Review Article

Prediction of the State of the Cardiovascular System based on the Allocation of the Boundaries of the Implementation of the Dynamic System

NT Abdullayev*, OA Dishin, ID Ibragimova and KhR Ahmadova

Azerbaijan Technical University, Azerbaijan

*Corresponding author: NT Abdullayev, Azerbaijan Technical University, Azerbaijan

Introduction

Recently, research methods for rare experimental events have been intensively developed [1-9]. For a quantitative assessment of rare events, the time intervals between occurrences of successive events above (or below) some threshold Q are usually considered. In this case, both the probability density function of these repeated intervals and their long-term dependencies are investigated (autocorrelation function, conditional repetition periods, etc.). In the numerical analysis necessarily considers not very large thresholds Q, which provide good statistical estimates of repeated intervals, and then these results are extrapolated to very large thresholds, for which the statistics are very poor [1-5].

In [6], based on 24-hour Holter-monitoring data, it was shown that linear and non-linear long-term memory, inherent in repeated heartbeat intervals, leads to a power law of PDF change. As a consequence, the power law will be satisfied by the probability WQ(t; Δt) of the fact that for Δt time units that have passed since the last repeated interval with an extreme event (record) greater than the threshold Q (briefly, Q-interval) , at least one Q-interval will appear, if for t time units before the last Q - a Q-interval of heartbeat has appeared. Using the probabilities WQ(t; Δt) using long-term memory, the procedure for predicting Q-intervals (RIA-procedure) was proposed in [6-14], which is an alternative to the traditional technique of recognition of specified patterns, the so-called PRT-recognition technique (pattern recognition technique) based on short term memory. The RİA approach does not require the limitedness of the used Q-interval statistics and in all cases gives the best result. Based on this approach, a procedure for predicting a larger (with a record exceeding a certain large threshold Q) repeated Q-interval of heartbeat was developed in [15] using the preliminary selection of repeated Q-intervals with persistent (steadily growing) records based

on the transformation [16,17] of the original signals (in this case, repeated Q-intervals) in a series of times to achieve a given threshold

of change. However, the implementation of the approach to predicting complex signals proposed in [17] based on the selection of the dynamic system implementation boundary provides for a preliminary selection of a "stable" segment at the end of the time series where the predictor is trained taking into account the latest dynamics of the process under study. In this regard, in [18,19], a computational procedure was proposed for segmentation of non-stationary time signals with fractal properties.

However, the implementation of the approach to predicting complex signals proposed in [17] based on the selection of the dynamic system implementation boundary provides for a preliminary selection of a "stable" segment at the end of the time series where the predictor is trained considering the latest dynamics of the process under study.

Minimal cCoverage and Fractality Index

This dimension is introduced by Hausdorff by the formula

$$D = \lim_{\delta \to 0} \frac{\ln N(\delta)}{\ln(\frac{1}{\delta})} \ (1.1),$$

where N (δ) is the minimum number of balls of radius δ covering the set A. The basis for the definition (1.1) is the asymptotics for δ, which is defined by the expression for fractal sets.

$$N(\delta) \sim (1/\delta)^D \text{ при } \delta \to 0 \ (1.2)$$

The fractal dimension of the time series can be calculated directly through the cell dimension D_c, which is also called the Minkowski dimension or the box dimension [20].

To determine the value of D, the plane on which the time series graph is defined is divided into cells of size δ, and the number of cells N (δ) is determined, where at least one point of this graph is located. It is then constructed by the points (y_i, x_i) $y_i = \ln N(\delta_i) x_i = \delta_i$ regression $y = ax + b$, whose coefficients are estimated using

the least squares method (MHK). Then $D_c = \hat{a}$ is the MHK-estimate of the coefficient a. However, to reliably calculate D_c in this way, a too large representative scale containing several thousand data is required [21]. Within this scale, a time series can change many times the nature of its behaviour.

In order to associate the dynamics of the process corresponding to the series $\{X(t_i)\}$, with the fractal dimension of this time series, it is necessary to determine the dimension D locally. For this, it is necessary to find an approximation sequence that is optimal in a certain sense for each fixed δ. For this purpose, the following procedure for minimal coverages is proposed in [3]. To calculate from the fractal dimension of a time series, or the graph of a real function of one scalar variable $y = f(t)$, defined on a certain interval [a, b], you can directly use the procedure for determining cell dimensions, considering cell coverage as a special case of covering with rectangles.

We introduce a uniform partition of the segment [a, b] by the points $t_1, \dots t_{m-1}$

$$\omega_m = [a = t_0 < t_1 \dots < t_{m-1} < t_m = b] \; \delta = \frac{b-a}{m} \; (1.3)$$

and construct the graph coverage of the function $y = f(t)$ with rectangles with base δ and height $A_i(\delta)$, Figure 1 [19].
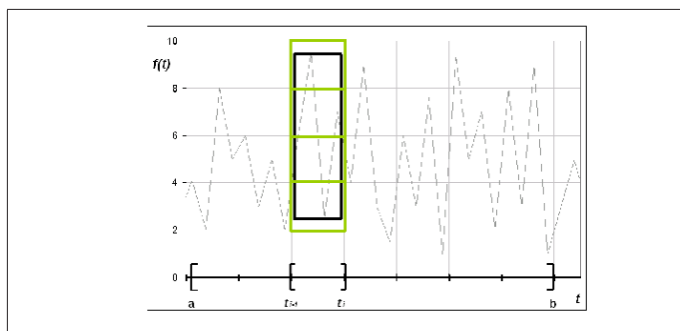

Figure 1: Fragment of cellular (gray rectangle) and minimal (black rectangle) coverings of the graph of the fractal function on the segment [t_(i-1),t_i]. Obviously, the minimum coverage is more accurate cellular

Here $A_I(\delta) = \overline{f_i}(\delta) - \underline{f_i}(\delta) \; f_i(\delta) = \max\limits_{t \in [t_{i-1}, t_i]} f(t)$
$\underline{f_i}(t) = \min\limits_{t \in [t_{i-1}, t_i]} f(t)$

We introduce further the value

$$V_f(\delta) \equiv \sum_{i=1}^{m} A_i(\delta) \; (1.4)$$

and we call $V_f(\delta)$ the amplitude variation of the function f(t) corresponding to the partition scale δ. The total coverage area $S\mu(\delta)$ is written as

$$S\mu(\delta) = V_f(\delta)\delta \; (1.5)$$

It is obvious that $S\mu(\delta)$ is the minimum area of the graphics coverage of the class of coatings with rectangles. Therefore, such a coating is called minimal [19].

After multiplying both parts in (1.2) by δ^2, for the cell division area $S_c(\delta)$ we get

$$S_c(\delta) \sim \delta^{2-D} \text{ at } \delta \to 0 \; (1.6)$$

and in particular for rectangular coverings

$$S_c(\delta) \sim \delta^{2-D_\mu} \; \delta \to 0 \; (1.7)$$

From (1.5) and (1.7) we find

$$V_f(\delta) \sim \delta^{2-D_\mu} . \delta^{-1} = \delta^{1-D_\mu} = \delta^{-\mu}$$

i.e. $V_f(\delta) \sim \delta^\mu$ at $\delta \to 0$ (1.8)

where

$$\mu = D_\mu - 1 \; (1.9)$$

The index μ is called [19] the fractality index, and the dimension D_μ is the dimension of the minimum coverage. The value of μ is related to the stability of the time series: the more stable the behavior of the original series (i.e., oscillations occur near one level), the greater the value of μ, while the opposite is true.

If the function f(t) is called stable (quasistationary) on the interval (a, b), when its volatility is invariant on this interval, which has long been evaluated [22] by the standard deviation of changes in the selected time window (a, b), then the degradation point time series, i.e. The point at which stability is violated (quasi-stationarity) can be determined by the jump in volatility. For the evaluation, change in volatility points that divide the original time series into quasistationary portions (segments) corresponding jumps volatility functions developed [18] computing segmentation procedure generalizes iterative method centered cumulative sum of squares (ICSS) [23]. In general, the problem of segmentation, non-stationary time series, i.e. splitting it into non-intersecting adjacent fragments that will be statistically homogeneous (or at least possess this property in a general degree than the initial data) is known as the task of finding the chance-point problem [24,25].

The main advantage of the fractality index μ compared to other fractal indicators (in particular, with the Hurst index H) is that the corresponding value $V_f(\delta)$ has a quick access to the power asymptotic mode (1.7). The Hurst index H is determined based on the assumption that

$$\langle | f(t + \delta) - f(t) | \rangle \sim \delta^H \text{ при } \delta \to 0 \; (1.10)$$

Angle brackets here mean averaging over the interval (a, b). For comparison of μ and H, we introduce the definition of the average amplitude $< A_i(\delta) >$ on the partition (1.3)

$$< A_i(\delta) > \equiv \frac{1}{m} \sum_{i=1}^{m} A_i(\delta) \; (1.11)$$

Multiply (1.4) by $\frac{1}{m} \sim \delta$ and substitute it into (1.8). Then we get

$$\langle A_i(\delta) \rangle \sim \delta^{H_\mu} \text{ at } \delta \to 0 \; (1.12)$$

As is known [26], if f(t) is a realization of a Gaussian random process, then the Hurst index H is related to the dimension D_μ and therefore to the index μ by the relation

$$H = 2 - D_\mu = 1 - \mu \; (1.13)$$

Therefore, in this case $H = H_\mu = 1 - \mu$. However, real physiological time series, generally speaking, are not Gaussian and therefore H_μ and H can vary greatly. Indeed, in formula (1.12) a power law holds for the average amplitude of the function f(t) over an interval of length δ, while in equation (1.10) the power law holds

for the average difference between the initial and final value f(t) on the same interval. So, in general, $H \neq H_\mu$.

The aforementioned property of the rapid output of $V_f(\delta)$ to the power asymptotic mode makes it possible to use μ as a local characteristic that determines the dynamics of the original process, because the representative scale for its reliable determination can be considered to have the same order with the characteristic scale of the main process states. It is natural to introduce the function μ(t) with the values of μ defined on the minimum interval τ_μ preceding t. At this interval, the value of μ can still be calculated with acceptable accuracy. Since in practice a time series always has a finite length, the interval τ_μ has a finite length. When calculating the rhythmogram of the heartbeat, we have taken $\tau_\mu = 24$ hours, $\delta = t_i - t_{i1} = 1$ hour, b-a = 30x24 = 720 hours, m = 30.

We divide the interval (a, b) of the values of the time series f(t) into intersecting intervals of length τ_μ, shifted relative to each other by 1 point (i.e., by the value of δ). In each interval of length τ_μ, we construct minimal coverings Π_i from rectangles with bases $\delta_i = i * \delta (i = 1,....,n; n = \tau_\mu)$ with amplitude variation $V_f(\delta_i)$ of function f(t). Denote $y_i = \ln V_f(\delta_i)$, $x_i = \ln(\delta_i)$ and by points $(y_i, x_i)(i = 1,......n)$ construct the regression $y = a_1 x + a_0$. Denote by points and construct a regression. Then MHK-estimates a_1 and a_0 of the coefficients a and b will be obtained by

$$V_f(\delta) = b * \delta^{-\mu} \quad (1.14)$$

where $b = e^{\hat{a_0}}$ $H\mu = -\hat{a_1}$

Denote $\tau_\mu^{(j)}(j = 1,...,m, m = 30)$ τ j-th rolling interval length τ_μ; $\tau_\mu^{(j)} = [t_0, t_{\tau_\mu}]$, $\tau_\mu^{(m)} = [t_{m-\tau_\mu}, t_m]$, and each starting point of the interval $\tau_\mu^{(j)}$ let's compare a pair of numbers (b_j, μ_j). Thus, at all points $t = t_0, t_1,....t_{m-\mu}$, we obtain the corresponding values b(t) and μ(t), defining relation (1.14). Using the resulting arrays of these values, we calculate the following characteristics: average values $\langle \mu(t) \rangle$ и $\langle b(t) \rangle$ of μ and b values, maximum and minimum values of $\mu_{max}, \mu_{min}, b_{max}, b_{min}$ and $\delta_\mu = 0.5(\mu_{max} - \mu_{min}) < \mu(t) >$ and $\delta b = 0.5(b_{max} - b_{min})/ < b(t) >$ values, which characterize the variation of values (variability) μ (t) and b (t). Using the resulting arrays of these values, we calculate the following characteristics: average values of μ and b values, maximum and minimum values as well as quantities and characterizing the spread of values (variability) μ(t) and b(t). Repeating the above calculations several times, we calculate the averages over all the obtained realizations of the values $< \mu(t) >, < b(t) >, < \delta_\mu > and < \delta_b >$, which will give an idea of the behavior of the local characteristics μ and b for the time series under study. Experimental studies conducted in [19] for various types of fractal time series show a low variability of the values of μ and b with a lower variability of the value of b.

Since the value of b is closely related to the distribution of the increments of the time series and, therefore, to the volatility of the time series, for to apply the computational segmentation procedure [18] based on the detection of the jump of the volatility function, it is advisable to use to divide the non-stationary series into "quasistationary" parts with the initial series f(t), the series

of values of b(t) and as an estimate of the points of change of the studied series to take the average value of the points of change obtained for the original series f(t) and the number of the values of b(t), respectively.

Algorithm for Detecting Changes in Volatility

We give a brief description of the theoretical part of the method for estimating the point of change in the volatility of the time series used in [18].

Consider a time series model.

$$Y_i = \mu(t_i) + \sigma(t_i)\varepsilon_i \, i - 1, 2, ...T \quad (2.1),$$

where $\{y_i\}$ is the sequence of random variables $\{\varepsilon_i\}$ is the sequence of stationary errors of observation of the quantity Y at the instants of time $t_i c$ $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = 1$; μ(t) and G(t) are the regression function (conditional average) and the volatility function (conditional variation), respectively. Model (2.1) is a special case of the non-parametric model

$$Y_i = \mu(X_i) + \sigma(X_i)\varepsilon_i \quad (2.2)$$

when X =t and, therefore, Xi=ti.

Following the procedure for finding a single point change (single change-point) volatility [23], we set

$$\sigma(t) = \begin{cases} \tau_1 * \sigma_0(t), & \text{если } t \leq k_0 \\ \tau_2 * \sigma_0(t), & \text{если } t > k_0 \end{cases} \quad (2.3)$$

where $\tau_1, \tau_2 and k_0$ are unknown parameters. For simplicity, we put $k_0 = [T * \theta_0], 0 < \theta < 1$, where [a] means the largest integer not exceeding a.

Let the hypothesis H0 mean that volatility has no points of change. When this hypothesis is fulfilled, it follows from (2.1) that $E(Y_i - \mu(t_i))^2 = \sigma^2(t_i)$. We introduce the notation $Z_i = (Y_i - \mu(t_i)) / \sigma_0(t_i)$

$$S_T = \sum_{i=1}^T Z_i^2, S_k = \sum_{i=1}^k Z_i^2, S_{T-k} = \sum_{i=k+1}^T Z_i^2 \quad (2.4)$$

and determine the statistics

$$V_k = (\frac{k(T-k)}{T^2})^{1/2}(\frac{1}{T-k}.S_{T-k} - \frac{1}{k}S_k) \quad (2.5)$$

If $\mu(t)$ and $G_0(t)$ are known, the least-squares method (MHK) for the volatility jump point k0 can be used to obtain the following estimate [27]

$$\hat{k} = \arg\max V_k \quad (2.6)$$

The value of V_k will be close to 0 when the hypothesis H0 is fulfilled and is non-zero if the volatility changes. Simple calculations lead to equality

$$V_k = (\frac{k(T-k)}{T^2})^{1/2}(\frac{1}{T-k}.S_{T-k} - \frac{1}{k}S_k) = (\frac{1}{k(T-k)})^{1/2}.S_T.D_k \quad (2.7)$$

where

$$D_k = \frac{k}{T} - \frac{S_k}{S_T} \quad (2.8)$$

The statistics Dk, defined by the formula (2.8), differs only

in sign from the statistics $D_k = \dfrac{S_k}{S_T} - \dfrac{K}{T}$ used in [23]. From (2.7) it follows that Dk can also be used as an estimate of the point of change k0.

$$\hat{k} = \arg \max_k \left| D_k \right| = \arg \max_k (k(T-k))^{1/2} * V_k \quad (2.9)$$

In [27], more general statistics for detecting changes in volatility are introduced.

$$V_k^v = \left(\frac{k(T-k)}{T^2} \right)^{1-v} \left(\frac{1}{T-k} S_{T-k} - \frac{1}{k} S_k \right) = \left(\frac{k}{T} \left(1 - \frac{k}{T} \right) \right)^{1-v} * V_k, \, 0 \le v \le \frac{1}{2} \quad (2.10)$$

From (2.10) we can obtain an estimate for the point of change k0

$$\hat{k}(v) = \arg \max_k \left| V_k^v \right| \quad (2.11)$$

By specifying $\hat{k}(v)$ one can obtain the following estimates of the quantities $\tau_1^2 \tau_2^2$

$$\tau_1^2(v) = \frac{1}{\hat{k}(v)} * S_{\hat{k}(v)}, \; , \; \tau_2^2(v) = \frac{1}{T-\hat{k}(v)} * S_{T-\hat{k}(v)}$$

To establish the asymptotic distribution of the statistics V_k^v with the known functions of regression μ(t) and conditional variation σ(t), the notation is introduced

$$V_T^v(t) = \begin{cases} 0, & \text{если} \quad 0 \le t < \frac{1}{T+1} \\ V_{[(T+1)]}^v, & \text{если } 1/(T+1) \le t < 1/(T+1) \\ 0, & \text{если} \quad \frac{T}{T+1} \, 0 \le t < 1 \end{cases} \quad (2.12)$$

If $\tau_1 = \tau_2 = 1$, i.e. hypothesis H0 is true, then under a certain condition [27] the limiting relation

$$\lim_{T \to \infty} \frac{\sqrt{T} * \sigma}{\sigma_w} * \sup_{\delta < t < 1-\delta} V_T^V(t) \underline{d} \sup_{\delta < t < 1-\delta} \left(t(1-t) \right)^{-v} B(t) \quad (2.13),$$

where $\underline{d}$ means convergence in distribution - any number from (0,1) $0 \le v \le 1$ B(t)- standard Brownian motion by [0,1]; $\sigma = \lim_{T \to \infty} S_T / T$ and

$$\sigma_w^2 = E\left(Z_1^2 - EZ_1^2 \right)^2 + 2 \sum_{i=2}^{T} E\left(Z_1^2 - EZ_1^2 \right)\left(Z_i^2 - EZ_i^2 \right)$$

The approximation of the random variable from the right-hand side of (2.13) with ν=1/2 is obtained using the formula [28]:

$$P\left\{ \sup_{h < t < 1-l} \left(\frac{B^2(t)}{t(1-t)} \right)^{1/2} \ge x \right\} = \frac{x \exp\left\{-x^2/2\right\}}{(2\pi)^{1/2}} \left\{ \log \frac{(1-h)(1-l)}{hl} - \frac{1}{x^2} \log \frac{(1-h)(1-l)}{hl} + \frac{4}{x^2} + \theta\left(\frac{1}{x^4} \right) \right\} \quad (2.14)$$

If there is a jump in the volatility of the function, i.e. heteroscedastic regression model $\sigma^2(t) \ne const$, the estimate of the point of change is defined as [27]:

$$\tilde{k} = \arg \max_k \left| \tilde{V}_k^v \right| \quad (2.15)$$

where

$$\left| \tilde{V}_k^v \right| = \left(\frac{k(T-k)}{T^2} \right)^{1-v} \left(\frac{1}{T-k} \tilde{R}_{T-k} - \frac{1}{k} \tilde{R}_k \right) \quad (2.16)$$

$$\tilde{R}_k = \sum_{i=1}^{k} \tilde{W}_i^2, \; \tilde{R}_{T-k} = \sum_{i=k+1}^{T} \tilde{W}_i^2, \; \tilde{R}_T = \sum_{i=1}^{T} \tilde{W}_i^2, \; \tilde{W}_i = Y_i - \hat{\mu}(t_i) \quad (2.17)$$

$\hat{\mu}(t)$ - evaluation of the regression function μ(t). The limiting relation (2.13) under certain conditions from [27] remains valid

when $\hat{k}(v)$ and V_k are replaced by $\hat{k}(v)$ and $\hat{V}_k^v$.

To estimate $\sigma_0(t)$ (2.3), it is advisable (especially when there is a certain amount of outliers or when the observation distribution function has "heavy tails") to use the absolute deviation estimate [29]

$$\hat{\sigma}_d^2(t) = \left(\sum_{i=1}^{n} K_{n,b}(t_i-t)\left|Y_i - \hat{\mu}(t_i)\right| \right) / \sum_{i=1}^{n} K_{n,b}(t_i-t) \quad (2.18),$$

where

$$K_{n,b}(t) = \frac{1}{b} K\left(\frac{t_i-t}{b} \right),$$

K(.)-nuclear function, and $b = b_n$-sequence of strip widths. As a nuclear function K(t), one can adopt the standardized Epechnikov core

$$K(t) = 3\left(1 - \frac{t^2}{5} \right) I(t^2 \le 5) / (4\sqrt{5}) \quad (2.19)$$

or Gauss core

$$K(t) = \frac{1}{\sqrt{2\pi}} \exp\left\{-t^2/2\right\} \quad (2.20)$$

Here I(A) is an indicator function of set A:

$I(A) = 1 \, at \, x \in A$ and I(A)=0 at $x \notin A$.

To estimate the widths of the bands in (2.18), we use the method using the χ^2 statistics of $R^2(y, \hat{\mu}, \hat{\sigma}_{bn})$, which is defined as [30]

$$R^2(y, \hat{\mu}, \hat{\sigma}_{b_n}) = \sum_{i=1}^{n} \frac{(Y_i - \hat{\mu}(t_i))^2}{\hat{\sigma}_{b_n}^2(t_i)} \quad (2.21),$$

where $y = (Y_1, ... Y_n)$; $\hat{\sigma}_{b_n}^2$ - is the estimate of $\hat{\sigma}_d^2(\cdot)$, defined by formula (2.17) with the strip width b=bn. The expected value of χ^2 - Pearson statistic (2.21), ER2, is approximately equal to the number of degrees of freedom n.

Since R2 depends on the estimated (conditional) variation, equality (2.21) can be used for the optimal choice of the bandwidth bn:

$$b_{n,opt} = \arg \min_{b_n} R^2(y, \hat{\mu}, \hat{\sigma}_{b_n}) \quad (2.22)$$

To detect the volatility points of the time series Y(t), we preliminarily construct the estimate $\hat{\mu}(t)$, the trend component μ(t) and allocate the irregular stochastic component $a(t) = Y(t) - \hat{\mu}(t)$ with the property of stationarity (in broad sense) i.e. with constant mean m and covariance function $K(t_1, t_2)$, depending only on the difference of the arguments. The verification of the a(t) process for stationarity is carried out using a sample autocorrelation correlogram function [31], which for a stationary time series must rapidly decrease with increasing delay time.

Consider a stationary time series $\{a_t\}$ with zero value and variances (variances) σ_t^2, t=1,...,T. Let $C_k = \sum_{k=1}^{n} a$ be the cumulative (accumulated) sum of the series $\{a_t\}$ and $D_k = \frac{C_k}{C_T} - \frac{k}{T}, k = 1,...,T; D_0 - D_T = 0$ (2.23) centered and normalized cumulative sum of squares.

For a constant variance $\sigma_t = \sigma$, the sequence $\{D_k\}$ oscillates with respect to zero. If sudden changes in the variance σ_t are possible, then the graph of the dependence of D_k on k will be concluded with a high probability within certain limits, which are

calculated from an analysis of the asymptotic distribution of D_k relative to some constant dispersion. Given this property, in [23] is

$$k^* = \arg\max_k D_k \quad (2.24)$$

If the maximum in (2.24) exceeds the above-mentioned boundary levels $|D_k|$ level of significance α, then according to the iterative method of centered cumulative sums of squares (ICSS), the k* value from (2.24) is taken as the estimate of the point of variation of the variance.

In [27], a modified ICSS algorithm was proposed (let's call it MICSS), in which the statistics D_k is replaced by V_k^v defined by expression (2.10). The critical values of $C_a^v(T)$ of the statistics V_k^v for v=0 are given in [23], and for v=1/2 in [28]. For the level of significance, α=0.05 is usually applied.

In the MICSS method in step 1, the statistics

$$M(t_1 : t_2) = \max_{t_1 \le k \le t_2} \sqrt{(t_2 - (t_2 - t_1 + 1))/2} / D_k(a[t_1 : t_2]) \quad (2.25)$$

replaced by statistics

$$M(t_1 : t_2) = \max_{t_1 \le k \le t_2} \frac{\hat{\sigma}}{\hat{\sigma}_w} \sqrt{t_2 - t_1 + 1} \left| \hat{V}_k^v(a[t_1 : t_2]) \right| \quad (2.26),$$

where a $[t_1; t_2]$ means that the sample of values of at is taken at $t_1 \le t \le t_2$, where $t_1 < t_2$ $\hat{V}_k^v a[t_1; t_2]$ -statistics (2.16) defined on the interval $[t_1; t_2]$.

In equality (2.26), the estimate σ is determined by the formula $\hat{\sigma} = \frac{1}{T} R_T$, where R_T is defined in (2.17). A consistent estimate for σ_w^2, which is the variance of the random variable $Z_t = (Y_t - \mu(t))/\sigma_0(t) = \varepsilon_t$, is obtained on the basis of analysis [32] of the sample stationary sequence $\{\hat{Z}_i, i = 1, ..., T\}, \hat{Z}_i = (Y_i - \hat{\mu}(t_i))/\hat{\sigma}_0(t_i)$ with $\hat{\sigma}_0(t_i) = \hat{\sigma}_d(t_i)$. For this, statistics are entered

$$B_{n,p} = \left\{ \frac{C_p}{n - l_n + 1} \sum_{j=0}^{n-l_n} \left(\frac{1}{\sqrt{\ln}} \left| S_j(l_n) - l_n \bar{\hat{Z}}_n \right| \right)^p \right\}^{1/p}, \quad p \ge 1 \quad (2.27),$$

where $\{l_n, n \ge 1\}$ is a sequence of positive integers with $1 \le l_n \le n$ and the notation is entered

$$S_j(k) = \sum_{i=j+1}^{j+k} \hat{Z}_i, \quad \bar{\hat{Z}}_l = \frac{1}{n} \sum_{i=1}^n \hat{Z}_i, \quad C_p = 2^{-p/2} \cdot \sqrt{\pi} \Gamma\left(\frac{p+1}{2}\right) \quad (2.28)$$

where $\Gamma(\bullet)$ – is the gamma function? If $l_n \to 0$ and $l_n / n \to 0$ as $n \to \infty$ then as $n \to \infty, B_{n,p}$ σ [32]. So, the estimate $\hat{\sigma}_w$ in (2.26) for sufficiently large n (large T) can be taken as $\hat{\sigma}_w = B_{n,p}$ (in the calculations, p=1 with c c1= Γ(1) and $l_n = \ln(n)$ were taken).

The critical values $C_{\alpha,0}(T^*)$ of statistics (2.25) for different $T^* = t_2 - t_1 + 1$ are defined in [23]. At a level of significance α =0.05, the asymptotic (for large T *) values of $C_{\alpha,0}(T^*)$ in [23] are determined to be 1.27; 1.30; 1.31; 1.33; 1.358 at T*=100; 200; 300; 400; 500; ∞ respectively. Linear interpolation for intermediate T* values gives the following values

$$C_{\alpha,0}(T^*) = C_{\alpha,0}(100 * i) + (\frac{T^*}{100} - i) * \left(C_{\alpha,0}(100(i+1) - C_{\alpha,0}(100 * i) \right)$$

if $100i < T^* < 100(i+1)$ at $i = 1, 2, 4$ и $C_{\alpha,0}(T^*) = 1,31$ at $300 \le T^* \le 400$

From (2.7), (2.10), (2.16), taking into account the estimate $\hat{\sigma} = S_T / T$, the statistics (2.26) with v=0 can be written as

$$M(t_1 : t_2) = \max_{t_1 \le k \le t_2} \frac{\hat{\sigma}^2}{\hat{\sigma}_w} \sqrt{t_2 - t_1 + 1} |D_k| \quad (2.29)$$

Therefore, the critical values $C_{\alpha,0}(T^*)$ of the statistics (2.26) with v=0 are written as $\tilde{C}_{\alpha,0}(T^*) = \frac{\hat{\sigma}^2 \sqrt{2}}{\hat{\sigma}_w}$

To determine the critical values $\tilde{C}_{\alpha,1/2}(T^*)$ of the statistics (2.26) with v=1/2, you can use the asymptotic relation (2.13), which with $\tilde{\delta} = \frac{\delta}{T}, \delta = \frac{1}{T(T+1)}$, $\tilde{t} = \frac{t}{T}, \hat{V}_T^v(\hat{t}) = V_{[\frac{T+1}{T}t]}^v$ and: $B(\hat{t}) = B(T\hat{t})$ is written as

$$\lim_{T \to 0} \frac{\sqrt{T} \cdot \hat{\sigma}}{\hat{\sigma}_w} \cdot \sup_{\tilde{\delta} < \tilde{t} < 1 - \tilde{\delta}} \left| \hat{V}_T^v(t) \right| \underline{d} \sup_{\tilde{\delta} < \tilde{t} < 1 - \tilde{\delta}} (\tilde{t}(1 - \tilde{t}))^{-v} \cdot \tilde{B}(\tilde{t}) \quad (2.30)$$

Solve x nonlinear equation

$$\frac{x \exp\left\{ -\frac{x^2}{2} \right\}}{\sqrt{2\pi} \left\{ 2\log\left(\frac{1}{\tilde{\delta}} - 1\right)\left(1 - \frac{1}{x^2}\right) + \frac{4}{x^2} \right\}} = \alpha \quad (2.31),$$

obtained from the curve of part (2.14) with $h = l = \tilde{\delta}$. Under the initial condition x=x0 (for example, x_0=100) we find the solution x* of equation (2.31) for fixed α in the MATLAB system. Then the value $\hat{C}_{\alpha,\frac{1}{2}}(T^*) = \sqrt{\frac{T^*}{T}} - x^*$ can be taken as the critical value of the statistics (2.26) with $v = 1/2$.

Modified Method for Predicting the Status of the Cardiovascular Systems on the Basis of Identifying the Boundaries of Quasistationary Sections of The Time Series

To solve the problem of predicting complex digital signals, it was proposed in [17] to use the last "stable" fragment of a time series for training a predictor, obtained by dividing into quasi-stationary regions. Such a task is known [33,34] as a problem of detecting a "debugging", i.e. changes in the probability properties of the signal. Such a definition can give for complex signals an inaccurate border of a true change in the dynamics of a signal. To detect the change in the dynamics of a complex signal, it was proposed to use an estimate of the local fractal dimension based on the fractality index μ of the time series [19].

As shown in section 1, it is most appropriate to characterize the non-stationary time series with two indicators b(t) and μ(t) included in the power representation (1.14) of the amplitude variation $V_f(\delta)$ of the function f(t) graph corresponding to the original time series. In this case, the value of b(t) is closely related to the volatility (conditional variation) of the time series, according to a jump in which in [18] an algorithm for segmentation of a non-stationary time series was developed.

Given this property of the function b(t), we can propose the following method for detecting the boundaries of quasistationary portions of the time series. Using the sliding window's original signal f(t), the derived estimation signal b(t), which characterizes the local fractal dimension of the time series, is plotted. In the

derivative and in the original signal according to the algorithm [18] are the points of disorder. The positions of these points are compared and the average is taken, which will be the boundary of the adjacent quasi-stationary areas. The fragment of the last generated quasi-stable signal dynamics is used as a learning set for the predictor.

To increase the efficiency of the Q-event forecast, i.e. a record of the rhythmogram of a heartbeat, a larger threshold Q, in the selected last quasi-stationary segment of the time series of the rhythmogram, a combination of the algorithm of repeated intervals using long-term memory [6] with the method of reaching the threshold of change [16] is used.

The RIA approach is based on the analysis of the nonlinear component of the long-term dependence of repeated intervals rj between events xi exceeding a certain defined threshold Q (Q-events).

Sometimes, instead of the threshold K, the average repeat interval or repetition period is set to $R_Q = 1 / \int_0^\infty P(x)dx$, where P(x) is the distribution of events $\{x_i\}$.

In this approach, to predict random signals with fractal properties, the mathematical apparatus of interval statistics is used when the probability $W_Q(t;\Delta t)$ that during the time Δt appears at least one Q-interval if during the time t for the last Q- event Q event has occurred.

$$W_Q(t;\Delta t) = \frac{(\delta(Q-1)\Delta t / R_Q)}{\dfrac{t}{R_Q} + (\delta(Q)-1)\Delta t / R_Q} \quad (3.1)$$

where $\partial(Q)$ characterizes the density distribution of the lengths r of repeated intervals

$$P_Q(r) \sim (r / R_Q)^{-\delta(Q)} \quad (3.2)$$

The method of reaching the threshold of change is based on the transformation of the original signal into a series of times to reach the threshold P:

$$x(t) \equiv \{x_1, x_2, ..., x_N\} \rightarrow x' \equiv \{x_1', x_2', ..., x_N'\}, x_i = \frac{x}{t_i},$$

$$i = 1,...N, x_j' = x(t_j'), j = 1,...,N(x_i' - x_{i-1}'), / x_{i-1}'| > p \quad (3.3)$$

$$p > 0, i = 2, 3, ..., N'\tau = \{\tau_1, \tau_2, ..., \tau_{N-1}\}, \tau_i = t_{i+1}' - t_i\, i = 1,...,N' - 1$$

where N is the number of samples in the original signal, x' is the converted signal, where only those values are left, the relative difference between which exceeds a certain predetermined threshold p, N' is the number of samples in the converted signal, τ-time series to reach the threshold of change p, where each value of τ-means the time it took the signal to exceed the threshold of change p. The coincidence of the signs of differences $x_i' - x_{i-1}'$ and $(\tau_i - \tau_{i-1})$ indicates the persistence (steady growth) of both the values x_i' and the time τ_i between records $x_i' and x_{i-1}'$. In application to records (extreme events) heartbeat means records and steady increase of intervals between consecutive records. Under the records of the heartbeat can be understood Q-events, and the intervals-Q intervals. Transformation (3.3) can be used to

pre-select heartbeat records in the Q-interval prediction procedure, based on the RİA technique, which allows speeding up the last procedure and increasing the reliability of the prediction obtained.

References

1. Bunde A, Eichner JF, Havlin S, Kantelhard JW (2003) The effect of long correlation on the return periods of rare events. Physica A 330: 1-7.

2. Bogachev MI, Eichner JF, Bunde A (2007) Effect of Nonlinear Correlations on the Statistic of Renturn Intervals in Multifractal Date Sets. Physical Review Letters 99(24): 240601.

3. Hllirberg S, Altmann EG, Holstein D, Kantz H (2007) Precursors of ectreme increments. Phys Rev E 75: 016706.

4. Bogachev MI, Eichner JF, Bunde A (2008) The effects of multifractality on the statistics of return intervals. The European Physical Journal Special Topics 161: 181-193.

5. Bogachev MI, Eichner JF, Bunde A (2008) On the Occurrence of Extreme Events in Long-term Correlated and Multifractal Data Sets. Pure appl Geophys 165: 1195-1207.

6. Bogachev MI, Kireenkov I, Nifontov E, Bunde A (2009) Statistics of return intervals between long heartbeat intervals and their usability for online prediction of disorders. New Journal of Physics 11: 063036.

7. Bogachev MI (2009) On the issue and predictability of emissions of dynamic series with fractal properties when using information about the linear and non-linear components of the long-term dependence. News Universities of Russia, Radio electronics, 5b: 31-40.

8. Bogachev MI (2010) Comparative analysis of the noise immunity of methods for predicting emissions of random signals with fractal properties using information about short-term and long-term dependencies. New Universities of Russia, Radio electronics, 1: 11-21.

9. Bogachev MI, Bunde A (2011) On the predictability of extreme events in records with linear and nonlinear long-range memory: Efficiency and noise robustness. Physical A 390: 2240-2250.

10. Ludescher J, Bogachev MI, Cantelhard JW, Schumann A, Bunde A (2011) On spurious and corrupted multifractality: The effect of additive noise, shorm-term memory and periodic trends. Physica A 390: 2480-2490.

11. Abdullaev NT, Dyshin OA, Ibragimova IJ (2016) The effectiveness of the use of long-term correlations in the differential diagnosis of cardiac-vascular system. Biotechnosfera 5: 27-33.

12. Abdullaev NT, Dyshin OA, Ibragimova IJ (2018) Comparative analysis of emission prediction methods in the dynamics of the heartbeat in the presence of additive random interference. Proceedings of the XIII International Scientific Conference FREME-2018. Vladimir-Suzdal, Book 1, p. 169-173

13. Abdullaev NT, Dyshin OA, Ibragimova IJ (2018) Prediction of emissions of time series with fractal properties of heartbeat intervals. Proceedings of the XIII International Scientific Conference FREME-2018. Vladimir-Suzdal, Book 1, p. 173-177.

14. Abdullaev NT, Dyshin OA, Ibragimova IJ (2016) Diagnosis of Heart Diseases Based on the Repeat Interval Statistics of Extreme Events of Heart Rate, Trudy FREME. Vladimir 1: S.43-46

15. Abdullayev NT, Dyshin OA, Ibragimova ID (2017) Prediction of the cardiovascular System on the assessment of Repeated Extreme Values Heartbeat Intervals and Times to Achieve Them in the Light of short-term and long-term Relationships. Journal of Biomedical Engineering and Medical Devices USA, Canada, Europe, Great Britain, 2(2): 1000126

16. Yegoshin AV (2007) Analysis of the time the signal reaches the threshold of change in the task neural network time series forecasting. Information-computational technologies and their applications. Collection of articles of the VI International Scientific and Technical Conference.: RIO PGSA, Penza, p.74-76.

17. Sidorkina IG, Egoshin FV, Shumnev DS, Kudrin PA (2010) Prediction of complex signals based on the selection of the boundaries of implementations of dynamic systems. Scientific and Technical Bulletin of the St. Petersburg State University of Information Technologies, Mechanics and Optics, 2(66): 49-53.

18. Abdullaev NT, Dyshin OA, Ibragimova ID, Akhmedova Kh R (2019) Segmentation of non-stationary physiological signals with fractal properties. Information Technologies 5: 300-312.

19. Starchenko NV (2005) Fractality index and local analysis of chaotic time series Dis. Candidate of Physics and Mathematics, Moscow, MEPI, pp. 119.

20. Kronover RM (2000) Fractals and chaos in dynamic systems. Postmarket, pp. 350.

21. Feder E (1991) Fractals. Mir pp. 374.

22. Mantegna RN, Stanley G Yu (2009) Introduction to econophysics. Correlations and complexity in finance. Pers eng, Book House, Librocom, pp. 192.

23. Inclan C, Tao G (1994) Use of Cumulative Sums of Square for retrospective detection of changes of varicence. Journal of American Statistical Association 89(427): 913-923.

24. Krishnaiah P, Miao B (1988) Revive about estimation of chance-points. Handbook of Statistics 7: 375-402.

25. (1994) Change-point problems. Lecture notes and monograph series. E Carlstein, HG Muller, D-Siegmunal (Eds.), Institute of Mathematical Statistics, Hargward, CA, Canada, 23: 385.

26. Mandelbrot BB (1982) The Fractal Geometry of Nature. WH Freeman, Sun-Francisco, pp. 468.

27. Chen G, Choi Y, Zhou Y (2005) Nonparametric estimation of structural change points in volatility models for time series. Journal of Econometries 126(1): 79-114.

28. Csorgo M, Horvath L (1997) Limit theorems in chance-point analysis. Wiley, New York pp. 438.

29. Xia Y, Tong H, Li WK (1988) Abstitude deviation estimation of volatility in nonparametric models. Research report No.177, Department of statistics and actuarial sciences the University Hong Kong, Hong Kong, pp.133-178.

30. Choin JM, Muler HG (1999) Nonparametric quasilikelihood. The annual of statistics 27: 36-64.

31. Smolentsev NK (2008) Fundamentals of the theory of wavelets. Wavelets in MATLAB, DMK Press, pp. 448.

32. Peligrad M, Shao QM (1995) Estimation of the variance of partial sums for ρ- mixing random variables. Journal of multivariate analysis 52: 140-157.

33. Shiryaev AN (1976) Statistical sequential analysis. Science pp. 272.

34. Darkhovski BS (1994) Nonparametric methods in change-point problems a general approach and some concrete algorithms. IMS Lecture Notes-Monograph Series (23): 99-107.