỐ MÔ HÌNH

Việc xác định các thông số của mô hình toán học rất quan trọng và ảnh hưởng trực tiếp đến kết quả tính toán. Mô hình tính toán dù đã áp dụng ở một số lưu vực cho kết quả rất tốt, nhưng rất có thể áp dụng được ở lưu vực chúng ta đang cần tính toán, nếu như chúng ta không tìm đúng giá trị các thông số của mô hình với những mô hình ít thông số, việc xác định các thông số tối ưu có thể làm bằng tay kết hợp với đồ thị, ví dụ tìm hai thông số x, k của phương pháp Muskingum) như khi thông số của mô hình tăng lên với hàng chục thông số thì việc tính toán các thông số tối ưu sẽ được thực hiện trên máy tính điện tử.

Mô hình hoá - đó là một phương pháp khoa học đầy hiệu lực giúp con người xâm nhập sâu vào bản chất của những hiện tượng tự nhiên hoặc xã hội phức tạp. Mục đích mô hình hoá là tạo dựng hiện tượng sao cho thông qua việc nghiên cứu nó, con người thu nhận được những thông tin mới cần thiết. Nếu việc dựng hiện tượng được thực hiện bởi tập hợp các hệ thức toán học (phương trình - bất đẳng thức, điều kiện lôgic, toán tử...) chúng ta có mô hình toán hiện tượng đó.

Trong 30 năm gần đây, đã diễn ra sự phát triển sâu rộng việc mô hình hoá những hiện tượng và hệ thống tự nhiên khác nhau. Mô hình hoá dòng chảy cũng nằm trong trào lưu đó. Ở nhiều nước đã hoàn thành công việc đồ sộ về xây dựng các mô hình toán dòng chảy. Vấn đề mô hình hoá dòng chảy được thảo luận trên nhiều hội nghị quốc tế. Số xuất bản về mô hình hoá dòng chảy đã lên đến con số vài trăm.

Trong những vấn đề then chốt của tính toán thuỷ văn là luôn luôn đánh giá lượng dòng chảy vì một lý do nào đó không trực tiếp đo đạc được. Khi thiết kế hồ nước hoặc một hệ thống thuỷ lợi, ngành thuỷ văn luôn luôn phải đánh giá " chuỗi dòng chảy tương lai ra sao, bao gồm những tổ hợp nhóm năm nhiều nước, ít nước thế nào, khả năng dòng chảy cực đoan là bao nhiêu. v.v." Chỉ khi có lời giải cho những câu hỏi này, chúng ta mới có thể đề xuất mô hình, kích thước công trình cần xây dựng. Không phải ngẫu nhiên mà hai nhà thuỷ lợi Xô Viết nổi tiếng X.L. Kristky và M.F. Menkel đã phát biểu " bản chất kinh tế nước này nằm ngay trong quá trình dòng chảy". Nhà quản lý thuỷ lợi và hệ thống thuỷ lợi luôn luôn phải băn khoăn, " có thể chờ đón dòng chảy bằng bao nhiêu trong một vài ngày tới". Dự đoán chính xác điều này nâng cao đáng kể hiệu quả hoạt động của công trình. Điểm chung của các vấn đề nêu trên là nhà thuỷ văn luôn luôn phải đánh giá " có thể chờ đợi những gì ở tự nhiên?" Tóm lại, ta cần phải mô hình hoá những hiện tượng thuỷ văn.

Mô hình hoá dòng chảy - đó là chế tạo dòng chảy, còn mô hình toán- quy trình, công nghệ của việc chế tạo đó. Cần khẳng định một điều :" Mô hình toán không thể nào trùng hợp hoàn toàn với mô hình thực, (hiện tượng)". Do vậy, mô hình toán hoàn toàn không phụ thuộc đơn trị vào hiện tượng nghiên cứu. Điều này cắt nghĩa vì sao trong vài chục năm gần đây đã ra đời hàng chục mô hình dòng chảy cùng mô phỏng một hiện tượng.

Nói chung, việc giải bài toán tối ưu gồm 3 giai đoạn :

1. Lập mô hình toán hoặc để mô tả các quá trình thực tế

2. Lựa chọn hàm mục tiêu, tức là chọn tiêu chuẩn đánh giá kết quả.

3. Xác định các giá trị tối ưu của các thông số.

Giai đoạn đầu đã được xét ở các tiết trước, bây giờ chúng ta nghiên cứu tiếp giai đoạn cuối.

2.4.1. Các tiêu chuẩn đánh giá mô hình

Hiện nay tiêu chuẩn đánh giá mô hình được công nhận là kết quả tính toán theo mô hình cần phải phù hợp với quan trắc kiểm nghiệm, độ nhạy của mô hình phải tốt. Hay sử dụng nhất là hàm mục tiêu.

Hàm mục tiêu được dùng phổ biến nhất trong thuỷ văn có dạng :

$$F = \sum_{i=1}^{n} (Q_d - Q_t)_i^2 \qquad (2.37).$$

Với(Qđ-Qt) là chênh lệch giữa giá trị đo và giá trị tính toán ở thời điểm t=i.Δt với i= 1,2,3...n.

Đánh giá theo hàm mục tiêu dạng (2.37) rất đơn giản, dễ dàng nhưng có nhược điểm là nó coi sai số tính toán gây ra bất kì ở thời điểm nào cũng có ý nghĩa như nhau. Thực tế khi tính toán lũ, những sai số gây ra ở phần thấp không quan trọng lắm, còn sai số gây ra ở phần đỉnh lũ thì gây tác hại lớn hơn, do đó người ta chọn hàm mục tiêu có dạng :

$$F = \sum_{i=1}^{n}\left[\left(\frac{1}{m}\sum_{j=1}^{m}(Q_d - Q_t)_j^2 + 2(Q_{dm} - Q_{tm})^2 + 5(T_d - T_t)\right)\right]_i \quad (2.38)$$

Hoặc có dạng :

$$F = \sum_{i=1}^{n}\left[\frac{|Q_{dm} - Q_{tm}|}{Q_{dm}} + \frac{|T_d - T_t|}{T_d} + \frac{|L_d - L_t|}{L_d}\right]_i \quad (2.39)$$

Trong đó i là số trận lũ được tính i= 1,2...n còn j là số thời đoạn tính toán trong 1 trận lũ j= 1,2...m.

(Qđ-Qt)là chênh lệch giữa lưu lượng thực đo Qđ và lưu lượng tính toán Qt ở thời điểm t=jΔt tính từ khi bắt đầu trận lũ. Qdm là lưu lượng đỉnh lũ thực đo, còn Qtm là lưu lượng đỉnh lũ tính toán.

Td, Tt tương ứng là thời gian lũ thực đo và tính toán .

Lđ,Lt là thời gian kéo dài của trận lũ thực đo và tính toán.

Nói chung tất cả hàm mục tiêu sử dụng trong thuỷ văn đều là phi tuyến của các thông số, do đó việc lựa chọn các thông số tối ưu thường phải tính qua nhiều lần lặp.

2.4.2. Lựa chọn thông số tối ưu

Có hai phương pháp thường hay sử dụng nhất:

2.4.2.1 Phương pháp dò tìm theo hướng dốc nhất

Cho hàm mục tiêu F với n thông số : x1, x2,..., xn.

F = F(x1, x2,..., xn) = F(x).

Để cho gọn ta dùng toán tử ∇. Nếu f là một hàm số nào đó trong không gian ba chiều x,y,z thì ∇f là một vectơ.

$$\nabla f = \frac{\partial f}{\partial x} i + \frac{\partial f}{\partial y} j + \frac{\partial f}{\partial z} k$$

với i,j,k là ba đơn vị chỉ phương các trục 0x, 0y, 0z trong hệ trục toạ độ Đề các.Hàm mục tiêu F có n thông số nên nó được biểu diễn trong không gian n chiều. Người ta đã chứng minh rằng nếu như hàm mục tiêu F là liên tục và ∇F tại Xk là xác định thì vectơ ∇F(Xk) biểu thị phương ngắn nhất đi về phía cực trị của hàm F(x).

Quá trình tìm thông số để hàm F(x) nhỏ nhất đã trình bày ở phần trước.

2.4.2.2 Theo phương pháp Rosenbroc

Phương pháp này công bố vào năm 1969 và đang được ứng dụng rộng rãi trong nhiều ngành khác nhau.

Nội dung của thuật toán là xét hàm mục tiêu dưới dạng ma trận n chiều từ đó giải ma trận tìm định thức phù hợp qua các phép tính lặp để lựa chọn các thông số để hàm mục tiêu F(x) đạt giá trị nhỏ nhất.

2.5 GIỚI THIỆU CÁC MÔ HÌNH TẤT ĐỊNH THÔNG DỤNG

2.5.1. Mô hình Kalinhin - Miliukốp - Nash

Năm 1958, khi nghiên cứu sự lan truyền sóng xả ở hạ lưu các trạm thuỷ điện, G.P.Kalinhin và P.I.Miliukov đã chia đoạn sông ra n đoạn nhỏ dưới tên gọi "các đoạn sông đặc trưng". Các đoạn sông đặc trưng được chọn có độ dài sao cho tồn tại mối quan hệ đơn trị tuyến tính giữa lượng nước trong nó với lưu lượng chảy ra. Như vậy thực chất "đoạn sông đặc trưng" là một bể tuyến tính, mà cơ chế hoạt động được mô tả bởi:

$$\frac{dW_i}{dt} = Q_{i-1} - Q_i$$

$$W_i = \tau_i Q_i$$

trong đó τi - thông số mang ý nghĩa thời gian chảy truyền trên "đoạn sông chảy truyền đặc trưng thứ i"

Hai phương trình trên tương đương với một phương trình:

$$\tau_i \frac{dQ_i}{dt} + Q_i = Q_{i-1}$$

Như vậy toán tử Ai trong trường hợp này có dạng:

$$A_i = \tau_i \frac{d}{dt} + 1$$ với $ai = \tau i$, $bi=1$

Mắc nối tiếp n "đoạn sông đặc trưng" tương tự nhau, phương trình (10.17)trở thành:

$$(\tau_1 \frac{d}{dt} + 1)^n Q = Q_0$$ với $\tau i=\tau 1$ và $bi=1$;

Các nghiệm riêng của phương trình thuần nhất có dạng:

$$Q_i(t) = t^{i-1} e^{-\frac{1}{\tau_1}}$$

và hàm ảnh hưởng trở thành:

$$P(t-\tau) = \left(\frac{1}{\tau_1 (n-1)!}\right)\left(\frac{t-\tau}{\tau_1}\right)^{n-1} e^{-\frac{t-\tau}{\tau_1}} \tag{2.40}$$

Công thức tương tự cũng được Nash tìm ra khi giả thiết rằng lưu vực được cấu tạo từ n bể chứa tuyến tính với quan hệ đơn trị - tuyến tính giữa thể tích nước và lưu lượng.

Như đã phân tích, hàm ảnh hưởng Kalinhin - Miliucốp - Nash có hai thông số n và τ là trường hợp riêng của hàm ảnh hưởng 3 thông số. Việc đưa thêm thông số b vào làm ảnh hưởng "dẻo" hơn, ngoài việc dễ thích nghi với việc xét tác dụng điều tiết của lóng sông còn khả năng xét được cán cân nước (các tổn thất bốc hơi, mất nước...).

2.5.1.1. Đường lưu lượng đơn vị.

Phương pháp lần đầu tiên do Sherman đề nghị vào năm 1932 , sau này được nhiều tác giả khác phát triển và hoàn thiện. Nội dung của phương pháp dựa trên 3 luận điểm:

a. Đường quá trình lưu lượng, được hình thành từ lượng mưa hiệu quả 1 đin (25,4 mm) rơi đều trên khắp khu vực trong một đơn vị thời gian, là đặc trưng không đổi của một khu vực (Đường quá trình đó được gọi là đường lưu lượng đơn vị).

b. Đường quá trình lưu lượng, được hình thành từ n đin rơi đều trên khắp lưu vực trong một đơn vị thời gian , có thể nhận được bằng cách nhân tung độ đường lưu lượng đơn vị với n.

c. Đường quá trình lưu lượng, được hình thành từ lượng mưa hiệu quả rơi đều trên khắp lưu vực trong 1 số đơn vị thời gian, có thể nhận được bằng cách cộng các đường quá trình được hình thành do lượng mưa từng đơn vị thời gian.

Phân tích 3 luận điểm trên thấy rằng chúng hoàn toàn tương đương với nguyên lý xếp chồng và việc tính dòng chảy tại mặt cắt cửa ra từ quá trình mưa hiệu quả với điều kiện đơn vị thời gian $\Delta t \rightarrow 0$ hoàn toàn theo biểu thức:

$$Q(t) = \int_{t_0}^{t} P(t-\tau)q(\tau)d\tau$$

Trong đó $P(t-\tau)$ - đường lưu lượng đơn vị ; $q(\tau)$ - quá trình mưa hiệu quả.

Như vậy, thực chất đường quá trình lưu lượng đơn vị là hình ảnh của hàm ảnh hưởng trong mô hình "hộp đen" và chúng được phân biệt với các mô hình "hộp đen" khác bởi tính độc đáo riêng biệt trong việc xác định hàm ảnh hưởng thông qua đờng lưu lượng đơn vị.

Cách đơn giản nhất xác định đường lưu lượng đơn vị được rút ra từ chính định nghĩa của nó: Chọn những trận lũ do lượng mưa rơi đều trong một đơn vị thời gian, rồi chia từng tung độ cho tổng lượng lũ.

2.5.2 Mô hình TANK

Mô hình TANK ra đời năm 1956 tại trung tâm quốc gia phòng chống lũ lụt Nhật, tác giả là M. Sugawar. Từ đó dến nay mô hình được hoàn thiện dần và ứng dụng rộng rãi nhiều nơi trên thế giới.

2.5.2.1 Cấu trúc mô hình Tank

Lưu vực được diễn tả như một chuỗi các bể chứa sắp xếp theo 2 phương thẳng đứng và nằm ngang. Giả thiết cơ bản của mô hình là dòng chảy cũng như dòng thấm và các hàm số của lượng nước trữ trong các tầng đất. Mô hình có hai dạng cấu trúc đơn và kép.

1. Mô hình TANK đơn

Dạng này không xét sự biến đổi của độ ẩm đất theo không gian, phù hợp với những lưu vực nhỏ trong vùng ẩm ướt quanh năm.

Lưu vực được diễn tả bởi bốn bể chứa xếp theo chiều thẳng đứng. Mỗi bể chứa có một hoặc một vài của ra ở thành bên và một của ra ở đáy. Lương mưa rơi xuống mặt đất đi vào bể trên cùng, Sau khi khấu trừ tổn thất bốc hơi một phần sẽ thấm xuống bể dưới theo cửa ra ở đáy, một phần cung cấp cho dòng chảy trong sông theo các cửa ra ở thành bên.

Quan hệ giữa lượng dòng chảy qua các cửa với lượng ẩm trong các bể là tuyến tính:

$$Y=\beta(X-H); \qquad (2.41)$$

$$Y0=\alpha.X \qquad (2.42)$$

Trong đó:β,α -hệ số của ra thành bên và đáy, H- độ cao cửa ra thành bên.

Theo cấu trúc trên, mô hình TANK mô phỏng cấu trúc ẩm trong các tầng đất của lưu vực. Lượng dòng chảy hình thành từ các bể thể hiện đặc tính các thành phần dòng chảy mặt sát mặt và dòng chảy ngầm . Dòng chảy hình thành từ tất cả các bể chứa mô tả sự biến dạng dòng chảy do tác dụng điều tiết của dòng sông là lớp nước có sẵn ban đầu trong sông.

2. Hệ thức cơ bản của mô hình

<u>a, Mưa bình quân lưu vực (P)</u>

$$P = \sum_{i=1}^{n} W_i . x_1 / \sum_{i=1}^{n} W_i \qquad (2.43)$$

Trong đó: n-số điểm đo mưa; Xi lượng mưa tại điểm thứ i, Wi-trọng số của điểm mưa thứ i. Theo M.Sugawara Wi sẽ được trọn là một trong bốn số sau: 0,25; 0,5;0,75;1,0.

<u>b, Bốc hơi lưu vực (E)</u>

$$E = \begin{cases} 0{,}8EVT & Khi\quad XA - PS - E \geq 0 \\ 0{,}75(0{,}8EVT - h_f) + h_f & Khi\quad XA - PS - E < 0 \\ & va\quad XA - PS - H_f > 0 \\ 0{,}6EVT & XA < PS \end{cases} \qquad (2.44)$$

c, Cơ cấu truyền ẩm bể chứa trên cùng được chia làm hai phần:trên và dưới, giữa chúng xảy ra sự trao đổi ẩm. Tốc độ truyền ẩm từ dưới lên T1 và trên xuống T2 được tính theo công thức:

$$T_1 = TB_0 + (1 - \frac{XA}{PS})TB \qquad (2.45)$$

$$T_2 = TC_0 + (1 - \frac{XS}{SS})TC \qquad (2.46)$$

Trong đó: XS,SS - lượng ẩm thực và lượng ẩm bão hoà phần dưới bể A, TBo,TB, TCo, TC-các thông số truyền ẩm, theo MSugawar chúng nhân những giá trị:

TB=TB0 = 3 mm/ ngày đêm

TC = 1mm/ ngày đêm

TC0 =0,5mm/ ngày đêm

d) Dòng chảy từ bể A. Lượng nước đi vào bể A là mưa (P). Dòng chảy qua các cửa bên(YA1, YA2)và của đáy (YA0) được xác định theo các công thức sau:

$$H_f XA+ P-PS \quad (2.47)$$

$$YA_0=H_fA_0 \quad (2.48)$$

$$YA_1 = \begin{cases} (H_f - HA_1); & khi \quad H_f > HA_1 \\ 0 \quad khi & H_f \le HA_1 \end{cases} \quad (2.49)$$

3.Phát triển mô hình TANK trên nền tảng học thuyết độ ẩm đất và học thuyết dòng chảy sườn dốc.

Như các mô hình nhận thức khác, mô hình Tank chứa một lượng thông số khí hậuá lớn. Trong tác phẩm của M.Sugawar những thông số này chưa được miêu tả về mặt vật lý. Do vậy, như K.Linsley nhận định mô hình chỉ có thể được thiết lập cho một lưu vực sau nhiều lần thử sai. Điều này đòi hỏi người sử dụng phải có đủ kinh nghiệm và có mức am hiểu mô hình nhất định. Phần này giới thiệu những hoàn thiện mô hình về mặt vật lý, nhằm giúp người sử dụng lựa chọn thông số có cơ sở và dễ dàng hơn.

Bể A mô phỏng bề mặt lưu vực và các tầng đất trong vùng thoáng, trong bể A có đặt ra những mức ẩm khác nhau của lưu vực (HS, HA3, HA2, HA1, PS, SS).

Trong quá trình chuyển động trên mặt lưu vực hướng về lòng sông một phần nước được giữ lại tạm thời trên sườn dốc.

Hiển nhiên có thể giả định rằng những phần khác nhau trong bể A mô phỏng những dạng trữ nước khác nhau trên mặt sườn dốc.

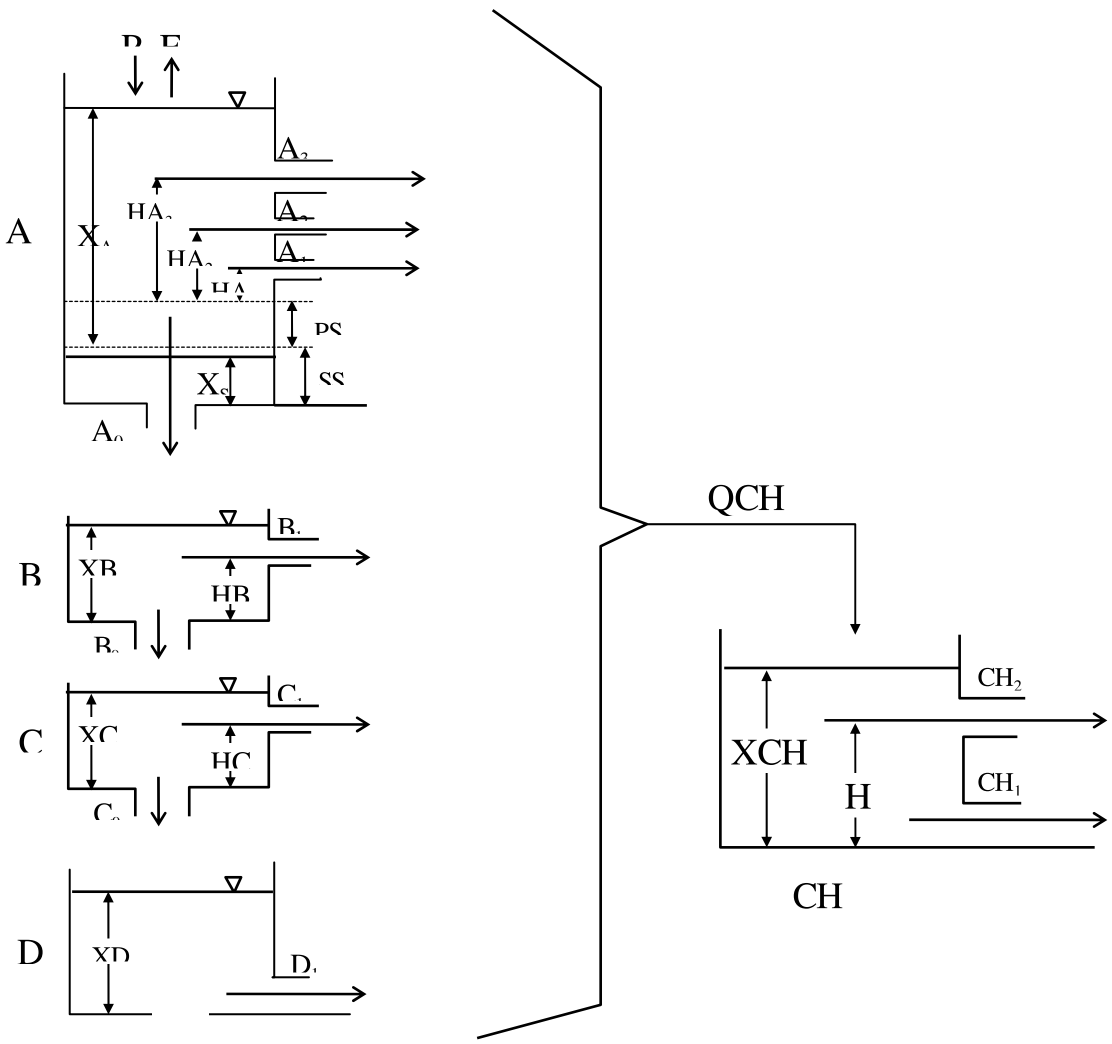
Theo các kết quả thí nghiệm của I.X. Vaxiliep và A.P. Ivanop, sau khi tưới bão hoà cho đất, phân phối độ ẩm theo chiều thẳng đứng có dạng như sau: phần dưới của tầng thổ nhưỡng có độ ẩm khá cao, gần đạt độ ẩm toàn phần (ĐATP), vì rằng nó thuộc tầng mao dẫn. Lên trên, độ ẩm giảm dần và cách mặt thoáng của nước ngầm 1 khoảng nào đó (càng lớn khi thành phần hạt càng nặng), độ ẩm đạt một trị số nhỏ nhất và không đổi độ ẩm đồngruộng (ĐAĐR). Nước chứa trong tầng thổ nhưỡng khi độ ẩm chưa đạt đến độ ẩm đồng ruộng luôn ở trong trạng thái treo và mất khả năng chảy xuống dưới.

Dường như, lượng ẩm chứa trong tầng thổ nhưỡng bão hoà đến độ ẩm đồng ruộng không có khả năng di chuyển. Nhưng thực tế không như vậy. Các kết quả nghên cứu của A.F. Bonsacop, M.M. Abramôva khẳng định trong quá trình bốc hơi, lượng ẩm treo chuyển động lên trên thành dòng, có nghĩa là có tính liên tục. Tính liên tục tồn tại không chỉ với độ ẩm đồng ruộng mà còn có thể nhỏ hơn nhiều. Nhưng chỉ đến một giới hạn nhất định. M,M. Abramôva gọi độ ẩm mà lượng ẩm treo mất khả năng di chuyển lên trên dưới tác dụng của bốc hơi là độ ẩm gián đoạn mao dẫn hay còn gọi là độ ẩm cây héo (ĐACH).

Giả định "phần dưới" của bể A (hình 4.5) mô phỏng tầng đất từ sát mặt sườn dốc đến giới hạn trên của tầng mao dẫn (TMD). Đó là vùng độ ẩm treo. Bản chất vật lí của thông số SS - độ ẩm đồng ruộng (ĐAĐR). Bản chất của lượng ẩm XS - nước mao dẫn. Cơ chế duy nhất tiêu hao được lượng ẩm XS là bốc hơi:

$$(DACH) \le XS \le SS \le (DADR) \quad (2.50)$$

Hiệu số SS -XS xác định lượng tổn thất không hoàn lại do đất giữ, và được thực hiện bởi quá trình truyền ẩm từ trên xuống T2. Bản chất quá trình là giai đoạn đầu của quá trình thấm giai đoạn thấm không ổn định . Giai đoạn này diễn ra khá nhanh. Như vậy quá trình T2 chỉ là quá trình truyền ẩm từ tầng trên xuống tầng dưới của bể A và kết thúc khi tầng dưới đạt đến độ ẩm đồng ruộng sau đó là quá trình thấm ổn định được thực hiện qua các cửa đáy ở các bể. Bản chất các lượng ẩm XB, XC, XD nước trọng lực.



Hình 2.4 Mô hình TANK đơn

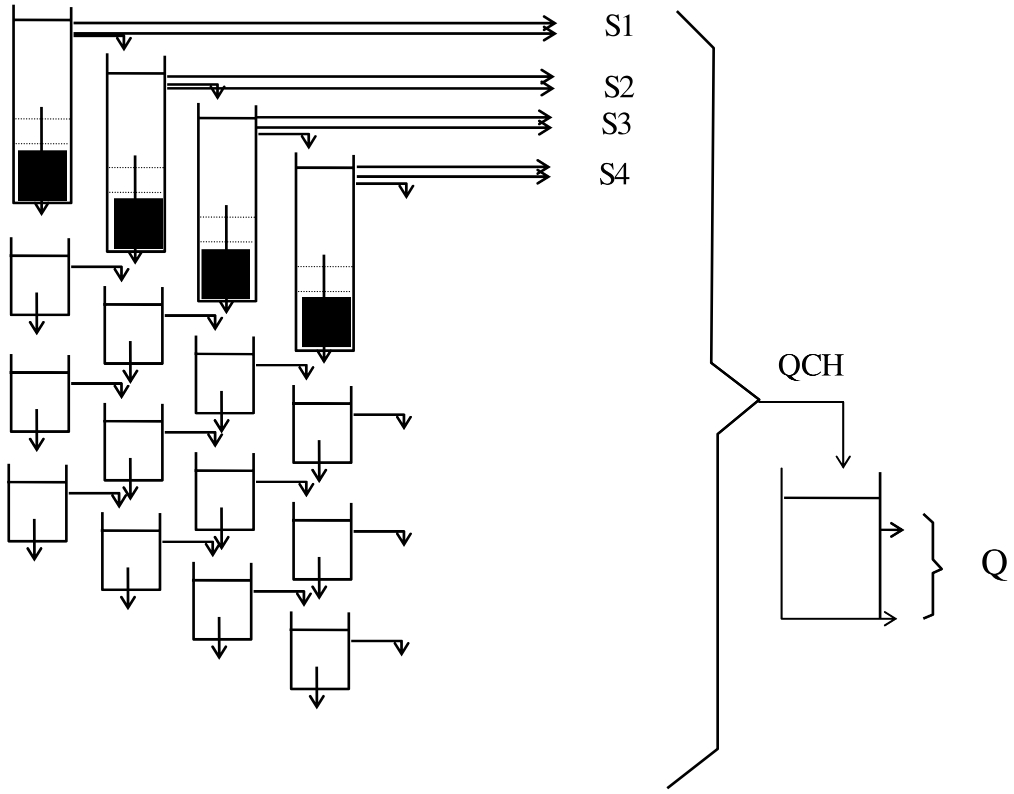
Trực tiếp ngay trên bề mặt sườn dốc tồn tại một lớp mỏng từ đó lượng ẩm thoát đi do bốc hơi và bốc hôi qua lá. Lớp mỏng này được mô phỏng bởi phần trên của bể A và đặc tính của nó được đánh giá bởi thông số PS.

Thông số PSC còn bao hàm cả lượng nước điền trũng trên mặt lưu vực. Nếu không có lớp nước điền trũng, giá trị của PS chỉ xấp xỉ lớp bốc hơi trong thời đoạn tính toán Δt. Bản chất quá trình truyền ẩm từ dưới lên T1 là quá trình bốc thoát hơi nước từ các tầng đất khác nhau thông qua con đường mao dẫn. Đây là điểm tương tự của mô hình TANK với mô hình Stanford.4, khi cho rằng lượng nước trong các tầng đất có sự trao đổi hai chiều.

Quá trình T1 không xảy ra khi và chỉ khi

$XA \geq PS + E$ (2.51)

có nghĩa là khi lượng ẩm làm bão hoà phần trên bể A, điền trũng và bốc hơi. Nguồn ẩm cung cấp cho quá trình T2 là XA, nguồn cung cấp cho quá trình T1 lấy từ các bể B,C,D (XB,XC,XD).



Hình 2.5 Mô hình TANK kép

Như vậy 5 quá trình trao đổi ẩm theo phương thẳng đứng đều có thể xảy ra song song, mối quá trình đều có những điều kiện tồn tại riêng , quy luật diễn biến riêng, chúng bổ sung ẩm cho nhau hoặc tiêu hao ẩm của nhau:

. Mưa

. Bốc hơi

. Thấm qua các cửa đáy

. Truyền ẩm lên T1

. Truyền ẩm xuống T2

Trong các dạng tổn thất còn chưa đề cập đến vai trò của thảm phủ thực vật. Hoàn toàn hợp lý có thể coi rằng thông số HA1 đảm nhận chức năng đó.

Dòng chảy mặt chỉ xuất hiện khi XA >PS + HA1 thông số HA2, HA3, xác định đặc điểm cấu tạo riêng biệt của sườn dốc và không có ý nghĩa vật lý cố định , biểu thức (PS+HA1-XA+SS-XS) xác định lớp tổn thất ban đầu.Giá trị của HA1,xấp xỉ với lớp nước mưa không đủ gây ra lũ và điều này hoàn toàn có thể xác định được khi đối chiếu giữa quá trình mưa và quá trình dòng chảy.

Các thông số HB, HC ,HD đánh giá các tổn thất ban đầu trên các tầng không thấm tương đối. Theo sự nghiên cưu của giáo sư A. N.Bephany cùng các cộng sự của ông , quá trình thấm qua tầng không thấm tương đối triết giảm rất nhanh theo thời gian. Sự thấm ổn định đạt được chỉ sau 15 -30 phút ngay cả trong trường hợp các tầng đất hoàn toàn khô. Trong thực tế thời đoạn tính toán Δt thường lớn hơn nhiều thời gian này và điều đó cho phép coi HB, HD là các hằng số. Giá trị của HB, HC, HD chỉ vào khoảng vài mm.

Trong mô hình, tác dụng điều tiết của sườn dốc đã tự động được xét thông qua các bể chứa xếp theo chiều thẳng đứng. Nhưng hiệu quả của tác động này không đủ mạnh và có thể coi tổng dòng chảy qua các cửa bên của bể YA2+YA1+YB2+YC1+YD1 chỉ là lớp cấp nước tại một điểm. Đây là một yếu điểm của mô hình TANK so với các mô hình khác như SSARR. Bản thân tác giả M.Sugawara nhận thức rõ điều này và khắc phục nó bằng cách cho phép dịch chuyển nhân tạo đỉnh lũ đi 1 thời gian T.

Có thể sử dụng thêm một bể chứa tuyến tính XK để mô phỏng tác động điều tiết sườn dốc. Như vậy, tổng dòng chảy (YA2+YA1+YB2+YC1+YD1) trước khi vào bể điều tiết lòng sông CH phải qua bể điều tiết sườn dốc XK. Cơ chế hoạt động của bể XK như sau :

Tính lớp cấp nước tại một điểm tại thời điểm

$$CK(I) = YA2+YA1+YB2+YC1+YD1 \qquad (2.51)$$

$$QCH = XK1.CK(I-1) + XK2.\ CK(I) +XK3.QCH \qquad (2.52)$$

Trong đó XK1, XK2, XK3 là các thông số và đảm bảo điều kiện XK1+ XK2+ XK3=1 Hiển nhiên, nếu trong (10.49) choXK2=1; XK1=XK3 =0 thì bể XK mất tác dụng và trở lại nguyên bản mô hình TANK ban đầu

4. Mô hình TANK dạng kép

Trong cấu trúc kép có sự biến đổi độ ẩm của đất theo không gian như hình(10.5). Lưu vực được chia thành các vành đai có độ ẩm khác nhau . Một vành đai được điễn tả bằng một mô hình TANK đơn. Về nguyên tắc số lượng vành đai có thể bất kỳ, trong thực tế tính toán thường lấy 4 vành đai mỗi vành đai có 4 bể , tổng cộng toàn mô hình chứa 16 bể .

Với sự mô phỏng này trên toàn lưu vực có những phần ẩm phần khô biến đổi theo quy luật nhất định . Khi mưa bắt đầu, phần lưu vực ẩm ướt sẽ phát triển từ khu hẹp ven sông lan dần đến những vùng cao hơn theo thứ tự S4, S3, S2, S1(Si biểu thị vành đai thứ i so với toàn lưu vực). Ngược lại khi mùa khô bắt đầu do lượng ẩm ướt cung cấp ít dần hoặc không có, lưu vực sẽ khô đần từ những vành đai cao nhất đến vành đai thấp hơn theo thứ tự S1, S2, S3, S4. Trong cấu trúc kép, lớp nước tự do trong mỗi bể được chuyển động theo hai hướng : thẳng đứng và nằm ngang. Mỗi bể chứa nhận được nước từ phía bể trên cùng vành đai và từ phía trái cùng tầng. Trong dạng này, mô hình có thêm các thông sốSi(i=1,2,3,4).

5. Chiến lược dò tìm thông số.

Trong hội nghị quốc tế về lũ và tính toán lũ (15-12 tháng 8 - 1976Leningrat) M. Sugawara nhận định" Do cấu trúc phi tuyến với các bể chưa sắp xếp theo chiều thẳng đứng, chưa có phương pháp toán học hữu hiệu nào để xác định các thông số của mô hình TANK, cách duy nhất là thử sai". Quan điểm này được một số nhà ứng dụng tán đồng.

Phương pháp thử sai không gây khó khăn gì lớn đối với những người đã có kinh nghiệm sử dụng mô hình. Nhưng đối với những ai chưa quen mô hình, khi sử dụng cách thử sai sẽ rất lúng túng và gặp phải điều khó khăn. Giáo sư L. C. Kuchmen và V.I.Koren cũng bày tỏ rằng mô hình TANK hiện nay được coi là một trong những mô hình tốt nhất, nhưng do có quá nhiều thông số, trong đó có những thông số cỡ phần nghìn (0.001)đã gây phần e ngại và khó khăn với người sử dụng chưa quen mô hình. Ngoài cách thử sai, cần thiết phải xây dựng những thuật toán khách quan dò tìm thông số. Năm 1979, M. Sugawara đề xuất phương pháp "lựa chọn tự động thông số mô hình". Sự lựa chọn tự động được thực hiện không phải bằng các phương pháp tối ưu hoá (tìm kiếm cực trị phiếm hàm mục tiêu) mà bằng cách thử sai, nhưng được tự động trên máy tính. Năm 1984 chúng tôi vận dụng phương pháp tối ưu hoá Rosenbroc kết hợp với nguyên lý"phức tạp hoá dần mô hình"do giáo sư L.C.Kuchmen đề xuất.

1.Phương pháp thử sai.

Phương pháp thử sai đòi hỏi người sử dụng phải nắm vững tính năng hoạtđộng của từng thông số. Toàn bộ các thông số của mô hình TANK có thể chia làm 2 loại: thông số có thứ nguyên (HS, PS, SS, HA3, HA2, HA1, HB, HC, HD, H, TB, TB0,TC, TC0) và thông số không thứ nguyên (A1, A2, A3, A0, B1, B0, C1, C0, D1, D0, XK1, XK2, XK3, CH4, CH2). Hiển nhiên là các thông số thứ nguyên sẽ thay đổi theo thời đoạn tính toán Δt. Bản chất của các thông số này là các thông số tổn thất, khi kết hợp với các thông số cửa đáy sẽ gây lên hiệu quả trễ trong quá trình dòng chảy. Các thông số cửa bên (A1, A2, A3, B1, C1, D1) trực tiếp tác động đến độ lớn đỉnh lũ, trong đó A1, A2, A3 tác động đến các đỉnh lũ lớn.

Tính năng hoạt động của các thông số cửa bên và các thông số cửa đáy có thể được mô tả tổng quát như sau:

a-Để làm thay đổi dạng đường quá trình, cần phải điều chỉnh $(\alpha+\beta)$. Thí dụ, muốn đường quá trình nhọn hơn, phải tăng $(\alpha+\beta)$ và ngược lại.

b-Để làm thay đổi tổng lượng dòng chảy trận lũ, cần điều chỉnh $\beta/(\alpha+\beta)$. Thí dụ, muốn làm tăng lượng dòng chảy mà không biến đổi dạng quá trình, cần phải tăng β và giảm α, giữ $(\alpha+\beta)$ không đổi và ngược lại.

Trong quá trình thử sai, phải luôn luôn theo dõi sự cân bằng nước hợp lý trong từng bể. Lượng ẩm trong từng bể (XA, XS, XB, XC, XD,XCH)liên tục biến đổi trong quá trình tính toán, sau một chu kỳ các lượng ẩm này phải đạt được những trị số hợp lý. Thí dụ, chu kỳ hoạt động của bể nước ngầm D là một năm (từ cuối mùa kiệt năm nay đến đầu mùa lũ năm sau), sau 1 năm hoạt động , XD cuối mùa kiệt phải đạt trị số hợp lý phù hợp với phương trình cân bằng nước viết cho một năm $(X=Y+ Z \pm \Delta U)$. Chênh lệch giữa XD đầu và cuối năm phải phù hợp với $\pm\Delta U$. Trong cả một chuỗi năm hoạt động XD không được nhỏ hơn 1 giá trị tương ứng vơi 1 lưu lượng dòng ngầm ổn định. Nếu bể D có xu hướng trữ nhiều hơn tháo, XD sẽ có xu thế lớn dần theo thời gian, dòng chảy kiệt các năm càng về sau càng lớn và ngược lại. Bất kỳ sự phá vỡ cân bằng nước nào trong các bể đều dẫn đến sự không ổn định của bộ thông số và sự bất hợp lý trong thành phần dòng mặt, dòng sát mặt và dòng ngầm. Khi tiến hành thử sai, cần phải nắm được đày đủ các thông tin về các thành phần dòng chảy, về các thành phần trong phương trình cân bằng nước từng bể, động lực các diễn biến cùng nguyên nhân gây ra sự mất cân bằng, từ đó có sách lược hiệu chỉnh thích hợp. Các bể C,B,A sẽ có các chu kỳ hoạt động ngắn hơn. Ngay trong bể A chu kỳ hoạt động của phần trên và phần dưới rất khác nhau. Phần trên của chu kỳ tương đương với thời gian một trận lũ, phần dưới có chu kỳ hoạt động xấp xỉ một năm. Nếu thấy XS sau khi đã đặt đến trạng thái bão hoà SS rồi không thay đổi nữa thì chứng tỏ PS chọn quá lớn, lượng ẩm trong phần trên luôn luôn đủ để bốc hơi.

2. Lựa chọn tự động thông số mô hình theo M. Sugawara

Chế độ này chỉ áp dụng đối với các thông số cửa bên và cửa đáy. Thoạt đầu, các thông số cửa bên và đáy nhận những giá trị sau: A1=A2=A0=0,2; B1=B0=0,05; C1=C0=0,01; D1=0,001 Quy ước ký hiệu dòng chảy qua các cửa bên A2, A1, B1, C1, D1 ần lượt tương ứng là Y1, Y2, Y3, Y4, Y5(H.10.6).

Toàn bộ quá trình dòng chảy được chia làm 5 thời đoạn 1,2,3,4,5 tương ứng với sự hoạt động của 5 cửa bên: A2, A1, B1, C1, D1. Quy tắc chia thời đoạn như sau:

Thời đoạn 1: những ngày mà dòng chảy qua cửa A2 đóng vai trò chính sẽ thuộc thời đoạn 1, nghĩa là khi tỷ số giữa Y1với tổng dòng chảy lớn hơn C (C - một hằng số).

$$Y1 \geq C(Y1 + Y2 + Y3 + Y4 + Y5) = CY$$

Thời đoạn 2: khi

$$Y1 < CY \quad \text{và } (Y1 + Y2) > CY$$

Thời đoạn 3: khi

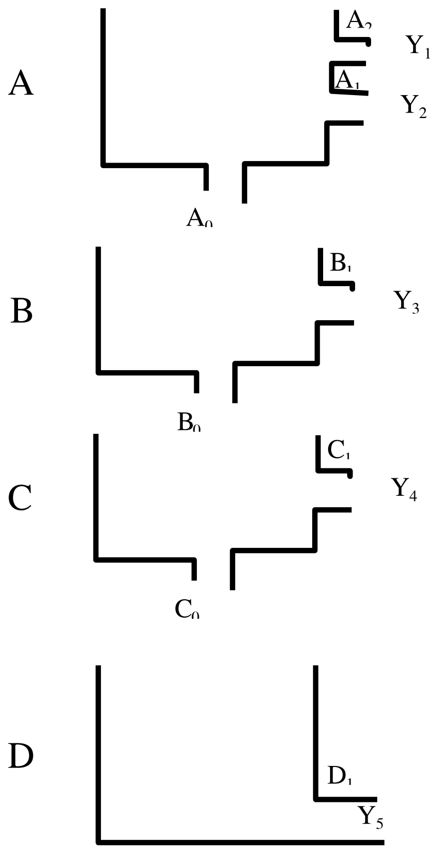
$$(Y1 + Y2) < CY \text{ và } (Y1 + Y2 + Y3) > CY$$

Thời đoạn 4: khi

$$(Y1 + Y2 + Y3) < CY \text{ và } (Y1 + Y2 + Y3 + Y4) > CY$$

Thời đoạn 5: phần còn lại .

C có thể được chọn trong các giá trị sau: 0; 0,5; 0,25; 0,1; 0,05.



Hình 2.6 Dòng chảy từ các bể A,B,C,D

Giá trị C=0,1 tỏ ra tốt đối vối các sông của Nhật. Trong từng thời đoạn 1,2,3,4,5 tổng lượng dòng chảy và hình dạng đường nước rút của quá trình thực đo và tính toán được đánh giá bởi các tiêu chuẩn sau:

$$RQ(I) = \sum_N \widetilde{Q}(N) / \sum_N Q(N) \qquad I = 1,\ldots,5.$$

$$RD_1 = \frac{\sum_N \left[\log \widetilde{Q}(N-1) - \log \widetilde{Q}(N)\right]}{\sum_N \left[\log Q(N-1) - \log Q(N)\right]} \qquad I = 1,\cdots,5$$

Trong đó Q là lưu lượng thực đo , $\tilde{Q}$ là lưu lượng tính toán, I là chỉ số của các thời đoạn, N là số ngày của mỗi thời đoạn I mà có hiệu số [Q(N-1) - Q(N)] dương.

Nguyên lý của việc tự động điều khiển thông số như sau:

a, Khi RQ(I)>1 RQ(I)<1, phải giảm (tăng)thông số cửa bên, và tăng(giảm)thông số cửa đáy. Việc này được thực hiện tự động bằng cách chia thong số cửa bên cho $\sqrt{RQ(I)}$ và nhân thông số của cửa đáy với $\sqrt{RQ(I)}$

b, Khi RD(I)>1 RD(I)<1, phải giảm (tăng) cả hai thông số như nhau. Việc điều khiển này được thực hiện bằng cách chia cả hai thông số cửa bên cho RD(I). Nguyên lý điều khiển nêu trên đưa đến các công thức điều khiển sau:

A0=A0(($\sqrt{RQ(1)}$/RD(1) + ($\sqrt{RQ(2)}$/RD(2)).(1/2)

AM1=A1/ ($\sqrt{RQ(2)}$.RD(2)).

A2=(A1 + A2)/($\sqrt{RQ(1)}$.RD(1) - AM1

A1=AM1

B0=B0 $\sqrt{RQ(3)}$ / RD(3)

B1=B1 /$\sqrt{RQ(3)}$ / RD(3)

C0=C0 .$\sqrt{RQ(4)}$ / RD(4)

C1=C1 /$\sqrt{RQ(4)}$ / RD(4)

D1= D1/RD(5)

Cần kiểm tra lượng nước được cung cấp từ các bể trên. NếuRQ(5) >1, RQ(5) <1, phải giảm (tăng) các thông số cửa đáy của các bể trên. Sự điều khiển lượng nước cung cấp cho bể D được thực hiện bằng điều khiển C0 của bể C, sau đó sự biến đổi trong bể C cho việc điều khiển C0 gây ra sẽ được bù trừ bởi các điều khiển B0 và v.v.. Với cách thức như vậy, sẽ có các công thức điều khiển tiếp sau:

C0= C0/RD(5)

B0=B0/ $\sqrt{RQ(5)}$

A0=A0 $\sqrt[4]{RQ(5)}$

Trong một số trường hợp, giá trị của RQ(I) và RD(I) có thể rất khác 1. Khi xuất hiện những trường hợp đó, chúng ta giới hạn RQ(I) và RD(I) trong phạm vi (1/2, 2)có nghĩa là giá trị RQ(I) và RD(I) lớn hơn 2 sẽ được lấy bằng 2, và những giá trị nhỏ hơn 1/2 sẽ được lấy bằng 1/2 . Trong quá trình điều khiển cần lưu ý hệ điều khiển nêu trên có thể không hội tụ. Có nghĩa là sau một vài lần tính lặp (thường có ít hơn 15 lần). Kết quả thu được khá tốt, nhưng sau đó kết quả lại tồi đi không phục hồi lại được. Một trong những nguyên nhân là RD(I) chịu tác động của nhiều yếu tố ngẫu nhiên kém tin cậy. Để giảm tác động của RD(I) có thể thay RD(I) = $\sqrt{RD(I)}$ hoặc RD(I) = $\sqrt[4]{RD(I)}.RD(5)$ là kém tin cậy nhất, do đó việc điều khiển thông số bể D phải rất thận trọng . Rất nhiều trường hợp RD(5) đã phá hỏng toàn bộ hệ điều khiển thông số nêu trên.

3. Tối ưu hoá thông số mô hình.

Bộ thông số mô hình được thiết lập theo phương pháp Rosenbroc vối hàm mục tiêu của quá trình điều khiển thông số nêu trên.

$$K = \sum_{i=1}^{n} \int_{0}^{T} \left[Q(t) - Q(t, A) \right]^2 dt \rightarrow \min$$

Trong đó: n -số quá trình đưa vào tốt ưu ; T -thời gian 1 quá trình, A - véc tơ thông số được mã số theo bảng sau:

A	1	2	3	4	5	6	7	8	9	10	11	12
T.S	A_1	A_2	A_3	HA_1	HA_2	HA_3	A_0	B_1	C_1	D_1	HB	HC

A	13	14	15	16	17	18	19	20	21	22	23	24
T.S	HD	B_0	C_0	D_0	XK_1	XK_2	XK_3	H	CH_1	CH_1	α	TB

A	25	26	27	28	29	30	31	32	33	34	35	36
T.S	TB_0	T	C	TC_0	PS	SS	KZ	XA	XS	XC	XD	XCH

Phương pháp tối ưu hoá không thể thành công nếu đưa tất cả những thông vào tốt ưu đồng thời. ở đây, tốt ưu hoá được coi là thử sai tự động theo hàm mục tiêu K với thuật toán Rossenbroc. Điều đó có nghĩa là thuật toán tối ưu phải đủ mềm dẻo cho phép lựa chọn được các thông số mong muốn đưa vào tối ưu, do vậy các thông số đều được gắn nhãn như bảng trên. Quá trình tối ưu thông số mô hình phải tuân theo những nguyên tắc đã được trình bày ở trên.

4. Một số nhận xét.

Mô hình TANK được nhiều cơ quan nghiên cứu ứng dụng. Trường Đại học thuỷ lợi, Viện khí tượng - Thuỷ văn, Viện thiết kế thuỷ lợi quốc gia, Công ty khảo sát thiết kế điện 1, Cục dự báo thuỷ văn v.v.. Trong quá trình ứng dụng nổi bật một số vấn đề:

1. Mô hình khó thể hiện sự "trễ" của dòng chảy so với mưa. Với đặc điểmnày, mô hình thích ứng với các lưu vực nhỏ. Điều này có thể khắc phục được bằng cách nối tiếp thêm một số bể tuyến tính và kênh tuyến tính biểu diễntác dụng điều tiết của lưu vực và lòng sông. Hoàn toàn có thể sử dụng lớp mô hình" hộp đen" nêu trên trong công việc này.

2. Do mô hình được cấu tạo từ các bể tuyến tính, các thông số cửa ra trong một số trường hợp tỏ ra kém nhậy. Trên một số lưu vực, dòng chảy mặt đóng vai trò đáng kể (lũ lên nhanh, rút nhanh), có sự phân hoá rõ rệt trong sự hình thành các cấp lưu lượng, quá trình dòng chảy đối nhạy cảm với quá trình mưa, nên sử dụng bể nước mặt (bể A) dưới dạng phi tuyến. Thí dụ. cửa A1 có thể là bậc 2, cửa a2, a3, có thể là số bậc cao hơn.

3. Xét điều kiện ban đầu.Trong mô hình, tất cả các quá trình thành phần như bốc thoát nước, tổn thất trên thảm thực vất tao đổi ẩm giữa các vùng và các bể, thấm, điều trũng, thảm thành dòng mặt, dòng sát mặt, dòng ngầm, diễn toán lủ tên sườn dốc và trên sườn dốc và trong sông được liên kết vơi nhau thông qua việc biến đổi các độ ẩm XA, XS, XB, XC, XD, XCH trong từng bể. Rất quan trọng việc xét các độ ẩm này tại đầu thời kỳ tính toán. Việc xét điều kiện ban đầu có thể tiến hành theo thủ pháp sau:

a) Để xét các độ ẩm ban đầu trong phần trên, phần dưới bể A(XA0, XS0) nên chọn thời điểm ban đầu tính toán là lúc đất đã được bão hoà, Độ thiếu hút ẩm trong đất coi như bằng 0 (thí dụ sau một trận mưa lớn gây lũ rõ rệt). Trong những trường hợp này có thể coi

XA0 = PS + HA1

XS0 = SS

b) Có đủ cơ sở để cho rằng XA, XS có quan hệ với độ ẩm lưu vực, do vậy, trước thời điểm tính toán, XA0, XS0 có thể được xác định qua mối ràng buộc của chúng đối với độ ẩm đất theo giáo sư N.Ph. Befanhi.

Jw= x1 + 0,7x2-4 + 0,5x5-9 +0,3x10-14 +0,2x15-30 + 0,1x31-60

Ở đây, x1 - lượng mưa 1 ngày trước thời điểm; x2-4 - lương mưa trong2, 3, và 4 ngày trước thời điểm tính toán v. v...

c) Để đánh giá độ ẩm ban đầu trong các bể khác (XB0, XC0, XCH0) hoàn toàn có thể giả định tồn tại các mối quan hệ bền vững giữa chúng với lưu lượng trước lũ Q0.

d) Độ ẩm XD0 ban đầu thiết lập theo vị trí số Q0 bằng cách tính ngược sau khi đã biết XA0, XS0, XB0, XC0, XCH0.

2.5.3 Mô hình SSARR

Mô hình SSARR do Roc-cờ-ut (Rockwood) đề xuất từ năm 1956. Khi xây dựng mô hình này người ta quan niệm rằng hệ thống sông ngòi dù phức tạp cũng chỉ gồm các thành phần cơ bản sau :

- Các lưu vực sông nhỏ

- Các hồ chứa tự nhiên và nhân tạo

- Các đoạn sông

Do đó người ta xây dựng mô hình toán học cho từng loại, sau cùng tập hợp lại ta sẽ có mô hình toán học của cả hệ thống sông. Các mô hình toán học thành phần đều sử dụng hai phương trình cơ bản là phương trình liên tục và phương trình trữ lượng.

Phương trình liên tục là:

1/2(I1 + I2)Δt - 1/2(Q1 + Q2)Δt = S2 - S1 (10.51)

trong đó I, I2 lưu lượng chảy vào ở đầu và cuối thời đoạn tính toán Δt ; Q1,Q2 - lưu lượng chảy ra ở đầu và cuối thời đoạn Δt; S1,S2 là dung tích hồ chứa ở đầu và cuối thời đoạn Δt.

Phương trình lượng trữ của hồ chứa là :

$$\frac{dS}{dt} = T_s \frac{dQ}{dt} \qquad (2.53)$$

hãy viết dưới dạng sai phân :

ΔS = TsΔQ (2.54)

Thay (10.53) vào (10.51) ta có :

$$\frac{I_1 + I_2}{2}\Delta t - \frac{Q_1 + Q_2}{2}\Delta t = T_s(Q_2 - Q_1) \qquad (2.55)$$

Đặt $I_m = \frac{I_1 + I_2}{2}$ và qua biến đổi ta có:

$$Q_2\left(T_s+\frac{\Delta t}{2}\right)=Q_1(T_s-\frac{\Delta t}{2})+I_m\Delta t$$

$$Q_2(T_s+\frac{\Delta t}{2})=Q_1(T_s+\frac{\Delta t}{2})-Q_1\Delta t+I_m\Delta t$$

$$Q_2=\frac{(I_m-Q_t)\Delta t}{T_s+\frac{\Delta t}{2}}+Q_1 \qquad (2.56)$$

Như vậy nếu biết được lưu lượng chảy vào trung bình Im lưu lượng chảy ra ở đầu thời khoảng tính toán Q1 và thời gian trữ nước của hồ Ts thì có thể tính được lưu lượng chảy ra ở cuối thời khoảng tính toán Q2 theo phương trình (2.56)

2.5.3.1. Mô hình lưu vực

- Lượng nước đến của một lưu vực kín gồm có lượng mưa và tuyết rơi (Hình 10.7). Một phần của lượng nước đến này được giữ lại trên bề mặt lưu vực làm ẩm đất, một phần bay hơi vào khí quyển, phần còn lại sẽ tạo thành 3 kiểu như sau :

- Chảy tràn trên mặt đất,

- Chảy ngầm trong đất và lớp đất ở phía trên

- Chảy ngầm trong lớp đất ở tầng sau (xem hình 2.7)

Người ta hình dung mỗi quá trình chảy kể trên như chảy qua một chuỗi các hồ kế tiếp nhau. Lượng nước chảy vào hồ chứa đầu tiên của chuỗi hồ chứa này chính là lượng chảy vào của hồ chứa tiếp theo. Tập hợp lượng nước chảy ra từ hồ chứa cuối cùng chính là lượng nước chảy ra của cả lưu vực.

Để tính được lượng nước chảy vào của các hồ chứa đầu tiên ta phải tính được toàn bộ lượng nước đến của lưu vực, sau đó tách riêng phần tham gia dòng chảy sát mặt và dòng chảy ngầm.

a, Tính lượng nước mưa trung bình trên lưu vực

Người ta thường tính lượng mưa trung bình ngày theo công thức :

$$X_N = \frac{1}{n}\sum_{i=1}^{n} a_i . x_i \qquad (2.57)$$

trong đó Xi - lượng mưa đo được ở trạm thứ i trong một ngày;

n - số trạm đo mưa trên toàn lưu vực;

ai - hệ số trung bình tính theo phương pháp hình nhiều cạnh hoặc lấy bằng tỉ số giữa lượng mưa trung bình hàng năm trên phần lưu vực tương ứng và lượng mưa trung bình hàng năm tại trạm đo mưa thứ i.

XN: lượng mưa trung bình ngày tính toán

Khi khoảng thời tính toán Δt ngắn hơn một ngày thì lượng mưa trung bình trong khoảng thời gian Δt là :

XΔt = b.XN (2.58)

Với b là hệ số chuyển đổi.

b, Tính độ ẩm của đất .

Hệ số dòng chảy phụ thuộc chủ yếu vào độ ẩm của đất trên lưu vực. Người ta chỉ dùng chỉ số độ ẩm A để biểu thị độ ẩm của đất.

A2 = A1 + (X-Y) - K1E (2.59)

Với A1, A2- chỉ số độ ẩm ở đầu và cuối khoảng Δt

X, Y - lượng mưa và lượng dòng chảy trong thời khoảng Δt

E - lượng bốc hơi ngày, tính trung bình trên toàn lưu vực. Nếu trên lưu vực có n trạm bốc hơi thì :

$$E = \frac{1}{n}\sum_{i=1}^{n} \gamma_i E_i \qquad (2.60)$$

γi - hệ số trung bình còn Ei lượng bốc hơi ngày do được trạm thứ i;

K1 - hệ số chuyển đổi, nó thay đổi theo độ ẩm của đất.

K1 = f1(A)

Trương hợp thiếu tài liệu bốc hơi hàng ngày thì dùng trị số bốc hơi trung bình tháng ET nhân với hệ số chuyển đổi K2. Lúc đó độ ẩm của đất tính theo công thức :

$$A_2 = A_1 + (X - Y) - \frac{\Delta t}{24} K_2 E_T \qquad (2.61)$$

c, Tính lớp dòng chảy

Lớp dòng chảy, tổng cộng là Y = αX

Với α là hệ số dòng chảy phụ thuộc vào độ ẩm của đất :

Lớp dòng chảy tổng cộng này được phân chia thành 3 thành phần ứng với dòng chảy mặt, dòng chảy sát mặt và dòng chảy ngầm.

Lớp dòng chảy ngầm trong 1 giờ là :

$$Y_{ng} = K_3 \frac{Y}{\Delta t} \qquad (2.62)$$

K3 - là hệ số chảy ngầm, nó phụ thuộc vào chỉ số thấm P :

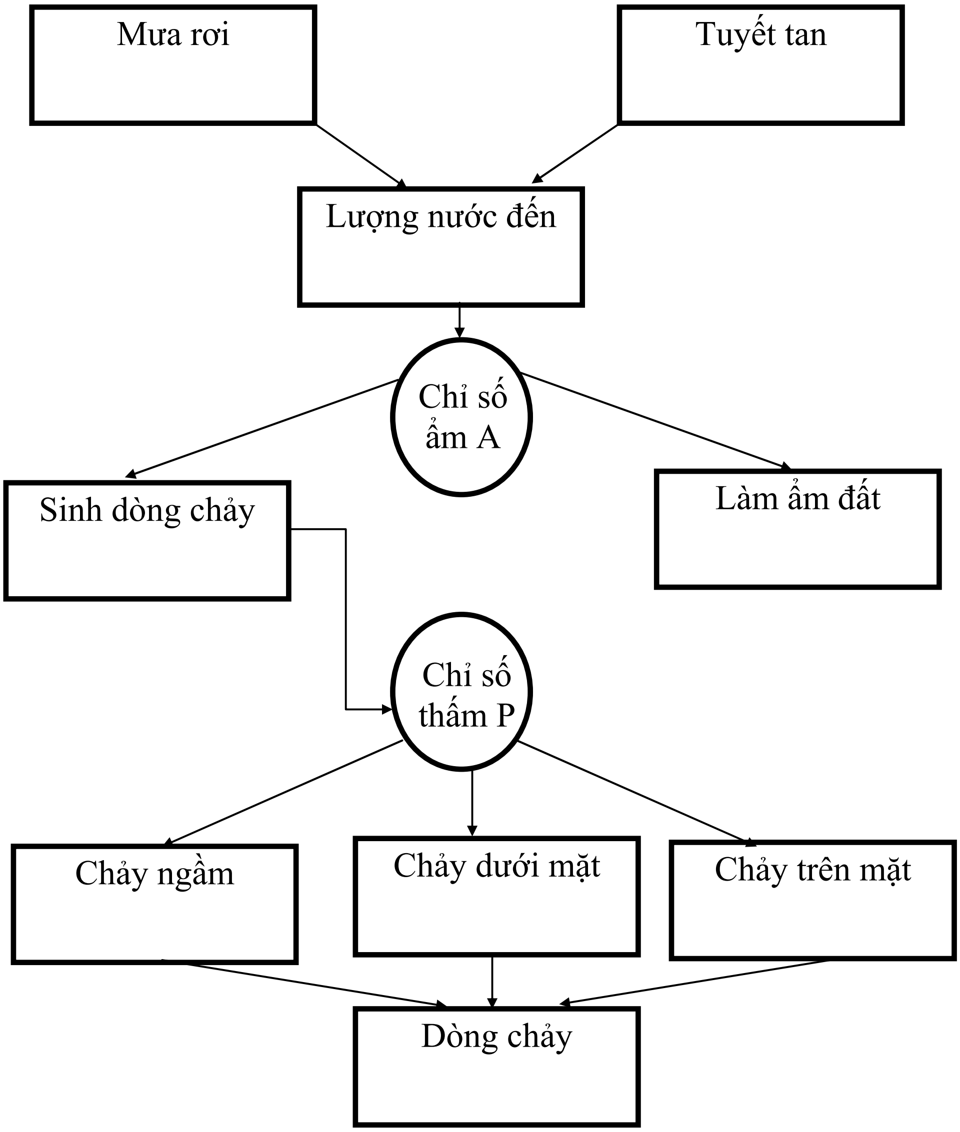
K3 = f3(P)

Chỉ số thấm P tính như sau :

$$P_2 = P_1 + \left(24\frac{Y}{\Delta t} - P_1\right)\frac{\Delta t}{T + \frac{\Delta t}{2}}$$

P1, P2 - chỉ số thấm ở đầu và cuối thời khoảng Δt

T - thời gian trữ nước biến đổi từ 30 đến 60 giờ



Hình 2.7 Sơ đồ mô hình lưu vực của mô hình SSARR $\alpha = f_2(A)$

Việc phân chia thành dòng chảy mặt Ym và dòng chảy sát mặt Ysm dựa vào các giả thiết sau :

- Dòng chảy mặt đạt trị số lớn nhất Ymmax và giữ nguyên vị trí số đó khi G lớn hơn 200% của Ymmax

- Dòng chảy mặt nhỏ nhất Y mmin bằng 10% của G

Với G = Ym + YSm = Y - Yng

Khi đó lớp dòng chảy mặt trong một giờ là :

Ym = f4(G)

Khi Ym < Ymmax thì:

$$Y_m = \left(0{,}1 + 0{,}2.\frac{G}{Y_{m\max}}\right)G$$

Nếu Ym ≥ Ymmax thì lấy Ym = Ymmax

YSm = G - Ym

d, Tính lưu lượng chảy ra của lưu vực :

Sau khi thực hiện phân chia lượng mưa hiệu quả thành 3 phần : lượng nước tham gia dòng chảy mặt, sát mặt và dòng chảy ngầm, ta coi đó là lượng nước chảy vào của 3 hồ chứa đầu tiên trong 3 hồ chứa tưởng tượng với 3 cách tạo thành dòng chảy. Nếu biết số hồ chứa của từng chuỗi n1, n2, n3và thời gian trữ nước TS1, TS2, Ts3 ta có thể tính được lưu lượng chảy ra từ hồ cuối cùng bằng cách sử dụng liên tiếp công thức (10.58). Lưu lượng chảy ra của lưu vực là tổng của các lưu lượng chảy ra từ 3 hồ chứa sau cùng.

e, Điều chỉnh thông số.

Các thông số có mô hình lưu vực là

- Các thông số để tính mưa bình quân trên lưu vực ai, b
- Các thông số để tính bốc hơi K1, K2, γi
- Các thông số n1,n2, n3, TS1, TS2, Ts3 và T.
- Quan hệ giữa hệ số dòng chảy và độ ẩm α=f2(A)
- Quan hệ để tính lớp dòng chảy ngầm K3 = f3(P)
- Quan hệ để phân chia dòng chảy mặt và dòng chảy ngầm Ym = f4(G).

Các thông số và quan hệ kể trên được lựa chon giá trị tối ưu thông qua việc tính thử dần sao cho sự sai khác giữa lưu lượng thực đo và lưu lượng tính toán là nhỏ nhất. Cho tới nay, việc điều chỉnh các thông số của mô hình SSARR còn chưa được tự động hoá, vì thế nó còn là một công việc rất phức tạp và phụ thuộc nhiều vào kinh nghiệm của người điều chỉnh mô hình. Ở trên đã kể ra nhiều thông số và quan hệ, nhưng chỉ có 4 loại sau ảnh hưởng nhiều nhất tới kết quả tính toán.

- Các hệ số tính mưa trung bình lưu vực ai, b

- Hệ số Ts1 của dòng chảy mặt

- Quan hệ hệ số dòng chảy và độ ẩm α =f2(A)

- Quan hệ của hệ số chảy ngầm với chỉ số thấm K3=f3(P)

Người ta chọn các thời kỳ có đường quá trình biến đổi nhiều (mùa lũ năm nước lớn) để điều chỉnh thông số, sau đó thử lại cho các năm khác.

2.5.3.2. Mô hình dòng chảy trong sông

Dòng sông được coi như bao gồm một chuỗi hồ chứa kế tiếp nhau, mỗi hồ chứa ứng với một đoạn sông dài từ 6 đến 10 km. Thời gian trữ nước Ts của đoạn sông tính theo quan hệ.

$$T_s = \frac{K4}{Q^n}$$

Với K4, n là các hằng số thực nghiệm.

Cũng có thể tính Ts theo quan hệ Ts = f(Q) lấy ra từ tài liệu thực đo. Lưu lượng chảy ra từ đoạn này được dùng làm lưu lượng chảy vào ở đoạn tiếp theo. Việc lựa chọn các giá trị của K4, n và chiều dài tính toán của các đoạn sông được làm theo cách thử dần.

2.5.3.3. Mô hình hồ chứa

Đối với hồ chứa tự nhiên, lưu lượng chảy vào hồ coi như đã biết, nếu tính được thời gian trữ nước Ts thì tính được lưu lượng chảy ra theo phương trình (10.58). Ts biến thiên theo mực nước hồ : Ts = f(H). Với mọi hồ chứa quan hệ Ts =f(H) đã được xác định sẵn từ trước, do đó biết lưu lượng chảy vào thì tính được ngay lưu lượng chảy ra.

Ở các hồ chứa nhân tạo, ngoài đường cong Ts = f(H) còn cần phải biết thêm Hmax, Hmin, đường cong H~Q khi H>Hmax và khả năng tháo qua hồ ứng với các cấp mực nước, nếu là hồ chảy theo chế độ có điều tiết thì phải tính đến sự điều tiết này. Lưu lượng chảy ra tính toán phải nhỏ hơn khả năng tháo qua của hồ và mực nước tính toán phải lớn hơn Hmin.

2.5.3.4. Mô hình hệ thống sông

Hệ thống sông bao gồm các lưu vực nhỏ, các hồ chứa và các đoạn sông. Những mô hình thành phần này đã biết, khi ghép lại trong mô hình hệ thống sông còn phải chú ý đến ảnh hưởng của nước vật, hoặc lượng nước lấy ra để tưới ruộng và lượng nước chảy thêm vào đoạn sông do mưa trên đồng ruộng, hoặc do nước sau khi đã tưới ruộng xong được tháo ra sông. Tất cả quá trình tính toán đã được thực hiện trên máy tính theo các chương trình mẫu.

2.5.4. Mô hình diễn toán châu thổ

Ở những dòng chảy qua đồng bằng, nếu không có đê bao bọc thì lúc mùa lũ đến, nước sông sẽ dâng lên và chảy tràn ra đồng ruộng hai bên bờ, khi đó dòng chảy không chỉ theo chiều dòng sông mà còn theo chiều vuông góc với dòng sông nữa. Để mô tả quá trình này, rõ ràng không thể dùng hệ phương trình Saint -Venant, vì hệ phương trình này chỉ mô tả quá trình chuyển động không ổn định, biến đổi chậm của nước chảy 1 chiều trong sông. Người ta có ý định mô tả sự chuyển đi của dòng lũ qua vùng đồng bằng ngập lụt giống như sự truyền đi của thuỷ triều ngoài biển vì chúng đều truyền đi theo cả hai chiều vuông góc với nhau. Phương trình biểu diễn sự truyền đi của thuỷ triều ở biển đã được lập ra từ lâu với giả thiết là một vùng biển thì độ sâu của nước không chênh lệch nhau quá nhiều. Ở vùng đồng bằng ngập lụt giả thiết này không còn đúng nữa, cho nên không thể có được kết quả nếu như áp dụng phương trình truyền thuỷ triều để tính truyền lũ qua vùng đồng bằng. Trong các năm 1962 - 1966, khi khảo sát vùng đồng bằng hạ lưu sông Mê- Công, người ta đã nhận thấy rằng chiều sâu của vùng "đồng bằng bị ngập hoàn toàn" là rất khác nhau, đến mức không thể coi chiều sâu nước ở mọi nơi là những đại lượng cùng cấp. Ngay cả khi lũ lớn nhất, đồng bằng vẫn còn nhiều nơi không bị ngập và hình thành nhiều ổ chứa nước ranh giới giữa các ô ngưỡng tràn. Căn cứ vào thực tế địa hình người ta chia bề mặt lưu vực thành nhiều ô, các ô này lại được xếp thành các tầng liên tiếp nhau sao cho một ô chỉ trao đổi nước với các ô khác ở cùng tầng và những ô ở tầng trên kề trước và sau nó. Đây là một giải pháp sáng tạo cho phép mô tả gần đúng dòng chảy hai chiều ở đồng bằng mà khối lượng tính toán lại giảm đi rất nhiều so với việc dùng phương trình truyền thuỷ triều. Cách chia lưu vực thành nhiều ô và tính toán trao đổi nước giã các ô như đã nái trên chính là nội dung mô hình Đen-ta (delta)do Prâysman (Preissman) và Cunge đưa ra.

Sau khi chia bề mặt lưu vực thành nhiều ô, người ta thừa nhận hai giảthiết là :

- Thể tích nước trong mỗi ô là hàm bậc nhất của mực nước trong ô.

- Lưu lượng chảy giữa hai ô là hàm bậc nhất của mực nước ở hai ô ấy ở cùng một thời điểm. Nghĩa là bỏ qua lực quán tính tác động tới lưu lượng chảy giữa hai ô. Người ta đã chứng minh ở vùng đồng bằng, sai số do bỏ qua lực quán tính là rất nhỏ.

Phương trình cân bằng nước viết cho ô thứ i là:

$$S_i \frac{dz}{dt} = P_i + \sum_{k=1}^{n} Qi,k \qquad (2.63)$$

Trong đó Pi là lượng mưa hiệu quả trên mặt ô thứ i, nó thay đổi theo thời gian t; Pi=f1(t). Giá trị Pi biết được từ tài liệu đo đạc mưa và thấm.

Si - diện tích mặt nước ô thứ i ứng với độ sâu thay đổi thì cũng biến đổi theo Si = f2(zi)

Qi,k là lưu lượng nước chảy từ ô thứ i vào ô thứ k, theo giả thiết Qi,k là hàm bậc nhất của Zi và Zk.

$$Qi,k = f_3(Z_i, Z_k)$$

Lượng nước chảy giữa hai ô liền nhau có thể tuân theo các qui luật chảy loại sông và loại bờ tràn. Khi chảy loại sông, dòng chảy không chảy tổn thất cục bộ và lưu lượng tính theo công thức Stric-lơ.

Qi,k =αAR2/3J1/2 (2.64)

Với A: là diện tích mặt cắt ướt giữa hai ô trứ i và thứ k.

R: bán kính thuỷ lực của A và J là tốc độ mặt nước, là hằng số. Vì A, R,J, đều là hàm số của mực nước giữa hai ô i và k cho nên.

$$Q_{i,k} = f(\overline{Z}_{i,k}) \qquad (2.65)$$

Với $\overline{Z}_{i,k} = \beta Z_i + (1-\beta) Z_k$. Ở đây hằng số β≤1.

Khi chảy loại bờ tràn thì thường gặp hơn cả loại chảy qua đập tràn đỉnh rộng. Lưu lượng qua ngưỡng cửa tràn phụ thuộc vào kích thước cửa tràn, mực nước thượng lưu và mực nước hạ lưu. Các công thức tính toán đã trình bày trong các giáo trình thuỷ lực.

Giả sử lựa chọn mức thời gian tính toán là Δt, ở thời điểm đầu t=n.Δ đã biết điều kiện đầu là giá trị độ sâu mực nước ở tất cả các ô, vậy là đã biết ta tính được Q_i^n, k bằng cách lấy tổng cộng lưu lượng chảy qua các mặt xung quanh của ô thứ i. Chỉ số n ở các kí hiệu Q_i^n, k, Z_i^n biểu thị các đại lượng Q, Z ở thời điểm t=nΔt

Lấy tích phân phương trình (2.63) trong khoảng thời gian Δt ta có :

$$S_i \Delta Z_i = P_i(\tau)\Delta t + \Delta t \sum_{k=1}^{n} Q_{i,k}(\tau) \qquad (2.66)$$

Với τ là thời điểm nằm giữa n.Δt và (n+1)Δt

n.Δt < τ < (n+1)Δt

Còn lưu lượng chảy từ ô thứ i sang ô thứ k là :

$$Q_{i,k}(\tau) = \beta Q_{i,k}^{n+1} + (1-\beta) Q_{i,k}^{n}$$

hằng số tự chọn trong khoảng 0 ≤ β ≤ 1

Nếu chọn β=0 thì $Q_{i,k}(\tau) = Q_{i,k}^n$, do đó tất cả các số hạng ở vế phải của phương trình(11.65) là đã biết, ta tính ngay được giá trị ΔZi ở vế trái, từ đó tính ra Zi ở thời điểm (n + 1)Δt theo công thức:

$$Z_i^{n+1} = Z_i^n + \Delta Z_i$$

Về mặt cơ cấu tính toán, chọn β = 0, sơ đồ rất đơn giản. Nhưng để β = 0 ta phải chọn Δt đủ nhỏ, sao cho có thể coi lưu lượng Qi,k không thay đổi nhiều trong khoảng Δt, bảo đảm điều kiện

$$Q_{i,k}(\tau) = Q_{i,k}^n$$

Thường ta phải chọn Δt< 30 phút. Việc chọn Δt nhỏ, dẫn tới thời gian tính trên máy tính tăng lên nhiều. Người ta thường chọn β ≠ 0 để có thể lựa chọn Δt dài hơn (từ 6 giờ đến 72 giờ). Khi chọnβ ≠ 0 phương trình (10.65) được giải theo phương pháp sơ đồ ẩm.

Nếu chọn β =1 ta có :

$$Q_{i,k}{}^{(\tau)} = Q_{i,k}^{n+1} \qquad (2.67)$$

$Q_{i,k}^{n+1}$ -là lưu lượng chảy từ ô thứ i sang ô thứ k ở thời điểm t=(n+1)Δt ta chưa biết được cho nên dùng phép khai triển Taylo để chuyển $Q_{i,k}^{n+1}$ thành một chuỗi các giá trị ở thời điểm t=nΔt đã biết. Khi bỏ qua các vô cùng bé bậc cao,khai triển Taylo của $Q_{i,k}^{n+1}$ là:

$$Q_{i,k}^{n+1} = Q_{i,k}^n + \frac{\partial Q_{i,k}^n}{\partial Z_i}\Delta Z_i + \frac{\partial Q_{i,k}^n}{\partial Z_k}\Delta Z_k \qquad (2.68)$$

Thay (2.67) và (2.68) vào (2.66) và sắp xếp lại các ẩn số ΔZi , ΔZk ta có:

$$\left(-\frac{S_i}{\Delta t}+\sum_K \frac{\partial Q_{i,k}^n}{\partial Z_i}\right)\Delta Z_i+\left(\sum_K \frac{\partial Q_{i,k}^n}{\partial Z_k}\right)\Delta Z_k+L_i \tag{2.69}$$

Ở phương trình (11.68), ΔZi, ΔZk là sự thay đổi mực nước ở các ô thứ i và ô thứ k chính là các ẩn số phải tìm, còn lại tất cả các thành phần khác đã biết ở thời điểm nΔt.

Ứng với mỗi ô ta viết được một phương trình tuyến tính dạng (2.69). Nếu lưu vực gồm m ô thì ta viết một hệ m phương trình tuyến tính bậc nhất với m ẩn số. Hệ phương trình này lúc nào cũng giải được bằng các phương pháp quen biết.

2.5.5 Một số kết quả ứng dụng mô hình tất định ở Việt Nam

Những thành tựu cơ bản trong lĩnh vực ứng dụng, nghiên cứu mô hình toán thuỷ văn ở Việt Nam được phản ánh khá đầy đủ trong Hội thảo Quốc gia về ứng dụng mô hình toán thuỷ văn và thuỷ lực trong phát triển và quản lý tài nguyên nước tại Hà Nội năm 1988. Mô hình được hoàn chỉnh và sớm có ứng dụng tại Việt Nam là mô hình SSARR đầu tiên ở lĩnh vực thuỷ văn công trình và sau đó được nghiên cứu ứng dụng cho dự báo lũ ở khu vực đồng bằng sông Cửu Long có tính đến ảnh hưởng triều và các pha lũ tràn bờ. Mô hình SSARR cũng được cải tiến và ứng dụng để dự báo lũ cho sông Hồng - một hệ thống sông phức tạp của đồng bằng Bắc Bộ, bước đầu cho những kết quả đáng khích lệ.

Mô hình TANK được ứng dụng ở Việt Nam vào cuối những năm 1980. Mô hình tương đối đơn giản, có ý nghĩa vật lý trực quan, thích hợp với các lưu vực sông suối vừa và nhỏ. Một số mô hình truyền thống đã được áp dụng từ trước như mô hình Kalinhin - Miuliacốp, phương pháp diễn toán lượng gianhập khu giữa được vận dụng khá linh hoạt trong các lĩnh vực tính toán và dự báo thuỷ văn.

Việc kết hợp các phương pháp truyền thống và các mô hình SSARR, TANK, NAM ... đang được triển khai trong nhiều dự án nghiên cứu ứng dụng.Kết quả sử dụng các mô hình SSARR, TANK, NAM cho các lưu vực sông suối nhỏ cho thấy các đặc trưng trung bình của dòng chảy năm, dòng chảy mùa và các tháng cũng như phân phối dòng chảy tính từ các mô hình trên đạt yêu cầu về độ chính xác cho giai đoạn qui hoạch.

Chương 3
MÔ HÌNH NGẪU NHIÊN

3.1 CẤU TRÚC NGUYÊN TẮC CỦA MÔ HÌNH NGẪU NHIÊN

3.1.1 Nguyên tắc mô phỏng

3.1.1.1. Quá trình ngẫu nhiên

Như đã phân tích trong chương 1, mô hình ngẫu nhiên mô phỏng quá trình dao động của bản thân quá trình thủy văn mà không chú ý đến các nhân tố đầu vào tác động của hệ thống. Các quá trình thủy văn tiến triển trong không gian và thời gian theo một cách thức mà trong đó có một phần mang tính tất định và một phần mang tính ngẫu nhiên. Trong một số trường hợp tính biến đổi ngẫu nhiên nổi trội hơn hẳn tính biến đổi tất định và khi đó nó được coi là một quá trình ngẫu nhiên thuần tuý. Trong những quá trình như vậy giá trị quan trắc của quá trình không có tương quan gì với các giá trị quan trắc trước đó, và các đặc trưng thống kê của tất cả các quan trắc là như nhau.

Khi các giá trị quan trắc không có tương quan với nhau, sản phẩm đầu ra của hệ thống thủy văn sẽ được xử lý như một mô hình ngẫu nhiên không gian độc lập và thời gian độc lập như đã chỉ ra trong 'cây phân loại' 1.4. Cách xử lý này rất thích hợp với các quan trắc của những sự kiện thủy văn cực đoan như dòng chảy lớn nhất hay các số liệu trung bình trong một khoảng thời gian dài của một số quá trình như lượng mưa trung bình năm.

Tuy nhiên với các đặc trưng thủy văn, mà rõ nét là các dòng chảy thời đoạn (năm, tháng, tuần) tình hình không hoàn toàn như vậy. Các kết quả nghiên cứu cho thấy rằng trên hầu hết các sông ngòi thế giới, tồn tại mối tương quan của dòng chảy trung bình các năm kề nhau với hệ số tương quan vào khoảng 0,2- 0,3. Với các thời đoạn ngắn hơn như dòng chảy tháng, dòng chảy tuần, mối liên hệ này càng rõ nét với hệ số tương quan thực sự lớn, có thể đạt đến 0,8-0,9. Mối liên hệ tương quan này có ảnh hưởng tới các kết quả tính toán thủy văn, thủy lợi mà không thể bỏ qua. Ratcovich Đ.IA. đã chỉ ra rằng khi tính đến tương quan giữa các năm kề nhau (R(1)= 0,3) thì dung tích hồ chứa tính được sẽ tăng lên 1,5 lần so với khi coi chúng là những đại lượng ngẫu nhiên độc lập. Còn khi R(1) = 0,5 thì dung tích tăng gấp 2 lần.

Sự tồn tại mối liên hệ tương quan này có nhiều nguyên nhân liên quan đến sự chuyển đổi lượng trữ ẩm trên lưu vực. Bản chất của nó liên quan đến chu kỳ hoạt động của mặt trời. Nghiên cứu chi tiết hơn trên cơ sở số liệu rộng rãi của các sông ngòi trên thế giới, Ratkovich cho rằng nguyên nhân cơ bản của mối liên hệ này là dao động bốc hơi trên bề mặt lưu vực. Ông cho thấy hệ số tương quan giữa các năm kề nhau R(1) có liên hệ rõ nét với môđun dòng chảy năm M_o và ở mức độ nhỏ hơn là hệ số biến đổi của dòng chảy năm C_v.

Mối liên hệ dòng chảy của các thời đoạn ngắn hơn như tháng, tuần càng rõ nét hơn, và cũng liên quan chặt chẽ với sự thay đổi lượng trữ ẩm trên lưu vực. Trong mùa kiệt mối liên hệ này liên quan chặt chẽ với quá trình rút nước lưu vực, có thể biểu thị bằng phương trình:

$$Q_t = Q_o e^{-(t-t_0)/k} \qquad (3.1)$$

Trong đó Q_0 là lưu lượng tại thời điểm t_0 và k là hệ số triết giảm.

Lượng dòng chảy các tháng mùa lũ có mối tương quan kém chặt chẽ hơn, tuy nhiên cũng liên quan đến sự thay đổi lượng trữ ẩm trên lưu vực qua các thời đoạn. Như vậy có thể thấy rằng dao động dòng chảy trung bình các thời đoạn không thể coi là một quá trình ngẫu nhiên thuần tuý. Và do đó cũng không thể dùng các quy luật thống kê với hàm phân bố xác suất một chiều để mô phỏng nó. Khi đó phải dùng các mô hình khác để mô phỏng dao động có tính đến mối quan hệ tương quan này. Thông dụng nhất hiện nay là mô hình tự hồi qui tuyến tính (hay mô hình Markov). Mô hình Markov có mô hình đơn hoặc phức. Mô hình Markov đơn chỉ xét mối liên hệ tương quan giữa các số hạng kề nhau, có hàm phân bố xác suất là hàm phân bố 2 chiều. Còn mô hình Marcov phức xét đến mối liên hệ xa hơn và hàm phân bố xác suất là hàm phân bố với số chiều lớn hơn 2.

Nói cách khác tập hợp dòng chảy trung bình các thời đoạn biểu thị một quá trình ngẫu nhiên. Để mô phỏng toán học các quá trình ngẫu nhiên người ta sử dụng các hàm phân bố và các thông số thống kê. Các phương pháp thống kê được xây dựng trên cơ sở các nguyên lý toán học miêu tả đặc tính biến động ngẫu nhiên của chuỗi quan trắc của một quá trình. Trong các phương pháp này người ta tập trung chú ý vào bản thân các kết quả quan trắc hơn là dựa trên các quá trình vật lý đã tạo ra các kết quả đó.

Các quá trình ngẫu nhiên có thể dừng hoặc không dừng, egôđích hoặc không egôđích. Quá trình ngẫu nhiên dừng là quá trình không bao hàm xu thế và chu kỳ, nó chỉ dao động xung quanh kỳ vọng. N ghĩa là trong quá trình ngẫu nhiên dừng các đặc trưng thồng kê như kỳ vọng, phương sai, hàm tự tương quan và mật độ phổ không thay đổi khi thay đổi thời gian tính toán. Quá trình ngẫu nhiên không dừng thì ngược lại, các đặc trưng này đều có thể thay đổi theo thời gian.

Về nguyên tắc quá trình ngẫu nhiên không dừng là tổng quát và chung nhất, các quá trình ngẫu nhiên dừng chỉ là một trường hợp riêng. Nhiều quá trình ngẫu nhiên không dừng có thể coi là dừng trong một khoảng thời gian gián đoạn hữu hạn nào đó. Ví dụ chuỗi dòng chảy trung bình tháng, tuần là chuỗi không dừng, còn chuỗi dòng chảy trung bình năm có thể coi là dừng vì khi đó qui luật bên trong năm bị loại trừ. Các quá trình ngẫu nhiên không dừng có thể trở thành dừng nhờ một số phép biến đổi(lọc sau đây:

a. Phép lọc sai phân.

- Nếu chuỗi có xu thế nhưng không có chu kỳ có thể dùng phép biến đổi sai phân(bậc 1 hoăc bậc 2)

Sai phân bâc 1 là chênh lệch giữa 2 giá trị kề nhau trong chuỗi

$$\Delta Q_t = Q_t - Q_{t-1} \qquad (3.2)$$

Trong đó : $Z_t = \Delta Q_t$ là giá trị sai phân bậc 1

Q_t, Q_{t-1} là giá trị thời đoạn trước và thời đoạn sau .

Nếu sai phân bậc 1vẫn còn thể hiện xu thế thì lấy sai phân lần nữa, được sai phân bậc 2. Sai phân bậc 2 chính là sai phân của sai phân bậc 1:

$$Z_t = \Delta^2(Q_t) = \Delta Q_t - \Delta Q_{t-1} = (Q_t - Q_{t-1}) - (Q_{t-1} - Q_{t-2}) \qquad (3.3).$$

Nếu sai phân bậc 2 chưa đạt được tính dừng ta có thể tiếp tục lấy sai phân bậc 3 hoặc cao hơn. Tuy nhiên trong thực tế chỉ cần lấy đến sai phân bậc 2 là đủ.

Trong nhiều trường hợp để đảm bảo tính tuyến tính thường dùng phép lọc logarit trước khi lấy sai phân, nghĩa là:

$$Z_t = \Delta(\ln Q_t) = \ln Q_t - \ln Q_{t-1} \qquad (3.3')$$

- Khi chuỗi có tính chu kỳ (mang tính mùa) phải dùng sai phân mùa để đưa chuỗi về dừng. Sai phân mùa là chênh lệch giá trị của hai quan trắc của L thời khoảng, trong đó L số thời khoảng mùa trong một năm. Chẳng hạn nếu là số liệu dòng chảy tháng ta có L =12. Do đó sai phân mùa bậc 1 là:

$$Z_t = Q_t - Q_{t-L2} = Q_t - Q_{t-12} \qquad (3.4)$$

Cũng có thể lấy sai phân bậc 2 của sai phân mùa bậc 1:

$$Z_t = (Q_t - Q_{t-L}) - (Q_{t-L-1}) \qquad (3.5)$$

b. Phép lọc loga.

Dùng phép lọc loga để tuyến tính hoá quan hệ hồi quy:

$$Z_t = \frac{\lg Q_t - \lg \overline{Q}_j}{\sigma_j} \qquad (3.6)$$

Trong đó Q_t là dòng chảy thời đoạn

$\overline{Q}_j$ là dòng chảy trung bình thời khoảng thứ j.

σ_j là khoảng lệch quân phương thời khoảng thứ j.

c. Phép lọc căn thức:

$$Z_t = \frac{\sqrt{Q_t} - \overline{Q}_j}{\sigma_j} \qquad (3.7)$$

Phép lọc này thường dùng cho lượng mưa tháng.

d. Cũng có thể dùng một số phép lọc đơn giản hơn để có chuỗi dừng:

-Để làm cho kỳ vọng bằng 0, ta có:

$$Z_t = Q_t - \overline{Q}_t \qquad (3.8)$$

-Để làm cho kỳ vọng bằng 1, ta có:

$$Z_t = \frac{Q_t}{\overline{Q}_y} \qquad (3.9)$$

3.1.1.2. Nguyên tắc mô phỏng .

Chuỗi dòng chảy thực đo Q_t có thể được đặc trưng bằng các tham số thống kê $\theta_1, \theta_2, ..., \theta_m$, trong đó phải kể đến các tham số quan trọng nhất là kỳ vọng (trị số trung bình), phương sai, hệ số biến đổi, hệ số lệch, hàm tự tương quan và mật độ phổ. Từ chuỗi quan trắc ta luôn thu được các ước lượng của các tham số θ_i. Mô phỏng chuỗi bằng mô hình toán tức là tạo ra chuỗi dòng chảy $Z_1, Z_2, ..., Z_n$ sao cho đảm bảo sự tương tự thống kê:

$$Z = \frac{M(P)}{N(D)} Q \qquad (3.10)$$

Các toán tử M(D), N(D), phụ thuộc vào bộ thông số thống kê θ_i được lựa chọn làm tiêu chuẩn đánh giá mô hình.

Chuỗi dòng chảy được mô phỏng bằng mô hình có độ dài N đủ lớn, phải đảm bảo có các tham số thống kê bằng với bộ tham số θ_i được tính từ chuỗi thực đo. Như vậy các tham số thống kê có từ chuỗi quan trắc đóng vai trò đặc trưng cho tổng thể, làm cơ sở cho phương pháp mô phỏng chúng. Hiển nhiên là mô hình hoá không làm chính xác thêm các thông số mà chỉ đưa ra thêm các thể hiện của quá trình ngẫu nhiên. Các thông số thống kê này là cơ sở để xác định các thông số của mô hình toán. Các thông số thống kê được xác định từ chuỗi thực đo thường có sai số do mẫu ngắn. Nó được chính xác hoá nhờ phương pháp bổ xung số liệu bằng các mô hình tất định nêu trong chương 2 hoặc các thuật toán sử lý thống kê đã được trình bày trong các giáo trình thuỷ văn khác.

Tiêu chuẩn đánh giá đồng thời cũng là lựa chọn mô hình tùy thuộc vào mục đích và phương pháp mô hình hoá. Trong các bài toán tính toán và dự báo thủy văn, bộ thông số thường được chọn là các đặc trưng thống kê cơ bản của chuỗi số như kỳ vọng, phương sai, hệ số lệch, hàm tự tương quan và mật độ phổ của toàn bộ chuỗi số cũng như từng thời đoạn tính toán. Trong các bài toán thủy lợi người ta còn quan tâm đến tổ hợp các nhóm năm nhiều hay ít nước, thời kỳ gián đoạn cấp nước và hệ số dung tích.

Sau đây chúng ta sẽ xem xét các thông số thống kê quan trọng biểu thị một chuỗi số ngẫu nhiên:

3.1.1.3. Các đặc trưng thống kê của chuỗi thủy văn.

Khi phân tích các chuỗi ngẫu nhiên thủy văn ta còn chú ý đến các thông số thống kê theo toàn bộ chuỗi và các thông số của các thời đoạn trong chuỗi:

a. Các thông số thống kê của toàn chuỗi:

-Kỳ vọng: $M(Q) = \dfrac{\sum_{i=1}^{n} Q_i}{n}$

-Phương sai(Varian): $D(Q) = \sigma^2(Q) = \dfrac{\sum_{i=1}^{n} \left(Q_i - \overline{Q}\right)^2}{n}$

-Hệ số biến đổi(Change coeficient): $C_v(Q) = \dfrac{\sigma(Q)}{M(Q)}$

-Hệ số lệch: $C_S(Q) = \dfrac{\sum \left(Q - \overline{Q}\right)^3}{\sigma^3(Q)}$ (3.11)

b. Các thông số thống kê của các thời đoạn

- $M(j) = \dfrac{1}{n}\sum_{t=1}^{n} Q(tj)$

- Phương sai: $D(j) = \sigma^2(j) = \dfrac{1}{n}\sum_{t=1}^{n} \left[Q_{tj} - M(j)\right]^2$

- Hệ số biến đổi: $C_v(j) = \dfrac{\sigma^2(j)}{M(j)}$

- Hệ số lệch: $C_S(j) = \frac{1}{n} \cdot \frac{\sum_{t=1}^{n}\left[Q_{tj} - M(j)\right]^3}{\sigma^3(j)}$ (3.12)

c. Một số đặc trưng đánh giá chất lượng mô hình:

- Sai số tuyệt đối trung bình MAE (mean absolute error)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i| \quad \text{với } e_i = Q_i - Q'_i \qquad (3.13)$$

Trong đó: y_i là giá trị thực,

y'_i là giá trị theo mô hình.

- Sai số tương đối trung bình MAPE (mean of the absolute percentage error):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|e_i|}{Q_i} \qquad (3.14)$$

- Sai số bình phương trung bình: MSE (mean square error)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}e_i^2 \qquad (3.15)$$

- Sai số chuẩn RMSE (the root mean square error- Standard error)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}e_i^2}{n}} \qquad (3.16)$$

- Sai số quân phương kiểm tra:

$$S_e = \sqrt{\frac{\sum_{i=1}^{n}e_i^2}{n-m}} = \sqrt{\frac{\sum\left(Q_i - \overline{Q}\right)^2}{n-m}} \qquad (3.17)$$

Trong đó m là số bậc tự do của mô hình (số các thông số)

- Hệ số tất định R^2:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} Q_i^2 - \frac{\left(\sum Q_i\right)^2}{n}} \tag{3.18}$$

- Hệ số tất định phù hợp:

$$R_{ph}^2 = \frac{S_e^2}{\sigma^2} = \frac{\sqrt{\frac{\sum e_i^2}{n-m}}}{\sqrt{\frac{\sum\left(Q_i - \overline{Q}\right)^2}{n-1}}} \tag{3.19}$$

d. Hệ số tự tương quan và hàm tự tương quan

Đặc trưng thể hiện mối liên hệ giữa các giá trị trong chuỗi là covarian và hệ số tương quan giữa chúng.

* Covarian (hiệp phương sai) được xác định theo biểu thức:

$$\gamma_K = \text{cov}\left[Q_i, Q_{i-K}\right] = \left[\left(Q_i - M(Q)\right)\left(Q_{i+K} - M(Q)\right)\right] \tag{3.20}$$

i= 1,2,...,n ; K= 1,2,...,m<n

Khi K = 0 thì covarian chính là phương sai của chuỗi số:

$$\gamma_0 = \sum\left[\left(Q_i - M(Q)\right)\right]^2 = \sigma^2(Q) = D(Q)$$

* Hệ số tự tương quan và hàm tự tương quan

Đó là tỷ số covarian bậc K và covarian bậc 0 của chuỗi số :

$$S_K = \frac{\gamma_K}{\gamma_0} = \frac{\sum\left[\left(Q_i - M(Q)\right)\left(Q_{i-K} - M(Q)\right)\right]}{\sum\left(Q_i - M(Q)\right)^2} = \frac{\gamma_K}{\sigma(Q)} \tag{3.21}$$

Trường hợp chuỗi số là rời rạc và hữu hạn ta có covarian bậc K là :

$$C_K = \frac{\sum_{i=1}^{n-K}\left(Q_i - \overline{Q}\right)\left(Q_{i-K} - \overline{Q}\right)}{n} \tag{3.22}$$

và hệ số tự tương quan là:

$$r = \frac{C_K}{C_0} \qquad (3.23)$$

Tổng hợp các giá trị của r_K lập thành hàm tự tương quan, $r_K \to 0$ khi K đủ lớn. Vì vậy giá trị trung bình của hệ số tự tương quan thường không được chú ý mà người ta chỉ quan tâm đến phương sai (varian) của chúng. Khi mẫu lớn thì r_K có phân bố chuẩn với trung bình bằng 0 và phương sai là:

$$Var(r_K) = \frac{1}{n}\left[1 + 2\left(r_1^2 + r_2^2 + ... + r_{K-1}^2\right)\right] \qquad (3.24)$$

Độ lệch chuẩn của hệ số tự tương quan là:

$$SE(r_K) = \sqrt{\mathrm{var}(r_K)} \qquad (3.25)$$

Để kiểm tra mức độ ý nghĩa của hàm tự tương quan cho một mẫu có độ dài khá lớn người ta sử dụng chỉ tiêu thống kê t:

$$t = \frac{r_K - 0}{\sqrt{\mathrm{var}(r_K)}} \qquad (3.26a)$$

hay $$\mathrm{t} = \frac{\mathrm{r_K}}{\frac{1}{\mathrm{n}}\sqrt{1 + 2\sum \mathrm{r_K^2}}} \qquad (3.26b)$$

Nếu $|t| < t_\alpha$, trong đó t_α là giá trị xác định theo bảng Fisher với mức ý nghĩa α thì hệ số tự tương quan r_K là không có ý nghĩa hay nói cách khác là không có tự tương quan bậc k (r_K= 0). Trong kiểm tra thống kê thường lấy mức ý nghĩa α = 5% = 0.05, khi đó có thể lấy t_α = 2 là giá trị tiêu chuẩn để kiểm tra. Khi $|t| < 2$ có thể coi là r_K=0 và cần chú ý khi xây dựng mô hình để mô phỏng.

*** Hệ số tự tương quan riêng và hàm tự tương quan riêng**

Hệ số tự tương quan riêng là để đo mối liên hệ giữa hai giá trị Q_i và Q_{i-K} khi hiệu quả của các giá trị khác bị loại trừ hay không đổi.

Giá trị hệ số tự tương quan riêng được xác định:

$$r_{KK} = \frac{r_K - \sum (r_{K-1,j})(r_{K-j})}{1 - \sum (r_{k-1,j})(r_j)} \tag{3.27}$$

Trong đó: j =1,2,...,k-1;

r_K là hệ số tự tương quan tại K bước

r_{Kj} là hệ số tự tương quan tại K bước về trước khi hiệu quả của j bước tiếp theo bị loại trừ.

$$r_{Kj} = r_{K-1,j} - (r_{KK})(r_{K-1,K-j})$$

(3.28)

Chú ý rằng: $r_{11} = r_1$ (3.29)

Các hệ số tự tương quan riêng r_{KK} cũng lập thành một hàm gọi là hàm tự tương quan riêng.

Cũng tương tự như hệ số tự tương quan, để kiểm tra mức độ ý nghĩa của từng hệ số r_{KK} người ta sử dụng chỉ tiêu thống kê:

$$t' = \frac{r_{KK} - 0}{\sqrt{\mathrm{var}(r_{KK})}} \tag{3.30}$$

với Var(r_{KK})=1/n (Quenouille, 1949) (3.31)

Với mức ý nghĩa α = 0,05 cũng có thể lấy giá trị tiêu chuẩn t_α =2

Khi t' tính được theo (3.23) mà $|t'| < t_\alpha$ thì hệ số tự tương quan riêng r_{KK} là không có ý nghĩa.

*** Hàm mật độ phổ G(f):**

$$G(f) = \frac{P(f)}{\gamma_0} = 2\left[1 + 2\sum_{K=1}^{\infty} r_K \cos 2\pi f_K . K\right] \tag{3.32}$$

Trong đó : r_K là giá trị hàm tự tương quan.

f_K là tần số phổ xác định theo công thức:

$$f_K = \frac{K}{2\tau_m \Delta t} \tag{3.33}$$

Một số nghiên cứu cho thấy hàm phổ kinh nghiệm tính toán cho các chuỗi có độ dài khác nhau ít thay đổi hơn hàm tự tương quan, do vậy người ta thích dùng hơn. Tuy nhiên cần nhấn mạnh rằng hàm phổ là biến đổi Fourie của hàm tự tương quan, do đó những yếu tố gây mất ổn định cho hàm tự tương quan cũng ảnh hưởng đến hàm mật độ phổ.

e. Hàm phân bố xác suất:

Như đã nêu ở phần trước, với chuỗi các đại lượng ngẫu nhiên độc lập chúng ta có hàm phân bố xác suất một chiều, còn với chuỗi ngẫu nhiên tương quan chúng ta có hàm phân bố xác suất nhiều chiều.

*Hàm phân bố xác suất một chiều:

Chuỗi dòng chảy cực hạn và hầu hết chuỗi dòng chảy năm được coi là các đại lượng ngẫu nhiên độc lập và được mô phỏng bằng các mô hình phân bố xác suất một chiều. Giá trị của các đại lượng chỉ phụ thuộc vào xác suất vượt (tần suất) hay thời kì xuất hiện lại mà không phụ thuộc vào thời gian và các giá trị xuất hiện trước đó.

Có rất nhiều dạng hàm phân bố một chiều được đề xuất để mô phỏng phân bố xác suất của các đại lượng thuỷ văn. Bằng cách thử lựa một phân bố xác suất cho vừa khớp với số liệu thuỷ văn, rất nhiều thông tin thống kê của mẫu có thể tổng kết một cách cô đọng trong hàm số và trong các thông số của hàm. Các thông số thống kê cho ta những thông tin thiết yếu của một tập số liệu thu gọn một dãy số thành một tập nhỏ hơn. Chúng là các số đặc trưng của tổng thể. Đó là các thông số kì vọng μ ,phương sai σ^2, hệ số biến đổi C_v, hệ số lệch C_S và đôi khi có cả hệ số nhọn C_e.

Lựa chọn một hàm phân bố xác suất cho phù hợp với số liệu thực nghiệm có thể tiến hành bằng phương pháp mômen hoặc phương pháp khả thi lớn nhất.

***Hàm phân bố xác suất nhiều chiều:**

Khi xem xét đến mối liên hệ tương quan giữa các giá trị trong chuỗi ta phải xem xét đến hàm phân bố xác suất có điêu kiện, nghĩa là xác suất xuất hiện của số hạng sau phụ thuộc vào giá trị của số hạng trước. Nghĩa là khi đó phải xem xét hàm phân bố xác suất nhiều chiều.

Hàm phân bố xác suất đồng thời của hai đại lượng ngẫu nhiên x,y có dạng sau:

$$F(x,y) = P\{X \le x, Y \le y\} \qquad (3.34)$$

$$\text{hay}: F(x,y) = \int_0^X \int_0^Y f(x,y)dxdy \qquad (3.35)$$

trong đó: F(x,y) là hàm mật độ xác suất.

Từ hàm phân bố đồng thời (3.35) có thể suy ra các hàm một chiều và có điều kiện tương ứng.

$$F_1(x) = F(x,\infty) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(x,y)dx \qquad (3.36a)$$

$$F_2(y) = F(y,\infty) = \int_{-\infty}^{y} \int_{-\infty}^{\infty} f(x,y)dy \qquad (3.36b)$$

$$F_3(x/y) = \frac{F(x,y)}{F_2(y)} \qquad (3.37a)$$

$$F_4(y/x) = \frac{F(x,y)}{F_1(x)} \qquad (3.37b)$$

Trong khi mô phỏng phân bố xác suất các thời đoạn ngắn trong năm các hàm phân bố thống kê một chiều thường dùng không còn phù hợp nữa. Kartvelixvenli đã giới thiệu một phân bố có tính mềm dẻo hơn cho dòng chảy tháng, đó là một biến dạng của phân bố chuẩn:

$$Z(u) = \alpha \ln(1+u) + \beta \ln u + \gamma \sqrt{u}$$

Trong đó: u là giá trị ban đầu .

α,β,γ là các thông số.

Tuy nhiên coi rằng tất cả các tháng có cùng một phân bố cũng chưa hợp lý, vì bản chất hình thành quá trình thủy văn của các tháng không như nhau. Đa số sông ngòi có tỷ số Cs/Cv của các tháng dao động trong phạm vi khá rộng. Vì vậy khó có thể tiếp nhận một hàm phân bố chung với cùng tỷ số Cs/Cv. Hàm phân bố P.III tỏ ra mềm dẻo hơn, nhưng cũng không đáp ứng do độ phân bố quá rộng của dòng chảy các tháng. Xvanhidze (1977) dề nghị sử dụng hàm JonsonB, giới hạn hai đầu làm hàm phân bố xác suất chung cho tất cả các tháng để mô hình hoá cho dòng chảy tháng. Hàm phân bố JonsonB mềm dẻo và đa dạng hơn, đồng thời có liên hệ bằng một quan hệ đơn giản. Tuỳ thuộc vào giới hạn dưới và giới hạn trên có thể nhận được một giá trị bất kì của tỷ số Cs/Cv. Khi thay đổi giới hạn của phân bố cũng làm thay đổi cả tính đối xứng của nó. Phân bố JonsonB chiếmn một khu vực rộng bao hàm có phân bố chuẩn, phân bố gamma và một số phân bố khác.

Khi xem xét mối liên hệ xa hơn, ta phải xem xét hàm phân bố xác suất đồng thời nhiều chiều. Hàm này có dạng:

$$F(x_1, x_2, ..., x_m) = P\{X_1 \le x_1, X_2 \le x_2, ..., X_m \le x_m\}$$

hay
$$F(x_1, x_2, ..., x_m) = \int_{-\infty}^{X_1} \int_{-\infty}^{X_2} ... \int_{-\infty}^{X_m} f(x_1, x_2, ..., x_m) dx_1 dx_2 ... dx_m \qquad (3.39)$$

Trong đó $f(x_1,x_2,...,x_m)$ là hàm mật độ xác suất nhiều chiều.

Hàm phân bố có điều kiện biểu thị:

$$F_1(x_1 / x_2, x_3, ..., x_m) = \int_{-\infty}^{X_1} \int_{-\infty}^{\infty} f(x_1, x_2, ..., x_m) dx_2 dx_3 ... dx_m \qquad (3.40)$$

Với các thời đoạn nhỏ hơn một tháng người ta chưa tiến hành khảo sát hàm phân bố xác suất của chúng vì sự phức tạp của chuỗi số liệu. Tuy nhiên có thể sử dụng một số hàm phân bố của dòng chảy tháng để áp dụng mô phỏng.

f. Sự phân bố các nhóm năm.

Trong khi mô phỏng dao động của dòng chảy năm, một đặc trưng rất cần lưu ý, đặc biệt trong tính toán thủy lợi là sự hình thành các nhóm năm nhiều, ít nước.

Có nhiều quan điểm khác nhau về sự hình thành các nhóm năm. Một số nhà nghiên cứu cho rằng sự xuất hiện của các nhóm năm mang tính chu kỳ, tương ứng với hoạt động của vết đen mặt trời, một số khác không thừa nhận. Tuy nhiên sự xuất hiện các nhóm năm là có thực, nó gây ảnh hưởng không nhỏ đến các hoạt động của nền kinh tế quốc dân. Ratcovich cho rằng hợp lý hơn nên coi rằng sự xuất hiện các nhóm năm cũng mang ý nghĩa xác suất tức là cũng hình thành hàm phân bố với cùng các thông số thống kê của chúng. Một số đặc trưng quan trọng của quy luật phân bố nhóm năm là độ dài nhóm năm và độ lặp lại của nó. Thay cho việc phân tích nhóm năm nhiều hay ít nước đối với giá trị trung bình nhiều năm, người ta thường hay sử dụng các nhóm năm đối với một phân vị xác suất (Probablity Quantile).

* Độ dài trung bình nhóm năm:

$$\overline{m} = \frac{m_0}{1 + 1/np} \qquad (3.41)$$

* Hệ số biến đổi:

$$C_v = \sqrt{\frac{1 - p - 1/n}{1 + 1/np}} \qquad (3.42)$$

Trong đó: n là dung lượng mẫu quan trắc

p là điểm phân vị xác suất

m_0 là độ dài trung bình nhóm năm của các đại lượng ngẫu nhiên độc lập: $m_0 = \frac{1}{p}$.

* Độ lặp lại của nhóm năm có độ dài m là:

$$P = \frac{\sum p(i)m(i)}{N} \qquad (3.43)$$

Trong đó: P là độ lặp lại,

P(i) là xác suất xuất hiện nhóm năm có độ dài m(i)

N là dung lượng mẫu.

Như vậy dưới tác động của nhiều nhân tố, các quá trình thủy văn nói chung là một quá trình ngẫu nhiên. Nguyên tắc cơ bản để mô phỏng của mô hình ngẫu nhiênlà phải đảm bảo đúng đắn quy luật dao động của các quá trình thủy văn, đảm bảo sự phù hợp của các đặc trưng thống kê của cả tổng thể toàn chuỗi, cũng như từng thời đoạn trong đó.

3.1.2. Cấu trúc của mô hình ngẫu nhiên

Như trên đã trình bày, các quá trình thuỷ văn là một quá trình ngẫu nhiên. Ngoài một số quá trình riêng biệt được coi là các đại lượng ngẫu nhiên độc lập, còn trong đa số các quá trình ngẫu nhiên, giữa các số hạng của chúng có mối liên hệ tương quan với nhau, đặc biệt là giữa các số hạng kề nhau. Ta có thể thấy rằng mỗi giá trị tại thời điểm t được xác định từ các số hạng đứng trước nó t-i (i=1,2,...p;t=1,2,...n) nghĩa là ta có phương trình:

$$Q_t = f(Q_{t-1}, Q_{t-2}, ..., Q_{t-P}) \qquad (3.44)$$

Phương trình (3.44) có thể là phi tuyến. Tuy nhiên trong trường hợp quá trình là dừng hoặc thông qua phép biến đổi đưa về quá trình dừng thì có thể coi Q_t là tổ hợp tuyến tính của các Q_{t-I}:

$$Q_t = \sum_{i=1}^{P} a_i Q_{t-i} \qquad (3.45)$$

Trong đó: a_i là các hệ số.

Đây chính là một quá trình tự hồi qui cấp p. Nếu quá trình (3.45) là một quá trình Gauxơ thì nó có thể coi là một quá trình Markov theo nghĩa rộng (Xvanhidze,1977).

Khi coi Q_t là một tổ hợp tuyến tính của các Q_{t-i} thì giá trị xác định theo (3.45) so với giá trị thực đo sẽ gặp một sai số. Để hiệu chỉnh sai số này người ta đưa vào một thành phần ngẫu nhiên $R_t(\xi)$. Thành phần ngẫu nhiên này cũng thay đổi tương ứng với Q_t. Khi đó (3.45) có dạng:

$$Q_t = \sum_{i=1}^{P} a_i Q_{t-i} + \varepsilon_t(\xi) \qquad (3.46a)$$

Cũng có thể hiểu theo một nghĩa khác, giá trị Q_t tính được theo (3.45) chỉ là giá trị trung bình có điều kiện của Q_t khi chịu ảnh hưởng của các Q_{t-i}. Giá trị thực của Q_t sẽ lệch khỏi giá trị trung bình có điều kiện này một độ lệch xác suất Q_t. Khi đó phương trình được viết lại thành:

$$Q_t = \sum_{i=1}^{P} Q_i Q_{t-i} + \phi a_0 \sigma \qquad (3.46b)$$

với $a_0 = \sigma\sqrt{\dfrac{D}{D_{ii}}}$.

Bản thân $\varepsilon_t(\xi)$ cũng có thể là một tổ hợp tuyến tính của các ε_{t-j} đứng trước

$$\varepsilon_t = \sum b_j \varepsilon_{t-j} \qquad (3.47)$$

Như vậy một mô hình ngẫu nhiên gồm có hai thành phần:

-Thành phần tự hồi quy có thể coi là thành phần tất định,

- Thành phần ngẫu nhiên.

Sau đây chúng ta xem xét kỹ hơn các thành phần này.

3.1.2.1. Thành phần tự hồi quy.

Thành phần tự hồi quy được xác định từ mối liên hệ tuyến tính giữa giá trị Q_t và các giá trị trước đó. Thường thì phải thông qua một phép biến đổi để đưa về một chuỗi dừng Z_t. Như vậy tổ hợp tuyến tính bây giờ là giữa các đại lượng đã biến đổi .

$$Z_t = \sum_{i=1}^{P} a_i Z_{t-i} \qquad (3.48)$$

Các phép biến đổi để đưa về chuỗi dừng đã được trình bày trong mục (3.1.1).

Kết quả cuối cùng phải đưa trở lại Q_t theo phép biến đổi ngược lại.

Việc lựa chọn bậc hồi quy p (hay số các số hạng liên hệ) là rất quan trọng. Tiêu chuẩn chung để lựa chọn bậc hồi quy p là cực tiểu phương sai của giá trị tính theo mô hình σ_P và khi tỷ số $\left(\frac{\sigma_P}{\sigma_{P-1}}\right)^2$ gần bằng 1, nghĩa là tăng thêm một bậc hồi quy thì phương sai không thay đổi (Xvanhiđze, 1977), hoặc tại đó hàm tự tương quan có bước nhảy đột ngột (Box-Jenkin).

Một số mô hình khác lại lựa chọn từng phương trình riêng cho từng thời đoạn mô phỏng, tức là chỉ xét đến bậc hồi quy p=1 (Thormat-Fiering). Giá trị tính được của thời đoạn này lại được coi là giá trị thực để tính toán cho thời đoạn tiếp theo. Và cứ tiếp tục như vậy cho từng thời đoạn.

Khi chỉ xét mối liên hệ với một số hạng đứng trước ta có mô hình tự hồi quy bậc 1 (AR(1)) hay là xích Markov đơn. Còn khi xét đến mối liên hệ đến p số hạng đứng trước ta có mô hình tự hồi quy bậc p (AR(p)), hay xích Markov phức.

Như vậy dạng tổng quát của mô hình tự hồi quy bậc 1 là:

$$Z_t = a_1 Z_{t-1} + a_2 Z_{t-2} + \dots + a_p Z_{t-p} + \varepsilon_t \qquad (3.49)$$

3.1.2.2. Thành phần ngẫu nhiên.

Thành phần ngẫu nhiên chính là thành phần sai số hay phần dư giữa giá trị thực và giá trị tính được theo mô hình tự hồi qui. Thành phần này khi làm dự báo là sai số , còn khi tại chuỗi mô phỏng nó là số ngẫu nhiên.

Việc xác định thành phần ngẫu nhiên ε_t tuỳ thuộc vào ý đồ và tiêu chuẩn mô phỏng của mô hình. Về cơ bản theo nguyên tắc mô phỏng là đảm bảo cho các thông số thống kê của chuỗi số không đổi. Như vậy:

$$\varepsilon_t = \alpha \xi_t \qquad (3.50)$$

Trong đó: ξ_t là số ngẫu nhiên có phân bố chuẩn với trung bình bằng 0 và phương sai bằng 1.

Trong nhiều trường hợp đại lương ε_t có tương quan với q các giá trị ε_{t-j}

(=1,2,...,q) trước đó, khi đó ta có quan hệ MA(q):

$$\varepsilon_t = b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + ... + b_q\varepsilon_{t-q} \quad (3.51)$$

Trường hợp tổng quát ta có mô hình ARIMA(p,q)

$$Z_t + a_1Z_{t-1} + ... + a_pZ_{t-p} = b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + ... + b_q\varepsilon_{t-q} - \varepsilon_t \quad (3.52)$$

Thực chất giá trị tính được Q'_t theo quan hệ tự hồi qui chỉ là giá trị trung bình có điều kiện. Giá trị thực sẽ lệch khỏi giá trị Q'_t một độ lệch xác xuất nào đấy, tuỳ thuộc dạng hàm phân bố xác xuất. Trong trường hợp đó thay cho thành phần ngẫu nhiên ε_t là phần dư ta coi ε_t là độ lệch xác xuất. Theo Chow(1964) ta có quan hệ:

$$Q_t = Q'_t + K_t\sigma_t \quad (3.53)$$

Trong đó: Q'_t được xác định từ quan hệ tự hồi quy (3.52),khi lấy giá trị thực Q.

K_t là độ lệch xác xuất

σ_t là khoảng lệch quân phương (phương sai) có điều kiện của đại lượng Q_t.

Hai giá trị K_t và σ_t được xác định tuỳ thuộc dạng hàm phân bố có điều kiện, các đặc trưng thông kê của nó và vào dạng tương quan giữa các đại lượng ngẫu nhiên.

Người ta thừa nhận một giả thiết rằng(Kritski-Menken,1977), trong trường hợp các đại lượng ngẫu nhiên có phân bố chuẩn thì tương quan giữa chúng là tương quan chuẩn, còn với các đại lượng ngẫu nhiên có phân bố Gamma thì tương quan giữa chúng là tương quan gamma. Các tương quan này sẽ chi phối biểu thức xác định các đặc trưng thống kê của các hàm phân bố có điều kiện, cũng tức là chi phối thành phần ngẫu nhiên trong mô hình. Chúng ta sẽ xem xét chi tiết hơn các vấn đề này trong mục xác định các thông số mô hình.

Có thể thấy mối tương tự về hình thức giữa mô hình tất định và mô hình ngẫu nhiên. Thật vậy với mô hình tự hồi qui AR(p) ta có:

$$\varphi(B)Z_t = \varepsilon_t \quad (3.53)$$

Trong đó: $\varphi(B)= 1-a_1B- a_2B^2-...-a_pB^p$ (3.54)

Còn với mô hình trung bình trượt MA(q) có:

$$Z_t= \theta(B).\ \varepsilon_t \quad (3.55)$$

Trong đó: $\theta(B)=1-b_1B-b_2B^2-...-b_qB^q$ (3.56)

Như vậy ta thấy $\varphi(B)$, $\theta(B)$ là các hàm chuyền hay hàm lọc. ε_t đóng vai trò của hàm vào và Z_t đóng vai trò là hàm ra. Dãy số ngẫu nhiên ε_t lọc qua hàm truyền ta được dãy số Z_t. Về hình thức các mô hình ngẫu nhiên trên không khác gì mô hình tất định và tương ứng với mô hình hệ thống thủy văn(1.4), nhưng về hình thức có sự khác nhau rất lớn. Trong mô hình tất định, mưa là hàm vào, lọc qua hàm truyền ta được hàm ra là dòng chảy. Còn ở mô hình ngẫu nhiên hàm vào là dãy ngẫu nhiên ε_t lọc qua các hàm truyền $\varphi(B)$, $\theta(B)$ để có hàm ra Z_t nhưng không thể coi dãy ngẫu nhiên ε_t gây ra dòng chảy Z_t. Về bản chất mô hình ngẫu nhiên không giải thích nguyên nhân và kết quả như mô hình tất định.

3.2. CÁC LOẠI MÔ HÌNH NGẪU NHIÊN

Như đã phân tích trong chương 1, theo cây phân loại (1.4) có mô hình ngẫu nhiên tương quan và mô hình ngẫu nhiên độc lập. Theo thứ tự của cây phân loại ta xét tương quan không gian trước khi xem xét tương quan thời gian. Tuy nhiên trong thực tế ứng dụng người ta lại quan tâm trước hết đến các mô hình ngẫu nhiên độc lập và tương quan theo thời gian, sau đó mới mở rộng xét đến tương quan không gian. Vì vậy trong phần này, phân tích tập trung vào 2 loại này là chủ yếu, còn mô hình tương quan không gian được đề cập đến trong một mức độ cần thiết. Nói chung chuỗi số liệu thủy văn lập thành một quá trình ngẫu nhiên có liên hệ tương quan. Tuy nhiên một số đại lượng, chẳng hạn như dòng chảy cực hạn, lại có thể coi là dãy các đại lượng ngẫu nhiên độc lập. Khi đó mô hình mô phỏng là các hàm phân bố xác xuất 1 chiều. Kết quả tính toán chỉ phụ thuộc vào xác xuất vượt(hay thời kỳ xuất hiện lại) mà không liên quan đến thời gian xuất hiện chúng. Tuy vậy các kết quả tính toán này rất có giá trị trong các bài toán thiết kế , qui hoạch công trình .

3.2.1. Mô hình ngẫu nhiên độc lập thời gian

Để mô phỏng phân bố xác xuất của các đại lượng ngẫu nhiên độc lập người ta dùng hàm phân bố 1 chiều. các hàm này đã được nhắc đến trong nhiều tài liệu, giáo trình về thủy văn, ở đây chỉ liệt kê lại một số hàm chủ yếu, mà không nhắc lại toàn bộ các thuật toán xác định các tham số của nó.

3.2.1.1. Hàm phân bố chuẩn.

Hàm phân bố chuẩn được xuất phát từ định lý giới hạn trung tâm , định lý này phát biểu rằng nếu các đại lượng ngẫu nhiên Q_i là dãy độc lập và phân bố đồng nhất với trung bình là μ và phương sai là σ^2 thì phân bố của tổng các biến ngẫu nhiên đó $y=\sum_{i1}^{n} Q_i$ sẽ tiệm cận về phân bố chuẩn với trung bình là nμ và phương sai là $n.\sigma^2$ khi n trở thành rất lớn. Điều quan trọng ở đây là định lý luôn nghiệm đúng cho dù biến ngẫu nhiên Q có hàm phân bố xác xuất như thế nào, chẳng hạn phân bố xác xuất của số trung bình cộng của mẫu $\overline{Q}=\frac{1}{n}\sum_{i1}^{n} Q_i$ có thể được xấp xỉ bằng một phân bố chuẩn có trung bình là μ và phương sai là $\frac{\sigma^2}{n}$, kết luận này không phụ thuộc gì vào dạng hàm phân bố xác xuất của Q. Các biến lượng thuỷ văn như lượng mưa năm nếu được tính toán như tổng các hiệu ứng của nhiều biến cố độc lập thì có xu thế tuân theo luật phân bố chuẩn. Tuy nhiên việc sử dụng phân bố chuẩn để mô phỏng các hiện tượng thủy văn gặp phải những hạn chế. Đó là biến ngẫu nhiên trong phân bố chuẩn biến thiên liên tục trong phạm vi (-∞,∞), trong khi hầu hết các đại lượng thủy văn là các đại lượng không âm. Mặt khác phân bố chuẩn là một phân bố đối xứng quanh giá trị trung bình trong khi các số liệu thủy văn biểu hiện rõ xu thế phân bố lệch.

Nhưng bằng cách biến đổi chuẩn hoá ta có thể đưa các đại lượng thủy văn về dạng phân bố chuẩn theo dạng sau:

$$Z = \frac{Q_P - \mu}{\sigma} \tag{3.57}$$

Hàm mật độ phân bố chuẩn có dạng sau:

$$F(Q) = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{(Q-\mu)^2}{2}} \tag{3.58}$$

Và hàm phân bố xác xuất là:

$$F(Q) = \int_{-\infty}^{x} f(Q)dQ = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{(Q-\mu)^2}{2}}dQ \qquad (3.59)$$

Trong dạng chuẩn hoá chúng ta có hàm phân bố và hàm mật độ:

$$\begin{aligned} f(z) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} \\ F(z) &= \int_{-\infty}^{Z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \end{aligned} \qquad (3.60)$$

Với biến u là biến hình thức. Chúng ta không có biểu thức giải tích của tích phân trên. Giá trị của nó có thể tra theo bảng tính sẵn, hoặc bằng công thức sau (Abramowitz và Stegun):

$$B = \frac{1}{2}\left[1 + 0.196854\ |Z| + 0.115194\ |Z|^2 + 0.000344\ |Z|^3 + 0.019527\ |Z|^4\right]^{-4} \quad (3.62)$$

Trong đó $|Z|$ là giá trị tuyệt đối của biến chuẩn hoá Z và hàm phân bố chuẩn lấy giá trị:

$$f(z) = \begin{cases} B & víi\ Z < 0 \\ 1 - B & víi\ Z > 0 \end{cases} \qquad (3.63)$$

Sai số tính Z theo công thức này nhỏ hơn 0.00025. Ngược lại ta có thể xác định được giá trị Z khi biết tần xuất (xác suất vượt) P=f(z) hay biết thời kỳ xuất hiện tại T (vì P = 1/T) thông qua một biến trung gian W:

$$W = \left[\ln\left(\frac{\mathbf{1}}{p^2}\right)\right]^{1/2} \qquad \text{vớ i } (\mathbf{0} < p \le \mathbf{0.5}) \qquad (3.64)$$

$$Z = W - \frac{0.515517 + 0.802853\,W + 0.010328\,W^2}{1 + 1.432788\,W + 0.189629\,W^2 + 0.001308\,W^3} \qquad (3.65)$$

Khi p > 0.5 thay p bằng (1-p) trong (3.64) và giá trị Z sau khi tính được bằng công thức (3.65) được gán dấu âm. Sai số tính Z theo công thức này nhỏ hơn 0.00045(Abra mowitz và Stegun).

Giá trị thực tế Q_P ứng với tần xuất p được suy từ công thức (3.57).

$$Q_p = \mu + \sigma Z \qquad (3.66)$$

Tuy nhiên có nhiều hàm phân bố không thể trực tiếp nghịch chuyển từ p=F(Q) sang Q, do vậy Chow đã đề nghị một công thức chuyển đổi sau:

$$Q_p = \mu(Q) + \phi\sigma(Q) \qquad (3.67)$$

Trong đó μ là kỳ vọng và σ^2 là phương sai. Biểu thức này có thể xấp xỉ bởi biểu thức:

S là khoảng lệch chuẩn mẫu,

ϕ_p là hệ số lệch xác xuất tương ứng với p.

Trong trường hợp phân bố chuẩn $\phi_p \equiv Z$.

3.2.1.2. Hàm phân bố log - chuẩn.

Đại lượng ngẫu nhiên Q có phân bố log-chuẩn nếu đại lượng Y=logQ có phân bố chuẩn. Theo Chow phân bố này có thể áp dụng cho các đại lượng thủy văn được tạo thành bởi tích của nhiều đại lượng thủy văn khác. Bởi vì nếu $Q = Q_1, Q_2 .... Q_n$ thì $Y=\lg Q=\sum_{i=1}^{n}\lg Q_i = \sum_{i=1}^{n} Y_i$ sẽ có phân bố dẫn đến chuẩn khi n đủ lớn khi các đại lượng Q_i là độc lập và có phân bố đồng nhất. Phân bố log-chuẩn có nhiều lợi thế hơn phân bố chuẩn vì nó bị chặn dưới (Q>0) và phép biến đổi logarit có xu thế làm giảm nhỏ hệ số lệch dương, là điều thường gặp trong các chuỗi số thủy văn. Bởi vì khi lấy logarit các số lớn bị thu nhỏ theo tỷ lệ lớn hơn nhiều so với các số nhỏ.

Hàm phân bố có dạng:

$$f(Q) = \frac{1}{Q\sigma\sqrt{2\pi}}\exp\frac{(y-\mu y)^2}{2\sigma_y^2} \text{ với } Q > 0 \qquad (3.69)$$

Với Y=logQ ;

μ_y và σ_y^2 là trung bình và phương sai của đại lượng đã logarit Y=logQ.

Hạn chế của phân bố này là ở chố nó chỉ có 2 tham số và đòi hỏi các giá trị logarit của chuỗi số phải đối xứng xung quanh giá trị trung bình.

Đối với hàm log chuẩn vẫn áp dụng các thủ tục tính toán như đối với phân bố chuẩn theo các công thức (3.64), (3.65) nhưng không phải trực tiếp với các đại lượng Q mà đối với các logarit của nó. Trong các công thức tính toán (3.67) và (3.68) lấy giá trị trung bình và độ lệch chuẩn theo số liệu đã logarit hoá:

3.2.1.3. Phân bố mũ

Một số các sự kiện thủy văn, chẳng hạn quá trình xẩy ra mưa có thể được coi là các quá trình Poisson, trong đó sự kiện xuất hiện một cách tức thời và độc lập với nhau trong một lớp thời gian hoặc dọc theo tuyến. Khoảng thời gian giữa các sự kiện hay còn gọi là khoảng thời gian đến trung gian được mô tả bởi một phân bố mũ với tham số λ ,tham số này được định nghĩa như là tốc độ trung bình của sự xuất hiện sự kiện. Người ta thường dùng phân bố mũ để mô tả các khoảng thời gian đến trung gian của các đột biến ngẫu nhiên trong hệ thống thủy văn, chẳng hạn dòng nước bị ô nhiễm chảy vào sông khi mưa rửa trôi các thành phần gây ô nhiễm trên mặt.

Hàm phân bố có dạng:

$$f(Q) = \lambda e^{-\lambda Q} \quad (Q \geq 0) \qquad (3.70)$$

Phân bố mũ có ưu điểm là dễ dàng xác định tham số λ từ các số liệu quan trắc ($\lambda = \frac{\mathbf{1}}{\overline{Q}}$). Phân bố này cũng rất thích hợp với các nghiên cứu lý thuyết như mô hình xác xuất của bể chứa tuyến tính ($\lambda = \frac{\mathbf{1}}{K}$ với K là hằng số lượng trữ). Nhược điểm của phân bố này là đòi hỏi các sự kiện xảy ra phải hoàn toàn độc lập, đó là một điều khó đạt được đối với các biến thủy văn, vì vậy nhiều nhà nghiên cứu (Kawas, Delleur) đã đề xuất một quá trình Poisson phức hợp trong đó λ được coi là một biến ngẫu nhiên thay vì là hằng số trong quá trình Poisson đơn.

Giá trị Q_p cũng có thể nghịch chuẩn trực tiếp từ (3.70):

$$Q_p = \frac{\ln\left(\lambda/p\right)}{\lambda} \qquad (3.71)$$

3.2.1.4. Phân bố gamma.

Phân bố Gamma là phân bố của một tổng gồm β biến ngẫu nhiên độc lập và có phân bố mũ đồng nhất. Đồ thị hàm phân bố gamma có dạng không đối xứng, rất phù hợp để mô phỏng các đại lượng thủy văn mà không cần logarit hoá, chẳng hạn để mô phỏng phân bố xác xuất của độ sâu mưa trong một trận mưa rào.

Hàm phân bố có dạng:

$$f(Q) = \frac{\lambda^{\beta} Q^{\beta-1} e^{-\lambda Q}}{\Gamma(\beta)} (x \geq 0) \qquad (3.72)$$

Trong đó Γ(β) là hàm gamma

Γ(β) = (β-1)! nếu β là nguyên dương. Trong trường hợp chung nó được xác định bằng biểu thức:

$$\Gamma(\beta) = \int_0^{\infty} U^{\beta-1} e^{-u} du \qquad (3.73)$$

Hàm này có 2 thông số (β và λ) nên thường được gọi là phân bố gamma 2 thông số hay phân bố nhị thức và bị chặn dưới tại giá trị 0. Đây là một bất lợi khi áp dụng cho các đại lượng thủy văn vì nói chung các đại lượng này có giới hạn dưới lớn hơn 0.

Giá trị Q_p cúng có thể xác định theo công thức (3.69) thông qua một bảng tính sẵn gần đúng Q_p.

3.2.1.5. Phân bố Pearson loại III

Phân bố Pearson loại III hay còn gọi là phân bố Gamma 3 thông số, được xác định bằng cách đưa thêm thông số thứ 3 là giới hạn dưới của biến ngẫu nhiên. Sử dụng phương pháp mô men có thể xác định được 3 tham số của phân bố là λ, β và ε. Đây là một phân bố rất mềm dẻo, có thể mang các hình dáng khác nhau tuỳ theo dự biến đổi của λ, β và ε.

Hệ thống phân bố Pearson gồm 7 loại, tất cả đều là nghiệm của f(Q) trong một phương trình vi phân có dạng:

$$\frac{df(Q)}{dQ} = \frac{f(Q)(Q-d)}{C_0 + C_1 Q + C_2 Q^2} \qquad (3.74)$$

Trong đó: d là số đông của phân bố(tức là tại Q thì f(Q) lấy giá trị cực đại)

C_0 ; C_1 ; C_2 là các hằng số cần được xác định. Khi $C_2 = 0$ nghiệm của (3.74) là phân bố Pearson loại III và hàm mật độ có dạng:

$$f(Q) = \frac{\lambda^{\beta}(Q-\varepsilon)^{\beta-1} e^{-\lambda(Q-\varepsilon)}}{\Gamma(\beta)} (Q \geq \varepsilon) \qquad (3.75)$$

Khi $C_1 = C_2 = 0$ nghiệm của (3.78) sẽ là phân bố chuẩn. Như vậy phân bố chuẩn chỉ là một trường hợp riêng của phân bố Pearson III. Có thể thiết lập được mối quan hệ λ , β và ε với các đặc trưng thống kê $\overline{Q}, \sigma$ và C_s.

Foster(1924) là người đầu tiên áp dụng phân bố Pearson III trong thủy văn khi mô phỏng phân bố xác suất của đỉnh lũ lớn nhất hằng năm.

Giá trị Q_p cũng được xác định từ công thức (3.67) , trong đó ϕ_p được tra từ bảng tính sẵn do Foster và Rưpkin thiết lập phụ thuộc vào p và c_s với $\phi_p = f(p, c_s)$ (gọi là bảng Foster-Rưpkin).

Khi sử dụng đường Pearson III cần lưu ý điều kiện của c_s như sau:

$$2C_V \leq C_S \leq \frac{2C_V}{1-K_{\min}} \qquad (3.76)$$

trong đó: $K_{\min} = \frac{Q_{\min}}{\overline{\alpha}}$

3.2.1.6. Phân bố Log-Pearson loại III

Đại lượng ngẫu nhiên Q có Phân bố log-Pearson loại III nếu như logQ tuân theo phân bố Pearson III. Trong trường hợp đặc biệt khi biến ngẫu nhiên Q có phân bố đối xứng thì phân bố log-Pearson loại III được đưa về phân bố chuẩn.

Hàm mật độ có dạng:

$$f(Q)=\frac{\lambda^{\beta}(y-\varepsilon)^{\beta-1}e^{-\lambda(y-\varepsilon)}}{Q\Gamma(\beta)},(\lg Q \geq \varepsilon) \qquad (3.77)$$

Vị trí của giới hạn ε trong phân bố này phụ thuộc độ lệch của chuỗi số liệu. Nếu phân bố của mẫu lệch dương thì $\log Q \geq \varepsilon$ và ε sẽ là giới hạn dưới. Ngược lại nếu mẫu lệch âm $\log Q < \varepsilon$ và ε là giới hạn trên. Phép logarit hoá làm giảm nhỏ độ lệch của chuỗi số liệu và có thể biến đổi chuỗi số liệu vốn lệch dương thành chuỗi lệch âm. Trong trường hợp đó, khi áp dụng phân bố log-Pearson III còn phải đặt thêm một giới hạn trên nhân tạo.

Giá trị Q_p có thể suy từ giá trị $y_p = \log Q_p$, trong đó y_p được suy từ công thức (3.67) với ϕ được tra từ bảng Foster-Rưpkin. Cũng có thể thấy khi $C_s = 0$, ϕ trùng với giá trị biến chuẩn hoá Z . Còn khi $C_s \neq 0$ có thể tính ϕ bằng phương trình của Kite (1977).

$$\phi = Z + (Z^2-1)K + \frac{1}{3}(Z^3-6Z)K^2 - (Z^2-1)K^3 + ZK^4 + \frac{1}{3}K^5 \qquad (3.78)$$

3.2.1.7. Phân bố Kritski-Menkel.

Khi $C_s < 2C_v$ thì mô hình phân bố Pearson III cho giá trị âm, điều này không phù hợp với các hiện tượng thủy văn. Vì vậy Kritski và Menkel đề nghị biến đổi đường Pearson III để được đường Kritski-Menkel bằng cách đặt một biến mới $X = aQ^b$ trong đó Q tương ứng phân bố Pearson III .

Hàm mật độ có dạng:

$$f(Q)=\frac{\lambda^{\beta}}{\Gamma(\beta)}a^{\beta}Q^{\beta-1}e^{-\beta(Q/a)^{1/6}},Q \geq 0 \qquad (3.79)$$

Phân bố Kritski-Menkel cúng đặc trưng bằng 3 thông số X, C_v và C_s, trong đó $C_s = mC_v$ với m > 0.

Giá trị nhỏ nhất luôn luôn bằng 0 với mọi tỷ số C_s/C_v.

Giá trị Q_p cúng được xác định theo công thức (3.67), trong đó K_p được tra từ bảng tính sẵn do Kritski-Menkel thành lập phụ thuộc vào p và C_v:

$$Q_p = K_p.\overline{Q} \qquad (3.80)$$

3.2.1.8. Phân bố giá trị cực hạn.

Giá trị cực hạn là giá trị lớn nhất hay nhỏ nhất được lựa chọn từ chuỗi số liệu thực nghiệm, chẳng hạn chuỗi lưu lượng lớn nhất hằng năm. Fisher và Tippett (1928) đã chứng minh rằng phân bố các giá trị cực hạn được lựa chọn từ những tập mẫu của bất kỳ phân bố xác xuất nào, khi số phân tử lựa chọn đủ lớn sẽ hội tụ về một trong 3 dạng phân bố giá trị cực hạn được gọi là loại I, loại II và loại III.

Gumbel (1941) đã nghiên cứu sâu hơn phân bố giá trị cực hạn loại I và đề ra phân bố Gumbel.

Hàm mật độ phân bố Gumbel có dạng:

$$f(Q) = \frac{1}{\alpha}\exp\left[-\frac{Q-u}{\alpha} - \exp\left(-\frac{Q-u}{\alpha}\right)\right], \quad -\infty < Q < \infty \qquad (3.81)$$
$$\alpha > 0$$

u là số đông của phân bố.

Giá trị Q_p cũng được suy ra từ công thức (3.67) với ϕ_p được xác định từ biểu thức :

$$\phi_p = \frac{\sqrt{6}}{\pi}\left\{0,5772 + \ln\left[\ln\frac{T}{T-1}\right]\right\}$$
$$(3.82)$$
$$\phi_p = \frac{\sqrt{6}}{\pi}\left\{0,5772 + \ln\left[\ln\frac{1}{1-p}\right]\right\}$$

3.2.2. Mô hình ngẫu nhiên tương quan

Các mô hình ngẫu nhiên mô phỏng chuỗi các đa lượng ngẫu nhiên độc lập đã đưa ra ở phần trên chỉ phù hợp với một số quá trình thủy văn thực tế. Nói chung các quá trình thủy văn lập thành một chuỗi đại lượng ngẫu nhiên có tương quan. Mối liên hệ tương quan này có thể là giữa các số hạng trong chuỗi, tức là tương quan theo thời gian, có thể là giữa các giá trị của cùng một thời đoạn khác nhau nhưng ở vị trí khác nhau, tức là tương quan không gian. Chúng ta sẽ xem xét các chuỗi có tương quan thời gian trước, vì các mô hình này được quan tâm nhiều và được chỉnh lí đầy đủ hơn. Sau đó sẽ đề cập đến chuỗi có tương quan không gian.

3.2.2.1. Mô hình ngẫu nhiên tương quan thời gian.

Có nhiều phương pháp và nhiều mô hình thực hiện việc mô phỏng toán học chuỗi thời gian thuỷ văn có tương quan. Có thể tổng hợp thành các nhóm sau:

1.Nhóm các mô hình theo phương pháp tổng hợp, trong đó mô hình là một mẫu kép, gồm nhiều thành phần tổng hợp thành chẳng hạn như mô hình Fragment.

2.Nhóm mô hình hoá trực tiếp các giá trị các biến thủy văn bao gồm các thành phần chu kỳ và xu thế như các mô hình Markov đơn hoặc phức.

3.Mô hình các giá trị đã biến đổi của các biến thủy văn để đạt một yêu cầu nào đấy, chẳng hạn đưa về chuỗi dừng, chuỗi không có giá trị âm hay chuỗi không có tính xu thế như mô hình ARIMA.

a. Mô hình Markov.

Mô hình Markov thực chấtlà mô hình tự hồi qui tuyến tính. Cùng với sự ra đời của phương pháp Monte-Carlo mô hình Markov ngày càng được sử dụng rộng rãi để mô phỏng các quá trình thủy văn. Mô hình Markov có ưu thế ở chỗ không chỉ rõ ràng và logic,mà các sơ đồ của nó được chỉnh lí chi tiết mà còn có thể tổng hợp cho trường hợp mô hình hoá theo nhóm, khi mô hình hoá đồng thời có chuỗi thuỷ văn trên nhiều vị trí có liên hệ tương quan.

Tuy nhiên xích Markov chỉ là dạng gần đúng ban đầu để mô tả chuỗi thuỷ văn (Ratkovich,1977), tuỳ theo từng trường hợp cụ thể có các biến dạng khác nhau và cần có các giả thiết bổ xung về hàm phân bố đồng thời và phân bố có điều kiện nhiều chiều. Chẳng hạn đối với phân bố chuẩn, Kartvelixvili(1981) đưa ra giả thiết sau (giả thiết N):

Nếu hàm phân bố một chiều của các đại lượng ngẫu nhiên tương quan có dạng phân bố chuẩn và ma trận tương quan xác định dương thì phân bố nhiều chiều cũng là chuẩn.

Còn Kritxki-Menken(1979) đưa ra giả thiết sau đây đối với các đại lượng ngẫu nhiên có phân bố gamma (giả thiết G) như sau:

Phân bố nhiều chiều của số hữu hạn đại lượng ngẫu nhiên là một phân bố Gamma nếu phân bố một chiều có dạng gamma và ma trận tương quan xác định dương.

Mô hình Markov gồm có mô hình Markov đơn và mô hình Markov phức .

a-1. Mô hình Markov đơn

Mô hình Markov đơn chỉ xét tương quan của hai số hạng kề nhau đối với xích Markov đơn cần có hàm phân bố đồng thời và phân bố có điều kiện hai chiều. Mô hình Markov đơn được thực hiện bằng phương trình tự hồi qui tuyến tính ứng với các mẫu tương quan khác nhau. Theo Ratcovich, có 5 biến dạng sau đây của mô hình Markov đơn.

*Trường hợp1: Mô hình của dãy các đại lượng ngẫu nhiên độc lập. Đây là trường hợp đặc biệt đã xem xét ở phần trên.

*Trường hợp2: Tương quan chuẩn giữa các đại lượng ngẫu nhiên phân bố chuẩn. Khi đó hàm phân bố có điều kiện cũng là hàm phân bố chuẩn .

Trong mô hình này phương sai có điều kiện liên hệ với phương sai không điều kiện theo biểu thức:

$$\sigma_{i+1} = \sigma\sqrt{1 - r^2} \qquad (3.84)$$

Nghĩa là phương sai có điều kiện $ó_{i+1}$ không phụ thuộc vào số hạng đứng trước nó.

*Trường hợp 3: Mô hình các đại lượng ngẫu nhiên có phân bố Gamma nhưng với tương quan gần chuẩn. trong trường hợp này phân bố có điều kiện là phân bố

Gamma nhưng phương sai có điều kiện có cùng quan hệ như tương quan chuẩn, nghĩa là cũng có:

$$\sigma_{i+1} = \sigma\sqrt{1 - r^2} \qquad (3.85)$$

*Trường hợp 4: Mô hình các đại lượng ngẫu nhiên có phân bố gamma và giữa chúng có tương quan gamma . Phương sai có điều kiện khác với tương quan chuẩn liên hệ theo biểu thức :

$$\sigma_{i+1} = \sigma\sqrt{(1 - r^2) + 2k_i r(1 - r)} \qquad (3.86)$$

Nghĩa là phương sai có điều kiện phụ thuộc vào số hạng đứng trước K_i.

*Trường hợp 5: Tương quan giữa tần suất của các sô hạng kề nhau. Như vậy chuỗi được tạo thành là chuỗi tần suất, sau đó mới chuyển sang giá trị theo phân bố xác suất đã cho. Cấu trúc của mô hình khác hẳn các trường hợp trước, nó không phụ thuộc dạng và giá trị của hàm phân bố xác suất. Hệ số tương quan biểu thị mối liên hệ giữa các đại lượng ngẫu nhiên phân bố đều.

a-2. Mô hình Markov phức

Mô hình Marov phức là chung nhất để mô phỏng dao động của quá trình thuỷ văn. Trong trường hợp này ta cần có hàm phân bố đồng thời và phân bố có điều kiện nhiều chiều.

* Xvanhidde (1977) giới thiệu một phương pháp giải tích để giải quyết bài toán mô hình hoá, nhưng đến giai đoạn cuối lại phải thực hiện bằng phương pháp số.

* Một mô hình do Rednhicovxki (1969)đề nghị mô hình hoá trực tiếp đại lượng ngẫu nhiên dựa vào ma trận tương quan:

$$Q_i = \overline{Q} + \sum_{j=1}^{P}(Q_{i-j} - \overline{Q}_{i-j})\frac{\sigma_i}{\sigma_{i-j}}\frac{D_{i,i-j}}{D_{ii}} + \Phi_i\sigma_i\sqrt{\frac{D}{D_{ii}}} \qquad (3.87)$$

Trong đó : Q_{i-j} là giá trị của chuỗi ở thời đoạn về phía trước.

D là định thức của ma trận tương quan .

$D_{ii}, D_{i,i-j}$ là định thức con trong D tương ứng với các phần tử r_{ij} và $r_{i,i-j}$

Khi quá trình là dừng ta có quan hệ;

$$Q_i = \overline{Q} - \sum_{j=1}^{P}(Q_{i-j} - \overline{Q})\frac{D_{i,i-j}}{D_{ii}} + \Phi\sigma\sqrt{\frac{D}{D_{ii}}} \tag{3.88}$$

Mô hình này cũng dựa trên tương quan chuẩn của các đại lượng ngẫu nhiên phân bố chuẩn .

**** Mô hình Thormat-Fiering cũng có dạng tương tự :***

$$Q_{i+1} = \overline{Q}_{j+1} + a_{1,j}(Q_i - \overline{Q}_j) + a_{2,j}(Q_{i-1} - \overline{Q}_{j-1}) + t_j\sigma_{j+1}\sqrt{1-r_j^2} \tag{3.89}$$

trong đó $a_{1,j}$ và $a_{2,j}$ là các hệ số hồi qui.

$\overline{Q}_j$ và $\overline{Q}_{j-1}$ là trung bình của tháng j và j+1.

2. Mô hình theo phương pháp tổng hợp:

Mô hình theo dạng này thường gồm 2 hay nhiều mẫu (thành phần) tổng hợp thành. Điển hình của loại mô hình này là mô hình Fragment. Mô hình Fragment là một mô hình mẫu kép gồm 2 thành phần:

-Thành phần trung bình năm $\overline{Q}(t)$ (mẫu 1).

-Thành phần Fragment là giá trị trong năm q(t) có thể là tháng, tuần,... nghĩa là thể hiện qui luật phân bố trong năm (mẫu2).

1. Mô hình hoá thành phần trung bình năm $\overline{Q}(t)$: Thành phần này có thể tiến hành mô hình hoá theo mô hình Markov (đơn hoặc phức) đã trình bày ở phần trên.

2. Mô hình hoá thành phần Fragment q(t):

Thành phần Fragment nhận được bằng cách chia tất cả tung độ của thuỷ đồ trong năm cho giá trị trung bình năm tương ứng (ta được hệ số môđun K). Các Fragment có thể là liên tục hay gián đoạn, bước gián đoạn có thể là tháng, tuần hay nhỏ hơn. Khi chia cho một số không đổi thì hình dạng đường quá trình không bị phá hoại và giữ lại được tất cả các mối liên hệ vốn có của quá trình. Trong mô hình này không cần thêm một giả thiết nào.

Số Fragment có thể lấy chính bằng số đường quá trình (số năm) quan trắc. Tuy nhiên có thể phân loại Fragment theo một tiêu chuẩn nào đấy và đưa vào các "hộp đựng". Tiêu chuẩn phân loại có thể là mức độ nhiều hay ít nước theo giá trị tần suất. Chẳng hạn có 3 loại hộp đựng tương ứng với: năm nước lớn ($P<0,33$), năm nước trung bình ($0,3<P<0,66$) và năm nước nhỏ ($P>0,66$). Cũng có thể sử dụng hệ số điều tiết φ hoặc hệ số phân bố không đều trong năm d để phân loại.Tuy nhiên quan hệ giữa φ và d với tần suất p của năm chưa thật rõ ràng.

Vì vậy theo kinh nghiệm thực tế số "hộp đựng" nên chọn từ 3 đến 5. Tăng thêm số "hộp đựng"không làm tăng thêm độ chính xác khi mô hình hoá.

Bằng cách nhân giá trị trung bình năm $\overline{Q}(t)$ đã được mô hình hoá ở trên với Fragment q(t) lựa chọn ta, được chuỗi số mô hình hoá Q(t) theo các thời đoạn định trước với độ dài tuỳ ý:

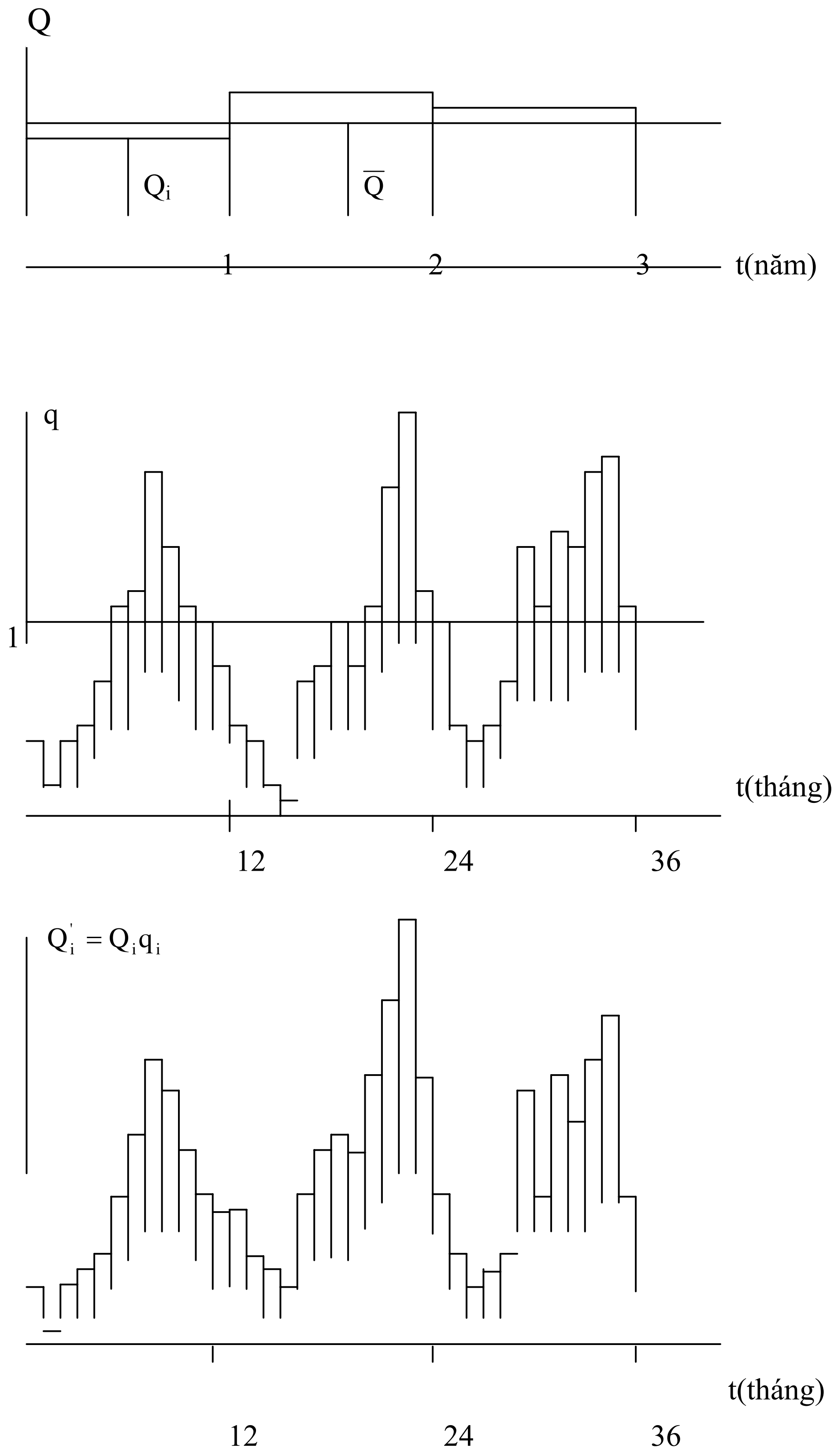
$$Q(t) = \overline{Q}(t).q(t) \tag{3.90}$$

Hình (3.1) cho thấy sơ đồ mô hình hoá theo mô hình Fragment.

Như vậy muốn mô hình hoá theo phương pháp này cần có hai dãy số ngẫu nhiên. Dãy số thứ nhất γ_1 để mô hình hoá giá trị trung bình năm theo mô hình Markov (đơn hoặc phức). Dãy số thứ hai γ_2 để lựa chọn các Fragment tương ứng theo sơ đồ rút hú hoạ quả cầu có đánh số sau đó hoàn trả lại.

Các dãy số γ_1, γ_2 được xác định từ dãy số ngẫu nhiên phân bố đều.

Mô hình Fragment cho phép nhận được thuỷ đồ đa dạng hơn mặc dù chúng ta không thoát ly khỏi quy luật phân bố thực tế của quá trình thủy văn trong quá khứ. Xác suất lặp lại nguyên vẹn một đường quá trình rất nhỏ và bằng 1/n.m trong dó m là số Fragment (hay số năm quan trắc) và n là độ dài chuỗi mô hình hoá. Thí dụ chuỗi 1000 năm được tạo ra từ 50 Fragment có xác suất lặp lại một đường quá trình là 0,0005.



Hình 3.1. Sơ đồ mô hình hoá theo Fragment.

Tuy nhiên khi độ dài chuỗi quan trắc không lớn (số Fragment không nhiều) mô hình có thể làm giảm độ chính xác trong khu vực tần suất bé và tần suất lớn. Nhưng do sự đơn giản trong thuật toán mô hình hoá nên mô hình thường hay dược sử dụng.

3. Mô hình tự hồi quy tuyến tính.

Đây là dạng thông dụng nhất để mô phỏng dao động của các chuỗi thủy văn. Nhưng khác với mô hình Markov, các hệ số hồi quy xác định theo chuẩn hội tụ, sai số còn lại được đưa vào thành phần ngẫu nhiên, không xác định theo xác suất có điều kiện. Thành phần ngẫu nhiên này khi dự báo chính là sai số dự báo của thời khoảng trước, còn khi tạo chuỗi nó là chuỗi số ngẫu nhiên sao cho đảm bảo đặc trưng thống kê (σ)không thay đổi. Mô hình thực hiện với chuỗi dừng, khi chuỗi là không dừng có thể dùng các phép biến đổi để đưa về chuỗi dừng, sau đó mới áp dụng mô hình. Khi tính toán hoặc dự báo phải đưa trở lại giá trị thực của đại lượng ban đầu.

* Mô hình có dạng tổng quát thường gọi là mô hình tự hồi quy-trung bình trượt ARIMA (Autoregresive integrated moving average).

$$Z_t = a_1 z_{t-1} + a_2 z_{t-2} + ... + a_P z_{t-P} + \varepsilon_t - b_1 \varepsilon_{t-1} - ... - b_q \varepsilon_{t-q} \quad (3.91)$$

Trong đó $a_1, a_2, .., a_P$ và $b_1, b_2, ..., b_q$ là các hệ số.

$\varepsilon_t, \varepsilon_{t-1}, ..., \varepsilon_{t-q}$ là các sai số ngẫu nhiên

Nếu dùng toán tử dịch chuyển B có thể viết (3.91) ở dạng gọn hơn:

$$\begin{aligned} &\Phi(B) Z_t = \theta(B) \varepsilon_t \\ \text{với } &\Phi(B) = 1 - a_1 B - a_2 B^2 - ... - a_P B^P \\ &\theta(B) = 1 - b_1 B - b_2 B^2 - ... - b_q B^q \end{aligned} \quad (3.93)$$

Trong đó B là quan hệ

$$BZ_t = Z_{t-1} \ , B^n Z_t = Z_{t-n} \quad (3.94)$$

* Tuỳ từng trường hợp cụ thể ta có các mô hình gọn hơn. Nếu chỉ có thành phần tự hồi quy ta có mô hình tự hồi quy AR(p)

$$Z_t = a_1 Z_{t-1} + a_2 Z_{t-2} + ... + a_P Z_{t-P} + \varepsilon_t \quad (3.95)$$

Còn khi chỉ có thành phần trung bình trượt ta có mô hình trung bình trượt MA(q):

$$Z_t = \varepsilon_t - b_1\varepsilon_{t-1} - b_2\varepsilon_{t-2} - \dots - b_q\,\varepsilon_{t-q} \qquad (3.96)$$

+ Các hệ số của mô hình được xác định dựa theo nguyên tắc tổng bình phương độ lệch giữa giá trị thực và giá trị tính theo mô hình là nhỏ nhất, và đó chính là tiêu chuẩn hội tụ. Các hệ số này thoả mãn hệ phương trình Yule_Walker và có thể xác định trực tiếp theo công thức truy hồi Durbin :

$$Q_{k+1,(k+1)} = \frac{r_{k+1} - \sum_{j=1}^{k} a_{k1(j)} r_{k+1-j}}{1 - \sum_{j=1}^{k} a_{k(j)} r_j} \qquad (3.96)$$

$$a_{k+1,j} = a_{k,j} - a_{k+1,k+1}.\ a_{k,k+1-j} \qquad (3.97)$$

+ Thành phần ngẫu nhiên được xác định để đảm bảo cho các đặc trưng thống kê không đổi :

$$\varepsilon_t = \alpha.\xi_t$$

Trong đó α là hệ số để cho các đặc trưng thống kê không đổi.

* Xvanhidde (1977) đề nghị một mô hình khác khi xét đến tính không dừng của chuỗi thủy văn. Chúng ta sẽ xem xét kỹ hơn ở phần sau(mục 3.3).

Ngoài ra còn một số mô hình khác nhưng về cấu trúc không có sự khác biệt nhiều lắm so với các loại mô hình vừa trình bày ở trên .

3.2.2.2. Mô hình ngẫu nhiên tương quan không - thời gian .

Các đại lượng thuỷ văn nói chung có liên hệ với các đại lượng tại các vị trí khác. Do vậy mô phỏng một chuỗi thủy văn có liên hệ tương hỗ với các chuỗi thủy văn khác là bài toán được quan tâm nghiên cứu, và thường gọi là mô hình hoá theo nhóm. Một số kết quả thu được nhờ các nghiên cứu của Anturin(1974), Grudinov(1958), Bansev (1972). Tuy nhiên phương pháp rất phức tạp và kết quả chưa thật mĩ mãn. Phương pháp đơn giản hơn là mô hình hoá thống kê bằng phương pháp Monte-Carlo. Phương pháp mô hình hoá nhóm như thế dựa trên một công cụ toán học như khi mô phỏng chuỗi thuỷ văn riêng biệt. Trong trường hợp này cũng phải đưa ra một giả thiết về phân bố nhiều chiều, bởi vì số hiệu thực không đủ để xây dựng nó. Nhiều nghiên cứu đã được tiến hành theo hướng này (Xvanhiđze(1964), Reznhicovski(1969), Raticovich(1975). Kritxki-Menken(1965) đã đưa ra một sơ đồ chung để mô hình hoá theo nhóm bao gồm các bước sau:

+ Xác định các thông số thống kê của các chuỗi thuỷ văn có liên hệ tương quan theo các phương pháp đã biết.

+ Thành lập ma trận tương quan giữa các chuỗi theo các số liệu quan trắc.

+ Thành lập và giải hệ phương trìnhh để xác định các hệ số hồi qui có liên hệ theo nhóm.

+ Xác định các thông số phân bố có điều kiện. Tần suất của chuỗi mô hình hoá được xác định theo phương pháp tạo số ngẫu nhiên và giá trị của chuỗi tương ứng được xác định tuỳ thuộc dạng hàm phân bố có điều kiện tính đến các mối liên hệ tương quan trong chuỗi và trong nhóm.

+ Chuỗi được mô hình hoá bằng phương pháp Monte-Carlo thoả mãn hàm phân bố có điều kiện và các thông số đã xác định.

+ Tất cả các chuỗi đã mô phỏng được chia ra thành các dãy (mẫu) với số số hạng bằng số năm quan trắc. Theo các dãy số này xác định các thông số để xây dựng hàm phân bố xác suất thoả mãn giả thiết ban đầu. So sánh các giá trị nhận được theo phân bố vừa xây dựng và thực tế để kết luận sự phù hợp của mô hình.

Sau đây trình bày hai phương pháp mô hình hoá theo nhóm. Tuy nhiên vì sự phức tạp của vấn đề nên chỉ đưa ra những nội dung cơ bản.

a. Mô hình hoá theo phương pháp Monte-Carlo.

Giả thiết rằng chuỗi thuỷ văn có phân bố Gamma. Có thể xác định giá trị của nó theo quan hệ:

$$Q_j = \overline{Q}_j + \sigma_j \Phi(c_{sj}, p_j) \qquad (3.99)$$

Trong đó j là chỉ số trạm , j=1,m

Với σ_j,ξ_j cho trước thì Q_j phụ thuộc vào tần suất p_j. Thực tế tính toán cho thấy nếu Q_j của trạm j và Q_k của trạm k có liên hệ tương quan với hệ số r_{jk} thì giữa tần suất p_j và p_k cũng có hệ số tương quan ρ_{jk} và có thể coi $r_{jk} \approx \rho_{jk}$.

Như vậy việc tạo chuỗi mô hình cho Q_j có liên hệ tương quan với các chuỗi $Q_k$khác (k=1,2,..m-1) theo ma trận tương quan $[r_{jk}]$ có thể chuyển thành bài toán tạo các chuỗi số ngẫu nhiên p_j có phân bố đều trên đoạn (0,1) và có tương quan với nhau theo ma trận tương quan $\{\rho_{jk}=r_{ik}\}$ và sau đó chuyển thành chuỗi Q_j theo (3.99). Việc mô hình hoá được tiến hành theo các bước sau:

Bước1: -Trước hết tạo dãy số ngẫu nhiên có phân bố đều trên đoan (0,1) theo các chương trình đã có trên máy tính hay các công thức gần đúng đã biết.

Bước2: -Tạo các chuỗi số ngẫu nhiên phân bố chuẩn độc lập có kỳ vọng băng 0 phương sai bằng 1 theo phương pháp Muller[].

Bước3: -Tạo các chuỗi $X'_j= (Q'_1,Q'_2,...,Q'_n)$ ngẫu nhiên phân phối đều có tương quan với nhau theo ma trận tương quan r_{jk} bằng phép biến đổi tuyến tính.

$$X'_1=a_{11}Q_1$$

$$X'_2=a_{21}Q_1+a_{22}Q_2$$

....

$$X'_n=a_{n1}Q_1+a_{n2}Q_2+...+a_{nn}Q_n. \qquad (3.100)$$

Trong đó các chuỗi $Q_1,Q_2,...,Q_n$ là các chuỗi số phân bố chuẩn dộc lập có kỳ vọng bằng 0 phương sai bằng 1 vừa xác định được theo bước 2.

+ Các hệ số a_{jk} được xác định thoả mãn hệ thức sau:

$$a_{jj} = \sqrt{r_{jj} - \sum_{k=1}^{j-1} a_{jk}^2}$$

$$a_{jk} = r_{jk} - \sum_{i=1}^{j-1} a_{jk} a_{ik} \text{ với } i > j \tag{3.101}$$

Bước 4: -Tạo chuỗi tần suất P_j=($p,p_2,...,p_n$) có tương quan theo ma trận tương quan $\{\rho_{jk}\}$ và theo công thức hàm phân bố chuẩn đối với X'_j:

$$P_j = F(X'_j) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X'_j} e^{\frac{-u^2}{2}} du \tag{3.102}$$

Trong đó X'_j là dãy xác định được ở bước 3.

+Giá trị của chuỗi mô hình hoá được xác định theo (3.99) tương ứng với chuỗi p_j vừa xác định được theo bước 4.

b. Mô hình hoá theo mô hình Thormat-Frering.

Về nguyên tắc mô hình hoá theo mô hình Thormat-Frering mở rộng có tính đến tương quan giữa các chuỗi là xác định lần lượt giá trị theo từng tháng có xét đến tương quan giữa các chuỗi. Để việc trình bày được đơn giản ta chỉ xét hai chuỗi có tương quan.

Giả sử có hai chuỗi thuỷ văn đồng bộ :

$Q_1^{(1)}, Q_2^{(1)}, ... Q_N^{(1)}$ của chuỗi (1)

$Q_2^{(2)}, Q_2^{(2)}, ... Q_N^{(2)}$ của chuỗi (2)

Tiến hành mô hình hoá theo các bước sau.

Bước1: Chuẩn hoá hai chuỗi trên thành hai chuỗi mới.

$$Y_{ij}^{(1)} = \frac{Q_i^{(1)} - \overline{Q}_j^{(1)}}{S_j^{(1)}} \quad víi \quad \overline{Q}_j^{(1)} = \frac{\sum_{i=1}^{n} Q_{ij}^{(1)}}{n}$$

$$Y_{ij}^{(2)} = \frac{Q_i^{(2)} - \overline{Q}_j^{(2)}}{S_j^{(2)}} \quad víi \quad \overline{Q}_j^{(2)} = \frac{\sum_{i=1}^{n} Q_{ij}^{(2)}}{n} \tag{3.103}$$

$\left(n = \frac{N}{12} \text{ là số năm quan trắc} \right)$

$j = 1,2,...,12$

$i = 1,2,...,N$

-Mô hình Thormat-Fering mở rộng cho tháng thứ j là :

$$\begin{aligned} Y_{ij}^{(1)} &= a_{11,j} Y_{i-1}^{(1)} + a_{12,j} Y_{i-1}^{(2)} + \varepsilon_i^{(1)} \\ Y_{ij}^{(2)} &= a_{21,j} Y_{i-1}^{(1)} + a_{22,j} Y_{i-1}^{(2)} + \varepsilon_i^{(2)} \end{aligned} \tag{3.104}$$

hoặc viết dưới dạng ma trận:

$$Y_i = AY_{i-1} + \varepsilon_i \tag{3.105}$$

hay là :

$$\begin{bmatrix} Y_i^{(1)} \\ Y_i^{(2)} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} . \begin{bmatrix} Y_{i-1}^{(1)} \\ Y_{i-1}^{(2)} \end{bmatrix} + \begin{bmatrix} \varepsilon_i^{(1)} \\ \varepsilon_i^{(2)} \end{bmatrix} \tag{3.106}$$

Bước 2: Xác định các thông số tức là các phần tử a_i của ma trận A

$$A = \frac{(B.C)}{D} \tag{3.107}$$

trong đó :

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{3.108}$$

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} b_{jj}^{(22)} & -b_{jj}^{(12)} \\ -b_{jj}^{(21)} & b_{jj}^{(11)} \end{bmatrix} \tag{3.109}$$

D là định thức của ma trận $B_{j,j+1}$:

$$B_{j,j+1} = \begin{bmatrix} b_{jj+1}^{(11)} & b_{j,j+1}^{(12)} \\ b_{j,j+1}^{(21)} & b_{jj+1}^{(22)} \end{bmatrix} \tag{3.110}$$

Các ma trận A,B,C được xác định cho từng tháng. Chỉ số (11),(12) ở dưới biểu thị thứ tự theo tháng. Còn chỉ số (11),(12) ở phía trên biểu thị sự tương ứng với các hệ số a_{11}, a_{22}...

Bước 3: Xác định các phần tử ngẫu nhiên của ma trận ε_i:

$$\varepsilon_i^{(1)} = \xi_1 \sqrt{\mathrm{var}(Y^{(1)})}$$

$$\varepsilon_i^{(2)} = \xi_1 \frac{\mathrm{cov}(Y^{(1)}Y^{(2)})}{\mathrm{var}(Y^{(2)})} + \xi_2 \sqrt{\frac{\mathrm{var}(Y^{(2)}) - [\mathrm{cov}(Y_1^{(1)}Y^{(2)})]}{\mathrm{var}(Y^{(1)})}} \qquad (3.111)$$

Trong đó : ξ_1, ξ_2 là số ngẫu nhiên chuẩn với kỳ vọng bằng 0 và phương sai bằng đơn vị.

$$\mathrm{var}\, Y^{(1)} = b_{j+1,j+1}^{(11)} - a_{11} b_{j,j+1}^{(11)} - a_{12} b_{j,j+1}^{(12)}$$

$$\mathrm{var}\, Y^{(2)} = b_{j+1,j+1}^{(22)} - a_{21} b_{j,j+1}^{(21)} - a_{22} b_{j,j+1}^{(22)}$$

$$\mathrm{cov}(Y^{(1)}Y^{(2)}) = b_{j+1,j+1}^{(12)} - a_{11} b_{j,j+1}^{(21)} - a_{12} b_{j,j+1}^{(22)} \qquad (3.112)$$

Bước 4: Lặp lại các bước từ (1) đến (3) cho 12 tháng, bắt đầu từ tháng 1 (j = 1) để tính các phần tử a_{iJ} cho từng tháng.

Bước 5: Tạo chuỗi thủy văn tháng:

$$Y_t = AY_{t-1} + \varepsilon_t.$$

Lúc đầu cho $Y^{(1)}{}_1 = Y^{(2)}{}_1 = 0$. Ta có cho năm thứ hai :

$$Y^{(1)}{}_2 = a_{11}Y^{(1)}{}_1 + a_{12}Y^{(2)}{}_1 + \varepsilon^{(1)}{}_2$$

$$Y^{(2)}{}_2 = a_{21}Y^{(1)}{}_1 + a_{22}Y^{(2)}{}_1 + \varepsilon^{(2)}{}_2 \qquad (3.113)$$

Vì $Y^{(1)}{}_1 = Y^{(2)}{}_1 = 0$ nên :

$Y^{(1)}{}_2 = \varepsilon^{(1)}{}_2$

$$Y^{(2)}{}_2 = \varepsilon^{(2)}{}_2 \qquad (3.114)$$

Tiếp tục cho số hạng năm thứ 3:

$Y^{(1)}{}_3 = a_{11}Y^{(1)}{}_2 + a_{12}Y^{(2)}{}_2 + \varepsilon^{(1)}{}_3$

$$Y^{(2)}{}_3 = a^{21}Y^{(1)}{}_2 + a_{22}Y^{(2)}{}_2 + \varepsilon^{(2)}{}_3 \qquad (3.115)$$

Lưu ý rằng Y_2 tính từ Y_1 theo (3.114) thì các hệ số của ma trận A tính theo tháng 1 và tháng 2. Còn (3.115) tính Y_3 theo Y_2 nên ma trận A tính theo tháng 2 và tháng 3. Do đó a_{11} ở (3.113) khác a_{11} ở (3.115).

Như vậy khi số trạm nhiều hơn, bài toán trở nên phức tạp hơn nhiều.

3.3 PHƯƠNG PHÁP XÁC ĐỊNH THÔNG SỐ

3.3.1. Tiêu chuẩn đánh giá mô hình

- Tiêu chuẩn chung để đánh gái mô hình là sự tương ứng phù hợp giữa chuỗi mô hình hoá và chuỗi quan trắc. Yêu cầu trước hết là các thông số trong mô hình được xác định sao cho tổng bình phương sai số là nhỏ nhất , nghĩa là:

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Q_i - Q_i^{'})^2 \rightarrow \min \qquad (3.116)$$

Trong đó : Q_i là giá trị quan trắc

$Q_i^{'}$ là giá trị tính theo mô hình.

Ngoài ra đối với mô hình ngẫu nhiên còn phải đảm bảo các đặc trưng thống kê của chuỗi mô hình hoá cùng các đặc trưng thống kê của từng thời đoạn mô phỏng phải không đổi, cùng với một số yêu cầu khác.

Một mô hình tốt cần đảm bảo các điều kiện sau:

- Quá trình tính toán phải hội tụ, nghĩa là quá trình tính sẽ dừng lại khi không có bộ giá trị nào khác của các thông số làm cho tổng bình phương sai số nhỏ hơn nữa.

- Phần dư (hay sai số dự báo) phải là ngẫu nhiên hay xấp xỉ phân bố chuẩn.

Để kiểm tra sự tương quan của phần dư e_t với các phần dư e_{t-1} trước đó ta dùng chỉ tiêu Durbin-_Watson:

$$Q = \frac{\sum (e_t - e_{t-1})^2}{\sum e_t^2} \qquad (3.117)$$

Chỉ tiêu d thường nằm giữa hai giá trị lớn nhất $d_u = 4$ và nhỏ nhất $d_L = 0$.

+Nếu $d<d_L$ hoặc $d > (4-d_L)$ thì có tương quan,

+Nếu $d_u<d<(4-d_u)$ thì không có tương quan,

+Nếu ở giữa d_L và d_u hoặc giữa $(4-d_u)$ và $(4-d_L)$ thì chưa chắc chắn cần khảo sát thêm. Cũng có thể kiểm tra các hệ số tự tương quan của phần dư. Nếu không có hệ số nào có ý nghĩa là không có tương quan.

Để kiểm tra phân bố chuẩn ta phải xây dựng hàm mật độ phân bố của phần dư e_t và so sánh với phân bố chuẩn.

- Các thông số phải đảm bảo điều kiện dừng hay thuận nghịch, nghĩa là phải thỏa mãn biểu thức:

$$a_1+a_2+...+a_P < 1 \quad \text{(Điều kiện dừng)} \qquad (3.118)$$

$$b_1+b_2+...+b_q < 1 \quad \text{(Điều kiện thuận nghịch)} \qquad (3.119)$$

Điều kiện này đảm bảo các phương trình đặc trưng sau đây có nghiệm:

$$Z^P- a_1Z^{P-1}- a_2Z^{P-2} -...-a_P = 0 \qquad (3.130)$$

$$Z^q- b_1Z^{q-1}- b_2Z^{q-2}-...-b_q = 0 \qquad (3.131)$$

- Tất cả các thông số đều phải có ý nghĩa thống kê. Muốn kiểm tra ta phải tính chỉ số thống kê của từng thông số t_a :

$$t_a = \frac{\text{Gía trị của thông số}}{\text{Độ lệch chuẩn của thông số}} = \frac{a}{S_a} \qquad (3.122)$$

Trong đó :

$$S_a = \sqrt{\frac{S_{ea}^2}{\sum Q_i^2 - \frac{(\sum Q_i)^2}{n}}} \quad (3.123)$$

$$S_{ea}^2 = \frac{\sum e_i^2}{n-m} = \frac{\sum (Q_i - \hat{Q}'_i)^2}{n-m} \quad (3.124)$$

Với: n là số giá trị dùng tính toán,

m là số thông số trong mô hình.

Các giá trị này được so sánh với giá trị chuẩn t_α, tra theo bảng Fisher ứng với mức ý nghĩa α (thường lấy α= 0,05) và số bậc tự do là n-m.

Nếu với thông số a nào đó mà $|t_a| < t_\alpha$ thì thông số đó bị loại bỏ và các tính toán trên mô hình được xác định lại trên cơ sở những thông số còn lại.

- Mô hình phải tiết kiệm nhất, hay số thông số phải ít nhất. Các thông số thừa sẽ bị loại bỏ, nghĩa là có thể chọn bậc hồi quy thấp hơn. Cơ sở để xem xét dư thừa thông số là xét ma trận tương quan giữa chúng. Khi giữa các thông số có hệ số tương quan cao ($|r|$ > 0,8-0,9) thì các thông số thừa cần loại bỏ. Trong các thông số có tương quan cao ta chỉ giữ lại những thông số mà tương quan giữa nó với yếu tố mô hình hoá là lớn nhất.

- Một mô hình tương ứng tốt còn phải đảm bảo phản ảnh đúng quy luật dao động của các quá trình thủy văn, nghĩa là đảm bảo về các đặc trưng thống kê hàm phân bố xác suất v.v.

Sự phân bố của chuỗi thuỷ văn bao giờ cũng tuân theo một qui luật xác suất nào đấy với các thông số tương ứng của nó. Do vậy chuỗi mô phỏng phải đảm bảo giá trị các thông số thống kê toàn chuỗi, cũng như các thông số thống kê của các thời đoạn mô phỏng (ví dụ theo tháng). Đồng thời cũng phải đảm bảo tính tương tự của hàm tự tương quan và mật độ phổ.

Nói chung chuỗi mô hình tương ứng và chuỗi thực đo nếu bảo đảm tiêu chuẩn tương tự thống kê :

$$\theta_i^{'} \to \theta_i \quad (i=1,2,3,...,n) \text{ khi } n \to \infty \qquad (3.125)$$

trong đó :

θ_i là giá trị đã cho của các thông số

$\theta_i^{'}$ là các giá trị xác định theo chuỗi mô hình hoá

Tuỳ theo từng bài toán cụ thể yêu cầu tương tự và số thông số yêu cầu kiểm tra cũng khác nhau.

Hàm phân bố xác suất của các đại lượng thủy văn thường có dạng Pearson III hay Log_Pearson III, và đối với phân bố dòng chảy tháng thường có dạng Jonson B. Chuỗi mô hình hoá cũng phải thoả mãn yêu cầu này.

- Đồng thời phải so sánh hàm tự tương quan và hàm mật độ phổ của chuỗi mô phỏng để thấy rõ qui luật liên hệ và dao động của các đại lượng thuỷ văn. Đối với mô hình Markov hàm tự tương quan thường có dạng :

$$R(\tau) = R^{\tau}_{(1)} \qquad (3.126)$$

Trong đó: τ là bước trượt

$R_{(1)}$ là hệ số tự tương quan của hai số hạng kề nhau. Biểu thức (3.126) có nghĩa là hàm tự tương quan giảm dần khi tăng số bước trượt τ.

- Đối với tính toán thuỷ lợi, Ratkovich đưa ra một số tiêu chuẩn khác để xem xét , đó là sự phân bố nhóm năm nhiều nhiều hay ít nước đối với tần suất p và đặc trưng chủ yếu là độ dài trung bình nhóm năm hay độ dài trung bình tương đối của nó ứng vơi tần suất p:

$$\tau_m = \frac{\overline{m}}{m_0} \qquad (3.127)$$

Trong đó $\overline{m}$ là độ dài nhóm năm tính theo chuỗi mô hình và chuỗi quan tắc.

m_o là độ dài trung bình của các đại lượng ngẫu nhiên độc lập theo lý thuyết :

$$m_0 = \frac{1}{p} \qquad (3.128)$$

Mặt khác có thể so sánh các đặc trưng thuỷ lợi giữa chuỗi quan trắc và chuỗi mô hình, đó là dung tích tương đối η_β và độ cấp nước tương đối η_α:

$$\eta_\beta = \frac{\overline{\beta}}{\beta_0} \qquad (2.129)$$

$$\eta_\alpha = \frac{\overline{\alpha}}{\alpha_0}$$

Trong đó: $\overline{\alpha_0}, \overline{\beta}$ là hệ số cấp nước và dung tích tương đối trung bình.

α_0, β_0 là hệ số cấp nước và dung tích tương đối tính với điều kiện các đại lượng ngẫu nhiên là độc lập.

Có thể xảy ra trường hợp là một số mô hình cùng cho mức tương tự như nhau của chuỗi mô hình với chuỗi quan trắc. Khi đó nên chọn mô hình nào cho tổng bình phương sai số là nhỏ nhất.

3.3.2. Phương pháp xác định thông số mô hình

Các thông số của mô hình phải được xác định sao cho thoả mãn các tiêu chuẩn đã nêu ở trên.

3.3.2.1. Xác định bậc sai phân d và các bậc hồi qui p,q:

a. Xác định bậc sai phân

Các mô hình ngẫu nhiên thường được mô phỏng đưới dạng các đại lượng ngẫu nhiên dừng, do vậy chuỗi thực tế thuỷ văn phải thông qua các phép biến đổi như đã trình bày trong ♣ 3.1 để đưa về chuỗi dừng.

-Trước hết tiến hành sai phân bậc một .

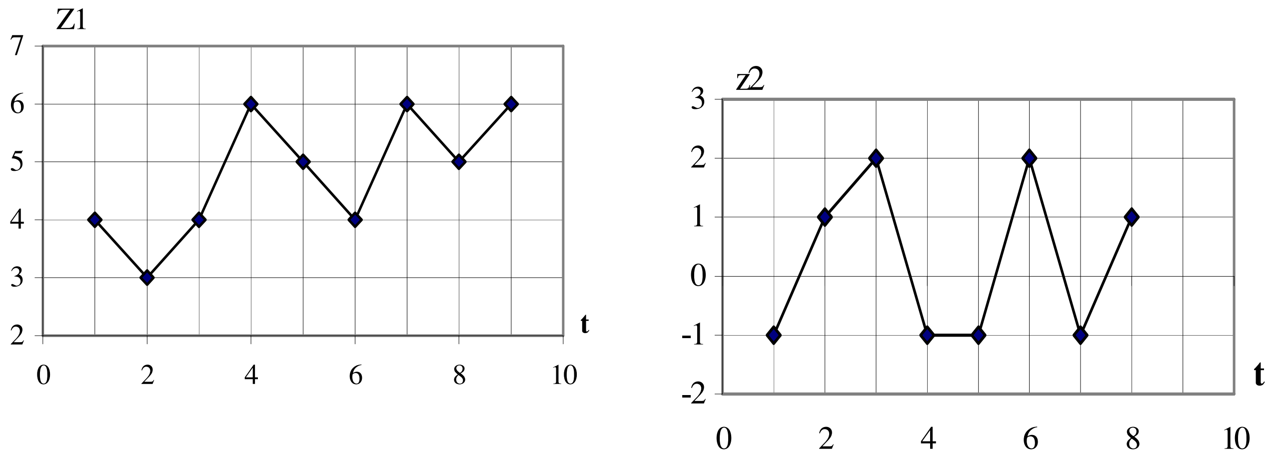
$$\Delta Q^{(1)} = Q_i - Q_{i-1} \qquad (3.130)$$

Nếu quá trình sai phân Δy vẫn còn thể hiện xu thế hoặc chu kỳ (ví dụ hình 3.2a) thì ta tiếp tục tiến hành sai phân bậc hai:

$$\Delta Q^{(2)} = \Delta Q_i - \Delta Q_{i-1} = (Q_i - Q_{i-1}) - (Q_{i-1} - Q_{i-2}) \qquad (3.131)$$

Nếu $\Delta Q^{(2)}$ cũng vẫn còn thể hiện xu thế và chu kỳ thì ta tiến hành sai phân bậc 3 và cứ tiếp tục như thế đến khi nào sai phân bậc d, ($\Delta y^{(d)}$) chỉ còn là những dao động ngẫu nhiên, có hàm TTQ và TTQR không tắt dần thì khi đó d được chọn (hình 3.2b).

Với bậc sai phân mùa P_s phương pháp làm cũng tương ứng tuy nhiên bước sai phân bây giờ là L=12.



(a): Sai phân bậc 1 (b): Sai phân bậc 2

Hình 3.2: Các dạng sai phân

b. Xác định bậc hồi quy

Bậc hồi qui p,q biểu hiện mối liên hệ giữa các đại lượng trong chuỗi thuỷ văn. Có nhiều quan điểm khác nhau khi chọn bậc hồi qui.

- Một quan điểm thường phổ biến nhất là xem xét dáng điệu của hàm tự tương quan (TTR) và tự tương quan riêng (TTQR). Kiểm tra ý nghĩa các hệ số tự tương quan theo công thức :

$$t_{r_k} = \frac{r_k}{\sqrt{\operatorname{var}(r_k)}} = \frac{r_k}{\sqrt{\frac{1}{n}}} \qquad (3.132)$$

So sánh giá trị t_{r_k} tính toán với giá trị chuẩn t_α từ phân bố Fisher ứng với mức ý nghĩa α (α=0,05). Nếu $\left|t_{r_k}\right| < t_\alpha$ thì hệ số TTQ đó không có ý nghĩa. Trong một mức độ gần đúng có thể lấy giá trị chuẩn t_α=2 với mọi bậc tự do.

- Tiếp theo xem xét dáng điệu của hai hàm này. Có mấy trường hợp điển hình như sau:

+ Hệ số tự tương quan tại tất cả các bước đều bằng 0 nghĩa là $\left|t_{r_k}\right| \prec t_\alpha$ chứng tỏ chuỗi chỉ bao hàm các thành phần ngẫu nhiên.

+ Hệ số TTQ cho các bước k=1,2,3(k<5) khá lớn và có ý nghĩa $\left(\left|t_{r_k}\right| \succ 2\right)$ sau đó giảm nhanh, chúng ta nói rằng hàm này bị ngắt tại bước k . Trong đa số các trường hợp hàm TTQ và TTQR bị ngắt tại k =2.

+ Hàm tự tương quan và TTQR tắt dần không có đỉnh.

Bậc hồi qui p,q sẽ được chọn tại nơi mà hàm TTQ và TTQR có bước nhảy đột ngột (hay bị ngắt).

Khi có thành phần mùa thì việc xác định các bậc hồi qui mùa P_s,Q_s cũng tương tự như vậy.

Tuy nhiên cũng có quan điểm khác để lựa chọn bậc hồi qui p. Theo Xvanhidze(1977), tiêu chuẩn lựa chọn là làm cực tiểu phương sai dự báo $\left[\sigma_m^{(p)}\right]^2$. p được chọn khi mà tăng thêm một bước trượt (tức là p+1) thì tỉ số $\left[\frac{\sigma_m^{(p+1)}}{\sigma_m^{(p)}}\right]$ dần đến 1. Chúng ta thừa nhận tỉ số $\frac{\sigma_m^{(p+1)}}{\sigma_m^{(p)}}$ có phân bố Fisher với (n-p-1) bậc tự do. Giá trị của phương sai $\sigma_m^{(p)}$ giảm đi khi tăng số bậc p. Tuy nhiên sự giảm này không như nhau theo các thời đoạn (tháng trong năm). Vì vậy trong mô hình sẽ lựa chọn p khi mà sau bậc đó phương sai ngẫu nhiên không giảm thực sự. Kết quả chỉ ra rằng khi tăng p thì sự trùng nhau của hàm TTq tính theo chuỗi mô phỏng và chuỗi thực đo cũng tăng lên và khi p=11 thì sự tương ứng hoàn toàn thoả mãn. Một số nghiên cứu cho rằng khi p>3 thì không làm tốt hơn mà thậm chí còn làm giảm đi sự trùng nhau giữa các thông số của chuỗi quan trắc và chuỗi mô hình hoá. Tuy nhiên các nghiên cứu khác cũng cho thấy khi p=3 thì sự trùng hợp chỉ tương ứng với 3 bước đầu tiên, còn ở các bước sau nó ảnh hưởng thực sự đến hàm TTq của trung bình thángvà đặc biệt quan trọng là trung bình năm. Và đối với chuỗi dòng chảy đưa đến sự giảm nhỏ dung tích hồ chứa. Vì vậy theo Xvanhidze nên lấy p=11 cho mọi chuỗi dòng chảy.

Riêng với dòng chảy năm, nhiều kết qủa nghiên cứu(Ratkovich, Xvanhidze, Kritski_Menken) cho thấy chỉ nên chọn bậc hồi qui là 1, và khi ấy ta có mô hình là Markov đơn. Tuy nhiên cũng có ý kiến (Drujnhin,1968) cho rằng hàm TTQ có ý nghĩa ở các bước xa hơn, phụ thuộc vào chu kỳ của dòng chảy năm và mô hình mô phỏng là xích Markov phức.

3.3.2.2. Xác định các thông số tự hồi qui a_i và trung bình trượt b_j.

Có mô hình chỉ có dạng tự hồi qui AR(p) hay chỉ có dạng trung bình trượt cũng có mô hình dạng hỗn hợp . Tuy về nguyên tắc việc xác định chúng không thực sự khác nhau, nhưng về chi tiết cũng nhiều cách sử lí riêng. Do vậy ở đây sẽ trình cho từng dạng riêng biệt.

a.- Xác định các thông số a_i:

Các thông số a_i thoả mãn hệ phương trình Yule_Walker:

$$C_K = a_1 C_{k-1} + a_2 C_{k-2} + ... + a_P C_{k-P} \qquad (3.133)$$

$$k = 1,2,...,p$$

Ứng với các giá trị k ta có hệ p phương trình. Đây là hệ phương trình tuyến tính hệ số hằng số, giải ra được nghiệm là các hệ số a_i. Có thể trực tiếp tính các a_i theo công thức truy hồi Durbin, suy ra từ hệ (3.133):

$$a_{k+1,(k+1)} = \frac{r_{k+1} - \sum_{j=1}^{k-1} r_{kj} r_{k-j+1}}{1 - \sum_{j=1}^{k-1} a_{kj} r_j} \qquad (3.134)$$

$$a_{k+1,(j)} = a_{kj} - a_{k+1,(k+1)} - a_{k,(k-j+1)} \qquad (3.135)$$

với $a_{1,(1)} = r_1$

Trong đó chỉ số trước, không có dấu ngoặc chỉ bậc hồi quy của mô hình, còn chỉ số sau có dấu ngoặc đơn chỉ thứ tự các hệ số a_i chẳng hạn $a_{2,(1)}$ là hệ số thứ nhất a_1 của mô hình tự hồi quy bậc 2, còn $a_{2,(2)}$ là hệ số thứ hai a_2 của mô hình này.

Cho k thay đổi thừ 1 đến p, tính lần lượt từ $a_{2,(1)}$ đến hệ số khác $a_{k,(k)}$

Ví dụ với AR(2) ta có p = 2

Khi k = 0 thì $a_{1,(1)} = r_1$

Khi k = 1 thay vào (3.134) được:

$$a_{2,(2)} = \frac{r_2 - a_{1,(1)} r_1}{1 - a_{1,(1)} r_1} = \frac{r_2 - r_1 r_1}{1 - r_1 r_1} = \frac{r_2 - r_1^2}{1 - r_1^2} \qquad (3.136)$$

$a_{2,(2)}$ chính là hệ số a_2 của mô hình AR(2)

Cho j = 1, thay vào (3.135) có :

$$a_{2,(1)} = a_{1,(1)} - a_{2,(2)} - a_{1,(1)} = r_1 \frac{r_2 - r_1^2}{1 - r_1^2} r_1 = r_1 \frac{1 - r_2}{1 - r_1^2} \qquad (3.137)$$

và $a_{2,(1)}$ chính là hệ số a_1 của mô hình AR(2).

Với các a_i khác của mô hình AR(p) cũng giải quyết tương tự.

b. Xác định các thông số b_i của MA(q)

Các thông số b_i thoả mãn một hệ phương trình tương tự như hệ Yule_Walker, được suy ra từ quan hệ:

$$r_k = \frac{\gamma_k}{\gamma_0} = \frac{-b_k + b_1 b_{k-1} + ... + b_q b_{k-q}}{1 - b_1^2 + b_2^2 - ... + b_q^2} \qquad (3.138)$$

Cho k = 1,2,...,q ta được một hệ phương trình phi tuyến, việc giải gặp nhiều khó khăn khi q > 2.

Ví dụ đối với MA(2) ta có q = 2

Khi k = 1, thay vào (3.138) được:

$$r_1 = \frac{-b_1 + b_1 b_2}{1 + b_1^2 + b_2^2} \tag{3.139}$$

Khi k = 2, thay vào (3.138) được :

$$r_2 = \frac{-b_2}{1 + b_1^2 + b_2^2} \tag{3.140}$$

Giải kết hợp (3.139) và (3.140) ta được các hệ số b_1,b_2. Khi bậc q > 2 thì hệ phương trình trở nên phức tạp hơn nhiều.

c.-Xác định các thông số a_I, b_i của ARMA(p,q)

Việc xác định đồng thời các hệ số a_i và b_i trong mô hình ARIMA hay ARMA có một số bước khác so với trường hợp đơn lẻ AR(p) hay MA(q)

*Để giải quyết độc lập các giá trị a_i theo công thức truy hồi Durbin hay hệ phương trình Yule_Walker thì hệ thức (3.133) được viết bắt đầu từ k > q tức là khi ấy các $b_i = 0$. Do đó ta có hệ :

$$\begin{aligned} C_{q+1} &= a_1 C_q + a_2 C_{q-2} + \ldots + a_p C_{q+1-p} \\ C_{q+2} &= a_1 C_{q+1} + a_2 C_q + \ldots + a_p C_{p+2-p} \\ &\qquad \ldots\ldots\ldots\ldots\ \ldots\ldots\ldots\ldots\ \ldots\ldots \\ C_{q+p} &= a_1 C_{q+p-1} + \ldots + a_p C_q \end{aligned} \tag{3.141}$$

Hệ phương trình (3.141) cũng là hệ tuyến tính bậc nhất, hệ số hằng số và có thể giải ra tìm các nghiệm a_1 , a_2 , a_p.

*Để tìm các hệ số b_i ta cũng xuất phát từ quan hệ (3.138), nhưng khác với mô hình MA(q), ở đây γ_k và γ_0 (hay C_k và C_0) không phải là của Z_t mà là của thành phần ngẫu nhiên ε_t . Nghĩa là $\gamma_k=\gamma_{k\varepsilon}$, $C_k=C_{k\varepsilon}$. Quan hệ giữa $\gamma_{k\varepsilon}$ và γ_{kz} hay $C_{k\varepsilon}$ và C_{kz} có dạng :

$$C_{k\varepsilon} = \sum_{i=0}^{p} a_i^2 C_{kz} + \sum_{i=1}^{p} \sum_{h=0}^{p-i} a_h a_i d_k \tag{3.142}$$

với $d_k = C_{k+i} + C_{k-i}$ (1.143)

C_{kz} có thể xác định theo các công thức ở phần đầu chương. Đó chính là covarian bậc k của z. Sau khi có C_{kz} tính được $C_{k\varepsilon}$ theo (3.142) và từ đó thay vào (3.138) được một hệ phương trình, giải hệ này ta được các hệ số b_i. Sau đây đưa ra một ví dụ minh hoạ cho một chuỗi dòng chảy có các đặc trưng sau khi chuẩn hoá là :

$$\overline{Z} = 0{,}0001\ ; \qquad \sigma^2 = 0{,}9804$$

*Tính các giá trị $C_{kz}r_{kz}$ và σ_Z^2. Sau khi phân tích hàm TTQ và TTQR , chọn các bậc hồi qui là p=1 và q=1, nghĩa là ta có mô hình ARMA(1,1).

*Để tìm a_i thay vào hệ (3.141) ta được :

$$C_2 = a_1.\ C_1 + 0,\ \text{do đó}\ a_1 = \frac{C_2}{C_1} = 0{,}8656$$

*Để tìm b_1 thay k lần lượt vào (3.142)

+Vớik=0 có: $C_{0\varepsilon} = C_{0Z} + a_1^2 C_{0Z} - 2a_1 C_{1Z} = C_{0Z}(1 + a_1^2) - 2a_1 C_{1Z}$ (3.145)

+Với k = 1 có: $C_{1\varepsilon} = a_0^2 C_{1Z} + a_1^2 C_{1Z} + a_0 a_1 \left[C_{(1+1)Z} + C_{(1-1)Z}\right]$ (3.146)

Vì a_0 = -1 nên $C_{1\varepsilon} = C_{1Z} + a_1^2 C_{1Z} - a_1(C_{2Z} + C_{0Z})$ (3.147)

+Với q=1, thay vào(3.138) được:

$$\gamma_{1\varepsilon} = C_{1\varepsilon} = (-b_1 + 0)\sigma_\varepsilon^2 \ \text{ hay }\ \sigma_\varepsilon^2 = \frac{-C_{1\varepsilon}}{b_1} \qquad (3.148)$$

$$\gamma_{0\varepsilon} = C_{0\varepsilon} = (1 + b_1^2)\sigma_\varepsilon^2 \ \text{ hay }\ \sigma_\varepsilon^2 = \frac{C_{0\varepsilon}}{1 + b_1^2} \qquad (3.149)$$

Cân bằng (3.148) và (3.149) ta có :

$$\frac{C_{0\varepsilon}}{1 + b_1^2} = \frac{-C_{1\varepsilon}}{b_1} \qquad (3.150)$$

Ta tính được $C_{0\varepsilon}$ = 0,8074 và $C_{1\varepsilon}$ = 0,3244, do đó thay vào (3.150) được phương trình :

$$0{,}8074b_1 - 0{,}3244(1+b_1^2) = 0.$$

Giải phương trình bậc hai này được hai nghiệm.

$$b_{11}=0{,}5037 \text{ và } b_{12}=1{,}985.$$

Do điều kiện hội tụ nên phải có $|b_1|<1$, do đó loại bỏ giá trị b_{12}. Cuối cùng ta có các hệ số a_1=0,8656 và b_1=0,5037 và mô hình có dạng :

$$Z_t = 0{,}8656Z_{t-1} + \varepsilon_t - 0{,}5037\varepsilon_{t-1} \qquad (3.151)$$

Lưu ý rằng các hệ số a_i,b_i phải thoả mãn điều kiện hội tụ và thuận nghịch, nghĩa là có: $\sum_{i=1}^{p} a_i < 1$ và $\sum_{i=1}^{q} b_i < 1$.

d. Xác định các thông số a_i, b_i bằng tối ưu hoá.

Về nguyên tắc ta có thể xác định được các hệ số a_i, b_i theo các bước trình bày ở trên. Nhưng khi bậc p,q>2 thì bài toán trở nên phức tạp. Do vậy trong thực tế hiện nay khi tính trên máy tính người ta thường dùng thuật toán tối ưu giống như ở mô hình tất định. Các phương pháp tối ưu hoá đã được trình bày trong chương 2(mục ←2.3). Hiện nay trong việc tìm các thông số của mô hình ngẫu nhiên thường dùng các thuật toán sau:

- Phương pháp đơn hình.
- Phương pháp Hooke Jeever.
- Phương pháp Quasi-Newton.
- Phương pháp Rosenbrock.

Mỗi phương pháp có một ưu điểm riêng khi giải quyết bài toán. Trong một số chương trình mẫu hiện nay (ví dụ Statistica) người ta sử dụng phương pháp Quasi-Newton. Phương pháp này xét đến cả đạo hàm bậc 1 và bậc 2, nghĩa là xét đến cả tốc độ biến thiên theo hướng dò tìm mục tiêu. Vì vậy phương pháp này mang lại kết quả tốt hơnvà cho tốc độ hội tụ nhanh hơn so với các phương pháp hiện có. Để sử dụng các thuật toán tối ưu hoá phải cho giá trị ban đầu $a_i^{(0)}, b_i^{(0)}$, sau đó thông qua tối ưu hoá để tìm bộ thông số tối ưu. Tuy nhiên để các bước tính giảm bớt và cũng đảm bảo điều kiện ổn định, thường chọn giá trị ban đầu cho các hệ số a_i, b_i là 0,1 nghĩa là có a_i=0,1 và b_i=0,1.

Giá trị ban đầu cho các hằng số trong mô hình (nếu có) là μ và δ được cho bởi quan hệ:

$$\mu^{(0)} = \overline{Z} \quad v\mu \quad \sigma^{(0)} = \mu^{(0)} - \left(1 - \sum_{i=1}^{p} a_i^{(0)}\right) \qquad (3.152)$$

Giá trị ban đầu được đưa vào chương trình. Một quá trình tính lặp được thực hiện để đạt được hàm mục tiêu. Muốn việc tính toán được kết thúc cần phải đưa ra tiêu chuẩn hội tụ. Tiêu chuẩn hội tụ cho biết độ chính xác của thông số tương ứng với tổng bình phương sai số. Phép lặp dò tìm thông số tối ưu số dừng lại khi mà sự biên đổi của các thông số ước lượng qua hai bước lặp liên tiếp nhỏ hơn giá trị này. Thường thì tiêu chuẩn này được ngầm định trong các phần mềm ứng dụng (ví dụ, với Statistica nó là 0,0001). Tuy nhiên cũng có thể được ấn định bởi người sử dụng bằng cách đưa vào từ bàn phím.

Có thể xảy ra trường hợp quá trình ước lượng không hội tụ (penanty). Có một số nguyên nhân, trong đó có một nguyên nhân chính là các thông số ban đầu của mô hình có sai số lớn và có tương quan lớn. Đó thường là mô hình chưa chuẩn. Khi đó cần xem xét lại hàm TTQ và TTQR để tìm ra một mô hình đơn giản hơn với các thông số nhỏ hơn. Hoặc nên định lại giá trị ban đầu.

Các thành phần mùa được sử lý đồng thời và các thông số sai phân mùa cũng được xác định đồng thời trong quá trình tối ưu hoá.

Hiện nay một phần mềm thống kê khá thông dụng của hãng Microsoft là Statistica, trong đó có phần mềm ARIMA. Sử dụng phần mềm đó ta có thể xác định đồng thời các thông số a_i, b_i của thành phần mùa và thành phần không mùa, các kết quả kiểm tra đánh giá sai số và sự phù hợp của mô hình cũng như kết quả dự báo được đưa ra ngay dưới dạng bảng và đồ thị. Trong Statistica cũng có phần mềm "ước lượng phi tuyến” (nonlinear estimasion) cho ta một loạt các phương pháp ước lượng thông số của một mô hình bất kỳ, bao gồm cả phương pháp đơn hình, Hooke Jeever, Quasi Newton, Rosenbrock .v.v.). Theo phần mềm này ta cũng có thể nhanh chóng thay đổi và lựa chọn các bậc hồi qui, sai phân p,d,q và P_s,D_s,Q_s.

3.3.2.3. Xác định hành phần ngẫu nhiên ε_t.

Thành phần ngẫu nhiên ε_t thực chất là phần dư hay sai số tính toán :

$$\varepsilon_t = Q_t - \overline{Q_t} \tag{3.153}$$

Tuy nhiên để mô hình mô phỏng đúng qui luật dao động của các quá trình thuỷ văn, về nguyên tắc, thành phần ngẫu nhiên được xác định sao cho các đặc trưng thống kê của chuỗi thuỷ văn không thay đổi. Trong các bài toán dự báo vì ε_t chưa biết nên trong mô hình coi ε_t=0. Còn trong các bài toán tính toán và tạo chuỗi mô hình hoá ,thành phần ε_t được xác định tuỳ theo mô hình và phương pháp xử lý.

-Trong mô hình tự hồi qui AR(p) ta có quan hệ :

$$\sigma_\varepsilon^2 = \sigma_z^2\left(1 - a_1 r_1 - a_2 r_2 - \ldots - a_p r_p\right) \tag{3.154}$$

$\varepsilon_i = \alpha\xi_i$ với α=σ_ε thoả mãn (3.155), ta được:

$$\alpha = \sigma_\varepsilon = \frac{\sigma_z}{\sqrt{1 - a_1 r_1 - a_2 r_2 - \ldots - a_p r_p}} \tag{3.155}$$

Trong đó ξ_i là dãy số ngẫu nhiên phân bố chuẩn có $\bar{\xi}$ =0 và Dξ =1.

- Trong mô hình MA(q) thành phần ngẫu nhiên cũng được xác định theo quan hệ

$$\varepsilon_i=\alpha\xi_i \quad \text{với} \quad \alpha = \sigma_\varepsilon = \frac{\sigma_z}{\sqrt{1 + b_1^2 + b_2^2 + \ldots + b_q^2}} \tag{3.156}$$

ξ_i cũng là dãy số ngẫu nhiên như trên.

*Trong mô hình ARIMA(p,q) cũng có quan hệ tương tự:

$\varepsilon_i=\alpha\xi_i$ với $\alpha=\sigma_\varepsilon$.

Tuy nhiên σ_ε ở đây xác định phức tạp hơn. Ta có :

$$\gamma_{0\varepsilon} = C_{0\varepsilon} = \left(1 + b_1^2 + b_2^2 + ... + b_q^2\right)\delta_\varepsilon^2 \qquad (3.157)$$

Do đó : $$\sigma_\varepsilon^2 = \frac{C_{0\varepsilon}}{1 + b_1^2 + b_2^2 + ... + b_q^2} \qquad (3.158)$$

Theo (3.148) thì $C_{0\varepsilon} = C_{0z} + a_1^2 C_{0z}$.

*Cũng có thể xác định theo công thức truy hồi Durbin để có phương sai dư δ_ε :

$$\sigma_{\varepsilon_t}^2 = \sigma_{\varepsilon_k}^2 = \sigma_{\varepsilon_{k-1}}^2 \left[1 - \left(a_k\right)^2\right] \qquad (3.159)$$

Trong đó: $\sigma_{\varepsilon_{k-1}}^2, \sigma_{\varepsilon_k}^2$ là phương sai được tính khi lấy bậc hồi qui là k-1 và k (k=1,2,3,...,p).

a_k là hệ số thứ k của mô hình hồi qui bậc k.

Như vậy có thể thấy rằng thành phần ngẫu nhiên ε_t luôn luôn gắn với dãy số ngẫu nhiên ξ_t có phân bố chuẩn với $\bar{\xi}$ =0 và $\sigma\xi$ =1. Dãy số ngẫu nhiên ξ_t này lại được suy ra từ dãy ngẫu nhiên có phân bố đều γ_t trên đoạn [0,1].

3.3.3. Phương pháp tạo chuỗi mô hình hoá

Một vấn đề quan trong của mô hình hoá ngẫu nhiên là tạo ra một chuỗi số có độ dài đủ đáp ứng yêu cầu tính toán phục vụ, nhưng vẫn bảo đảm các thông số thống kê và hàm phân bố xác suất không đổi. Phương pháp tạo chuỗi thường dùng nhất hiện nay là phương pháp Monte-Carlo.

3.3.3.1. Phương pháp Monte-Carlo.

a. Khái niệm phương pháp Monter- Carlo

Phương pháp Monte-Carlo hay gọi là phương pháp "phép thử thống kê", là phương pháp mô hình hoá các đại lượng ngẫu nhiên tuân theo một qui luật xác suất với các thông số thống kê định trước.

Phương pháp Monte-Carlo ra đời từ bài báo "The Monte-Carlo method" của Metropolis N. và Ulam S.(1949). Người tạo ra phương pháp này chính là các nhà toán học Mỹ Phon Nêiman và Ulam S. Cơ sở của phương pháp đã có từ lâu trong lí thuyết xác suất, tuy nhiên nó chỉ phát trển rộng rãi khi xuất hiện máy tính điện tử (Computer). Tên gọi Monte-Carlo xuất phát từ thành phố Mote-Carlo là nơi có các con quay, một dụng cụ để nhận được số ngẫu nhiên.

Sơ đồ chung của phương pháp Mote-Carlo như sau:

Cho rằng chúng ta cần xác định một số chưa biết nào đấy m. Ta hãy lấy một đại lượng ngẫu nhiên ξ sao cho $M\xi=m$, và giả thiết rằng $D\xi=b^2$.

Xem xét N đại lượng ngẫu nhiên $\xi_1,\xi_2,...\xi_N$ có phân bố trùng với phân bố của ξ. Nếu N đủ lớn thì theo định lý giới hạn trung tâm phân bố của tổng :

$\zeta_N = \xi_1+\xi_2+...+\xi_N$ sẽ có phân bố gần chuẩn với kì vọng a=N.m và phương sai $\sigma^2=N.b^2$. Từ nguyên tắc "3 xigma" ta có :

$$P\{Nm - 3b\sqrt{N} < \zeta_N < Nm + 3b\sqrt{N}\} = 0{,}997 \qquad (3.160)$$

Theo bất đẳng thức trong dấu {} cho N, nhậnh được bất đẳng thức tương đương và có cùng xác suất :

$$P\left\{m - \frac{3b}{\sqrt{N}} < \frac{\zeta_N}{N} < m + \frac{3b}{\sqrt{N}}\right\} \approx 0{,}997 \qquad (3.161)$$

Hoặc viết gọn lại:

$$P\left\{\left|\frac{1}{N}\sum_{j=1}^{N}\xi_j - m\right| \prec \frac{3b}{\sqrt{N}}\right\} \approx 0{,}977 \qquad (3.162)$$

Đây là một quan hệ đặc biệt quan trọng của phương pháp Monte-Carlo. Nó cho phép tính m và đánh giá sai số. Quả vậy, chúng ta tìm N giá trị của đại lượng ngẫu nhiên ξ (hay tìm N đại lượng ξ cũng vậy vì chúng có cùng phân bố xác suất nên trùng nhau). Từ (3.162) thấy rằng giá trị trung bình số học gần bằng m. Với một xác suất lớn, sai số không vượt quá giá trị $\frac{3b}{\sqrt{N}}$. Dĩ nhiên sai số sẽ dần đến 0 khi tăng N.

Phương pháp Monte-Carlo tỏ ra đơn giản và có tính tổng quát. Nhược điểm chính của nó là sự hội tụ chậm chạp và sự hội tụ của nó là sự hội tụ theo xác suất. Tuy nhiên đây chưa hẳn là nhược điểm vì trong thực tế nhiều khi phương pháp xác suất lại tỏ ra hợp lý.

Phương pháp Monte-Carlo được đưa vào ứng dụng trong thuỷ văn từ những năm 1960. Bằng cách tạo ra một dãy các số hạng có phân bố đã cho, cùng với xích Markov, phương pháp này cho phép mô phỏng dao động của các quá trình thuỷ văn, tạo nên nhiều thể hiện mà chuỗi dòng chảy tự nhiên không có được, đồng thời vẫn giữ nguyên các đặc trưng thống kê cơ bản của chuỗi dòng chảy thực. Như vậy phương pháp Monte-Carlo không làm tăng độ chính xác các thông số ban đầu mà chỉ cung cấp cho ta những phiên bản mới để sử dụng cho các bài toán tính toán thuỷ văn và thuỷ lợi khác. Chuỗi mô hình hoá theo phương pháp này cũng không phải là các giá trị "dự báo".

b. Mô hình hoá theo phương pháp Monter-Carlo

Muốn mô hình hoá theo Monte-Carlo cũng phải xuất phát từ dãy số ngẫu nhiên có phân bố đều , sau đó mới chuyển thành dãy có phân bố bất kỳ.

3.3.3.2. Phương pháp tạo chuỗi có phân bố đều

Đại lượng ngẫu nhiên γ có phân bố đều trên đoạn [a,b] là đại lượng ngẫu nhiên có mật độ phân bố như sau:

$$P(\gamma) = \begin{cases} \dfrac{1}{b-a} & khi \quad a < \gamma < b \\ 0 & khi \;\; \gamma > b \, v\mu \, \gamma < a \end{cases} \qquad (3.163)$$

Để tạo ra chuỗi số phân bố đều, và với các bài toán thuỷ văn thường lấy trên đoạn (0,1), thường có 3 phương pháp sau:

a. Bảng số ngẫu nhiên.

Bảng số ngẫu nhiên có từ trước khi phương pháp Monte=Carlo ra đời. Tippet(1954) là người đầu tiên lập ra bảng số này. Kendan và Babinton-Smith đã tạo ra một con quay đánh số để có được dãy số ngẫu nhiên và lập thành bảng. Hiện nay đã có bảng số ngẫu nhiên với 1000000 chữ số(Hãng REND,1955). Bảng số ngẫu nhiên có thể nạp trực tiếp vào máy tính. Trong khi tính toán ta sẽ lấy các số này lần lượt theo một nguyên tắc nào đấy , chẳng hạn lấy theo hàng ngang, hoặc hàng dọc, lấy từ dưới lên hoặc lấy từ trên xuống, lấy cách 2 hay 3 chữ số. Ví dụ lấy hai chữ số đầu là 86 và coi nó là tần suất $\gamma_1=p_1=0{,}86$. Hai chữ số sau sẽ là 51 và $\gamma_2 = p_2 = 0{,}51$. Tương tự có $\gamma_3=0{,}59$, $\gamma_4=0{,}07$ ... Tuy nhiên bảng số ngẫu nhiên chiếm một phần không nhỏ của bộ nhớ máy tính, vì vậy nó ít được dùng.

b. Phương pháp truyền số ngẫu nhiên.

Số ngẫu nhiên được tạo ra nhờ một thiết bị đặc biệt, giống như bánh xe điện, hoạt động theo nguyên tắc vật lý, ở mỗi thời điểm của máy tính, thiết bị truyền một số ngẫu nhiên vào một ô trong bộ nhớ củamáy. Tuy nhiên thiết bị không có khả năng lặp lại các phép tính từ đầu, do đó không thể kiểm tra lại và điều quan trọng hơn là khả năng trục trặc của thiết bị gây ra sự sai lệch không kiểm tra được. Vì vậy phương pháp này cũng ít dùng.

c. Phương pháp tạo số giả ngẫu nhiên.

Đó là phương pháp thông dụng nhất hiện nay. Bằng một thuật toán nào đó tạo ra được dãy các giá trị có phân bố đều , các số này được gọi là số giả ngẫu nhiên.

Thuật toán tạo số giả ngẫu nhiên do FonNeiman đề suất. Sau đó một loạt các phương pháp khác ra đời. Đại đa số các phương pháp đều xuất phát từ hệ thức hồi qui, chẳng hạn nhờ hệ thức:

$$X_{i+1}=(a\ \ X_i-c)\ \ (\mathrm{mod}\ P) \qquad (3.164)$$

$0\le a<P$ với P là số nguyên dương; a,c là các số cho trước, và X_{it1} là số dư của phép chia $(aX_i+c)/m$.

Như vậy ưu thế của số giả ngẫu nhiên là hoàn toàn rõ ràng. Trước hết là nhờ ở thuật toán đơn giản, do vậy tốc độ tạo số ngẫu nhiên có cùng tốc độ của máy tính. Chương trình chỉ chiếm một bộ nhớ rất nhỏ. Các số ngẫu nhiên có thể tái tạo lại bất kỳ khi nào. Và cuối cùng dãy số chỉ cần một lần kiểm tra là có thể sử dụng cho nhiều lần. Nhược điểm cơ bản của phương pháp này là sự “giới hạn” trữ lượng số giả ngẫu nhiên. Tuy nhiên có nhiều phương pháp cho phép nhận được số lớn các dãy số như thế, chẳng hạn thay đổi số hạt giống ban đầu γ_o.

Trong các máy tính hiện nay ở bất kỳ ngôn ngữ nào cũng có thể gọi chương trình để thu được dãy số ngẫu nhiên có phân bố đều.

3.3.3.3. Tạo chuỗi số có phân bố bất kỳ.

Từ chuỗi số ngẫu nhiên có phân bố đều có thể xác định được dãy số có phân bố bất kỳ.

Thuật toán để tìm một số ngẫu nhiên ξ có phân bố f(x) từ số ngẫu nhiên có phân bố đều γ là giải phương trình:

$$F_{\xi}(x) = \int_{-\infty}^{\xi} f(x)dx = \gamma \qquad (3.165)$$

hay $$\xi = F_{\xi}^{-1}(x) \qquad (3.166)$$

Về nguyên tắc phương trình (3.165) có thể giải được nếu f(x) là hàm liên tục và có dạng giải tích tường minh. Tuy nhiên việc tìm nguyên hàm của tích phân (3.165) không phải lúc nào cũng làm được, ngay cả đối với phân bố chuẩn. Nếu hàm $F_{\xi}(x)$ là đủ trơn thì có thể giải phương trình (3.165) với một tập hợp nào đó của γ. Lập bảng giá trị hàm ngược $F_{\xi}^{-1}(x)$ tức là dãy số rời rạc của ξ. Sau đó dùng phép nội suy tính ra các giá trị ξ tương ứng với các giá trị γ.

Nói riêng với trường hợp phân bố chuẩn ta phải giải phương trình:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\xi} e^{\frac{-x^2}{2}} dx = \gamma \qquad (3.167)$$

Tích phân (3.167) không có nguyên hàm trong dạng hiện, do đó phải giải gần đúng bằng phương pháp số. Chẳng hạn Muller đề ra phương pháp sau:

$$\xi_i = \sqrt{-2\log\gamma_1}\cos 2\pi\gamma_2 \qquad (3.168)$$

$$\xi_{i+1} = \sqrt{-2\log\gamma_1}\sin 2\pi\gamma_2 \qquad (3.169)$$

Dãy số ξ_i , ξ_{i+1} sẽ có phân bố chuẩn với $M\xi = 0$ và $D\xi = 1$

Hoặc có thể dùng công thức gần đúng (3.65) của AbranoWits và Stegun (1975) để xác định $F_\xi(x)$ theo tập hợp các giá trị ξ (bảng hàm phân bố chuẩn). Sau đó dùng phép tra ngược tìm ξ khi biết $F_\xi(x) = \gamma$, ta sẽ được chuỗi phân bố chuẩn ξ cần tìm.

Dĩ nhiên số ngẫu nhiên $\alpha = a + \sigma\xi$ (3.170)

sẽ có $M\alpha = a$ và $D\alpha = \sigma^2$

Với các phân bố khác ta cũng xuất phát từ dãy số phân bố đều và giải(3.165) để được dãy số mới có phân bố đã cho.

Sau khi có dãy số với phân bố đã cho (nói riêng là phân bố chuẩn) đưa vào thành phần ngẫu nhiên trong các mô hình đã nói ở trên. Với một tập hợp tuỳ ý của dãy số nhẫu nhiên ξ ta có thể tạo ra một chuỗi dòng chảy có độ dài tuỳ ý và đảm bảo đặc trưng thống kê không thay đổi nhờ các mô hình nói trên.

3.4. MỘT SỐ MÔ HÌNH NGẪU NHIÊN THÔNG DỤNG HIỆN NAY.

Hiện nay có nhiều mô hình ngẫu nhiên đang được áp dụng ở trên thế giới, trong số đó có một mô hình đang được áp dụng có kết quả tại Việt Nam.

3.4.1. Mô hình tự hồi quy trung bình trượt ARIMA (AUTOREGRESIVE INTERGRATED MOVING AVERAGE MODEL)

Mô hình ARIMA hay còn gọi là mô hình Box_Jenkin do Box và Jenkin đề xuất năm 1970. Mô hình áp dụng cho chuỗi dừng, có dạng tổng quát:

$$Z_t = \mu + a_1 Z_{t-1} + a_2 Z_{t-2} + \ldots + a_p Z_{t-p} + \varepsilon_t - b_1\varepsilon_{t-1} - b_2\varepsilon_{t-2} - \ldots - b_q Z_{t-q} \qquad (3.171)$$

Trong đó Z_t , Z_{t-1} ,..., là các giá trị của chuỗi dừng

ε_t , ε_{t-1} ,..., là các sai số

μ , a_1 , a_2 ,..., b_1 , b_2 ,... là các thông số của mô hình.

Để xây dựng mô hình ARIMA còn cần thực hiện các bước sau:

• Nhận dạng: Trong bước này chuỗi được biến đổi thành chuỗi dừng. Trên cơ sở phân tích hàm tự tương quan (TTQ) và tự tương quan riêng (TTQR) đưa ra một mô hình thử nghiệm ban đầu.

•Ước lượng thông số: Trong bước này tính toán các giá trị ban đầu của các thông số. Sau đó bằng chương trình tối ưu hoá xác định được đồng thời các thông số tự hồi qui và trung bình trượt của thành phần mùa và không mùa.

•Kiểm định mô hình: Sau khi có mô hình, cần kiểm tra sự phù hợp của mô hình với số liệu thực đo. Việc kiểm tra được tiến hành bằng cách phân tích sai số (phần dư), ý nghĩa các thông số mô hình. Nếu một kết quả nào đó không chấp nhận được thì mô hình phải được sửa chữa và lặp lại các bước tính toán trước đây.

•Dự báo và tạo chuỗi: Sau khi mô hình để được kiểm định có thể tiến hành dự báo và tạo chuỗi. Các giá trị dự báo phải được kèm theo với giới hạn tin cậy.

Sau đây chúng ta sẽ xem xét chi tiết các bước ở trên. Trong quá trình giải quyết phải tách ra thành phần có tính mùa và thành phần không mang tính mùa.

3.4.1.1.Nhận dạng.

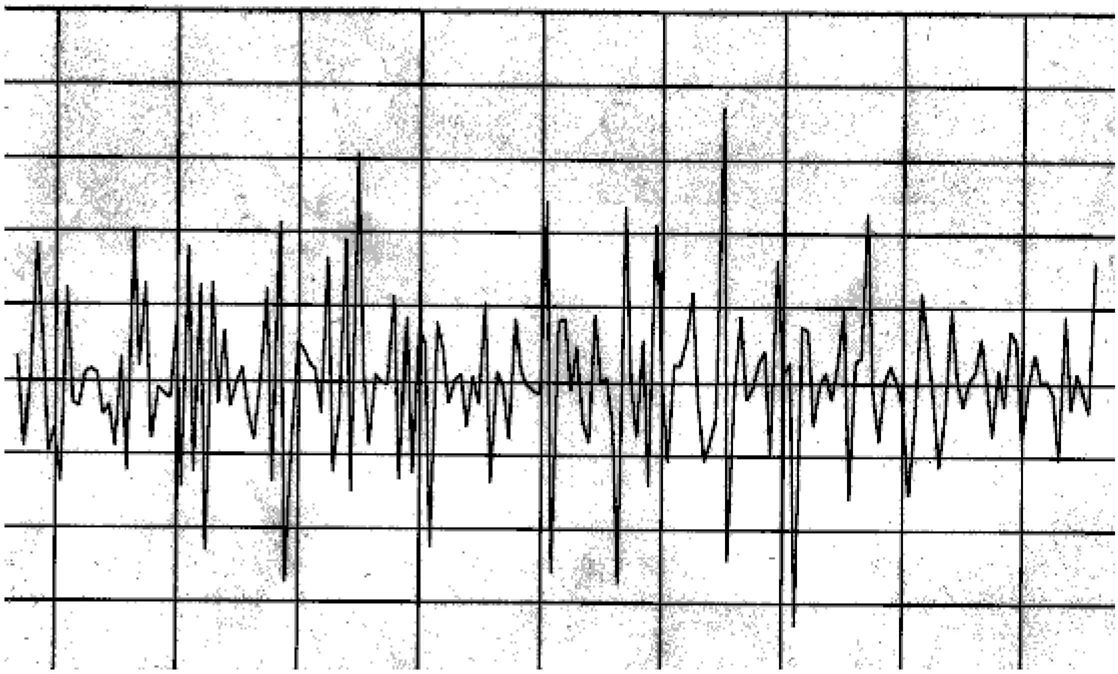
Phân tích hàm TTQ và hàm TTQR.

Trước khi định dạng mô hình phải đưa chuỗi về dạng chuỗi dừng. Việc biến đổi có thể thực hiện thông qua lôgarit hoá, sai phân bậc 1, bậc 2,... như đã trình bày trong ←3.2. Chỉ khi nào quá trình chuỗi biến đổi dao động xung quanh giá trị trung bình, có dạng hình răng cưa thì khi ấy mới coi là đúng. Cũng có thể thông qua hàm TTQ và TTQR. Chuỗi là dừng khi hàm TTQ và TTQR ở tất cả các bước bằng 0 hoặc chỉ khác không ở một vài bước. Còn ở các bước mùa (L, 2L, 3L) chúng bị ngắt hoặc tắt dần. Thực hiện với chuỗi dòng chảy trạm Kontum sông Sesan, chuỗi chỉ dừng sau bước logarit hoá và sai phân bậc 1.

Sau khi có chuỗi dừng ta tính lại hàm TTQ và TTQR của chuỗi dừng này theo các công thức (3.22), (3.23) và (3.27),(3.28). Đồng thời dùng chỉ tiêu thống kê t (3.30) để kiểm tra. Như đã trình bày ở trên, ta lấy t_α=2,0 làm giá trị tiêu chuẩn so sánh. Nếu $|t| < t_\alpha$ thì hệ số TTQ và TTQR đó không có ý nghĩa.

Với thành phần mùa, giá trị tiêu chuẩn t_α=1,25. Khi nào ở các bước mùa L, 2L, 3L,... có $|t| < t_\alpha$ =1,25 thì cũng coi là không có ý nghĩa.

Hàm TTQ và TTQR có một số dạng điển hình như phân tích trong mục ←3.2. Lưu ý rằng dáng điệu của thành phần mùa bị che lấp bởi thành phần không mùa.



Hình 3.3: Chuỗi dòng chảy trạm Kontum sau khi biến đổi(chụp lvsv).

b. Các dạng mô hình .

Tương ứng với dáng điệu hàm TTQ và TTQR của các thành phần mùa và không mùa ta có thể có các dạng sau:

- Khi hàm TTQ tắt dần, còn hàm TTQR bị ngắt, có đỉnh ở bước k=1 và không có tương quan ở bước khác thì chỉ có một thông số tự hồi qui (p=1), tức là có mô hình AR(1).

- Khi hàm TTQ tắt dần, còn hàm TTQR có đỉnh ở bước 2 và không có tương quan ở các bước khác thì có hai thông số tự hồi qui (p=2), tức là có mô hình AR(2).

• Khi hàm TTQ có đỉnh ở bước 1, không có tương quan ở các bước khác, còn hàm TTQR tắt dần thì có 1 thông số trung bình trượt (q=1), tức là có mô hình MA(1).

• Khi hàm TTQ có đỉnh ở bước 2, không có tương quan ở các bước khác, còn hàm TTQR có dạng tắt dần thì có hai thông số trung bình trượt (q=2), tức là có mô hình MA(2).

• Khi hàm TTQ tắt dần từ bước 1, và hàm TTQR cũng tắt dần từ bước 1 thì có một thông số tự hồi qui (p=1) và một thông số trung bình trượt (q=1), tức là có mô hình ARMA(1,1). Nếu phép biến đổi đưa về chuỗi dừng là sai phân bậc 1 (d=1) ta có mô hình ARIMA(1,1,1). Khi phép biến đổi là sai phân bậc 2(d=2) ta có mô hình ARIMA(1,2,1).

Các trường hợp khác ít xảy ra. Tuy nhiên cũng có khi cả hai hàm TTQ và TTQR đều bị ngắt như đã phân tích ở ←3.2. Khi ấy phải chọn dạng hàm nào bị ngắt đột ngột hơn thì coi là bị ngắt, còn hàm kia là tắt dần, từ đó có các mô hình tương ứng.

Khi phân tích hàm TTQ và TTQR phải xem xét ở cả các bước mùa. Có thể tính và vẽ riêng các hàm này hoặc cũng có thể phân tích chung trên một hàm TTQ, nhưng lưu ý riêng tại các bước mùa 1L, 2L, 3L,...Khi xét thành phần mùa ta có mô hình dạng ARIMA(p,d,q,P_S,D_S,Q_S) trong đó P_S,D_S,Q_S tương ứng với thành phần mùa. Có thể đưa ra bảng tổng kết sau đây về các trường hợp mô hình (bảng 3.1).

Có thể thấy dáng điệu của hàm TTQ và TTQR như dẫn ra trong hình 3.4.

Việc xác định đúng đắn số bậc hồi quy p,q là khó khăn. Mặc dù hầu hết các mô hình hỗn hợp có p = 1 và q= 1, nhưng cũng có khi mô hình bậc cao hơn (nghĩa là p= 2-3 hoặc q= 2-3) thích hợp hơn. Khi đó cần thông qua bước kiểm định mô hình để chọn bậc p,q tốt nhất.

3.4.1.2. Ước lượng thông số mô hình.

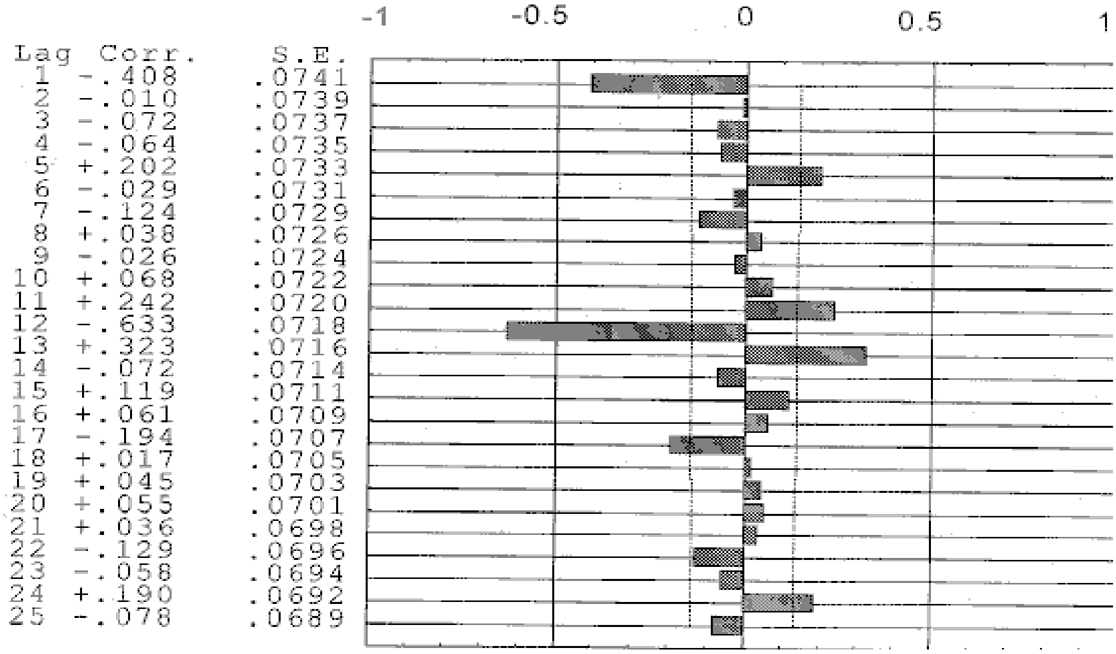
Trong bước này ta xác định đồng thời các thông số mô hình thành phần không mùa a_i , b_i và thành phần mùa a_{i2} , b_{i2}.

Ta có thể sử dụng hệ phương trình Yule_Walker và công thức Durbin để tìm các hệ số này. Tuy nhiên việc xác định đồng thời chúng chiếm nhiều thời gian.

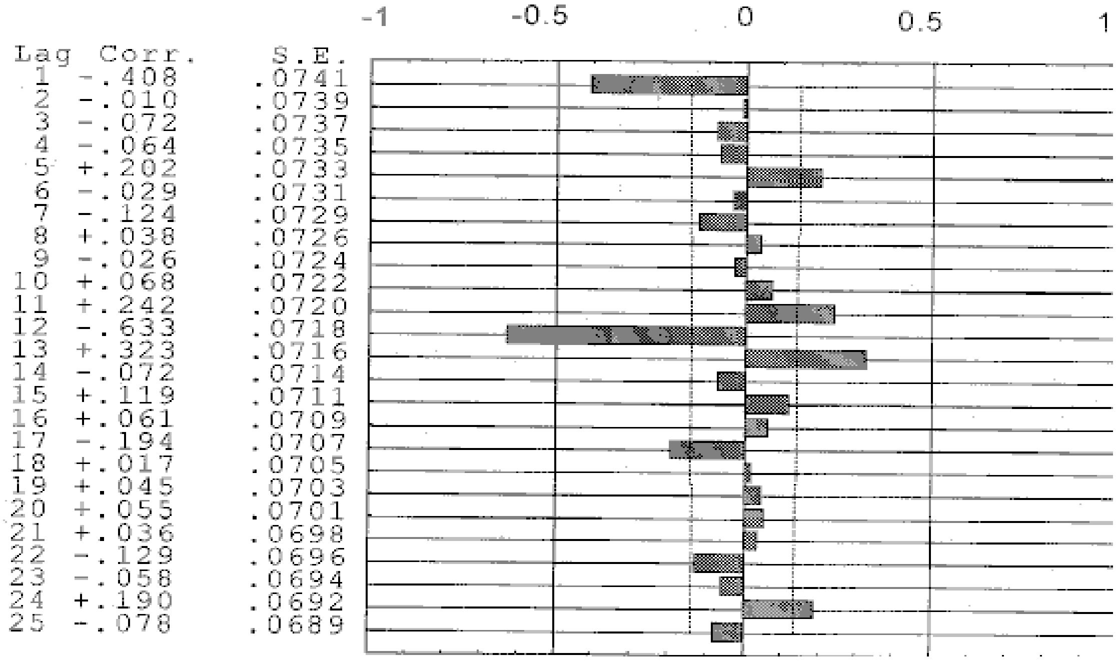
Nguyên tắc xác định thông số vẫn là bình phương tối thiểu, tức là đảm bảo cho tổng bình phương sai số(hàm mục tiêu) là nhỏ nhất. Khi ấy các thông số xác định được theo hệ phương trình Yule_Walker cũng chỉ là thông số thô ban đầu, cần thông qua phương pháp tối ưu hoá để được bộ thông số tốt nhất.

Bảng 3.1: Các đặc trưng và dạng cơ bản của mô hình ARIMA (p,d,q, P_S,D_S,Q_S)

Hàm TTQ	Hàm TTQR	Mô hình
Bị cắt ở bước 1 hoặc 2,không có đỉnh mùa có ý nghĩa	Tắt dần	Trung bình trượt không mùa(q=1 hoặc 2) $Z_t = \mu - \Theta_1\varepsilon_{t-1} + \varepsilon_t$
Bị cắt sau bước mùa L,không có đỉnh ở bước không mùa	Tắt dần	Trung bình trượt mùa(Q_S=1) $Z_t = \mu - \Theta_{1L}\varepsilon_{t-L} + \varepsilon_t$
Tắt dần	Bị cắt sau bước 1 hoặc 2 không có đỉnh ở mức mùa	Tự hồi quy không mùa(p=1) $Z_t = \delta + \Phi_1 Z_{t-1} + \varepsilon_t$
Tắt dần	Bị cắt sau bước mùa L,không có đỉnh ở mức không mùa	Tự hồi quy mùa(P_S=1) $Z_t = \delta + \Phi_{1L} Z_{t-L} + \varepsilon_t$
Tắt dần	Tắt dần	Hỗn hợp(ARIMA) +Không mùa: $Z_t = \delta + \Phi_1 Z_{t-1} + \Theta_1\varepsilon_{t-1} + \varepsilon_t$ + Mùa: $Z_t = \delta + \Phi_{1L} Z_{t-L} - \Theta_{1L}\varepsilon_{t-L} + \varepsilon_t$



Hình 3.4a :Dáng điệu hàm tự tương quan trạm Kontum s. Se san



Hình 3.4b :Dáng điệu hàm tự tương quan riêng trạm Kontum s. Se san

Các thông số của thành phần mùa cũng như không mùa phải đảm bảo điều kiện khả nghịch và dừng, nghĩa là phải thoả mãn các quan hệ (3.118) và (3.119)

$$a_1 + a_2 + ... + a_p < 1$$

và $$b_1 + b_2 + ... + b_q < 1$$

Giá trị hằng số cho thành phần mô hình AR(p) có thể chọn $\overline{m} = \overline{Z}$, tức là giá trị trưng bình của chuỗi dừng. Còn hằng số cho mô hình hỗn hợp ARMA(p,q) có thể là

$\sigma = \mu^2(d - a_1 - a_2 - a_p)$. Các hằng số cho thành phần mùa cũng tương tự .

Để giải quyết nhanh chóng bài toán thường dùng phương pháp tối ưu hoá. Phương pháp này cho phép xác định đồng thời các thông số thành phần mùa, thành phần không mùa và đảm bảo cực tiểu hàm mục tiêu. Trong phần mềm ARIMA của Statistica thường dùng phương pháp Quasi Newtơn. Chương trình tự động giải quyết các bước tính lặp để cho kết quả cuối cùng. Như vậy các giá trị ban đầu không nhất thiết phải tính toán như trên mà chỉ cần đảm bảo yêu cầu khả nghịch và dừng. Do đó có thể lấy ngay giá trị ban đầu là $a_1 = a_2 = ... = a_p = 0{,}1$ và $b_1 = b_2 = ... = b_q = 0{,}1$.

Sau một số bước phép lặp sẽ hội tụ và chương trình dừng lại thông báo kết quả giá trị các thông số tối ưu cùng giá trị hàm mục tiêu (tổng bình phương sai số) tương ứng. Tiêu chuẩn hội tụ thường được ngầm định (trong Statistica lấy bằng 0,0001). Cũng có thể ấn định bằng bàn phím do nguồn sử dụng yêu cầu.

Trường hợp Penanty, tức là trường hợp các giá trị ban đầu không đưa đến hội tụ thì có thể sử dụng các công thức sau để ước lượng giá trị ban đầu. Ví dụ đối với mô hình AR(2):

$$a_1^0 = \frac{r_1(1 - r_2)}{1 - r_1^2} \qquad (3.171)$$

$$a_2^0 = \frac{r_2 - r_1^2}{1 - r_1^2} \qquad (3.173)$$

$$\text{và } r_1 = \frac{-b_1}{1 + b_1^2} \qquad (3.174)$$

Ví dụ bằng phần mềm ARIMA trong Statistica xác định được bộ thông số đối với dòng chảy tháng của trạm. Trung nghĩa sông Sesan như sau:

$a_1 = p(1) = 0$; $b_1 = q(1) = 0,61236$

$a_2 = p_S(1) = 0$; $b_2 = q_S(1) = 0,76417$

Tương ứng với nó là mô hình:

$$Z_t = \varepsilon_t - 0,61236 - 0,76147\varepsilon_{t-L} \qquad (3.175)$$

3.4.1.3. Kiểm định mô hình.

Đây là giai đoạn đánh giá chất lượng mô hình, đánh giá sự phù hợp của mô hình với chuỗi thủy văn thực tế. Việc kiểm định có thể thông qua các đặc trưng sau:

a. Phân tích sai số

Phân tích sai số là bước quan trọng khi đành giá sự tương ứng của mô hình Box _Jenkin với tài liệu thực. Sai số tình toán trong mô hình ARIMA phải có phân bố chuẩn và độc lập đối với nhau (ngẫu nhiên).

- Để kiểm tra tính độc lập của sai số ta xem xét hàm TTQ của sai số. Nếu sai số độc lập thì giữa chúng không có tương quan, hay nói cách khác các hệ số tương quan giữa chúng xấp xỉ bằng không, nằm trong phạm vi nhỏ hơn sai số tính các hệ số tương quan. Nếu không đạt tiêu chuẩn này thì giữa các sai số có mối liên hệ, chứng tỏ mô hình không tương ứng với chuỗi số liệu. Khi đó hàm TTQ và TTQR của sai số cho ta định hướng để điều chỉnh mô hình.

Ví dụ từ sự phân tích hàm TTQ và TTQR ta có mô hình trung bình trượt của một chuỗi số liệu trạm C là :

$Z_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1}$. Nhưng khi phân tích các hàm TTQ và TTQR của sai số cho thấy không phải tất cả chúng đều bằng không, nghĩa là sai số không phải ngẫu nhiên. Như vậy mô hình đã chọn là không thực sự phù hợp (bảng 3.2).

Phân tích chi tiết hơn cho thấy hàm TTQ và TTQR của chuỗi số đều tắt dần, do đó chọn mô hình hỗn hợp cho kết quả tốt hơn**.**

Bảng 3,2, Hàm TTQ và TTQR của sai số trạm C

Bước trễ	Hàm TTQ		Hàm TTQR	
	r_k	t_k	r_{kk}	t_{kk}
1	-0,246	-2,255	-0,246	-2,255
2	-0,134	-1,160	-0,207	-1,897
3	0,015	0,128	-0,084	-0,769
4	-0,293	-2,496	-0,381	-3,492
5	-0,073	-0,580	-0,380	-3,483
6	0,259	2,051	-0,077	-0,706
7	-0,142	-1,072	-0,320	-2,932
8	0,178	1,326	-0,144	-1,320
9	0,111	0,810	-0,066	-0,650
10	-0,225	-1,629	-0,211	-1,934
11	0,104	0,730	-0,035	-0,321
12	-0,043	-0,300	-0,087	-0,797

- Cũng có thể kiểm tra ý nghĩa theo nhóm các giá trị riêng biệt của hàm tự tương quan theo chỉ tiêu Box_Pierce, hay còn gọi là chỉ tiêu Lijung_Box:

$$Q^* = n(n+2)\sum_{K=1}^{K} \frac{r_k^2}{n-K} \qquad (3.176)$$

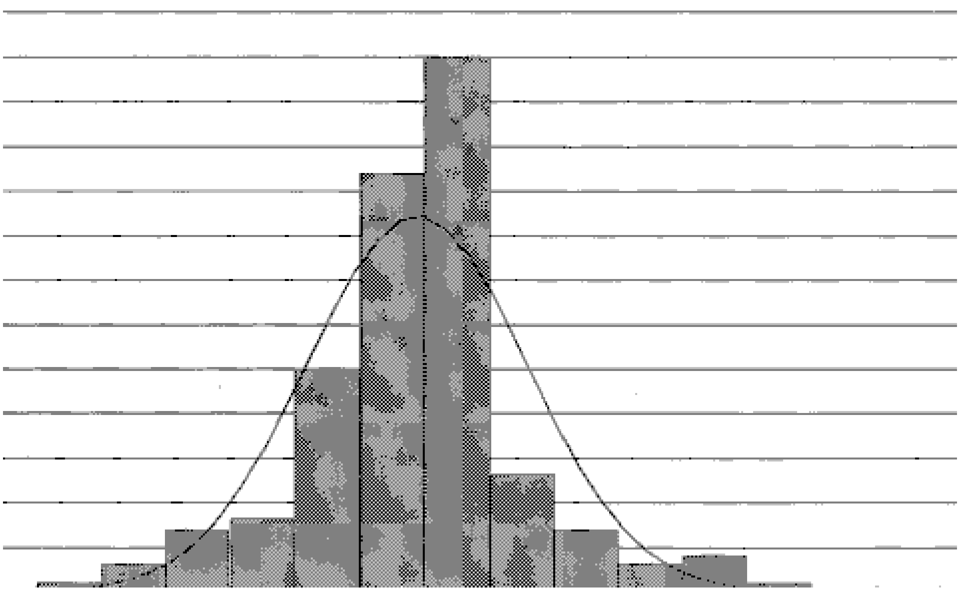
Trong đó: r_k là hệ số TTQ của sai số tại các bước k (k = 1,2,...,k), n là số giá trị của chuỗi dừng để tính toán hay số số hạng của chuỗi gốc. k là số giá trị TTQ trong tổng.

Phân bố của chỉ tiêu Lijung_Box có dạng xấp xỉ phân bố χ^2 với số bậc tự do k-p- q cho thành phần không mùa và k- p- q- P_S- Q_S cho thành phần mùa. Nếu giá trị Q^* tính được nhỏ hơn giá trị tra trong bảng thì có thể coi các giá trị TTQ của nhóm đó bằng 0 và sai số là ngẫu nhiên.

Trong hầu hết các chương trình đều tính Q^* cho các bước L=12,24,36 và 48. Giá trị nhỏ nhất của k được chọn đủ lớn để cho p-q có một hiệu quả đáng kể trong bậc tự do. Thường k=12 hoặc 24 được coi là thoả mãn.

Ví dụ: Trong mô hình trạm C kể trên, chỉ tiêu Box-Pierce tính được là 35,86. Với số bậc tự do là m=11 và mức ý nghĩa α=0,04 tra bảng χ^2 được giá trị tiêu chuẩn Q_α=19,6751. Như vậy $Q^*>Q_\alpha$, chứng tỏ sai số là không ngẫu nhiên và mô hình không tương ứng với tài liệu thực đo, cần phải hiệu chỉnh lại.

- Để kiểm tra tính chuẩn của phân bố sai số ta phải xây dựng hàm phân bố của sai số và so sánh với phân bố chuẩn lí thuyết (hình 3.6).



Hình 3.6: So sánh phân bố của sai số và phân bố chuẩn lý thuyết trạm Kontum s. Sê san

Cũng có thể kiểm tra bằng so sánh độ lệch xác suất. Đồ thị độ lệch xác suất được xây dựng như sau:

+ Các giá trị của sai số được xắp xếp theo thứ tự.

+ Tính giá trị độ lệch chuẩn ϕ của phân bố chuẩn.

+ Vẽ quan hệ giữa ϕ và sai số ε. Nếu sai số có phân bố chuẩn thì các điểm tập trung quanh một đường thẳng đứng.

b. Kiểm định các thông số.

Không phải mọi thông số đều có ý nghĩa thống kê. Khi đó ta phải tiến hành kiểm định ý nghĩa thống kê của các thông số vừa ước lượng. Chỉ tiêu thống kê có dạng (3.122).

$$t_{ai} = \frac{a_i}{S_{ai}}$$

Trong mô hình ta chỉ giữ lại những thông số có ý nghĩa thống kê, nghĩa là các thông số mà giá trị tuyệt đối của nó nhỏ hơn giá trị tiêu chuẩn tra trong bảng Fisher $|t| < t_\alpha$ (α=0,05). Có thể chọn t_α=2 cho mọi bậc tự do. Các số hạng không có ý nghĩa thống kê sẽ bị loại trừ và các bước tính toán được tính lại theo số các số hạng được giữ lại.

Bảng (3.3) cho thấy các kết quả tính toán xác định các thông số của trạm Kontum sông Senanê(Thực hiện bằng phần mềm ARIMA của Statistica). Các thông số ước lượng đều có ý nghĩa thống kê

c. Xác định thông số thừa.

Mô hình ARIMA đưoc chọn là mô hình có số thông số ít nhất. Trong quá trình phân tích đã loại đi nhiều trường hợp. Tuy vậy vẫn có thể có 2,3 mô hình cùng phù hợp với số liệu thực tế . Nhưng ta chỉ lựa chọn mô hình với các thông số ít hơn. Để xem xét khả năng có thông số thừa hay không ta xác định ma trận tương quan của các thông số. Ma trận tương quan giữa các thông số được xác định từ ma trận Hessian. Trong phần mềm Statistica đã cho ngay kết quả thực hiện tính ma trận này. Nếu giữa các thông số có tương quan cao (r>0,8-0,9) thì có khả năng dư thừa thông số. Khi ấy một mô hình bậc thấp hơn (tức là có số thông số ít hơn) cho khả năng phù hợp tốt với số liệu thực.

Bảng 3.3: Kết quả tính toán thông số của trạm Kontum sông Senan.

Mô hình ARIMA(0,1,1)(0,1,1) Trạm Kontum Sông Sê san Các bước biến đổi: Ln(x),D(1),D(12)			
Số quan trắc	Tổng bình phương sai số ban đầu	Tổng bình phương sai số cuối cùng	Sai số quân phương
179	48,309	20,527	0,11663
	q(1)	Q(1)	
Giá tri thông số	0,62558	0,75977	
Sai số	0,07346	0,04077	

d. Lựa chọn mô hình tốt nhất.

Mô hình tốt nhất là mô hình có số thông số ít nhất và cho độ lệch quân phương kiểm tra nhỏ nhất. Độ lệch quân phương được xác định theo công thức (3.16). So sánh các mô hình sau đây cho trạm Kontum ta thấy :

+ Mô hình ARIMA(1,1,1)(1,1,1) cho S_1=0,16539

+ Mô hình ARIMA(0,1,1)(0,1,1) cho S_2=0,11663

Như vậy mô hình ARIMA (0,1,1)(0,1,1) có số thông số ít và độ lêch quân phương kiểm tra là nhỏ nhất. Vì vậy mô hình này được chọn để dự báo.

3.4.1.4. Dự báo

a. Dự báo.

Mô hình sau khi đã lựa chọn xong sẽ được dùng để dự báo. Khi viết mô hình để dự báo phải đưa về dạng chuỗi gốc. Ví dụ với mô hình tự hồi qui bậc 2 của sai phân bậc 1 ta có:

$$\overline{Z_t} = a_0 + a_1 Z_{t-1} + a_2 Z_{t-2} \tag{3.177}$$

ở đây: $\overline{Z}_t = Q_t - Q_{t-1}$; $Z_{t-1}=Q_{t-1}-Q_{t-2}$; $Z_{t-2}=Q_{t-2}-Q_{t-3}$

Thay các biểu thức này vào (3.177) được:

$Q_t-Q_{t-1}=a_0+a_1(Q_{t-1}-Q_{t-2})+a_2(Q_{t-2}-Q_{t-3})$.

Sau khi biến đổi ta được biểu thức cuối cùng để dự báo:

$$Q_t=a_0+(1+a_1)Q_{t-1}+(a_2-a_1)Q_{t-2}-a_2Q_{t-3} \tag{3.178}$$

Bắt đầu từ một thời điểm nào đó, tiến hành dự báo cho thời kì sau (các tháng sau chẳng hạn). Theo số liệu quan trắc ta đã có các giá trị Q_{t-1},Q_{t-2},Q_{t-3}. Thay các giá trị này vào (3.178) ta xác định được Q_t cần dự báo.

Tiếp tục coi giá trị Q_t vừa tính được là Q_{t-1} và coi Q_{t-1} là Q_{t-2} .v.v.. Thay vào (3.178) xác định được Q_t mới, và tiếp tục như vậycho đến hết các thời khoảng cần dự báo. Dĩ nhiên các giá trị dự báo cho các thời đoạn sau xa hơn sẽ gặp sai số nhiều hơn, vì nó dựa vào các dự báo trước đó. Sai số chung gặp phải là:

Việc dự báo cho các thời kì xa hơn (t+L) có thể tiến hành nhưng phải đảm bảo rằng các thông số của chuỗi không thay đổi theo thời gian và các thông số không cần tính lại. Nếu khi dự báo thấy sai nhiều thì phải thiết lập một mô hình mới.

b. Xác định khoảng tin cậy.

Khoảng tin cậy được xác định với biên 95%(tức là tương ứng với mức ý nghĩa α=0,05).

$$Q_t - t_{\frac{\alpha}{2}}.\frac{S}{\sqrt{n}} < Q_t < Q + t_{\frac{\alpha}{2}}.\frac{S}{\sqrt{n}} \tag{3.179}$$

Trong đó: $t_{\alpha/2}$ là giá trị hệ số tin cậy xác định được tương ứng với α. Thường chọn $t_{\alpha/2}$=1,96 (khi đã biết phương sai σ) hoặc $t_{\alpha/2}$=2,365 (khi chưa biết σ),

n là dung lương mẫu dự báo.

Lưu ý rằng độ lệch quân phương dự báo kiểm tra càng nhỏ thì phạm vi khoảng tin cậy càng nhỏ và trong hầu hết các trường hợp khoảng tin cậy nhỏ cho ta một mô hình chính xác hơn.

Có thể vẽ đồng thời giá trị dự báo, giá trị thực và sai số dự báo lên cùng một đồ thị để phân tích so sánh dao động của chúng, nếu có chỗ nào đó không tương ứng thì ta cần tìm nguyên nhân để hiệu chỉnh hoặc định lại mô hình.

3.4.1.6. Tạo chuỗi dòng chảy.

Việc tạo chuỗi mô hình hoá được thực hiện theo dạng mô hình đã lựa chọn ở trên. Tuy nhiên thành phần ngẫu nhiên ε_t được xác định từ đại lượng ngẫu nhiên ξ_t có phân bố chuẩn với trung bình $M\xi= 0$ và phương sai $D\xi=1$ sao cho các đặc trưng thống kê của chuỗi số mô hình hoá không đổi. ξ_t có thể xác định theo công thức(3.165).

Đưa thành phần ε_t xác định được vào mô hình đã chọn sẽ tính được Z_t. Tiếp tục như vậy sẽ tạo nên được chuỗi Z_t với độ dài tuỳ ý. Từ chuỗi Z_t ta chuyển ngược lại để được các đại lượng gốc.

Ví dụ với chuỗi dòng chảy trạm Trung nghĩa ta có mô hình :

$$Z_t= \varepsilon_t - 0,61236\varepsilon_{t-1} - 0,76417\varepsilon_{t-L-1} \qquad ; \text{với } L=12 \qquad (3.175)$$

Z_{t-1} đã biết, ε_t và ε_{t-1} là thành phần ngẫu nhiên tính được từ số ngẫu nhiên ξ. Thay vào (3.180) tính được Z_t. Tiếp tục như vậy được chuỗi có độ dài tuỳ ý(thường lấy 1000 số hạng). Chuỗi này sẽ cho các đặc trưng thống kê không thay đổi.

3.4.1.7. Mô hình ARIMA tích bội.

Có thể phối hợp tất cả các thành phần mùa và không mùa vào một dạng chung gọi là mô hình ARIMA tích bội (multiplicative). Cũng sử dụng toán tử dịch chuyển viết được dạng chung của mô hình là:

$$\phi_P(B)\phi(B^L)Z_t = \delta + \theta_q(B)\theta_Q(B^L)\varepsilon_t \qquad (3.182)$$

Trong đó:

$\phi(B)=1-\phi_1 B-\phi_2 B^2-...-\phi_P B^P$ là thông số tự hồi quy không mùa

$\phi_{PS}(B^2)=1-\phi_{1L}B^{1L}-\phi_{2L}B^{2L}-...-\phi_{PL}B^{PL}$ là thông số tự hồi qui mùa

$\theta_q(B)=1-\theta_1 B-\theta_2 B^2-...-\theta_q B^q$ là thông số trung bĩnh trượt không mùa

$\theta_{QS}(B)=1-\theta_{1L}B^{1L}-\theta_{2L}B^{1L}-...-\theta_{QL}B^{QL}$ là thông số trung bĩnh trượt mùa

σ là hằng số.

Ví dụ với mô hình trung bình trượt không mùa bậc 2 (q= 2) và mùa bậc 1 (Q_S=1) có thể viết như sau:

$$\phi_P(B)\phi_P(B^2)Z_t=\delta+\theta_q(B)\theta_Q(B^{1L})\varepsilon_t \qquad (3.182)$$

Trong đó: $\phi_P(B)=(1-0)$ không có tự hồi qui không mùa.

$\phi_P(B^{1L})=(1-0)$ không có tự hồi qui mùa.

$\theta_2(B)=1-\theta_1 B-\theta_2 B^2$ cho $q=2$.

$\theta_2(B^{1L})=1-\theta_{1L}(B^{1L})$ cho $Q=1, L=12$ (tháng).

Thay thế vào (3.182) ta được:

$$(1-0)(1-0)Z_t=\delta+(1-\theta_1 B-\theta_2 B^2)(1-\theta_{1L}B^{1L})\varepsilon_t.$$

Sử dụng toán tử dịch chuyển biến đổi ta được dạng cuối cùng:

$$Z_t=\delta+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\theta_{1L}\varepsilon_{t-4}+\theta_1\theta_{1L}\varepsilon_{t-5}+\theta_2\theta_{1L}\varepsilon_{t-6} \qquad (3.183)$$

3.4.2. Mô hình MARKOV (MARKOV MODEL)

Mô hình Markov (trong trường hợp biến rời rạc gọi là xích Markov) để mô phỏng các quá trình thuỷ văn là trường hợp riêng của quá trình Markov. Xích Markov gồm có đơn và phức.

3.4.2.1 Xích Markov đơn. (Simple Markov chain).

Xích Markov đơn là xích có bước chuyển bằng 1 tức là chỉ xét mối tương quan của các số hạng kề nhau. Hai dãy số hạng kề nhau lập thành hai chuỗi đại lượng X và Y. Khi đó hàm phân bố đồng thời của chúng có dạng (3.34 và (3.35)

Trong dạng các đại lượng ngẫu nhiên rời rạc ta được:

$$P(x,y)=\sum_{i\le x}\sum_{j\le y}P(X<x,Y<y) \quad (3.184)$$

trong đó tổng được lấy cho tất cả các giá trị của các đại lượng ngẫu nhiên i<x và j<y.

- Các đặc trưng của hệ thống 2 đại lượng ngẫu nhiên như sau:

+Kì vọng toán học của đại lượng x.

$$\begin{aligned}\bar{x}&=M(x)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy\\ \bar{y}&=M(y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}ú(x,y)dxdy\end{aligned} \quad (3.185)$$

+ Phương sai:

$$\begin{aligned}D(x)&=\sigma_x^2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\bar{x})^2f(x,y)dxdy\\ D(y)&=\sigma_y^2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y-\bar{y})^2f(x,y)dxdy\end{aligned} \quad (3.186)$$

+ Hệ số tương quan giữa hai đại lượng:

$$r_{xy}=r=\frac{M[(x-\bar{x})(y-\bar{y})]}{\sigma_x\sigma_y} \quad (3.187)$$

Từ hàm phân bố đồng thời (3.34) có thể xác định được hàm phân bố xác suất một chiều (3.36):

$$\begin{aligned}F_1(x)&=F(x,\infty)=\int_{-\infty}^{X}\int_{-\infty}^{\infty}f(x,y)dxdy\\ F_2(y)&=F(y,\infty)=\int_{-\infty}^{Y}\int_{-\infty}^{\infty}f(x,y)dxdy\end{aligned} \quad (3.36)$$

và hàm mật độ tương ứng :

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy$$
$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx \qquad (3.188)$$

Chú ý rằng vì 2 đại lượng x và y là phụ thuộc nên:

$$F(x,y) \neq F_1(x).F_2(y)$$
$$f(x,y) \neq f_1(x).f_2(y). \qquad (3.189)$$

Cũng từ hàm phân bố đồng thời (3.34) có thể xác định được các hàm phân bố có điều kiện(hàm chuyển) như sau:

$$F_3(x)_y = \frac{F(x,y)}{F_2(y)} \quad \text{và} \quad F_4(y)_x = \frac{F(x,y)}{F_1(x)}$$
$$f_3(x)_y = \frac{f(x,y)}{f_2(y)} \quad \text{và} \quad f_4(y)_x = \frac{f(x,y)}{f_1(x)} \qquad (3.190)$$

Trong thực tế thuỷ văn không đủ tài liệu để xây dựng các hàm mật độ 2 chiều. Vì vậy người ta thường đưa ra những giả thiết về chúng. Hàm này phải đủ đơn giản để dễ dàng xác định xác suất có điều kiện, đồng thời phải đảm bảo sự phù hợp với hàm phân phối một chiều nghĩa là:

$$F(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$
$$víi\ f(x,y) \geq 0 \qquad (3.191)$$

Thường trong mô hình hoá người ta thừa nhận rằng, với các đại lượng ngẫu nhiên chuẩn hoá hoặc gần chuẩn thì hàm phân bố xác suất đồng thời 2 chiều và phân bố có điều kiện cũng là phân bố chuẩn. Còn đối với các đại lượng ngẫu nhiên phân bố gamma thì chúng là phân bố gamma.

- Để mô hình hoá ta phải xác định giá trị của số hạng sau Q_{i+1}(tương ứng với đại lượng ngẫu nhiên y). Khi đã biết số hạng trước Q_i(tương ứng với đại lượng ngẫu nhiên x). Như vậy ta phải xác định các đặc trưng của hàm phân bố có điều kiện y(x).

+ Kì vọng toán học (trung bình) có điều kiện $\overline{y}_x$ được xác định từ quan hệ hồi qui của hai đại lượng y và x.

$$\overline{y}_x = f(x) \tag{3.192}$$

Quan hệ (3.192) không nhất thiết là tuyến tính, tuy nhiên dạng tuyến tính là đơn giản nhất và cho hàm mật độ f(x,y) hợp lý nhất. Do vậy $\overline{y}_x$ được xác định theo hệ thức:

$$\overline{y}_x = \overline{y} + r\frac{\sigma_y}{\sigma_x}(x - \overline{x}) \tag{3.193}$$

trong đó: σ_y, σ_x là phương sai không điều kiện.

$\overline{x}, \overline{y}$ là kì vọng không điều kiện.

Vì là cùng một chuỗi số nên ta có các quan hệ sau:

$$\sigma_y = \sigma_x = \sigma$$
$$\overline{x} = \overline{y} = \overline{X}.$$

Có thể coi $\overline{y} = \overline{x} = 1$ mà không giảm tính tổng quát. Khi ấy (3.193) trở thành :

$$\overline{x}_{i+1} = \overline{y}_x = 1 + r(x_i - 1) \tag{3.194}$$

+ Phương sai có điều kiện $\sigma(y/x)$ và khoảng lệch xác suất có điều kiện $\phi(y/x)$ được xác định tuỳ thuộc vào các dạng tương quan được xem xét.

- Ratkovich(1977) đã phân tích và tổng hợp thành 5 dạng khác nhau của xích Markov đơn để mô tả các quá trình dòng chảy trung bình năm.

*Dạng 1: Mô hình cho các đại lượng ngẫu nhiên độc lập. Dạng này là đơn giản nhất. Với các thông số đã cho $\overline{Q}, C_v, C_S$ và tiếp nhận dãy số ngẫu nhiên phân bố đều ξ trên đoạn (0,1) là tần suất p_i, theo hàm phân bố đã chọn (chuẩn, gamma) xác định được chuỗi thủy văn với phân bố đã cho.

$$Q_i = \overline{Q} + \phi_i(\xi_i, C_S)\sigma \tag{3.195}$$

*Dạng 2: Mô hình cho các đại lượng ngẫu nhiên phân bố chuẩn, giữa chúng có tương quan chuẩn. Khi đó ta có các đặc trưng:

+ Kỳ vọng có điều kiện xác định theo (3.194).

+ Phương sai có điều kiện:

$$\sigma_{i+1} = \sigma\sqrt{1 - r_0^2} \tag{3.196}$$

Trong đó: σ là phương sai không điều kiện

r_0 là hệ số tương quan giữa các đại lượng ngẫu nhiên phân bố chuẩn + Việc mô hình hoá có thể thực hiện theo hệ thức:

$$Q_{i+1} = 1 + r_0(Q - 1) + \phi_{i+1}C_v\sqrt{1 - r_0^2} \tag{3.197}$$

Trong đó: C_v là hệ số biến đổi không điều kiện

ϕ_{i+1}là khoảng lệch xác suất có điều kiện: $\phi_{i+1} = f(\xi_i)$

*Dạng 3: Mô hình cho các đại lượng ngẫu nhiên có phân bố gamma nhưng có tương quan gần chuẩn nghĩa là khi ấy phương sai có điều kiện cũng quan hệ như tương quan chuẩn :

$$\sigma_{i+1} = \sigma\sqrt{1 - r_g^2} \tag{3.198}$$

+ Việc mô hình hoá có thể thực hiện theo công thức:

$$Q_{i+1} = 1 + r_g(Q_i - 1) + \phi_{i+1}C_v\sqrt{1 - r_g^2} \tag{3.199}$$

Trong đó r_g là hệ số tương quan của các đại lượng ngẫu nhiên có phân bố gamma.

*Dạng 4: Mô hình cho các đại lượng ngẫu nhiên có phân bố gamma và giữa chúng có phân bố gamma. Trong trường hợp này phương sai có điều kiện phụ thuộc vào số hạng đứng trước nó:

$$\sigma_{i+1} = \sigma\sqrt{(1 - r_g^2) + 2Q_i r_g (1 - r_g)} \quad (3.200)$$

Như vậy khác với tương quan chuẩn, ở tương quan gamma phương sai có điều kiện sẽ tăng lên khi số hạng đứng trước tăng. Tuy nhiên kỳ vọng của chúng bằng nhau và có giá trị:

$$M\left[\sigma^2 (y / x)\right] = \sigma^2 (1 - r^2) \quad (3.201)$$

Giá trị lớn nhất của hệ số biến đổi có điều kiện trong cả hai trường hợp tương quan đạt được khi x= 0 và bằng:

$$\text{Max}\left[C_v (y / x)\right] = C_v \frac{1 + r}{1 - r} \quad (3.202)$$

Nói chặt chẽ hơn thì phân bố có điều kiện không thực sự là phân bố gamma mà chỉ hướng đến nó khi $r \to 0$. Khi $x \to 0$ thì các thông số có điều kiện tiệm cận đến các thông số phân bố không điều kiện nghĩa là khi x nhỏ thì phân bố có điều kiện gần đến phân bố gamma.

Việc mô hình hoá được thực hiện theo công thức:

$$Q_{i+1} = 1 + r_g (Q_i - 1) + \phi_{i+1} C_v \sqrt{(1 - r_g^2) + 2 r_g Q_i (1 - r_g)} \quad (3.203)$$

*Dạng 5: Mô hình hoá cho các tần suất giữa các số hạng tức, là mô phỏng chuỗi tần suất của các số hạng mà không phải là mô phỏng trực tiếp các giá trị của chúng.

Nguyên tắc cơ bản của việc mô hình hoá này khác hẳn với các mô hình vừa trình bày ở trên. Khi đó kì vọng toán học có điều kiện thay đổi phụ thuộc vào độ lệch tần suất của các số hạng so với giá trị trung bình. Còn tương quan là mối liên hệ giữa các đại lượng ngẫu nhiên phân bố đều trên đoạn (0,1). Cấu trúc mô hình không phụ thuộc vào cả dạng phân bố, cả về giá trị bằng số của các thông số. Chuỗi này có thể coi như một biến dạng của tương quan gamma. Mật độ phân bố có điều kiện trong trường hợp này luôn luôn bằng 1.

Hàm phân bố có điều kiện có các thông số :

+ Kì vọng toán học: $\overline{p}_{i+1} = \frac{1}{2} + r_0(p_i - \frac{1}{2})$ (3.204)

+Phương sai : $\sigma_{i+1} = \sqrt{\frac{1-r_0^2}{12}}$ (3.205)

Còn hàm phân bố không điều kiện khi đó có:

+ Kì vong: $p = \frac{1}{2}.$ (3.206)

+ Phương sai: $\sigma = \frac{1}{\sqrt{12}}.$

Hàm phân bố có điều kiện được khai triển dưới dạng một chuỗi, trong mức độ chính xác cho phép có thể lấy 5 số hạng và có thể dùng nó trực tiếp để mô hình hoá, xác định các số hạng tần suất p_{i+1}:

$$F = 1 - p_{i+1} = \int_0^{p_{i+1}} f(p_i, p_{i+1})dp_{i+1} = p_{i+1} + 3r_p p_{i+1}(p_{i+1}-1)(2p_i-1) +$$

$$+\frac{5}{2}p_{i+1}(p_{i+1}-1)(2p_{i+1}-1)\left[3(2p_i-1)^2-1\right] \; +7r_p^3 p_{i+1}(p_{i+1}-1)[1+5p_{i+1}(p_{i+1}-1)]$$

$$(2p_i-1)[10p_i(p_i-1)+1]+9r_p^4 p_{i+1}(p_{i+1}-1)(2p_{i+1}-1)$$

$$[1+7p_{i+1}(p_{i+1}-1)]\{10p_i(p_i-1)[7p_i(p_i-1)+2]+1\} \quad (3.207)$$

Sau đó chuyển từ tần suất p_{i+1} sang dãy số hạng Q_{i+1} bằng cách giải phương trình(3.165) đối với p_{i+1}.

$$\int_0^{x_{i+1}} f(Q_{i+1})dx_{i+1} = p_{i+1}$$

trong đó $f(Q_{i+1})$ là hàm phân bố xác suất lựa chọn, có thể là chuẩn, PearsonIII, Kritxki-Menken, .v.v.

Sơ đồ tiến hành mô hình hoá theo mô hình Markov đơn gồm các bước sau:

(1). Xác định các đặc trưng thống kê không điều kiện $\overline{x},\sigma,C_v$ theo tài liệu đã có.

(2) Tạo chuỗi ngẫu nhiên γ có phân bố đều trên đoạn (0,1)theo các chương trình đã có trên máy tính.

(3) Lấy tần suất có điều kiện bằng chuỗi số ngẫu nhiên vừa tạo ra : $p_i=\gamma_i$.

(4) Xác định các đặc trưng thống kê có điều kiện:

+ Kì vọng có điều kiện theo công thức (3.193) hay (3.194).

+ Phương sai có điều kiện theo công thức (3.196),(3.198) hay (3.200).

Bảng 3.4: So sánh thông số thống kê của chuỗi quan trắc và mô hình hoá

Chỉ số	Các đặc trưng thống kê				
	Cv	Cs	Cv/Cs	$r_p(1)$	$r_g(1)$
Giá trị thực của sông VN	0,175	0,262	1,5	0	0
Giá trị theo mô hình dang(1)	0,173	0,331	1,9	-0,005	-0,056
Giá trị thực của sông Mekong	0,125	0	0	0,40	0,40
Giá trị theo mô hình dạng (2)	0,120	-0,119	-0,099	0,35	0,36
Giá trị theo mô hnìh dang (5)	0,122	0,047	0,038	0,38	0,36

(5) Cho giá trị ban đầu x_1 tuỳ ý. Chẳng hạn có thể tiếp nhận $Q_1=k_1=0{,}45$.

(6) Xác định độ lệch xác suất có điều kiện ϕ_{i+1} từ dạng hàm phân bố thông qua bảng tra hay công thức tính gần đúng.

(7) Xác định giá trị của chuỗi mô hình hoá Q_i theo các công thức (3.195), (3.197), (3.199), (3.203) và (3.207).

Ví dụ đối với chuỗi dòng chảy năm của sông ngòi Việt Nam, có hệ số tương quan giữa các năm kề nhau của lưu vực sông Mê Công tương đối lớn: r_1= 0,3-0,4, còn đối với các sông khác khá nhỏ: $r_2 \approx$ 0,0-0,15. Các hệ số tương ứng khác là :$C_{v1} \approx$

0,175; C_{S1}= 1,5C_v ; C_{v2} = 0,125; $C_{S2} \approx 0$. Tiến hành mô hình hoá theo 5 dạng mô hình của xích Markov đơn cho các kết quả sau với chuỗi 1000 số hạng :

3.4.2.2. Xích Markov phức (Complex Markov chain)

Trên cơ sở giả thiết hàm phân bố 1 chiều của tất cả các đại lượng ngẫu nhiên là như nhau và hàm phân bố có điều kiện cũng có dạng như hàm không điều kiện, Xvanhiđde(1977) đề nghị một phương pháp mô hình hoá như sau:

- Giả thiết hàm phân bố nhiều chiều có dạng :

$$f_{n+1}(Q_1, Q_2, ..., Q_n) = C_{0n} \prod_{k=1}^{n+1} f_1(Q_k) + \sum_{1 \le i \le k \le n+1}^{n+1} C_{ik} \sigma(Q_i - Q_k) \prod_{m=1}^{n+1} f_2(Q_m) \quad (m \neq k)$$

(3.208)

Trong đó :

+ $C_{ik} = r(Q_i, Q_k)$ là hệ số tương quan của hai đại lượng Q_i và Q_k .

+ $f_1(Q)$ là hàm phân bố nhiều chiều cho tất cả các đại lượng ngẫu nhiên. Nó tồn tại nếu $C_{ik} \ge 0$ và $C_{0n} \ge 0$.

+ C_{0n}= 1- $\sum C_{ik}$

- Hàm phân bố có điều kiện có dạng:

$$f(Q_{n+1} / Q_1, Q_2, ..., Q_n) = \begin{cases} F_1(Q_{n+1}) \text{ nếu } Q_i - Q_k = 0 \text{ dù chỉ với một cặp } i, k \\ \dfrac{C_{0n}}{C_{0,n-1}} F_1(Q_{n+1}) + \dfrac{1}{C_{0,n+1}} \sum C_{i,n+1} I(Q_{n+1} - Q_i) \\ \qquad \text{nếu} Qi - Qo \neq 0 \text{ với mọi } i, k \end{cases} \quad (3.208)$$

Trong đó I(Q) là hàm Hevisaiđa

$$I_x(Q) = \begin{cases} 0 \text{ nếu } Q < 0 \\ 1 \text{ nếu } Q \ge 0 \end{cases} \quad (3.209)$$

Theo (3.208) và (3.209) có thể tiến hành mô hình hóa lần lượt các số hạng của chuỗi theo sơ đồ xích Markov phức.

- Trong thực tế sử dụng phương pháp đơn giản hơn. Đó là mô hình do Redơnhicovxki(1969) đề xuất, xuất phát từ giả thiết các đại lượng ngẫu nhiên có phân bố chuẩn :

$$Q_i = \overline{Q}_i - \sum_{j=1}^{p}(Q_{i-j} - \overline{Q}_{i-j})\frac{\sigma_i}{\sigma_{i-j}}.\frac{D_{i(i-j)}}{D_{ii}} + \phi_i\sigma_i\sqrt{\frac{D}{D_{ii}}} \qquad (3.210)$$

Trong đó: Q_{i-j} là giá trị của chuỗi trong j bước về phía trước (j= 1,2,...,p)

p là bậc hồi qui

$\overline{Q}_{i-j}$ là giá trị trung bình của thời khoảng (i-j)

σ_i, σ_{i-j} là phương sai của chuỗi thời khoảng i và (i-j)

D là định thức của ma trận tương quan.

D_{ii} và $D_{i,(i-j)}$ là phần phụ đại số tương ứng với các phần tử r_{ij} và $r_{i,(i-j)}$ của định thức D ma trận tương quan.

ϕ_i là độ lệch xác suất được xác định theo dạng hàm phân bố khi tiếp nhận dãy số ngẫu nhiên γ_i có phân bố đều trên đoạn (0,1) làm tần suất p_i.

Nếu quá trình là dừng (đối với dòng chảy năm) thì $\overline{Q}_{i-j} = \overline{Q}_i = \overline{Q}$ và $\sigma_{i-j} = \sigma_i = \sigma$, do đó ta có mô hình dạng đơn giản hơn:

$$Q_i = \overline{Q} + \sum_{j=1}^{p}(Q_{i-j} - \overline{Q})\frac{D_{i,(i-j)}}{D_{ii}} + \phi\sigma\sqrt{\frac{D}{D_{ii}}} \qquad (3.211)$$

Như vậy các bước tiến hành mô hình hoá theo mô hình Markov phức như sau:

(1) Theo chuỗi số liệu quan trắc xác định được các đặc trưng $\overline{Q}, \overline{Q}_{i-j}, \sigma_i, \sigma_{i-j}$.

(2) Phân tích ma trận tương quan để chọn số bậc hồi qui p. Chỉ chọn bậc hồi qui ứng với các hệ số tương quan có ý nghĩa thống kê.

(3)Theo ma trận tương quan xác định định thức D và các phần phụ đại số $D_{ii}, D_{i,(i-j)}$.

(4) Tiếp nhận số ngẫu nhiên γ_i phân bố đều trên đoạn (0,1) là giá trị tần suất p_i . Theo dạng hàm phân bố xác định được ϕ_i.

(5) Theo các công thức (3.210) và (3.211) tính được Q_i

(6) Coi Q_i là Q_{i-1} và lại tính lại theo bước(5). Lần lượt như vậy ta được chuỗi cần mô phỏng.

- Nhiều tác giả đề nghị sử dụng mô hình tự hồi qui tuyến tính, dạng AR(p) nhưng đối với chuỗi không dừng và thành phần ngẫu nhiên sử lý theo mô hình Markov.

$$Q_{n,m} = a_{m1}Q_{n,m-1} + a_{m2}Q_{n,m-2} + ... + a_{mp}Q_{n,m-p} + C_{m0}Q_n + C_{m1}Q_{n-1} + ...$$
$$+ C_{mn}Q_{n-m} + \sigma_{mn}\xi_{nm}$$

Với: $$Q_n = d_1Q_{n-1} + d_2Q_{n-2} + ... + d_qQ_{n-q} + \sigma_n\xi_n \qquad (3.212)$$

Trong đó Q_n , Q_{nm} là giá trị của năm thứ n và của thời đoạn thứ m trong năm thứ n nghĩa là thời đoạn thứ (n-1)T+m theo cách đánh số thông thường. Như vậy nếu là quá trình tháng thì $Q_{n,m-i}$ sẽ là tháng thứ (T+m-i) của năm thứ n với T= 12.

$\xi_{n,m}$ và ξ_n là đại lượng ngẫu nhiên phân bố chuẩn với các thông số (0,1).

$a_{mi}(1 \le i \le k), C_{ml}(0 \le i \le l)$ là các hệ số hồi qui biểu thị mối liên hệ của giá trị thời đoạn thứ m với các (m-k) thời đoạn trước, với năm hiện tại và (n-l) năm trước.

d_k là hệ số hồi qui biểu thị liên hệ của năm thứ n với (n- k) năm trước.

σ_{mn}, σ_n là phương sai có điều kiện của tháng thứ m trong năm thứ n và của năm thứ n .

Các chỉ số p,l,q chỉ ra bậc hồi qui trong xích Markov phức . Khi thay đổi p,l,q thì thay đổi các hệ số a_{mp}, C_{ml} và d_q.

Trong mô hình (3.212) không yêu cầu các đại lượng phải có phân bố chuẩn.

Khi thành phần trung bình năm được coi là dừng và $Q_{n,m}$ là quá trình có chu kỳ và có liên hệ dừng với Q_n (trường hợp tháng trong năm) thì mô hình có dạng đơn giản hơn:

$$Q_{nm} = a_{m1}Q_{n,m-1} + a_{m2}Q_{n,m-2} + ... + a_{mk}Q_{n,m-k} + \sigma_m \xi_{nm} \qquad (3.213)$$

Nghĩa là chỉ xét và mô hình hoá cho từng năm một .

Các hệ số a_{mi} có thể xác định theo công thức truy hồi Durbin. Còn phương sai có điều kiện σ_m có thể xác định nhờ công thức truy hồi phương sai dư:

$$\left[\sigma_m^{(k)}\right]^2 = \left[\sigma_m^{(k-1)}\right]^2 \left[1 - a_{m,k}^{(k)}.a_{m+1,1}^{(k)}\right] \qquad (3.214)$$

$\xi_{n,m}$ là dãy đại lượng ngẫu nhiên độc lập chuẩn với thông số (0,1) . Để nhận được các giá trị mô hình của năm đầu tiên ta đưa vào (3.212) và (3.213) các giá trị của năm đầu tiên. Để đơn giản trong quá trình mô hình hoá Xvanhiđde đề nghị chọn k= 11 khi mô hình hoá theo tháng, nghĩa là gọn trong một năm.

3.4.3. Mô hình động lực thống kê Aliôkhin *(Statistic dynamical model)*

Mô hình này căn cứ vào giả thiết rằng đại lượng thủy văn trung bình thời khoảng có liên hệ với các thời khoảng trước.

$$Q_t = f(Q_{t-1}, Q_{t-2}, ..., Q_{t-p}) \qquad (3.215)$$

Đưa quan hệ (3.215) về dạng tuyến tính ta được:

$$Q_t = k_1Q_{t-1} + k_2Q_{t-2} + ... + k_pQ_{t-p} = \sum_{\tau=1}^{p} k(\tau)Q(t-\tau) \qquad (3.216)$$

Trong đó : Q_t là giá trị tính theo mô hình.

Q_{t-1}, Q_{t-2}, ..., Q_{t-p} là giá trị các hời khoảng trước.

$k(\tau)$ là các hệ số hồi qui được xác định theo nguyên tắc bình phương nhỏ nhất, làm cực tiểu tổng bình phương sai số.

Hệ số k(τ) có hể xác định theo công thức:

$$k(\tau) = \frac{D_{0\tau}}{D_{00}} \qquad (3.217)$$

Trong đó D_{00} là định thức của ma trận tương quan cấp p.

$$D_{00} = \begin{vmatrix} 1 & r_1 & r_2 \;.... & r_p \\ r_1 & 1 & r_1 & r_{p-2} \\ & .... & & \\ r_{p-1} & r_{p-2} \;.... & 1 & \end{vmatrix}$$

(3.218)

$D_{0\tau}$ là định thức con nhận được từ D_{00} bằng cách thay cột τ của nó bằng cột số hạng tự do của (3.218).

Trong mô hình này bậc hồi qui p được xác định tuỳ thuộc vào hệ số tương quan bội lý thuyết $\overline{R}(p)$ và tiêu chuẩn ngẫu nhiên σ(p).

$$\overline{R}(p) = \sqrt{1 - \frac{D_P}{D_{P-1}}} \qquad (3.219)$$

$$\sigma(p) = \frac{\sigma_{Sp}^2}{\sigma_p^2} \qquad (3.220)$$

Trong đó : D_p và D_{p-1} là định thức con bậc p và (p-1) của định thức ma trận tương quan.

σ_{sp}^2 là phương sai của chuỗi mô hình.

σ_p^2 là phương sai của chuỗi thực tế.

p được chọn khi có $\overline{R}(p)$ lớn nhất và σ(p) nhỏ nhất.

Để nâng cao thêm độ chính xác của mô hình (3.216) có thể sử dụng phương pháp trực giao hoá các đại lượng ngẫu nhiên Q_t , Q_{t-1} ,..., Q_{t-p}. Đồng thời có thể thêm vào mô hình (3.216) một thành phần ngẫu nhiên ε_t sao cho các đặc trưng thống kê không đổi như đã làm với mô hình ARIMA.

3.4.4. Mô hình THORMAT-FIERING.

a. Mô hình Thormat- Fiering bậc 1

Mô hình Thormat-Fiering là trường hợp riêng của mô hình ARIMA với phép lọc đơn giản:

$$Z_t = Q_t - \overline{Q}$$

-Mô hình thường dùng cho chuỗi số liệu tháng, dạng chung của mô hình là:

$$Z_{i+1} = \overline{Z}_{j+1} + a_j(Z_i - \overline{Z}_j) + \xi_i \sigma_{j+1}\sqrt{(1 - r_j^2)} \qquad (3.221)$$

Trong đó : Z_i , Z_{i+1} là giá trị tháng thứ i và (i+1) trong chuỗi mô phỏng

(i= 1,2,..,N).

$\overline{Z}_{j+1}, \overline{Z}_j$ là giá tị trung bình tháng thứ j và (j+1) trong năm.

(j= 1,2,...,12)

ξ_i là số ngẫu nhiên phân bố chuẩn có các thông số (0,1) hoặc là độ lệch xác suất chuẩn ứng với số ngẫu nhiên γ_i

σ_{j+1} là phuơng sai của tháng thứ (j+1)

r_j là hệ số tương quan giữa tháng thứ j và tháng j+1

a_j là hệ số hồi qui của tháng thứ j:

$$a_j = \frac{r_j \sigma_{j+1}}{\sigma_j} \qquad (3.222)$$

Như vậy mỗi tháng có một mô hình Thormat-Fiering. Mô hình Thormat-Fiering bậc 1 gần giống mô hình AR(1) và thực chất là mô hình Markov đơn. Như vậy 12 tháng có 12 mô hình AR(1).

b. Mô hình Thormat- Fiering bậc 2

*Mô hình Thormat-Fiering bậc 2 có dạng gần giống AR(2) hay có dạng gần giống mô hình Markov phức với số bậc hồi qui là 2.

$$Q_{i+1} = \overline{Q}_{j+1} + a_{1j}(Q_i - \overline{Q}_j) + a_{2j}(Q_{i-1} - \overline{Q}_{j-1}) + \xi_i \sigma_i \sqrt{1 - r_j^2} \qquad (3.223)$$

Trong đó: a_{1j}, a_{2j} là hệ số hồi quy xét đến mối liên hệ của hai số hạng về phía trước. Các hệ số này cũng có thể xác định theo công thức truy hồi Durbin hay chương trình tối ưu hoá.

Việc mô hình hoá theo mô hình Thormat - Fiering không khác biệt gì lắm với việc mô phỏng theo ARIMA hoặc mô hình Markov. Lưu ý rằng việc tính toán thường bắt đầu từ tháng đầu năm (tháng 1). Giá trị Q_{i+1} tính theo (3.222) và (3.223) được coi là Q_i cho việc mô hình hoá tiếp theo.

Dĩ nhiên mô hình tiếp theo vẫn có dạng (3.222) và (3.223) nhưng với các hệ số hồi quy khác đi.

Mô hình Thormat - Fiering cũng được dùng để mô hình các chuỗi thuỷ văn có quan hệ tương hỗ như đã trình bày trong mô hình không- thời gian ở ♣3.2.

Ngày nay, mô hình ngẫu nhiên đang được cải tiến. Thành phần ngẫu nhiên đang được xem xét trong mối liên hệ với tác động của các nhân tố địa vật lý. Và một lớp mô hình ngẫu nhiên tất định đã hình thành góp phần tạo nên một kết quả mô phỏng tốt hơn và dự báo chính xác hơn.

Chương 4
ỨNG DỤNG CỦA MÔ HÌNH TOÁN THUỶ VĂN

Các mô hình toán thủy văn ngày càng tỏ ra có nhiều ứng dụng hiệu quả trong các lĩnh vực sản xuất và đời sống cùng với sự phát triển nhanh chóng của công nghệ máy tính và phương pháp tính, các mô hình ngày càng được hoàn thiện hơn và nâng cao độ chính xác, giải quyết có hiệu quả các bài toán tính toán, dự báo, quy hoạch và quản lý tài nguyên nước có thể phânlàm hai lĩnh vực ứng dụng chính. Đó là ứng dụng trong dự báo tính toán thủy văn và trong tính toán thủy lợi.

4.1. ỨNG DỤNG TRONG TÍNH TOÁN THUỶ VĂN

Trong lĩnh vực này các mô hình tính toán được ứng dụng nhiều và thực sự có hiệu quả. Một số bài toán chính như sau:

4.1.1. Sử lý và quản lý số liệu thủy văn

Các mô hình tất định xem xét đến tất cả các nhân tố hình thành dòng chảy vì vậy nó phát hiện được nguồn gốc gây ra sai số đo đạc. Từ đó nó giúp cho việc phân tích đánh giá chất lượng số liệu, đánh giá tính hợp lý của mạng lưới các trạm quan trắc thủy văn. Mạng lưới các trạm khí tượng thủy văn có nhiệm vụ thu thập số liệu, cung cấp thông tin về biến đổi theo không gian, thời gian của các nhân tố địa vật lý, cảnh quan đến dòng chảy. Về lý thuyết, càng tăng số trạm thì số thông tin thu thập được càng nhiều. Tuy nhiên cũng chỉ nên tăng đến một mức độ nào đó, vì tăng quá nhiều cũng không mang lại độ chính xác cao hơn mà chi phí cho hoạt động lại tốn kém, khó khăn. Mạng lưới trạm cần phải có số lượng và sự phân bố hợp lý. Bằng mô hình toán có thể xác định được tỷ lệ hợp lý của lưới trạm cho từng mục đích khác nhau. Chẳng hạn lưới trạm để dự báo khác với lưới trạm để thu thập, đều tra cơ bản, Johanxon dùng mô hình Standford IV nghiên cứu yêu cầu lưới trạm đo mưa ở Mỹ và cho thấy kết quả như ở bảng (4,1). Nghiên cứu cũng cho thấy số trạm cần thiết để đạt độ chính xác nhất định là độc lập đối với diện tích lưu vực.

- Mô hình toán cũng giúp cho việc khôi phục và kéo dài tài liệu có kết quả. Mô hình tất định thực hiện việc mô phỏng chi tiết, nhận thức các quá trình hình thành dòng chảy và các quá trình thủy văn khác. Đồng thời trong quá trình giải quyết đã sử dụng máy tính và các phương pháp tính hiện đại nên độ chính xác được nâng cao. Do đó kết quả khôi phục và kéo dài tài liệu theo mô hình có thể coi là đáng tin cậy và được sử dụng như đầu vào của các bài toán tính toán thủy lợi. Mô hình ngẫu nhiên cho phép xem xét sự biến đổi của dòng chảy trong một thời kỳ dài, cho phép kiểm tra tính hợp lý của các tài liệu thu thập được. Trên cơ sở mô hình hoá lượng mưa ta cũng có thể bổ sung cho chuỗi số liệu thủy văn.

Bảng 4.1: Sai số mô phỏng khi dùng lưới trạm đo mưa khác nhau

Đặc trưng	Sai số mẫu		Sai số hiệu chỉnh thông số			F lưu vực
Số trạm mưa	1	4	1	4	10	
Dòng chảy năm	4%	Nhỏ	19%	7,5%	Nhỏ	2500
Dòng chảy mặt	-	-	38%	-	-	32-500
Đỉnh lũ Qmax	10%	-	40%	-	-	100-500

4.1.2. Dự báo và tính toán thủy văn

Một lĩnh vực ứng dụng có hiệu quả rõ rệt của mô hình toán là tính toán và dự báo các đặc trưng thủy văn theo yêu cầu phòng chống lũ lụt, tính toán các dặc trưng thủy văn thiết kế cho các công trình và quản lý vận hành hệ thống lưu vực

Các mô hình toán cho phép phân tích ảnh hưởng của nhiều nhân tố, diễn biến của nhiều quá trình tương hỗ mà trước đây bằng các phép tính tương quan giản đơn, các công thức kinh nghiệm không giải quyết được. Dự báo theo mô hình có thể cho kết quả của nhiều đặc trưng trên nhiều địa điểm cùng một lúc với chất lượng được nâng cao rõ rệt, góp phần tích cực vào việc phòng chống lũ lụt trên phạm vi rộng. Các mô hình tất định như mô hình SSARR, TANK, NAM, HEC-1 đang dần thể hiện ưu thế và thay thế các phương pháp truyền thống trong công tác tính toán, dự báo thủy văn. Các mô hình ngẫu nhiên góp phần dự báo hạn dài và hạn vừa phục vụ cho công tác vận hành hồ chứa. Đồng thời nó cũng giúp cho việc quản lý, qui hoạch nguồn nươc trên một lưu vực từ thượng nguồn đến cửa sông được thuận lợi. Các mô hình 1 chiều, 2 chiều, 3 chiều mô tả tác động nhiều chiều đa dạng các yếu tố thủy văn, cho phép giải quyết nhiều bài toán phức tạp về diễn biến chất lượng nước, độ mặn của vùng cửa sông.

Mô hình toán cũng giúp cho việc xác định các đặc trưng thủy văn, trong đó có các đặc trưng thống kê, được chính xác hơn. Các thông số thống kê xác suất xác định trên cơ sở tài liệu quan trắc thường không đảm bảo độ chính xác cần thiết. Mô hình toán cho phép đánh giá mức độ tin cậy của các thông số này để từ đó đưa ra các hệ số hiệu chỉnh cho từng khu vực, làm tăng độ chính xác xác định các đặc trưng thủy văn cho thiết kế công trình.

Mô hình toán cho phép tạo ra một chuỗi có độ dài tuỳ ý, phản ánh nhiều mối liên hệ mà trong chuỗi số liệu quan trắc ngắn không thể hiện được. Theo chuỗi mô hình hoá có thể đánh giá được độ lệch có thể của những giá trị, xác định theo mẫu và tổng thể. Chuỗi mô hình hoá với độ dài đủ lớn có thể coi là một tổng thể. Chia chuỗi này thành các mẫu ngắn hơn, xác định các đặc trưng thống kê tương ứng. So sánh các giá trị theo mẫu và tổng thể, có thể đưa ra những nhận xét và kết luận về mối quan hệ này. Từ đó đưa ra các hệ số hiệu chỉnh thích hợp cho các thông số xác định trên cơ sở chuỗi quan trắc ngắn để có giá trị chuẩn của tổng thể, Blokhinov(1976) ứng dụng chuỗi mô hình hoá để nghiên cứu hiệu quả của các phương pháp xác định thông số thống kê và cho rằng phương pháp thích hợp tối đa cho kết quả tốt nhất.

Giá trị trung bình số học là ước lượng vững, không chệch và có hiệu quả của kỳ vọng toán học, nghĩa là $\overline{X} = M\{x\}$. Kết quả mô hình hoá cho thấy trung bình số học thực tế là không lệch. Tuy nhiên phương sai của trung bình số học xác định theo mẫu tăng lên khi tăng hệ số tương quan giữa các số hạng kề nhau. Phân bố xác suất của giá trị trung bình mẫu có dạng lệch dương.

Kritxki và Menken đã đưa ra biểu thức để đánh giá độ lệch của giá trị trung bình số học theo mẫu như sau:

$$\sigma_{\overline{Q}} = \frac{\sigma_x}{\sqrt{n}} \sqrt{\frac{1 + \dfrac{2r}{n(1-r)}\left(n - \dfrac{1-r^n}{1-r}\right)}{1 - \dfrac{2r}{n(n-1)(n-r)}\left(n - \dfrac{1-r^n}{1-r}\right)}} \qquad (4.1)$$

- Khi r nhỏ có biểu thức:

$$\sigma_{\overline{X}} = \frac{\sigma_x}{\sqrt{n}} \sqrt{\frac{1+r}{1-r}} \qquad (4,2)$$

- Còn khi r =0 thu được:

$$\sigma_{\overline{X}} = \frac{\sigma_x}{\sqrt{n}} \qquad (4,3)$$

Trong đó : r là hệ số tương quan giữa các số hạng kề nhau,

σ_x là độ lệch quân phương tính theo mẫu,

$\sigma_{\overline{X}}$ là độ lệch quân phương của giá trị trung bình đã hiệu chỉnh,

n là dung lượng mẫu.

Từ kết quả mô hình hoá cho thấy sự khác nhau của độ lệch tính theo công thức (4.1) và theo chuỗi mô hình hoá là không thực sự lớn. Tuy nhiên khi tăng r thì sự khác nhau này tăng lên, Với r =0,7; C_v =1,0 và n= 10 thì sự khác nhau đạt tới 10-20%. Khi r không lớn (r <0,5) và n>10 thì công thức (4,1) cho kết quả đủ tin cậy, không cần hiệu chỉnh.

Độ lệch quân phương (phương sai) và hệ số biến đổi xác định theo mẫu có phân bố lệch âm. Khi tồn tại mối tương quan trong chuỗi thì độ lệch này tăng lên. Quan hệ giữa hệ số biến đổi xác định theo mẫu C_{vm} và theo tổng thể C_v có dạng:

$$C_{vm} = \left[C_v - \frac{C_v}{4n}\left(1 + 3C_v^2\right)\right]\sqrt{1 - \frac{2r}{n(n-1)(n-r)}\left(n - \frac{1-r^n}{1-r}\right)} \qquad (4.4)$$

Trên cơ sở chuỗi mô hình hoá cũng cho thấy giá trị C_{vm} theo mẫu của phân bố Kritxki-Menken có độ lệch âm nhỏ hơn giá trị nhận được theo đường cong Pearson III. Sự khác nhau này tăng lên thực sự khi dung lượng mẫu n nhỏ và khi hệ số tương quan r nhỏ.

Tương tự theo chuỗi mô hình hoá cũng xác định được quan hệ của hệ số lệch C_{sm} theo mẫu và C_s theo tổng thể như sau:

$$C_s = \frac{\left(n + 4 + 4C_v\right)\left(1 + 2r^2\right)}{n} C_{sm} \qquad (4.5)$$

Bằng chuỗi mô hình hoá có thể xác định được mối tương quan giữa các đặc trưng thống kê của các đại lượng thủy văn.

*Đối với phân bố Gamma có các quan hệ:

$$
\begin{aligned}
&R(M_x, M_y) = R_{xy} \\
&R(\sigma_x, \sigma_y) = \frac{R_{xy}(R_{xy} + 2C_{vx}C_{vy})}{\sqrt{(1 + 2C_{vx}^2)(1 + 2C_{vy}^2)}} \\
&R(C_{vx}, C_{vy}) = R_{xy}^2 \\
&R(C_{sx}, C_{sy}) = R_{xy}^3 \\
&R(R_x, R_y) = R_{xy}^2
\end{aligned}
\qquad (4.6)
$$

*Còn đối với phân bố chuẩn nhận được:

$$
\begin{aligned}
&R(M_x, M_y) = R_{xy} \\
&R(\sigma_x, \sigma_y) = R_{xy}^2 \\
&R(C_{vx}, C_{vy}) = \frac{R_{xy}(R_{xy} + 2C_{vx}C_{vy})}{\sqrt{(1 + 2C_{vx}^2)(1 + 2C_{vy}^2)}}
\end{aligned}
\qquad (4.7)
$$

Trong đó: $R_{xy} = R(1)$ là hệ số tương quan giữa hai số hạng kề nhau

$M_x, M_y, C_{vx}, C_{vy}, C_{sx}, C_{sy}$ là các đặc trưng thống kê của các đại lượng thủy văn x và y.

Như vậy các hệ thức liên hệ của phân bố gamma và phân bố chuẩn khác nhau ở chỗ hai hệ thức của phương sai (σ_x, σ_y) và hệ số biến đổi (C_{vx}, C_{vy}) đổi chỗ cho nhau, Còn các hệ thức khác là hoàn toàn tương tự.

*Đối với hệ số tương quan ta biết rằng độ chính xác của nó khi bước trượt τ tăng là không cao, có sai số lớn. Bằng chuỗi mô hình hoá ta cũng xác định được quan hệ như sau:

$$
R = R_m \frac{1 + \frac{0{,}7}{1 + C_v} + 3R_m}{n} \qquad (4.8)
$$

Trong đó: R_m là giá trị xác định theo mẫu.

Kết quả nghiên cứu cũng cho thấy biên độ dao động của $R(\tau)$ sẽ giảm dần khi tăng dung lượng mẫu n.

Từ những kết quả nghiên cứu theo mô hình ở trên người ta đưa vào hệ số hiệu chỉnh cho từng đặc trưng xác định cho chuỗi quan trắc ngắn, và đưa ra các chỉ dẫn khi sử dụng.

*Dao động của dòng chảy trong năm thể hiện rõ tính chu kỳ rõ rệt, gắn liền với chu kỳ quay của quả đất quay quanh mặt trời. Về tính chu kỳ của dao động nhiều năm của dòng chảy thì còn có nhiều ý kiến khác nhau. Tuy nhiên trong thực tế vẫn xuất hiện các nhóm năm nhiều và ít nước với những xác suất nhất định. Đặc điểm này ảnh hưởng rõ rệt đến việc tính toán và điều hành hồ chứa. Chuỗi thủy văn quan trắc có độ dài ngắn không đủ làm sáng tỏ quy luật này. Chuỗi mô hình hoá với độ dài lớn có thể đánh giá xác suất xuất hiện của những nhóm năm với lượng nước khác nhau. Trên cơ sở nghiên cứu chuỗi mô hình hoá Ratkôvich Đ. IA.(1984) cho thấy ở vùng ẩm ướt xác suất xuất hiện một năm ít nước là 3%, 2 năm liên tục ít nước chỉ là 0,3%. Trong khi đó vùng khô hạn các chỉ số này là 14% và 12%. Có nghĩa là khả năng xuất hiện nhóm năm ít nước ở vùng khô hạn cao hơn vùng ẩm ướt. Nhóm năm với độ dài 1-2 năm có độ lặp lại lớn hơn nhóm năm có độ dài lớn hơn. Ở sông có mô đun dòng chảy lớn (>10-20 l/s km^2) có nhiều khả năng xuất hiện những nhóm năm nhiều nước với độ dài lớn. Ở các sông có mô đun dòng chảy nhỏ (< 1l/s km^2) khả năng xuất hiện những nhóm năm nhiều nước với độ dài ≥ 7 năm là khá nhỏ.

Với sự tăng lên của hệ số tương quan R(1), độ dài nhóm năm cũng tăng lên. Đối với tần suất p ≥ 50% ta có các quan hệ sau:

$$\overline{m} = a + bR(1) \qquad (4.9)$$

Trong đó: a, b là những hệ số.

Sự phân bố của nhóm năm nhiều nước và ít nước đối với một điểm phân vị xác suất(quantile) là đối xứng gương, nghĩa là độ lặp lại của nhóm năm ít nước đối với tần suất p cũng bằng độ lặp lại của nhóm năm nhiều nước đối với tần suất (1-p), Tuy nhiên cần phải thấy rằng một mô hình dù hoàn hảo đến đâu, dù các kỹ thuật sử lý có làm cho số liệu phù hợp tốt với tài liệu đã quan trắc, thì kết quả đó vẫn chưa phải là hoàn toàn tin cậy, đưa ngay ra sủ dụng được. Các kết quả phải được kiểm nghiệm, so sánh với các phương pháp truyền thống, với các lưu vực tương tự. Mỗi mô hình có những giả thiết của mình, đòi hỏi những điều kiện phù hợp với nó. Do vậy khi sử dụng mô hình cần phải chú ý các vấn đề sau đây:

a.Về lưu vực ứng dụng.

Lưu vực ứng dụng vào mô hình phải đáp ứng đầy đủ tài liệu đầu vào theo yêu cầu của mô hình. Mô hình được chỉnh lý tốt nhưng đầu vào không đảm bảo thì kết quả dĩ nhiên là không chính xác. Đối với mô hình tất định, đầu vào là số liệu của các nhân tố ảnh hưởng đến đặc trưng thuỷ văn cần tính toán, dự báo. Còn đối với mô hình ngẫu nhiên đó là các thông số thống kê và dạng hàm phân bố xác suất. Đây là vấn đề then chốt, quyết định chất lượng và hiệu quả mô hình.

Trong quá trình sử dụng mô hình cũng cần phải quan tâm đến tính đại biểu và đồng bộ của tài liệu đầu vào. Dòng chảy tại mặt cắt khống chế phản ảnh ảnh hưởng tổng hợp của các nhân tố phân bố trên toàn lưu vực. Sự phân bố không gian của các yếu tố chi phối các kết quả tính toán. Các điểm đo phân bố đều trên lưu vực sẽ phản ánh đúng và đầy đủ tác động của các nhân tố hình thành dòng chảy. Các thông số và kết quả tính toán theo mô hình do vậy sẽ có độ chính xác cao.

b.Về mô hình ứng dụng.

Mỗi mô hình đều có những giả thiết và phạm vi, điều kiện ứng dụng riêng. Các mô hình mô phỏng tốt các qua stình thủy văn trên lưu vực nhỏ, nhưng lại không đem lại kết quả khi áp dụng cho lưu vực vừa và lớn. Mô hình phức tạp, nhiều thông số phản ánh được nhiều quá trình, nhưng lại có thể tích luỹ sai số, và nhất là khi số liệu đo đạc không đủ, không tương ứng. Do vậy không phải cứ mô hình phức tạp là cho kết quả có độ chính xác cao hơn. Khi tài liệu không đầy đủ thì mô hình đơn giản ít thông số có khi lại cho kết quả tốt hơn. Mô hình tất định tập trung thích hợp để mô phỏng cho lưu vực nhỏ, còn mô hình với thông só phân bố lại tỏ ra có ưu thế đối với các lưu vực vừa và lớn. Mô hình tất định có kết quả tốt với các bài toán có thời gian dự kiến ngắn. Trong khi đó các mô hình ngẫu nhiên sử dụng để dự báo hạn dài và vừa lại cho hiệu quả cao hơn. Do vậy khi chọn mô hình để áp dụng cần xem xét điều kiện bài toán, cấu trúc mô hình, phạm vi ứng dụng, đồng thời phải xét đến khả năng tài liệu quan trắc tương ứng.

c.Về đánh giá và sử dụng kết quả.

Các phương pháp toán học hiện đại có thể chỉnh lí làm cho kết quả phù hợp tốt với tài liệu quan trắc, tuy nhiên các kết quả tính theo mô hình chưa hẳn đã đủ tin cậy để đem vào ứng dụng trong thực tế.

Qua thực tế ứng dụng, tài liệu đầu vào không phải bao giờ cũng thoả mãn đầy đủ và chính xác. Nó chỉ đáp ứng được một mức độ nào đó nào đó của mô hình. Các tiêu chuẩn đánh giá mô hình cũng chỉ phản ánh một phần độ chính xác. Kết quả tính toán có thể phù hợp với đỉnh nhưng lại không phù hợp với tổng lượng hay thời gian xuất hiện. Các phương pháp tối ưu hoá và thử sai có những ưu điểm rõ rệt nhưng cũng có những nhược điểm nhất định không dễ gì khắc phục, và cũng có phần nào đó phụ thuộc chủ quan người hiệu chỉnh thông số.

Mặt khác lưu vực không phải là bất biến. Có rất nhiều nguyên nhân tự nhiên và nhân tạo làm thay đổi cảnh quan lưu vực. Chẳng hạn việc phá rừng, trồng rừng hay xây dựng các công trình thuỷ lợi làm thay đổi quá trình tập trung dòng chảy, thay đổi quá trình tổn thất, do đó làm thay đổi các thành phần cân bằng nước. Các điều kiện đó đã làm sai lệch các kết quả tính toán so với thực tế, đôi khi sự sai lệch này là đáng kể. Do đó các kết quả tính theo mô hình cần kiểm tra so sánh, sử lí hiệu chỉnh trước khi đưa ra thực tế. Đánh giá so sánh kết quả có thể dùng các phương pháp sau:

+Theo phương trình cân bằng nước: So sánh số liệu tính toán và thực đo.

+ Theo sự biến đổi không gian: So sánh với các lưu vực lân cận hay lưu vực tương tự cùng sử dụng số liệu.

+ Theo các chỉ tiêu thống kê: Đánh giá theo chuỗi tính toán và thực đo.

Các sai lệch của kết quả tính theo mô hình cần được lí giải làm rõ nguyên nhân. Sau đó sử lí, hiệu chỉnh và đưa ra các chỉ dẫn cần thiết khi áp dụng vào thực tế. Nếu có sai lệch đáng kể phải xác định lại thông số trên nền tài liệu mới. Nếu cần thiết phải thay đổi lại cấu trúc, thậm chí chọn một mô hình khác.

4.2. ỨNG DỤNG TRONG TÍNH TOÁN THUỶ LỢI.

Trong lĩnh vực tính toán thuỷ lợi, qui hoạch và quản lí tài nguyên nước, mô hình toán cũng được áp dụng rộng rãi,

4.2.1. Đánh giá các đặc trưng thống kê

Một công trình thuỷ lợi hoạt động trong một hệ thống có những tác động tương hỗ với nhau. Mô hình toán cho phép đánh giá ảnh hưởng của công trình đến các mặt khác trong hệ thống, lấy đó làm cơ sở cho việc thiết kế các công trình mới trong cùng hệ thống. Với mô hình toán có thể tiến hành đánh giá, so sánh về các chỉ tiêu thiết kế tại các vị trí khác nhau, cũng như tại một vị trí nhưng với các phương án có qui mô khác nhau tương ứng với các điều kiện nước đến, yêu cầu nước dùng và và mức bảo đảm cấp nước khác nhau. Thông qua các kết quả mô phỏng theo nhiều phương án khác nhau, mô hình toán cho phép đánh giá mức độ ảnh hưởng của hệ thống cung cấp và tiêu thoát nước đến các chỉ tiêu thiết kế. Từ đó đưa ra kiến nghị về việc phát triển hợp lý các công trình trong hệ thống lưu vực.

Thông qua chuỗi mô hình hoá có thể đánh giá các đặc trưng thiết kế. Các đặc trưng thuỷ lợi thường được quan tâm nhiều là dung tích hồ chứa V(hay hệ số dung tích β), yêu cầu dùng nước α và mức bảo đảm cấp nước P. Các đặc trưng này xác định từ chuỗi quan trắc có độ dài giới hạn thường dẫn đến những kết quả sai lệch. Sự sai lệch này có mấy nguyên nhân. Đó là do các thông số thống kê đưa vào tính toán chưa được hiệu chỉnh gây ra sai số hệ thống. Đó cũng có thể do sự dao động của các đặc trưng thuỷ lợi chưa được đề cập đến gây nên. Thông qua chuỗi mô hình hoá ta có thể hiệu những chỉnh độ lêch này. Chuỗi mô hình hoá với độ dài lớn cho ta nhiều tổ hợp có khả năng mà trong chuỗi quan trắc ngắn không có được. Cũng chia chuỗi mô hình hoá thành các mẫu với độ dài khác nhau. Tính toán các đặc trưng thuỷ lợi và so sánh nó với với các đặc trưng tính theo tổng thể (chuỗi dài mô hình hoá).

Kết quả cho thấy khi khối lượng phép thử (hay độ dài chuỗi mô hình) lớn thì giá trị các đặc trưng thủy lợi thực tế trùng với kì vọng toán học. Ở đây kì vọng toán học hay chuẩn được hiểu là giá trị trung bình trong một thời gian tương đối dài, khi mà chế độ thuỷ văn không đổi.

Theo kết quả nghiên cứu thấy rằng trong đa số các trường hợp giá trị mẫu của dung tích có độ lệch dương. Độ lệch này xác định được theo biểu thức:

$$\sigma_\beta = 5\frac{\beta}{n^{2/3}} \tag{4.10}$$

Giá trị của độ cấp nước hay yêu cầu dùng nước α không lệch và có độ phân tán không lớn. Độ lệch của nó được xác định theo biểu thức:

$$\sigma_\alpha = \frac{a\alpha + bp + c\beta}{n^{1,1-0,5p-0,3C_V}} \tag{4.11}$$

Trong đó: σ_α và σ_β là độ lệch chuẩn của dung tích β và khả năng cấp nước α,

` a, b, c là các hệ số phụ thuộc vào sự thay đổi của dòng chảy,

C_V là hệ số biến đổi của dòng chảy.

Các đặc trưng thuỷ lợi có thể xác định theo các phương pháp trong tính toán thuỷ lợi. Chuỗi mô hình hoá cho phép tận dụng ưu thế của phương pháp thứ tự thời gian khi yêu cầu nước dùng α thay đổi, bằng cách thực hiện tính toán liên tục theo phương pháp điều tiết toàn liệt trên cơ sở một chuỗi mô hình dài.

Căn cứ vào chuỗi mô hình hoá Ivanov và Redơnhicovxki(1969) cho rằng kết quả tính toán điều tiết theo chuỗi quan trắc có độ lệch lớn. Khả năng cấp nước α có sai số đạt tới 40% so với giá trị thực (Khi C_V=1,0 R(1)=0,5 ; β=2 và n=25). Để nhận được các giá trị không lệch cần có các công thức để hiệu chỉnh chúng. Các hệ thức này cũng được xác định thông qua chuỗi mô hình hoá.

Chuỗi mô hình hoá cho phép thực hiện tính toán thủy lợi với điều kiện nước dùng thay đổi cũng như tổn thất thay đổi. Khi có sự thay đổi lớn của bốc hơi thì cần có chuỗi mô hình hoá của bốc hơi với hệ số tương quan lấy bằng hệ số tương quan giữa các số hạng của dòng chảy để tính toán. Chuỗi mô hình hoá cũng làm sáng tỏ độ lặp lại của các khoảng thời gian gián đoạn cấp nước cũng như mức độ thiếu nước trong những khoảng gián đoạn này.

Trên cơ sở chuỗi mô hình hoá, đã xây dựng được các toán đồ, giúp cho việc xác định các đặc trưng thủy lợi được nhanh chóng và đạt độ chính xác cao. Ví dụ các toán đồ Pleskov cho các đại lượng độc lập, các toán đồ của Gugly, Xvanhiđde và Ratkovich cho các đại lượng ngẫu nhiên có tương quan. Bằng chuỗi mô hình hoá Ratkovich cũng cho thấy khi tính đến hệ số tương quan giữa các số hạng, với trường hợp điều tiết sâu, mức bảo đảm cấp nước lớn thì dung tích hồ chứa cần tăng cao hơn. Chẳng hạn đối với sông có M =4-10l/s km^2 và R(1) =0,3 thì dung tích tăng lên 1,5 lần. Còn với sông M <1l/s km^2 và R(1)= 0,5 thì dung tích tăng lên 2 lần.

4.2.2. Quy hoạch và điều hành hệ thống nguồn nước

Trong lĩnh vực này mô hình toán được ứng dụng để giải quyết có hiệu quả các bài toán của thực tế. Có mấy nội dung chính như sau:

a.Về quy hoạch hệ thống nguồn nước.

Các mô hình cho phép mô phỏng hệ thống lưu vực với các phương án khác nhau, và từ đó rút ra các kết luận về số các công trình cần xây dựng, vị trí công trình cũng như quy mô kích thước của nó trong hệ thống. Hiệu quả của mô hình được cân nhắc trên tác động tổng hợp của các nhân tố trên lưu vực. Mô hình toán cho phép xét đến tác động tổng hợp này, đồng thời nó cung cấp đầu vào cho các bài toán quy hoạch đáng tin cậy.

b.Về điều hành hệ thống.

Các công trình hoạt động trên lưu vực có liên hệ với nhau, vì vậy điều hành hệ thống nguồn nước là một bài toán tổng hợp phức tạp. Mô hình toán cho phép xem xét đến các giải pháp cụ thể bằng cách phân tích chi tiết các khả năng nước đến, yêu cầu nước dùng, lợi ích kinh tế xã hội và khả năng đảm bảo của công trình. Mô hình toán cũng đảm bảo dự báo nguồn nước đến, một đầu vào đặc biệt quan trọng để có thể điều chỉnh biểu đồ điều phối, nâng cao hiệu quả hoạt động của công trình. Mô hình toán tất định cũng như ngẫu nhiên làm tăng độ chính xác các dự báo phục vụ vận hành các công trình thủy lợi. Nếu thực hiện việc nối mạng, thu thập và truy cập thông tin nhanh chóng thì hiệu quả điều hành hệ thống càng được nâng cao.

c.Về quản lý lưu vực.

Mô hình toán cho phép tính toán các nguôn nước của các lưu vực trong các điều kiện khai thác khác nhau, cũng như khi tác động của con người lên các cảnh quan của lưu vực. Về mặt này mô hình toán có thể thay thế cho mô hình vật lý, thay thế cho các bãi dòng chảy thực nghiệm tốn kém, làm sáng tỏ vai trò của các nhân tố địa vật lý đến dòng chảy cũng như ảnh hưởng của dòng chảy đến các đặc trưng của lưu vực. Từ các điều kiện khai thác lưu vực, mô hình toán giúp cho việc dự báo tính toán các quá trình xói trên lưu vực, khả năng bồi lấp hồ chứa. Từ đó xây dựng các phương án phòng chống có hiệu quả, bảo vệ lưu vực và tăng tuổi thọ công trình. Trên cơ sở phân tích bằng mô hình toán, đề xuất các biện pháp xây dựng công trình đảm bảo khai thác lưu vực hợp lý và bền vững.

Một lưu vực không chỉ nằm trong một nước mà thường bao gồm nhiều quốc gia. Việc khai thác sử dụng của một nước phụ thuộc rất nhiều vào chủ quan của con người và các hoạt động của các quốc gia trên cùng lưu vực. Mô hình toán giúp ta tìm được lời giải tổng hợp cho việc lợi dụng nguồn nước chung, cũng như ảnh hưởng của từng hoạt động của từng quốc gia đến lưu vực. Từ đó có một sự hợp tác liên quốc gia lâu dài, có giải pháp phối hợp chung để khai thác lưu vực có lợi nhất, không làm ảnh hưởng lẫn nhau.

Các mô hình toán còn là một "công cụ" rất thuận tiện để nghiên cứu thủy văn, nhất là đánh giá các nhân tố ảnh hưởng đến dòng chảy không thua kém gì các mô hình vật lý. Các lời giải từ mô hình có thể định hướng cho những công trình nghiên cứu có giá trị trong thực tế.

Mô hình toán nhận thức quá trình hình thành dòng chảy. Vì vậy nội dung mô hình toán có thể là bài giảng có tính khoa học, tổng hợp và hiện đại của thủy văn học.Tuỳ theo yêu cầu, mức độ có thể đưa vào các chương trình đào tạo, các bài giảng chuyên đề ở bậc đại học và sau đại học.

4.3. BÀI TẬP ỨNG DỤNG.

4.3.1. Bài tập số 1: ỨNG DỤNG MÔ HÌNH SSARR.

1. Cho số liệu diễn toán bao gồm:

- Số liệu mưa: Các trạm K T Plâycu, Kontum và Ban mê thuột để dự báo dòng chảy cho khu giữa từ Pakse- Kratie.

- Số liệu dòng chảy các trạm Pakse, Kratie , Tân Châu, Châu Đốc để diễn toán theo 2 nhánh sông Tiền và sông Hậu.

- Số liệu mực nước lớn nhất ngày tại các trạm Pakse, Kratie, Tân Châu, Châu Đốc.

Tất cả tài liệu được cho trong bảng (4.3)

Yêu cầu dự báo lũ vùng đồng bằng sông Cửu Long tại hai trạm thủy văn Tân Châu và Châu Đốc bằng mô hình SSARR.

Bảng 4.3. Số liệu mưa - dòng chảy khu vực trung, hạ lưu sông MêKông

TT	Dòng chảy Q(m^3/s)				Mưa khu giữa X(mm)		
	Pakse	Kratie	T.châu	C.đốc	K.tum	Pleiku	B.m.t.
1	167470	255230	191	150	102	0	130
2	182850	223540	195	156	20	220	0
3	210370	260240	199	161	893	590	50
4	287450	340060	227	183	221	85	230
5	289020	409360	269	203	0	20	20
6	286850	387700	268	207	120	82	0
7	294250	378380	268	207	110	21	100
8	301690	379130	277	220	1080	34	20
9	302130	401290	290	230	50	101	180
10	288390	392290	294	235	0	3	80
11	277030	377390	294	238	153	230	3
12	290320	358210	291	236	20	260	60
13	293810	364050	290	233	510	410	10
14	294030	365510	288	231	460	390	0
15	300370	373410	293	236	550	600	420
16	304760	390290	301	242	630	140	360
17	316340	439490	315	250	730	180	230
18	314990	466320	326	262	890	770	140

19	317010	463220	335	271	481	262	50
20	341080	485750	343	278	190	390	100
21	378650	518380	353	288	30	320	13
22	397810	545920	366	297	70	176	630
23	401400	568440	380	309	203	70	5
24	389380	562910	397	327	230	662	20
25	374260	549410	410	340	140	40	380
26	357450	537020	418	353	120	700	206
27	344180	527650	424	367	4	1	650
28	317730	508990	427	377	0	3	310
29	284780	460740	430	381	4	310	40
30	254430	420530	429	384	10	930	180
31	234540	370180	429	387	0	190	198
32	206010	378620	429	390	0	0	20
33	202560	330420	429	391	8	120	0
34	201360	312300	429	390	680	70	0

2. Các bước thực hiện

(1).Công nghệ dự báo:

∗Chương trình SSARR được tổ chức khá linh hoạt bao gồm 4 chương trình chính kiểm soát toàn bộ hoạt động của các chương trình con, tạo ra một bộ chương trình thống nhất tổ chức theo khối (hình 4.1)

∗Tạo file số liệu phục vụ cho tính toán dự báo.

-File số liệu bao gồm 2 phần: Phần bảng quan hệ tham số và phần số liệu thủy văn (bảng 4, 3).

+Các tham số được tham khảo từ các lưu vực tương tự để có bộ tham số sơ bộ ban đầu đưa vào tính toán. Sau đó chúng được điều chỉnh trong quá trình chạy mô hình.

+Số liệu được đưa vào tiếp theo sau phần tham số.

- File số liệu được tổ chức dưới dạng bìa.

- Các số liệu cũng có thể nhập trực tiếp vào máy, nhưng chỉ được lưu giữ tạm thời.

Chương trình nhập số liệu	Trao đổi giữa người và máy
(Input data)	(Interractive Drive)
Chương trình tổng hợp dòng chảy	File tính toán
từ mưa trên lưu vực	(Work file)
(Watershed Model)	
Chương trình diễn toán	File trình tự chạy
đoạn sông và hồ chứa	(Bulk file)
(River and Reservoir model)	
Chương trình kết suất kết quả	File vẽ đồ thị
(Output Reports)	(Drawer file)

Hình 4.1 Sơ đồ công nghệ mô hình SSARR,

∗Xây dựng sơ đồ hình thái lưu vực diễn toán:

- Sơ đồ hình thái lưu vực diễn toán gồm 2 dòng song song (hình 4,2)

- Theo sơ đồ bên phải:

+Bên phải là sơ đồ dòng chảy thực đo diễn toán dòng sông.

+Bên trái là sơ đồ dòng chảy tổng hợp từ mưa rồi diễn toán trên các đoạn sông tương ứng.

- Theo sơ đồ bên trái:

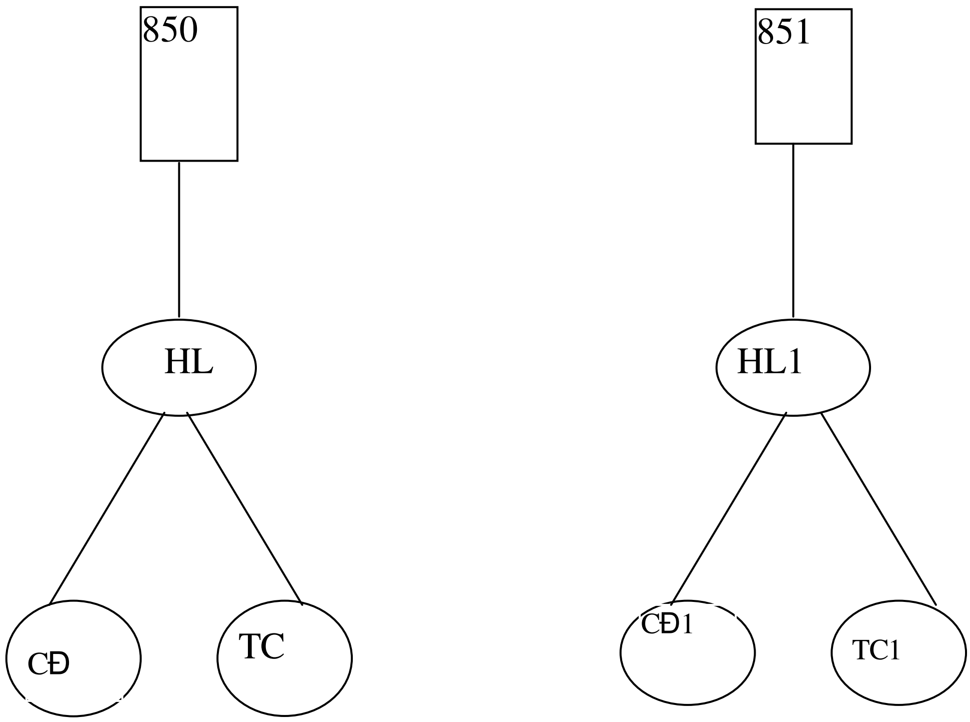
+Dòng chảy được tổng hợp từ mưa cho khu giữa Pakse-Kratie (KG PA-KR)

+Lưu lượng Pakse (PA) được diễn toán qua đoạn sông (650) rồi cộng với kết quả dòng chảy khu giữa PA-KR, để được dòng chảy tạ Kratie (KR).

+Diễn toán từ Kratie về hạ lưu (HL), Thông qua quan hệ 3 biến lưu lượng diễn toán (HL1), (HL) và mực nước tại trạm vào thời điểm dự báo để tính ra mực nước Tân Châu (TC) và Châu Đốc (CD) dự báo.

+Lưu lượng thực đo Pakse (TDPA) diễn toán qua đoạn sông (651) về điểm truyền (670). Sau đó lấy lưu lượng thực đo Kratie (TDKR) trừ đi điểm truyền (670) để được dòng gia nhập "thực đo" của khu giữa Pakse-Kratie. Kết quả được dùng để so sánh với kết quả KGPA-KR từ mưa, nhằm điều chỉnh các tham số cho KGPA-KR.

+Diễn toán lưu lượng thực đo Kratie (TDKR) về hạ lưu (HL1), sau đó thông qua quan hệ ba lưu lượng diễn toán và mực nước tại trạm vào thời điểm dự báo để suy ra mực nước Tân châu (TC1) và Châu đốc (CD1).



Hình 4.2: Sơ đồ hình thái lưu vực sông Mêkông.

Kết quả được đem ra so sánh với kết quả sơ đồ tính để điều chỉnh toàn bộ tham số cho đoạn sông.

∗ Xác định tham số:

- Các tham số được xác định theo phương pháp thử sai cho từng lưu vực bộ phận, sao cho kết quả tính toán phù hợp với tài liệu thực đo nhất.

- Việc điều chỉnh các tham số được tiến hành có lưu ý đến mức độ ảnh hưởng của từng tham số đến kết quả tính toán.

+Tham số trọng số trạm mưa W_I ảnh hưởng đến mức độ đóng góp của trạm thứ i cho lưu vực.

+Liên hệ SMI-ROP: là liên hệ thông số quan trọng nhất, ảnh hưởng đến lượng chảy tập trung và đường quá trình tính toán. Khi ROP tăng thì lưu lượng đỉnh và thể tích lũ tăng lên.

+Liên hệ RGS-RS: ảnh hưởng trực tiếp đến đỉnh và thời gian tập trung nước.

+Chỉ số bốc thoát hơi ETI: làm thay đổi lượng dòng chảy nhưng không lớn.

+Liên hệ BII-BFP: làm thay đổi dòng chảy ngầm, đường lũ xuống, nhưng không ảnh hưởng đến thể tích.

+Số lần trữ nước N và thời gian trữ nước T_S: ảnh hưởng rất lớn đến dạng đường quá trình tính. Khi N và T_S của dòng chảy mặt càng lớn thì đường càng lệch về bên phải và đỉnh lũ giảm.

+Thời gian trữ nước ngầm T_SBII: làm phần dòng chảy ngầm thay đổi, chỉ ảnh hưởng đến phần rút nước.

Trong các tham số mô hình lưu vực, ba thông số W_I, SMI-ROP và T_S là nhạy cảm nhất.

+Số lần trữ nước đoạn sông N:ảnh hưởng đến đỉnh cũng như thời gian tập trung nước, N lớn thì đỉnh càng giảm.

+Hệ số trữ nước đoạn sông KT_S càng lớn thì đường quá trình càng lệch về bên phải.

Khi điều chỉnh mô hình đầu tiên điều chỉnh quan hệ SMI-ROP, Sau đó điều chỉnh trọng số W_I để lượng lũ phù hợp. Cuối cùng điều chỉnh N và T_S để được quá trình tương ứng.

- Kết quả có bộ tham số sau:

+ Mô hình lưu vực khu giữa Pakse-Kratie.

+ Bốc thoát hơI EIT.

+ Quan hệ BII-BFP.

+ Thông số diễn toán dòng chảy:

- Mặt: N = 3, $T_S = 24^h$
- Sát mặt: N = 3, $T_S = 52^h$
- Ngầm: N = 2, $T_S = 500^h$

+ Thời gian trữ nước ngầm $T_SBII = 50^h$

+ Hệ số tỷ trọng trạm mưa W_I = 1,5 (trung bình của cả 3 trạm)

+ Giá trị BII ban đầu: 0,80 cm/ngày

+ Giá trị SMI ban đầu: 35 cm

+ Đoan sông diễn toán Paksê- Kratie

- Số lần trữ nước N=4
- Hệ số đường cong (T_S=KTS/a^n): 0,33
- KTS= 600

+ Đoạn Kratie- Tên chân

- N=5
- n=-0,20
- KTS=5

Với thông số này đỉnh xuất hiện ở Chân Đốc sớm hơn hay muộn hơn ở Tân Chân 2-3 ngày

* Kiểm định mô hình:

- Kiểm định trên số liệu phụ thuộc:

+ Kết quả mô phỏng dòng chảy cho 5 năm số liệu phụ thuộc 1964,1973, 1984, 1986, 1991 tương đối khả quan. Sai số trung bình ở cấp mực nước trên 3,5m tại Tân

Châu là 13cm, tại Châu Đốc là 14,9cm. Đỉnh lũ lớn nhất năm tại Tân Châu sai 9,5cm, thời gian suất hiện đỉnh sai 0-2 ngày, (hình 4.3)

+ Nếu lấy sai số cho phép 15cm thì kết quả mô phỏng Tân Châu chỉ đạt 67%, Châu Đốc đạt 74%. Điều này có thể là do mới chỉ xét 3 trạm mưa trên thượng lưu sông Srepok.

Kết quả theo chỉ tiêu chất lượng $\frac{S}{\sigma}$ đưa ra trong bảng 4.4

+ Bằng bộ thông số đã xác định được ở trên, tiến hành dự báo kiểm tra cho các năm 1994, 1995, 1997.

+ Kết quả cho thấy về định tính các đường quá trình tính và thực đo có sự tương đồng. Song tại một số trận lũ của năm 1995 đường tính toán thiên lớn, năm 1994 các đỉnh của 2 đường có sự lệch pha. Đó là do chưa xét được lượng ra nhập khu giữa Kratie-Tân Châu.

+ Về định lượng sai số trung bình ở mực nước trên 3,5 m tại Tân Châu là 12,4cm, tại Châu Đốc là 13cm. Đỉnh lũ lớn nhất tại Tân Châu sai 10cm, thời gian xuất hiện đỉnh sai 0-2 ngày. Tại Châu đốc sai 15cm, đỉnh sai 0-3 ngày,(Hình 4.4).

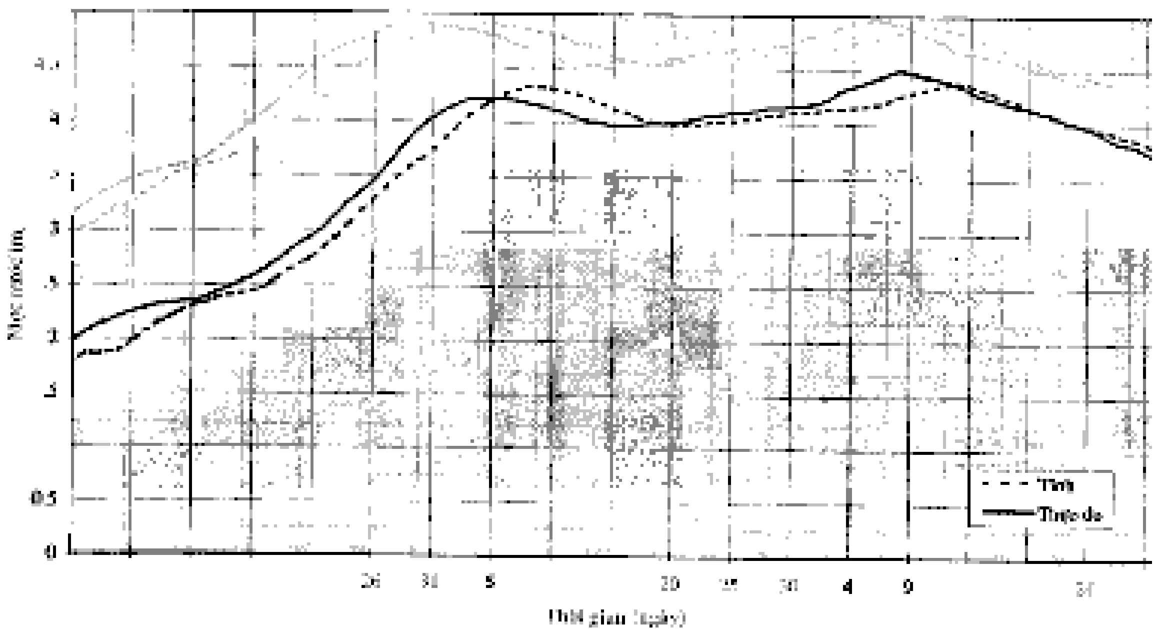
- Kiểm tra trên số liệu độc lập,

+Kiểm tra theo chỉ tiêu $\frac{S}{\sigma}$ đưa ra trong bảng 4.5.

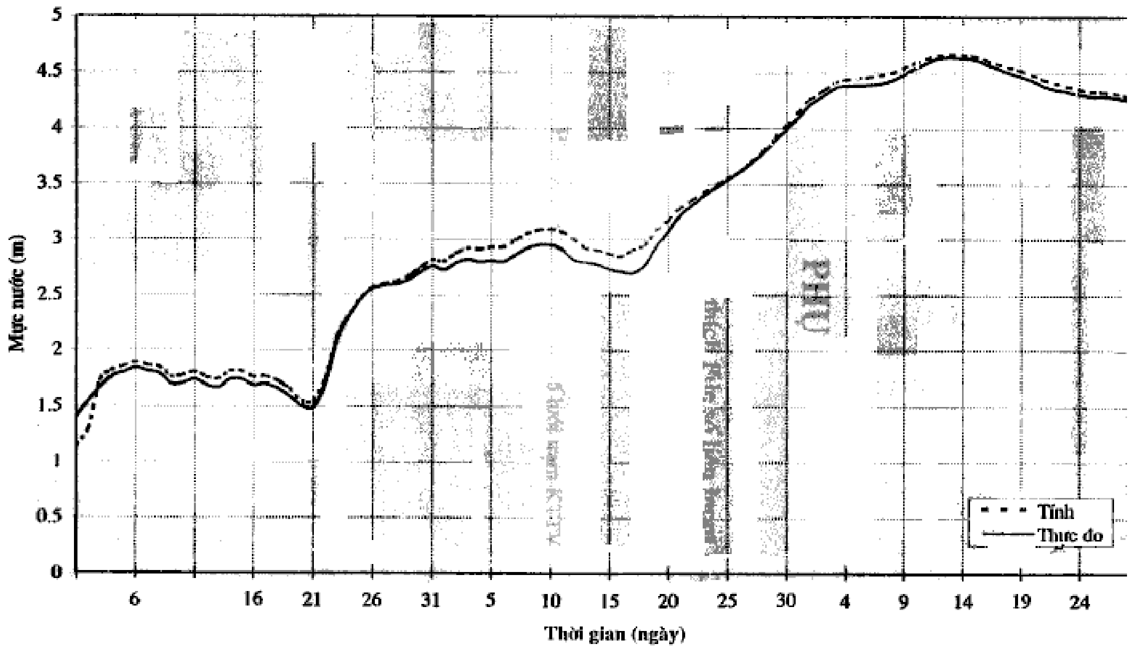
- Nhận xét: Trong dự báo tác nghiệp cần phân tích các điều kiện KTTV trên lưu vực, cập nhập trạng thái lưu vực, tham số và sai số để hiệu chỉnh kịp thời nâng cao chất lượng dự báo.

Bảng 4.4: Kết quả đánh giá mô hình theo số liệu phụ thuộc

Năm	Tỷ số S/σ		Nhận xét
	Trạm Tân Châu	Trạm Châu Đốc	
1964	0,34		Tốt
1978	0,337	0,32	Tốt
1984	0,234	0,14	Tốt
1986	0,23	0,22	Tốt
1991	0,159	0,184	Tốt



Hình 4.3: So sánh quá trình dự báo và thực tế



Hình 4.4: Kết quả dự báo

Bảng 4.5: Kết quả dự báo kiểm tra trên số liệu độc lập

Năm	Tỷ số S/σ		Nhận xét
	Trạm Tân Châu	Trạm Châu Đốc	
1994	0,24	0,24	Tốt
1995	0,54	0,554	Đạt
1997	0,43	0,456	Tốt

4.3.2. Bài tập số 2: ỨNG DỤNG MÔ HÌNH ARIMA.

1. Cho số liệu dòng chảy tháng của các trạm Kontum, Trung Nghĩa và Yaly sông Sê San (bảng 4). Yêu cầu dự báo dòng chảy tháng đến hồ chứa Yaly.

2; Các bước giải:

- Chỉ thực hiện theo ARIMA cho trạm Kontum, sau đó bằng quan hệ hồi quy giữa Kontum và Yaly sẽ tính được dòng chảy Yaly. Với trạm Trung nghĩa có thể tiến hành tương tự.

Bảng 4.5: Số liệu dòng chảy tháng các trạm Kontum, Trung nghĩa,Yaly

Năm-tháng	Kontum	Tr.nghĩa	Yaly	Năm-tháng	Kontum	Tr.nghĩa	Yaly
1994-1	61,1	48,6	135	7	57,0	104	277
2	45,6	37	112	8	102	166	394
3	37,4	30,2	84,5	9	124	208	455
4	40,3	37,7	97,8	10	172	190	424
5	48,9	47,1	127	11	227	202	498
6	59,2	81	196	12	148	108	327
7	161	348	630	1996-1	64,4	58,3	186
8	170	383	722	2	52,3	44,5	142
9	345	602	963	3	37,2	36	118
10	142	231	494	4	40	42	105
11	89,4	127	258	5	58,9	60,3	152
12	72,2	87,9	287	6	64	75,4	229
1995-1	53,1	59,9	140	7	101	195	352
2	46,1	44,2	116	8	144	334	533
3	35,6	34,4	102	9	290	432	714
4	30,1	29,8	81,3	10	215	277	617
5	31,8	39,7	101	11	414	416	917
6	36,3	53,8	123	12	235	210	484

Công nghệ: gồm 4 bước:

* Định dạng mô hình:

-Trước hết để định dạng mô hình cần đưa chuỗi về dạng dừng bằng các phép biến đổi.

+Logarit hoá Ln(x), Sau khi logarit hóa, vẽ lại đường quá trình thấy chúng vẫn còn có tính chu kỳ(hình 4.6).

+Sai phân với bước trễ một tháng (D(1)) và với bước trễ 12 tháng D(12). Vẽ lại đường quá trình sau các biến đổi trên (hình3.3,chương 3) thấy chúng có dạng răng cưa không có quy luật. Chuỗi như vậy chứng tỏ là chuỗi dừng.

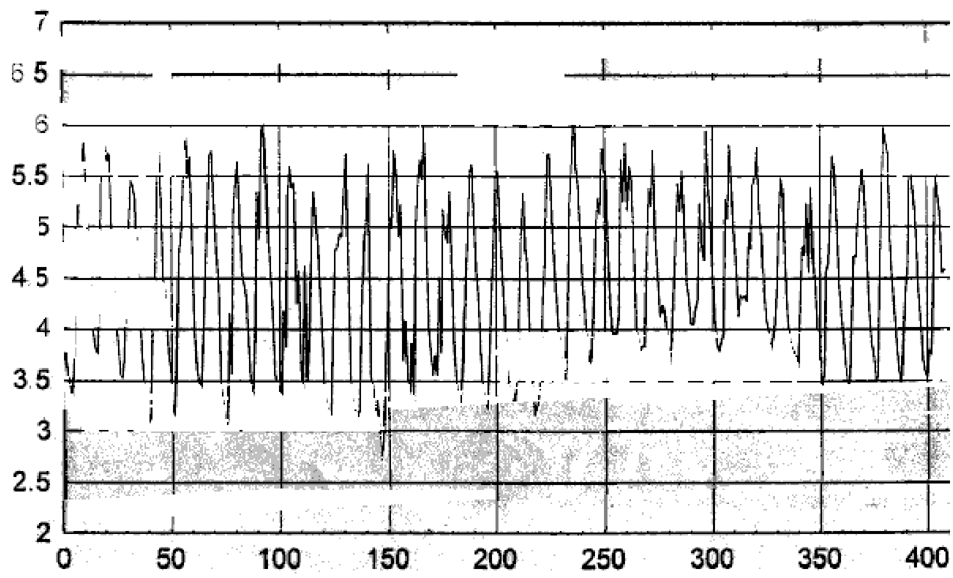
+ Phân tích hàm tự tương quan (TTQ) và tự tương quan riêng (TTQR) của chuỗi đã biến đổi thấy chúng có dạng gần tắt dần. Như vậy có thể chọn mô hình hỗn hợp ARIMA (p,d,q) (p_S, d_S, q_S) để mô phỏng chuỗi.

+ Phân tích hàm TTQ và TTQR (hình 3.4,chương 3) cùng thấy chúng có bước nhảy sau bước trễ một tháng và 12 tháng. Như vậy chúng là hàm ARIMA (1,1,1) (1,1,1) tức là có các bậc của thông số như sau:

+Bậc tự hồi quy tháng: p =1

+Bậc sai phân tháng: d =1

+Bậc trung bình trượt tháng: q =1



Hình 4.6: Quá trình lưu lượng tháng trạm Kontum sau khi log hoá

+Bậc tự hồi quy mùa: p_S=1

+ Bậc sai phân mùa: d_S=1

+ Bậc trung bình trượt mùa: q_S=1

* Xác định các thông số: Bằng phương pháp tối ưu hóa Quasi-Newtơn xác định được các thông số cơ bản cho trạm Kontum như sau(bảng 4.6):

Bảng 4.6: Các thông số của mô hình ARIMA cho trạm Kontum sông Se san

Mô hình ARIMA(1,1,1)(1,1,1) Trạm Kontum Sông Sê san Các bước biến đổi: Ln(x),D(1),D(12)				
Số quan trắc	Tổng bình phương sai số ban đầu	Tổng bình phương sai số cuối cùng	Sai số quân phương	
407	168,05	66,487	0,16359	
Các thông số(p,p_S: Tự hồi quy; q,q_S: Trung bình trượt)				
Thông số	p(1)	q(1)	p_S(1)	q_S(1)
Giá trị thông số	0,13416	0,68941	0,1497	0,8056
Sai số	0,08933	0,06819	0,05823	0,03329

* Kiểm định mô hình:

Việc kiểm định mô hình được thực hiện theo các chỉ tiêu sau:

- Phân tích sai số:

+Phân tích hàm TTQ và TTQR của sai số(hình 4.6) thấy ở các bước quan hệ đều có hệ số TTQ và TTQR nhỏ, nhỏ hơn độ lệch chuẩn của chúng, chứng tỏ chuỗi thực sự là dừng.

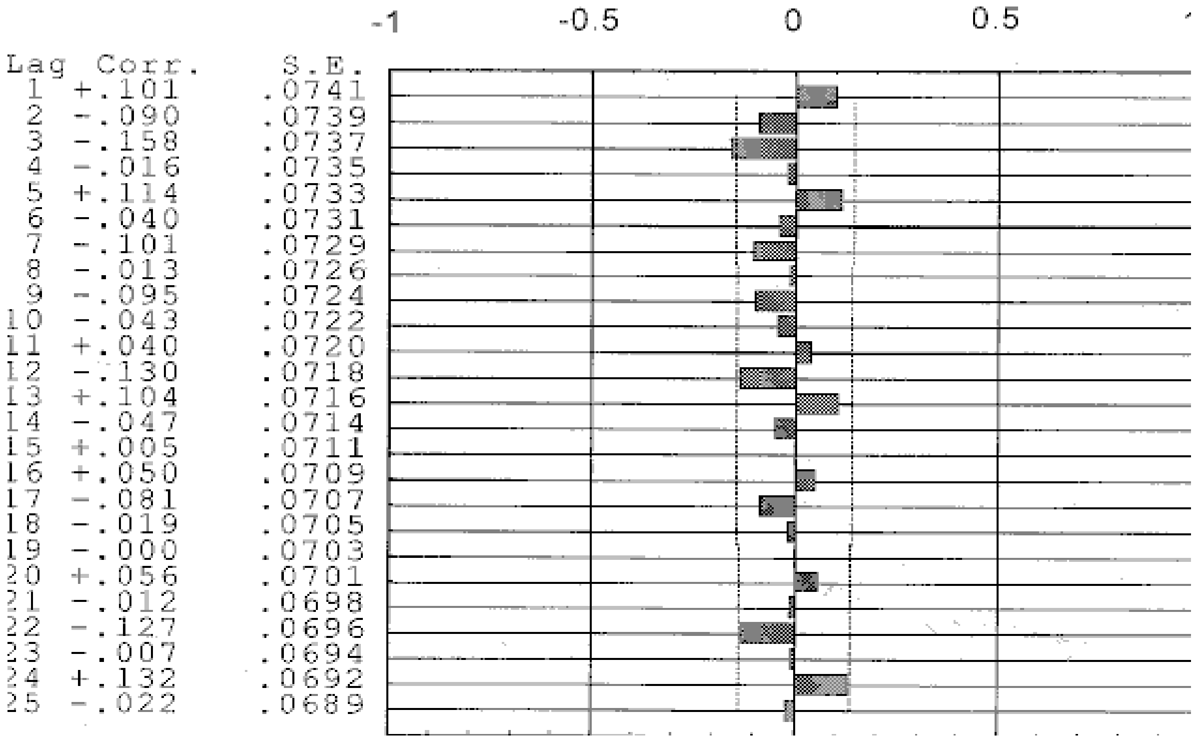
+Phân tích phân bố sai số (hình 3.6,chương 3) thấy chúng có dạng gần chuẩn.

- Kiểm định ý nghĩa các thông số: thấy rằng chỉ tiêu thống kê t đều lớn hơn nhiều t_α tra từ bảng Student với nước ý nghĩa α và số bậc tự do n- 4 = 403; $t_\alpha = 2$

$t_{P(1)}$= 16,3529 ; $t_{PS(1)}$= 21,2547

$t_{q(1)}$= 35,2131 ; $t_{qS(1)}$= 32,1632

-Kiểm định tương quan giữa các thông số cho thấy (bảng 4.7)



Hình 4.6: Hàm TTQ của sai số trạm Kontum

Bảng 4.7: Tương quan giữa các thông số

Thông số	p(1)	q(1)	$p_S(1)$	$q_S(1)$
p(1)	1	0,6852	0,0287	0,1394
q(1)	0,6852	1	0,0183	0,0278
$p_S(1)$	0,0297	0,0183	1	0,6352
$q_S(1)$	0,1394	0,0278	0,6352	1

Như vậy chứng tỏ có biểu hiện dư thừa thông số. Phân tích thêm thấy giá trị của các thông số P(1) và $P_S(1)$ có độ tin cậy kém hơn q(1) và $q_S(1)$ và giá trị thông số nhỏ hơn và sai số chuẩn lớn hơn. Do đó có thể loại bỏ hai thông số P(1) và $P_S(1)$, tức là có mô hình ARIMA (0,1,1)(0,1,1).

-Phân tích tổng bình phương sai số để dễ so sánh chọn lại mô hình ARIMA (0,1,1)(0,1,1) và cho kết quả như bảng 4.8

Bảng 4.8: Kết quả tính toán thông số của trạm Kontum sông Senan.

Mô hình ARIMA(0,1,1)(0,1,1) Trạm Kontum Sông Sê san Các bước biến đổi: Ln(x),D(1),D(12)			
Số quan trắc	Tổng bình phương sai số ban đầu	Tổng bình phương sai số cuối cùng	Sai số quân phương
179	48,309	20,527	0,11663
	q(1)	Q(1)	
Giá tri thông số	0,62558	0,75977	
Sai số	0,07346	0,04077	

So sánh hai mô hình ARIMA (1,1,1)(1,1,1) và ARIMA (0,1,1)(0,1,1) thấy chúng không có sự khác biệt lớn về tổng bình phương sai số. Theo nguyên tắc chọn mô hình sẽ chọn ARIMA (0,1,1)(0,1,1) làm mô hình dự báo, tức là có mô hình Z_t= 0,59526e_{t-1}+ 0,85216e_{t-12}.

* Dự báo: Để dự báo phải chuyển về dạng nguyên thủy.

- Trước hết đưa về dạng y= lnQ

+ Ta có: $Z_t = (y_t - y_{t-1}) - (y_{t-12} - y_{t-12-1})$

Trong đó:

$$e_{t-1} = y_{t-1} - y'_{t-1} \text{ là sai số dự báo trước 1 tháng}$$
$$e_{t-12} = y_{t-12} - y_{t-12-1} \text{ là sai số dự báo trước 12 tháng}$$

Vậy có: $y_t = y_{t-1} - y_{t-12} - y_{t-13} - 0{,}59526 e_{t-1} + 0{,}85216 e_{t-12}$

- Chuyển về dạng nguyên thủy:

$$Q_t = e^{yt}$$

- Theo các biểu thức trên dự báo kiểm tra cho các năm phụ thuộc, tức là các năm đưa vào để xác định thông số mô hình.

Bảng 4.9: Kết quả dự báo Q tháng Yaly theo ARIMA

TT	Năm-tháng	$Q_{tđo}$	$Q_{dự\ báo}$	TT	Năm-tháng	$Q_{tđo}$	$Q_{dự\ báo}$
1	1994-1	135	136,6	19	7	277	235,2
2	2	112	101,0	20	8	394	460,6
3	3	84,5	86,6	21	9	455	442,5
4	4	97,8	80,2	22	10	424	452,7
5	5	127	118,6	23	11	498	280,5
6	6	196	167,0	24	12	327	199,2
7	7	630	237,2	25	1996-1	186	134,1
8	8	722	464,7	26	2	142	99,2
9	9	963	447,0	27	3	118	85,1
10	10	494	456,8	28	4	105	78,7
11	11	258	283,1	29	5	152	116,4
12	12	287	201,4	30	6	229	164,1
13	1995-1	140	135,3	31	7	352	233,2
14	2	116	100,1	32	8	533	456,7
15	3	102	85,8	33	9	714	438,7
16	4	81,3	79,5	34	10	617	448,6
17	5	101	117,5	35	11	917	277,9
18	6	123	165,6	36	12	484	197,4

Kết quả đưa ra trong bảng 4.9

* Đánh giá chất lượng dự báo:

+ Tỷ số $\frac{s}{\sigma} = 0,50$

+ Mức bảo đảm: P = 83,3%

+ Nhận xét: Phương án thuộc loại tốt.

* Từ trạm Kontum có thể suy ra dòng chảy dự báo cho tuyến đập Yaly theo một phương trình hồi quy tuyến tính:

Q_{Yaly}= 2,447 Q_{Kontum} + 52,698 với hệ số tương quan R= 0,9385.