Statistical estimation of the ergodicity of the series of the average monthly runoff of the Desna River and its correlation characteristics

S. Moskalenko[1], L. Malytska[2]
[1] - Taras Shevchenko National University of Kyiv
[2] - Ukrainian Hydrometeorological Institute

Introduction

In hydrological studies, in order to identify in a more general form the patterns inherent in the time series of hydrological observations, the methods set forth in the theory of random functions are used. Random functions of time are sometimes called random processes. These methods are increasingly used in hydrological time series analysis and are intended to determine the structure of the location of hydrological data in the time of their occurrence. Among the hydrological problems solved using the methods of the theory of random functions: analysis of the structure of long-term variability of various hydrological characteristics, study of probabilistic characteristics of fluctuations, extrapolation of time series in order to predict river water flow and various geophysical processes, study of turbulent fluctuations of velocities, pressures and other parameters of water flow and the like (*Sikan, 2007*). When performing hydrological calculations or forecasts, it is necessary to clearly understand which mathematical model is used to describe the likely structure of hydrological sequences. If we assume that the sample under study is representative, then it can be used to make a certain satisfactory idea of the entire general population as a whole (*Kaisl, 1972; Babak et al., 2017*).

In the presented work, the object of research is the Desna River (near the hydrological post of the Chernigov city). The subject is the average monthly water discharges, which are considered as one rather long implementation of a random process. The purpose of the work is to conduct a study of the average monthly water discharge of the river, as an ergodic periodically correlated random process.

Method

The theory of random processes (functions) is a mathematical science that studies the patterns of random events and phenomena in the dynamics of their development. In the class of random processes, a subset of ergodic processes is distinguished, which are a constructive model for statistical processing of observational data. The ergodic property of a random process is that each of its individual implementations is, as it were, a «characteristic representative» of the entire set of possible implementations, or one implementation of sufficient duration replaces many implementations during processing (*Sikan, 2007; Novitsky and Us, 2011*). The time series was considered as one realization of a random process. In other words, for a priori given ergodic process with unknown characteristics, it is necessary to determine their statistical estimates and thereby confirm or refute the statistical hypothesis about the adequacy of the model. This problem statement is a partial case of a more general task of construction (development, creation, justification) of a mathematical model based on observational data of various kinds (*Babak et al., 2017*).

The main characteristics of an ergodic random process $X(t)$ with a period of correlation T are the periodic functions of mathematical expectation $m_X^*(t)$ and dispersion $D_X^*(t)$:

$$m_X^*(t) = \left(\sum_{i=0}^{i=k-1} X(t+iT) \right) \Big/ k$$

$$D_X^*(t) = \left(\sum_{i=1}^{i=k-1} \left[X(t+iT) - m_X^*(t+iT) \right]^2 \right) \Big/ (k-1)$$

where k is an integer, $kT=N$, N is the length of the implementation.

Periodic functions of mathematical expectation and variance are important, but not exhaustive characteristics, since they do not allow taking into account the nature of the relationship between individual members of the implementation (*Khristoforov,1994; Sikan, 2007*). To take into account this relationship, another important characteristic is determined - the correlation function of a random process $r^*(t,\tau)$:

$$r^*(t,\tau) = \left(\sum_{i=0}^{k-1-\tau} \left[\overset{o}{X}(t+iT) - \overset{o}{X}(t+iT+\tau) \right] \right) \Big/ \left((k-1-\tau) \sqrt{D^*(t+iT)} \sqrt{D^*(t+iT+\tau)} \right)$$

where τ - the distance between the sections (time shift), $\overset{o}{X}$- centered random process, $\overset{o}{X} = (X(t) - m_X^*(t))$.

To test the homogeneity and stationarity of the studied sequence, we used methods of testing statistical hypotheses based on F- and t- statistics, which are called, respectively, statistical criteria F - Fisher's and t - Student's.

RESULTS

The Desna River is the largest in length and the second largest basin, the left-bank tributary of the Dnieper. According to the physical and geographical conditions of water runoff formation, its basin is an elevated, slightly undulating plain with a slight slope. Temperate continental climate type is dominated. The water regime of the Desna is determined by spring floods and summer-autumn and winter low-water periods. In the years with average water content, 55-60% of the river water runoff falls in the spring period, in the summer-autumn period, it is 20-25% and in the winter it is 10-15%.

Monitoring the water flow on the Desna River near the city of Chernigov began in 1884. For the period from 1884 to 1894, data on average monthly expenses are in many cases missing. From 1895 to 1919, the data on the average monthly runoff are quite complete, but there are some gaps in 1919. Therefore, for this study, we selected period from 1920 to 2015, when there were no gaps in the observations of the runoff.

Thus, for the study, a time series of data for 96 years was formed with a discreteness step of one month, that is, one implementation was compiled, which consists of 1152 indicators of average monthly water flow rates.

Before performing the tasks, it is first necessary to confirm the hypothesis of the stationary nature of the studied process, followed by confirmation of the hypothesis of ergodicity of the process. Table 1 presents the results of such verification at 1% significance level.

Table 1 Results of test for stationarity of the average monthly water runoff of the Desna River (significance level 2α = 1%)

Stationarity criteria	Analytical statistics	Empirical statistics for each month of the year											
		Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Fisher's, F	2,22	1,14	1,25	1,61	1,16	3,65	1,12	1,70	1,63	1,51	2,14	1,17	1,16
Student's, t	2,63	4,66	5,05	1,75	2,64	2,56	1,49	2,53	2,20	2,40	5,62	4,58	4,50

Hypotheses about the constancy of the variance of the average monthly runoff over time (according to the Fisher's criterion) are accepted in all months, except for May. An additional check of the stability of the mathematical expectation in time (according to the Student's criterion) is necessary to increase the reliability of the stationarity of the studied sequence in relation to such test. Hypotheses about the constancy of the mathematical expectation are accepted for the period from March to September and not accepted - from October to February.

To obtain statistical estimates of the characteristics of the process under study, averaging over time was carried out over one implementation with a step of one month, while the correlation period was taken to be twelve months. In order to estimate the mathematical expectation and dispersion of a random process, we calculated these characteristics for each of the twelve months. Table 2 shows the characteristics of the investigated random process - the function of the mathematical expectation and the dispersion of the average monthly water discharge for the Desna River (the city of Chernigov) in the interval of twelve months with a discreteness step of one month, and in Fig. 1 graphical display of such assessments. The proposed in Fig. 2, the model of a random process was compared with the actual functions of the average monthly water discharge for a number of previous years. In principle, it well emphasizes the ergodicity of the process under study.

Table 2 Functions of mathematical expectation and dispersion of average monthly water consumption (m3/s) in the interval of twelve months within the framework of the model of random process, Desna river (Chernihov)

Functions	months											
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
$m_X^*(t)$	193	195	320	969	808	298	194	165	152	170	206	196
$D_X^*(t)$	8385	10995	37734	290870	196010	19182	9512	5064	5226	7613	8755	7070

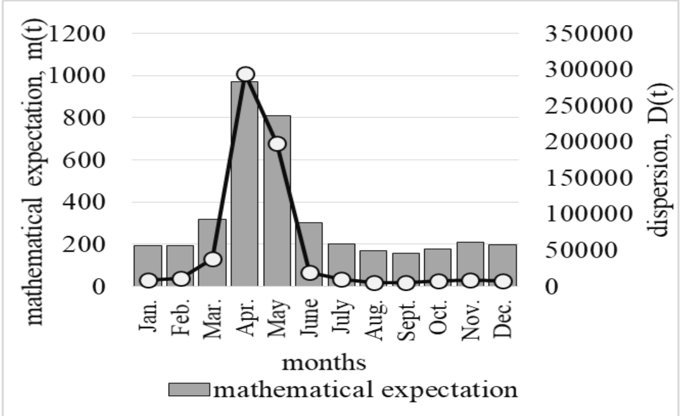

Figure 1. Estimates of the mathematical expectation and dispersion of the average monthly water consumption in the interval of 12 months; Desna River (Chernihov)

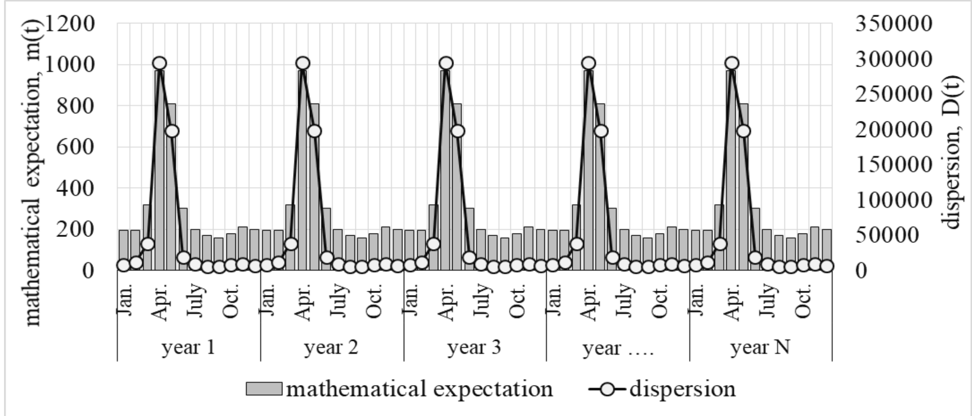

Figure 2. Estimates of the mathematical expectation and dispersion of the average monthly water flow in the N-year time interval within the random process model; Desna River (Chernihov)

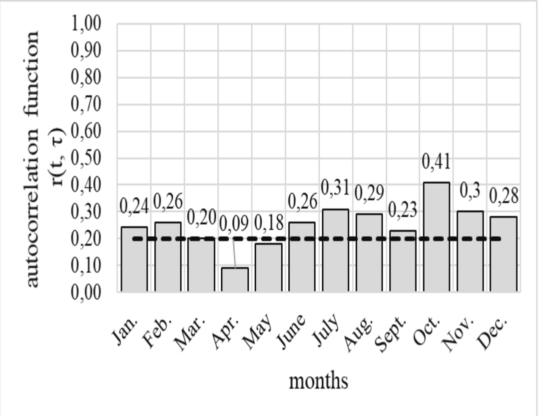

Figure 3. Correlation function r(t,τ) between the average flow of the current months from the corresponding months of the last year, Desna River (Chernihov)

Analyzing, from a hydrological point of view, the estimates of the mathematical expectation and dispersion (Fig. 1-2), we see that in the water regime of the Desna River the highest average monthly runoff (mathematical expectation) falls on April-May, which corresponds to the passage of spring floods in these months. For all other months, their average water runoff is several times less than in April-May, which is typical for summer-autumn and winter periods. Regarding the dispersion, the outlines of its change are similar to the mathematical expectation. If we calculate the coefficients of variation, which in hydrology serve as a measure of the variability of the series, then, in principle, they are in the range from 0.43 to 0.61. The highest values are in March-May (0.55-0.60).

Since in the practice of hydrological calculations the main issue is how present depends on past, not the future, we selected from the correlation matrix the corresponding rows - the ordinates of the autocorrelation function of the average monthly water discharge (Table 3). Thus, the autocorrelation function of a periodically correlated random process is a set of 12 autocorrelation functions for the average monthly water discharge of the Desna River near the city of Chernihov - for January, February, March, ..., December. Analyzing the ordinates (Table 3), it can be argued that, on the Desna River, relationships between the average monthly runoff with the index r * (t, τ) ≥ 0.20 are observed for the entire low-water period - from June to March of the next year. Depending on the magnitude of the shift for these months, the autocorrelation coefficient gradually decreases. For example, for the Desna River, with a shift of one month, the autocorrelation coefficients between the average water flow of the current month of the previous one change for different months from 0.63 to 0.86, and already with a shift of two months, the autocorrelation coefficients are in the range from 0.23 to 0.69, etc. (Table 3, Fig.3).

Table 3 Ordinates of the autocorrelation function of the average monthly water discharge *r*(t, τ)*, Desna River (Chernihov)

shift, τ	months											
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
0	1	1	1	1	1	1	1	1	1	1	1	1
1	0,80	0,77	0,63	-0,08	0,42	0,68	0,66	0,86	0,78	0,86	0,84	0,79
2	0,68	0,59	0,48	-0,18	-0,42	0,17	0,23	0,46	0,63	0,65	0,69	0,64
3	0,59	0,50	0,38	-0,1	-0,23	-0,16	-0,04	0,08	0,48	0,56	0,48	0,57
4	0,55	0,50	0,23	-0,08	-0,11	0,04	0,02	-0,11	0,14	0,43	0,42	0,44
5	0,40	0,52	0,25	0,05	0	0,21	0,23	0,07	-0,03	0,09	0,34	0,41
6	0,37	0,37	0,32	-0,01	0,2	0,28	0,33	0,2	0	-0,03	0,09	0,23
7	0,23	0,34	0,26	0	0,15	0,46	0,29	0,28	0,14	0,12	-0,01	-0,06
8	-0,06	0,24	0,17	-0,06	0,14	0,37	0,43	0,27	0,2	0,3	0,15	-0,04
9	-0,09	-0,08	0,12	-0,11	0,04	0,31	0,43	0,36	0,23	0,31	0,29	0,18
10	0,14	0,07	-0,14	-0,04	0,05	0,22	0,37	0,38	0,27	0,32	0,25	0,31
11	0,30	0,16	-0,14	0,18	0,11	0,24	0,34	0,31	0,31	0,36	0,29	0,27
12	0,24	0,26	0,20	0,09	0,18	0,26	0,31	0,29	0,23	0,41	0,3	0,28

Conclusions

Thus, the study proved that the average monthly water flow of the Desna River near Chernihov, if considered as one long-term implementation with a discreteness step of one month and a correlation period of twelve months, is an ergodic periodically correlated random process

Monitoring` 21
Kyiv, Ukraine

EAGE