MỞ ĐẦU

1. Các luận điểm chung.

Các phương pháp thống kê trong các nghiên cứu thuỷ văn được ứng dụng khi giải nhiều bài toán vì nhiều khi nó là con đường duy nhất để đánh giá định lượng các khía cạnh khác nhau của hiện tượng thuỷ văn. Phát biểu trên xuất phát từ bản chất đa nhân tố của quá trình thuỷ văn. Thực vậy người ta đã biết một cách rộng rãi rằng nhiều hiện tượng thuỷ văn là kết quả tác động của một số lớn các nhân tố, mức độ ảnh hưởng của mỗi trong các nhân tố đó lê sự hình thành của hiện tượng đang xét tính một cách trọn vẹn là điều không thể. Mô tả toán học các hiện tượng tương tự chỉ có thể bằng phưng pháp thống kê. Thí dụ, xét lưu lượng cực đại của nước, giá trị của nó xác định trực tiếp kích thước các thành phần quan trọng của công trình thuỷ. Dòng chảy cực đại được hình thành dưới tác động của các nhân tố khí tượng và đặc điểm của mặt đệm.

Các nhân tố khí tượng bao gồm mưa, lớp phủ tuyết, sự phân bố củ chúng theo diện tích bồn thu nước, cường độ và thời đoạn mưa và cấp nước của lớp phủ tuyết. Cũng ảnh hưởng tới dòng chảy cực đại của sông ngòi là độ ẩm trước đó của lưu vựcmà nó lại được xác định bởi một tổ hợp các yếu tố khí tượng và các điều kiện địa lý tự nhiên khác: mưa, bốc hơi từ bề mặt lưu vực, các tính chất thuỷ lý của lớp thổ nhưỡng và nhiều yếu tố khác. Các nhân tố địa lý tự nhiên bao gồm kích thước và dạng bồn thu nước, cấu trúc mạng lưới thuỷ văn, độ dốc sông ngòi và lưu vực, điều kiện địa chất và thuỷ đại chất của bồn thu nước, sự có mặt của điền trũng, ao hồ, đầm lầy, rừng, hồ chứa và v.v... Làm sáng tỏ các quy luật đặc trưng cho hiện tượng được hình thành như hệ quả của các mối quan hệ đa nhân tố chỉ có thể bằng phương pháp thống kê.

Áp dụng các phương pháp thống kê trong thuỷ văn có một vài đặc điểm chi phối đặc thù của hiện tượng đang xét trong thuỷ văn.

Đặc điểm thứ nhất là trong hành trangcủa nhà thuỷ văn thương có ít thông tin mà nó thường không thể tăng lên được nữa. Khi đó quan trọng nhất là vấn đề ước lượng thống kê các tham số lựa chọn của phân phối để tăng nhân tạo lượng thông tin (dẫn các dãy thuỷ văn ngắn về thời đoạn nhiều năm), lựa chọn mô hình toán tương đối phù hợp thở mãn tốt nhất số liệu thực nghiệm. Thực vậy, thường không biết trước được hàm phân bố nào sẽ mô tả đặc trưng thuỷ văn này hay kia. Khi đó mọi thông tin bổ sungvề dạng đường cong phân bố, ngoài số liệu quan trắc , tất nhiên là ngắn, đều chưa có. Nên sự lựa chọn đường cong phân bố thường được thực hiện xuất phát từ một vài quan niệm chung, thí dụ về các điều kiện biên cần thoả mãn sơ đồ được tiếp nhận. Mức độ tương ứng của tài liệu thực nghiệm với đường cong phân bố được lựa chọn sử dụng (đường đảm bảo) sau đó được kiểm tra bằng cách so sánh đường cong phân bố lý thuyết với thực nghiệm.

Trong nhiều trường hợp số liệu quan trắc về dòng chảy thường trùng lặp với một số đường phân bố giải tích. Trong những trừng hợp như vậy lựa chọn đường cong phân bố này hoặc khác trở thành một nhiệm vụ không xác định tất nhiên dẫn đến nhiều kết quả tính toán khác nhau.

Sau khi xác định qui luật phân bố mà nó mô tả hiện tượng thuỷ văn ta quan tâm, xuất hiện nhiệm vụ đánh giá các tham số phân bố tổng hợp theo tập mẫu và nó đến lượt lại được thực hiện với một mức độ chính xác nào đó phụ thuộc vào dạng đường cong phân bố và lượng thông tin khi thực hiện tính toán các tham số lựa chọn của phân bố. Do vậy đánh giá lựa chọn các tham số của phân bố được thực hiện thường xuyên với sai số này hoặc kia, xác định nó trong bất kỳ tính toán thuỷ văn nào là nhiệm vụ quan trọng bậc nhất. Bài toán này thường bị phức tạp hoá bởi sự hiện diện của sự bất đối xứng trong chuỗi thuỷ văn và mối quan hệ nội tại trong dãy. Đối với các trường hợp đó các phép giải tích của lý thuyết ước lượng tập mẫu tất nhiên là chưa có. Lời giải gần đúng các vấn đề đó trong nhiều trường hợp có thể nhận được trên cơ sở phương pháp Monte-Carlo - phương pháp thực nghiệm thống kê.[1]

Đặc điểm thứ hai của việc áp dụng các phương pháp thống kê trong thuỷ văn là ở chỗ dãy quan trắc về dòng chảy sông ngòi trong một số trường hợp là không đồng nhất cả thời gian lẫn không gian. Điều này làm phức tạp hơn việc mô tả thống kê tập hợp các đại lượng thuỷ văn. Cho nên, trước khi tính toán thống kê thường cần phải chọn lọc một cách kỹ lưỡng thông tin ban đầu từ quan điểm đồng nhất về mặt vật lý và thống kê. Không tính đến điều này có thể dẫn tới các kết luận không chính xác. Để minh hoạ điều đó có ví dụ sau đây. Giả sử xét dòng chảy cực đại của sông ngòi, trên đó trong một số năm xác định đã xây dựng hồ chứa để thực hiện điều tiết mùa dòng chảy sông ngòi. Trong trường hợp đó hoàn toàn tất nhiên là phân bố dòng chảy cực đại trước và sau khi xây dựng hồ chứa sẽ khác nhau và trộn hai phân bố vào một nhóm là không thể được. Thường rất khó xác định trước nguyên nhân phá vỡ trạng thái đồng nhất của chuỗi quan trắc. Trong những trường hợp như vậy đặc biệt cần thiết phải tính tới việc

[1] Lần đầu tiên phương pháp Monte - Carlo được trình bày bởi các nhà toán học Mỹ Dj. Neyman và S. Ulam. Ngày nay phương pháp này thường được gọi là phương pháp thực nghiệm thống kê .

sử dụng các tiêu chuẩn thống kê đồng nhất với việc phân tích vật lý kỹ lưỡng chuỗi quan trắc đang nghiên cứu.

Đặc điểm thứ ba của việc ứng dụng các phương pháp thống kê trong thuỷ văn liên quan tới sự có mặt của quan hệ nội tại các thành phần trong chuỗi, nó phá vỡ tính ngẫu nhiên của mẫu, kết quả là lượng thông tin độc lập giảm , tính bất ổn định của ước lượng thống kê tăng đồng thời thay đổi cấu trúc của chuỗi thuỷ văn. Những vấn đề này càng có ý nghĩa đặc biệt quan trọng khi điều tiết dòng chảy sông ngòi vì tính chất nhóm các năm ít và nhiều nước phần nhiều được xác định bởi quan hệ nội tại của chuỗi.

Các đặc điểm đã nêu của việc mô tả thống kê hiện tượng thuỷ văn được phản ánh trong các phần tương ứng của cuốn sách này.

Ngoài các luận điểm có tính nguyên tắc chung đã nêu, trong cuốn sách còn xét tới các thủ thuật cụ thể sử dụng đường cong phân bố và lưới xác suất áp dụng trong thuỷ văn , các phương pháp kéo dài chuỗi quan trắc ngắn về thời kỳ nhiều năm, phương pháp phân tích tính đồng nhất và quan hệ ngẫu nhiên của chuỗi thuỷ văn với việc sử dụng các khái niệm cuả lý thuyết hàm ngẫu nhiên. Xét đến cả phương pháp thực nghiệm thống kê (phương pháp Monte - Carlo) ứng dụng giải một vài bài toán thuỷ văn.

Giải quyết nhiều bài toán thuỷ văn thống kê sẽ không thực hiện được nếu không sử dụng máy tính điện tử.

Thực vậy, khó thể tưởng tượng nếu dẫn một chuỗi ngắn về thời kỳ nhiều nămvới việc sử dụng vài tương tự trên cơ sở toán học của phương pháp tuyến tính bôi mà không sử dụng máy tính điện tử.

Việc sử dụng rộng rãi phương pháp thực nghiệm thống kê khi phân tích nhóm các năm nhiều nước và ít nước, sử dụng nhiều phương pháp lý thuyết hàm ngẫu nhiên để mô tả như dao động dòng chảy nhiều năm của sông ngòi (tính toán hàm tự tương quan và tương quan quan hệ, tính hàm phổ và phổ quan hệ. tính toán đồng phân và sai phân của các pha dao động tuần hoàn) sẽ mất ý nghĩa nếu thiếu maý tính điện tử.

Việc tự động hoá tổng hợp các hệ thống lựa chọn, kiểm tra, xử lý, bảo tồn và khái quát thông tin thuỷ văn được thực hiện ngày nay tại Tổng cục KTTV đồi hỏi việc áp dụng rộng rãi các phương pháp thống kê cũng như các phương tiện hiện đại của kỹ thuật tính toán - máy tính điện tử. Tuy nhiên diều đó không phải là ưu thế chủ yếu của tự động hoá tổng hợp đo đạc thuỷ văn.

Thiết lập quỹ dữ liệu thuỷ văn trên các phương tiện kỹ thuật mang thông tin mở ra những khả năng to lớn giải quyết các bài toán thuỷ văn khác nhau theo một lãnh thổ rộng lớn, có thể là cả lãnh thổ Liên bang Xô viết, trên cơ sở sử dụng máy tính và các phương pháp thống kê hiện đại. Có thể tin rằng việc kết hợp các máy tính có tốc độ cao với các phương pháp phân tích thống kê hiện đại dẫn tới các sơ đồ tính toán và dự báo dòng chảy sông ngòi chất lượng cao.

Khi trình bày nhiều chương, cuốn sách sử dụng rộng rãi các kết quả tính toán thực hiện trên máy tính. Tuy nhiên, trình bày có hệ thống cơ sở áp dụng máy tính trong các nghiên cứu thuỷ văn còn thiếu vì nó nằm ngoài khuôn khổ nội dung cuốn sách này.

Hiện nay có rất nhiều tài liệu phổ biến theo lý thuyết xác suất và toán học thống kê, trong đó xem xét một cách khá trình tự cơ sở toán học của các thuật toán sử dụng khi giải các bài toán thuỷ văn nêu trên. Tuy nhiên khi sử dụng các phép toán đã được xử lý rộng rãi của lý thuyết xác suất trong các nghiên cứu và tính toán thuỷ văn khả năng áp dụng nó còn xa mới trọn vẹn, đôi khi thậm chí còn chưa chuẩn xác. Trong các trường hợp này việc làm sáng tỏ các đặc điểm xuất hiện khi áp dụng lý thuyết xác suất vào trong thuỷ văn và việc hình thành các thủ thuật phân tích thống kê trong thực tiễn có ý nghĩa quan trọng.

Tiến tới mục đích đó và để khai thác tốt hơn các tài liệu trong cuốn sách dẫn ra nhiều thủ thuật thu được từ hoạt động khoa học và thực tế hoặc được thành lập theo các tài liệu quan trắc. Tất nhiên, trong các thủ thuật này hoàn toàn chưa mở ra hết bản chất của các vấn đề xem xét, nó chỉ minh hoạ cho các tài liệu đang trình bày.

Các vấn đề lý thuyết thống kê toán học không được trình bày chi tiết mà chỉ sử dụng các kết quả cần thiết cho áp dụng thực tiễn. Để khai thác sâu hơn khía cạnh toán học của vấn đề đang xét cần tham khảo thêm các cuốn sách phổ cập khác. Trong cuốn sách chỉ trình bày các phương pháp thống kê thường hay sử dụng nhất trong thuỷ văn và các phương pháp (theo ý các tác giả) thường xuyên sử dụng nhất trong tính toán và dự báo thuỷ văn.

2. Một vài nét ngắn gọn về sự phát triển phân tích thống kê tài liệu thuỷ văn

Sử dụng các thuật toán xử lý thống kê tài liệu quan trắc thuỷ văn liên quan tới việc hoàn thành việc khái quát đầu tiên, có nghĩa là về mặt lịch sử tương ứng tới giai đoạn đầu tiên của phát triển thuỷ văn học. Khi đó để đặc trưng các đại lượng thuỷ văn chỉ có các tham số cơ bản nhất của chuỗi thống kê : giá trị trung bình, độ lệch quân phương và các ma trận khác nhau. Trong giai đoạn này, dễ thấy mô tả thống kê đầy đủ nhất là đường cong đảm bảo trạng thái mực nước (lưu lượng nước) trong năm. Người ta cũng đã sử dụng một ít phân tích tương quan.

Khởi đầu cho việc sử dụng rộng rãi các phép toán xác suất và thống kê toán học liên quan tới sự xuất hiện công trình của A. Hazen[152-153], lần đầu tiên sử dụng lý thuyết xác suất để nghiên cứu các qui luật thống kê dao động nhiều năm của dòng chảy sông ngòi.

A. Hazen tiếp nhận đường cong Gauxơ để mô tả phân bố thống kê chuỗi dòng chảy sông ngòi có tính chất đối xứng, chạy từ $-\infty$ đến ∞ và được đặc trưng bởi hai tham số: giá trị trung bình của đại lượng biến đổi và độ lệch quân phương của nó (hoặc hệ số biến đổi). Để xác định suất đảm bảo thực nghiệm Hazen sử dụng công thức

$$P = \frac{m - 0,5}{n},$$

với n - số thành viên của chuỗi; m- số thứ tự của thành viên chuỗi phân bố theo trật tự giảm (hoặc tăng) dần.

Các công trình của Hazen đã đặt nền móng cho việc xây dựng các lưới xác suất, cho phép làm thẳng các đường cong đảm bảovà đễ dàng cho việc ngoại suy. A. Hazen dựng lưới trên đó làm thẳng hoàn toàn đường cong phân bố chuẩn (đường cong Gauxơ).

Giai đoạn quan trọng tiếp theo trong việc sử dụng các thủ thuật thống kê trong thuỷ văn là các công trình của A. Phoster [149-151] và Đ. L. Xocolovski [131-132].

A. Phoster xác định rằng chuỗi dòng chảy thường không đối xứng và vì thế giới thiệu áp dụng cho việc xây dựng đường cong đảm bảo dòng chảy đường cong bất đối xứng Piêcson III. Ngoài ra, đường cong này với các giá trị xác định của tham số không mang giá trị âm, hơn hẳn so với phân bố chuẩn về tính tương ứng với bản chất hiện tượng đang xét.

Đối với khả năng sử dụng thực tiễn rộng rãi đường cong Piecson III, Phoster thiết lập bảng giá trị hàm cho phép theo các tham số cơ bản xác định bởi nó (giá trị trung bình, hệ số biến đổi và hệ số bất đối xứng) dựng mọi đường cong. Bảng Phoster được S. I. Rưpkin[117] hiệu đính và được sử dụng tốt trong tính toán thuỷ văn ở Liên Xô. Tiếp theo bảng này được mở rộng bởi GGI đối với các giá trị cao hơn của hệ số bất đối xứng (tới $C_s = 5,2$).

Việc các nhà thuỷ văn sử dụng rộng rãi các phép toán lý thuyết xác suất và thống kê toán học ở Liên Xô bắt đầu từ lúc xuất hiện công trình của Đ.L. Xocolovski [132], trong đó trình bày sơ đồ tính toán Phoster với đường cong Piecson III. Đồng thời Xocolovski còn đưa ra một thành phần hoàn toàn mới trong cấu trúc của Phoster, chỉ ra cách xác định đại lượng hệ số biến đổi theo công thức thực nghiệm đối với sông ngòi không có số liệu đo đạc thuỷ văn trực tiếp. Vào thời điểm xuất hiện công trình của Xocolovski cũng đã có đề xuất của Cotrerin để xác định chuẩn dòng chảy của sông ngòi chưa được nghiên cứu.

Như vậy, xuất hiện khả năng dựng đường cong đảm bảo của dòng chảy thậm chí đối với sông ngòi hoàn toàn chưa nghiên cứu thuỷ văn. Đối với việc đó chỉ cần nhận một vài tỷ lệ tiêu chuẩn giữa các đại lượng của hệ số biến đổi (C_v) và hệ số bất đối xứng (C_s). Tính cần thiết của cách giải như vậy được xác định bởi tình huống là đại lượng hệ số bất đối xứng (C_s) theo chuỗi dòng chảy ngắn đang có được xác định rất không chính xác. Áp dụng với việc tính toán đại lượng dòng chảy năm có suất đảm bảo khác nhau tỷ lệ này được đề xuất bằng hai lần ($C_s = 2C_v$), và tương ứng với giới hạn dưới của đại lượng ngẫu nhiên đang xét.

Tiếp về sau, Xocolovski [131] phổ biến nghiên cứu tính ứng dụng của đường cong Piecson III để tính toán lưu lượng cực đại suất đảm bảo khác nhau.

Lúc đầu việc áp dụng rộng rãi đường cong Piecson III đã có ý đến mong muốn loại bỏ nhược điểm của nó là nó nhận giá trị âm với các giá trị suất đảm bảo lớn khi mà hệ số bất đối xứng của chuỗi ngẫu nhiên nhỏ hơn hai lần giá trị hệ số biến đổi ($C_s<2C_v$). Tính chất nêu trên của đường cong đang xét dẫn tới nhận giá trị âm của dòng chảy (hoặc là một đại lượng dương cực lớn, đối với việc mô tả chuỗi thống kê bởi đường cong Piecson III) khi ngoại suy phần thấp của đường cong đảm bảo ngoài giới hạn quan trắc.

Thử nghiệm đầu tiên theo hướng này do G. N. Brocovits [26-29]. Hàm phân bố xác suất các đại lượng thay đổi trong khoảng $0<x<\infty$, ông thể hiện dưới dạng khai triển theo đa thức (nhiều thành viên) Lager. Thành phần đầu tiên của khai triển trùng với biểu thức của đường cong Piecson III với $Cs = 2Cv$ là mô hình xuất phát để biến đổi tiếp theo bằng cách nhập các thành phần tiếp theo của khai triển.

Đề nghị của Brocovits áp dụng với ba tham số (giá trị trung bình, hệ số biến đổi, hệ số bất đối xứng) được xem xét bởi Velicanov[33], ông thực hiện biến đổi đường cong Piecson III thành biểu thức tổng quan hơn bằng cách nhân tung độ của đường cong xuất phát với một thừa số nhiễu có dạng nhiều thành phần Ao + A_1x + A_2x^2 + A_3x^3 + ... Tuy nhiên lời giải mà Velicanov thu được, như đã được G.A. Alecxayev chứng minh, không hoàn toàn loại trừ được nhược điểm đã nêu của đường cong Piecson III .

Cách do Brocovits và Velicanov khởi xướng được E.Đ. Xapharov chọn [119]. Xuất phát từ biểu thức chung của đường cong phân bố xác suất, xuất hiện khi khai triển một hàm bất kỳ (thay đổi trong khoảng 0,∞) theo đa thức Lager, Xapharov đi đến phương trình đường cong phân bố xác suất trùng với Velicanov khi biến đổi đường cong Piecson III bằng phương pháp biến đổi bằng thừa số nhiễu . Khi ứng dụng phương trình này Xapharov lập một bảng chuẩn để dựng đường cong đảm bảo với C_v thay đổi trong khoảng từ 0,05-1,0 và với các tỷ lệ C_v/C_s khác nhau. Đồng thời ông đề xuất thuật toán đồ giải giải tích xác định hệ số biến đổi (C_v) và hệ số bất đối xứng (C_s) áp dụng cho phân bố xác suất đang nghiên cứu.

Do các biến đổi đã nêu không loại trừ được nhược điểm cơ bản đã nêu ở trên của đường cong Piecson III , khi Cs ~ 2Cv thì dẫn tới kết quả tính toán không khác mấy đường cong Piecson III nên chúng không nhận được sự ứng dụng rộng rãi.

Nhiệm vụ biến đổi đường cong Piecson III để loại bỏ nhược điểm bản chất của nó là giá trị âm khi Cs < 2Cv được giải quyết bởi S.N. Krixki và M. Ph. Menkel[78], họ thực hiện biến đổi biến ban đầu x (dấu hiệu phân bố) bằng biến thế z theo hệ thức $z=ax^b$, với a và b - tham số phụ thuộc vào đại lượng hệ số biến đổi và hệ số bất đối xứng dãy thực nghiệm của biến x ban đầu.

Việc áp dụng thực tế đường cong Krixki - Menkel được nhận tên gọi là phân bố gamma ba tham số[1] trở nên khả thi sau ấn phẩm của Đ.V. Korenhistov [64] bảng tung độ các đường cong này đối với các giá trị khác nhau của hệ số biến đổi Cv và hệ thức Cv/Cs. Trong các bảng này gồm giá trị hệ số biến đổi Cv từ 0,10 đến1,20. E.G. Blokhinov và N.V. Nhicolskaia [24] đã mở rộng bảng tới Cv=2,0.

Bên cạnh việc biên soạn hướng tới việc loại bỏ nhược điểm đã nêu của đường cong Piecson III, người ta còn nghiên cứu với mục đích là tạo ra các sơ đồ khác mô tả

[1] Tên gọi này tuy nhiên là chưa đầy đủ, vì mô hình xuất phát - đường cong Piecson III là phân bố gamma, biểu diễn ở dạng chung cũng qua ba tham số (x, Cv, Cs). Cho nên hoàn toàn không nhất thiết như Krixki và Menkel đã làm là trói buộc khái niệm đường cong Piecson III với điều kiện Cs = 2Cv. Nói riêng, áp dụng vào nhiệm vụ tính toán dòng chảy trong nhiều trường hợp hoàn toàn hợp lý với Cs ≥ 2Cv .

các qui luật thống kê mang tính chất của chuỗi. ta đã biết những cố gắng thể hiện hàm mật độ xác suất f(x) của đại lượng x thay đổi trong khoảng từ -∞ đến ∞ ở dạng chung hơn so với đường cong phân bố chuẩn[1] (đường cong Gauxơ). Cách giải quyết này đã được M. V. Mialcovski [90] áp dụng và ông sử dụng phương pháp khai triển hàm f(x) về chuỗi Gramm-Sarle trong dạng đa thức (đa thành viên) Ermit. Thành viên đầu tiên của khai triển này trùng với biểu thức của qui luật phân bố chuẩn. Do đó phương pháp này, về bản chất, dẫn đến biến đổi (biến dạng) của qui luật phân bố chuẩn Gauxơ ở dạng phân bố bất đối xứng bằng cách xét thêm các thành viên chuỗi phụ thuộc vào các mômen bậc cao hơn so với các mômen xác định đường cong chuẩn xuất phát.

Phép biến dạng đường cong chuẩn đã nêu với sự trợ giúp của khai triển hàm phân bố đại lượng (x), thay đổi trong khoảng -∞ <x<∞, về chuỗi Gramm-Sarle có thể xem như là chuyển đổi từ hàm phân bố ban đầu (Gauxơ) đến một qui luật phân bố chung hơn (như là phân bố bất đối xứng) bằng cách nhân tung độ mô hình phân bố xuất phát với một hàm f(x) nào đó, gọi là nhiễu. Hàm này thường được biểu thị dưới dạng một liệt đại số.

Tương tự, có thể xem phép biến đổi trên của việc biến đổi phân bố Piecson III với sự trợ giúp của khai triển về chuỗi theo đa thức Lager.

Lời giải với sự sử dụng đường cong phân bố Sarle , Mialcovski đạt đến giai đoạn thuận tiện cho các tính toán thực tiễn. Ông đã thành lập các bảng cho phép xác định tung độ đường cong đảm bảo phụ thuộc vào hệ số bất đối xứng và độ nhọn - chính là các tham số của đường cong này.

Mặc dù vậy, đề xuất này không phổ biến trong thực tiễn tính toán thuỷ văn do việc xây dựng đường cong dựa trên việc ước lượng các đại lượng hệ số bất đối xứng và độ nhọn mà chúng theo chuỗi thực nghiệm được xác định với độ chính xác rất thấp.

Ngoài ra, biến đổi do Mialcovski đề xướng trong một số trường hợp không loại bỏ được khả năng nhận giá trị âm với một vài giá trị nào đó của biến. Như A.M. Basin[16] đã chứng minh, đường cong Sarle không chiếm ưu thế nào và chỉ có tính chất thực nghiệm.

---

[1] Thuật ngữ đường cong "chuẩn" là do Piecson đề xuất và ông nói:" Nhiều năm trước đây tôi gọi là đường cong chuẩn là đường cong Gauxơ - Laplas. Tên gọi này thuận tiện vì để lại gốc quốc tế và không thuận tiện vì coi như các phân bố khác bị hiểu là không chuẩn. Tất nhiên điều này không đúng."[99, tr.21].

Bên cạnh những vấn đề đã nêu, sử dụng đường cong chuẩn để giải các bài toán thuỷ văn gắn liền với biến đổi logarit hoặc là phương trình đường cong chuẩn, hoặc là giá trị dòng chảy ban đầu. Rõ ràng phân bố xác suất chuẩn logarit chỉ có các đại lượng ngẫu nhiên dao động trong miền giá trị dương (thí dụ như lưu lượng nước trong sông ngòi), dologarit không có giá trị âm. Trong trường hợp thứ nhất phương trình qui luật chuẩn được thế biến x bởi logx. Kết quả là thu được một phân bố chuẩn logarit (chuẩn-loga) bất đối xứng bắt đầu từ 0 và không bị chặn trên. Trong trường hợp thứ hai, có nghĩa là sử dụng chuỗi đầu vào không phải là x mà là logx, giá trị dao động của các giá trị hiển nhiên dương $0 \leq x < \infty$ đạt được $-\infty < \lg x < \infty$, làm trơn tính bất đối xứng của chuỗi và sau đó mô tả bởi đường cong phân bố chuẩn.

Hướng gắn với biến đổi logarit dựa trên phân tích toán học được nhà toán học Đan Mạch A. Phiser thực hiện; Sleyd [155] áp dụng cho tính toán dòng chảy sông ngòi . Thông tin về điều này chứa trong bài báo của S.N. Krixki và M. Ph. Menkel [84].

Khả năng sử dụng đường cong logarit chuẩn để mô tả các qui luật dao động thống kê lũ do mưa đã được các nhà bác học Mỹ Berdon và Kumperon nghiên cứu.

Ở Liên Xô vấn đề về khả năng sử dụng đường cong chuẩn để đánh giá độ lặp lại của lũ do mưa trong trường hợp biến đổi đại lượng chuỗi đầu vào thành logarit được E.G. Blokhinov nghiên cứu chi tiết. Trong công trình đó Blokhinov đưa ra đề nghị sử dụng hoàn thiện các đường cong loga-chuẩn. Cần nhận thấy rằng trong lĩnh vực tính toán thuỷ văn đề xuất về biến đổi logarit các đại lượng biến đôỉ thuộc về S. I. Rưpkin[116], người đã sử dụng nó để xây dựng sơ đồ tính toán lưu lượng cực đại của nước với các xác suất an toàn khác nhau.

Kết quả phân tích đó Rưpkin đi đến kết luận về khả năng mô tả qui luật thống kê của logarit các đại lượng cực đại lưu lượng nước nhờ đường cong Piecson III.

Một trong những ưu thế của đề nghị này, theo Rưpkin, là ở chỗ đường cong Piecson III bị chặn trên khi $P \rightarrow 0$. Tuy nhiên tính chất này của đường cong nói trên cũng như các đường cong khác bị chặn trên sẽ rất khó dùng trrong thực tế vì xác định giới hạn trên của đại lượng biến đổi thường gắn với việc phải thực hiện đủ một số động tác bất kỳ khi ngoại suy đường cong đảm bảo hoặc khi sử dụng phương pháp nhóm điểm.

Bên cạnh các sơ đồ phân bố xác suất kể trên các nhà thuỷ văn xô viết còn xét tới một vài cách khác đối với khả năng sử dụng để ước lượng dao động ngẫu nhiên dòng chảy sông ngòi. Thế nên, G. A. Alecxayev[10] đã ủng hộ phân tích chi tiết đường cong Gudrits. Ông xem xét sơ đồ xác suất lý thuyết thoả mãn qui luật phân bố Gudrits và thành lập bảng chuẩn các tung độ chuẩn hoá cho phép dựng đường cong đảm bảo trên cơ sở đánh giá 3 tham số: giá trị trung bình, hệ số biến đổi và hệ số bất đối xứng. G. A. Alecxayev chứng minh rằng đường cong Gudrits, khác với đường cong Piecson III, không âm ngay cả với Cs < 2Cv (thậm chí với giá trị Cs âm) nếu như Cs > 2Cv -0,9. Tuy nhiên do thiếu những ưu thế cơ bản so với đường cong Piecson III và đường cong Krixki-Menkel nên đường cong Gudrits không được phổ biến trong thực tiễn tính toán thuỷ văn để ngoại suy đường cong đảm bảo.

Thời gian gần đây, G.G. Svanhidze và G. L. Grigolia [44,124] đã nghiên cứu khả năng sử dụng phân bố Jonshon bị chặn cả trên lẫn dưới. Để ước lượng tham số phân bố đã cho họ đã sử dụng lần đầu tiên bốn mômen. Giới hạn trên và dưới của phân bố này xác định theo cực tiểu của chỉ tiêu phù hợp χ^2 với các giới hạn phân bố khác nhau. Khi đó xác định ảnh hưởng của các giới hạn lên tham số phân bố $(\bar{x}, Cv, Cs)$ và hệ số tương quan giữa các thành viên trong chuỗi.

Sự tập trung lớn trong các tài liệu thuỷ văn dành cho việc giải thích cơ sở sử dụng các sơ đồ thống kê khác nhau (cụ thể là đường cong Piecson III và phân bố gamma ba tham số) để đánh giá các đặc trưng khí tượng thuỷ văn độ lặp hiếm, có nghĩa là trong vùng ngoại suy.

Bản chất của các nghiên cứu này là ở chỗ không thể chứng minh sự tương ứng của các qui luật phân bố chuỗi dòng chảy sông ngòi bởi sơ đồ thống kê này hoặc kia với các cấu trúc lý thuyết. Kết luận như thế với mức độ tin cậy này hoặc kia có thể được chỉ trên cơ sở phân tích chuỗi thống kê đang có của đại lượng nghiên cứu.

Kiểm tra sự tương ứng của các đường cong thực nghiệm và lý thuyết theo tài liệu quan trắc ở một số tuyến đo khí tượng thuỷ văn được Xocolovski [132] tiến hành khi trình bày phương pháp Phoster. Tuy nhiên, do độ dài thời đoạn quan trắc không lớn ở một số tuyến đo, đặc biệt là vào thời kỳ Xocolovski thực hiện công trình (hoặc là tính hạn chế rõ ràng của mẫu về ý nghĩa thống kê), việc so sánh như vậy không thể được coi là đủ cơ sở tin cậy về tính áp dụng của sơ đồ lý thuyết đường cong phân bố đại lượng ngẫu nhiên trong thuỷ văn.

Cho nên vào năm 1941 G. N. Brocovits và G.N. Velicanov [30] đã cố gắng mở rộng khả năng phương pháp so sánh trực tiếp đường cong đảm bảo thực nghiệm và giải tích bằng cách sử dụng số liệu về lưu lượng nước theo vài tuyến trộn vào một tập hợp để xây dựng đường cong đảm bảo thực nghiệm (phương pháp trạm năm). Khi đó trong một tổ hợp nhập các chuỗi lưu lượng có các giá trị hệ số biến đổi ít khác nhau, từ kết quả phân tích trên Brocovits và Velicanov đi tới kết luận về khả năng sử dụng đường cong Piecson III với Cs = 2Cv để mô tả qui luật dao động ngẫu nhiên của lưu lượng nước (cụ thể là lưu lượng cực đại).

Tiếp theo S.N. Krixki và M. Ph. Menkel [83] đã tiến hành nghiên cứu rộng rãi tính ứng dụng của đường cong Piecson III cũng như đường cong mang tên họ là gamma ba tham số để đánh giá dao động ngẫu nhiên của dòng chảy sông ngòi. Kết quả của sự phân tích này được thực hiện với việc sử dụng phương pháp trạm năm và một vài chỉ tiêu đồng nhất thống kê xác định rằng đường cong Krixki và Menkel là sơ đồ tiện ích để mô tả các qui luật thống kê dao động dòng chảy sông ngòi. Tương tự đối với đường cong Piecson III với Cs ≥ 2Cv.

G. P. Kalinhin [58] đã thử xây dựng mô hình phân bố thống kê mới của dao động ngẫu nhiên dòng chảy năm và cực đại. Các nghiên cứu do ông thực hiện là hoàn thành bảng tung độ đường cong đảm bảo khái quát lưu lượng nước cực đại và trung bình năm. Các bảng này là các trường hợp riêng của các đường cong Piecson III trong giới hạn khoảng biến đổi của hệ số bất đối xứng nhỏ. Do vậy, kết quả nghiên cứu này chỉ có thể coi như thêm một khẳng định (trên tài liệu thực nghiệm) khả năng sử dụng đường cong Piecson III để mô tả dao động ngẫu nhiên của dòng chảy năm và cực đại.

Krixki và Menkel dành sự chú ý nhiều cho vấn đề đánh giá thống kê độ chính xác việc xác định mẫu các tham số đường cong phân bố[78, 79].

Trước khi xuất hiện các công trình của Krixki và Menkel với việc xác định các sai số ngẫu nhiên của các ước lượng mẫu các tham số thống kê của chuỗi các đại lượng thuỷ văn người ta sử dụng các mối quan hệ dùng cho các tập tuân theo qui luật phân bố chuẩn Gauxơ. Krixki và Menkel [78] dựa trên phương pháp mômen và xuất phát từ luật phân bố nhị thức khi Cs = 2Cv nhận được biểu thức sai số ngẫu nhiên xác định độ lệch quân phương (chuẩn), hệ số biến đổi và hệ số bất đối xứng, độ nhọn và tung độ đường cong đảm bảo.

Vào năm 1968, Krixki và Menkel [79] đã công bố các công thức hiệu chỉnh sai số chuẩn ước lượng mẫu hệ số biến đổi và tung độ đường cong đảm bảo Piecson III nhận được có tính đến hệ số tương quan giữa các ước lượng mẫu trung bình và chuẩn (độ lệch quân phương).

Sự phát triển nghiên cứu tiếp theo trong lĩnh vực đánh giá độ chính xác của ước lượng mẫu tham số các đường cong phân bố cũng như giải quyết hàng loạt vấn đề khác liên quan tới lĩnh vực giải thích các qui luật thống kê đặc thù cho chuỗi các đặc trưng thuỷ văn , liên quan tới việc phổ cập vào thực tiễn tính toán thuỷ văn và thuỷ lợi là phương pháp Monte-Carlo (mô hình hoá toán học). Lần đầu tiên cơ sở phương pháp này được trình bày khá đầy đủ trong các công trình của G.G. Svanhide [123].

Dựa trên phương pháp Monte-Carlo, E. G. Blokhinov [18] khi sử dụng khả năng thực nghiệm số trên máy tính điện tử thu được biểu thức đối với sai số ngẫu nhiên với hiệu chỉnh về sự trộn ước lượng mẫu các tham số chỗi thống kê các đặc trưng thuỷ văn.

Trong các công trình G.G Svanhide [120, 129] và Khomerika [141] đưa ra phương án mô hình hoá chuỗi thuỷ văn có tính đến phân phối trong năm của dòng chảy. Nó dựa trên phương pháp chọn hai lần: lượng nước trong năm và đường quá trình quan trắc thực (lát cắt). Khi đó tính đến cả các quan hệ ngẫu nhiên giữa dòng chảy các thời kỳ khác nhau.

Ngày nay, người ta đưa ra khá nhiều các phương pháp mô hình hoá thống kê chuỗi thuỷ văn , từ đó mối quan tâm lớn nhất là phương án dựa trên luật phân bố chuẩn với chuyển đổi tới phân bố đã cho [125]. Con đường này của mô hình hoá thống kê là hữu hiệu hơn cả khi mô hình nhóm các chuỗi thuỷ văn với ma trận cacs hệ số tương quan kép cho trước. Khi đó có thể tính đến cả tương quan trong chuỗi.

Vào năm 1941, Krixki và Menkel [77] đưa ra đề nghị sử dụng để ước lượng các tham số thống kê chuỗi các đại lượng thuỷ văn bằng phương pháp tương tự tối đa, cơ sở toán học của nó được soạn thảo bởi nhà toán học Anh R. Phiser. Khả năng sử dụng phương pháp tương tự tối đa trong tính toán thuỷ văn được Blokhin xem xét kỹ[22]. Các lời giải mà ông nhận được đưa tới khả năng sử dụng thực nghiệm với sơ đồ phân bố gamma ba tham số.

Kết thúc tổng quan ngắn về sử dụng các đường cong phân bố lý thuyết trong thực tế tính toán thuỷ văn nhận thấy rằng mô tả thống kê tương tự các số liệu thuỷ văn xuất phát từ giả thuyết thiếu một qui luật nào đó trong liệt đại lượng ngẫu nhiên được nghiên cứu.

Tuy nhiên, trường phái này là không thật chặt chẽ, vì trong công trình của P.A. Ephimovits đã chứng tỏ sự hiện diện của quan hệ tương quan trong chuỗi tại các tập thống kê đại lượng dòng chảy năm. Sự có mặt quan hệ tự tương quan, cụ thể trong chuỗi dòng chảy năm không phủ nhận khả năng sử dụng đường cong phân bố lý thuyết trong thuỷ văn , nhưng xác định được tính cần thiết xét tới vấn đề này, đặc biệt khi xét các bài toán sau:

1) khi nghiên cứu dao động tuần hoàn dòng chảy sông ngòi, gồm nghiên cứu nhóm các năm nhiều nước và ít nước.

2) khi soạn thảo phương pháp dự báo đặc trưng dòng chảy sông ngòi với hạn dài (1 năm và nhiều hơn) trên cơ sở sử dụng hàm tự tương quan và phương pháp tương quan tuyến tính bội.

3) khi nghiên cứu qui luật dao động theo thời gian và không gian của dòng chảy sông ngòi.

Đặc biệt việc soạn thảo tích cực theo các hướng trên bắt đầu từ hai chục năm gần đây và tiếp diễn tới bây giờ. Điều này liên quan tới sự phát triển lý thuyết hàm ngẫu nhiên và chủ yếu với việc sử dụng rộng rãi máy tính điện tử.

Mô tả toán học dao động nhiều năm dòng chảy sông ngòi dựa trên các giả thuyết cơ bản sâu đây, mà các vấn đề thảo luận khó có thể nói là đã kết thúc được hiện nay.

1. Giả thiết về sự độc lập hoàn toàn của dao động dòng chảy sông ngòi nhiều năm. Áp dụng tới sự nghiên cứu dao động dòng chảy năm và đặc biệt là nhóm các năm ít nước và nhiều nước, giả thuyết này, tất nhiên, là bị phủ nhận, còn trong quan hệ nhóm với các đặc trưng thuỷ văn khác (như lưu lượng nước cực đại và cực tiểu) sự thừa nhận nó vẫn chưa thống nhất.

2. Giả thiết về sự hiện diện quan hệ tương quan tuyến tính gữa thể tích dòng chảy các năm hỗn hợp (xich Markov đơn). Giả thiết này nhận được sự thừa nhận rộng rãi và được sử dụng để đánh giá độ chính xác việc xác định các tham số đường cong phân bố , trong tính toán điều tiết dòng chảy và v.v... Sự hiện diện quan hệ giữa các giá trị năm giữa của các đặc trưng thuỷ văn khác được nghiên cứu ít hơn so với chuỗi dòng chảy năm.

3. Giả thiết tương ứng của dao động nhiều năm của dòng chảy năm của mô hình quá trình ngẫu nhiên dừng với thời đoạn không liên tục. Mô hình này tìm thấy được một vài ứng dụng khi lập phương pháp tính toán và dự báo các đặc trưng khác nhau (chủ yếu là các giá trị cực đại và cực tiểu) của dòng chảy năm. Giả thiết đang xét không có thể coi là đã được chứng minh hay loại bỏ do sự thiếu chuỗi các đại lượng thuỷ văn, thể hiện tập đủ lớn.

4. Giả thiết tương ứng dao động các đại lượng dòng chảy sông ngòi của mô hình toán dạng quá trình ngẫu nhiên không dừng. Nghiên cứu tính ứng dụng của giả thiết này mới chỉ bắt đầu.

5. Giả thiết về tính egodic dao động dòng chảy sông ngòi , dự đoán khả năng thay thế quan trắc theo thời gian (ở một số điểm không gian) bằng đặc trưng thuỷ văn nào đó được quan trắc trong không gian, hoặc ngược lại. Sự thực hiện giả thiết này của các đặc trưng thuỷ văn có nghĩa là khả năng xét tổng hợp chuỗi các đặc trưng thuỷ văn trong giới hạn một vùng nào đó, nơi mà giả thiết đang thực hiện.

Khả năng sử dụng giả thiết này trong thuỷ văn còn chưa được chứng minh.Sự xuất hiện đa dạng về số lượng các mô hình toán mô tả dao động nhiều năm của dòng chảy sông ngòi mức độ nào đó liên quan tới chuỗi quan trắc thậm chí có độ dài lớn nhất vẫn không đủ để khẳng định một cách tin tưởng về sự đúng đắn của việc lựa chọn mô hình này hoặc mô hình kia.

Ngoài ra, trong hàng loạt công trình, tiến hành không chính xác việc ước lượng kết quả nhận được trên cơ sở sử dụng các mô hình toán trong các công trình đó, dẫn đến việc đánh giá cao khả năng của sơ đồ thử nghiệm do vậy dẫn đến việc khó có cơ sở để tuyên truyền chúng.

Các giả thuyết kể trên thuộc về tập các đặc trưng dòng chảy năm của sông ngòi (lưu lượng nước trung bình năm, cực đại và cực tiểu), chúng có thể được diễn toán bằng hàm ngẫu nhiên dừng với mức chính xác như tập các đặc trưng dòng chảy sông ngòi có thể biểu diễn qua một tập rời rạc. Nếu như dao động dòng chảy sông ngòi được xét với quan điểm cao hơn như tính liên tục của phổ dao động theo thời gian, thì trong trường hợp đó các qui luật thống kê của các dao động này có thể thể hiện chỉ ở dạng các quá trình không dừng.

Không đề cập đến ở đây mọi công bố theo vấn đề được nêu, chỉ đề cập đến các công trình ở một mức độ nào đó có dạng cơ bản để thể hiện tốt trong lĩnh vực đang xét.

Các công trình sớm nhất sử dụng các phép lý thuyết hàm ngẫu nhiên dừng đưa đến các phép giải bằng số thuôcj về Iu. M. Alekhin [12-14]. Trong các công trình này chứa đựng một tài liệu rộng rãi tính toán các hàm tự tương quan dao động nhiều năm của dòng chảy sông ngòi với thời đoạn $\tau \leq 30$ năm.

Hàm tự tương quan thu được được tác giả sử dụng cho nền của phương pháp gọi là động lực - ngẫu nhiên dự báo dòng chảy sông ngòi. Bản chất đề nghị này dẫn đến việc ngoại suy tuyến tính đại lượng dòng chảy năm theo biểu thức dạng:

$$Q_{i+1} = k_1Q_i + k_2Q_{i-1} + k_3Q_{i-2} + ... + k_nQ_{i-(n-1).}$$

với Q_i - đại lượng dòng chảy năm; k_i (i=1,2,3...,n) - hệ số ngoại suy xác định theo số liệu thực tế.

Hàm tự tương quan để phan tích các qui luật thống kê dao động tuần hoàn của dòng chảy năm được I. P. Druzjinhin sử dụng. Kết quả nghiên cứu theo ý của Druzjinhin và những người khác [51], chứng tỏ về tính không dừng của dao động thể tích dòng chảy năm. Thừa nhận quan điểm này họ, tất nhiên, giả thiết khả năng sử dụng phép hàm ngẫu nhiên dừng để nghiên cứu qui luật thống kê dao động tuần hoàn dòng chảy sông ngòi gọi các dao động này là qúa trình dừng ....

Khi nghiên cứu các qui luật dao động tuần hoàn dòng chảy năm với việc sử dụng hàm phổ, G. P. Kalinhin và A. I. Đavưđova [59] đi tới kết luận về sự có mặt của chu kỳ thời đoạn khác nhau.

Khi đánh giá khả năng nghiên cứu dao động nhiều năm của dòng chảy bằng các thủ thuật sử dụng trong các công trình của Alekhin, Druzjinhin, Kalinhin, Đavưđova và những người khác, cần thiết đánh giá kỹ lưỡng và đầy đủ các kết luận thu được, nếu thiếu nó việc sử dụng thực tế chúng, như để dự báo đại lượng dòng chảy năm chỉ có tính tạm thời. Chi tiết hơn vấn đề này sẽ được xét khi trình bày nội dung chính của cuốn sách.

Một hướng nghiên cứu khác của vấn đề đang xét dựa trên việc sử dụng sơ đồ xích Markov đơn. Mô hình toán học này được tiếp nhận trong nhiều nghiên cứu qui luật dao động dòng chảy năm của S. N. Krixki và M. Ph. Menkel [67, 69, 78, 82], E.G. Blokhinov [25], Đ. Ia. Ratkovits [95-97], A. S. Reznhicovxki [37,98] G.G. Svanhide [120-124] và những người khác. Cách làm này có tính khách quan lớn của sự phân tích và dựa trên phương tiện toán học được xử lí tốt hơn.

Trong các công trình tiếp theo khi mô tả dao động nhiều năm của dòng chảy sông ngòi người ta sử dụng rộng rãi lý thuyết hàm ngẫu nhiên (lý thuyết qúa trình xác suất). Trong số các nghiên cứu chi tiết của hướng này có chuyên khảo của N. A. Cartvelisvili [61], Đ. I. Kazakevits [56] và G.A. Alexayev [9].

Gần đây, sự tập trung nghiên cứu dao động không gian và thời gian các đặc trưng khác nhau của chế độ thuỷ văn. Các nghiên cứu này dựa trên lý thuyết trường đồng nhất và đẳng hướng. Kết quả phân tích dùng để giải các bài toán nội suy các giá trị đại lượng thuỷ văn theo lãnh thổ để chọn các chỉ tiêu khách quan phân bố mạng lưới đài trạm thuỷ văn, để đánh giá độ chính xác của việc xác định các đại lượng thuỷ văn. Cơ sở thống kê trong các nghiên cứu này được coi là các hàm tương quan không gian. Sự trình bày tốt nhất các vấn đề này có trong chuyên khảo của G.A. Alecxayev [9].