

Methods of nonlinear dynamics as a hybrid tool for predictive analysis and research of risk-extreme levels

Elena Popova^{a,*}, Luís de Sousa Costa^b, Alfira Kumratova^a and Daria Zamotajlova^a

^a*Kuban State Agrarian University, Krasnodar, Russia*

^b*CIMO – Centro de Investigação de Montanha, Departamento Ambiente e Recursos Naturais Polytechnic Institute of Bragança, Campus de Santa Apolónia, Bragança, Portugal*

Abstract. The purpose of this research is to develop and adapt a complex of hybrid mathematical and instrumental methods of analysis and risk management through the prediction of natural time series with memory. The paper poses the problem of developing a constructive method for predictive analysis of time series within the current trend of using so-called “graphical tests” in the process of time series modeling using nonlinear dynamics methods. The main purpose of using graphical tests is to identify both stable and unstable quasiperiodic cycles (quasi-cycles). Modern computer technologies which allow to study in detail complex phenomena and processes were used as a toolkit for the implementation of nonlinear dynamics methods. Authors propose to use for the predictive analysis of time series a modified R/S -analysis algorithm, as well as phase analysis methods for constructing phase portraits in order to identify cycles of the studied time series and confirm the forecast. This approach differs from classical forecasting methods by implementing trends accounting and appears to the authors as a new tool for identifying the cyclical components of the considered time series. Using the proposed hybrid complex, the decision maker has more detailed information that cannot be obtained using classical statistics methods. In this paper, authors analyzed the time series of Kuban mountain river runoffs, revealed the impossibility of using the classical Hurst method for their predictive analysis and also proved the consistency of using the proposed hybrid toolkit to identify the cyclic components of the time series and predict it. The study acquires particular relevance in the light of the absence of any effective methods for predicting natural-economic time series, despite the proven need to study them and their risk-extreme levels. The work was supported by Russian Foundation for Basic Research (Grant No 17-06-00354 A).

Keywords: Nunoffs, floods, forecasting, time series, phase portrait, quasi-cycles, phase analysis, R/S -analysis

1. Introduction

Hitherto none of the published works does not give an idea of floods on the scale of the globe. The classification of floods according to the scale of social and environmental damage, the scientific basis for rational use of areas prone to flooding, the system concept of measures that need to be carried out in flood-hazardous areas in the periods before, during and after floods are also not developed.

The predicted climate warming and the inevitable growth of the further development of river valleys will undoubtedly lead to an increase in the frequency and destructive power of floods. Therefore, the urgent task is to develop effective measures to prevent floods and protect against them, since this will reduce the costs of eliminating the consequences of the disasters caused by them by 50–70 times. A set of measures in flood-hazardous areas, including forecasting, planning and implementation of activities, should be carried out before the onset of the flood, during its passage and after the termination of the natural disaster.

This study is devoted to the tools of predictive analysis and the study of risk-extreme levels in natural-

*Corresponding author: Elena Popova, Kuban State Agrarian University, 13 Kalimina Str. 350044 Krasnodar, Russia. E-mail: elena-popov@yandex.ru.

economic time series. The relevance of the subject of this study has acquired special economic importance after the floods that occurred along the river Kuban in June 2002, the losses from which are estimated at about \$ 0.5 billion.

In the conclusion of the commission [2] it is said that as a result of heavy precipitations that fell in the mountainous areas of the Kuban River Basin, a flood occurred. It has no analogue for almost a century of observations both in terms of maximum expenditures and rise in levels, and in damage caused people and businesses. As a result, 246 localities suffered in the south of Russia, more than 110 kilometers of a gas pipeline, 269 bridges, 1,490 kilometers of roads were destroyed, 102 people died. The total number of victims in the Southern Federal District reached 340 thousand people and the material damage exceeded 15 billion rubles.

As for damage directly to the regions, it should be noted that the damage to the Krasnodar Territory caused by the flood exceeds the material damage of Karachay-Cherkessia 10 times and is more than three billion rubles. 76 bridges were damaged in the Kuban (46 completely destroyed); more than 2.5 thousand households were destroyed; 58 settlements were flooded. The damage is especially high in the agro-industrial complex of the region, where 15 thousand hectares of arable land turned out to be under water; more than 100 thousand heads of poultry, 2.5 thousand heads of pigs were killed; cattle losses also were high. The amount of damage in agriculture is more than 250 million rubles.

The main focus of the study is on two tools of predictive analysis related to non-linear dynamics methods: fractal analysis of time series (TS) [4,5,12] and phase analysis of TS (see [5] for a comment on [12]).

2. Modified sequential R/S -analysis algorithm for estimating the depth of memory of a time series' beginning

It should be noted that the founder of fractal analysis is the British hydrologist Kh. E. Hurst. Exploring the statistics of river flows, he proposed a new statistical methodology to distinguish between random and non-random systems, the constancy of trends, and the duration of cycles, if any of them exists. Hurst showed that most natural phenomena, including river runoffs, temperatures, precipitation, sunspots, follow a "biased random walk", i.e. trend with noise. Trend stability and noise level can be assessed by how the normalized span

of TS changes over time. In other words, it shows how much the value entered by it (Hurst index, $H \in (0; 1)$) exceeds the value of 0.5.

If the levels of TS reflect a "purely random process" (they are independent random variables), then in accordance with the classical statistics for such a TS the value of the Hurst index is $H \approx 0.5$. The natural TS studied by Hurst and later by Mandelbrot et al. (precipitation, sunspots, tree rings, etc.) have a so-called long-term memory [4,5], so that for any of these TS Hurst index has a value that exceeds 0.7 (see Figure 4.1 in [6]).

It should be particularly noted that by the time the monographs [4,5] were published, the generally accepted opinion was already established that for the vast majority of natural systems the value of H deviates significantly from 0,5 to the black noise area [5].

This fact deserves special attention due to the fact that in the process of fractal analysis of the TS of the runoffs of the mountain rivers of the North Caucasus, the Hurst index H was found to be in the range of smaller than 0.5 values. The obtained empirical result can be explained by the fact that the considered natural series of TS do not have long-term memory, or the fact that the method of normalized scope (R/S -analysis) proposed by Hurst for calculating an estimate is not universal and not adequate for all natural and non-natural TS. One of the results of this paper is the confirmation of the second assumption formulated above. Hence the need to build a new author's approach for the implementation of the fractal analysis of TS is justified.

For more than half a century, the R/S -analysis of natural, economic, etc. TS was carried out by researchers on the basis of the algorithm published in [8], which in English literature is called "Hurst Standardized Range (SR) Method" (in [6] it is called " R/S -Analysis: A Step by Step Guide").

In this work, TS of monthly runoffs volumes in the upper reaches of the Kuban River (the geographical point of the beginning of the Great Stavropol Canal) for the period from 1926 to 2003 inclusive was investigated. A graphic representation of a part of this TS (from 1988 to 2003) is given in Fig. 1. Note that similar types of histograms are characteristic to TS of the runoffs volumes of other mountain rivers: Bolshoi Zelenchuk, Teberda, Aksaut and Marukha.

The authors carried out a fractal analysis (built H - and R/S -trajectories for time series) of the runoffs volumes of mountain rivers: Kuban, Bolshoi Zelenchuk, Teberda, Aksaut and Marukha.

Due to the lack of data in the war period (1940–1945) and taking into account the fact of climate warming [1,11], the original Kuban TS $Z = z_i, i = 1, 2, \dots, N$ was conditionally divided into three parts: the period from January 1926 to December 1939, the period from January 1946 to December 1987 and the period from January 1988 to December 2003. Let's denote them respectively by K_1, K_2 and K_3 .

A description of the computational scheme of the proposed algorithm for sequential R/S -analysis in order to estimate the depth of memory of TS beginning will be presented using the example of TS

$$Z = z_i, i = 1, 2, \dots, N, \tag{1}$$

which elements are TS K_3 data, i.e. further explanations and descriptions of the calculations are based on TS K_3 data.

First, in the considered TS its initial segments $Z^\tau = z_1, z_2, \dots, z_\tau, \tau = 3, 4, \dots, n$ are sequentially formed. For each of them the current average $\bar{z}_\tau = \frac{1}{\tau} \sum_{i=1}^{\tau} z_i$ is calculated.

Further, for each fixed $Z^\tau, \tau = 3, 4, \dots, n$ the accumulated deviation for its length segments t is calculated: $X_{\tau,t} = \sum_{i=1}^t (u_i - \bar{u}_\tau), t = \bar{1}, \bar{\tau}$. Then the span $R = R(\tau) = \max_{1 \leq t \leq \tau} (X_{\tau,t}) - \min_{1 \leq t \leq \tau} (X_{\tau,t})$ is calculated. It is normalized, i.e. represented as a fraction R/S , where $S = S(\tau)$ – standard deviation for TS segment $Z_\tau, 3 \leq \tau \leq n$.

In contrast to stage 7 of the Hurst's SR algorithm, in the present computational scheme, the estimate of the Hurst index is calculated on the basis of the "Hurst's empirical law", which in [5,8] is represented by the formula

$$H = H(\tau) = \frac{\log\left(\frac{R(\tau)}{S(\tau)}\right)}{\log\left(\frac{\tau}{2}\right)}. \tag{2}$$

Based on Eq. (2), two trajectories in Cartesian logarithmic coordinates are constructed. One of them, so-called H-trajectory, consists of points with coordinates (x_τ, y_τ) , where $x_\tau = \log \tau, \tau = 3, 4, \dots, n$ and y_τ is calculated according to Eq. (2): $y_\tau = H(\tau)$.

The second, so-called R/S -trajectory, consists of points with coordinates (x_τ, y_τ^0) , where $x_\tau = \log \tau$ and $y_\tau^0 = \log\left(\frac{R(\tau)}{S(\tau)}\right)$. For clarity, in the graphical representation of these trajectories each pair of neighboring points are connected by a segment.

Figure 2 shows the R/S - and H -trajectories resulting from the application of a sequential R/S -analysis algorithm to the TS of the Kuban River runoffs (for a graphic representation of this TS see Fig. 1). If the considered TS has a long-term memory, its R/S -trajectory

demonstrates the exhaustion of the memory of the beginning of the series by the so-called disruption from the trend or, in another terminology, changing the trend direction along which a certain number of initial points of the R/S -trajectory follow [5,6]. The above term "change of trend" implies that the points of the R/S -trajectory following the point of trend change already "do not return" to the original trend. Based on a massive computer experiment for numerous TS, the authors formulated the following definition of a trend-proof initial segment of TS, ending with an end to this trend:

1. A certain number of $l \geq 2$ points related to the beginning of the R/S -trajectory follow along a linear trend.
2. After the point l , the R/S -trajectory changes the trend, and the subsequent points of this trajectory "do not return" to the original trend.
3. The time series of ordinates $y_\tau, r = 1, 2, \dots$ of the points of the H -trajectory when moving from y_l to y_{l+1} gets a negative increment; the point of the l H -trajectory is in the zone of black noise, i.e. Hurst value $y_l = H(l) > 0.5$.

Note 1. When visualizing the R/S -trajectory and the H -trajectory, it is necessary to take into account that R/S -analysis does not compute the coordinates (x_τ, y_τ) corresponding to the first two levels of the investigated TS, i.e. for $\tau = 1$ and $\tau = 2$. Therefore, it is proposed to number the points of the indicated trajectories, starting with number 3 (see Fig. 2).

Note 2. In the general case, the trend change point of the R/S -trajectory appears with a lag, whereby the trend change point number of this trajectory is the upper estimate of the depth of memory about the beginning of the considered TS.

According to the above points 1–3 and notes 1 and 2 from the visualization of Fig. 4 it follows that the depth of memory about the beginning of the considered TS Z is estimated above by a number 16.

The most important conclusion, which follows from the established fact of the presence of long-term memory in the TS of mountain river runoffs, is that there is base for developing a system for the short- and medium-term forecast of the runoffs volumes of these mountain rivers. The system of short- and medium-term forecast is planned to be developed on the mathematical apparatus of the theory of cellular automaton and the theory of fuzzy sets [3,14].

The amount of memory used by the cellular automaton and, ultimately, the complexity of the computational prediction scheme substantially depend on the

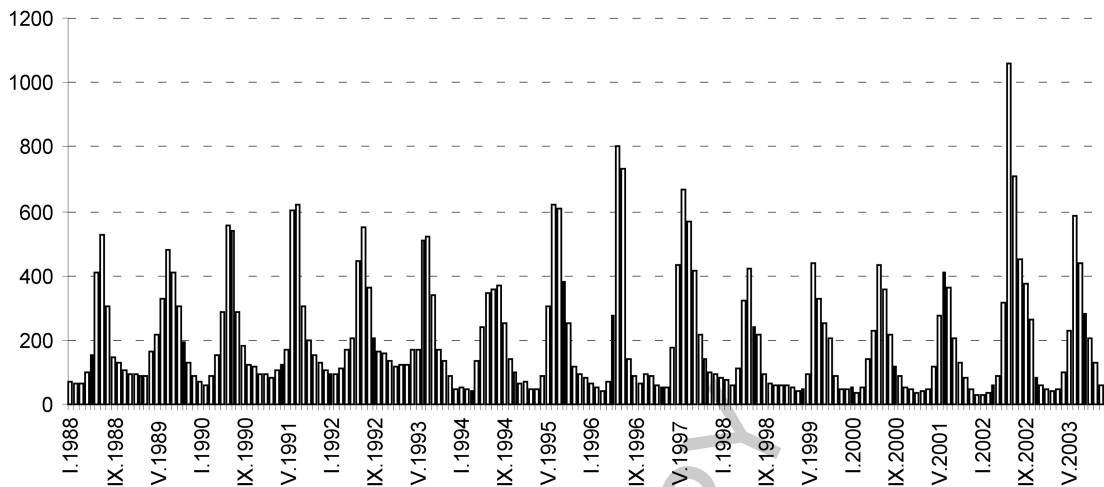


Fig. 1. Histogram of TS Z for the period from 01.1988 to 12.2003.

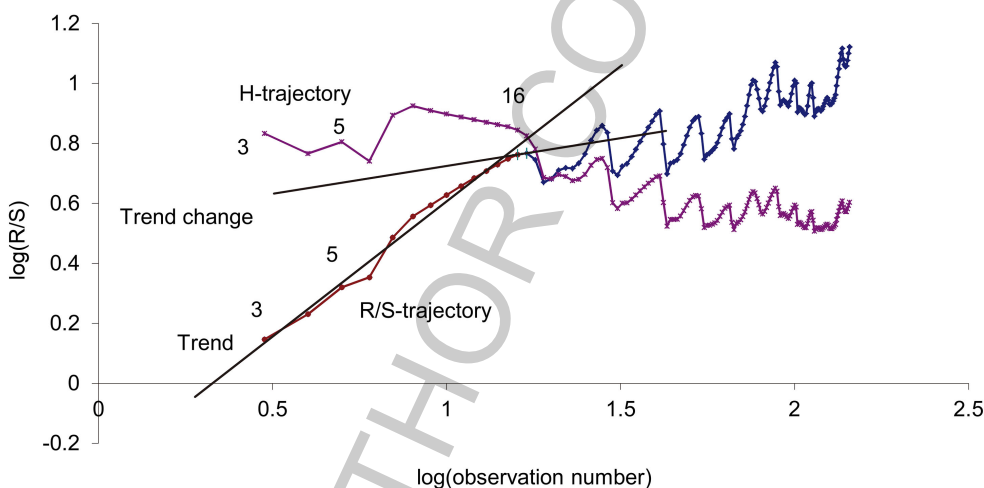


Fig. 2. R/S -trajectory and H -trajectory of Kuban River runoffs TS for the period from 1926 to 2003.

depth of the memory of the forecasting TS. Therefore, in the present work, numerical calculations are implemented with sufficient completeness in order to substantiate the upper estimate of the memory depth of the considered TS. These calculations in the present study are performed using the algorithm presented below, which determines the presence of such memory and evaluates its depth numerically, presenting it as a fuzzy set. The operation of the algorithm consists of two stages.

Stage 1. Formation on the basis of TS (1) of a family $S(z) = \{Z^r\}$, $Z^r = Z_i^r, i = 1, 2, \dots, n_r, r = 1, 2, \dots, m$, which consists of m TS, where by index i the elements of r series derived from $(r - 1)$ TS Z^{r-1} by deleting its first element Z_1^{r-1} . Here m is defined as the highest index value r , such that the $Z^m = z_i^m, i =$

$1, 2, \dots, n_m$ series still has a trend reversal point in its R/S -trajectory; the original TS (1) also belongs to the $S(Z)$ family in which it is assigned with the index value $r = 1$.

Stage 2. Realization of R/S -analysis of TS from the $S(Z)$ family and formation of a fuzzy set of values of the upper memory depth estimating the beginning of the series for each TS of this family.

For each TS $Z^r, r = \overline{1, m}$ as a result of its R/S -analysis, a R/S -trajectory and a H -trajectory are constructed. They determine the number l^r of such a point at which the trend has changed, i.e. l^r is the number of the first white noise zone point, where H -trajectory received a negative increment, and the R/S -trajectory changed the trend.

Table 1

Depth of memory for TS of mountain river Kuban monthly runoffs for the period from January 1988 to December 2003

Depth of memory	12	13	14	15	16	17	18	19	20	21
Amount of memory depth	1	6	12	16	17	17	9	9	8	5
Share of memory depth	0.01	0.06	0.12	0.16	0.17	0.17	0.09	0.09	0.08	0.05
Memory depth function	0.05	0.3	0.64	0.85	0.90	0.90	0.47	0.47	0.42	0.26

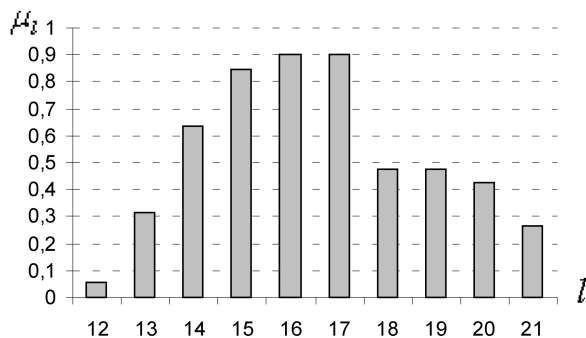


Fig. 3. Geometric representation of the FS depth of memory for TS Z.

Let’s give a description of the computational scheme of stage 2, which implements the fractal analysis of TS using the “algorithm of sequential *R/S*-analysis”. The following notation will be used: $N(l)$ – number of all series $Z^r = z_i^r, i = \overline{1, n}$ from the family $S(Z)$, each of which has a trend point number l^r that equals 1, $l^0 = \min_{1 \leq r \leq m} l^r, L^0 = \max_{1 \leq r \leq m} l^r; m = \sum_{l=4}^{L^0} N(l)$ – number of series from the family $S(Z)$; $d(l) = \frac{N(l)}{m}$ – the proportion of such series in $S(Z)$, each of which has a memory loss at depth l ; $L(Z) = \{l\}$ – the set of values of elements’ number of the trend change in the series of the family $S(Z)$; $M(Z) = \{(l, \mu(l))\}$ – fuzzy set (FS) of memory depth for TS (1) as a whole. Table 1 presents the result of the algorithm α_1 for TS Z.

Presented in the last row of Table 1 $\mu(l)$ – values of the function of l elements’ belonging to a fuzzy set $M(Z)$ are proportional to numbers $s(l), l \in L(Z)$; they are obtained by rationing the values of the shares $d(l)$ so that $\mu(l) < 1$ for every $l \in L(S)$. The result of stage 2 for TS Z is presented in Table 1. The values of the elements of the last row in Table 1 $\mu(l)$ are calculated as follows. First the maximum fraction $d^* = \max_{l \in L(Z)} d(l)$ is found (for example, in Table 1 value $d^* = 0, 17$) and the corresponding depth l^* ($d^*(l) = l^*$, in Table 1 value $l^* = 16$).

Further, for this depth l^* , the value of the membership function $\mu^* = \mu(l^*)$ is set by experts (value in Table 1 $\mu^* = \mu(16) = 0,9$). After that, for the remaining elements $l \in L(Z)$, the corresponding values of the membership function $\mu(l)$ are calculated by the for-

mula $\mu(l) = \frac{\mu^*}{d^*} d(l)$. The formation of FS $M(z)$ is carried out by pairwise combining the elements of the first and last rows of Table 1. The obtained depth assessment of TS Z memory is represented as the following fuzzy set:

$$M(Z) = \left\{ \begin{array}{l} (12; 0, 05), (13; 0, 3), (14; 0, 34), \\ (15; 0, 85), (16; 0, 9), \\ (17; 0, 9), (18; 0, 47), (19; 0, 47), \\ (20; 0, 42); (21; 0, 26) \end{array} \right\} \quad (3)$$

For clarity, Fig. 3 shows the geometric image of the FS of the considered TS Z.

The conclusions arising from the results of the calculations are as follows:

1. The memory depth of a particular TS is not a fixed number; its value changes along the considered TS, i.e. for its various segments, it is different; for example, as it can be seen from Table 1 for TS Z, the numerical values of the memory depth vary in the segment of the natural series 12, 14, ..., 21.
2. For the numerical representation of the memory depth of the considered TS (1) Z, the most appropriate is the mathematical apparatus of the fuzzy sets theory, i.e. the estimated depth is a fuzzy set $M(Z) = \{(l, \mu(l))\}, l \in \{l^0, l^0 + 1, \dots, L^0\}$, where l is the numerical value of the encountered memory depth, $\mu(l)$ is the value of the membership function for this depth.

As noted in [6], *R/S*-analysis is a tool for identifying cycles, both periodic and non-periodic. Figure 4 shows the image in the form of a histogram of the *R/S*-trajectories trend change points’ dynamics of TS $Z^r \in S(Z)$ for indexes r set in accordance with the months from January 1988 to April 1996.

Based on the results of the calculations, two notes on the new additional “cyclic” capabilities of the “*R/S* sequential analysis algorithm” compared to the “Hurst norm range algorithm” are formulated.

Note 3. To present the conclusions arising from the visualization of Fig. 3, we need to use the concepts of “cycle” and “quasi-cycle”. Let’s assume that Fig. 4 gives a graphical representation of the TS $L = l_i, i = 1, 2, \dots, N$. According to the generally accepted definition, the term “cycle” means a segment

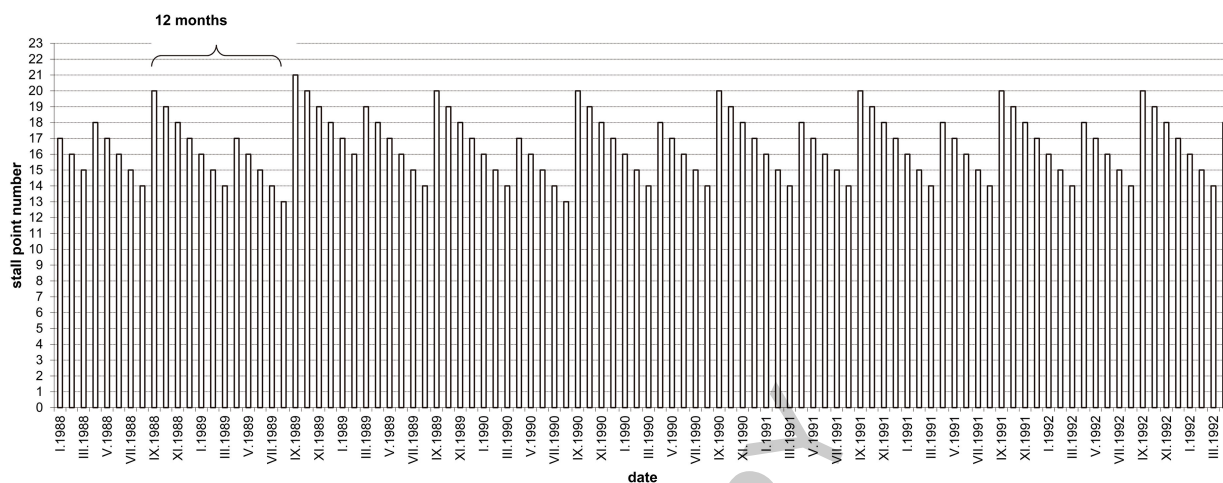


Fig. 4. Histogram of trend change numbers of R/S -trajectories of $TS Z^T$ from the family $S(Z)$.

$$l_i, l_{i+1}, \dots, l_{i+r}, \quad (4)$$

consisting of two parts; first part is increasing (decreasing), second – decreasing (increasing). Moreover, the specified segment (4) is locally maximal in its length ($r + 1$).

Generally, it can be said that the considered TS has the cyclic nature in case when this TS contains cycles of the same length. If TS contains cycles of different lengths, then it should be named “quasi-cycle”.

Presented definition of the cycle is not exhaustive and can be modified or supplemented. It is proposed to do so in case of analyzing the cyclical characteristics of the considered mountain rivers’ runoffs. First, note that the histogram in Fig. 4 consists of periodically repeating segments of the form

$$l_i, l_{i+1}, \dots, l_{i+11} \quad (5)$$

in each of which the element l_i represents a local maximum ($i = 9, 21, 3, \dots, 9 + 12k$), and the element l_{i+11} is a local minimum. In addition, each segment (5) consists of a quasi-cycle of the “decrease-increase” type and a quasi-cycle of the “increase-decrease” type, and these two quasi-cycles intersect, i.e. both contain a pair of common elements in the form of an intermediate minimum l_{i+t-1} and an intermediate maximum l_{i+t} , $t \in \{6, 7, 8\}$. Due to segments’ (5) structure (in the form of a union of two quasi-cycles) and length equality with value of 12, let’s call every segment of the form Eq. (5) by the term “cycle”.

Based on the above definitions, from the visualization of Fig. 2 we obtain the following statements.

Note 4. The information obtained at the output of the sequential R/S -analysis algorithm on the dynamics of

the trend change points of the R/S -trajectory not only establishes the fact of the cyclical nature of the evolution of the considered TS, but also gives the exact value of the boundaries of the cycles as part of this TS. We note in particular that in the considered TS of the flow of mountain rivers the indicated boundary points are September (beginning) and August (end). A pair of points (local minimum and local maximum), representing the intersection of two quasi-cycles that make up the cycle Eq. (5), fluctuates within three months (March, April, May).

Note 5. As mentioned in [5], to obtain sufficiently accurate estimates of the fractal characteristics of TS using the Hurst normalized span algorithm, series with a length of several thousand or even about ten thousand observations are required. The method of sequential R/S -analysis presented above essentially removes this condition, since it seems possible to use it to obtain estimates of the fractal characteristics of a limited segment of the considered TS.

It is also worth mentioning that the above notes 3–5 actually represent an element of verification of the proposed algorithm for sequential R/S -analysis. The presence of strict cyclicity is important, both in terms of pre-forecast analysis, and in terms of building predictive models.

The adapted method of sequential R/S -analysis is registered by the Russian Agency for Patents and Trademarks and has a certificate No. 2003611093.

Detection of long-term memory in mountain river runoffs is not a final goal of this research by itself. It should serve as an objective substantiation of the fundamental possibility of building a predictive model, in the course of which all significant factors are taken

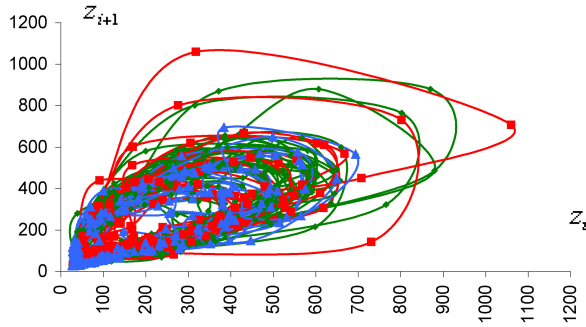


Fig. 5. Phase portrait of Kuban river runoffs' TS for the period from 1926 to 2003.

into account (for example, the presence of cycles of the form Eq. (4)) that caused this memory. In the context of the forecasting problem, it is appropriate to note the already established, i.e. which has become the classic main point of the decomposition analysis of time series [7]. According to this provision, in general, TS can be divided into 4 components: a) trend, b) cyclic component, c) seasonal fluctuation, d) irregular or residual component. In this case, the cyclical component can carry very significant information for the compilation of the forecast.

3. Phase portraits toolkit for identifying time series cycles and prediction confirmation

Studying the runoffs of mountain rivers, the construction of phase portraits of TS (1) in the phase space [5] of the dimension 2 $F(Z) = \{(z_i, z_{i+1})\}$, $i = 1, 2, \dots, n - 1$ is efficient. This type of phase trajectory mountain river Kuban runoffs' TS for the period from 1926 to 2003 is presented in Fig. 5. Note that for ease of visualization, the triple coloring of these phase portraits is chosen, which corresponds to time periods: from January 1926 to December 1939 (blue color), from January 1946 to December 1987 (green color) and from January 1988 to December 2003 (red color).

Note 6. Following Peters [5], Packard [12] and many other researchers (see references in [5]) for TS (1) as its phase space we use the simplest version of the form

$$\Phi_\rho(Z) = \{z_i, z_{i+1}, \dots, z_{n-\rho+1}\}, \quad (6)$$

$$i = 1, 2, \dots, n - \rho + 1$$

As known, constructing the phase space Eq. (6) for a specific time series, the issue of its dimension is of fundamental importance ρ . This dimension must be not

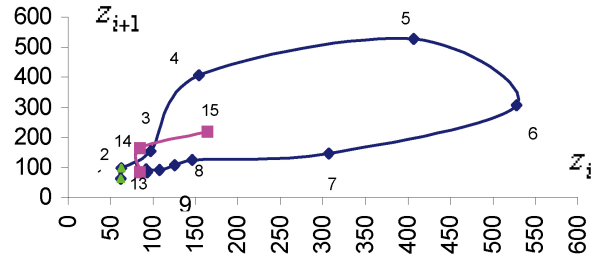


Fig. 6. The first quasi-cycle of the phase portrait, including the lag -13, 14, 15; points 1 and 2 are missing in the R/S -trajectory.

less than the dimension of the attractor of the observed series. In turn, as known, the fractal dimension C of its series can be used as the dimension of the attractor with a fairly acceptable accuracy. The value of this dimension, as noted in [5], is calculated by the formula

$$D = 2 - H \quad (7)$$

Since for the value analyzed in this work $H \in (0, 1)$, the estimate $D < 2$ is obtained [5]. Thus, for the purposes of this research, it is sufficient to use the phase space of dimension $\rho = 2$.

Let's consider this phase portrait as a trajectory, i.e. a sequence of points in which each adjacent pair is connected by a link (cut or curve). In this trajectory, we also distinguish its segments, which are called the term "phase quasi-cycles" or briefly "quasi-cycles". The definition of a phase quasi-cycle is in a certain sense close to the definition of a classical cycle. The difference between these two concepts is that the starting and ending points of the phase quasi-cycle do not have to be the same. The end point of a quasi-cycle is determined by its entry into the neighborhood of the initial point. In this case, self-intersection of the initial and final links of the quasi-cycle is allowed, if this leads to the maximum convergence of the initial and final points. In reality, there are such time series of evolutionary processes, in which the phase portraits contain such pairs of non-adjacent points (observations), in which the coordinates in the phase space actually coincide. The presence of such pairs of points actually destroys the cyclic structure of the phase trajectories.

A remarkable and very important feature of forecasting the mountain rivers' runoffs considered by the TS is that the phase portrait consists of a sequence of non-intersecting quasi-cycles, whose length is 12 months. In general, the trajectory of the phase portrait for a time series consists of 16 non-intersecting quasi-cycles C_r , $r = 1, 2, \dots, 16$. All quasi-cycles are built from February to January, thereby forming 12-month cycles (i.e. year).

Table 2
Dimensions of quasi-cycles and numbers of points of breakdown in the R/S -trajectories

C_k	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}
L_k	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
N_k	15	15	16	16	16	16	15	16	15	15	16	15	15	15	16

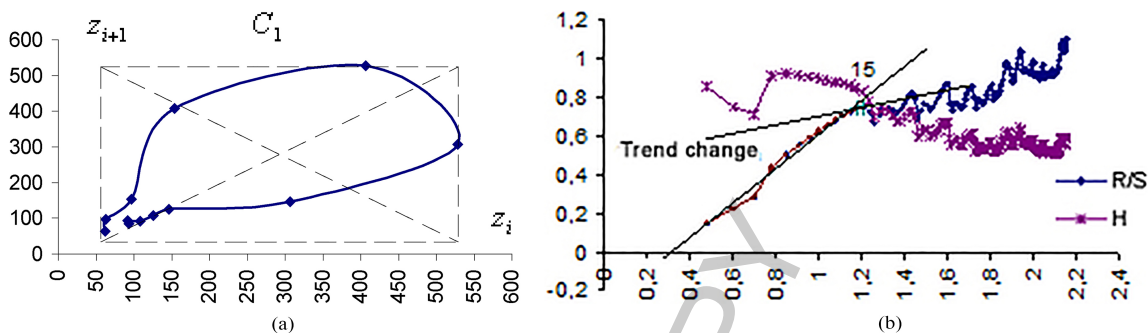


Fig. 7. The first 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

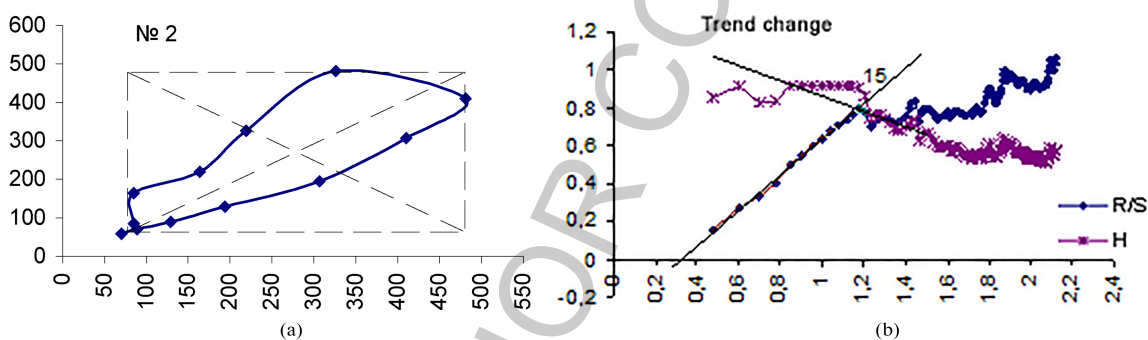


Fig. 8. The second 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

Note 7. In Fig. 7a, in terms of phase analysis tools, a separate annual cycle belonging to TS Z (1) is presented as a typical one. The R/S -trajectory and the H -trajectory presented in Fig. 7b demonstrate the exhaustion of the above cycle in Fig. 6. This exhaustion is demonstrated by the breakdown of the R/S -trajectory at a point $i = 15$ that exceeds by 3 the length of the annual cycle.

The latter is due to the existence of lags in the work of the sequential analysis algorithm. In this case, the size of this lag is 3. In Fig. 6, the indicated lag is represented by three dots colored pink; these are points 13, 14 and 15. The first two points painted green are not reflected in Fig. 7b due to algorithmic features the process of building an R/S -trajectory and H -trajectory.

The dimensions L_k of all 16 quasi-cycles and the number of the point of breakdown in the R/S -trajectories are presented in Table 2.

Let's denote by Z_k such segment of TS Z , which is obtained by removing from all points of observation

related to quasi-cycles C_1, C_2, \dots, C_{r-1} ; according to this definition $Z_1 = Z$.

Let's compare the memory depth of the considered TS, represented by a fuzzy set Eq. (5), with the dimensions of the quasi-cycles presented in the second row of Table 2. From this comparison it follows that the presence of long-term memory in the considered TS along with other factors is also due to the cyclic component of this TS.

Note 8. For each quasi-cycle the concept of "dimensional rectangle of a quasi-cycle C_r " should be defined. Through the points (z_i, z_{i+1}) with the maximum and minimum abscissa values (ordinates) straight lines parallel to the ordinate axis (abscissas) should be drawn. The intersection of these two pairs of parallel straight lines obtained (in Fig. 7a are represented by a dotted line) forms the dimensional rectangle of the quasi-cycle C_r . In other words, a dimensional rectangle is such a minimal convex hull of the points of a

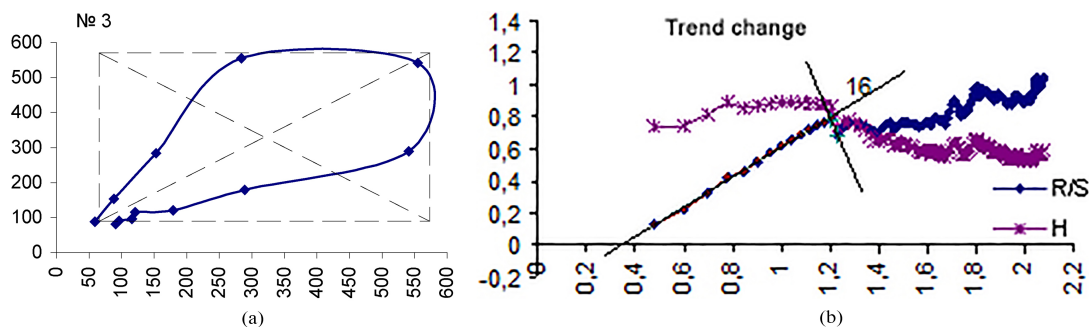


Fig. 9. The third 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

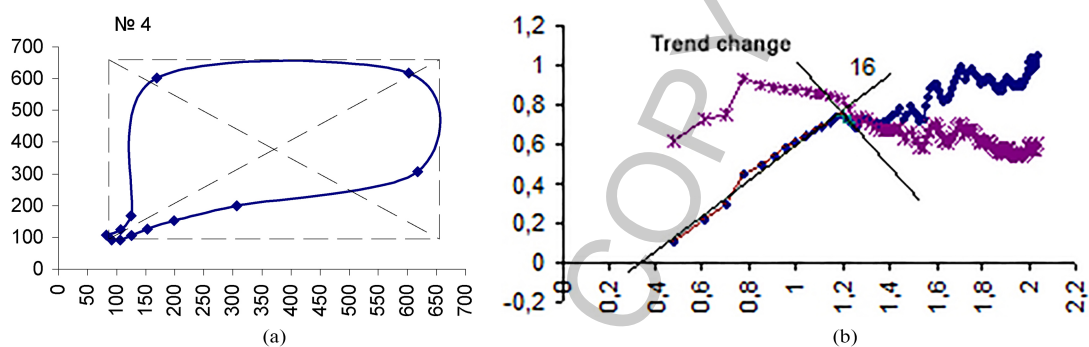


Fig. 10. The fourth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

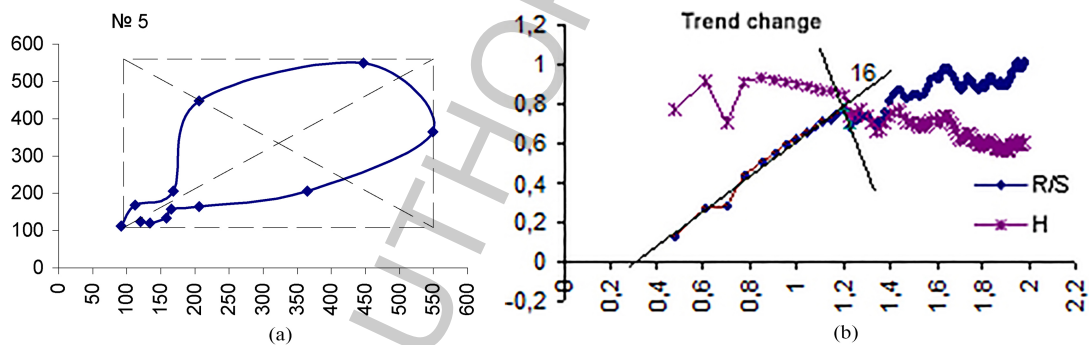


Fig. 11. The fifth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

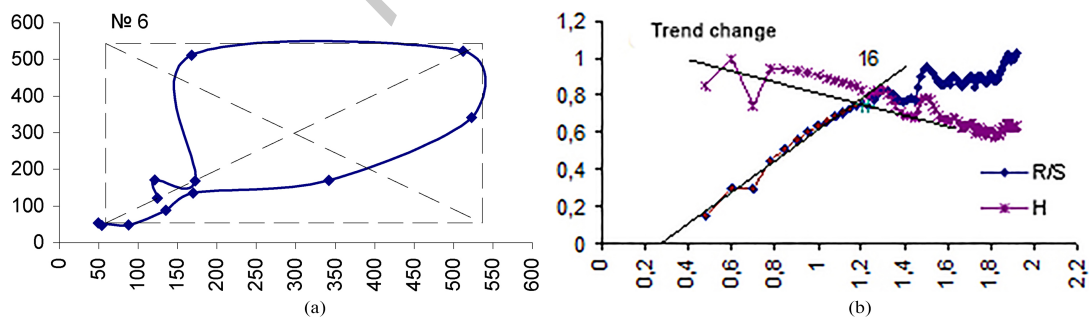


Fig. 12. The sixth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

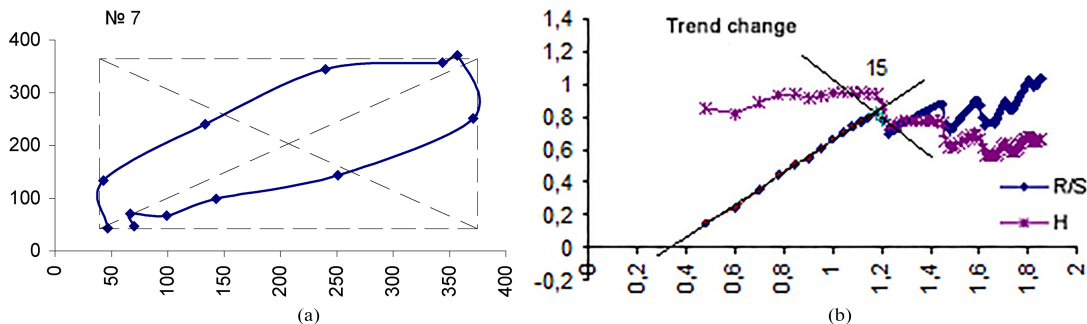


Fig. 13. The seventh 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

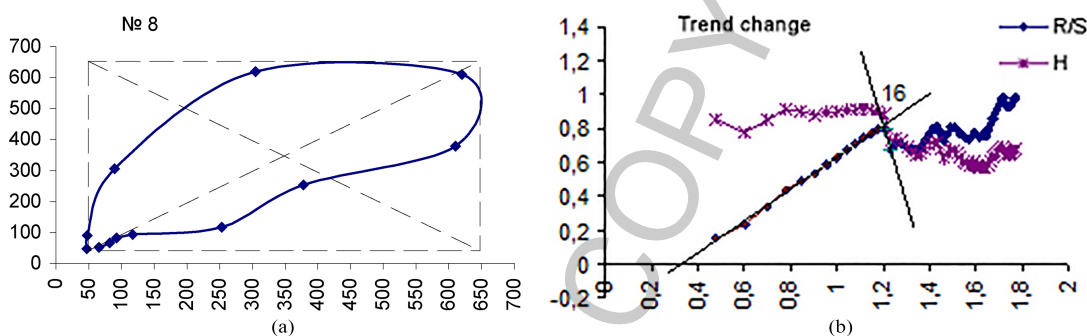


Fig. 14. The eighth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

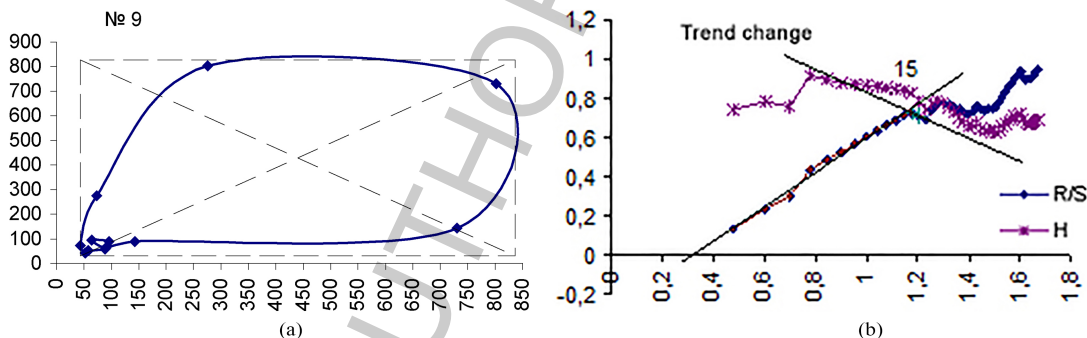


Fig. 15. The ninth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

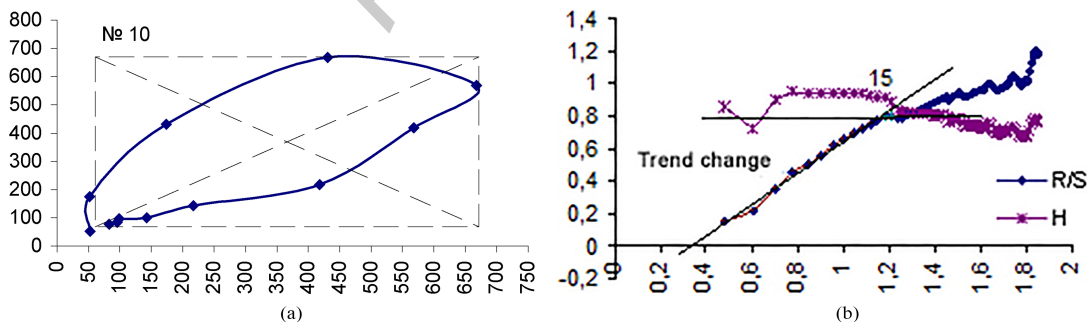


Fig. 16. The tenth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

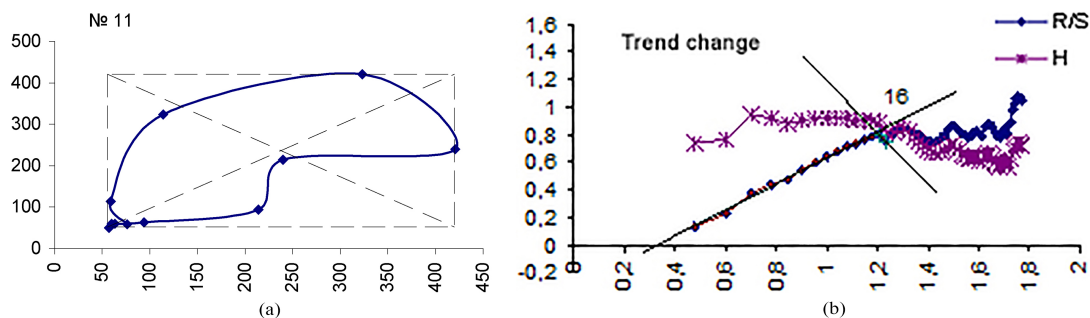


Fig. 17. The eleventh 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

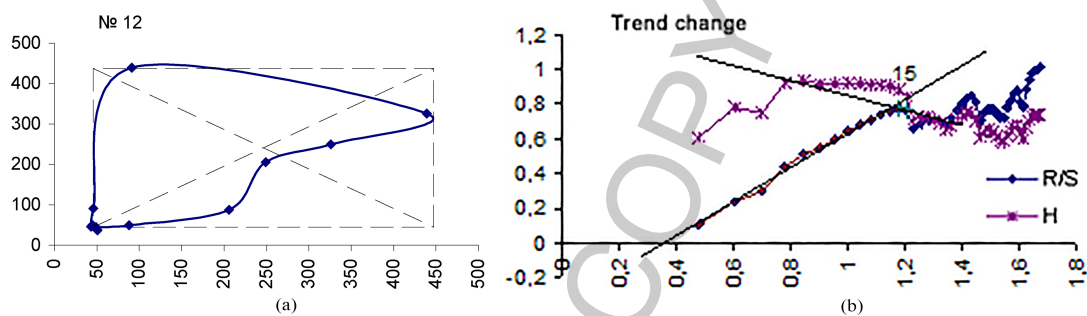


Fig. 18. The twelfth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

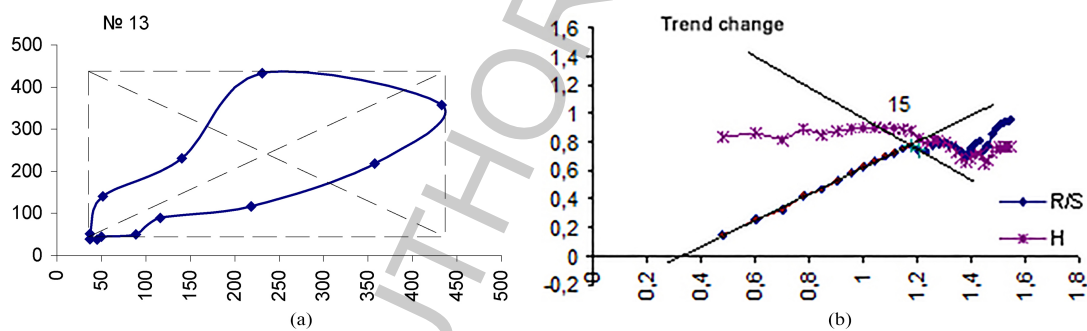


Fig. 19. The thirteenth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

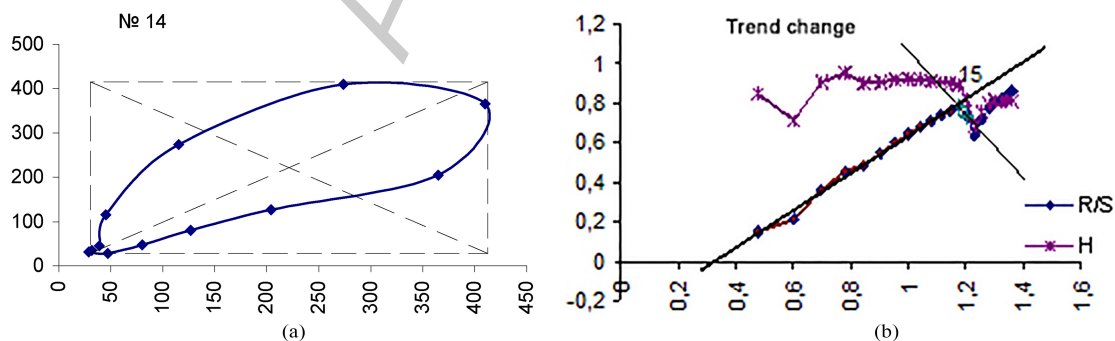


Fig. 20. The fourteenth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

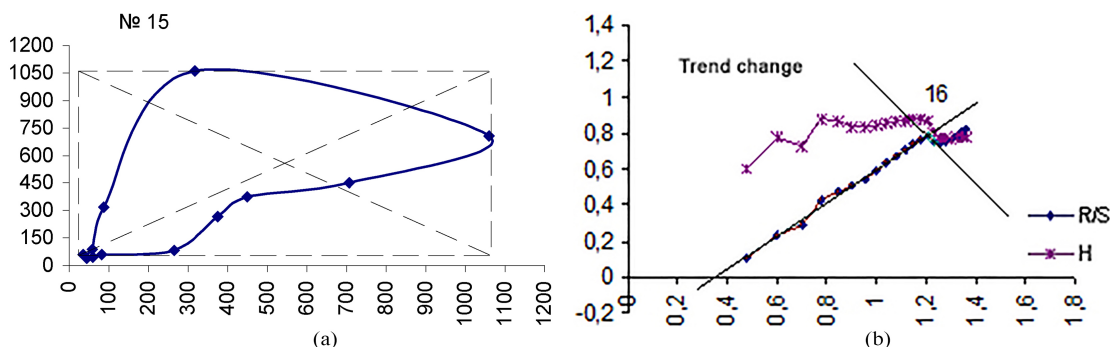


Fig. 21. The fifteenth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

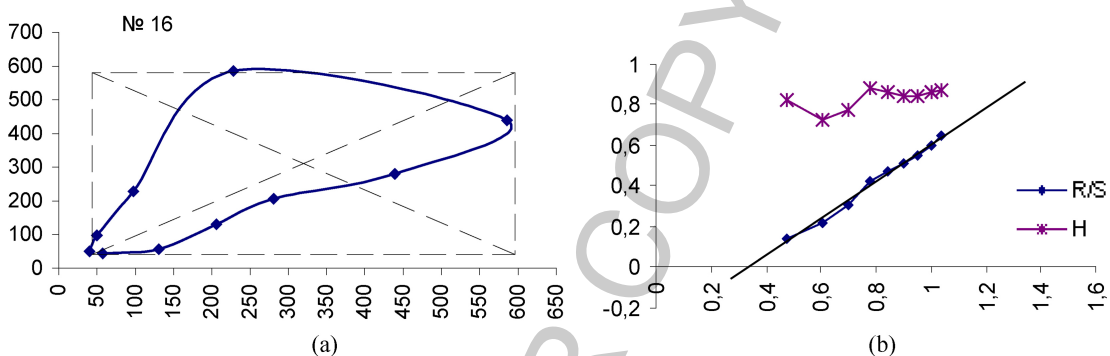


Fig. 22. The sixteenth 12-month quasi-cycle of TS Z and its detection using the R/S - and H -trajectories.

quasi-cycle C_r , which is a rectangle with sides parallel to the axes of coordinates. The intersection of the diagonal of the dimensional rectangle determines the so-called center of rotation of the quasi-cycle O_r , the coordinates of which are denoted as $O_r(x_r, y_r)$.

Let's analyze the characteristic features of the phase portrait of TS Z.

1. The phase portrait of TS Z is divided into quasi-cycles, which have dimension 12. This fact is in sufficient agreement with the results of a fractal analysis devoted to the estimation of the depth of memory of TS.
2. In each quasi-cycle, exactly every link has a direction of rotation clockwise. In this case, the dimensional rectangle can be divided into 4 sectors by straight lines parallel to the coordinate axes with the intersection in the center of the dimensional rectangle.
3. Centers of quasi-cycles $O_r(x_r, y_r)$, in the order of their numbering $r = \overline{1, 72}$, evolve along a definite trajectory, the points of which are located in a rather small neighborhood of the bisector of the positive orthant of the Cartesian coordinates.

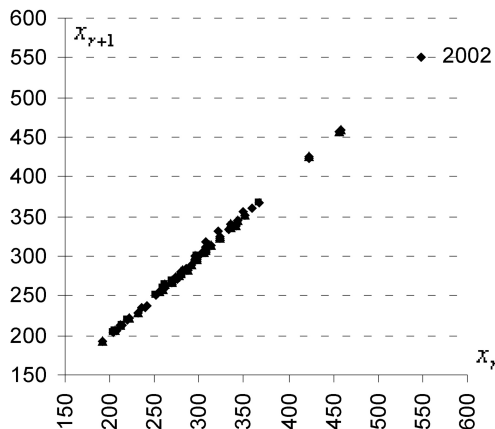


Fig. 23. Evolution of TS quasi-cycle centers for the period from 1926 to 2003.

As can be seen from Fig. 23, although the coordinates of the centers of all quasi-cycles define the points of the bisector of the positive orthant $x_r x_{r+1}$ of the Cartesian coordinates, the trajectory of movement of these coordinates is characterized by a significant range $R \approx 550 - 200 = 350$, which is more than 1.5 times greater than the minimum point. In the light of

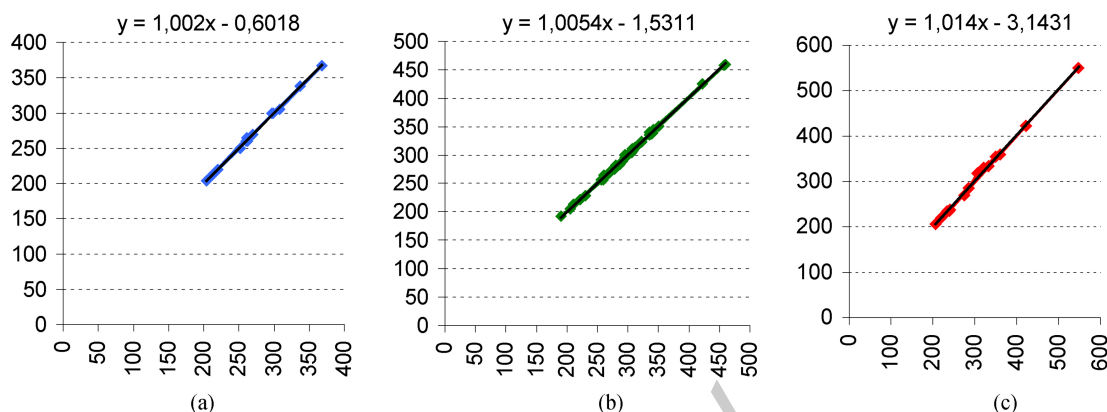


Fig. 24. Decomposition of the trajectories of the centers into temporary periods.

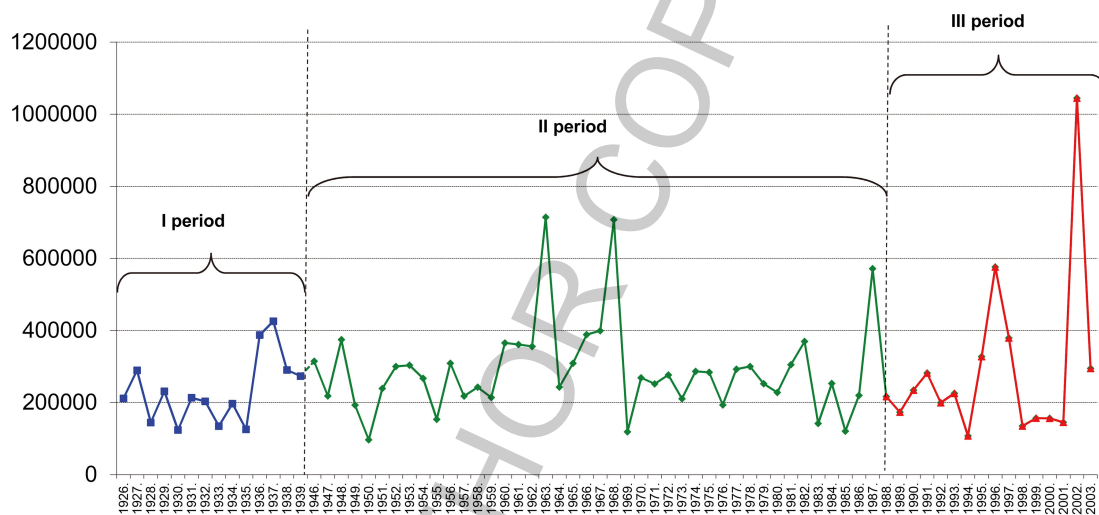


Fig. 25. Movement of areas of dimensional rectangles (taking into account time parameter).

this fact, it is interesting to identify long-term trends that govern the evolution of these centers of dimensional rectangles. Due to this, splitting up in Fig. 24 into three periods – 1926–1940 – Fig. 24a; 1946–1987 – Fig. 24b; 1988–2003 – Fig. 24c – was done. From the visualization of Fig. 23, the following tendency appears with sufficient certainty: with approximately the same value $\min \approx 200$, over time, the magnitude increases in the following ratio: $R_1 \approx 350 - 200 = 150$ (Fig. 24a), $R_2 \approx 450 - 200 = 250$ (Fig. 24b), $R_3 \approx 550 - 200 = 350$ (Fig. 24c). The given range of magnitude confirms the well-known statement of climatologists about the existence of a general trend of climate warming in the northern hemisphere, since the filling of mountain rivers, especially in the summer months, is determined by the intensity of the melting of glaciers.

From the visualization of Fig. 25, it follows that the evolution of the size (area) of the overall rectangles of quasi-cycles is cyclical.

Step 1. Fractal analysis of TS (1) in order to establish the presence of long-term memory and assess its depth. At the output a fuzzy set $L = L(Z) = \{(l, \mu_l)\}$ of TS Z depth of memory estimates is obtained.

Step 2. Construction of a phase portrait $\rho(Z)$ for a given TS.

Step 3. Decomposition of the phase portrait into quasi-cycles C_r , taking into account the fact that the initial and final quasi-cycles may be incomplete.

Step 4. Conducting an analysis of the evolution of quasi-cycle centers $O_r(x_r, y_r)$, the evolution of the size (area) of dimensional quasi-cycling rectangles, and the nature of the rotation of quasi-cycle links.

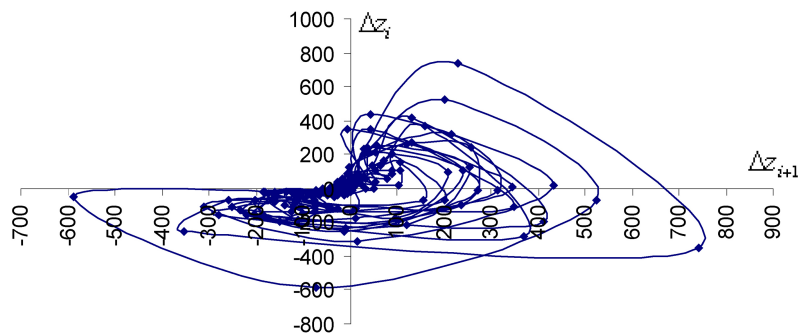


Fig. 26. TS phase portrait.

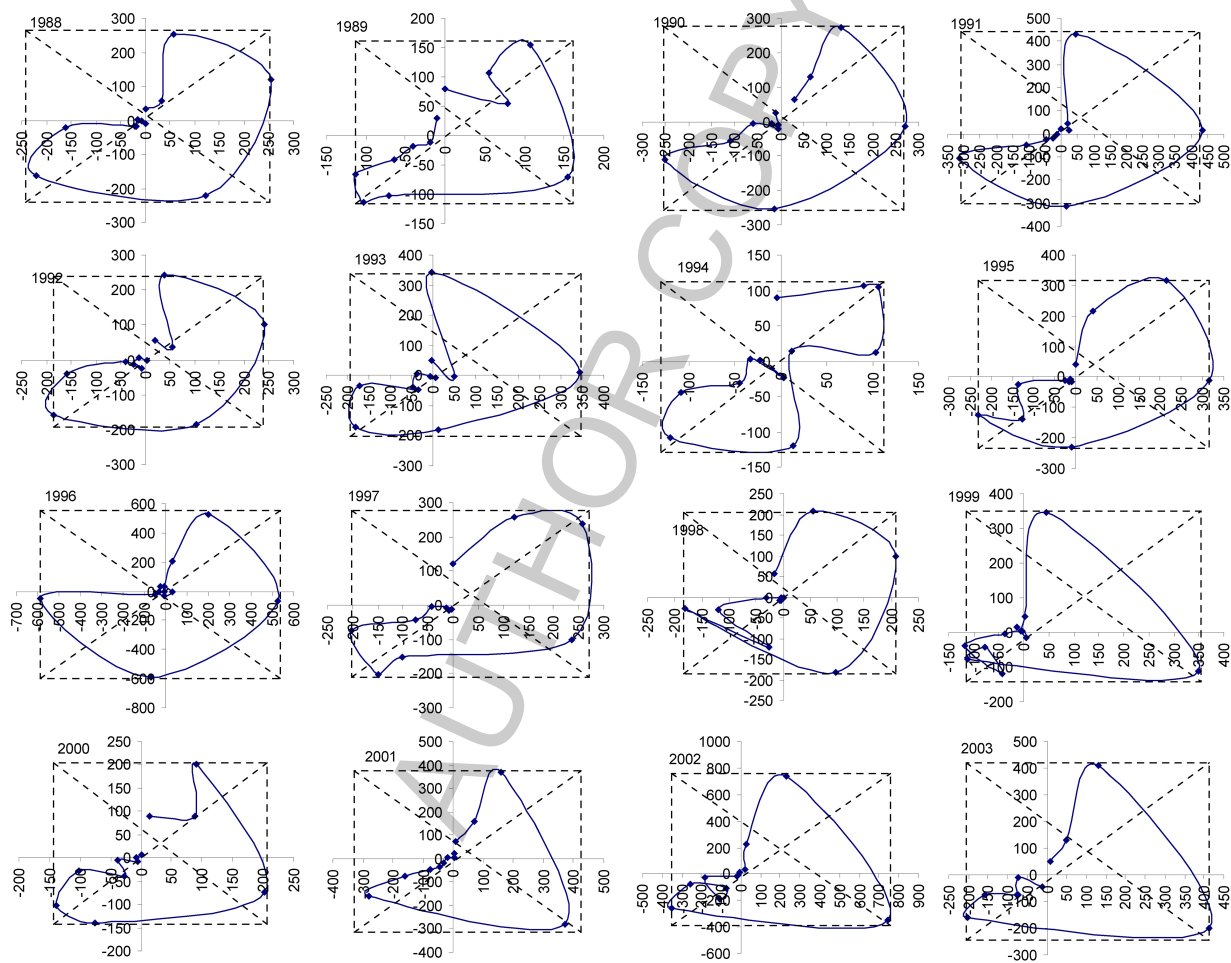


Fig. 27. Quasi-cycles of TS 3, isolated from the phase portrait of Fig. 26.

Step 5. Constructing a forecast on the principle of continuation (completion) of the corresponding quasi-cycle using the results of step 4 for two cases in which the last quasi-cycle is a) finished, b) unfinished.

In case a), we use the overall dimensions and the na-

ture of rotation of the quasi-cycles, taking into account which sector of the dimensional rectangle the predicted point belongs to.

In case b), the same procedures are carried out, but taking into account the evolution of centers and tran-

Table 3
Correlation coefficients between meteorological factors and TS of mountain rivers monthly runoffs

	Av t	MAX t	MIN t	Av Hum	Sun. Sum	Sum Prec	Kuban	B. Zel.	Teberda	Aksout	Marukha
Av t	1										
MAX t	0.87	1									
MIN t	0.91	0.62	1								
Av Hum	0.04	-0.24	0.27	1							
Sun. Sum	0.00	0.36	-0.29	-0.55	1						
Sum Prec	0.06	-0.17	0.21	0.38	-0.43	1					
Kuban	0.15	-0.04	0.27	0.42	-0.19	0.69	1				
B. Zel.	0.38	0.14	0.48	0.41	-0.23	0.74	0.86	1			
Teberda	0.30	0.25	0.31	0.48	-0.02	0.16	0.47	0.37	1		
Aksout	0.18	-0.02	0.27	0.47	-0.16	0.66	0.81	0.72	0.45	1	
Marukha	0.12	-0.08	0.27	0.45	-0.36	0.65	0.59	0.59	0.21	0.67	1

sitions from the final point of one cycle to the starting point of a new cycle.

Thus, the proposed approach differs from the classical methods of forecasting with the new implementation of taking trends into account (the evolution of the centers and dimensions of the dimensional rectangles), as well as with the new tools (phase portraits) for identifying the cyclic component of the considered TS.

Along with the presented above (in Fig. 5), there are other approaches to the construction of phase portraits of TS. Many researchers are building phase portraits in the form of “the level of the TS indicator is its first derivative”, i.e. these portraits are built in the phase space F'' , where z_i is the increment of the i -th element of the TS (1). Let’s introduce the designation of TS increments of the monthly runoffs of the mountain river Kuban for the period from January 1988 to December 2003:

$$\Pi_3 : \langle n_i \rangle, \overline{i = 1, 192}. \tag{8}$$

This kind of phase trajectory of mountain river Kuban runoffs TS K_3 is shown in Fig. 26. This trajectory consists of 16 quasi-cycles $C'_r, r = 1, 2, \dots, 16$. Figure 27 visually confirms the fact of self-similarity of TS Π_3 quasi-cycles.

It should be noted that the centers of the quasi-cycles TS Π_3 , in the order of their numbering, evolve along a certain trajectory, the points of which are located exactly on the bisector of the positive orthant of the Cartesian coordinates.

It is easy to see that the phase portraits presented in Fig. 27 and the numbers of the points of failure with the R/S -trajectory and the H-trajectory also confirm the presence of the cyclic component in the time series under consideration.

From the point of view of economic security among various natural factors, the most interesting is the statistics of precipitation amounts in combination with the statistics of monthly runoffs of the mountain river

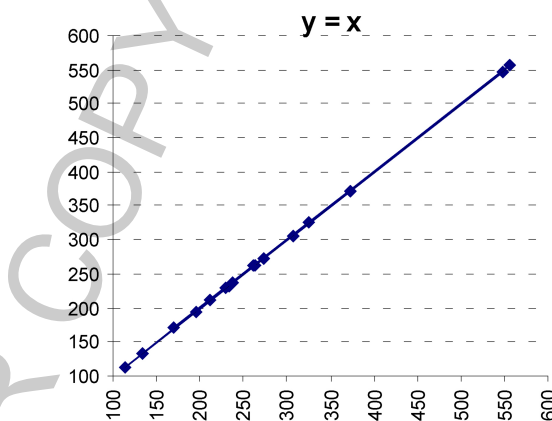


Fig. 28. Evolution of TS Π_3 quasi-cycles centers.

Kuban for the period from January 1969 to December 2003. In this context, the summer months deserve the most attention, first of all – June, which is rather strongly correlated with the precipitation index and at the same time has close to the maximum runoff volume of mountain rivers.

To identify investigative relationships between weather conditions (such as: average air temperature, maximum air temperature, minimum air temperature, average air humidity, sunshine duration, amount of precipitation per day) and the dynamics of monthly runoffs of mountain rivers Kuban, Bolshoi Zelenchuk, Teberda, Aksout and Marukha a correlation analysis was conducted.

When studying relationships in two-dimensional data, one should always remember the following: description and consideration of relationships. This is the most common goal, providing basic information, with which one can better understand the true structure of the world around us. When studying a complex system, it is very important to know which factors most closely interact with each other, and which factors generally affect each other. Availability of such information can

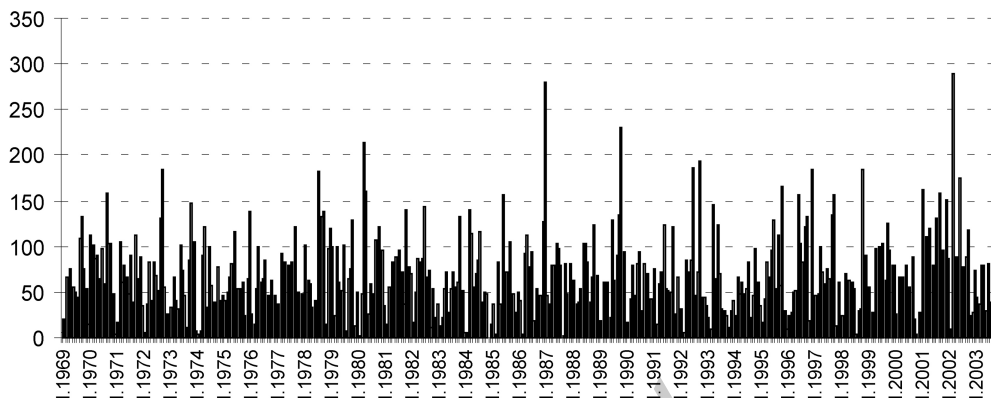


Fig. 29. Graphic representation of monthly precipitation amounts TS.

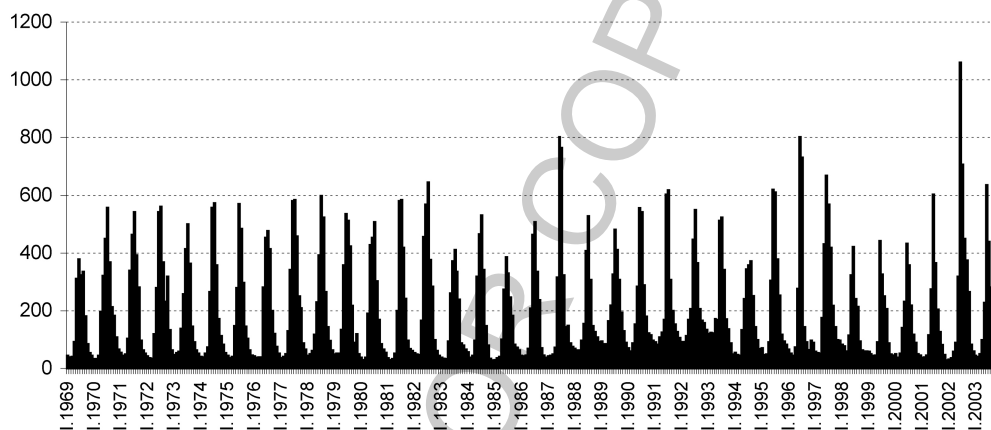


Fig. 30. Graphic representation of Kuban River monthly runoffs TS.

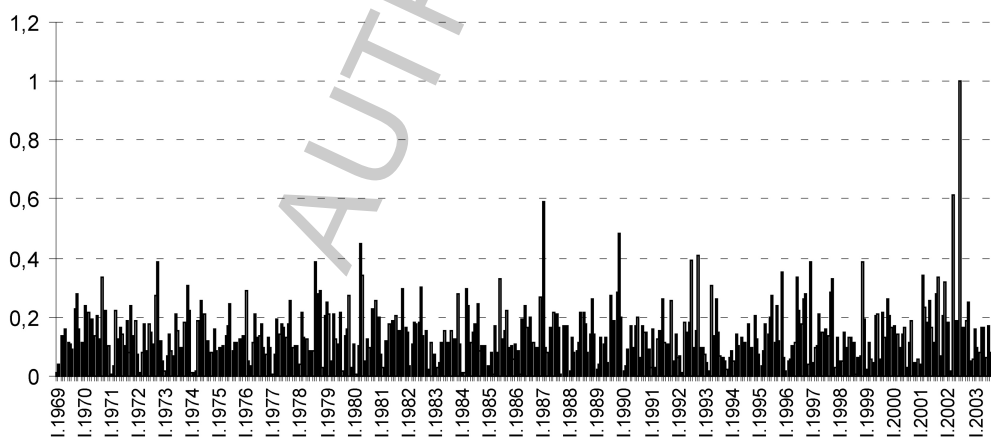


Fig. 31. Graphic representation of the normalized values of monthly precipitation TS.

be of considerable assistance in long-term planning and making other strategic decisions [7].

It is interesting to investigate the dependence of the behavior of mountain rivers under meteorological fac-

tors in June. Table 2.5 presents the values of the correlation coefficients.

In the context of the above, it seems expedient to construct TS of monthly precipitation and its com-

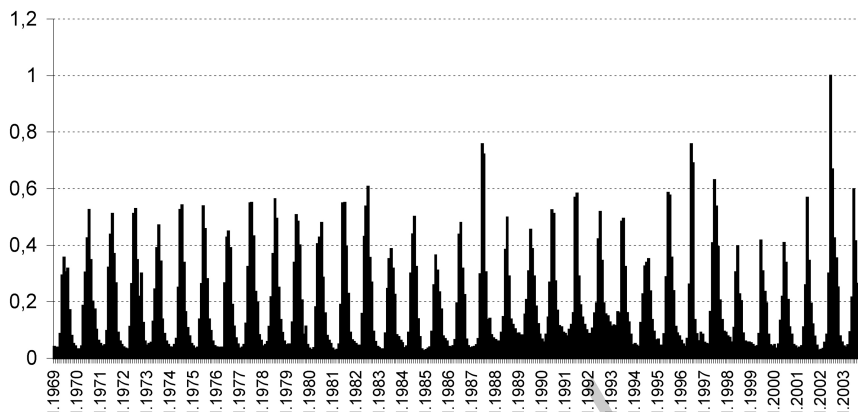


Fig. 32. Graphic representation of the normalized of the Kuban River monthly runoffs TS.

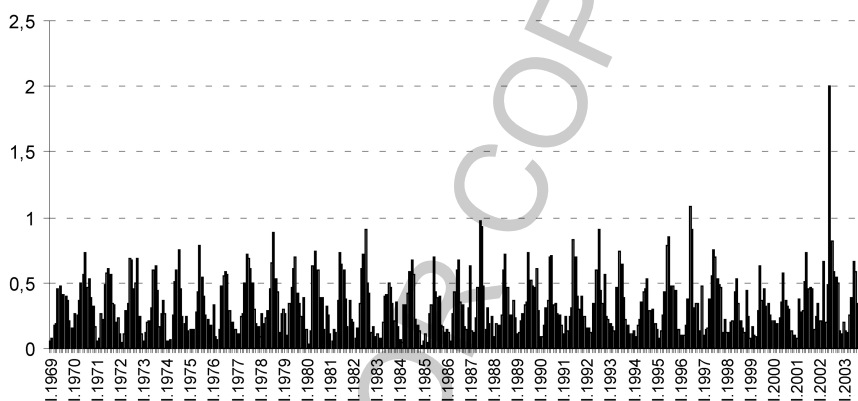


Fig. 33. TS histogram – the sum of the normalized total precipitation TS Kuban River monthly runoffs TS.

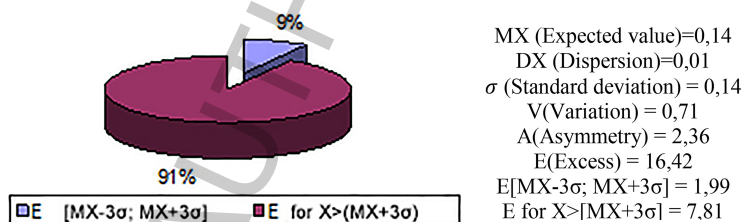


Fig. 34. Statistical parameters and the percentage of the weight of the “head” to the weight of the “tail” of TS X (monthly amounts of precipitation TS).

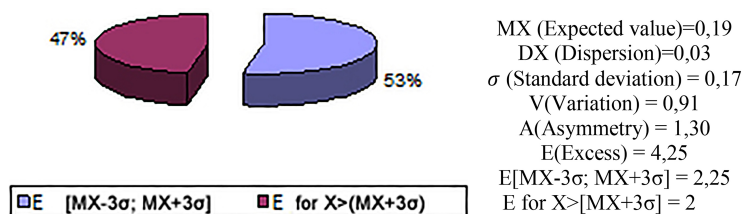


Fig. 35. Statistical parameters and the percentage of the weight of the “head” to the weight of the “tail” of TS Y (Kuban River monthly runoffs TS).

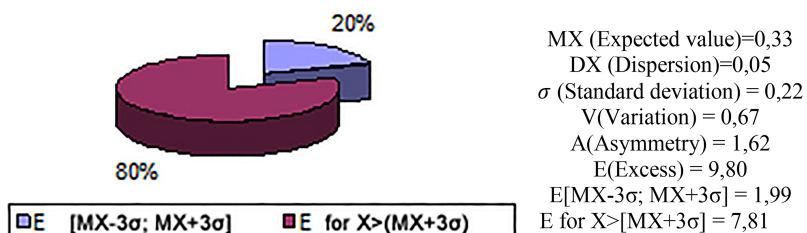


Fig. 36. Statistical parameters and the percentage of the weight of the “head” to the weight of the “tail” of TS BP Z (the sum of the normalized values of monthly precipitation TS and Kuban River monthly runoffs TS).

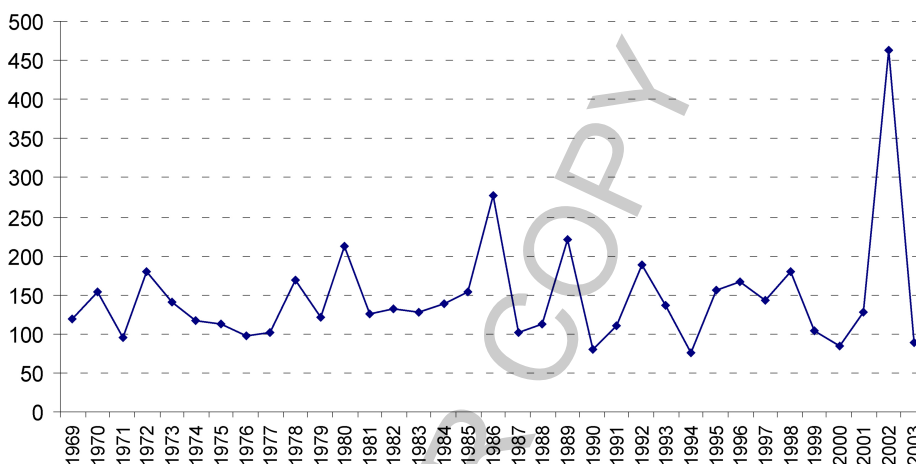


Fig. 37. Range $X_{max} - X_{min}$ for TS X (monthly precipitation TS).

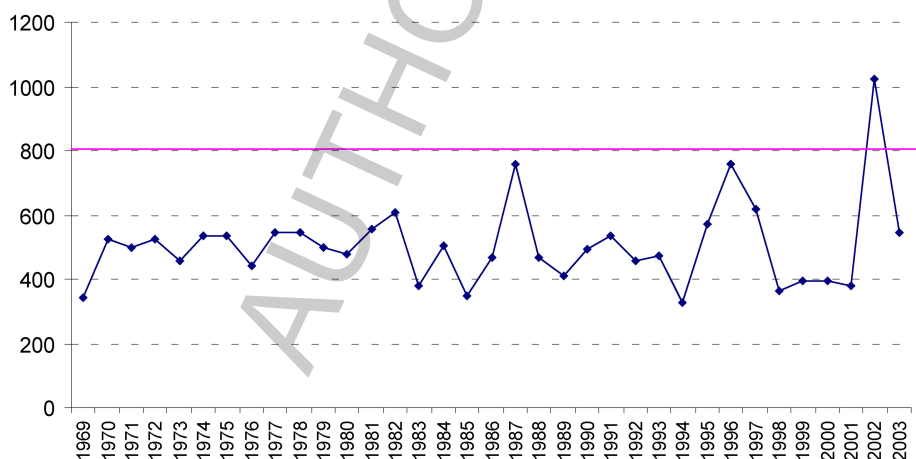


Fig. 38. Range $Y_{max} - Y_{min}$ for TS Y (Kuban River monthly runoffs TS).

parative analysis, together with Kuban mountain river runoffs TS. Graphical representations of these TSs are presented in Figs 29 and 30, respectively.

In order to achieve comparability in units of measurement, both TS are normalized on the basis of the rule the height of the maximum level is taken as 100%

value in both TS. Normalized values are obtained by dividing the numerical value of the considered level by the maximum value. A graphical representation of the normalized amounts of precipitation and runoff volumes is presented in Figs 31 and 32. Let’s introduce the following notation, corresponding to this graphic

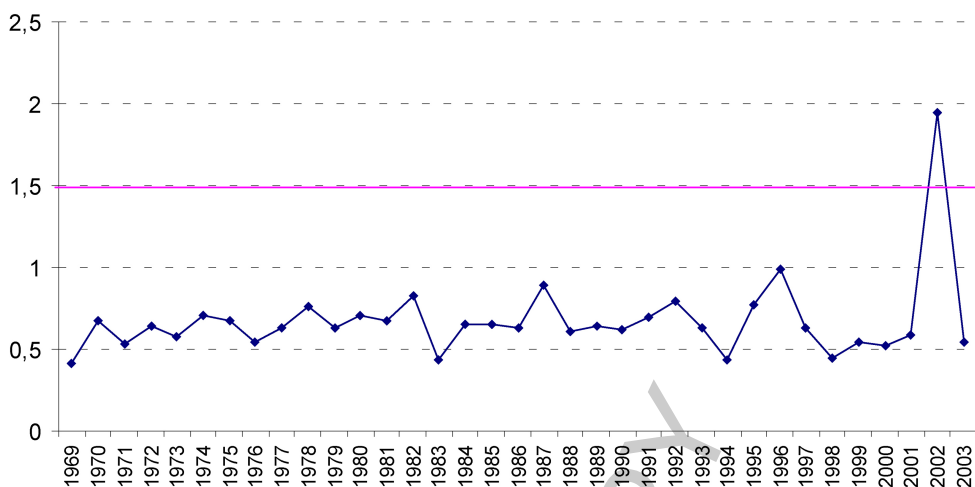


Fig. 39. Range $Z_{max} - Z_{min}$ for TS Z (sum of normalized values of monthly precipitation TS and Kuban River monthly runoffs TS).

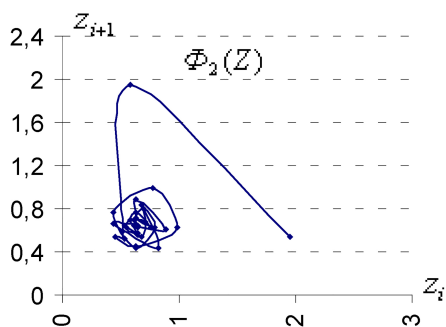


Fig. 40. Phase portrait of the TS Z (sum of normalized values of monthly precipitation TS and Kuban River monthly runoffs TS from 1969 to 2003).

representation:

$$X = \langle x_i \rangle, i = 1, 2, \dots, n, \tag{9}$$

$$Y = \langle y_i \rangle, i = 1, 2, \dots, n. \tag{10}$$

To assess the risk of the months of the year in relation to potential flooding, let's consider the amount of TS (9) and TS (10):

$$Z = X + Y = z_i, i = 1, 2, \dots, n, \tag{11}$$

where $z_i = x_i + y_i$.

Graphical representation of TS (11) is presented in Fig. 33.

For clarity, Figs 34–36 present the statistical parameters of TS X, Y and Z.

It is easy to see that each of the considered TS has a “heavy tail” [5], with the most “heavy tail” – 91% for TS X – monthly precipitation TS (see Fig. 34). In fact, to the same extent, the “heavy” tail is for TS Z (see Fig. 36).

Table 4
Dimensions of TS Z quasi-cycles

C_k	C_1	C_2	C_3	C_4	C_5	C_6
L_k	6	4	5	6	5	7

It is of interest to investigate in the data of TS X, Y and Z the behavior of the range of variability from the highest value (precipitation in TS X and maximum runoffs in TS Y) to the minimum received during the year (12 months). Figures 37–39 show graphic images of oscillations.

Fundamentally possible achievement of threshold values and determination of the of local extremes' measure of remoteness from this threshold value is the task of ensuring the economic security of the region. Interval from which the danger begins is expertly set. Using interval analysis, one can establish that the threshold values are located above the level of 1.5 in TS $Z = X + Y$.

It is of interest to investigate phase portraits of range $Z_{max} - Z_{min}$ TS.

The dimensions L_k of six quasicycles are presented in Table 3.

The memory depth of the considered TS Z, represented by a fuzzy set of quasi-cycles' dimensions (see Table 3), is consistent with the fact of the presence of an eleven-year solar cycle. For visualization, Figure 42 presents a graphical representation of the dimensions of TS Z quasi-cycles.

Presented above numerical values of the results obtained, related primarily to the presence and weight of “heavy tails”, suggest a high degree of “riskiness” of the considered mountain Kuban river runoffs in the context of the region's economic security.

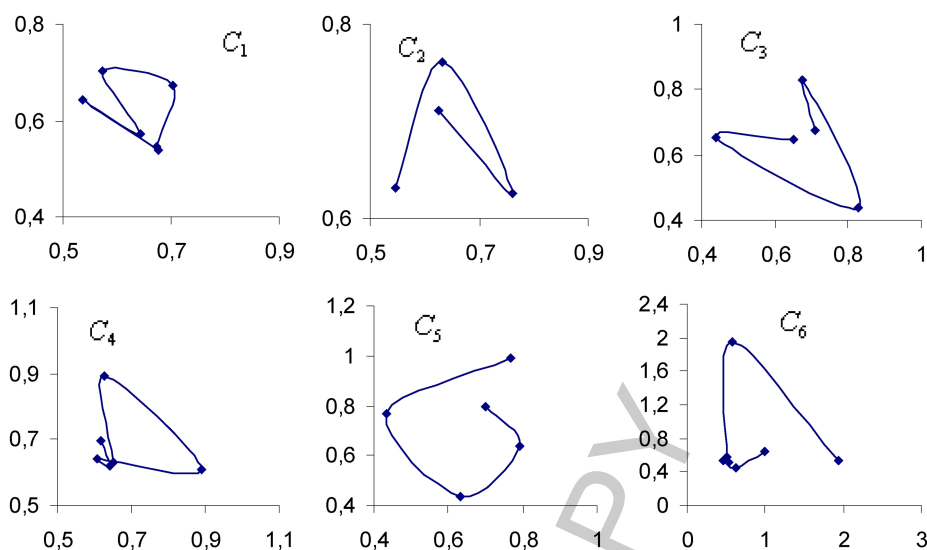
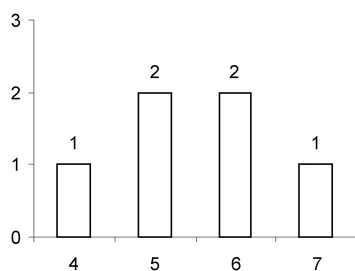


Fig. 41. TS quasi-cycles, isolated from the phase portrait of Fig. 40.

Fig. 42. Frequency of quasi-cycles of dimensions 4, 5, 6 and 7 for TS Z .

4. Conclusions

As the main results of the research, let's note the following points:

- The research was conducted on the basis of nonlinear dynamics methods. Foresight analysis tools are used – fractal time series analysis and phase analysis. The results obtained in the course of the research using these methods not only correlate, but also complement each other.
- The established fact of the presence of long-term memory in the time series of the mountain rivers' runoffs gives basis to the development of a system for a short- and medium-term forecast of the runoffs of these mountain rivers. The system of short- and medium-term forecast is planned to be developed on the mathematical apparatus of the theory of cellular automaton and the theory of fuzzy sets [3]. Moreover, the amount of memory used by the cellular automaton and, ultimately,

the complexity of the computational prediction scheme significantly depend on the memory depth of the predicted TS. The upper estimate of the depth of memory of the considered TS is justified and clearly calculated.

- R/S -analysis [6] is a tool for identifying cycles, both periodic and non-periodic. Based on the results of the calculations, new additional possibilities of the “algorithm of sequential R/S -analysis” developed by the authors compared to the “Hurst norm range algorithm” for detecting cycles are shown.
- The method of sequential R/S -analysis presented by the authors can be used to obtain estimates of the fractal characteristics of the considered TS limited segment. This is a distinctive feature, because to obtain sufficiently accurate estimates of the fractal characteristics of TS using the Hurst normed span algorithm, series of several thousand or even about ten thousand observations are required.
- A detailed step-by-step representation of the sequential R/S -analysis algorithm and a computational experiment on real data are in fact the verification of this algorithm. Identification of the presence of strict cyclicity is important, both in terms of pre-forecast analysis, and in terms of building predictive models.
- Detection of long-term memory in mountain river runoffs by an objective substantiation of the fundamental possibility of building a predictive model, in the course of which all essential fac-

tors (for example, the presence of cycles of the form Eq. (4)), which determine the presence of this memory, are taken into account. Given the current basic position of the decomposition analysis of time series of 4 components: a) the trend, b) the cyclical component, c) the seasonal fluctuation, d) the irregular or residual component, we note that the identified cyclical component contributes very significant information to make an accurate forecast.

- The author’s approach which differs from the classical methods of forecasting with the implementation of the trend accounting is proposed (evolution of the centers and dimensions of the dimensional rectangles), as well as with the new tools (phase portraits) to identify the cyclic component of the considered TS.
- Investigative relationships between weather conditions (such as: average air temperature, maximum air temperature, minimum air temperature, average air humidity, sunshine duration, amount of precipitation per day) and the behavior dynamics of monthly runoffs of mountain rivers: Kuban, Bolshoi Zelenchuk, Teberda, Aksaut and Marukha were revealed. Description and consideration of the relationship is the most common goal, providing basic information, with which one can better understand the true structure of the world around us. When studying a complex system, it is very important to know which factors most closely interact with each other, and which factors generally affect each other. Knowledge of this information provides significant assistance in long-term planning and other strategic decisions [7].

Thus, in this work methodically and step by step is presented the first part of the author’s research – predictive analysis (also in [13]). This research is basic for the next step – to obtain an accurate forecast of the considered mountain river Kuban runoffs, which has a high degree of “riskiness” in the context of the region’s economic security. Using developed models and methods as a base, authors plan to create a DSS, as an integral part of the solution package within activities carried out in flood-hazardous areas.

References

- [1] M.I. Budyko, Changing of the climate, Gidrometeoizdat, Leningrad, 1974, p. 280.
- [2] The results of the commission, Issue 24, July 27, 2002. Available online: <http://www/admkrai.krasnodar.ru> (accessed on 12.02.2019).
- [3] V.A. Perepelitsa, M.D. Kasaeva, F.B. Tebuyeva and L.G. Temirova, Using tools of cellular automaton for the formation of predictable fuzzy values of yield on the basis of a time series, News of universities, North Caucasus region, *Natural Sciences* **4** (2003), 5–11.
- [4] V.A. Perepelitsa and E.V. Popova, Mathematical models and methods for assessing the risks of economic, social and agricultural processes, Rostov-on-Don University, Rostov-on-Don, 2002, p. 202.
- [5] E. Peters, Chaos and order in the capital markets: a new analytical view of cycles, prices and market volatility, Wiley, New Jersey, 2000, p. 288.
- [6] E. Peters, Fractal analysis of financial markets: application of the theory of chaos in investment and the economy, Internet-Trading, Moscow, 2004, p. 304.
- [7] E. Siegel, Practical business statistics, Williams, Moscow, 2004, p. 1056.
- [8] E. Feder, Fractals, Mir, Moscow, 1991, p. 260.
- [9] A.M. Yangishieva, V.A. Perepelitsa, E.V. Popova and A.D. Salpagarov, The use of nonlinear dynamics methods for the predictive analysis of the runoffs volumes of mountain rivers, *Ecological Bulletin of Scientific Centers of the Black Sea Economic Cooperation* **1** (2005), 73–84.
- [10] K. Holden, D.A. Peel and J.L. Thompson, Economic forecasting: introduction, Press Syndicate of the University of Cambridge, Cambridge, United Kingdom, 1990, p. 231.
- [11] Natural Disasters in the World. Statistical Trend on Natural Disasters. National Land Agency, Japan, IDNDR, Promotion Office, 1994.
- [12] N. Packard, J. Crutchfield, D. Farmer and R. Shaw, Geometry from a time series, *Physical Review Letters* **45** (1980), 712–716.
- [13] E. Popova, L. de Costa and A. Kumratova, Hybrid instrumental means of predictive analysis of the dynamics of natural and economic processes, *Advances in Intelligent Systems and Computing* **923** (2019), 31–39.
- [14] O. Kingsley, N. Usman and I. Syed, Semantic fuzzy mining: Enhancement of process models and event logs analysis from syntactic to conceptual level, *International Journal of Hybrid Intelligent Systems* **14**(1–2) (2017), 67–98.