Nordic Hydrology 3, 1970, 181-205
Published by Munksgaard, Copenhagen, Denmark

GROUNDWATER TIME SERIES
An Exercise in Stochastic Hydrology

ERIK ERIKSSON

International Meteorological Institute
in Stockholm

The term "stochastic hydrology" implies a statistical approach to hydrologic problems as opposed to classic hydrology which can be considered deterministic in its approach. During the International Hydrology Symposium, held 6-8 September 1967 at Fort Collins, a number of hydrology papers were presented consisting to a large extent of studies on long records of hydrological elements such as river run-off, these being treated as time series in the statistical sense. This approach is, no doubt, of importance for future work especially in relation to prediction problems, and there seems to be no fundamental difficulty for introducing the stochastic concepts into various hydrologic models. There is, however, some developmental work required – not to speak of educational in respect to hydrologists – before the full benefit of the technique is obtained.

The present paper is to some extent an exercise in the statistical study of hydrological time series – far from complete – and to some extent an effort to interpret certain features of such time series from a physical point of view. The material used is 30 years of groundwater level observations in an esker south of Uppsala, the observations being discussed recently by Hallgren & Sandsborg (1968).

DEFINITIONS AND CONCEPTS

A time series can be a series of evenly time-spaced observations on some natural object. This definition is probably not entirely unsatisfactory from a statistical point of view but this need not concern us here. If the series contains some

Contribution No. 198 from the International Meteorology Institute at Stockholm.

random elements, the variable making up the time series is considered a stochastic variable (*cf.* Hannan 1962, p. 1).

The concept of randomness in natural events requires some qualifications. Everybody familar with physical laws would be inclined to regard all events as strictly deterministic. However, since some events are extremely complicated from a physical point of view and details of them are even uninteresting, it is often of great advantage to use a simple physical model for the process itself to get, so to speak, a bird's-eye view, and to regard the deviations observed as due to random elements in the process. How far one can go in making up the simplified model then depends on the type of observations available.

A time series is called stationary if it has a finite variance and all covariances – means of products like $x_t\, x_{t-k}$ depend only on k not on t. This definition is rather strict since probably no time series in hydrology is stationary in this sense. It may therefore be convenient to refer to quasi-stationary time series, for instance when means of products vary periodically with t. By suitable transformation of the variables the series can be made strictly stationary. Time series which are not stationary are classified as evolutionary, a simple example being the random walk process which may picture diffusion processes.

The further discussion here will be limited to stationary or quasi-stationary series.

A stationary time series may contain deterministic components. A deterministic component is completely predictable. Trends are sometimes considered as deterministic components. This is useful as long as one has reason to believe that the trend will continue for a reasonable future period, making this part of the time series completely predictable. In many time series in hydrology one can suspect that trends are parts of very slow normal fluctuations, in which case they are certainly not deterministic although their prediction may be relatively reliable.

Time series may also contain periodic deterministic components like the daily variation and the seasonal variation. If the variation is not strictly sinusoidal, higher harmonics are also to be regarded as deterministic components as long as there are physical reasons for their existence. The run-off in temperate latitudes is often influenced by the fact that winter precipitation mostly occurs as snow. This will introduce a deterministic periodic component into the run-off rate which can only be described by several harmonics of the yearly period. Again, the term "deterministic" implies complete predictability.

The mean, trend, and periodic deterministic components can be regarded as linearly superimposed on a time series generated by some linear stochastic process. Thus, if they are subtracted from the series, the remainder represents entirely this linear stochastic process.

The term "linear" does not imply that the process in nature is strictly linear, only that the process is conveniently described as a linear process. The term "stochastic" does not imply that the process in a physical sense contains a random element, only that the physical process responsible for the variation is so complicated that it is of advantage to refer to a random element in a mathematical sense.

There are various kinds of linear stochastic processes, one being the moving average of a random independent variable. The type of linear stochastic processes useful as models in hydrology are processes which are called autoregressive or, more recently, linear Markov processes (see e. g. Quimpo & Yevjevich 1967). The simplest example is the so-called first-order linear Markov process reading

$$x_t = ax_{t-1} + \mu_t \tag{1}$$

where a is a constant less than unity and μ_t is a random independent variable. Hence, the random element enters into the time series through μ_t. The constant a is the autocorrelation for time lag one, also called first autocorrelation coefficient. It is an expression for the *dependency* of the process, a dependency which is akin to a memory effect.

As will be shown later this dependency can have a physical reality. It can, for instance, be anticipated that storage of water in a lake will introduce a dependency in the run-off rate even if the inflow rate is entirely random and independent. The same is true for a groundwater reservoir and for a number of different "reservoirs" of matter in nature. Hence, autocorrelation can be expected in many time series.

A slightly more complicated process has been called the second-order linear Markov process and is written

$$x_t = ax_{t-1} + bx_{t-2} + \mu_t \tag{2}$$

the only difference being that the dependence goes back two time lags. The constants a and b are functions of the autocorrelation coefficients for one and two time lags.

Also higher order linear processes can be formulated in the same way. The physical significance of the linear processes in eqs. (1) and (2) will be apparent from the following example. Consider the first-order linear differential equation

$$\frac{dx}{dt} + ax = b(t) \tag{3}$$

x representing a property of a system and b(t) an action from outside on the system. If this is written in finite difference form, it reads

$$x_t - x_{t-1} + \frac{a\Delta t}{2}(x_t + x_{t-1}) = \frac{\Delta t}{2}(b_t + b_{t-1})$$

or with terms rearranged

$$x_t = \frac{2 - a\Delta t}{2 + a\Delta t} x_{t-1} + \frac{\Delta t}{2 + a\Delta t}(b_t + b_{t-1}) \tag{4}$$

If now b_t, the forcing function, so to speak, should be random and independent, we could as well put the last expression equal to μ_t and obtain eq. (1), i. e. the first-order linear Markow process, in which r_1 is given by

$$r_1 = \frac{2 - a\Delta t}{2 + a\Delta t} \tag{5}$$

Starting with a second-order linear differential equation we should obtain eq. (2).

The linear stochastic processes can therefore be said to have the same physical foundation as the corresponding linear differential equations.

A stationary stochastic process will generate time series which show fluctuations or oscillations. It is, of course, tempting to describe these in terms of frequencies and this has led to the concept of the frequency spectrum distribution of a stochastic variable. Mostly the differential of this funtion is used and is called the frequency spectrum density which tells in what manner the total variance of a stochastic variable is distributed on various frequencies. For a random independent variable the variance is expected to be evenly distributed over all possible frequencies; this is referred to as a "white noise" spectrum. Further, for all stochastic variables which have no deterministic components the spectrum density function is continuous. Considering a process like that in eq. (1), most of the variance will be found in the low frequency part of the spectrum when r_1 is positive. Again the similarity between eq. (1) and eq. (3) is recognized since eq. (3) can describe the behaviour of an electric filter which damps high frequencies – a low pass filter. The second-order linear process in eq. (2) may behave similarly but may also concentrate the variance to a certain band of frequencies exhibiting a kind of resonance behaviour, all depending upon the magnitudes of the coefficients a and b.

Spectrum density functions are often computed from time series and can be said to give a condensed picture of the statistical properties of a time series. The same can, of course, be said about the so-called correlogram which shows the autocorrelation as a function of the time lag used, although the spectrum density function is to some extent more easily interpreted.

Since computed spectra represent samples of a time series, the spectrum density will show fluctuations which represent conditions in the time interval studied but which are not significant in the statistical sense. That is, if another part of the time series is studied, its spectrum density would not necessarily show the same pattern of fluctuations. The most widely used significance test for sampling fluctuations in computed spectral densities is that devised by Blackman & Tukey (1958), although there are other tests also available.

THE DATA

As mentioned earlier Hallgren & Sandsborg (1968) recently published a study of records of groundwater levels in an esker south of Uppsala, covering a period of 30 years. Records were also available on the water level of the adjacent river Fyris and on temperature and precipitation for the same period. All data were normalized to a time interval of 10 days (1st, 10th, and 20th of every month). Thus, groundwater and river data represent levels at these times whereas temperature represents averages for these periods and precipitation the collected amount for each interval. These data were entered on punched cards which were made available to the author.

The position of the observation wells is shown on the map in Fig. 1 where the outline of the esker is shown by the dashed line. A discussion of the features of the area is given by Hallgren & Sandsborg (1968). On an average the water level decreases from Well 1 and south, indicating a southward flow of the groundwater. A natural exit point is the place marked Ultunakällan on the map.

COMPUTATIONS

The computational procedure was designed to separate the various components of a stochastic variable in steps. If the time series as it existed is denoted $\{x_t\}$, the first step consisted in isolating the mean and the linear trend in the form

$$\alpha + \beta \cdot t \quad (t \equiv 1, 2, \ldots\ldots N)$$

by the least square procedure. Then a new series was formed by the transformation.

$$y_t = x_t - \alpha - \beta \cdot t \qquad (6)$$

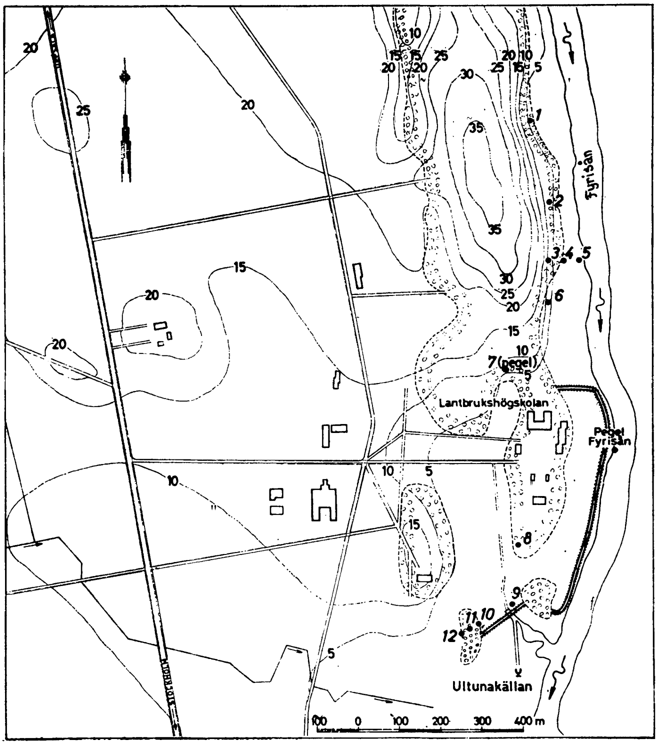

Fig. 1.

Map of the Ultuna area south of Uppsala. The esker is delinated by the dashed line. Elevation of contour lines is given in metres. Observation well numbers are shown by the italic numerals. Pegel Fyrisån refers to the measurement position of the river water level. (From Hallgren & Sandsborg (1968) reproduced by their kind permission.)

The next step consisted in estimating the periodic deterministic component, i. e. the deterministic seasonal variation by isolating

$$P(t) = A \cos \frac{2\pi t}{36} + B \cdot \sin \frac{2\pi t}{36}$$

using again the least square technique and, depending on the procedure, also higher harmonics. Then a new time series $\{z_t\}$ was formed by the transformation

$$z_t = y_t - P(t) \tag{7}$$

Now the first autocorrelation coefficient of $\{z_t\}$ was computed, i. e. the correlation between z_t and z_{t-1}. If this is denoted by r_1, a new time series $\{u_t\}$ was constructed by the transformation

$$u_t = z_t - r_1 z_{t-1} \tag{8}$$

This series now consisted of 1079 values of u_t and was used for computing covariances, i. e. $u_t u_{t-m}$ from $m = 0$ to $m = 109$, the covariance for $m = 0$ being, of course, the variance of the series. From the autocovariances the spectrum density estimates were computed according to the procedure suggested by Blackman & Tukey (1958) including the smoothing ("hanning"). Thus spectral density estimates were obtained for 109 frequency intervals covering a frequency range from 0 to 0.5 cycles/10 days i. e. from zero to 18 cycles/year. If this spectral density estimate is denoted by $s'(m)$, the spectral density of the time series $\{z_t\}$ is obtained from

$$s(m) = \frac{s'(m)}{1 + r_1^2 - 2 r_1 \cos \frac{\pi m}{109}} \tag{9}$$

which also was listed with $s'(m)$ and the autocovariances $R'(m)$ of $\{u_t\}$.

With respect to the separation of the seasonal deterministic component three procedures were used:

Procedure 1. Only the first harmonic was computed and subtracted.

Procedure 2. The harmonics 1 to 4 were computed and subtracted.

Procedure 3. The harmonics 1 to 18 were computed and subtracted.

Actually, in procedure 3 the subtraction was achieved by first computing the mean of y_t for a certain day of the different years, thus a mean of 30 values, then subtracting this mean from these variables, repeating the procedure for all the 36 observations of a year. Means were also listed and will be discussed later. Such a procedure was also used by Quimpo & Yevjevich (1967).

COMPUTED SPECTRAL DENSITIES

Fig. 2 shows a diagrammatic representation of the spectral densities $s(m)$ and $s'(m)$ computed from the data from Well 1 using procedure 1. Thus, the mean, trend, and the first harmonic of the seasonal variation have been removed from the spectrum. The ordinate is represented by a logarithmic scale. On the abscissa the corresponding m-value has been converted into frequency as cycles/year. It is seen that the spectral density $s(m)$ is concentrated to the low frequency

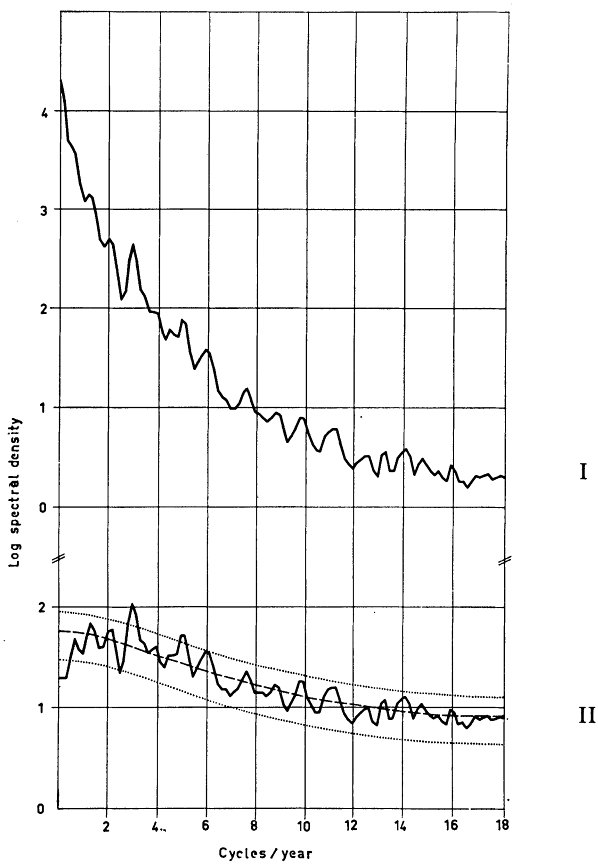

Fig. 2.

Estimated spectral density from observations of Well 1 (curve I) and spectral density after removal of the autocorrelation effect (curve II). Note that effects of mean, trend, and basic yearly harmonic were removed prior to the analysis. The smooth curve is the expected spectral density of a first-order Markov process with autocorrelation equal to that of the series for one time lag and the variance of the series (giving curve II). Dashed lines are the 90 per cent confidence limits computed according to Blackman & Tukey (1958).

end which can be anticipated since the first autocorrelation coefficient in this case is computed to be 0.974, thus rather high. When the effect of this autocorrelation is removed by the procedure described, the lower curve, s′(m), is obtained. Also here the spectral density is somewhat concentrated in the low frequency range. The first autocorrelation coefficient of the corresponding series, denoted r_1' ($= R_1'/R'_0$), is 0.44, thus not exceptionally high.

If one assumes that $\{u_t\}$ is generated entirely by the process

$$u_t = 0.44\, u_{t-1} + \mathit{u}_t \tag{10}$$

i. e. a first-order linear process, one can compute the expected frequency spectrum density of $\{u_t\}$ from

$$s'(m) = \frac{R_0' \cdot (1 - 0.44^2)}{1 + 0.44^2 - 2 \cdot 0.44 \cos \frac{\pi m}{109}} \tag{11}$$

where R_0' is the variance of the $\{u_t\}$ series. The function above has also been entered on the diagram as the dashed line. It is seen that it would represent the obtained spectrum density function fairly closely if it had been smoothed. This brings us to the spectral density fluctuations which occur in all such cases since it is a sample from a finite time series. If the time series used in the computation had been extremely long, all the frequences could have been better represented so the sampling fluctuations would have been smaller and we may have approached the dashed line. To test whether the variations in spectrum density are significant, i. e. would occur under any circumstances, or are sampling fluctuations, a procedure has been worked out by Blackman & Tukey (1958).

In this procedure so-called confidence intervals are constructed which will contain a certain fraction of the sampling fluctuations, all based on probality theory. Thus 90 per cent confidence intervals should contain, say, 90 per cent of the spectrum density. The 90 per cent confidence interval constructed around the spectrum expected from the first order linear process just considered is also shown in Fig. 2 by the dotted lines. It is seen that broadly speaking the computed spectrum is well contained within these limits so that one could – provided one believes that this significance test is proper – say that the spectrum density function computed from the data most likely is a sample of the spectrum density function represented by the dashed line. Looking closer at the curve there are a couple of places where we could suspect a significant deviation. One is at the very beginning, i. e. for very low frequencies where the spectrum density found seems to be too small compared with the expected. This is not unlikely since the removal of trend made previously certainly also affects the very low frequencies, i. e. removes part of these too. This is, of course, unavoidable so one should not pay too much attention to this discrepency.

The other possible deviation is at a frequency of 3 cycles a year where a rather pronounced peak appears. This is most likely the third harmonic of the deterministic component, the seasonal variation. Also the fifth harmonic may be represented but is not very pronounced.

As a conclusion we may say that curve II is likely to represent the spectrum density function of the linear process of the form

$$u_t \equiv 0.44\ u_{t-1} + \mu_t \tag{12}$$

where μ_t is a random independent variable of zero mean and variance equal to R_0' $(1\text{-}0.44^2)$. A somewhat better estimate of the process may be obtained by removing the effect of the 3rd and possibly the 5th harmonic of the seasonal variation although the improvement may not be striking. Considering the value of $r_1 \equiv 0.97$ we can further state that the linear process representing curve *I,* thus the time series after removal of mean, trend, and first seasonal harmonic, is of the form

$$z_t = 1.41\ z_{t-1} - 0.45\ z_{t-2} + \varepsilon_t \tag{13}$$

since $a \equiv r_1 + r_1'$ and $b \equiv - \dfrac{r_1'}{r_1}$

Here ε_t is a random independent variable of zero mean and the variance given earlier. The expected spectral density function of this process in the interval considered should be

$$s(m) \equiv \frac{R_0' \cdot (1 - 0.44^2)}{\left(1 + 0.44^2 - 2 \cdot 0.44 \cos \dfrac{\pi m}{109}\right)\left(1 + 0.97^2 - 2 \cdot 0.97 \cos \dfrac{\pi m}{109}\right)} \tag{14}$$

Hence, the stochastic process generating the observed data can be described by eq. 13 adding the mean, trend, and deterministic seasonal variation component.

Before finishing this case we should pay some attention to the fluctuations of spectral density which were classified as not significant. The fluctuations are, of course, real in a sense since they certainly are not entirely a result of computational procedures. They represent, of course, some smoothing of the real fluctuations since the computation procedure gives only average spectral densities for frequency intervals; but apart from this they would certainly not be generated by the computation procedure. Hence, they represents events during the time span studied and are thus a description of the course of events, admittedly in a rather unintelligible language. They are of some interest if the same pattern occurs in other time series, in which case we can suspect a considerable cross-covariance or, stated in other words, a spatial dependency. This is of special interest if we want to assess what a time series from an observation well represents in terms of area. This point will be covered in a second paper but we may already now make some qualitative comparisons between spectrum density functions from the different wells even if they are not significant for the infinite time series.

In order to demonstrate the nature of spectral density fluctuations further the data from observation Well 1 were divided in two portions representing 15 years each; thus one series containing data from the period 1938-1952 and the other period 1953-1967. These were analysed by procedure 1 except that the number of autocovariances computed was only half keeping their number to about 10 per cent of the total number of variables. The results are shown in Fig. 3 as the s′(m) spectral density, i. e. after removal of the effect of the first

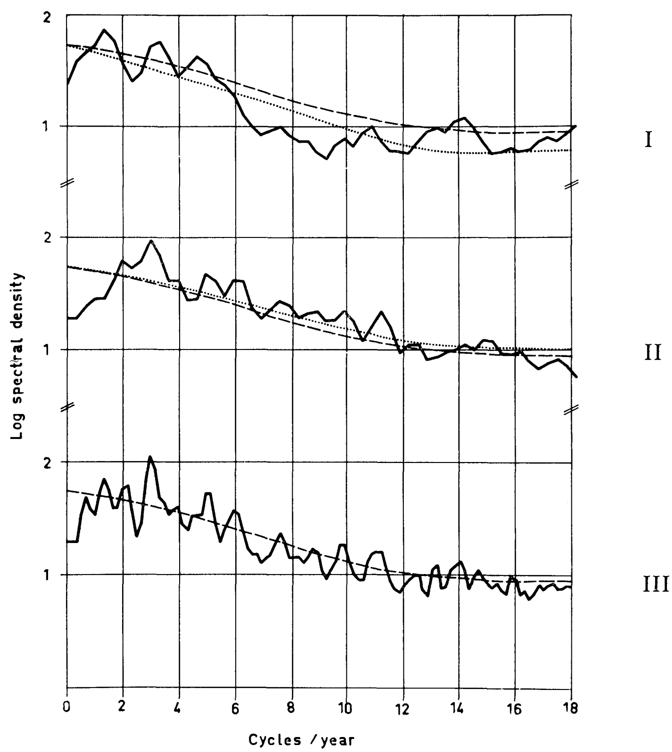

Fig. 3.

Spectral densities from data on Well 1 computed according to procedure 1. I is the spectral density of the whole series, 1938-1967, II the same for 1938-1952, and III for 1953-1967. The smooth lines show the corresponding expected spectral densities computed from the autocorrelation time-lag one of the corresponding series.

autocorrelation coefficient. In the Figure the corresponding curve from Fig. 2 is also entered. The expected s′(m) from Fig. 2 is entered on all three curves by the dashed lines. In addition the expected s′(m):s for the two 15-year series were also computed from the R_0' and R_1' as described and are shown by the dotted lines. As to the spectral density fluctuations, there is better correspondence between the 1938-1952 series and the complete series than between the 1953-1967 series and the complete series. This may be interpreted to mean that the complete series is to a considerable extent dominated by the "events" in the first 15-year period. Something seems to have happened during the second 15-year period – actually a change in the groundwater regime as also indicated by the difference in the expected spectral densities. According to Hallgren & Sandsborg (1968), this well is strongly influenced by the increasing water exploitation of the esker by the town of Uppsala. It is, however, not evident that this would effect the spectrum density function. There may have been other changes in the area, e. g. in the vegetation – being a forest – which have brought about this change.

What seems to have happened from a spectral density point of view is that the spectrum density in the high frequency range has decreased except at the very end. In the low frequency range there is hardly any similarity between the three spectra apart from the peak at 3 cycles/year which is most probably the third harmonic of the seasonal variation.

Comparisons of spectral densities computed by the three procedures mentioned is of interest for deducing the effect of possible higher harmonics of the deterministic seasonal variation. This was done on the time series from Well 1. The spectral densities s′(m) computed by the different procedures are shown in Fig. 4. The lowest curve was obtained by procedure 1, the middle one by procedure 2, i. e. removal of the effect of four harmonics, and the uppermost is obtained by procedure 3, i. e. after removal of 18 harmonics of the seasonal variation. It is seen that the procedure 2 mainly effects the 2 and 3 cycle/year periods which disappear. Very little is gained by removing higher harmonics – only the 5th is noticeably affected. The effects are probably better brought out in Fig. 5 which shows the differences for (A) procedure 1 minus procedure 2 and (B) procedure 1 minus procedure 3. Here the effects are brought forward more distinctly. If the third harmonic is regarded as significant, procedures 2 and 3 do not differ significantly. The pattern of spectrum density fluctuations is not affected by the various procedures. Of special interest in Fig. 5 are the negative (on a log basis) differences which occur and which are due to imperfections in the estimation procedure of the spectral densities. In this procedure a weighting function is introduced for assessing the average spectral density in the various frequency intervals. The properties of various weighting functions have been discussed

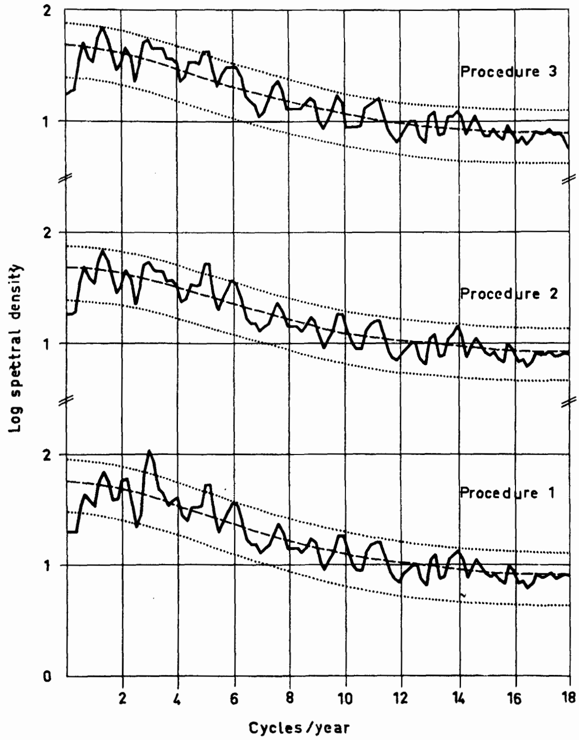

Fig. 4.
Spectral density from data on Well 1, from procedures 1, 2, and 3. In all cases the effect of mean, trend, and seriel correlation are removed. Procedure 1 removes in addition the 1st harmonic of the seasonal variation, procedure 2 the 1st to 2nd, 3rd and 4th harmonic, and procedure 3 up to the 18th harmonic of the seasonal variation.

thoroughly by Blackman & Tukey (1968). All weighting function which are of practical interest introduce some negative weight in the neighbourhood of frequency density peaks. It is these negative weights which show up as negative values in Fig. 5. For very strong peaks of spectral densities such errors would be noticeably large. It is therefore an advantage to make the computation of spectral densities on data where not only deterministic components but also strong dependency has been removed as in the present case when the effect of the first autocorrelation is removed. This procedure is known as "prewhitening" although generally a number of around 0.7 is used instead of the first au-

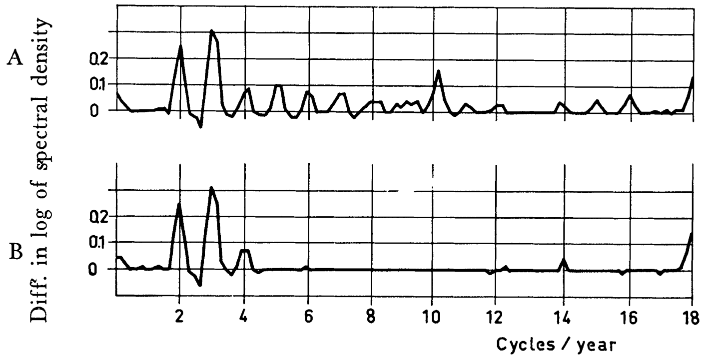

Fig. 5.
Difference between spectral densities of the curves in Fig. 3 A: curve II minus curve I. B: curve III minus curve I.

tocorrelation coefficient. However, since the value of r_1 is of special interest, one might as well use this.

To compare spectral densities from different wells, three were selected, Wells 1, 6, and 11 and procedure 3 was used. The corresponding $s'(m)$ functions are shown in Fig. 6 together with computed expected spectral densities and the 90 per cent confidence limits.

It is seen that the spectral density fluctuations of Wells 6 and 11 agree fairly well up to a frequency of about 10 cycles/year but for higher frequencies there is hardly any similarity. The intepretation of this is that any one of these wells would represent fluctuations of frequencies up to 10 cycles/year in the area; their representativity is thus good up to these frequencies. Higher frequencies are more "local". As to Well 1 there is some similarity with the others up to about 4 cycles/year after which the similarity becomes vague.

As to the expected spectral densities they seem to indicate a good agreement with the actual observations of the spectral density fluctuations. A few points are found outside the confidence limits, especially at about 3.5 and 4.5 cycles/ year but could be excessive fluctuations. To test this further the test designed by Grenander & Rosenblatt (1956, p. 195) was used on the spectrum of Well 6. This test, however, applies to the frequency spectrum, i. e. to the integrated spectrum density functions.

The integral was obtained by successive summing of the spectral densities, and the other parameter used in this test, *H*, was obtained by summing the squares of the autocovariances. Chosing again 90 per cent confidence limits

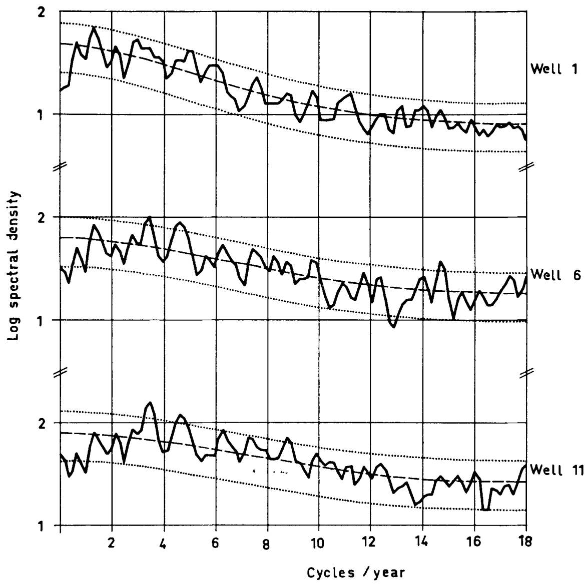

Fig. 6.
Spectral densities from data on Wells 1, 6, and 11 obtained by procedure 3 with computed expected spectral densities according to the same procedure and confidence limits used by Blackman & Tukey (1958).

around the expected spectrum computed from H and the variance, it was possible to judge the significance of the computed spectrum. The result is shown in Fig. 7 where the full-drawn curve is the corresponding expected spectrum computed from H and the variance, and the dotted lines are the 90 per cent confidence limits around this expected frequency spectrum. It is seen that the spectrum is entirely contained within the confidence limits.

Finally, similar spectral density functions as in Fig. 6 were computed for the River Fyris water level, for the total ten-day precipitation at Ultuna, and for ten-day means of air temperature at Ultuna. The results are shown in Fig. 8. Expected spectral densities and 90 per cent confidence limits around these, computed according to Blackman & Tukey (1958), are also shown, all the computations were done using procedure 3.

As to River Fyris the first autocorrelation coefficient became 0.85, indicating a rather strong dependency. When the effect of this dependency is removed, the first autocorrelation coefficient becomes 0.21 indicating as also seen in the

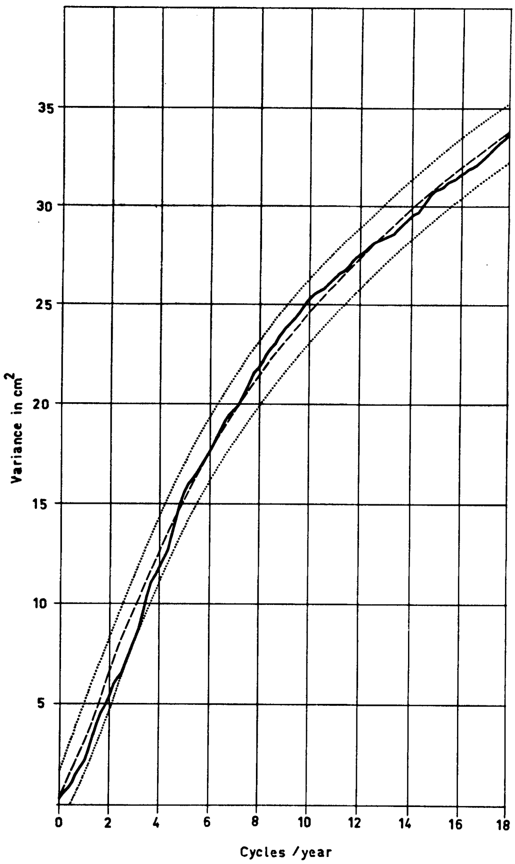

Fig. 7.

The frequency spectrum of Well 6 (i. e. the integrated spectral density) after the effects of deterministic components and autocorrelation time lag one are removed according to procedure 3; and the expected frequency spectrum computed from the autocorrelation time lag one and the corresponding time series. Confidence limit computed by the procedure given by Grenander & Rosenblatt (1956).

Figure a rather weak dependency although certainly significant. Thus, the river water stage seems to be a second-order Markov process as also found by many others (see Quimpo & Yevjevich 1967).

The precipitation records give a rather small autocorrelation – 0.12 – and when the effect of this is removed there is only a white noise spectrum left. The expected density of the remainder is thus a constant as indicated in the Figure.

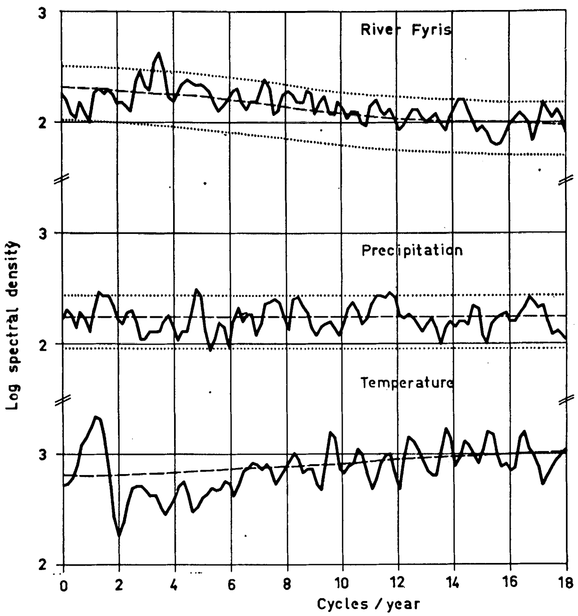

Fig. 8.
Spectral densities of data from water levels of river Fyris; 10-day sums of precipitation at Ultuna and 10-day means of temperature at Ultuna, computed by Procedure 3.

The fluctuations in spectral density are well contained within the 90 per cent confidence limits. The linear Markov process is thus simple and should read

$$x_t = 0.12\, x_{t-1} + \mu_t \tag{15}$$

the dependency being very weak. Precipitation summed over 10 days behaves very much like a random independent variable. From this point of view it should be very difficult to predict.

The poor predictability of precipitation is, of course, of great interest since precipitation represents an input to a hydrologic system.

Temperature behaves in a rather surprising way. The first autocorrelation is rather high – 0.67 – but when this effect is subtracted, the first autocorrelation coefficient r_1' becomes negative – 0.12. As seen in the Figure the corresponding spectrum density function, the dashed line, is not a good estimate of the computed function, at least not for low frequencies. The linear process describing the temperature fluctuations is certainly not simple: it may contain a number of dependency terms which would be difficult to interpret. The complicating factor seems to be a high density of low frequency oscillations – the peak being at a period of about 9 to 10 months which at first does not seem to make much sense.

If one converted the spectrum density function to s(m) which contains the whole stochastic process the maximum density at the period 9-10 months would become very pronounced. Since the trend in this particular series was practically nil, no bias has been introduced into the low frequency end of the spectrum which makes the feature puzzling.

The ordinary Blackman & Tukey significance test would strongly support the reality of the maximum. However, this is not the place to venture into this phenomenon although there is a peak in the groundwater density spectra at the same place. In a later section this peculiar spectrum density function will be discussed further.

The effect of the different computing procedures on the autocorrelation coefficients r_1 and r_1' is shown in Table 1. For the groundwater data, removal of higher harmonics of the seasonal variation increases r_1 slightly and decreases r_1' to a fair extent. It looks as if this removal brings the corresponding linear stochastic process closer to one of first order, i. e. the simpler type. For river Fyris and temperature the effect is the same. Precipitation, however, shows a decrease in r_1 as higher harmonics are removed.

To judge the justification of the degree to which deterministic components are removed before a study of the remaining stochastic process in carried out the mean values of the different dates of the whole period were computed and are shown graphically in Fig. 9. These curves correspond to a periodic function of 18 harmonics and their shape may therefore indicate whether it is reasonable to regard them as strictly deterministic, i. e. expected to have the same shape for any 30-year period apart from fluctuations due to statistical estimation errors. It is seen from the Figure that insofar as groundwater data are concerned the deterministic assumption is probably not too had although quite a good description would have been obtained with fewer harmonics, say four or five. For river Fyris the wiggles are unlikely to be deterministic. Also here fewer harmonics would have been appropriate. Finally, the prepicitation is definitely "overdone". A couple of harmonics would probably have been sufficient to describe the deterministic part of this series. Thus, in conclusion we can say that for groundwater and river water data procedure 3 is reasonable for removing the deterministic component; it can hardly introduce any bias in the estimates which could effect their interpretation.

Table 1.
Autocorrelations coefficients r_1 and r_1' with procedures 1, 2, and 3

Observation series	r_1			r_1'		
	Procedure 1	Procedure 2	Procedure 3	Procedure 1	Procedure 2	Procedure 3
Well 1	.9741	.9757	.9765	.45	.41	.42
Well 2			.9668			.36
Well 5	.9591		.9619	.44		
Well 6			.9548			.29
Well 8	.9200		.9240	.32		.25
Well 9	.9224		.9271	.32		
Well 10	.9144		.9170	.29		
Well 11	.9137	.9144	.9176	.34	.29	.28
Well 12			.9795			
Fyris river	.8537		.8532	.21		.17
Precipitation	.1388		.1249	.00		.00
Temperature	.6407	.6634	.6725	–.12	–.12	–.12

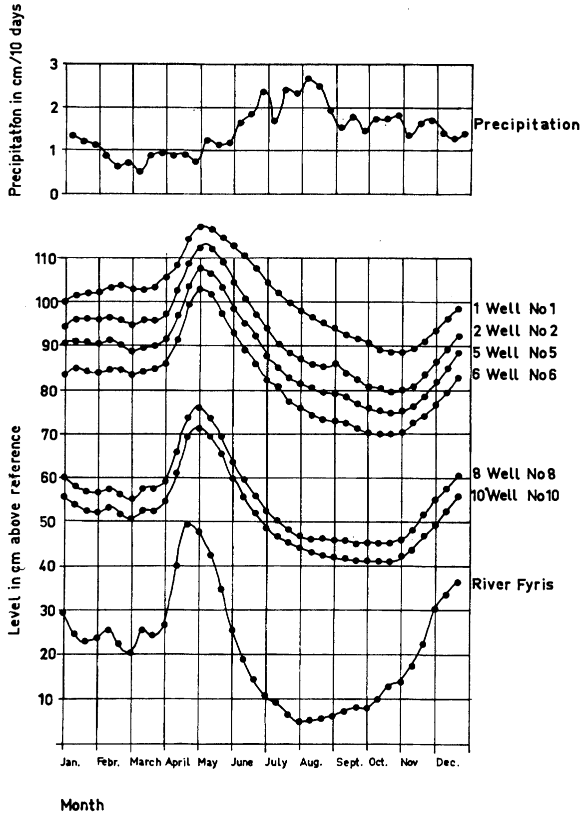

Fig. 9.
The "deterministic" part, except trend of time series, on river Fyris and on Wells 1, 2, 5, 6, 8, and 10 and precipitation obtained by procedure 3.

A SIMPLE STOCHASTIC MODEL OF GROUNDWATER RESERVOIRS

The simplest model of the present esker is that of a reservoir which is fed from infiltration water from above and drains at a rate which is proportional to the head of water (above the outlet). This would lead to a balance equation reading

$$p\frac{dH}{dt} = I - \alpha (H - H_0) \qquad (16)$$

where p is the porosity of the material, H the groundwater level, H_0 the level of the outlet, and α a proportionality coefficient. Written in finite difference form the equation reads

$$p\,(H_t - H_{t-1}) = \frac{1}{2}(I_t + I_{t-1}) - \frac{\alpha}{2}(H_t + H_{t-1}) + \alpha H_0 \tag{17}$$

which can be rearranged to read

$$(2\,p + \alpha)\,H_t = (2\,p - \alpha)\,H_{t-1} + I_t + I_{t-1} + 2\,\alpha H_0 \tag{18}$$

Considering deviations of H_t from a mean level, we put

$$H_t = \overline{H} - H_t' \qquad I = \overline{I} + I_t'$$

where the bars are used for means and the primes for deviations from the mean. Introducing this we get

$$(2\,p + \alpha)\,H_t' = (2\,p - \alpha)\,H'_{t-1} + I'_t + I'_{t-1} + 2\,\alpha\,(H_0 - \overline{H}) + 2\,\overline{I}\,. \tag{19}$$

But

$$2\,\alpha(H_0 - \overline{H}) + 2\,\overline{I} = 0 \tag{20}$$

so

$$H'_t = \frac{2\,p - \alpha}{2\,p + \alpha}\,H'_{t-1} + \frac{I'_t + I'_{t-1}}{2\,p + \alpha}. \tag{21}$$

If now I_t' is a random independent variable – we consider trends and deterministic components to have been removed from all observations – the equation above represents a first-order linear Markov process of the type

$$z_t = r_1\,z_{t-1} + \mu_t. \tag{22}$$

If I_t' should not be independent, we can write the equation

$$z_t = r_1\,z_{t-1} + u_t \tag{23}$$

and would find that

$$u_t = r_1'\,u_{t-1} + \varepsilon_t. \tag{24}$$

Considering precipitation we found a slight dependency (*cf.* Table 1) expressed by $r_1 = 0.12$. Because of storage of precipitation in the ground we could expect a somewhat greater dependency for I_t', probably of the same order as the r_1':s determined from the groundwater level series. This implies that $\frac{2\,p - \alpha}{2\,p + \alpha}$ should be interpreted as the corresponding r_1, the first autocorrelation coefficient. Hence, we can attempt to evaluate $\frac{2\,p - \alpha}{2\,p + \alpha}$ by equating it to r_1. Now only one of these variables can be determined. Hence we must assume a likely value of, say, p which is easier to guess than α. Likely values in this aquifer may range between 0.2 and 0.3. Computing α for these values will thus give a likely range of α:s. We find that α can be expressed by

$$\alpha = 2\, p \frac{1 - r_1}{1 + r_1} \tag{25}$$

or expressing, for convenience, α per year

$$\frac{\alpha}{p} = 72 \frac{1 - r_1}{1 + r_1} \tag{26}$$

where α is fraction of head/year.

From eq. (20) it is seen that the mean infiltration rate is given by

$$\overline{I} \equiv \alpha(\overline{H} - H_0)$$

Thus, if H_0 can be determined and $\overline{H}$ is known, $\overline{I}$ can be computed. As to H_0 this is somewhat difficult to assess. The mean river stage is 21 cm above the reference level and since the spring "Ultunakällan" empties into the river if must be somewhat higher. A value of 30 cm was therefore chosen for H_0. Table 2 summarizes the computations of α/p and α for the alternatives $p = 0.3$ and $p = 0.2$ and $\overline{I}$ based on the estimated heads $\overline{H}$-H_0 again for the two alternatives of p. As seen the results now in cm water/year are quite reasonable. The average precipitation for the 30-year series is 52.9 cm. Considering the coarse soil of the esker, an infiltration of 30 per cent of the precipitation would not be excessive which corresponds to a porosity of 0.2 as an average value. The other alternative, $p = 0.3$, may give too high a mean infiltration rate.

A somewhat smaller value of H_0, say, the mean river water level with $p = 0.2$, gives the mean infiltration rates shown in the last column of the Table which are somewhat more uniform than the previous ones and somewhat higher although not unreasonable in any way.

The computations made certainly demonstrate the nature of the stochastic process discussed as well as the possibility of setting up stochastic models in hydrology.

REMARKS ON THE MAXIMUM IN SPECTRAL DENSITY OF THE TEMPERATURE SERIES

As mentioned when discussing the frequency spectrum density function of the temperature series, a very pronounced – and consequently significant – maxi-

Table 2.

Computation of α and $\overline{I}$ for various values of p and H_0

Obs. well	r_1	α/p	α (p = 0.2)	α (p = 0.3)	$\overline{H}$	$\overline{H}$-H_0 (H_0=30)	$\overline{I}$ (p = 0.2)	$\overline{I}$ (p = 0.3)	$\overline{I}$ (p = 0.2) (H_0=20)
1	0.977	.84	0.168	.252	101	71	17.9	17.9	13.4
2	0.967	1.21	0.242	.363	93	63	15.2	22.8	17.4
5	0.962	1.40	0.280	.420	88	58	16.2	24.4	18.7
6	0.955	1.66	0.332	.498	82	52	17.3	25.9	20.2
8	0.924	2.84	0.568	.852	56	26	14.8	22.2	19.9
9	0.927	2.73	0.546	.819	54	24	13.1	19.7	18.0
10	0.917	3.16	0.632	.948	51	21	13.3	19.9	19.0
11	0.918	3.07	0.614	.921	51	21	12.9	19.4	18.4
12	0.920	3.00	0.600	.900	54	24	14.4	21.6	19.8

mum was found for the region 1.1-1.3 cycles/year. The implication of this is that through interference with the seasonal variation it would lead to slow oscillation of periods of up to 10 years. The reasons for such periods would depend on some inherent property of the atmosphere – possibly interaction between oceans and atmosphere – introducing oscillations of limit cycle nature with periods in the range of 9-10 months into the atmosphere.

Direct observation of such cycles would be difficult because of the strong seasonal temperature variation. However, the interference frequency 5 years is easier to locate. One example which happens to coincide almost completely with the present time span used is the length of ice coverage of the northern part of the Baltic, more specifically the part between Åland and the Swedish coast. Records on ice coverage were compiled recently by Hydroconsult AB for a different purpose and the results are shown graphically in Fig. 10 obtained through

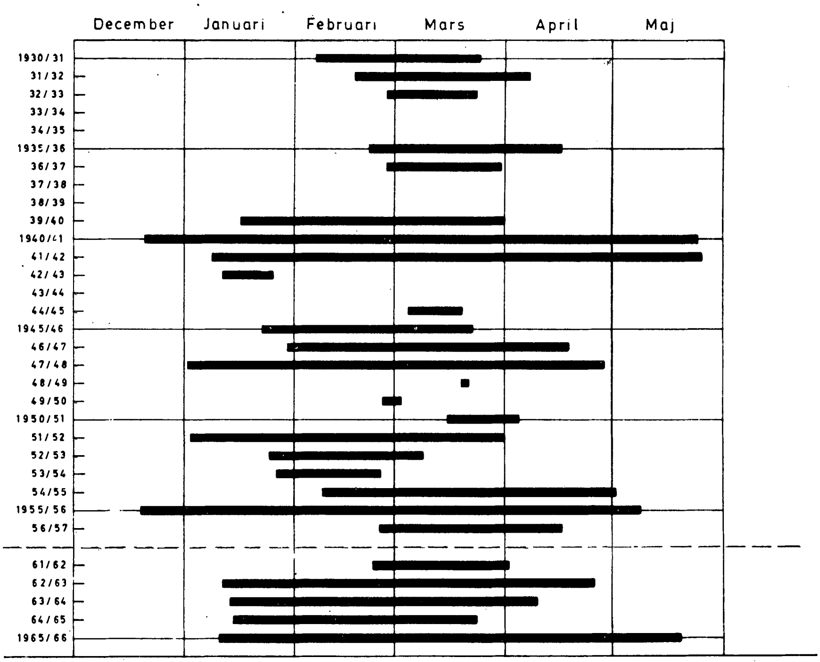

Fig. 10.
Ice coverage periods in the North Baltic area (Ålands hav).
(From Hydroconsult AB, Uppsala.)

the courtesy of Hydroconsult AB. It is seen that ice coverage varies rather periodically with maxima about every fifth year.

The type of interference expected from a 10-month period and a 12-month period is an alternation between the climatic types cold winters–warm summers and mild winters–cold summers.

With a time series of only 30 years it is impossible to increase the resolution in the frequency spectrum density in the region one cycle/year but longer temperature records would certainly be helpful in that respect.

ACKNOWLEDGEMENT

I am particularly indebted to Prof. G. Hallgren, Department of Agricultural Hydrotechnics, Uppsala, Sweden, for making the data on groundwater levels available.

REFERENCES

Blackman, R. B. & Tukey, J. W. (1958) *The measurements of power spectra from the point of view of communications engineering.* Dover Publ. Inc., New York.

Grenander, U. & Rosenblatt, M. (1956) *Statistical analysis of stationary time series.* Almquist & Wiksell, Stockholm.

Hallgren, G. & Sandsborg, J. (1968) Om grundvattenståndets variationer i Uppsalaåsen åren 1938-67. *Grundförbättring 21* (3), 85-131.

Hannan, E. J. (1962) *Time series analysis.* Methuen Co. Ltd., London, and John Wiley & Sons Inc., New York.

Quimpo, R. G. & Yevjevich, V. (1967) Stochastic description of daily river flow. *(Proc. Int. Hydr. Symp.* Sept. 6–8, 1967, Fort Collins, U.S.A.) Colorado State Univ., Fort Collins, Colo., pp. 290-297.