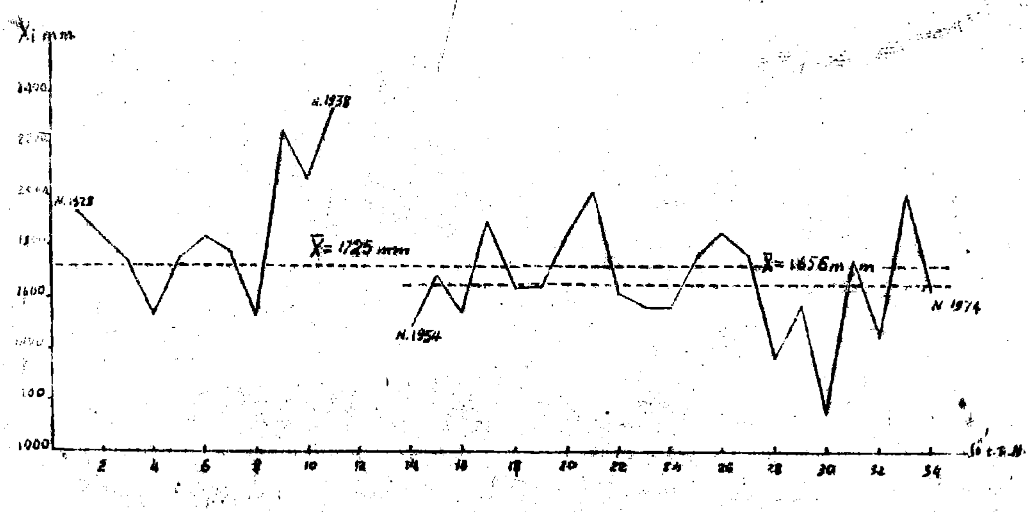
Xác định chu kỳ mưa năm
BẰNG PHƯƠNG PHÁP THỐNG KÊ HIỆN ĐẠI

Hà Văn Hiền
(Viện KTTV)

I. Đặt vấn đề .

Nghiên cứu chu kỳ của mưa trên toàn lãnh thổ có một vai trò rất quan trọng, bởi vì nó giúp cho chúng ta thấy được qui luật diễn biến của nguồn nước, thông qua đó đi đến sử dụng nước một cách hợp lý phục vụ các ngành kinh tế quốc dân trong từng giai đoạn cụ thể. Đặc biệt trong công tác phòng chống bão lụt và hạn hán việc nắm vững tính chu kỳ của mưa càng là vấn đề cấp thiết và có ý nghĩa thực tiễn to lớn.

Trong bài này xin giới thiệu một số nét cơ bản của việc nghiên cứu chu kỳ mưa năm trạm Buôn mê thuột có số liệu liên tục 21 năm (1954-1974) bằng phương pháp hàm tương quan và hàm mật độ phổ, đồng thời nêu lên sự vận dụng các phương pháp thống kê trong việc xác định chu kỳ của mưa và đưa ra các thử nghiệm thống kê để xác định mức ý nghĩa của vấn đề được xem xét (hình 1).



Hình 1. Đường qúa trình lượng mưa năm trạm Buôn mê thuột.

II. Các phương pháp thống kê sử dụng trong việc xác định chu kỳ của mưa năm

1. Cơ sở sử dụng các phương pháp thống kê.

Trong qúa trình nghiên cứu các yếu tố khí tượng thủy văn, phương pháp thống kê được áp dụng rất rộng rãi, điều đó không những nói lên tính ưu việt của phương pháp này, sự thích hợp của nó đối với công việc xác định và đánh giá các đặc trưng, mà còn phản ánh được sự phù hợp giữa tính chất của lượng thông tin với các điều kiện để áp dụng các phương pháp đó. G.X.Kavediax và L.M.Brennan /5/ đã chỉ ra rằng : sự tồn tại tính chất ngẫu nhiên của chuỗi số liệu là điều kiện rất quan trọng trong việc áp dụng các phương pháp thống kê. X.Xôlômôn /5/ có kết luận : trong các biến số thủy văn thì mưa là biến số độc nhất có dạng phân phối gần phân phối chuẩn hơn cả. Điều này nói lên rằng tính ngẫu nhiên của mưa cũng cao hơn tính ngẫu nhiên của các biến số thủy văn khác.

Như vậy điều kiện để áp dụng các phương pháp thống kê trong việc xác định chu kỳ của mưa năm được thỏa mãn. Đặc biệt khi áp dụng phương pháp hàm mật độ phổ thì chuỗi số liệu còn được coi là ngẫu nhiên dừng.

2. Xác định chu kỳ của mưa năm bằng phương pháp hàm tương quan và mật độ phổ.

a- Phương pháp hàm tương quan R (τ)

Hàm tương quan R (τ) hay r (τ) của mưa năm được xác định theo công thức :

$$r(\tau) = \frac{\sum \Delta_x(t) . \Delta_x(t+\tau)}{\sqrt{\sum \Delta_x^2(t) . \sum \Delta_x^2(t+\tau)}}$$

ở đây : $\Delta_x(t) = x(t) - \bar{x}(t)$

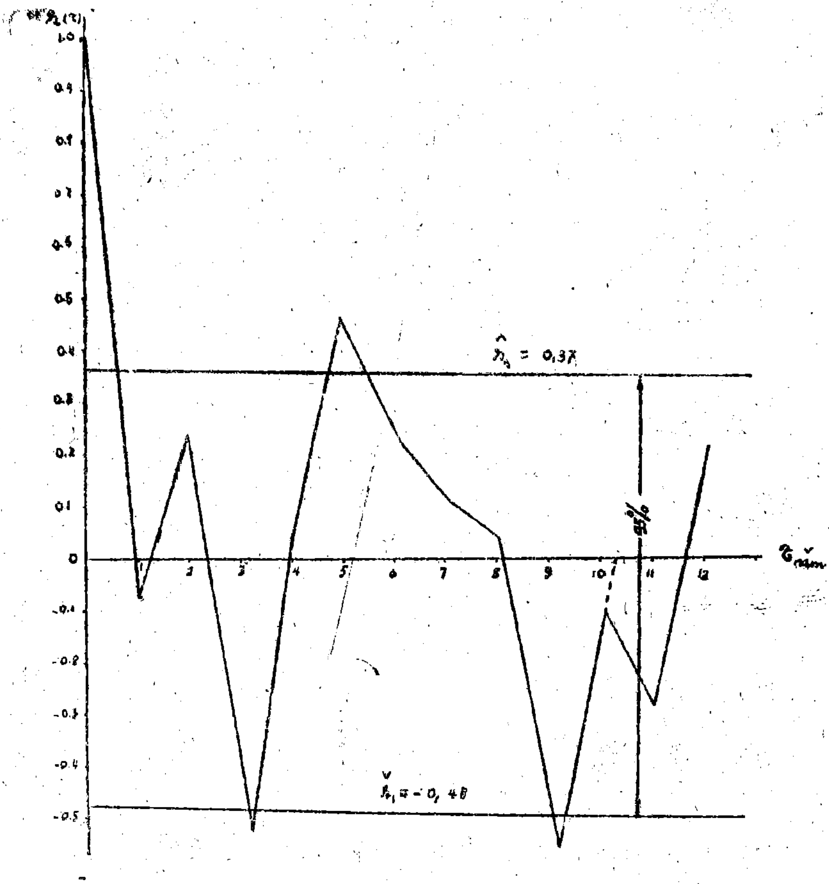
$\quad {}_x(t + \ ) = x(t + \ ) - \bar{x}(.t \ )$

τ - thời gian trượt ở đây cụ thể : τ = 1; 2; 3; ... ,12 năm.

Kết qủa tính r(τ) được trình bày trong bảng 1, và được biểu thị trên hình 2.

Bảng 1. Hệ số tương quan mưa năm trạm Buôn mê thuột

τ	1	2	3	4	5	6	7	8	9	10	11	12
$r(\tau)$	-0,06	0,23	-0,50	0,04	0,46	0,22	0,12	0,05	-0,54	-0,14	-0,34	0,22



Hình 2. Hàm tương quan lượng mưa năm trạm Buôn mê thuột.

Bằng việc xây dựng hàm tương quan r (τ) có thể rút ra chu kỳ của lượng mưa năm tại Buôn mê thuột có các giá trị sau:

1. Chu kỳ 3 năm
2. Chu kỳ 6 năm

b. Phương pháp hàm mật độ phổ G (f)

Phương pháp hàm mật độ phổ là phương pháp thống kê hiện đại. Nó là công cụ phân tích chuỗi số liệu và cho ta kết qủa chính xác đánh giá các hiện tượng đang nghiên cứu. Trường hợp xác định chu kỳ mưa năm đưa ra ở đây, cho ta thấy những giá trị chu kỳ mưa năm vừa được xác địnhbằng phương pháp hàm tương quan có thể chấp nhận được hay không.

Công thức tính hàm mật độ phổ :

$$\tilde{G}_x(f) = \frac{2}{m}\left[\tilde{R}_o + 2\sum_{r-1}^{m-1} R_r \mathrm{Cos}\,\frac{\pi.r.f}{f_c} + \tilde{R}_m \mathrm{Cos}\,\frac{\pi.mf}{f_c}\right.$$

Ở đây : $\tilde{R}_i$ - ước lượng của hàm tương quan

m - số bước trượt (ở đây $m = 12$); $r = 1 \div m - 1$

f - tần số và $f = \frac{K.f_c}{m}$ khi $K = m$ thì $f = f_c$

f_c - tần số tới hạn $f_c = \frac{1}{2}\Delta t$

Δt - có giá trị cụ thể ở đây là 1 năm.

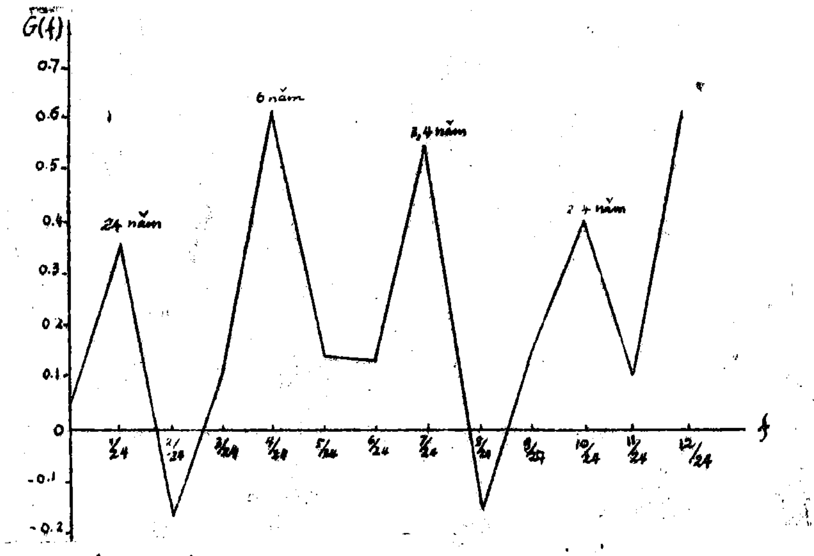
Kết qủa tính hàm mật độ phổ được trình bày trong bảng 2 và được biểu thị trên hình 3.

Bảng 2. Hàm mật độ phổ trạm mưa Buôn mê thuột.

K	0	1	2	3	4	5	6	7	8	9	10	11	12
f	0	1/24	2/24	3/24	4/24	5/24	6/24	7/24	8/24	9/24	10/24	11/24	12/24
G(f)	0,05	0,364	-0,153	0,110	0,623	0,146	0,130	0,564	-0,139	0,15[illegible]	0,413	0,106	0,623

Kết qủa phân tích hàm mật độ phổ thấy xuất hiện các chu kỳ sau :

1. Chu kỳ T = 2,4 năm
2. Chu kỳ T = 3,4 năm
3. Chu kỳ T = 6 năm
4. Chu kỳ T = 24 năm



Hàm 3. Hàm mật độ phổ lượng mưa năm trạm Buôn mê thuộc trước khi lọc.

Tuy vậy bằng kết quả phân tích và qua đồ thị của hàm mật độ phổ xuất hiện những giá trị âm không thể hiện được ý nghĩa vật lý của vấn đề đang xét, do vậy để có thể nhận được kết quả chính xác và hợp lý ta phải sử dụng hàm phổ có lọc.

$$\tilde{G}_K = \frac{2}{m} \left(\tilde{R}_o + 2 \sum_{r=1}^{m-1} D'_r \tilde{R}_r \cos \frac{\pi r K}{m} \right)$$

Ở đây D'_r là công thức Parden [3] có dạng :

$D'_r = 1 - 6\,(r/m)^2 + 6\,(r/m)^3$ với $r = 0,1 \ldots m/2$

$D'_r = 2\,[1-(r/m)]^3$ với $r = m/2 + 1, \ldots, m$

$D'_r = 0$ $\quad r = m$

Các giá trị D'_r được trình bày trong bảng 3.

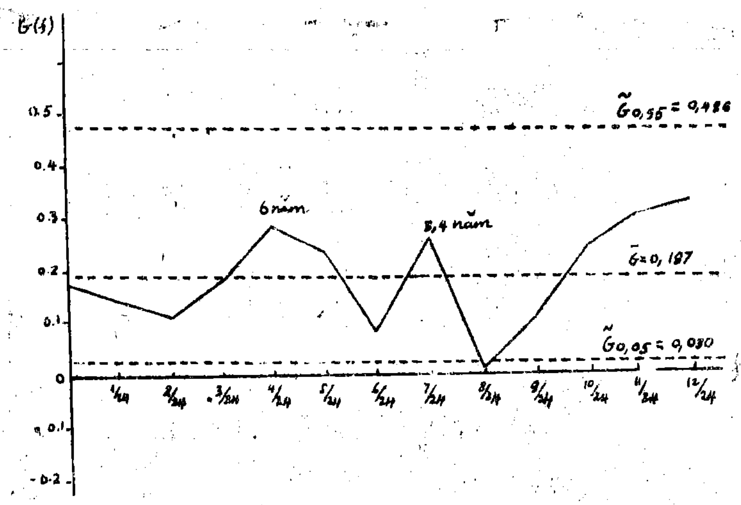
Bảng 3. Giá trị D'_r theo công thức Parden

r	1	2	3	4	5	6	7	8	9	10	11	12
D'_r	0,96	0,86	0,72	0,56	0,39	0,25	0,14	0,07	0,03	0,01	0,001	0

Kết qủa tính hàm mật độ phổ có sử dụng hàm lọc được trình bày trong bảng 4 và được biểu thị trên hình 4.

Bảng 4. Hàm mật độ phổ trạm mưa Buôn mê thuột sau khi lọc.

K	0	1	2	3	4	5	6	7	8	9	10	11	12
f	0	1/24	2/24	3/24	4/24	5/24	6/24	7/24	8/24	9/24	10/24	11/24	12/24
G(f)	0,18	0,142	0,103	0,190	0,296	0,242	0,092	0,258	0,023	0,131	0,250	0,313	0,338



Hình 4 . Hàm mật độ phổ lượng mưa năm trạm Buôn mê thuột sau khi lọc.

Kết qủa phân tích ở đây chothấy xuất hiện các chu kỳ :

1. Chu kỳ T = 3,4 năm
2. Chu kỳ T = 6 năm

Kết qủa này cũng xác nhận rằng các giá trị chu kỳ ở đây cũng mang đặc tính vững chắc của nó sau khi lọc, đồng thời chứng minh rằng những chu kỳ được xác định bằng phương pháp hàm tương quan ở trên là phù hợp với kết qủa của phương pháp hàm mật độ phổ.

III . <u>Thử nghiệm mức ý nghĩa của các phương pháp hàm tương quan và hàm mật độ phổ.</u>

Để đánh giá chính xác hơn các kết qủa chu kỳ của lượng mưa năm đã xét ở trên, chúng ta tiến hành thử nghiệm mức ýnghĩa của các hệ số tương quan và các giá trị của hàm mật độ phổ.

1. <u>Thử nghiệm mức ý nghĩa hàm tương quan r (τ).</u>

Thử nghiệm mức ý nghĩa cho các hệ số tương quan r_x (τ) được đưa ra bởi Anđécxơn [3] dựa trên họ chu kỳ ngẫu nhiên chuẩn có độ dài TN. Mặc dầu giả thuyết về tính chu kỳ ở đây có tính nhân đạo, thử nghiệm thống kê này là một trong số ít thử nghiệm hiện có để kiểm tra mức ý nghĩa của biểu đồ tương quan.

Thử nghiệm này có giới hạn tin cậy cho hệ số tương quan r_1 như sau :

$$G.H\ (r_1) = \frac{-1 \pm z_\alpha\ (N-2)^{1/2}}{N-1}$$

Ở đây : r_1 - hệ số tương quan có bước trượt đầu tiên τ = 1 năm.
z_α - biến số chuẩn ứng với mức ý nghĩa 5% và bằng 1,96.

Các giới hạn này được xác định trên hình 2.

Thực hiện việc kiểm tra này như sau : Hệ số tương quan r_1 nằm trong miền giới hạn tin cậy của nó điều đó khẳng định giả thuyết không về r = 0 là đúngvới mức 95% nghĩa là số liệu của chuỗi mà ta xét không tương quan với nhau. Như vậy độ dài chu kỳ mà ta vừa kết luận ở trên có thể được coi là chuẩn với mức 95%.

2. <u>Thử nghiệm mức ý nghĩa của hàm mật độ phổ G (f).</u>

Thử nghiệm của hàm mật độ phổ nhằm mục đích xác nhận tính chuẩn của mẫu số liệu đã chọn để tính chu kỳ, do đó ở đây cần phải xác định được giới hạn tin cậy của phổ tập hợp để từ đó có thể cho phép ta so sánh với phổ tính toán theo mẫu số liệu đã chọn.

Nếu như giả thiết rằng : chuỗi số liệu được tính là hoàn toàn ngẫu nhiên được lấy ra từ một tập hợp chuẩn thì những ước lượng phổ của nó phải phân phối theo phổ tập hợp tuân theo phân phối có dạng χ^2/ν . Ở đây : ν - số bậc tự do và bằng : $\frac{2N}{m}$, N là khối lượng mẫu chọn (cụ thể N = 21 năm); m là số bước trượt lớn nhất (m = 12).

Phổ tập hợp có thể là phổ trắng (đường thẳng nằm ngang) hoặc phổ Mác-cốp (phổ đỏ). Việc xác định phổ tập hợp thuộc loại nào lại phụ thuộc vào mức ý nghĩa thống kê của hệ số tương quan r_1.

Do ở trên ta đã có kết luận rằng $r_1 = 0$, do vậy chuỗi số liệu được coi là hoàn toàn ngẫu nhiên.(Mặc dầu rằng điều kiện $r_1 = 0$ mới chỉ là điều kiện cần thiết)

Xuất phát từ tính chất của chuỗi số liệu như vậy, phổ tập hợp ở đây phải là phổ trắng.

Ước lượng của nó được tính :

$$\tilde{S} = \frac{\sum_{0}^{12} \tilde{G}_i}{13} = 0,197$$

Để xác định giới hạn tin cậy của phổ tập hợp phải sử dụng phân phối χ^2.

$$\text{Do } \nu = \frac{2N}{m} = \frac{2.21}{12} = 3,5$$

Tra bảng phân phối χ^2 có $p = 0,95$ và $\nu = 3,5$:

$$\chi^2 (p = 0,95;\ \nu = 3,5) = 8,66 \Rightarrow \frac{\chi^2}{\nu} = \frac{8,66}{3,5} = 2,47$$

$$\chi^2 (p = 0,05;\ \nu = 3,5) = 0,532 \Rightarrow \frac{\chi^2}{\nu} = \frac{0,532}{3,5} = 0,152$$

Biên trên của phổ tập hợp :

$$\frac{\chi^2_{0,95}}{\nu} = \frac{\tilde{G}_{0,95}}{\tilde{S}} \Longrightarrow \tilde{G}_{0,95} = \frac{\tilde{S}.\chi^2_{0,95}}{\nu} = 0,197.2,47$$

$$\Longrightarrow \tilde{G}_{0,95} = 0,486$$

Biên dưới của phổ tập hợp :

$$\frac{\chi^2_{0,05}}{\nu} = \frac{\tilde{G}_{0,05}}{\tilde{S}} \Longrightarrow \tilde{G}_{0,05} = \frac{\tilde{S}.\chi^2_{0,05}}{\nu} = 0,197.0,15 = 0,030$$

Biên trên và biên dưới của phổ tập hợp được trình bày trên hình 4.

Như vậy ta thấy rằng : Mọi giá trị của phổ tính toán đều nằm trong miền tin cậy của phổ tập hợp, nghĩa là : sự phân phối của phổ tính toán tuân theo dạng phân phối của phổ tập hợp dạng χ^2/ν. Như vậy phổ tập hợp mà ta chọn ở trên là đúng.

Cũng từ đây ta có kết luận rằng : Mẫu số liệu được chọn để tính chu kỳ của trạm Buôn mê thuột mặc dầu còn ngắn, dao động còn mạnh nhưng nếu sử dụng hàm lọc thì vẫn cho ta kết qủa tin cậy như một mẫu số liệu hoàn toàn ngẫu nhiên được lấy ra từ một tập hợp chuẩn.

IV. Kết luận.

1. Điều quan trọng để thực hiện việc xác định chu kỳ của mưa ở đây là độ dài của chuỗi số liệu. Như trên đã nêu lên rằng chuỗi số liệu phải có tính ngẫu nhiên cao và được lấy ra từ một tập hợp chuẩn (mặt chất của chuỗi số liệu) nhưng về độ dài (về lượng) của chuỗi số ta còn chưa xét tới.

Trong nguyên tắc tính chu kỳ theo phương pháp này thường thì số bước trượt lớn nhất m_{max} chiếm tới 30% độ dài của chuỗi số liệu, thậm chí có thể lớn hơn chút ít.

Như vậy nếu ta trượt với m_{max} = 12 năm thì lượng số liệu cần có của chuỗi số khoảng 30 - 36 năm. Nhưng trên thực tế ta chỉ có chuỗi số 21 năm liên tục(1954 - 1974). Vậy vấn đề là phải xét xem 21 năm số liệu này có thể thay thế cho một chuỗi có độ dài từ 30 - 36 năm không.

Để giải quyết vấn đề này, xin đưa ra ở đây số liệu của trạm Buôn mê thuột giai đoạn : (1928 - 1938) có 11 năm số liệu liên tục, như vậy toàn bộ lượng số liệu của Buôn mê thuột ở đây là 32 năm (hình 1). Ta coi 32 năm số liệu này là một tập giá trị X. Tập giá trị này có các đặc trưng sau :

$$E(X) = a = 1725\text{mm} \quad \sigma = 246\text{mm}. \quad N = 32$$

chuỗi số (1954-1974) có : $\bar{X}$ = 1656mm S = 207mm. N = 21

ta cần kiểm tra giả thuyết H_o : Mẫu rút ra từ tập giá trị X đạt yêu cầu hay không .

a. Thực hiện chuẩn biến số ta có :

$$t = \frac{/\bar{X} - a/}{\frac{\sigma}{\sqrt{N}}} = \frac{/1656 - 1725/}{\frac{246}{\sqrt{32}}} = 1,56 < 1,96$$

b - So sánh hai phương sai :

$$t = \frac{/S^2 - \sigma^2/}{\sqrt{\frac{2.\sigma^4}{N}}} = \frac{/42816 - 60381/}{\sqrt{2.\frac{(60381)^2}{32}}} = 1,16 < 1,96$$

Vậy kết luận rằng : Mẫu số liệu (1954 - 1974) được rút ra từ tập X đạt yêu cầu với p = 95%.

2. Chúng ta đã có kết luận tính chất của mẫu số liệu đưa ra ở đây thỏa mãn theo ý nghĩa thống kê về mặt chất cũng như về mặt lượng có thể cho phép ta tiến hành nghiên cứu chu kỳ của mưa năm. Nhưng thực tế sự hạn chế của lượng số liệu cũng còn đặt ra nhiều vấn đề :

- Các giá trị chu kỳ được tính ra ở trên chưa phải là những chu kỳ duy nhất ở khu vực này.

- Lượng thông tin ngắn khiến cho ta không nghiên cứu chu kỳ trong lĩnh vực tần số điều đó làm cho kết qủa phản ánh chưa thật đầy đủ.

3. Công cụ tính toán theo phương pháp này thường là máy tính điện tử. Tuy nhiên trong trường hợp tính cho một trạm với lượng thông tin không nhiều, chúng ta có thể dùng các máy tính đơn giản hơn.

Tài liệu tham khảo :

1. G.A.Xéc-gây-ép. D.A.I-anút-sơ.
   Các phương pháp thống kê nghiên cứu các đối tượng thiên nhiên. NXB Khí tượng thủy văn, Lê-nin-grát, 1973.
2. E.I.Gu-rơ-xki.
   Tuyển tập các bài tập theo lý thuyết xác suất và toán thống kê.NXB " Trường cao đẳng ", Mát-xcơ, 1975.
3. Tr.Kai-xlơ.
   Phân tích chuỗi thời gian số liệu thủy văn. NXB Khí tượng thủy văn Lê-nin-grát, 1972.
4. I.U.A.Rô-za-nốp.
   Các qúa trình ngẫu nhiên. NXB " Khoa học " Mát-xcơ-va, 1971.
5. Phương pháp thống kê trong thủy văn. Bản tiếng Nga dịch từ tiếng Anh của M.I.Ru-xin-nốp. NXB Khí tượng thủy văn, Lê-nin-grát, 1970.