[redacted]LIMATIC FLUCTUATIONS STUDIED [redacted] BY USING ANNUAL FLOWS AND EFFECTIVE ANNUAL PRECIPITATIONS[1]

by

V. M. Yevdjevich

Civil Engineering Section, Colorado State University, Fort Collins, U.S.A.

INTRODUCTION

SUBJECT

Analyses of climatic fluctuations by using the data of annual flows and derived effective annual precipitations[2] from many river gauging stations around the world are the subject of this paper.

Many efforts have been made in the part either to discover the regularity or to prove the randomness[3] in the fluctuations of some climatic or hydrologic events on annual basis. The analysis here is intended to add some new elements to this old problem.

SELECTION OF PHENOMENA FOR THE STUDY OF FLUCTUATIONS

The study of climatic fluctuations was restricted to time series of annual flows and effective annual precipitations, because they are considered here as relatively the most reliable series for investigation of climatic changes related to the water resources problems. The run-off at a river gauging station integrates the effects of climatic factors on large areas, and some bias inherent to point measurements (of rainfall, temperature, and similar) is thus eliminated. The analysis of data of rainfall, temperature, tree rings, sediment deposits (varves) and others has shown either inconsistency[4] and non-homogeneity[5] in data, difficult to correct, or a built-in dependence model (biological model in tree rings, or sedimentation model in varves), which does not exist in data of annual flows and effective annual precipitation.

BASIC APPROACH

Generally, two opposite standpoints may be distinguished among the results of studies of climatic fluctuations, with many positions between the two opposite views.

Approach of oscillatory movements

This standpoint is based on assumptions that moon effect, sun-spot fluctuations and other solar activities, cosmic effects, persistence in ocean and air-mass movements, and the like, create a persistent or even oscillatory movement of high and low values, which combined with random effects make the future values of a time series depend on the previous ones. The hidden periodicities are often being advanced from investigation of small samples of climatic factors. The non-significance of differences between parameters of observed and stochastic time series is usually neglected by this standpoint.

Approach of randomness and stochastic processes

This standpoint starts from the fact that the natural fluctuations of annual values are very close to random sequence, if some influences of known regression effects are eliminated. The examples of these effects are: storage of water and heat in oceans, atmosphere, earth; regression effect created by the methods used in defining the time series; end effect of the time unit used; especially retardation of water due to water storage in river basins in annual flow fluctuations; etc. Very small departures, considered as non-significant, of computed statistical parameters from those of the random series or series derived from random time series by stochastic processes are emphasized by this point of view. The fact that those statistical parameters have consistent

1. This research in fluctuations of annual run-off was sponsored originally by the United States National Bureau of Standards and the United States Geological Survey; it is now sponsored by the United States National Science Foundation.
2. Effective annual precipitation defined as total annual precipitation minus total annual evapotranspiration in a river basin.
3. Randomness is deffned here in classical way, that there is no systematic link between successive values of an infinite time series.
4. Inconsistency is defined here as systematic errors due to measuring or computational techniques.
5. Non-homogeneity is defined here either as man-made effects, or accidents in nature (fires, land-slides) which change the measured values in comparison with the virgin values.

departures in the same direction from the same statistical parameters of random or stochastic series is neglected in this standpoint, even though those departures can be considered non-significant.

Many scientists agree that the random component in series of annual flows and effective annual precipitation is very high, but add that something exists beyond the known storage effects, because of that consistency in departures of statistical parameters between the observed and random or stochastic time series.

The controversy lies mainly in the way that these departures are explained and related to some physical or non-physical causes.

The approach of this paper is the analysis of the mentioned departures between the time series of observed annual flows, or derived effective annual precipitations, and the random or stochastic time series, and the explanations for these departures, which divide the two standpoints.

The use of the moving average by smoothing the original time series in order to study the fluctuations of a phenomenon is avoided in this study. From the studies of Slutzky (1937) it is well known that its use creates, of itself, fluctuations which may lead to erroneous conclusions. The random series becomes stochastic (non-random) series, when the method of the moving average is used. The series will be used here with their unchanged member values.

The random time series is considered in this paper as bench mark or yardstick series. The stochastic series as derived from random series by known processes are a further extension of bench mark random series.

STATISTICAL TECHNIQUES USED

The following statistical parameters were selected for use in this paper:

1. Serial correlation coefficients and correlograms. They give a simple method of studying the dependence in successive terms of a time series. They are convenient, and especially the first serial correlation coefficient, when the absolute values of serial correlation coefficients are small. This is the case usually with time series of annual river flows and effective annual precipitations.
2. Range. It is defined for a part of time series as the difference between previous maximum and previous minimum on the cumulative curve of departures.

DATA USED

Data used consists of annual flows and of stored water in river basins at the end of water years for 140 river gauging stations from many parts of the world, and namely: 72 from the United States of America, 13 from Canada, 37 from Europe, 10 from Australia, 1 from New Zealand, 4 from Africa, 2 from Japan and 1 from the Middle East.

The sizes of river basins range from 2.5 km^2 to 2.5 million km^2, but the majority are from 800 to 80,000 km^2. The annual average flows are in the range of 2.8-2,800 m.^{3}/sec., with an average yield up to 30 lit./sec./km.^{2}.

The length of records is 40-150 years. Seven stations of long record are specially studied: River Göta at Sjötorp-Vänersborg (Sweden), 150 years; River Rhine at Basle (Switzerland), 150 years; River Nemunas at Smolininkai (Lithuania, U.S.S.R.), 132 years; River Danube at Orshava in Iron Gates (Rumania), 120 years; St. Lawrence River, at Ogdensburg, New York (U.S.A.), 97 years; River Mississippi at St. Louis, Missouri (U.S.A.), 96 years; and River Neva at Petrokrepost (U.S.S.R.), 76 years.

DETERMINATION OF EFFECTIVE ANNUAL PRECIPITATIONS

FIRST APPROXIMATION DETERMINATION OF EFFECTIVE ANNUAL PRECIPITATION

The effective annual precipitation is defined here as:

$$P_e = P_i - E_1 = V_i + W_e - W_b = V_i - \Delta W_i \qquad (1)$$

where P_e = effective annual precipitation for a given water year and a river basin (net available water in each water year for a river basin, or net input of water from atmosphere into a basin); P_i = total annual precipitation on that river basin and for a given water year; E_i = total annual evapotranspiration (total water losses from the basin, from surface and underground waters, into the atmosphere); V_i = annual water flow volume of a river for a given water year; W_b = total stored water in a river basin at the end of a water year; W_e = total stored water in a river basin at the beginning of the corresponding water year; ΔW_i = difference of stored water for the end and beginning of a water year (positive in wet water years, negative in dry water years).

As there are usually great errors in the determination of P_i and E_i for a river basin, starting from precipitation and evaporation measurements on limited number of individual points, the effective annual precipitation was obtained for this paper exclusively by using W_b and W_e, or their difference ΔW_i. The value of ΔW_i was determined for the approximate values W_b and W_e.

The mean flow recession curve of average daily or average monthly flows was determined for each station, and for the season around the end or beginning of water years, and they were approximated by exponential functions, either of the type $Q = Q_o e^{-ct}$, or of the type $Q = Q_o e^{-ct^n}$, with Q_o the initial discharge, Q (any discharge of recession curve), and t (time) as variables, with c, or c and n, the parameters which characterize the equations of mean flow recession curves. The inte-

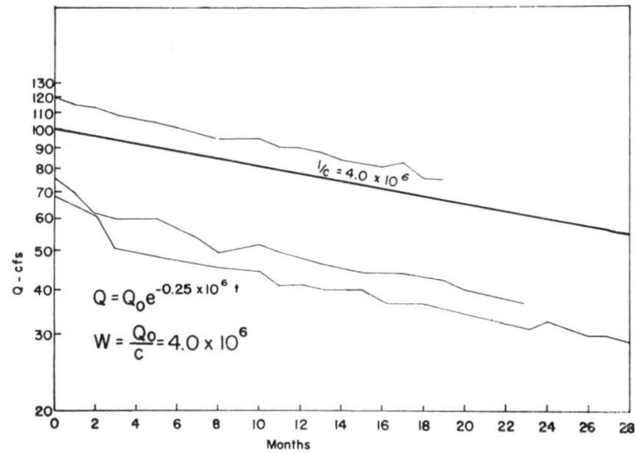

FIG. 1. Determination of the mean recession curve of Ashley Creek, Utah, by using monthly flows, for the purpose of compiling the water carry-overs from one water to another.

gration of the above functions of mean recession curves, from given Q_o for $t = 0$, to $t =$ infinite give the volume of water W, or the stored water in a river basin. In this case, W_b or W_e are functions (for given c, or for given c and n) of the discharge Q_o at the end of a water year, if this end occurs during a recession curve period (for the United States, Europe and Canada the end was on 30 September, and the beginning of the next water year on 1 October). In case a flood wave (rising level especially) occurs at the end of a water year, a special procedure was followed to obtain the stored water volume W.

The illustration of this method of determining W is given in Fig. 1 (Ashley Creek, Utah). The recession curve of monthly flows is a straight line in semi-log paper, with the average c-value equal to 0.25×10^{-6} for the simple exponential function. In this case the integration of the volume from $t = 0$ and Q_o to $t =$ infinity gives $W = Q_o/c$, or in this example $W = 4.0 \times 10^{-6} Q_o$, with $Q_o =$ discharge on 1 October. By using equation (1) for each water year, as $P_e = V_i - \Delta W_i$, the effective annual precipitations were computed as a first approximation for all river gauging stations and for all water years.

The computed series of annual flows in modular coefficients as $U_i = \frac{V_i}{V_a}$, where $V_a =$ average annual flows, and derived series of effective annual precipitations in modular coefficients as $Y_i = \frac{P_e}{V_a}$, are used as basic time series for analyses of climatic patterns.

As the effective annual precipitations represent the net input of water in a river basin for a water year, this is the water volume after the storage effects of the surface and underground reservoirs and other storage capacities in a river basin were excluded. As they represent annual flows without a carry-over of water from year to year, the interdependence among the successive values of effective annual precipitations should be smaller than for the annual flows, as will be be shown later.

TRUE VALUES OF EFFECTIVE ANNUAL PRECIPITATIONS

The true value of the effective annual precipitations, is, however,

$$P_i = V_i - \Delta W_i \pm (E_i + I_i + H_i + G_i) \qquad (2)$$

where $E_i =$ random error in annual flows (usually small for annual river flows); $I_i =$ inconsistency or systematic errors, produced by measuring and computation techniques, in general given as trends or jumps, or their combination; $H_i =$ non-homogeneity or change in the virgin flow, produced by different man-made influences or accidents in the river basin; $G_i =$ error in determining ΔW_i approximately by using in this study the average recession curves, instead of a recession curve for each individual water year and gauging station.

The experience shows that the influence of I_i, H_i, and G_i may be important, and that these factors and errors cannot be often neglected in treating the fluctuations of annual flow and of effective annual precipitation.

In some cases the patterns in fluctuations of flows determined as dependence of successive values of time series could be explained partly by the inconsistency and non-homogeneity in data. The sufficient accuracy of computed members of the two time series V_i and ΔW_i should not be assumed *a priori*.

RELATION OF EFFECTIVE ANNUAL PRECIPITATIONS AND ANNUAL RIVER FLOWS

Assuming that an input of water in all storage capacities in a river basin is random, and that the outflow of stored water follows an exponential function (which is close to reality for most of underground and uncontrolled surface storage spaces), the output of water will be a stochastic time series of the type

$$U_i = b_o Y_i + b_1 Y_{i-1} + b_2 Y_{i-2} + \dots \qquad (3)$$

where b_o, b_1, $b_2 \dots$ are monotonically decreasing and positive coefficients, with $\sum_{o}^{\infty} b_j = 1$. This equation is valid also, if the Y_i values are non-random, but an approximate exponential outflow law may be applied to storage capacities of a river basin, In nearly all cases, due to a fast outflow of water from storage spaces, the infinite series on the right side of equation (3) may be replaced by m members. Practically m does not pass 10 (the case of annual flows of the St. Lawrence River at Ogdensburg, New York, with a tremendous storage capacity of Great Lakes). In this case $\sum_{0}^{m-1} b_j = 1$, neglecting the insignificant tail beyond $(m-1)$ value of b_j.

The property of equation (3) is that the standard deviation of U series becomes smaller than that of Y series under given four conditions for b coefficients. Equation (3) is a general type of linear equation for moving average model (Markov chains), and this type of moving average attenuates the extremes of Y series in producing U series.

As the effective annual precipitation is the difference of precipitation P_i and evapotranspiration E_i, the properties of Y series, therefore, depend on both series P_i and E_i. Because P_i and E_i are usually dependent among them, and particularly when the ratio of effective and total precipitation $\frac{P_e}{P_t} = 1 - \frac{E_i}{P_i}$ is small, the characteristics of Y series depend on three factors: P_i, E_i, and $E_i = f(P_i)$. If P_i values would be of random sequence or close to it, and if $E_i = f(P_i)$ were of a complex relationship (not a simple linear function), the Y series would be non-random. Therefore, E_i plays an important effect on the type of time series which can approximate the series of effective annual precipitation.

PATTERNS IN CLIMATIC FLUCTUATIONS MEASURED BY FIRST SERIAL CORRELATION COEFFICIENT OF BOTH ANNUAL FLOWS AND EFFECTIVE ANNUAL PRECIPITATIONS

FREQUENCY DISTRIBUTIONS OF FIRST SERIAL CORRELATION COEFFICIENT

Procedure

The first serial correlation coefficient, defined as the correlation coefficient of successive pairs of annual flows or effective annual precipitations, was used as a measure of dependence of successive values in the two time series (flow and precipitation), or as an index of the possible climatic fluctuation patterns.

The unbiased first serial correlation coefficient is given in classical statistical books as:

$$r_i = \frac{\sum_{1}^{N} U_i U_{i+1} - \frac{1}{N-1} \sum_{1}^{N-1} U_1 \sum_{1}^{N-1} U_{i+1}}{(N-2)\, s_i s_{i+1}} \tag{4}$$

where U_i denotes any member of annual flow time series (Y_i of effective annual precipitations), U_{i+1} is the next member to U_i, so that (U_i, U_{i+1}) represents the successive pairs of members of the time series, N = total number of members in a time series, N — 1 is total number of correlated pairs, s_i is the standard deviation of (N — 1) first members, and s_{i+1} is the standard deviation of (N — 1) last members of the time series, with r_1 = first serial correlation coefficient.

The unbiased standard deviation is given in classical books of statistics as:

$$s_i = \left[\frac{1}{N-2} \sum_{1}^{N-1} U^2_i - \frac{1}{(N-2)(N-1)} \left(\sum_{1}^{N-1} U_i\right)^2\right]^{1/2} \tag{5}$$

For s_{i+1}, U_i is replaced by U_{i+1}.

The digital computer was used to determine the first serial correlation coefficients for all 140 stations, and s_i and s_{i+1} through equation (5) were used in this computation. The r_1 values were determined for both series: U (annual river flows), and Y (effective annual precipitations).

For a pure random time series (fluctuation of a random variable), considered as an open series (or that N — 1 pairs are used for the computation of the first serial correlation coefficient) in contrast with a circular time series for which the last term is supposed to be succeeded by the first term of the series (in which case there are N pairs), R. L. Anderson (1941) gives the expected value with the symbol $E(r_1)$ (the mean value of r_1 distribution) of the first serial correlation coefficient for circular time series as:

$$E(r_1) = -\frac{1}{N-1} \tag{6}$$

and the variance of distribution of first serial correlation coefficient as

$$\text{var}\, r_1 = \frac{N-2}{(N-1)^2} \tag{7}$$

which both converge toward zero by an increase of N. For $N \geq 40$ the open time series gives r_1 values close to those of circular series. As the practical minimum of years is 40 ($N_{min} = 40$), the maximum value of $E(r_1)$ is —0.0256, which is close to zero. The standard deviation s_r of r_1 is maximum for N minimum. In case of $N_{min} = 40$ it is $s_r(\max) = 0.158$, which is a relatively high value. For the maximum value of N = 150 years in the study of 140 stations, $s_r(\min) = 0.082$, also a rather large value. Therefore, it is to be expected that the values of r_1 for many stations, under the assumption of random fluctuations, would cover a relatively large range, both positive and negative values, with the mean close to zero.

The legth of time series for 140 stations is different, with the mean value $N_m = 55$, and extremes 37 and 150. The cumulative frequency distributions of 140 values of first serial correlation ocefficient for both U series and Y series are given in Fig. 2. The theoretical cumulative frequency distribution of random time series is plotted also, for $N_m = 55$, with the mean —0.018 and standard deviation 0.135.

Results

Many values of r_1 are negative: 16 (or 11.4 per cent) for U series, and 26 (or 18.6 per cent) for Y series. The approximate corrections for carry-over has nearly doubled the number of negative r_1 values in Y series as compared with U series. The mean r_1 value for U series is 0.176 and that for Y series 0.130. The difference represents 35 per cent of the value for Y series. The water carry-over in river basins is thus the cause of at least 35 per cent of the mean positive value of first serial correlation coefficient of U series. The median values are 0.160 and 0.115 respectively, with difference being 39 per cent of the median value for Y series. Therefore, one-third to two-fifths of positive first serial correlation coefficient of series of annual flows is explained by the water storage effect in river basins.

The first serial correlation coefficients of log U_i (or log Y_i) have shown the same patterns as the first serial correlation coefficients of U_i (or Y_i). The differences between r_1 values for log U_i series and U_i series (or for log Y_i series and Y_i series) are relatively small, and either positive or negative. There is no clear pattern to be distinguished among two sets of r_1 values (U_i versus log U_i, Y_i versus log Y_i).

The difference between the first serial correlation coefficients of U series and of Y series increases with an increase of parameter $\frac{W}{A}$, given in feet, as the ratio of mean annual carry-over (W) to the area of river basin (A). Generally the greater $\frac{W}{A}$ the greater is the difference $r_1(U) - r_1(Y)$.

A general trend is also derived, when the specific yield of river basins, q in lit./km.^{2}, is related to the decrease of the first serial correlation coefficients of U series to the corresponding value of Y series. The smaller the specific yield, the greater is, on the average, this difference of r_1 values. This can be easily explained by the fact that for given topographical and geological conditions in a river basin, for its given area, the available space for underground and surface water storage in relation to mean annual run-off is greater on the average in dry climates (small specific yields) than in humid climates (great specific yields).

Fig. 2 shows, that the slope of curves (1) and (2) in the range of 20-95 per cent of probability can be well fitted by straight lines, or by normal probability function. The slope of both these curves is the same and is equal to the slope of the cumulative distribution of an infinite set of random time series with N = 55 (equal to mean length of series U and Y).

The results of the analysis of 140 stations show that the arid regions have on the average the greater values of r_1 for both U and Y series than the humid regions. The example of the region of the Upper Colorado River Basin and around it (14 stations) shows this trend clearly, as given in Fig. 3. The legend of the figure explains all plotted cumulative curves. The difference of median r_1 values of this semi-arid region and of 140 stations are 0.064 for U series, and 0.069 for Y series.

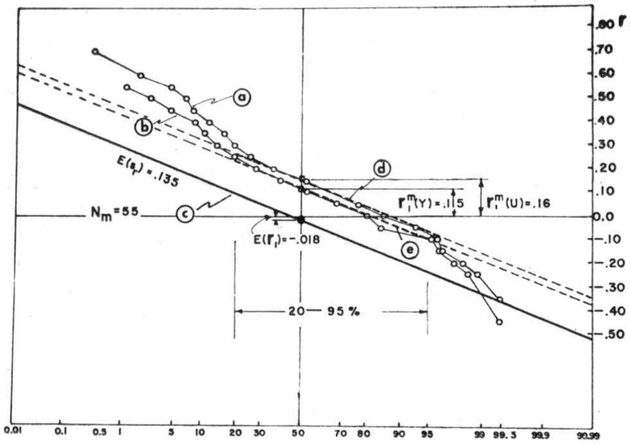

Fig. 2. Cumulative frequency distributions of first serial correlation coefficients (r_1) for 140 river stations, using carthesian-probability scales : (1) U_i series, computed; (2) Y_i series, computed; (3) random series for N = 55; (4) U_i series, fitting a straight line for range 20-95 per cent; (5) Y_i series, fitting a straight line for 20-95 per cent range of cumulative frequency.

Example of effect of carry-over

As an example of a significant impact of water carry-over from one water year to another on the first serial

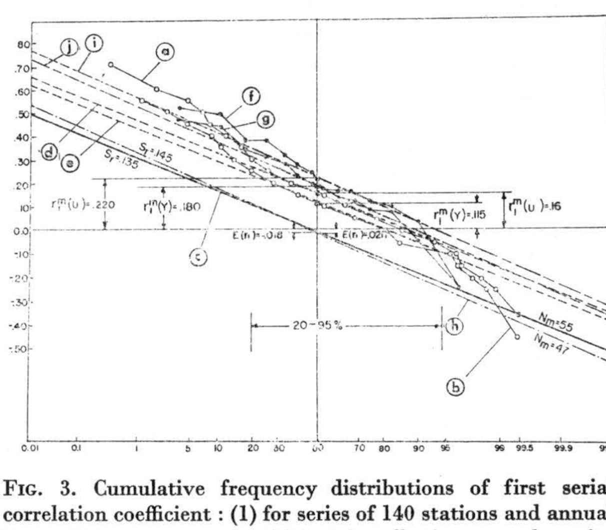

Fig. 3. Cumulative frequency distributions of first serial correlation coefficient : (1) for series of 140 stations and annual flows; (2) same as under (1) but for effective annual precipitation; (3) for random series, $N_m = 55$; (4) line parallel to line 3 through $r = 0.16$; (5) same as (4) through $r = 0.115$; (6) for series of 14 stations of Upper Colorado and annual flows; (7) same as under (6) but for effective annual precipitation; (8) for random series, $N_m = 47$; (9) line parallel to line 8 through $r = 0.22$; (10) same as (9) through $r = 0.18$.

correlation coefficient is the case of St. Lawrence River at Ogdensburg, New York. For the period of observations of 97 years, from 1860 to 1957, the first serial correlation coefficient of actual annual flows for St. Lawrence is 0.705. When the effect of stored water in the Great Lakes was taken into consideration, the first serial correlation coefficient of the effective annual precipitation was dropped to 0.094. If the stored water in the St. Lawrence River Basin outside the Great Lakes, which means in the small lakes, river and in the underground, would be taken into consideration also, then the first serial correlation coefficient of 0.094 would probably be decreased still further.

Discussion of results and conclusions

From the comparison of the distributions represented in Figs. 2 and 3 the following may be derived:

1. Cumulative frequency distributions of U series and Y series either for 140 or 14 stations (Upper Colorado) are very close to the straight lines (extremes excluded) on arithmetic-probability paper; or their first serial correlation coefficients are normally distributed.
2. Slopes of these lines, which represent the standard deviation of first serial correlation coefficient distributions, are also very close (at least for probabilities in the range 10-90 per cent) to the slopes of cumulative distributions of coefficients of random time series which have the same lengths as the mean length for U series and Y series.
3. Difference between the frequency distribution of first serial correlation coefficients for U and Y series, and the frequency distribution of random time series is only in the mean values of coefficients. The U and Y series have positive mean values of r_1, while the means of random series are negative but practically close to zero.
4. Differences between the means of first serial correlation coefficients for U and Y series, on one side, and the mean for corresponding random time series for 140 stations selected from many parts of the world are:

   $$\text{U series: } \Delta\bar{r}_1(\text{U}) = 0.160 + 0.018 = 0.178$$

   $$\text{Y series: } \Delta\bar{r}_1(\text{Y}) = 0.115 + 0.018 = 0.133$$

   or the sequence of effective annual precipitations are much closer to random series than the sequence of annual flows.
5. Main problem in detecting the patterns in fluctuations of annual flows and of effective annual precipitations by an analysis of distributions of first serial correlation coefficients is the interpretation of these differences in the means or medians of first serial correlation coefficients, first of U and Y series versus the random time series, and then of differences between U and Y series.
6. It seems that the arid and semi-arid regions have a greater difference between observed series and random series than the humid regions, as will be shown later in regional distributions of the first serial correlation coefficient.
7. Factors which cause the mean first serial correlation coefficients to be greater than zero in humid regions are more pronounced and emphasized in the semi-arid and arid regions. A complex relationship between evapotranspiration and precipitation in arid regions, as well as a complex law of evaporation of rainfall in the air before the rain falls on the earth in arid regions, may be partly responsible for this difference.
8. There are departures at the extremes of distributions of first serial correlation coefficients from the straight line passed through the median values with the slopes of corresponding random time series. Effects of glaciers and snow carry-over from year to year, sampling departures, and the use of mean length N of observed series for deriving the parameters of first serial correlation coefficient distributions for random time series may be partly responsible for these departures.
9. Supposing that the figures for U and Y series are true values, it can be concluded that the first serial correlation coefficients are substantially decreased by excluding the water carry-over from year to year. The carry-over is responsible for one large part of the greater values of median or mean first serial correlation coefficients of U series than those of random time series.
10. Equation (2) points out that the difference between the given values of effective annual precipitation and the true values of effective annual precipitations can be caused also by four types of errors: random errors, inconsistency, non-homogeneity, and the errors in determining the carry-over from year to year, apart from the carry-over effect (ΔW). The random errors only increase the standard deviation of the random time series. Random errors in annual flows are small and can be practically neglected. Any inconsistency (errors in one side which can change from place to place in time series or any inconsistency in the form of trends or jumps), and any non-homogeneity of data, and any error in computing the carry-over in a river basin, increase on the average the first serial correlation coefficients (Yevdjevich, in preparation, 1962).

REGIONAL DISTRIBUTION OF FIRST SERIAL CORRELATION COEFFICIENT

Fig. 4 shows the position of 72 river gauging stations in the United States. It is an example of regional distribution of first serial correlation coefficient. The solid lines divide 14 hydrological regions as designated by the United States Geological Survey. There are two

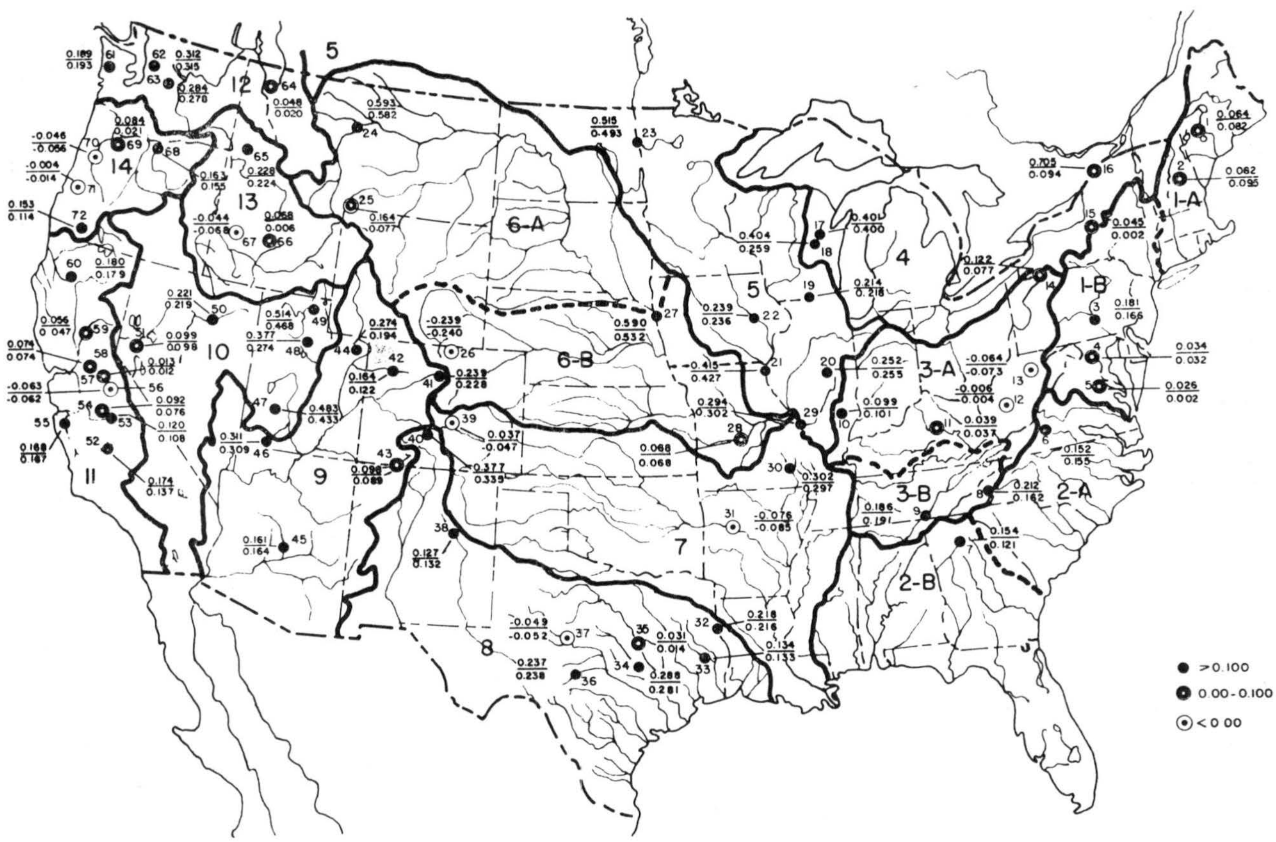

FIG. 4. Regional distribution of first serial correlation coefficient for annual flows (upper figure) and effective annual precipitations (lower figure) for 72 river gauging stations in the United States.

figures for each station: the upper figure is the first serial correlation coefficients for the annual flows, and the lower figure is the first serial correlation coefficients for the effective annual precipitations. Taking the effective annual precipitations as a measure, all stations having the first serial correlation coefficient greater than +0.10, and all stations having the first serial correlation coefficients lower than +0.10 are specially marked. The negative first serial correlation coefficient for the effective annual precipitations are given an additional sign. There are 48 stations which have the first serial correlation coefficient of Y series above and 37 stations below +0.10. There are 14 stations with negative first serial correlation coefficients.

The general conclusions from the results represented in Fig. 4 are:

1. The humid regions of the east and the west of the United States more frequently have the first serial correlation coefficients for the effective annual precipitation below +0.10 than above +0.10.
2. The dry regions in the Middle West and in the Rocky Mountains more frequently have the first serial correlation coefficients of the effective annual precipitations above +0.10 than below +0.10. The regions around the Gulf of Mexico would be considered as approximately having the same number of stations with first serial correlation coefficients above or below +0.10.

It can be concluded from this approximate analysis, that stations in arid regions are more likely to have greater first serial correlation coefficients for the effective annual precipitation than stations in humid regions.

It seems a quite attractive conclusion that, before any climatic reason is studied for explaining this difference in first serial correlation coefficients between arid and humid regions, or before any climatic reason is advanced for the positive mean first serial correlation coefficients in effective annual precipitation the effects of inconsistency, non-homogeneity, and errors in determining the volumes of carry-over should be first taken into consideration.

The man-made changes are very likely to affect more the flows of an arid region than of a humid region. However, the complex relationship relating the evaporation and precipitation, either in the air during rainfall, or especially on the earth surface, may be the other important factor affecting the above differences.

PATTERNS IN FLUCTUATIONS MEASURED BY CORRELOGRAMS

DEFINITION, GENERAL REMARKS AND PROCEDURE

Definition

The correlogram is defined as a graph of discrete points relating the serial correlation coefficients and the lag between successive correlated pairs of members of a time series.

General remarks

The correlogram is a measure and an indicator of independence or of the type of dependence among the members of a time series. A random time series, if it is sufficiently long, has a random sequence of coefficients in correlogram, but the serial correlation coefficients are confined within the confidence limits of a given probability.

The confidence limits on 95 per cent level for random time series were computed by using equation (8).

According to R. L. Anderson (1941), the confidence limits for random time series for 95 per cent level are approximately

$$R_{95\%} = \frac{-1 \pm 1.64\sqrt{N-k-2}}{N-k-1} \qquad (8)$$

with the meaning that 5 per cent of the points of correlogram should be on the average outside the confidence limits.

Procedure

The serial correlation coefficients for the annual flows and effective annual precipitations were compiled by using the following equations, similar to equations (4) and (5) in which r_1 is replaced by r_k and unity by k:

$$r_k = \frac{\sum_{1}^{N-k} U_i U_{i+k} - \frac{1}{N-k} \sum_{1}^{N-k} U_i \sum_{1}^{N-k} U_{i+k}}{(N-k-1)s_i s_{i+k}} \qquad (9)$$

with

$$s_i^2 = \frac{1}{N-k-1}\left[\sum_{1}^{N-k} U_i^2 - \frac{1}{N-k}\left(\sum_{1}^{N-k} U_i\right)^2\right] \qquad (10)$$

and for s^2_{i+k} the value i in equation (10) is replaced by $i+k$.

The computation of these unbiased serial correlation coefficients was carried out up to $m \leq \frac{N}{4}$, where N = length of time series, by using a digital computer.

RESULTS

Correlograms for individual river stations with long records

From seven river stations of long record, already mentioned above, and whose correlograms are computed and studied, only three will be given here.

Fig. 5 gives the correlogram for the River Göta at Sjötorp-Vänersborg (Sweden), N = 150 years; Fig. 6

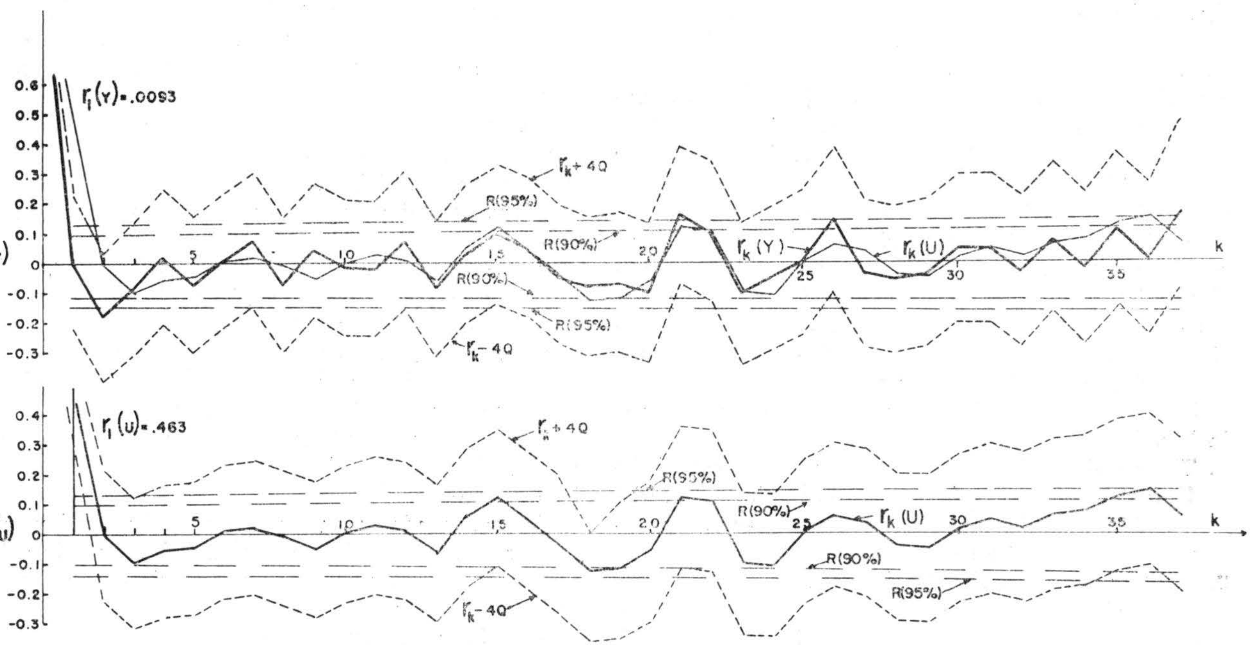

FIG. 5. Correlogram for River Göta at Sjötorp-Vänersborg.

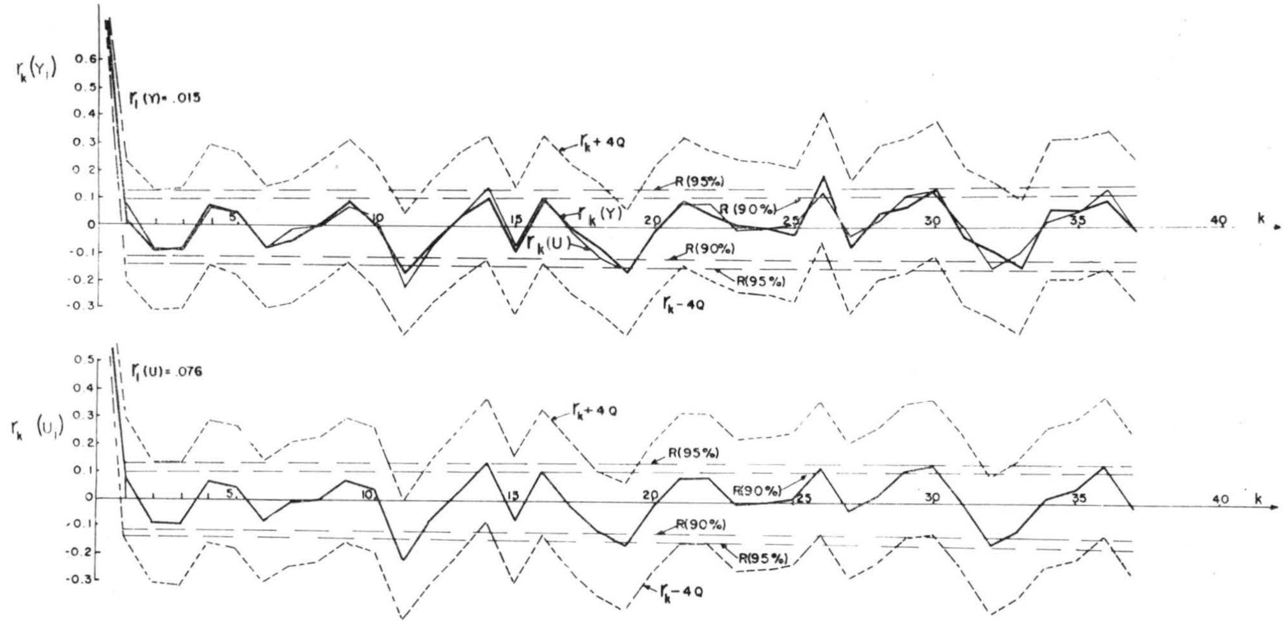

FIG. 6. Correlogram for River Rhine at Basle.

for the River Rhine at Basle (Switzerland), N = 150 years; and Fig. 7 for the St. Lawrence River at Ogdensburg, New York (U.S.A.), N = 97 years. The correlograms are given for both U and Y series, with confidence limit on 95 per cent level computed by equation (8), and similarly for 90 per cent level. The confidence limits of $r_k \pm 4Q$, where r_k = any correlation coefficient of lag k, and Q = probable error of r_k, are also given.

For the River Göta, Y series, the last value for $k = 37$ excluded, only two values of $r_k (k = 2, k = 21)$ exceed the 95 per cent confidence interval. The value $r_1 = 0.0093$ of Y series is very close to the expected value of random series of $r_1 = -0.0067$. By correction of water carry-over in river basin, the value $r_1 = 0.463$ of U series is reduced to value $r_1 = 0.0093$ of Y series.

For the River Rhine the correlograms of both series lie practically inside the confidence interval of 95 per cent level, because only two or three values of r_k exceed the limits. The $r_1 = 0.076$ of U series is reduced to $r_1 = 0.015$ of Y series by eliminating the effect of carry-over. The expected value of random series is $r_1 = 0.0067$.

The St. Lawrence River has three r_k values of Y series correlogram exceeding the confidence limits of 95 per cent level. The U series has a correlogram with significant positive r_k values from $r_1 - r_9$. The $r_1 = 0.705$ of U series is reduced by excluding river basin carry-over to $r_1 = 0.094$ of Y series.

Mean correlogram for 140 river gauging stations

Fig. 8 gives the average correlogram for 140 river gauging stations, obtained by using the mean value of each r_k. As time series are of different length N, and the number m of computed r_k is $\frac{N}{4}$, the mean number for each r_k is either $n = 140$ (up to $k = 10$), or smaller than 140, with only $n = 2$ for $k = 37$. The confidence limits are averaged in the same way, by summing up the limits for all stations and dividing the sum by the number n of the stations for a given lag k. Fig. 8 gives also the number n of r_k values for each k which were

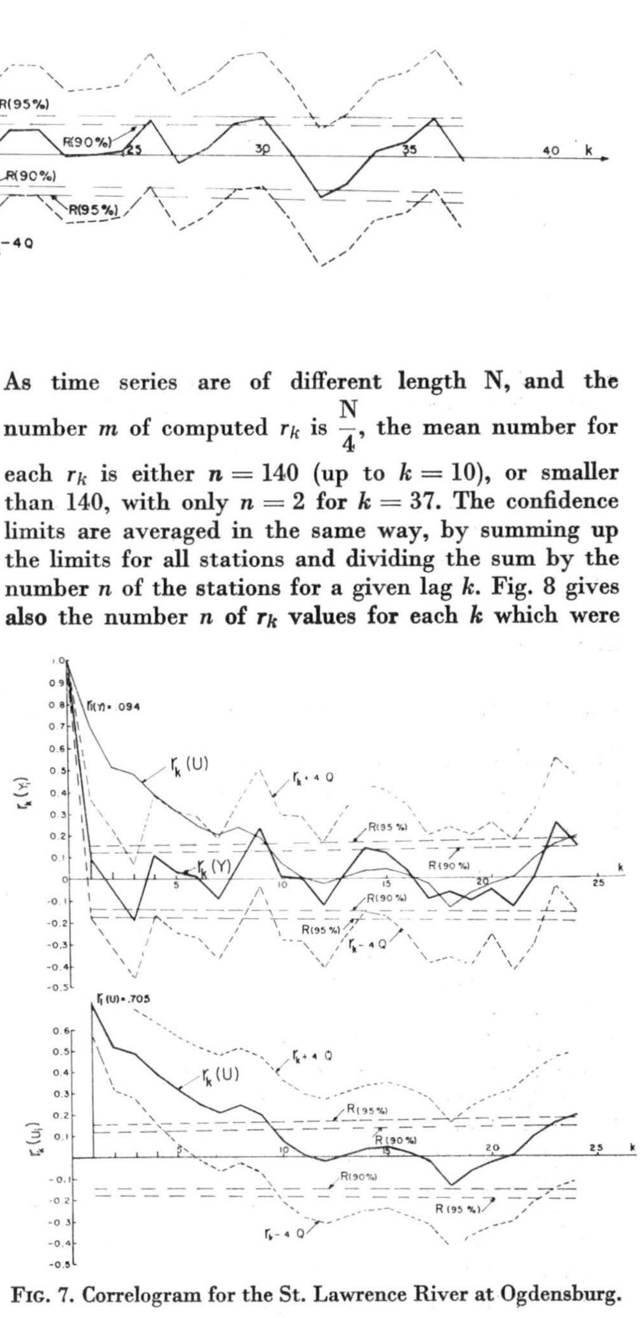

FIG. 7. Correlogram for the St. Lawrence River at Ogdensburg.

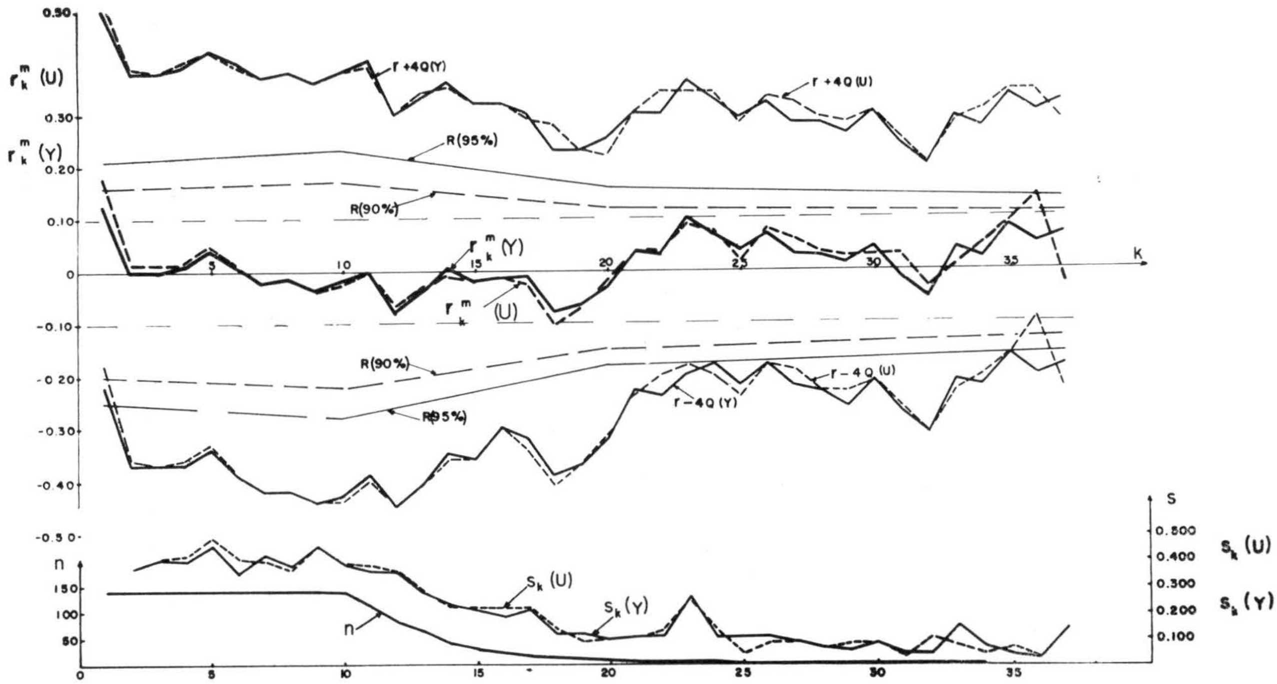

FIG. 8. Mean correlogram for 140 river stations, with mean confidence intervals (r_k 4Q), R (95 per cent) and R (90 per cent), with number n for each r_k, and the standard deviation s_k of r_k values around the mean value r_k.

used in computing the mean values, as well as the standard deviations s(U) and s(Y) of r_k values around the mean for each k. The confidence limits decrease with k, because of the averaging process, but theoretically the confidence limits should increase with an increase of k.

DISCUSSION AND CONCLUSIONS OF RESULTS

The analysis of correlograms, either of long-record stations or of the mean correlogram for 140 stations, points to the following:

1. The long-record stations (76-150 years of observations), regardless of some inherent inconsistency and non-homogeneity in data, have correlograms of Y series very close to random time series, because the number of r_k values exceeding the confidence limits is of the order of expected exceedances underlying the definition of limits.
2. The r_1 values of U series for seven long-record stations are mostly significantly different from zero, while this is not the case for Y series.
3. The St. Lawrence River is a typical example of the effect of water carry-over on the correlogram. The relationship of correlograms of U and Y series suggests the model of U and Y relationship of a moving average type, as expressed by equation (3), with b_j coefficients monotonically decreasing positive values, with their sum unity.
4. The long-record stations do not show any significant periodic movement (sun-spots, for example).
5. The mean correlogram for 140 stations clearly points out, that the values of r_k, except the mean value of r_1(U), are insignificantly different from zero, or from a random time series. The mean value of r_1(U) is greatly influenced by the effect of water carry-over.
6. The mean correlogram shows that neither the sun-spot average period of 11 years, nor the double sun-spot average period are of a significant influence on the effective annual precipitations. The lags 23-27 are, however, a little greater than zero (up to $r = 0.10$), but are inside the confidence interval.
7. Regardless which confidence limits are used (95 per cent level, 90 per cent level, $r_k \pm 4$Q), the correlograms of effective annual precipitations are not significantly different from random time series. If there would be a significant trend in climate changes all over the world, the correlogram would show this trend.

PATTERNS IN FLUCTUATIONS MEASURED BY RANGE

DEFINITION OF RANGE

The maximum range for a time series of length N and for the period N is defined as the difference of the

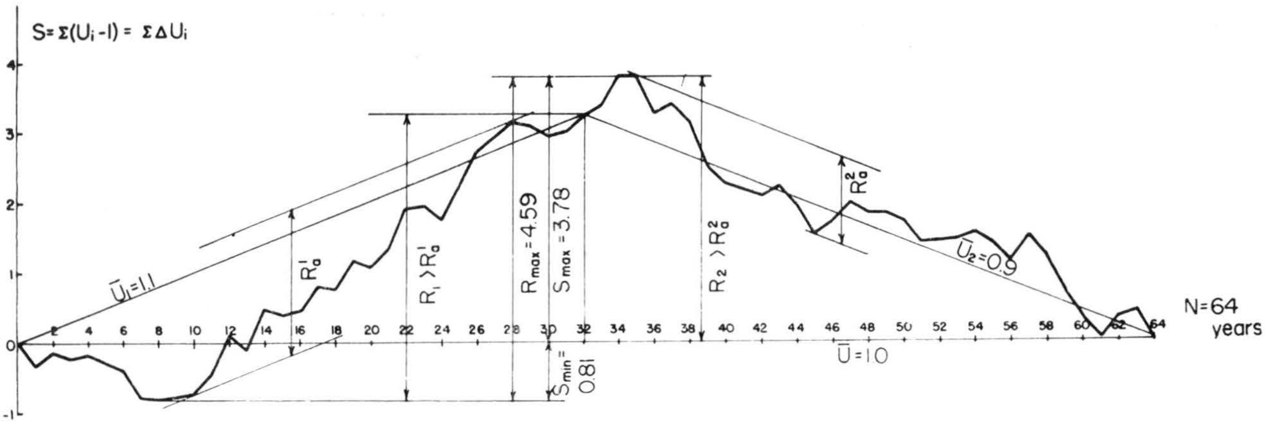

FIG. 9. Cumulation curve of departures of virgin annual flows (expressed as modular coefficients U_i) of Upper Colorado River at Lee Ferry, Arizona, from the mean $\overline{U} = 1$, for 64 years, with maximum range R_{max}, two adjusted ranges $R^1{}_a$ and $R^2{}_a$, and maximum ranges R_1 and R_2 for two periods of $n = 32$ years.

maximum and minimum values on the cumulative curve of departures.

Fig. 9 gives the cumulative curve of departures for modular coefficients U_i of virgin annual flows of the Upper Colorado River at Lee Ferry. The maximum range for N = 64 years is defined as $R_{max} = S_{max} - S_{min}$, where S represents the values of the cumulative curve of departures. According to H. E. Hurst (1951), the maximum range for annual flows can be conceived as the maximum accumulated storage when there is never a deficit in outflow (equal to the mean discharge), or as the maximum deficit, where there is never any storage, or as the sum of accumulated storage and deficit, when both storage and deficit exist.

The basic characteristics for the above definition of range is the use of departures from the mean value of flows for N years, also, for the determination of range for shorter periods than N.

In a broader sense, any constant value U_o different from unity may be used to determine departures and the corresponding ranges.

The adjusted maximum range is defined by W. Feller (1951) as the difference of maximum and minimum of the cumulative curve of departures, but with the changing mean. If a period has a length of n years, smaller than N, the mean of period n is used for computing the departures and the adjusted range. An example of this maximum adjusted range is given in Fig. 9 for two periods of $n = 32$ years (first half and second half of the total period). The lines of means $\overline{U}_1$ and $\overline{U}_2$ are plotted, and by using lines parallel to them S_{max} and S_{min} are obtained for both half periods, and then the adjusted ranges $R^1{}_a$ and $R^2{}_a$ are determined. The use of the mean $\overline{U} = 1$ for determining the ranges for both 32-year periods gives the maximum range values R_1 and R_2, also shown in Fig. 9.

All these definitions of ranges, based on $\overline{U} = 1$, or any U_o, or as the adjusted range can be used, depending on the type of problem at hand. In using the range as a statistical technique for comparing the observed series with the random series the range defined on the basis of the mean for the period of observation of N years will be used here, though the basic study (Slutzky, 1937) from which this paper is derived discusses all concepts of range.

DISTRIBUTION OF RANGE OF DIFFERENT PERIODS FOR RANDOM TIME SERIES

General distribution function

Let a period of length of n years be fixed for study (i.e., 5 years, 10 years, 25 years, etc.). Let also the total period of N years be divided in m smaller n year periods, so that $m.n + n_1 = N$, where $n_1 < n$ is a residual. If the range was determined for each of m periods on the basis of the mean for the total period N, there would be m values of the maximum range, one for each of n year periods. The distribution of these maximum ranges may be conceived as a statistical technique for analysis of fluctuations.

Fig. 10 shows the eight values of R_8 for $n = 8$ years for the Upper Colorado River at Lee Ferry. There is a distribution of R_8-values, in this case ranging from 0.82 to 1.65. This corresponds to a fixed value $n = 8$, and for the mean as basic reference for computing the departures.

Assuming n as a variable, the reference value U_o also as a variable, a general four variables function

$$F(R, p, n, U_o) = 0 \qquad (11)$$

with p = probability of the range R for given n and given U_o, can be determined either analytically (by using some approximation), or by numerical procedures in the case of a random series, or a stochastic series, with known basic distribution of U_i (or Y_i).

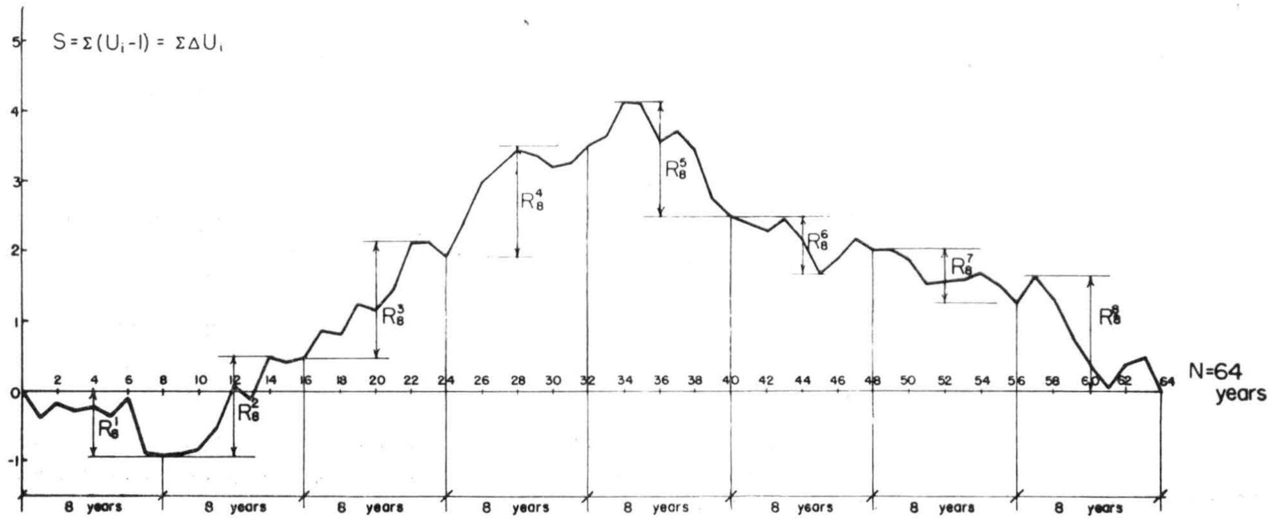

FIG. 10. Determination of ranges for eight periods, each 8 years long, for the homogeneous sample of annual river flows (reduced to depletion conditions in 1954-57) of Upper Colorado River at Lee Ferry, Arizona.

Distribution of ranges of random series as yardstick distribution

The range distribution theoretically developed for random series was used here as a yardstick to compare with the range distribution of annual flows and of effective annual precipitations.

The range distribution for both random and observed series was defined here by three statistical parameters: mean, coefficient of variation and skew coefficient. If the three parameters for given n and $\overline{U} = 1$ do not differ for the compared random and observed series in a significant manner, it is assumed that observed series is close to random series.

Distribution of range for a random series and large n

The asymptotic values for expected mean and for variance of the range of random series is given by W. Feller (1951)

$$\overline{R}_n = 1.5958 \ldots s\sqrt{n} \simeq 1.6\, s\sqrt{n} \qquad (12)$$

and

$$S^2_n = \mathrm{var}\,(R_n) = 4s^2 n \left((\ln 2 - \frac{2}{\pi} \right) = 0.2182\, s^2 n \qquad (13)$$

where s = standard deviation of time series of length N; n = length of period for which the mean range and variance of range are determined, $\overline{R}_n$ is the expected mean of range for period of length n, and S_n^2 is the variance of the range distribution.

It follows from equations (12) and (13) that the asymptotic value of the coefficient of variation of range distributions is a constant equal to 0.292.

The condition for the application of equations (12) and (13) is a large n value.

Distribution of range of a random series for small n

In case the random variable is normally distributed, the mean range for $n = 1$ is

$$\frac{\overline{R}_1}{s} = \int_0^\infty \sqrt{\frac{2}{\pi}}\, R_1\, e^{-R^2_1/2}\, dR_1 = \sqrt{\frac{2}{\pi}} \simeq 0.80 \qquad (14)$$

and variance of R_1

$$\frac{S_1^2}{s^2} = \int_0^\infty (R_1 - \overline{R}_1)^2\, p_1(R_1)\, dR_1 = \left(1 - \frac{2}{\pi}\right) = 0.363 \qquad (15)$$

where $\overline{R}_1/s$ and S_1^2/s^2 represent the mean and variance of a standardized variable with variance unity. When equations (14) and (15) are applied to modular coefficients, U_i and Y_i with mean unity and variance s^2, then

$$\overline{R}_1 = 0.80\, s\,; \text{ and } S_1^2 = 0.363\, s^2$$

The range is a truncated distribution of half the normal distribution, with a skew coefficient

$$C_s = \left(2 - \frac{\pi}{2}\right)\left(\frac{2}{\pi - 2}\right)^{2/3} \simeq 0.995$$

The distribution of R_1 is given in Fig. 11.

For $n = 2$ the probability of range R_2 is

$$p_2(R_2) = \frac{2}{\sqrt{\pi}}\, e^{-R^2_2/4} \left[\int_0^{R^2/\sqrt{2}} \frac{e^{-t^2/2}\, dt}{\sqrt{2\pi}} + \sqrt{2}\, e^{-R^2_2/4} \int_0^{R_2} \frac{e^{-t^2/2}\, dt}{\sqrt{2\pi}} \right] \qquad (16)$$

where t = a standardized variable; R_2 = any range for $n = 2$; $p_2(R_2)$ = probability of a given range R_2.

The distribution of R_2 and its statistical parameters are computed by numerical integration of equation (16), and the distribution is shown in Fig. 11.

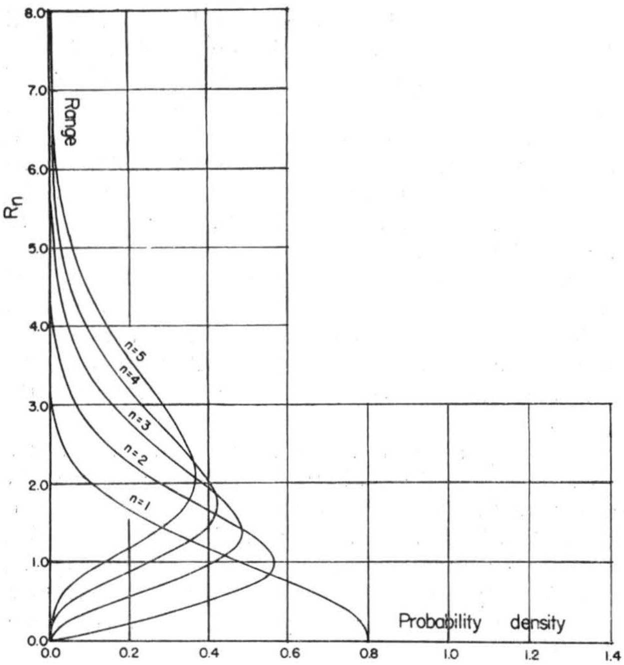

FIG. 11. Probability density distribution of ranges for $n = 1$, 2, 3, 4, and 5 for random time series of standardized normal variate.

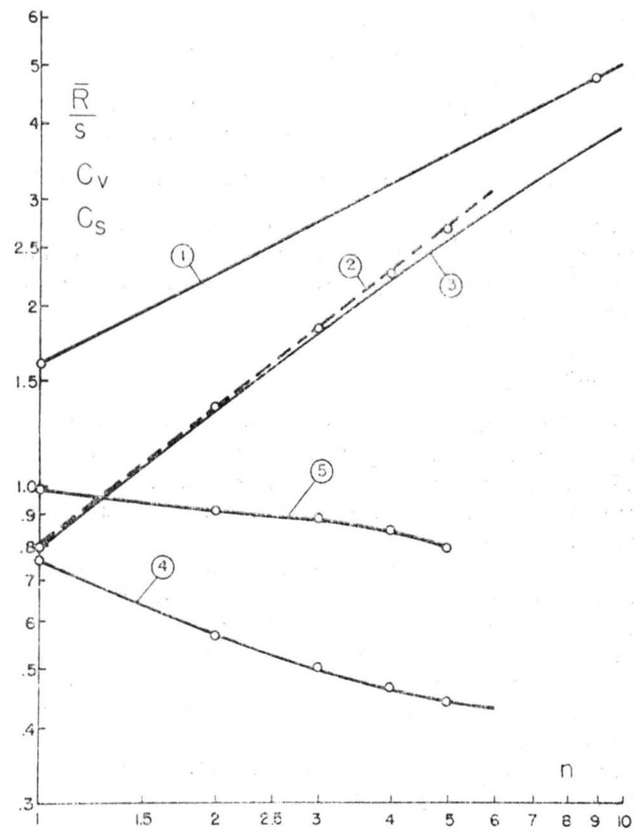

FIG. 12. The statistical parameters of range distribution of random time series of normal variable for small n (1-10) :

1. Mean range ($\bar{R}/s$), asymptotic values.
2. Exact values of mean range, obtained by numerical integration.
3. Mean range obtained by formula, given by Anis and Lloyd.
4. Exact values of coefficient of variation.
5. Exact values of skew coefficient.

The distributions and statistical parameters for R_3, R_4, and R_5 are also expressed in similar forms as equation (16), then numerically integrated or computed, and the distributions are shown in Fig. 11, and parameters in Fig. 12.

A. A. Anis and E. H. Lloyd (1953) give the mean value of the range for finite small n as

$$\bar{R}_n = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1/2} \tag{17}$$

with i integers from 1 to n.[1] As an example, for $n = 4$, $\bar{R}_4 = 2.22166$. The values $\bar{R}_i$ computed by equation (17) are given in Fig. 12, curve (3).

The comparison of the mean values of the range distribution for different small n is given for three types of values: (*a*) asymptotic values according to equation (12); (*b*) values obtained by numerical integration of exact distributions, equations (14), (15), (16) and similar; and (*c*) values obtained from analytical expression of equation (17). Though there are some departures among the curves (*b*) and (*c*), it could be assumed that equation (17) approximates closely the exact values given by curve (*b*) for very small n. The asymptotic values depart greatly from the exact values for very small n. For $n = 1$ the asymptotic value is double the exact value for the mean range.

DISTRIBUTION OF MAXIMUM RANGE OF EFFECTIVE ANNUAL PRECIPITATION FOR 14 RIVER GAUGING STATIONS OF THE UPPER COLORADO RIVER BASIN AND AROUND IT

Procedure

The cumulative curves of ΔY_i departures were plotted and the ranges for $n = 1, 2, 3, 4, 5, 7, 10, 15, 20, 25$ were determined for as many periods m of length n as they may be included in the total length N for each station, without overlapping of periods of length n. While the number m of R values were greater for R_1 (equal to

1. The authors (Anis and Lloyd, 1953) give the coefficient of equation (17) both as $\sqrt{2/\pi}$, and $1/\sqrt{2\pi}$, and A. A. Anis in two successive papers (*Biometrica*, vol. 42, 1955, p. 96-101, and vol. 43, 1956, p. 79-84) gives always the value of coefficient as $1/\sqrt{2\pi}$. The author of this paper has found out that to his approach the value of $\sqrt{2/\pi}$ of equation (17) was correct.

$N-1$), the number m decreased with an increase of n, so that for R_{25} this number was either one ($N < 50$) or two ($N \geq 50$) for each station.

The values R_i were divided by the standard deviation of Y series of the corresponding station, so that all values $\frac{R}{s}$ for all stations refer to a unique standardized variable with the variance $s^2 = 1$. This procedure enabled the pulling together of values R for given n of all 14 stations in one large sample (size 646 for R_1, and size 28 for R_{25}).

Results

The distributions of ranges for 14 stations, with all corresponding values pulled together, are given in Fig. 13, and for the first five values of n (1, 2, 3, 4, 5) the distributions of ranges of random series (Fig. 11) are plotted also.

The computed values of mean, coefficient of variation and skew coefficient are plotted in Fig. 14. The corresponding statistical parameters for random series are plotted in this figure also, and specifically: (*a*) mean and coefficient of variation of R-distribution for asymptotic range values; (*b*) mean, coefficient of variation and skew coefficient for five values of n, computed by numerical integration of range distribution; and (*c*) the mean, computed by equation (17).

Comparison of Y series and random series

Fig. 13 shows that the distributions of ranges for 14 river gauging stations for small n are very close to distributions of ranges of random series.

Though the values m (sample sizes for R distributions) are large for R_1 to R_5, the sampling stability of range is, however, relatively small. This is mainly due to the fact that the concurrent values of Y series (values for the same water year) of 14 stations in the Upper Colorado River Basin and around, pulled together, are not independent among them.

It can also be seen from the comparison of distributions in Fig. 13, that both the mean and the standard deviation of range distributions of Y series increase constantly with an increase of n, as it is the case with a random series.

Taking into consideration the following factors: (*a*) regional sampling; (*b*) errors in the computation of carry-overs from one water year to another; and especially (*c*) non-homogeneity of data (created by man-made changes in river basins), it can be assumed here that the distributions of range of the effective annual precipitations in the Upper Colorado River Basin are very close to those of random time series.

The comparison of statistical parameters of range distributions of Y series for 14 river gauging stations in the Upper Colorado River Basin and around it, with the statistical parameters of range distributions of random series shows that the mean values of ranges of Y series are nearly identical with mean values of range of random series, because the curves (2) and (4) of Fig. 14 are very close, at least for values $n = 2 - 15$.

The asymptotic values of mean range, curve (1) in Fig. 16, are much larger than the values of the mean ranges of Y series. The mean ranges computed by equation (17) approximate well the computed mean ranges of Y series.

The comparison of the coefficients of variation of range distributions of Y series with those of random series shows that the departures between them are not great, curves (5) and (7), Fig. 14, while the asymptotic constant value, curve (6), is much smaller than the observed values.

The comparison of skew coefficients of range distributions of Y series with those of random time series shows, in the limits of the sampling instability of the third statistical moment of distributions of Y series, that the closeness of two curves, (7) and (8), Fig. 14, is sufficiently good to derive the conclusion that even the skew coefficients are very close for the two range distributions.

RANGE DISTRIBUTIONS FOR SEVEN LONG-RECORD RIVER STATIONS

Dependence factor

Assuming that only the coefficients b_o through b_{m-1} (first m values) in equation (3) are significantly different from zero, then according to Cramer (1935)

$$r_k = \sum_{o}^{m-k} b_j b_{j+k} \Big/ \sum_{o}^{m-1} b_j^2 \qquad (18)$$

The value

$$D = \left(1 : \sum_{o}^{m-1} b_j^2\right)^{1/2} \qquad (19)$$

is defined here as the dependence factor of a time series. Then (Yevdjevich, in preparation)

$$D^2 = 1 + 2 \sum_{1}^{m-1} r_k \qquad (20)$$

With $D = \sqrt{1+2r_1}$, when only r_1 is significantly different from zero, and approximately $D^2 = 1/(1-r_1)$ when several other r_k values are significantly different from zero.

Dividing the mean range values, or other parameters of range distributions by D, the series of different dependence factors may be compared.

Comparison of parameters of range distribution for seven long-record stations with random series

Fig. 15 shows the mean relative range $\overline{R}$, divided by the product $D \times s$ (dependence factor of series multiplied by standard deviation of U or Y series), for seven

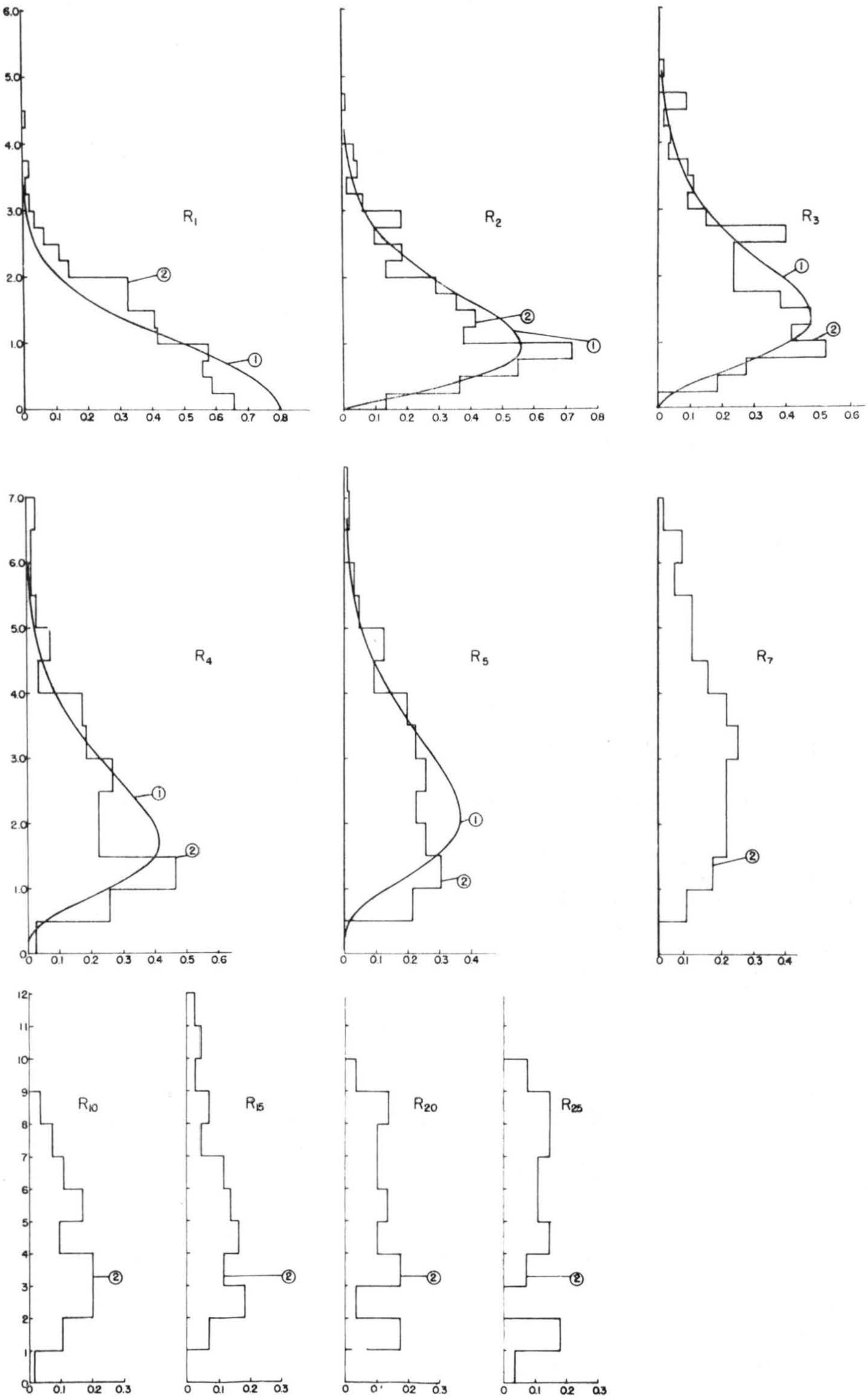

FIG. 13. Distributions of ranges for random time series (for $n = 1, 2, 3, 4, 5$), and for 14 river gauging stations from Upper Colorado River Basin and around it, for $n = 1, 2, 3, 4, 5, 7, 10, 15, 20$ and 25, for all 14 stations pulled together as a unique standardized variate of Y series (modular coefficients) of effective annual precipitations :
1. Random series.
2. Upper Colorado River Basin.

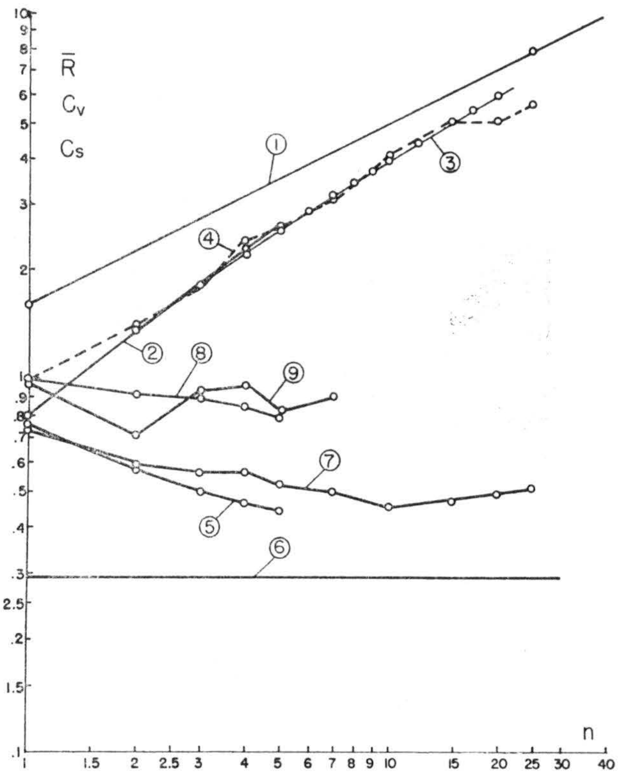

Fig. 14. Comparison of statistical parameters of range distributions for effective annual precipitations (Y series) of 14 river gauging stations in Upper Colorado River Basin and around it, with those of random time series :

1. Asymptotic values (1.6 *n*) of the expected mean of random series.
2. Exact values of means for random series.
3. Values of means for random series computed by equation (23).
4. Means of range distributions for 14 river gauging stations.
5. Exact values of coefficient of variation for random series.
6. Asymptotic constant value for coefficient of variation for random series.
7. Coefficient of variation of range distributions for 14 river gauging stations.
8. Exact values of skew coefficients for range distributions of random series.
9. Skew coefficients of range distributions for 14 river gauging stations.

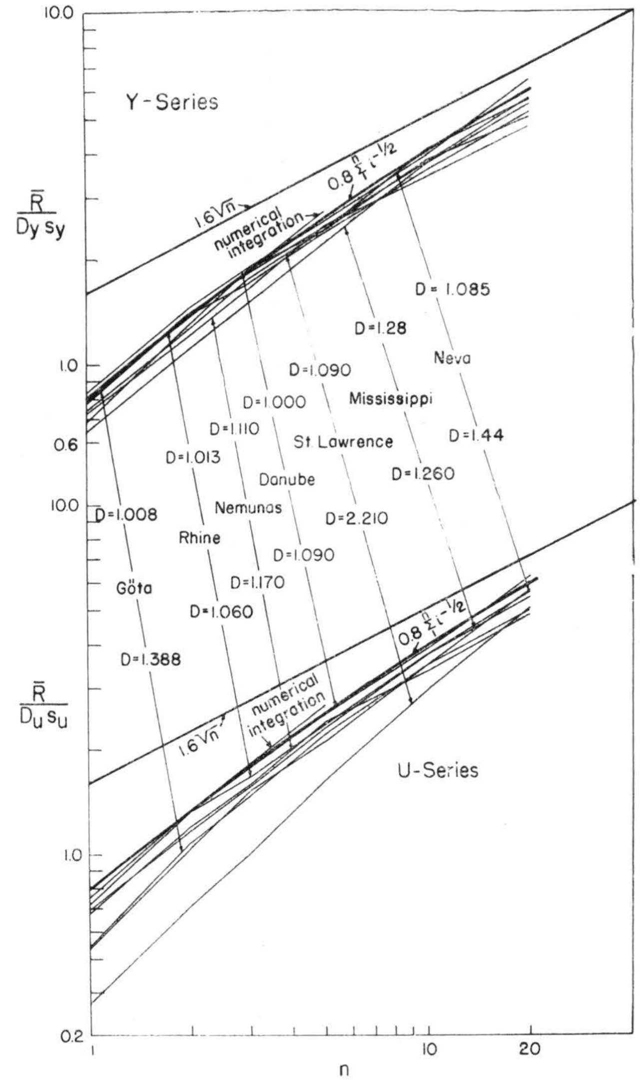

Fig. 15. Comparison of mean ranges divided by D × *s* (dependence factor multiplied by standard deviation) for seven rivers with long records, for both Y_i and U_i series, with the asymptotic mean values (Feller), mean values for small *n* (Anis-Lloyd), and for mean values determined by numerical integration of exact range distributions.

rivers in relation to *n*. The curves for these rivers are compared with series (D = 1): (*a*) asymptotic values of mean range, given by Feller; (*b*) mean ranges given by Anis-Lloyd formula, equation (17); and (*c*) mean ranges obtained by numerical integrations. The dependence factors are given for each river and U and Y series respectively.

The upper graph of Fig. 15, with relatively small D values (close to unity), clearly points out that the computed mean ranges for seven rivers are very close to the corresponding values of random time series for small *n*. For the Danube (120 years), the Göta (150 years), and the Rhine (150 years) it is difficult either to distinguish their mean ranges from the ranges of random series.

The lower graph shows that the greater is D (St. Lawrence) the greater is the departure of that series from the corresponding values of random series.

As distributions of U and Y series are skewed, the departures evidenced in Fig. 15 may be partly attributed to the skewness effect.

Fig. 16 shows a comparison between the coefficient of variation of range distributions, as function of *n*, with the same parameter of range distribution of random

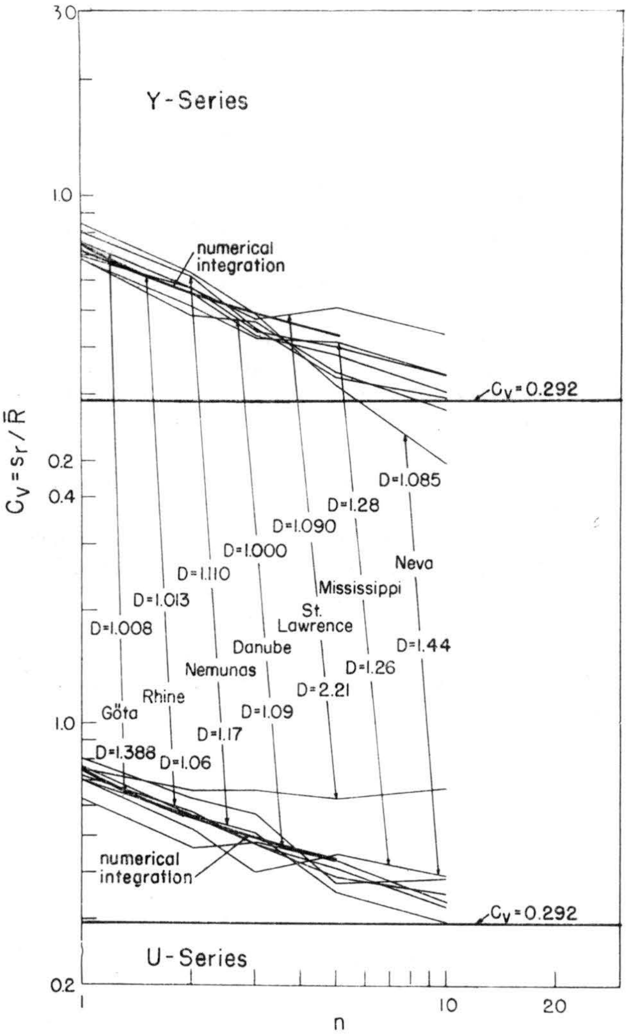

Fig. 16. Comparison of coefficient of variation of ranges for seven rivers with long records, for both Y_i and U_i series, with the asymptotic constant value $C_v = 0.292$ (Feller) and the coefficients of variation, determined by numerical integration of exact distribution of ranges.

series. There are two lines for random series: asymptotic constant value $C_v = 0.292$ of this coefficient, as given by Feller, and the values obtained by numerical integration of exact probability distribution functions. It points out that for small values of D the Y series approaches sufficiently close the random series.

GENERAL CONCLUSIONS

The analysis of hydrological characteristics in fluctuation of annual river flows and derived effective annual precipitations leads to the following conclusions:

1. Distribution of first serial correlation coefficient, correlogram analysis, and distribution of maximum range have shown that the sequence of effective annual precipitations is very close to random sequence.
2. Most of dependence between the successive values of annual flows can be explained: (*a*) by changing of water carry-over from one water year to another in the form of different water storage in a river basin; (*b*) by non-homogeneity in data; (*c*) by some systematic errors in compilation of annual flows; and (*d*) by error from regional sampling. After these factors are taken care of the room left for the causes of this linkage, which come either from the atmosphere or solar activities, remains small.
3. Dependence between successive values of effective annual precipitations can be explained: (*a*) by data inconsistency, data non-homogeneity, and error in computation of water carry-over in the corresponding river basin; (*b*) by an assumed complex relationship of evaporation and evapotranspiration to precipitation (non-linear relationship); (*c*) by a regression effect of stored moisture in atmosphere; (*d*) by regional sampling errors; and (*e*) by selection of the beginning of a water year.
4. There is no statistical evidence that the fluctuations of annual flows or effective annual precipitations may be composed of hidden periodicities, or of some regular patterns in the fluctuations, which can be extrapolated in the future with a reasonable expectancy that they would occur and would be verified by future flow records.
5. There is no evidence that the climatic factors as related to water resources have been changed significantly in the last 150 years.
6. Reliable forecasts of future annual flows (of order two to five years or more) by methods which are based on extrapolation or regular patterns in annual flow fluctuations (for instance, of hidden periodicities, or sun-spot cycles) do not seem possible.
7. Non-homogeneity (or inconsistency) in data of annual flows is an important hydrological characteristic of many river basins. It affects substantially the characteristics of available records of river flows, making the dependence of flows and effective precipitations greater than it would be without it. This is usually also the case with other series of climatic factors, related to water resources.
8. Carry-over of water from one water year to the next is an important hydrological characteristic of river basins, which greatly affects the linkage between the successive values of annual flows.

RÉSUMÉ

Étude des fluctuations climatiques au moyen des débits annuels et des précipitations annuelles effectives
(V. M. Yevdjevich)

La présente communication traite des fluctuations, dans les bassins fluviaux, tant des débits annuels que des précipitations effectives annuelles que l'on en déduit. Les valeurs des précipitations annuelles effectives (précipitation moins évapotranspiration) sont obtenues en corrigeant les débits annuels pour tenir compte des reports d'eau en fin d'année. Les données de base comprennent les valeurs des débits annuels enregistrés dans 140 stations fluviales de mesure situées dans de nombreuses régions du monde. Les caractéristiques statistiques des séries de débits annuels et des précipitations annuelles effectives sont rapprochées des caractéristiques de séries chronologiques aléatoires, ou de séries dérivées de séries chronologiques aléatoires. Les techniques appliquées consistent à établir des coefficients de corrélation, avec des corrélogrammes et les amplitudes des écarts accumulés à partir de la moyenne. La présente analyse indique qu'une partie importante de la différence qui apparaît entre les débits annuels observés et les séries chronologiques aléatoires peut s'expliquer par les effets de régression du report d'eau, de neige et de glace en fin d'année. Une autre partie de cette différence, en ce qui concerne les débits annuels et les précipitations annuelles effectives, peut être attribuée aux erreurs systématiques et au manque d'homogénéité des données et à des rapports complexes entre l'évaporation ou l'évapotranspiration et la précipitation. L'ensemble de ces facteurs explique en grande partie la différence constatée, un petit reste pouvant être considéré comme représentant l'effet d'autres facteurs qui peuvent influer sur la persistance des fluctuations climatiques. La suite des précipitations annuelles effectives est très proche des fluctuations aléatoires. Aucune donnée statistique ne montre que le climat, dans ses rapports avec les ressources hydrauliques disponibles, a changé d'une manière significative depuis cent cinquante ans, à en juger du moins par les fluctuations des débits fluviaux annuels.

BIBLIOGRAPHY / BIBLIOGRAPHIE

ANDERSON, R. L. 1941. Distribution of the serial correlation coefficient, *Mathematical Statistics Annals*, vol. 8, no. 1, p. 1-13.

ANIS, A. A.; Lloyd, E. H. 1953. On the range of partial sums of a finite number of independent normal variates, *Biometrica*, vol. 40, p. 35-42.

FELLER, W. 1951. The asymptotic distribution of the range of series of independent random variables, *Mathematical Statistical Annals*, vol. 22, p. 427.

HURST, H. E. 1951. Long-term storage capacity of reservoirs, *Trans. Amer. Soc. civ. Engrs*, vol. 116, no. 2447, p. 776.

SLUTZKY, E. 1937. The summation of random causes as the source of cyclic processes, *Econometrica*, no. 5.

YEVDJEVICH, V. M. *Fluctuations of annual river flows.* (To be published.)