Automation and Remote Control, Vol. 66, No. 12, 2005, pp. 1871–1896. Translated from Avtomatika i Telemekhanika, No. 12, 2005, pp. 3–30.
Original Russian Text Copyright © 2005 by Grebenyuk.

REVIEWS

Methods of Analyzing the Nonstationary Time Series with Implicit Changes in Their Properties

E. A. Grebenyuk

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

Received January 27, 2005

Abstract—The problems of analysis of the nonstationary time series with explicit and implicit changes in their properties were discussed. The main approaches and methods of analysis of the nonstationary processes were described. Consideration was given to the trend-stationary and difference-stationary processes. The algorithms to determine the process type hinge on verifying hypotheses like "initial process is a DS-process (TS-process)" against the alternative ones. The notions of false regression and cointegration were discussed. The problems arising at analysis of the varying-property processes and methods of their solution were reviewed. Consideration was given to the cases of explicit and implicit changes and algorithms of detection in the current and *a posteriori* modes.

1. INTRODUCTION

Studies of the observation time series describing economic and social processes represent an important part of analysis aimed at elucidation of the regularities in the processes under consideration, construction of models for prediction, monitoring, and control of the process dynamics, and generation of alarms at undesirable consequences. Prediction, monitoring, detection of the interconnections between individual process components, and analysis of the consequences of possible changes are topical in diverse fields such as macroeconomics, sociology, finances, and marketing studies. The nonstationary time series are exemplified by the series describing dynamics of the stock market, the macroeconomic indices, and the indices characterizing the state of the social sphere. The majority of the economic and social processes that obey the nonstationary time series include changes in the properties or a set of states described each by a certain model. The passages between states are characterized by changes in process behavior, may occur under the action of environmental factors at arbitrary, unknown-in-advance time instants, and are not necessarily noticeable without special statistical processing.

The nondetected changes give rise to erroneous conclusions at process analysis. In the case of possible changes, the process analysis differs from that in the case of no changes. The recent interest to analysis of the varying-property processes is due to the following reasons:

(1) Availability of an adequate[1] model of the observation time series describing system behavior is the necessary condition for analysis. If properties change within the model interval, then a model disregarding them is no more adequate.

(2) If the process properties may change, then, after constructing the model over a no-change interval, for successful prediction and monitoring one must check it for adequacy.[2]

(3) The power of the tests for process type and cointegration drops dramatically in the presence of changes in the processes.

[1] A model is regarded as adequate if it corresponds to the original process in what concerns reflection of its given properties. The error of such a model should not exceed the predefined limit values.

[2] In the sense of constraints on the model error.

The following section presents the main notions used to analyze the nonstationary time series and overviews the tests for cointegration and process type in the absence of changes in properties. Section 3 discusses different approaches to detection of changes and considers the methods of *a posteriori* and current detection of changes and process analysis.

2. ANALYSIS OF THE NONSTATIONARY TIME SERIES IN THE ABSENCE OF PROPERTY CHANGES

2.1. Trend-stationary and Difference-stationary Processes

2.1.1. Necessary definitions. In the general case, the random process is a function of two dissimilar variables, the random variable w and time t: $y(w,t)$, $t = 0, \pm 1, \pm 2, \ldots, \pm t, \ldots$. The time series results from the observations of the random process at a fixed time intervals. It is assumed that the time series is a sample y_t, $t \in T$, from a sequence of random variables $y(w,t)$, $t = 0, \pm 1, \pm 2, \ldots, \pm t, \ldots$.

The following problem is solved at modeling the real processes: given is a realization of a series, needed is to select a model capable of generating such a realization (data-generating model).

The properties of a stationary process do not vary in time. Its values oscillate about a constant mean value, the oscillation variance is constant, and the values of the autocorrelation function decrease with increase in the inter-observation time.

The mean value and/or variance of the nonstationary process depend on time, its variance with time tends to infinity, the autocorrelation function does not decrease with increased inter-observation time, and on the finite samples the sample autocorrelation function slowly fades out.

The typical time series may have four components [1]:

(1) trend or systematic motion;

(2) oscillations about the trend;

(3) seasonal changes;

(4) "nonsystematic" or "irregular" "random" component.

The mathematical description of the time series is a sum of one or more components. The present review does not discuss the seasonal changes. According to [1], by the trend is meant the stable systematic change in series direction over a long time period.

To describe processes with trends, the following representations were widely used till the 1970's:

$$y_t = f(t) + z_t, \tag{1}$$

where $f(t)$ is a purely deterministic component, z_t is a random stationary process representing a sum of the oscillatory and random components

$$z_t = \sum_{j=1}^{k} a_j z_{t-j} + \sum_{l=1}^{q} b_l \varepsilon_{t-l} + \varepsilon_t,$$

where a_j and b_l are constant coefficients and ε_t is a white-noise process with zero expectation and variance σ_ε.

Lack of a random factor in the mechanism of trend origination is the disadvantage of the models like (1). The processes where the trend can be both deterministic and stochastic obey the autoregression models with unit root:

$$y_t = \mu + \rho y_{t-1} + z_t, \tag{2}$$

where μ is a constant value defining the drift, $\rho = 1$, and z_t is a stationary process.

Relative to the deterministic trend, processes like (1) are called the trend-stationary processes (TS-processes). The stationary process is a special case of the zero-trend TS-process ($f(t) = 0$). Processes (2) are called the difference-stationary or integrated processes (DS-processes) [2, 3].

Definition. The time series y_t is called the stationary series relative to the deterministic trend $f(t)$ or trend-stationary if the series $(y_t - f(t))$ is stationary. The series that is stationary relative to some deterministic trend is called the TS-series.

Definition. The nonstationary process y_t is called the integrated process of the order of $d \geq 1$, (DS-series) if

(1) y_t is not stationary or stationary relative to the deterministic trend;

(2) the series $\Delta^d y_t$ resulting from the d-times differentiation of the series y_t is a trend-stationary series;

(3) the series $\Delta^{d-1} y_t$ resulting from the $(d-1)$-times differentiation of the series y_t is not a trend-stationary series.

The integrated process of the order d is denoted by $I(d)$.

2.1.2. Need for determining the process types and the attendant problems. Memory about the random actions of process (1) depends on the orders k and q of the process z_t, process (2) retaining infinitely long the memory of each random action ε_t:

$$y_t = \mu t + \sum_{i=1}^{t} z_i + z_0 = \mu t + \sum_{i=1}^{t} \left(\sum_{j=1}^{k} a_j z_{i-j} + \sum_{l=1}^{q} b_l \varepsilon_{i-l} + \varepsilon_i \right) + z_0.$$

Therefore, in the TS-series the effect of random actions fades out with time, whereas in the DS-series there is no decay, and each individual random action ε_t influences with identical force all the subsequent values of the series.

Determination of the type of process is a necessary step in constructing cointegration, stationary linear combination of nonstationary processes, because it is possible only if both processes are integrated processes of the same order.

Particular processing scenarios are used for each of the process types. Erroneous determination of the process type at analysis leads to undesirable consequences [4].

Construction of a parametric model of the nonstationary series is usually preceded by series "stationarization" done either by extracting the stationary trend, which lies in estimation of the trend function parameters by the LMS method and subtraction of its value from the original series, or by applying to the original series the difference operator

$$\Delta y_t = (1 - D) y_t, \quad \Delta^2 y_t = (1 - D)^2 y_t, \ldots, \Delta^d y_t = (1 - D)^d y_t,$$

where $D y_t = y_{t-1}, \ldots, D^d y_t = y_{t-d}$, D being the delay operator. Choice of the series "stationarization" procedure depends on the type of the nonstationary process. If the operation of extraction of the deterministic trend is applied to the DS-process, then, as was shown in [5, 6], it may result in false periodicity of the detrended series.

As follows form the definition of the DS-series, the series $\Delta^{d-1} y_t$, which is the result of the $(d-1)$-times differentiation of y_t, becomes stationary. However, if differentiation is applied to a stationary or trend-stationary process, then one may get as the result a process like the sliding mean for which no autoregression representation exists. In this case, the differentiated series is autocorrelated despite the fact that the original series is a sum of the deterministic linear trend and white noise (the Slutsky effect) [7].

The majority of the existing methods of determination of the process type are based on verifying the hypothesis that the original series is a DS-process or the "unit-root hypothesis" that can be formulated as follows.

Let the observation time series $y_1, y_2, \ldots, y_t, \ldots$ be described by the model

$$y_t = \mu + \sum_{j=1}^{k} \rho_j y_{t-j} + \varepsilon_t, \tag{3}$$

where ε_t is a stationary process, ρ_j are constants, and μ is the constant defining the value and direction of drift. Model (3) can be rearranged in

$$\Delta y_t = \mu + \alpha y_{t-1} + \sum_{j=1}^{k} \alpha_j \Delta y_{t-j} + \varepsilon_t, \tag{4}$$

where $\alpha = \sum\limits_{j=1}^{k} \rho_j - 1$. If the coefficient $\alpha = 0$, then y_t has a unit root and is a nonstationary or integrated first-order process ($I(1)$); if $\alpha < 0$, then it is an integrated zero-order, that is, stationary, process ($I(0)$). If $\alpha > 0$, then it is called the "explosive" process.[3]

The point estimate of regression in the case of processes $I(1)$ has four essential differences from the estimates of the processes $I(0)$ [8]:

(1) The estimates converge with the rate T and not $\sqrt{T}$ as in the case of stationary process. If $\alpha < 0$ and ε_t is a stationary process, then the limit distribution of the estimate $\widehat{\alpha}$ is the normal one [9]:

$$\sqrt{T}(\widehat{\alpha} - \alpha) \to N(0, \alpha^2). \tag{5}$$

If $\alpha = 0$, then the limit distribution of the estimate $\widehat{\alpha}$ has the form

$$T(\widehat{\alpha} - 1) \Rightarrow \frac{\left\{\int W dW - \frac{1}{2}(k-1)\right\}}{\int W^2 dW}, \tag{6}$$

where $W(t)$ is the standard continuous-time Brownian motion. This estimate was first established for the AR(1) model [10] and later for the general case [11]. Therefore, α is a "superconsistent" estimate converging with the rate T instead of $\sqrt{T}$.

(2) The limit distribution (6) was tabulated by Dickey [12] and Fuller [13]. It is a nonstandard asymmetrical distribution with asymptotic lower and upper 5% quantiles: 8.1 and 1.28.

(3) For the $I(1)$ process, the estimate $\widehat{\alpha}$ is a consistent estimate α even if the errors of regression of y_t to y_{t-1} are serially correlated and correlated with the regressors or their differences.

(4) The limit distribution of the estimate depends both on the true trend and the trend function and its estimated parameters.

Since the point estimates of the regression coefficients have different statistical properties, they prevent from discriminating the processes $I(0)$ and $I(1)$.

2.1.3. Criteria for process type. Many criteria such as those of Dickey–Fuller, Perron, Park, Kwiatkowski, *et al.* [14] were developed to determine the order of process integration.

Extended Dickey–Fuller criterion. Most popular is the set of Dickey–Fuller criteria [15, 16]. For model (4), verified is the hypothesis H_0: $\alpha = 0$ against the hypothesis H_1: $\alpha \neq 0$. If the

[3] Processes with $\alpha > 0$ are disregarded here.

hypothesis H_0 is rejected, then the process cannot be regarded as stationary. The t statistic $t = \dfrac{\widehat{\alpha}}{\widehat{\sigma}_\alpha}$ obeying the Dickey–Fuller distribution is used to verify the hypothesis H_0. If $|t| > t_{\text{crit}}$, then the zero hypothesis is rejected and the series is regarded as belonging to the TS type; otherwise, the zero hypothesis is accepted. The form of the Dickey–Fuller distribution, its quantiles and critical values are essentially dependent on whether the estimated model includes a constant or trend. There exist different procedures of using the Dickey–Fuller test of which that suggested in [17] is best known. Since it was described in detail in [2] and [3], we confine ourselves only to its main steps. It begins with analysis of the fullest model including the constant and trend:

$$\Delta y_t = \mu + \alpha y_{t-1} + \beta t + \sum_{j=1}^{k} \alpha_j \Delta y_{t-j} + \varepsilon_t.$$

For this model, one first verifies the hypothesis $H_0 : \alpha = 0$. If it is rejected, the series is classified with the TS type; if it is not rejected, then the parameters are estimated and significance of the coefficient β is verified for the equation

$$\Delta y_t = \mu + \beta t + \sum_{j=1}^{k} \alpha_j \Delta y_{t-j} + \varepsilon_t.$$

If the coefficient is significant, then the verification procedure is completed; otherwise, the hypothesis $H_0 : \alpha = 0$ is verified for the model of the form

$$\Delta y_t = \mu + \alpha y_{t-1} + \sum_{j=1}^{k} \alpha_j \Delta y_{t-j} + \varepsilon_t.$$

If this hypothesis is rejected, the series is an autoregression process belonging to the type TS; if the hypothesis is not rejected, then the parameters are estimated and significance of the free term is verified for the equation

$$\Delta y_t = \mu + \sum_{j=1}^{k} \alpha_j \Delta y_{t-j} + \varepsilon_t.$$

If the free term is significant, then the type determination procedure is completed; otherwise, the hypothesis $H_0 : \alpha = 0$ is verified for the models like

$$\Delta y_t = \alpha y_t + \sum_{j=1}^{k} \alpha_j \Delta y_{t-j} + \varepsilon_t.$$

If this hypothesis is rejected, then the series is an autoregression process belonging to the type TS; if the hypothesis is not rejected, then the series is classified with the DS-series.

The above procedure of determining the process type does not come only to a single application of the test, but includes verification of correct determination of model specification—availability of the free term and trend.

Low power which often leads to confirmation of the zero hypothesis in the cases where this is not true is the main disadvantage of the Dickey–Fuller and many other similar criteria. The power of criteria decreases with availability of structural changes in the processes. If in the course of process analysis the instants of such changes are determined and the intervals without them are extracted, then the power of Dickey–Fuller criteria is increased.

Phillips–Perron criterion of membership of the series y_t to the DS class [18] comes to verifying the hypothesis $H_0 : \varphi = 0$ for the model

$$\Delta y_t = \varphi y_{t-1} + \alpha + \beta t + \varepsilon_t, \quad t = 2, \ldots, T, \tag{7}$$

where the parameters α and β may be zero. However, in contrast to the Dickey–Fuller criterion, the random components ε_t with zero expectations can be autocorrelated (with sufficiently fast decrease of the autocorrelation function), have different variances (heteroskedasticity), and have distributions that need not be normal but are such that $E|\varepsilon_t|^\delta \le C < \infty$ for some $\delta > 2$.

The Phillips–Perron criterion is based on the t-statistic for verification of the hypothesis $H_0 : \varphi = 0$, but uses a variant of this statistic Z_t corrected to possible autocorrelatedness and heteroskedasticity of the series ε_t. At calculation of the statistic Z_t, estimated is the "long-term" variance of the series ε_t defined as

$$\lambda^2 = \lim_{T\to\infty} T^{-1} E(\varepsilon_1 + \ldots + \varepsilon_T)^2,$$

where $E(^*)$ is the expectation of *.

If ε_t are the residues of the estimated model (7), then the following estimate [19] can be taken as $(\lambda^2)^*$ for λ^2:

$$(\lambda^2)^* = \gamma_0^* + 2\sum_{j=1}^{l}\left[1 - \frac{j}{l+1}\right]\gamma_j^*, \tag{8}$$

where

$$\gamma_j^* = T^{-1}\sum_{t=j+1}^{l} \varepsilon_t^* \varepsilon_{t-j}^*$$

is the jth sample autocovariance ε_t. If l and T tend to infinity so that $(l/T)^{1/4} \to 0$, then, as was shown in [20], $(\lambda^2)^*$ is the consistent estimate of λ^2 and the asymptotic distributions of the statistic Z_t coincide with the corresponding asymptotic distributions of the statistic t_φ in the Dickey–Fuller criterion. The critical values of Z_t are either taken from tables [20] or calculated from formulas [21, 22].

Kwiatkowski–Phillips–Schmidt–Shin (KPSS) criterion [23] considers series membership to the TS type as the zero hypothesis. The series is assumed to follow the model

$$y_t = \beta t + w_t + \varepsilon_t,$$

where $w_t = w_{t-1} + \upsilon_t$ is the random walk with $\upsilon_t \sim N(0, \sigma_\upsilon^2)$ and ε_t being a stationary process that meets the same assumptions as the series ε_t in the Phillips–Perron criterion. The zero hypothesis assumes that the process variance ε_t is zero. The alternative hypothesis assumes that this variance is other than zero, therefore the series under consideration belongs to the class of DS series. The proposed criterion is that of maximum likelihood for verification of the zero hypothesis. Application of this criterion gives rise to the problem of choosing the window width l in the estimate [19] because the values of the criterial statistic are rather sensitive to l.

DF-GLS criterion [24] is more powerful asymptotically than the Dickey–Fuller criterion. It verifies the zero hypothesis for the model

$$\Delta y_t^d = \alpha_0 y_{t-1} + \alpha_1 \Delta y_{t-1}^d + \ldots + \alpha_p \Delta y_{t-p}^d + \varepsilon_t,$$

where y_t^d is a locally "detrended" series.

Despite the great diversity of the existing and developed methods, it deserves noting that the problem of discriminating the TS and DS processes in analysis of the actual series has not yet been solved once and for all, which is due to the low power of the developed procedures on short samples as well as to the presence of structural changes in the series under study which also make for reduction of the test power. Use of more than one criterion for analysis is not always a good issue because on the same series they mostly provide similar results.

2.2. Cointegration

2.2.1. A new approach to analysis of nonstationary series. The traditional approach to the problem of nonstationary data lay in formulating statistical models as relations between the first-order differences, that is, the growth rate. Yet the statistical model based exclusively on the differences can catch only the short-term process dynamics and does not allow one to analyze the long-term relations between the variables. The problem of developing methods following the possible long-term relations masked by the noise of short-term oscillations was solved in 1986 by Granger [25] who established that a certain combination of two or more nonstationary series can be stationary. The economic theory often makes predictions exactly of this kind: if there exist equilibrium relations between two economic variables, on short terms they can deviate from the equilibrium, but tend to equilibrium on longer terms.

For example, according to the theory of purchasing power parity, the motion of the exchange rates is defined by their tendency to establish equilibrium between different currencies on the assumption of their purchasing power. Therefore, the percentage rate of changes in the exchange rate must be equal to the difference between the percentage rates of changes in the process levels for the domestic and foreign commodities [26].

Stability of the demand for money is very important for carrying out an effective monetary policy. By the demand for money is meant the demand for real money remainders

$$\frac{M^d}{P} = f(I, \mathbf{R}),$$

where M^d is the demand for the nominal money remainders, P is the level of process, I is the real income, and $\mathbf{R}$ is the vector of profitability indices. In the empirical studies, various money aggregates are used as M^d, the values of GNP as I, the profitabilities $\mathbf{R}$ being represented by the interest rates. A large body of literature [27–29] exists on the empirical studies of relations between the money supply and the process level confirming or questioning the theoretical studies.

The notion of "cointegration" was introduced by Granger to denote a stationary combination of nonstationary variables. His studies initiated new approaches to the time series. The checks for stationarity and cointegration are mow standard procedures beginning specification of the dynamic econometric models. The cointegration analysis proved to be especially valuable for analysis of systems where great random disturbances affect the short-term dynamics, whereas the long-term oscillations are bounded by the general economic equilibrium relations.

Cointegration can exist only between nonstationary processes of the same order of integration, correlational bonds are possible between stationary processes, between processes of different types there are no bonds, and an attempt to establish them may result in a false regression [3, 26, 27].

2.2.2. False regression and cointegration. Granger and Newbold [30] estimated regression models like

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t, \tag{9}$$

Indices DJI and RTSI.

where x_t and y_t are unrelated random walks:

$$\Delta x_t = \varepsilon_{1t},$$
$$\Delta y_t = \varepsilon_{2t}.$$

Since x_t is not related with y_t, one may expect that the coefficient β_1 and the determination coefficient R^2 vanish. However, they established that often the zero hypothesis of no relation between the variables is not rejected for a very high value of R^2 and small values of the Durbin–Watson statistic, which suggested [30] that for small values of the Durbin–Watson statistic high values of R^2 may signal false regression. The fact established in [30] was explained analytically by Phillips [31] who demonstrated that the t-statistics in model (9) do not follow the Student distribution but tend to infinity with the sample volume. Whence it follows that any critical value of the t-statistic for rejection of the zero hypothesis $\beta_1 = 0$ increases with the sample volume. It was shown also in [4] that the Durbin–Watson statistic vanishes with increased sample if the variables are not interrelated or converges to a nonzero value, otherwise.

Preliminary stationarization of the DS-series by taking a finite number of differences renders impossible the analysis of long-term relations between the variables, although various researchers noted the fact that there exists a long-term relation between the nonstationary series. In particular, this concerns the financial indices and series of macroeconomic activities. We consider by way of example the graphs of the DJI (upper graph) and RTSI (lower graph) from June 1996 to October 1997. As can be seen from the figure, both indices increase and decrease actually synchronously. Over the period under study, the DJI and RTSI indices were related cointegrationally.

Cointegration was discovered by Granger [25, 32, 33], and the formal statistical analysis of the cointegrated processes was developed by R.F. Engle and C.W.J. Granger [34], S. Johansen [35, 36], and P.C.B. Phillips [37–40].

Definition ([2]). If a linear combination of integrated random processes of the order d is an integrated process of the order $d-1$, then they are referred to as the cointegrated processes, and the vector of linear combination coefficients, as the cointegrating vector.

The basic Granger's result lies in that in the case of cointegration $I(1)$ of the series y_t and x_t there exists a vector representation

$$x_t = c_1 + \sum_{j=1}^{k}(a_{1j}x_{t-j} + b_{1j}y_{t-j}) + \sum_{k=0}^{q} d_k\varepsilon_{1,t-k},$$
$$y_t = c_2 + \sum_{j=1}^{k}(a_{2j}x_{t-j} + b_{2j}y_{t-j}) + \sum_{k=0}^{q} d_k\varepsilon_{2,t-k}$$

and a representation in the form of the error correction model (ECM):

$$\begin{aligned}\Delta x_t &= \mu_1 + \alpha_1 z_{t-1} + \sum_{j=1}^{\infty}(y_{1j}\Delta x_{t-j} + \delta_{1j}\Delta y_{t-j}) + \sum_{k=0}^{\infty}\theta_k\varepsilon_{1,t-k},\\ \Delta y_t &= \mu_2 + \alpha_2 z_{t-1} + \sum_{j=1}^{\infty}(\gamma_{2j}\Delta x_{t-j} + \delta_{2j}\Delta y_{t-j}) + \sum_{k=0}^{q}\theta_k\varepsilon_{2,t-k},\end{aligned} \tag{10}$$

where $z_t = y_t - \beta x_t - E(y_t - \beta x_t) \sim I(0)$ is a stationary series with the zero expectation and $a_1^2 + a_2^2 > 0$. If the vector series $(x_t, y_t)^{\mathrm{T}} \sim I(1)$ is generated by the ECM, then the series x_t and y_t are cointegrated because in this case all components of the ECM, except for the series z_{t-1}, are stationary. Consequently, the series z_{t-1} is stationary as well. Therefore, if x_t, $y_t \sim I(1)$ and cointegrated, then there exists a long-term relation represented by $y_t = \alpha + \beta x_t$ and short-term dynamics represented by model (10), both models being coordinated with each other.

2.2.3. Vector autoregression models. To construct models of multidimensional time series, C.A. Sims [41] developed a construction named the vector autoregression (VAR) models. Three different types of the VAR-models—reduced, recursive and structural—are distinguished. They are all of them dynamic linear models relating the current and past values of the n-dimensional vector Y_t of the time series. No economic-theoretical constraints, except for the choice of variables, are used to construct the reduced and recursive VAR-models; construction of the structural VAR-model is based on the constraints established by the macroeconomic theory. The reduced VAR-model is a system of n equations representable in the matrix form as

$$Y_t = A_0 + A_1Y_{t-1} + \ldots + A_pY_{t-p} + E_t, \tag{11}$$

where A_0 is the constant vector, $A_1, \ldots, A_p$ are the coefficient matrices, and E_t is the vector of serially noncorrelated errors which are assumed to have mean zero and the covariance matrix Σ_E.

To estimate the parameters of the reduced VAR-model, the least squares method (LSM) is applied separately to each equation. The error covariance matrix Σ_E is estimated consistently by the sample covariance matrix of the residues obtained by LSM. The model order is determined using the Akaike information criterion (AIC) or the Schwatz information criterion (BIC) [42]. Watson considered four representations of model (11) with integrated variables [43].

(1) Vector model of error correction

$$\Delta x_t = \Pi x_{t-1} + \sum_{i=1}^{p-1} \Phi_i \Delta x_{t-i} + \varepsilon_t,$$

where $\Pi = I - \sum_{i=1}^{p} \Pi_i = \Pi(1)$ and $\Phi_i = -\sum_{j=i+1}^{p} \Pi_j$, $i = 1, \ldots, p-1$. Equations of this kind were introduced in [44].

(2) Representation in the form of moving average of the first-order differences

$$\Delta x_t = C(L)\varepsilon_t.$$

Equivalence of the vector model of error correction and the representation as moving average which was shown in [33] underlies the Granger theorem of representation [34].

(3) Representation in the form of common trends:

$$x_t = C(1)\xi_t + C^*(L)\varepsilon_t + x_0, \tag{12}$$

where $\xi_t = \sum_{s=1}^{t} \varepsilon_s$ and $C^*(L) = (1-L)^{-1}[C(L) - C(1)] = \sum_{i=0}^{\infty} C_i^* L^i$, where $C_i^* = -\sum_{j=i+1}^{\infty} C_j$ and $\varepsilon_i = 0$ for $i \le 0$. Equation (12) is a multidimensional Beveridge–Nelson decomposition [45] into the "long-term," $C(1)\xi_t + x_0$, and the "short-term," $C^*(L)\varepsilon_t$, components representable as

$$x_t = A\tau_t + C^*(L)\varepsilon_t + x_0,$$

where A is the full-rank $n \times k$ matrix of $(\tau_t - k)$ "long-term" components.

(4) Triangular representation (in the case of cointegration):

$$\Delta x_{1,t} = \mu_{1,t},$$
$$x_{2,t} - \beta x_{1,t} = \mu_{2,t},$$

where $x_t = (x_{1,t}, x_{2,t})$, $x_{1,t}$ of size $k \times 1$, $x_{2,t}$ of size $r \times 1$, $\mu_t = (\mu_{1t}, \mu_{2t})^{\mathrm{T}} = D(L)\varepsilon_t$.

2.2.4. Procedures for verifying cointegration. Procedures for verification of cointegration were developed by Engle and Granger as well as by Johansen.

Engle–Granger Procedure [34]. **Step 1.** Determination of the order of process component integration using the Dickey–Fuller criterion.

Step 2. If the processes are established to be integrated and of the same order, then regression is calculated by the ordinary LS method.

Step 3. The order of integration of the regression residues is checked.

To verify the residue integration order, the Dickey–Fuller can be used. However, other critical values of the test tables are used because the tested process was obtained using the LS method. The tables can be found in [20].

Johansen Procedure [35, 36]. Let us consider n cointegrated vectors $Y_t = \{y_t^1, \ldots, y_t^n\}$ described by model (11) in the form of the error correction model

$$\Delta Y_t = -\Pi Y_{t-1} + B_1 \Delta Y_{t-1} + \ldots + B_{p-1}\Delta Y_{t-p+1} + E_t,$$

where $\Pi = I - A_1 - \ldots - A_p$ and the matrices $B_1, \ldots, B_{p-1}$ are expressed in terms of the matrices $A_1, \ldots, A_p$ as $B_j = \sum_{i=j+1}^{p} A_i$. The number of cointegrated vectors is equal to the matrix rank Π. An exact proof of this fact can be found in [36].

Step 1. The VAR(p) model is estimated for the k-dimensional time series Y_t:

$$Y_t = \overline{\alpha} + t\overline{\beta} + A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + E_t.$$

Step 2. The matrix $\Pi = I - A_1 - \ldots - A_p$ is estimated in the representation

$$\Delta Y_t = \overline{\alpha} + t\overline{\beta} - \Pi Y_{t-1} + B_1 \Delta Y_{t-1} + \ldots + B_{p-1} \Delta Y_{t-p+1} + E_t.$$

Step 3. The zero hypothesis $H_0 : \operatorname{rank} \Pi < k$ is verified against the alternative hypothesis $H_1 : \operatorname{rank} \Pi = k$ or $H_0 : \operatorname{rank} \Pi = k + 1$.

3. ANALYSIS OF THE NONSTATIONARY PROCESSES WITH STRUCTURAL CHANGES IN PROPERTIES (STRUCTURAL DISCONTINUITIES)

3.1. Problems Caused by the Structural Discontinuities in Processes

The majority of the existing methods of verification of process type and cointegration assume that over the observation interval the process is constant. However, in the physical systems the process regularities may vary with time under the environmental actions. The changes in process properties will be referred to as the structural discontinuities. This term is widely used by the researchers who define it in different ways. We introduce the following definition.

Definition. The structural discontinuities are changes in the process properties reflected in the changes of the coefficients of the parametric process model and/or in the parameters of the distribution of the random component. The structural discontinuities can modify the type of process, the mean value of process or its differences, the process trend inclination, and the coefficients of its parametric model.

In the presence of structural discontinuities in the process,

—the parametric system model is modified; if the changes in the model parameters are disregarded, then the parameter estimates become insignificant, and the predictions lose precision;

—verification of the criteria for determination of process type become more complicated;

—verification of cointegration becomes more complicated.

In the presence of structural discontinuities, the standard procedures distort the conclusions about existence of the unit root. Perron [46] demonstrated that for the extended Dickey–Fuller criterion the unit-root zero hypothesis needs not be rejected if in the process a discontinuity occurs in the linear trend or the mean value is shifted. It was shown [47, 48] that, depending on the form of structural discontinuity, the zero hypothesis can be accepted even if it is not true. The discontinuities in the cointegration equation were shown to reduce the power of the cointegration test [49, 50].

Owing to the aforementioned reasons, much attention is paid by the analysts to the algorithms to determine the instants of occurrence of the structural discontinuities and analyze the processes in real conditions. Under the action of diverse, and often opposite, random factors, the process structural discontinuities can be masked by these random factors and, therefore, not necessarily be detectable from the observed variables.

Detection of structural discontinuities can be based either on *a posteriori* procedures detecting the structural changes after completion of the observations or on the real-time procedures detecting changes with the rate of current observations. Choice of procedure is defined by the requirements of application.

Specification of the instants of origination of the structural discontinuities with subsequent decomposition of the process into subprocesses where the properties do not change is a natural way of analyzing such processes. The existing procedures for change detection are intended only for the processes that are stationary within the intervals between the structural discontinuities. Therefore, prior to using them for analysis of nonstationary processes, the original process must be first stationarized. Another approach to analysis of the processes with changes in the properties is represented by the algorithms for analysis of the nonstationary processes (verification of the process type and cointegration) with regard for changes in the properties. Let us first consider the *a posteriori* procedures detecting structural discontinuities.

3.2. Methods of A Posteriori *Detection of the Instants of Process Structural Discontinuities*

The problem of specifying the no-change intervals in a process can be posed as that of verifying the no-change hypothesis H_0 against the hypothesis H_1 changes. If H_0 is satisfied, then the observed process has no changes in the properties; if H_1 is satisfied, then the process has at least two intervals and is described by an individual model over each interval.

If the instant of possible change in the process state is known, then the hypothesis that there are no changes in the properties at this instant is verified against the hypothesis of changes for the fixed-length (N) time series under study. If no assumptions about the process model exist, then verification of changes is carried out using the tests for comparison of two samples. The two-sample Kolmogorov–Smirnov test or other nonparametric tests [51] can be used, for instance.

The Kolmogorov–Smirnov test makes use of the statistic

$$KS = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \max_x |F_1(x) - F_2(x)|,$$

where N_1 and N_2 are the numbers of points in the time series, respectively, before and after the postulated property change and $F_1(x)$ and $F_2(x)$ are the empirical functions of distribution of the time series segments, respectively, before and after the change.

If the distribution F_0 before the change instant is known, then the tests for concordance such as the χ-square and Kolmogorov–Smirnov tests [51, 52] are used to verify that the distribution before and after the postulated change instant is F_0.

Different strategies for verification of discontinuities exist for the case where the series is described by the regression or autoregression model:

(1) test for a change in the regression coefficients using the F-statistic [53–60];

(2) determination of the boundaries of the uniformity interval where the process model can be regarded as constant [61–65];

(3) tests for generalized changes by residue analysis (the so-called CUSUM and MOSUM tests) [66–70].

3.2.1. Test for change of the regression coefficients using the F-statistic. *Test for discontinuity at the known time instant* τ. The original model is as follows:

$$Y_t = X_t^{\mathrm{T}} \beta_t + \mu_t. \tag{13}$$

The zero hypothesis is as follows:

$$\beta_t = \beta \quad \text{for all} \quad t.$$

The alternative hypothesis is as follows:

$$\beta_t = \beta \quad t \le k \quad \text{and} \quad \beta_t = \beta + \gamma, \quad t > k, \quad \text{where } k \text{ is the date of discontinuity.}$$

After getting the estimates of coefficients, the Chow test [53, 54] for equality of the regression coefficients in two samples is used. In the Wald form, this test is as follows:

$$F_T\left(\frac{k}{T}\right) = \frac{(\overline{u}'\overline{u} - u_1'u_1 - u_2'u_2)/k}{(u_1'u_1 + u_2'u_2)/(T-2k)},$$

where k is the instant of postulated discontinuities and $\overline{u}$, u_1, and u_2 are the sums of square residues of model (13) as estimated from the full sample, k first and $n-k$ last observations, respectively. This test statistic has an F-distribution with k and $n-2k$ degrees of freedom.

Test for discontinuity at the unknown time instant. To take into account the discontinuities in the constructed models at the unknown time instants, the dummy or "empty" variables are used which assume 0 before and 1 after the change. Let it be known that the discontinuity lies within the interval between τ_0 and τ_1.

After estimating the coefficients, the Chow test for equality of the regression coefficients in two samples is applied: the zero hypothesis of no-change is verified by the F-statistic for the hypothesis that the coefficients of all additional variables are zero. For detection of changes on average, the dummy variables are introduced directly in the equations, and for detection of changes in the trend, the dummy variables are added to the regression coefficients.

If the process has one possible instant of state change and this instant is unknown, the test is carried out for each point of the time series. The Chow test is constructed at each instant of the postulated discontinuities and the maximum F-statistic is taken. The distribution of this test is called the Quandt Likelihood Ratio [55, 56]

$$QLR = \mathrm{MAX}[F(\tau_0), F(\tau_0+1), \ldots, F(\tau_1)]$$

which depends on the number of tested constraints and on the positions of the end points τ_0 and τ_1 in the sample T. The limit distribution of the QLR-statistic is as follows [8, 57, 58]:

$$QLR = \sup_{\lambda\in[\lambda_0,\lambda_1]}\left\{\frac{B_k^\mu(\lambda)'B_k^\mu(\lambda)}{\lambda(\lambda-1)}\right\},$$

where

$$\lambda_i = \lim_{T\to\infty}\frac{r_i}{T}, \quad i = 0, 1.$$

The critical values of the functional QLR as calculated by means of statistical modeling can be found in [57]. A method of calculating the critical values was developed in [59] together with a function of their dependence on the number of model parameters and availability of trend and free term.

Estimation of the instant of discontinuity by maximizing the Chow criterion provides an acceptable accuracy only in the case of linear regression with homoskedactic covariance matrix. LSM is the most suitable method for estimation of the parameters, including the discontinuity date, in the regression models. For each possible discontinuity point, the model parameters are estimated by the least-squares method for each of the subsamples, and the sum of square errors is calculated and accumulated. A point where the sum of square errors is minimal over the entire sample is used as the estimate of the discontinuity instant. The algorithms to estimate the discontinuity instant by the least-squares method were suggested and examined in [50–63]. The works [61, 62] established the asymptotic distribution of the discontinuity estimator and demonstrated how to construct the confidence interval for the date of discontinuity.

Detection of more than one discontinuity. The majority of the classical methods of detection and modeling of structural changes assume that there exists only one discontinuity. Methods for

determination of the number and instant of occurrence of multiple discontinuities were developed in [63, 65, 66]. The idea of the algorithm of [65] is as follows. At the first step, existence of a single discontinuity is verified. If the zero hypothesis of no discontinuity is rejected, then the original sample is divided in two parts and the test is applied again to each of the samples. These steps are repeated until the test ceases to determine existence of discontinuities. Processing of the date of structural discontinuity as an unknown parameter provides both the date of discontinuity and its confidence interval. Successive procedures for estimation of multiple discontinuities were developed in [63, 66]. Their key concept lies in that in the presence of multiple discontinuities the sum of square errors may have a local minimum near each discontinuity. Consequently, the global minimum can be used as the discontinuity date, the rest of the local minima can be regarded as the candidates for estimation of the discontinuity date. Then the sample is decomposed in two parts at the point of estimation of the discontinuity date, and analysis is continued for each of the samples.

Specification of the uniformity intervals in context of the given linear model. Another approach to specifying the process states lies in determining the boundaries of the interval where the model may be regarded as invariable.

Let the time series Y_t follow the model

$$Y_t = X_t^{\mathrm{T}}\theta + \sigma\varepsilon_t, \tag{14}$$

where $t = 1, \ldots, N$, $X_t = (x_1, \ldots, x_{pt})^{\mathrm{T}}$, and $\theta = (\theta_1, \ldots, \theta_p)^{\mathrm{T}}$, σ is a constant defining the process variance.

Needed is to estimate the coefficients θ in (14), provided that their values are piecewise-constant functions over the interval $[1, \ldots, N]$. We follow [67] is calling the intervals of constant θ the uniformity intervals. The problem of determining the time interval where the model is adequate was solved in [68–70].

The basic concept of the adaptive estimation algorithm is as follows. Obviously, if θ is constant over the interval I, it is also constant over any interval $J \subset I$. The algorithm relies on specifying the greatest interval and estimating θ over it by the least-squares method.

The procedure of adaptive estimation includes a set of tests verifying whether the difference between $\widehat{\theta}_I$ and $\widehat{\theta}_J$ is appreciable, where $\widehat{\theta}_I$ and $\widehat{\theta}_J$ are the estimates of θ over the intervals I and J, respectively. If I is a uniform interval, then the following condition is met:

$$\left|\widehat{\theta}_{i,I} - \widehat{\theta}_{i,J}\right| \le \mu\delta\sqrt{w_{ii,I}} + \lambda\delta\sqrt{w_{ii,J}},$$

where $\widehat{\theta}_{i,I}$ and $\widehat{\theta}_{i,J}$ are the estimates of θ_i obtained, respectively, over the intervals I and J; $w_{ii,I}$ and $w_{ii,J}$ are the diagonal elements of the matrix $W = \left(\sum\limits_{t=1}^{N} X_t X_t^{\mathrm{T}}\right)^{-1}$ for $x \in I$ and $x \in J$, respectively, λ and μ are positive constants, and σ is the variance.

Tests for generalized deviations (CUSUM, MOSUM). In the tests for generalized changes, the parametric model is constructed by the ordinary least-squares method or the maximum likelihood method, and a statistic is generated which verifies availability of significant deviations in the parameter residues or estimates [71–73] or in the estimates constructed in the moving window from the estimates based on the full sample [74–76].

Consideration is given to the regression model

$$Y_t = X_t^{\mathrm{T}}\beta + \varepsilon_t,$$

where ε_t is an independent normally distributed variable with the zero mean. The tests for structural changes are concerned with verification of the no-change zero hypothesis

$$H_0 : H_0 : \beta_i = \beta_0, \quad i = 1, \ldots, n,$$

against the alternative hypothesis that the coefficient vector changes at some time instant t.

CUSUM tests. The cumulative sum of recursive residues is discussed in [72]. The coefficients β are estimated successively from the increasing volume of the sample

$$\widetilde{\beta}\left(\frac{r}{T}\right) = \left(\sum_{t=2}^{r} X_{t-1}X'_{t-1}\right)^{-1}\left(\sum_{t=2}^{r} X_{t-1}Y_t\right), \tag{15}$$

$$w_t = \frac{Y_t - \widetilde{\beta}\left(\frac{t-1}{N}\right)X_{t-1}}{f_t},$$

where

$$f_t = \left\{1 + X'_{t-1}\left(\sum_{s=2}^{t-1} X_{s-1}X^1_{s-1}\right)^{-1} X_{t-1}\right\}^{\frac{1}{2}}.$$

The algorithm is as follows:

$$CUSUM(\lambda) = \frac{1}{\widetilde{\sigma}\sqrt{T}}\sum_{i=k+1}^{[N\lambda]} w_i, \tag{16}$$

where

$$\widetilde{\sigma} = \left\{\frac{\sum\limits_{t=k+1}^{N} w_t - \overline{w}}{N}\right\}^{\frac{1}{2}}.$$

For the zero hypothesis, the limit process is a standard Brownian motion [77, 78]. If there exists a single point t_0 of structural change, then the recursive residues will have the zero mean only until t_0. The critical values for the statistic

$$\sup_{\lambda} |CUSUM(\lambda)/(1+2\lambda)|$$

were obtained in [72].

W. Ploberger and W. Kramer [73] proposed to study the cumulative sums of the LSM general residues:

$$W^0_n(t) = \max_{k\in[1,T]} \frac{1}{\widehat{\sigma}\sqrt{T}}\sum_{i=1}^{k}\widehat{\varepsilon}_i.$$

For $W^0_n(t)$, the limit process is a standard Brownian bridge $W^0(t) = W(t) - tW(1)$ having for one structural discontinuity the peak at the point t_0.

MOSUM tests. Another method to detect the structural changes lies in analysis of the moving sum of the same residues, that is, the sums of fixed numbers of residues h taken into consideration until the current time instant t. As was shown in [74], the sum of residues in the moving window (empirical MOSUM process) tends to the Brownian bridge.

To detect changes in the time series variance, C. Inclan and G.C. Tiao [79] proposed to use the statistic

$$\Delta = \sup_k \left|\sqrt{\frac{N}{2D_k}}\right|,$$

where

$$D_k = \frac{C_k}{C_T} - \frac{k}{N}$$

and $C_k = \sum_{i=1}^{k} \varepsilon_i^2$, $k = 1, 2, \ldots, N$, is the cumulative sum of the square ε_t. If ε_t are normally distributed random variables with zero mean $\varepsilon_t \sim N(0, \sigma^2)$, then the asymptotic distribution of the test obeys

$$IT \Rightarrow \sup_r |W^*(r)|,$$

where $W^*(r) = W(r) - rW(1)$ is the Brownian bridge and $W(r)$ is the standard Brownian motion. The main disadvantage of the IT test lies in the dependence of its asymptotic distribution on the assumption of normality, independence, and similarity of the distributions of the random variables ε_t.

Main types of a posteriori *algorithms for determination of the structural discontinuities.* The present review discusses the main types of the methods for detection of structural changes from the historical data. Numerous methods that can be classified with one or another of the aforementioned types are developed in this area. The proposed algorithms differ in the methods of their realization and the original models. For example, [80, 81] consider the CUSUM test for the case of dynamic model with serially correlated errors:

$$y_t = \alpha y_t + X_t'\beta + \mu_t,$$

where

$$\mu_t = \rho\mu_{t-1} + \varepsilon_t, \quad |\rho| < 1.$$

The main types of the methods for determination of the structural discontinuities are tabulated below.

3.3. Methods of Current Detection of the Instants of Structural Discontinuities in the Process

The above methods of analysis in the presence of property changes and estimation of the instants of origination of these changes are retrospective because they verify the accomplished events from the historical data. To detect changes in process properties from the current observations, methods of successive analysis are used. They are popular in engineering problems, yet, except for [82, 83], their application to econometrics is unknown. Two tests for monitoring the possible structural changes were proposed in [82]. In these tests, after getting a current observation, the estimates are calculated over all data and compared with the similar estimates obtained without this observation. The no-change hypothesis is rejected if their difference is too great.

The class of tests for monitoring of changes which includes the algorithms of [82] as a special case was considered in [83]. The distributions of these tests were shown to be defined by the increments of the generalized Brownian bridge, and the corresponding boundary functions for the proposed tests were shown to grow with the rate $\sqrt{\log t}$.

However, the algorithms of current detection [82, 83] are anything but optimal. To detect changes in the properties of stationary processes, the successive analysis theory and algorithm that are optimal under certain conditions were developed [84]. The algorithms rely on the procedure of successive verification of the hypotheses. The hypothesis H_0 of "the process distribution density

Methods for determining structural discontinuities

Methods	Input information	Brief description	Result	Author
Kolmogorov–Smirnov test	Data	Empirical distributions are constructed on the subsamples and the maximum of difference is sought	Point of change (one discontinuity)	Kolmogorov–Smirnov
Tests for concordance: Kolmogorov–Smirnov, χ square	Data, original distribution	Verification that the statistics of different samples have no significant differences	Probability of meeting the hypothesis	
Chow test	Data, discontinuity instant, regression variables	Two models are constructed, F-statistic is calculated and analyzed	Test for discontinuity at the given time instant	Chow
	Data, regression variables	Two models are constructed for all possible instants of discontinuities, F-statistics are calculated and the maximum one is taken	Test for discontinuity and determination of the discontinuity instant	
Determination of the model parameters and the discontinuity date by the generalized LSM	Data, regression variables	Models are constructed at each possible point of discontinuity by the generalized LSM method and the residues are calculated, the discontinuity point minimizing the sum of square residues	Test for discontinuities and determination of the discontinuity instant and the model	Bai, Andrews
Specification of uniform intervals	Interval of model constancy (rated)	Model is constructed over the interval, model constancy is tested over the extended interval	Instant of model property change	Liptser, Spokoiny
Test for generalized fluctuations (oscillations) CUSUM	Residues of the model constructed over the analyzed interval	Model is constructed, residues are calculated, exceedance is tested	Interval where the deviations of model from the mean value are observed	Brown, Durbin, Evans, Ploberger, Kramer
CUSUM in the moving window	Residues of the model constructed over the analyzed interval	Model is constructed, residues are calculated, exceedance is tested	Interval over which model variance deviations are observed	Inclan, Tiao
MOSUM	Residues of the model constructed in the moving window	Model is constructed in the moving window, window is shifted, residues are calculated, exceedance is tested	Interval over which model deviations from the mean value are observed	
MOSUM in moving window		Model is constructed in the moving window, window is shifted, residues are calculated, exceedance is tested	Interval over which model variance deviations are observed	Chu, Hornik, Kuan

is P_1" is verified against the alternative hypothesis H_1 of "the process distribution density is P_2." A vast majority of the successive detection algorithms were developed for the cases where the distribution functions P before and after detection differ only in their parameters:

$$P_1 = p(\theta_1) = p_{\theta_1}, \quad P_2 = p(\theta_2) = p_{\theta_2}.$$

To explain the formulation of the current detection problem, we define the main notions [85].

Definition. The random variable τ is called the stop instant or the Markovian instant for the process y_t if it assumes only integer values and the event $\{\tau = n\}$ is defined by the sequence $y_1, y_2, \ldots, y_n$.

Definition. The successive criterion for verification of the hypothesis H_0 against the alternative hypothesis H_1 is defined as the pair (g, τ), where τ is the stop instant and $g(y_1^\tau)$ is the decision function.

Definition. The criterion (g, τ) is called the successive probability ratio criterion (SPRC) for verification of the simple hypothesis H_0: $P = P_1$ against the alternative hypothesis H_1: $P = P_2$ if, as the result of applying it to the sequence y_1, $y_2, \ldots, y_n, \ldots$, executed at each step is the test $a < s_k^t < b$, where

$$s_k^t = \sum_{i=k}^{t} \ln \frac{f_{\theta_2}(s_i)}{f_{\theta_1}(s_i)} \tag{17}$$

and a and b are thresholds such that $-\infty < a < b < \infty$, and either of the following decisions is made:

—the hypothesis H_0 is accepted if $s_k^t \le a$,

—the hypothesis H_1 is accepted if $s_k^t \ge b$,

—observations are continued if $a < s_k^t < b$.

Quality of the statistical criterion is estimated by the probabilities of errors of the first and second order:

$$\alpha_0 = P\left(g(y_1^N) \ne H_0 \mid H_0\right), \quad \alpha_1 = P\left(g(y_1^N) \ne H_1 \mid H_1\right),$$

where y_1^N is the sequence of observations $y_1, \ldots, y_N$.

Let $Y_1, Y_2, \ldots, Y_{t_a}, Y_{t_a+1}, \ldots$ be a random sequence of independent variables, where $Y_1, Y_2, \ldots, Y_{t_a}$ are distributed identically with the distribution function P_{θ_1}, $Y_{t_{a+1}}, Y_{t_{a+2}}, \ldots$ are distributed identically with the distribution function P_{θ_2}, and θ_1 and θ_2 are the known distribution parameters before and after the property changes. Each observation is tested for changes by the SPRC. The following characteristics are used to estimate the quality of the successive criterion:

—mean time between the false alarms $E_1(\tau \mid \tau < t_a)$,

—mean time of detection delay $E_2(\tau - t_a \mid \tau \ge t_a)$.

The value of these characteristics depends on the chosen values of the thresholds a and b, parameters θ_1, θ_2 and the given confidence level defined by the probabilities α_0 and α_1.

There are two alternative formulations of the optimal compromise between the rate of detection and the level of false alarms:

(1) Lorden's formulation which requires to minimize the worst-case delay

$$\tau = \sup_{t_a \ge 1} \operatorname{essSup} E_{\theta_2}(t - t_a + 1 \mid y_1^{t_a},\ t > t_a), \tag{18}$$

under a constraint on the minimal time between false alarms

$$T = E_{\theta_1}(t \mid t < t_a) < \gamma, \tag{19}$$

where τ is the instant of detection of changes and $E_{\theta_i}(^*)$ is the expectation of $(^*)$ under the distribution $f_{\theta_i}(y_t)$, $i = 1, 2$.

(2) Shiryaev's Bayesian formulation according to which the instant of change has a geometrical distribution and the mean expected delay

$$\tau = \sup_{t_a \geq 1} E_2(\tau - t_a \mid \tau \geq t_a) \tag{20}$$

must be minimized under a constraint on the minimum time between false alarms.

The Shiryaev–Roberts algorithm has the form

$$\tau = \min(t > 1 : R_t \geq h), \tag{21}$$

where $R_t = \sum\limits_{i=1}^{t} \exp(\Delta_k^t)$, the statistic R_n satisfies the recurrent representation

$$R_t = (1 + R_{t-1})\frac{f_{\theta_2}(y_t)}{f_{\theta_1}(y_t)}, \quad R_0 = 0. \tag{22}$$

Algorithm (21), (22) minimizes the mean delay in detection (20) under constraint on the minimum time between the false alarms (19).

For the case where the parameters θ_1 and θ_2 are known precisely before and after the change in properties and the time of change t_a is unknown, E.S. Page [86] developed a SPRC-based algorithm of cumulative sums that was named CUSUM. Its basic idea lies in using $b = 0$ as the lower threshold and resuming SPRC each time as s_k^t reaches the lower threshold until the upper threshold a is exceeded. The instant of changes in properties is defined as the solution of the optimization problem:

$$\tau = \inf\left\{\tau \geq 1 : \ g = \max_{1 \leq k \leq t} s_k^r \geq a\right\}, \tag{23}$$

where s_k^τ obeys (17).

The decision function g admits the recurrent representation

$$g^t = \max\left(0, g^{t-1} + \ln\frac{f_{\theta_1}(y_t)}{f_{\theta_2}(y_t)}\right), \quad g^0 = 0.$$

The algorithm is representable as a set of SPRC's activated each at time t with the upper threshold a.

For the case of precisely known parameters before and after detection, algorithm (17), (23) was shown [87] to be optimal in the sense of the criterion for minimization of the mean "worst-case" detection delay (18) for the mean time before the false alarm T satisfying condition (19).

For all real θ_2, the mean detection delay time follows

$$E_{\theta_1}(\tau(\gamma)) \sim \frac{\ln(\gamma^{-1})}{K(\theta_2, \theta_1)}, \tag{24}$$

where $K(\theta_2, \theta_1)$ is the information measure called the Kullback–Leibler information which is defined as follows:

$$K(\theta_i, \theta_j) = E_{\theta_i}\left[\ln\frac{f_{\theta_2}(y_t)}{f_{\theta_1}(y_t)}\right], \quad i, j = 1, 2, \quad i \neq j. \tag{25}$$

CUSUM optimality in the sense of (23) was proved for the nonasymptotic formulation in [88, 89]. Algorithms of current detection of changes in nonstationary processes whose decision rules are based on the decision rules (17), (23) optimizing the mean time of delay in detection (18) under constrained minimum time between alarms (19) were considered in [90–92].

3.4. Methods of Verifying Process Type and Cointegration in the Presence of Structural Discontinuities

The problem of analysis of a process with structural discontinuities was solved above by specifying in the process the intervals where no changes in properties occur. After their determination, the problem comes to the ordinary problem of analysis without changes in properties. Yet almost all the above methods of specifying the no-change intervals, except for [90–92], started from the assumption that within the interval the process is stationary. One has to perform series "stationarization" before applying them to the nonstationary processes. Choice of the "stationarization" method depends on the original process that can be unknown at the initial stages of econometric study. In this section we consider another approach to analysis of processes with varying properties: determination of the process type and verification of cointegration in the presence of structural discontinuities.

It was assumed in [46, 93] that at a certain time instant t_b there is a discontinuity and the zero hypothesis presupposed that the process obeys one of the following models.

(1) Model with discontinuity on the average

$$y_t = \mu + \delta_1 DVI_t + y_{t-1} + \varepsilon_t,$$

where

$$DVI_t = \begin{cases} 1 & \text{if} \quad t = t_b + 1 \\ 0 & \text{if} \quad t \neq t_b + 1. \end{cases}$$

(2) Model with drift change

$$y_t = \mu + \delta_2 DV_t + y_{t-1} + \varepsilon_t,$$

where

$$DV_t = \begin{cases} 0 & \text{if} \quad t < t_b + 1 \\ 1 & \text{if} \quad t \geq t_b + 1. \end{cases}$$

(3) Model with discontinuity on the average with simultaneous change in drift

$$y_t = \mu + \delta_1 DVI_t + \delta_2 DV_t + y_{t-1} + \varepsilon_t.$$

The following hypotheses are alternatives to each of the above models.

(1a) Model of trend-stationary series with level change

$$y_t = \mu + \delta_2 DV_t + \beta t + \varepsilon_t.$$

(2a) Model of trend-stationary series with change in trend level

$$y_t = \mu + \delta_3 DVT_t^* + \beta t + \varepsilon_t,$$

where

$$DVT_t^* = \begin{cases} 0 & \text{if} \quad t < t_b + 1 \\ t & \text{if} \quad t \geq t_b + 1. \end{cases}$$

(3a) Model of trend-stationary series with change in trend inclination and drift

$$y_t = \mu + \delta_2 DV_t + \delta_3 DVT_t + \beta t + \varepsilon_t,$$

where

$$DVT_t = \begin{cases} 0 & \text{if} \quad t < t_b + 1 \\ t - t_b & \text{if} \quad t \geq t_b + 1. \end{cases}$$

To verify the zero hypothesis, corresponding regression models are constructed, and then the Dickey–Fuller test is applied to the regression residues. Perron proved that in the presence of structural changes the magnitudes of the critical values of the Dickey–Fuller criterion are greater and depend on the position of jump in the series. He calculated these critical values.

The asymptotic critical values of the t-statistic of the Perron criterion depend on the type of structural changes, the parameter $\lambda = \dfrac{t_b}{T}$, and on the postulated model—either that with an additive outlier where the structural change occurs unexpectedly or that with an innovative outlier where the structural change occurs gradually.

Zivot and Andrews proposed, instead of the conditional Perron criterion, a criterion determining the discontinuity date by calculating the t-statistic for verification of the unit-root hypothesis for every possible instant of discontinuity, the instant of the minimum $t_{\min}$ of this statistic being taken as that of discontinuity [94]. In extension of the results of Zivot and Andrews, Perron [95] studied the critical values of the statistic $t_{\min}$ vs. the choice of the number of retarded differences and developed an estimation procedure integrated in the RATS package.

The problems arising at determination of the series type are apparent from the example of statistical analysis of the 13 basic USA macroeconomic series based on the annual data and the quarterly series of GNP for the post-war period 1948–1987. For these series, C.R. Nelson and C.I. Plosser [96] verified the unit-root hypothesis and used the critical values of the Dickey–Fuller criterion [15]. At that they established that the unit-root hypothesis is not rejected for 13 of 14 series, the only exception being the series of employment rate logarithms. Perron [46] analyzed the same series considering a series with linear trend whose inclination changes at a certain time instant as an alternative to the DS-series and obtained results opposite to those of Nelson and Plosser: the unit-root hypothesis was rejected for 11 of 14 series. To determine the type of 14 series of macroeconomic indices, Zivot and Andrews [94] used an algorithm which determines the instant of change in trend inclination by selecting every possible instants of changes and calculating for each variant the t-statistic for verification of the unit-root hypothesis. The date for which the value of statistic is minimal is accepted as the estimated date. With this method of analysis, the unit-root hypothesis is not rejected for 11 series.

The problem of testing cointegration in the presence of structural changes in processes was considered in [50]. The power of the traditional test for cointegration was shown to drop dramatically in the presence of discontinuities.

Tests for cointegration in the presence of structural discontinuities were considered in [97], that is, consideration was given to a new type of cointegration, that in the presence of discontinuities. The zero hypothesis of no cointegration (existence of the unit root) was verified against the alternative hypothesis of existence of cointegration in the presence of changes either in the process mean or in the cointegration vector, the instant of change being assumed to be unknown.

Consideration was given to the model

$$y_{1t} = \mu + \alpha y_{2t} + e_t,$$

where y_{1t} and y_{2t} are, respectively, one-dimensional and m-dimensional vector processes of the type $I(1)$ and e_t is a process of the type $I(0)$.

The zero hypothesis assumes lack of cointegration.

Consideration is given to the following hypotheses regarded as alternatives:

$$y_{1t} = \mu + \mu_2\varphi_{t\tau} + \alpha y_{2t} + e_t \tag{26}$$

—change in the level;

$$y_{1t} = \mu + \mu_2\varphi_{t\tau} + \beta t + \alpha y_{2t} + e_t \tag{27}$$

—change in the level in the presence of trend;

$$y_{1t} = \mu + \mu_2\varphi_{t\tau} + \alpha_1 y_{2t} + \alpha_2 y_{2t}\varphi_{t\tau} + e_t \tag{28}$$

—change in the level and cointegration vector, where

$$\varphi_{t\tau} = \begin{cases} 0 & \text{if} \quad t \le [n\tau] \\ 1 & \text{if} \quad t > [n\tau] \end{cases}$$

is a dummy variable defining the instant of change occurrence and $[^*]$ is the integer part of *.

Depending on which of the alternative hypotheses is verified, the residues of models (26), (27) or (28) are estimated for every possible instant τ of change and are used to calculate the standard statistics for cointegration analysis without changes in the models. The value of τ for which the chosen statistic has the minimum value is taken as the possible instant of change. If the calculated statistic is smaller than its tabular value, the hypothesis is rejected.

It deserves noting that this type of hypothesis does not bear out occurrence or nonoccurrence of a structural change in the process because the alternative hypothesis includes as a special case the standard model of cointegration without structural change. Since the structural changes in one or more cointegrated processes violate the cointegration relations and may cause changes in other cointegration components, timely detection of these changes is an important problem.

4. CONCLUSIONS

The present review analyzed the state-of-the-art in the studies of stationary and nonstationary time series in the presence of changes in their properties, the structural discontinuities, defined the structural discontinuities, and specified the main stages of analyzing the processes with property changes. An analytical review of the methods for detection of process structural discontinuities, determination of the process type and analysis cointegration in the lack and presence of process structural discontinuities, and current detection of structural discontinuities was carried out.

As the result of this review, one can conclude that for successful prediction and preventive monitoring the process under study must be analyzed, the prehistory analysis including the following:

—determination of the instants (periods) of changes in the process properties or the process uniformity intervals;

—determination of the process type (stationary, nonstationary);

—construction of the process models (model bank) within the uniformity intervals;

—adjustment of the algorithms of current detection and verification of process uniformity.

Analysis of the current process state includes:

—application of the procedures of current detection of process changes to all new observations;

—analysis of the current observations after the last property change and construction of the process model or choice of a model from the bank.

REFERENCES

1. Kendall, M. and Stuart, A., *The Advanced Theory of Statistics, vol. 3: Design and Analysis, and Time Series*, London: Charles Griffin, 1973, 2nd ed. Translated under the title *Mnogomerny statisticheskii analiz i vremennye ryady*, Moscow: Nauka, 1976.
2. Nosko, V.P., *Ekonometrika. Vvedenie v regressionnyi analiz vremennykh ryadov* (Econometrics. Introduction to Regression Analysis of the Time Series), Moscow: Manuscript available at `www.iet.ru`, 2002.
3. Kantorovich, G.G., Analysis of Time Series, *Ekonom. Zh. VSHE*, 2002, no. 2, pp. 251–273.
4. Nelson, C.R. and Kang, H., Pitfalls in the Use of Time as an Explanatory Variable in Regression, *J. Business & Economic Stat.*, 1984, vol. 2, pp. 73–81.
5. Chan, K.H, Hayya, J.C., and Ord, J.K., A Note on Trend Removal Methods: The Case of Polynomial Versus Variate Differencing, *Econometrica*, 1977, vol. 45, pp. 737–744.
6. Nelson, C.R. and Kang, H., Spurious Periodicity in Inappropriately Detrended Time Series, *J. Monetary Econom.*, 1981, no. 10, pp. 139–162.
7. Slutsky, E., The Summation of Random Causes as the Source of Cyclical Processes, *Econometrica*, 1937, vol. 4, pp. 105–146.
8. Stock, J.H., Unit Roots, Structural Breaks and Trends, in *Handbook of Econometrics*, 1994, vol. IV, pp. 2740–2841.
9. Mann, H.B. and Wald, A., On Stochastic Limit and Order Relationships, *Ann. Math. Stat.*, 1943, vol. 14, pp. 217–277.
10. White, J.S., The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case, *Ann. Math. Stat.*, 1958, vol. 29, pp. 1188–1197.
11. Phillips, P.C.B., Time Series Regression with a Unit Root, *Econometrica*, 1987, vol. 55, pp. 277–301.
12. Dickey, D.A., Estimation and Hypothesis Testing in Nonstationary Time Series, *PhD Dissertation*, Ames: Iowa State Univ., 1976.
13. Fuller, W.A., *Introduction to Statistical Time Series*, New York: Wiley, 1976.
14. Hamilton, J.D., *Time Series Analysis*, Princeton: Princeton Univ. Press, 1994.
15. Dickey, D.A. and Fuller, W.A., Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *J. Am. Stat. Ass.*, 1979, vol. 74, pp. 427–431.
16. Dickey, D.A. and Fuller, W.A., Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica*, 1981, vol. 49, pp. 1057–1072.
17. Dolado, H., Jenkinson, T., and Sosvilla-Rivero, S., Cointegration and Unit Roots, *J. Econom. Surveys*, 1990, no. 4, pp. 243–273.
18. Phillips, P.C.B. and Perron, P., Testing for a Unit Root in Time Series Regression, *Biometrika*, 1987, vol. 75, pp. 335–346.
19. Newey, W. and West, K., A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 1987, vol. 55, pp. 703–708.
20. MacKinnon, J.G., Critical Values for Cointegration Tests, in *Longrun Economic Relationships: Readings in Cointegration, Ch. 13*, Engle, R.F. and Granger, C.W.J., Eds., Oxford: Oxford Univ. Press, 1991.
21. MacKinnon, J.G., Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests, *J. Business Econom. Stat.*, 1994, no. 12, pp. 167–176.
22. Ericsson, N.R. and MacKinnon, J.G., Distributions of Error Correction Tests for Cointegration, *Econom. J.*, 2002, no. 5, pp. 185–318.

23. Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y., Testing of the Null Hypothesis of Stationary against the Alternative of a Unit Root, *J. Econom.*, 1992, vol. 54, pp. 159–178.

24. Elliott, G., Rothenberg, T.J., and Stock, J.H., Efficient Tests for an Autoregressive Unit Root, *Econometrica*, 1996, vol. 64, pp. 813–836.

25. Granger, C.W.J., Developments in the Study of Cointegrated Variables, *Oxford Bull. Econom. Stat.*, 1986, vol. 48, pp. 213–228.

26. Pollak, R., *The Theory of the Cost-of-Living Index*, Oxford: Oxford Univ. Press, 1989.

27. Straim, S.S., Survey of Literature of Demand for Money: Theoretical and Empirical Work with Special Reference to Error—Correction Models, *IMF Working Papers/WP/99/64*, Washington Int. Monetary Fund, 1999.

28. Johansen, S. and Juselius, K., Maximum Likelihood Estimation and Inferences on Cointegration—with Applications to the Demand for Money, *Oxford Bull. Econom. Stat.*, 1990, vol. 52, pp. 169–210.

29. Pelipas', I., Demand for Money and Inflation in Belarus, *EKOVEST*, 2001, no. 1, pp. 6–63.

30. Granger, C.W.J. and Newbold, P., Spurious Regressions in Econometrics, *J. Econom.*, 1974, vol. 2, pp. 111–120.

31. Phillips, P.C.B., Understanding Spurious Regressions in Econometrics, *J. Econom.*, 1986, vol. 33, pp. 311–40.

32. Granger, C.W.J., Some Properties of Time Series Data and Their Use in Econometric Model Specification, *J. Econom.*, 1981, vol. 16, pp. 121–130.

33. Granger, C.W.J., Co-Integrated Variables and Error-Correcting Models, *UCSD Discussion Paper*, 1983, pp. 83–113.

34. Engle, R.F. and Granger, C.W.J., Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 1987, vol. 55, pp. 251–276.

35. Johansen, S., Statistical Analysis of Cointegration Vectors, *J. Econom. Dynam. Control*, 1988, no. 12, pp. 231–254.

36. Johansen, S., Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, 1991, vol. 59, pp. 1551–1580.

37. Phillips, P.C.B., Optimal Inference in Cointegrated Systems, *Econometrica*, 1991, vol. 59, pp. 283–306.

38. Phillips, P.C.B. and Ouliaris, S., Testing for Cointegration Using Principal Components Methods, *J. Econom. Dynam. Control*, 1988, no. 12, pp. 205–230.

39. Phillips, P.C.B. and Ouliaris, S., Asymptotic Properties of Residual based Tests for Cointegration, *Econometrica*, 1990, vol. 58, pp. 165–193.

40. Phillips, P.C.B. and Durlauf, S.N., Multiple Time Series Regression with Integrated Processes, *Rev. Econom. Studies*, 1986, vol. 53, pp. 473–495.

41. Sims, C.A., Macroeconomics and Reality, *Econometrica*, 1980, vol. 48, pp. 1–48.

42. Lütkepohl, H., *Introduction to Multiple Time Series Analysis*, Berlin: Springer, 1993.

43. Watson, M.W., Vector Avtoregression and Cointegration, in *Handbook of Econometrics*, Amsterdam: North-Holland, 1994, vol. 4, pp. 2844–2915.

44. Sargan, J.D., Wages and Prices in the United Kingdom: A Study in Econometric Methodology, in *Econometric Analysis for National Economic Planning*, Hart, P.E., Mills, G., and Whittaker, J.N., Eds., London: Butterworths, 1964.

45. Beveridge, S. and Nelson, C.R., A New Approach to Decomposition of Time Series in Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle', *J. Monetary Econom.*, 1981, vol. 7, no. 15, pp. 1–74.

46. Perron, P., The Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *Econometrica*, 1989, vol. 57, pp. 1361–1401.

47. Maddala, G., The Effects of Different Types of Outliers on Unit Root Tests, *Advance Econom.*, vol. 13, Greenwich: JAI Press, 1997.

48. Maddala, G. and Kim, I.M., *Unit Roots, Cointegration and Structural Change*, Cambridge: Cambridge Univ. Press, 1998.

49. Perron, P., Further Evidence on Breaking Trend Functions in Macroeconomic Variables, *J. Econom.*, 1997, vol. 80, pp. 355–385.

50. Gregory, A.W., Nason, J.M., and Watt, D., Testing for Structural Breaks in Cointegrated Relationships, *J. Econom.*, 1994, pp. 491–504.

51. Borovkov, A.A., *Matematicheskaya statistika* (Mathematical Statistics), Moscow: Nauka, 1984.

52. Wilks, S., *Mathematical Statistics*, New York: Wiley, 1962. Translated under the title *Matematicheskaya statistika*, Moscow: Nauka, 1967.

53. Chow, G.C., Tests of Equality between Sets of Coefficients in Two Linear Regressions, *Econometrica*, 1960, vol. 28, pp. 591–605.

54. Chow, G.C., A Comparison of the Information and Posterior Probability Criteria for Model Selection, *J. Econom.*, 1981, vol. 16, pp. 21–33.

55. Quandt, R.E., Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes, *J. Am. Stat. Ass.*, 1960, vol. 55, pp. 324–330.

56. Davies, R.B., Hypothesis Testing when a Nuisance Parameter is Present Only under the Alternative, *Biometrika*, 1977, vol. 64, pp. 247–254.

57. Andrews, D.W.K., Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica*, 1993, vol. 61, pp. 821–856.

58. Kim, H.J. and Siegmund, D., The Likelihood Ratio Test for a Change-Point in Simple Linear Regression, *Biometrika*, 1989, vol. 76, pp. 409–423.

59. Hansen, B.E., Approximate Asymptotic p Values for Structural-Change Tests, *J. Business & Econom. Stat.*, 1997, vol. 15, pp. 60–67.

60. Hansen, B.E., Tests for Parameter Instability in with Regression with $I(1)$ Processes, *J. Business Econom. Stat.*, 1992, no. 10, pp. 321–335.

61. Andrews, D.W.K. and Ploberger, W., Optimal Tests when a Nuisance Parameter is Present Only under the Alternative, *Econometrica*, 1994, vol. 62, pp. 1383–1414.

62. Bai, J., Least Squares Estimation of a Shift in Linear Processes, *J. Time Series Anal.*, 1994, vol. 5, pp. 453–472.

63. Bai, J., Estimating Multiple Breaks One at a Time, *Econom. Theory*, 1997, no. 13, pp. 551–563.

64. Bai, J., Estimation of a Change Point in Multiple Regression Models, *Rev. Econom. Stat.*, 1997, vol. 79, pp. 551–563.

65. Bai, J. and Perron, P., Estimating and Testing Linear Models with Multiple Structural Changes, *Econometrica*, 1998, vol. 66, pp. 47–78.

66. Chong, T.T.L., Partial Parameter Consistency in a Misspecified Structural Change Model, *Econom. Lett.*, 1995, vol. 49, pp. 351–357.

67. Hardle, W., Kleinow, T., and Stahl, G., *Applied Quantitative Finance*, New York: Springer, 2002.

68. Lepski, O., One Problem of Adaptive Estimation in Gaussian White Noise, *Theory Probab. Appl.*, 1990, vol. 35, pp. 459–470.

69. Lepski, O. and Spokoiny, V., Optimal Pointwise Adaptive Methods in Nonparametric Estimation, *Ann. Stat.*, 1997, vol. 25, pp. 2512–2546.

70. Liptser, R. and Spokoiny, V., Deviation Probability Bound for Martingales with Applications to Statistical Estimation, *Stat. & Prob. Lett.*, 1999, vol. 46, pp. 347–357.

71. Kuan, C.M. and Hornik, K., The Generalized Fluctuation Test: A Unifying View, *Econom. Rev.*, 1995, vol. 14, pp. 135–161.

72. Brown, R.L., Durbin, J., and Evans, J.M., Techniques for Testing the Constancy of Regression Relationships over Time with Comments, *J. Royal Stat. Soc.*, 1975, series B, vol. 37, pp. 149–192.

73. Ploberger, W. and Kramer, W., The CUSUM Test with OLS Residuals, *Econometrica*, 1992, vol. 60, pp. 271–286.

74. Chu, C.S.J., Hornik, K., and Kuan, C.M., MOSUM Tests for Parameter Constancy, *Biometrika*, 1995, vol. 82, pp. 603–617.

75. Ploberger, W., Kramer, W., and Kontrus, K., A New Test for Structural Stability in the Linear Regression Model, *J. Econom.*, 1989, vol. 40, pp. 307–318.

76. Chu, C.S.J., Hornik, K., and Kuan, C.M., The Moving-Estimates Test for Parameter Stability, *Econom. Theory.*, 1995, no. 11, pp. 699–720.

77. Billingsley, P., *Weak Convergence of Probability Measures*, New York: Wiley, 1968.

78. Karatzas, I. and Shreve, S.E., *Brownian Motion and Stochastic Calculus*, New York: Springer, 1991.

79. Inclan, C. and Tiao, G.C., Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance, *J. Am. Stat. Ass.*, 1994, vol. 89, pp. 913–923.

80. Kramer, W., Ploberger, W., and Alt, R., Testing for Structural Change in Dynamic Regression Models, *Econometrica*, 1988, vol. 56, pp. 1355–1369.

81. Tran, K.C., Testing for Structural Change in the Dynamic Adjustment Model with Autoregressive Errors, *Empirical Econom.*, 1999, vol. 24, pp. 61–76.

82. Chu, C.S.J., Stinchcombe, M., and White, H., Monitoring Structural Change, *Econometrica*, 1996, vol. 64, pp. 1045–1065.

83. Leisch, F., Hornik, K., and Kuan, C.M., Monitoring Structural Changes with the Generalized Fluctuation Test, *Econometric Theory*, 2000, vol. 16, pp. 835–854.

84. Wald, A., *Posledovatel'nyi analiz* (Successive Analysis), Moscow: Fizmatgiz, 1960.

85. Shiryaev, A.N., *Statisticheskii posledovatel'nyi analiz. Optimal'nye pravila ostanovki* (Successive Statistical Analysis. Optimal Stop Rules), Moscow: Nauka, 1976.

86. Page, E.S., Continuous Insrection Schemes, *Biometrika*, 1954, vol. 41, pp. 100–115.

87. Lorden, G., Procedures for Reacting to a Change in Distribution, *Ann. Math. Stat.*, 1971, vol. 42, pp. 1897–1908.

88. Moustakides, G.V., Optimal Stopping Times for Detecting Changes in Distributions, *Ann. Stat.*, 1986, vol. 14, pp. 1379–1387.

89. Ritov, Y., Decision Theoretic Optimality of the CUSUM Procedure, *Ann. Satist.*, 1990, vol. 18, pp. 1464–1469.

90. Grebenyuk, E.A., Monitoring of Nonstationary Processes: Analysis and Study of Changes in the Staitonarity Properties, *Probl. Upravlen.*, 2004, no. 3, pp. 15–20.

91. Grebenyuk, E.A., Detection of Changes in the Properties of Nonstationary Random Processes, *Avtom. Telemekh.*, 2003, no. 12, pp. 44–59.

92. Grebenyuk, E.A., Analysis and On-line Diagnosis of Systems Described by Nonstationary Random Processes, *Probl. Upravlen.*, 2003, no. 4, pp. 23–29.

93. Perron, P. and Vogelsang, T., Nonstationarity and Level Shifts with an Application to Purchasing Power Parity, *J. Business Econom. Stat.*, 1992, no. 10, pp. 301–320.

94. Zivot, E. and Andrews, D.W.K., Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *J. Business Econom. Stat.*, 1992, no. 10, pp. 251–270.

95. Perron, P., Further Evidence on Breaking Trend Functions in Macroeconomic Variables, *J. Econom.*, 1997, vol. 80, pp. 355–385.

96. Nelson, C.R. and Plosser, C.I., Trends and Random Walks in Macroeconomic Time Series, *J. Monetary Econom.*, 1982, no. 10, pp. 139–162.

97. Gregory, A.W. and Hansen, B.E., Residual-based Tests for Cointegration in Models with Regime Shifts, *J. Econom.*, 1996, vol. 70, pp. 99–126.

This paper was recommended for publication by V.A. Lototskii, a member of the Editorial Board

Copyright of Automation & Remote Control is the property of Springer Science & Business Media B.V.. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.