

ABOUT NON-STATIONARITY IN TIME SERIES OF  
FIELD MEASUREMENTS AND METHODS FOR THEIR  
PRIMARY ANALYSISV.S. PETRAKOVA *Представлено П.П. ПЕТРОВЫМ*

**Abstract:** The paper is aimed at analyzing the difficulties that arise when checking time series of field measurements for stationarity and identifying the main characteristics of the series. Attention is paid to the correctness of using the most popular methods of primary analysis for real data. The author's simple procedure for determining the changes in mean and variance of time series. The proposed procedure is tested on synthetic data. The methods described in the paper are applied to real time series, presented in the form of measurements of the concentration of PM2.5 suspended particles in the atmospheric boundary layer of the urban city.

**Keywords:** primary time series analysis, stationarity tests, correlogram analysis, variance ratio test.

---

THE WORK WAS PERFORMED ACCORDING TO THE GOVERNMENT RESEARCH ASSIGNMENT FOR SOBOLEV INSTITUTE OF MATHEMATICS SB RAS, PROJECT FWNF-2024-0002.

© 2024 ПЕТРАКОВА В.С..

Работа выполнена в рамках государственного задания Института математики им. С.Л. Соболева СО РАН, проект FWNF-2024-0002 "Обратные некорректные задачи и машинное обучение в биологических, социально-экономических и экологических процессах".

*Поступила 1 января 2023 г., опубликована 31 декабря 2023 г.*

## 1 Introduction

A time series is a collection of observations of the time evolution of some variable. Time series occur in a wide variety of fields, from economics to the physical sciences, and having prior knowledge of the causes and strength of the relationship between the dependent and independent variable (time) distinguishes a set of elements of a time series from a sample of values of a random variable, making time series analysis more labor-intensive process. Also, the behavior of a variable reflected by a time series cannot be reproduced using repeated observations, since only one implementation occurs at each point in time, and the presence of a cause-and-effect relation makes each time series unique.

The most important problem in time series analysis (TSA) is their forecasting. Indeed, a large number of studies in various areas of science and technology has been devoted to the development of methods for predicting the further development of the area represented by a set of data and assessing the result of such a forecast (see review articles [1] and [2]). The popular forecasting methods here remain methods of analyzing time series of the ARIMA type [3], Holt-Winters model [4], regression and classification models [5], as well as neural networks and other machine learning type methods [2]. In the first half of 2024 alone, according to the Google Scholar service, more than 20 thousand articles appeared aimed at developing this topic. However, it is known that models based on identifying various kinds of trends give a significant error if trend changes [4]. This leads to the development of break-points detection methods in time series analysis [6]. On the other hand, methods based on machine learning are dependent on the choice of the training sample, which, in the case of a complex time series (for example, with implicit seasonality, imbalances, physically significant anomalies), atypical scenarios can be attributed, adding an error that cannot be assessed due to the characteristics of machine learning. Thus, prior knowledge about the properties of a series is crucial in its prediction.

At the same time, the (weak) stationarity of a series or its subsequence is actually the only guarantee that the resulting forecast will be the most accurate, regardless of the model used, since stationarity implies the preservation of the main statistical characteristics of the series over time. In a mathematically strict sense, the desire to deal with a stationary series is caused by the possibility of justifying forecast models for such a series using the Wold theorem [7] on decomposition, according to which any stationary process can be uniquely represented as the sum of two uncorrelated processes: a deterministic one, the forecast of which for any time in advance error-free, and purely random (white noise). In practice, this means a simple property: the researcher understands how the system behaves, and this behavior persists over time. This is the main reason why the test for the stationarity of a series is the first step in its forecasting. However, even such a simple test can cause a number of difficulties due to the characteristics of the

available statistical criteria. This paper presents an overview of some tools for assessing the stationarity of time series of field measurements and their functional interpretations using the example of real measurements of the concentration of suspended particles with a diameter less than 2.5 microns (PM2.5) in Krasnoyarsk in 2020. Limitations on the applicability of some popular approaches are shown.

The paper is organized as follows. Firstly, in Section 2 it is described the data of field measurements under research and problems which arise in this area. The time series considered in the work are measurements of the pollutant concentration in the boundary layer of the atmosphere of Krasnoyarsk. After in Section 3 it is proposed the overview of the most popular methods of primary time series analysis. Here are described (a) graphical methods: ACF and PACF analysis; (b) statistical tests: Dickey-Fuller, Phillips–Perron, Kwiatkowski–Phillips–Schmidt–Shin ones; (c) procedures for assessing the behavior of statistical characteristics of a series: Cochrane procedure. Then the described methods are applied to available data in section 4. The section also mentions periodogram analysis, which is a primary analysis method and allows additional conclusions to be drawn about the system under study. Section 5 has a description of the author’s proposed procedure and the tests of it on synthetic and real data. It is based on the analysis of histograms of the distribution of means and variances of individual subsequences of the time series and aimed at obtaining additional information about the structure of time series. And the final section is aimed at conclusions and discussions. Thus, sections 3 and 4 are overviews and contain information about already known more or less frequently used methods of estimating the stationarity of a time series and the results of their application to real data. Section 5 is devoted to a new method for primary TSA and an analysis of its applicability to synthetic and real data.

## 2 Description of the data under study

A good example of a time series encountered in practice is environmental measurements, which may include both meteorological measurements and air quality data. A generally accepted marker and, at the same time, one of the most harmful air pollutants in the ground layer of the atmosphere of modern cities is particulate matter with a diameter of 2.5 microns or less (PM2.5). The concentration of suspended particles PM2.5 is a basic indicator of pollution in a city or region, affecting the health of its residents, as well as a criterion for its attractiveness for living. This necessitates the need to identify typical scenarios for the dynamics of periods of increased PM2.5 concentrations and to develop new methods for predicting such periods. Concentrations of fine particular matters in the atmospheric boundary layer are known to be sensitive to meteorological factors (e.g., air temperature and relative humidity) [8] and human activities, and tend to show changes depending on the day of the week and season [9]. Understanding temporal

trends is critical to accurately determining PM<sub>2.5</sub> concentrations and mitigating human exposure to air pollution. Now, the methodology commonly used to estimate environmental parameters is based on classical descriptive statistics, but this is of rather limited value due to the large variability associated with air quality data and the low signal-to-noise ratio of available measurements. Time series analysis may be a good approach to avoid these difficulties by allowing the identification of underlying deterministic behavior and thus contributing to the understanding of cause-and-effect relationships in environmental problems.

Krasnoyarsk is one of the Russian cities where atmospheric air quality is monitored at stationary observation posts. The Ministry of Ecology and Rational Natural Resources Management of the Krasnoyarsk Territory maintains the regional departmental information and analytical data system on the state of the environment of the Krasnoyarsk Territory (RDIAS). Nine automated observation posts (AOP) of RDIAS are located in Krasnoyarsk. Once every 20 minutes, automatic measurements of meteorological parameters and the concentration of pollutants in the surface layer of the atmosphere are performed. To monitor PM<sub>2.5</sub> concentrations, RDIAS uses dust analyzers model E-BAM (Met One Instruments Inc., USA, see [10]), the operating principle of which is based on measuring the absorption of  $\beta$  radiation by dust particles deposited on a filter tape. This method is certified by the U.S. EPA (United States Environmental Protection Agency, see [11]). Analyzers of this class are recommended for measuring the content of PM<sub>10</sub> and PM<sub>2.5</sub> fractions in the atmosphere, certified and accredited in many countries around the world, including Russia (No. 57884-14 in the State Register of Measuring Instruments). The available data set contains measurements from 00:20 January 1, 2019 to 23:40 December 31, 2022. Note that the data contains a very large number of gaps due to failures of dust analyzers. The longest subsequence of the time series from the “Kirovsky” post (link to the geoportal with the location <https://air.krasn.ru/map.html?2=>) is 181 days in the period from 2020-02-23 11:20 to 2020-08-23 07:40 ( $181 \times 72 = 13032$  measurements).

Note that at present, dozens of works by researchers are devoted to the analysis of data on PM<sub>2.5</sub> concentrations from a network of spatially distributed sensors and subsequent forecasting of the development of the environmental situation in the region, for example, see studies [3, 9, 12, 13, 14].

Figure 1 shows a graphical representation of the data under study. Here, Figure 1a shows the original data (measurements every 20 minutes) and averaged measurements per day. And Figure 1b displays logarithm of data on PM<sub>2.5</sub> concentrations and their averages per day. To simplify the text, below it will be called the measurements obtained directly from the post as “original data”; averaged original data for a day as “daily original data”; and logarithmic original measurements (and their averaging) as “logarithmic data” (“daily logarithmic data”). Briefly describe the data and the problems associated with them.

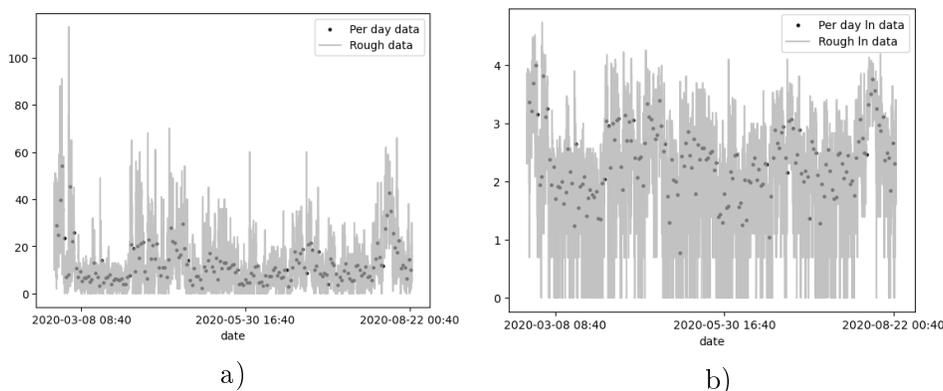


FIG. 1. a) “original data” (gray line), “daily original data” (black dots); b) “logarithmic data” (gray line) and “daily logarithmic data” (black dots)

To check the studied data for normality, the Kolmogorov-Smirnov, Shapiro-Wilk, Jarque-Bera, Anderson-Darling, and D’Agostino-Pearson tests were performed on each sample. Here it is not proposed the description and comparison of the used tests with each other; it is only noted that there are many works on their comparison on various synthetic data (see, for example, the works [15]–[19]). Note that the mentioned studies don’t contain an unambiguous conclusion on the optimality of one of the selected methods when used to study the normality of a sample, when the distribution of sample elements is not known for certain. When applied to the data studied in this work, all tests performed indicate the need to reject the null hypothesis of normal distributions, except for “daily logarithmic data”. For the mentioned time series, all tests, except the Kolmogorov-Smirnov test, indicate a high probability of accepting the null hypothesis about the normality of the sample distribution. The obtained result can be explained by the fact that the Kolmogorov-Smirnov test doesn’t work well on large samples with a weakly skewed distribution. Note that visual analysis of histograms of distributions of the studied data confirms the test results. Consideration of logarithmic data separately, as is already clear, is due to the log-normality of the original data.

The main problem with this type of data is the physicality of outliers that cannot be excluded from the sample. An increase in PM<sub>2.5</sub> concentration can develop both in short periods, associated, for example, with street traffic during rush hours, and in long-term excesses (several days, several hundred measurements), associated with emissions from manufacturing companies and lack of wind. And all the specified templates are embedded in the data. The second problem is related to numerous data gaps due to hardware failure. There can be up to a dozen such gaps per year (several hundred

measurements). And for the correct use of data in the future, these missing data must be restored.

The primary assumption about the series depicted in Figure 1 is that the “original data” (and “daily original data”) are most likely not stationary, since trend change points which don’t fit within the standard deviation are visible. A visual analysis of the logarithmic series shows that it may turn out to be stationary, which allows to use almost any model to predict it. One of the goals of the work is to test this thesis in different ways and identify associated difficulties.

### 3 The “stationarity” concept

A time series is called stationary if its main statistical characteristics do not depend on time. In mathematical statistics, there are two types of stationarity: strong and weak one. They are defined as follows [20]: a random process  $X_t = \{x_t\}$  with  $t \in \{t_1, t_2, \dots, t_N\}$  is said to be strongly stationary if for any set of times  $t_1, t_2, \dots, t_n$  and any integer  $k$  joint probability distributions  $x_{t_1}, x_{t_2}, \dots, x_{t_n}$  and  $x_{t_1+k}, x_{t_2+k}, \dots, x_{t_n+k}$  are the same. In the language of dynamical systems, this means that in every conceivable embedding space, the statistical properties of the phase flows relating to different parts of the time series are the same. A less stringent requirement is weak stationarity: in which (a) the theoretical expectation and variance of the time series do not depend on time; (b) the covariance between observations at any two points in time depends only on the time lag between them. In what follows, by “stationarity” of a time series, as in most cases described in the literature, we will understand weak stationarity.

Regarding non-stationary time sequences in the analysis of time series, it is customary to distinguish between TS-series (trend stationary), which are stationary with respect to a deterministic trend, and DS-series (difference stationary), which are not related to TS series and are reduced to stationary only as a result of their single or multiple differentiation. The fundamental difference between these two classes of series is that in the case of a TS series, subtracting the corresponding deterministic trend from the series leads to a stationary series, whereas in the case of a DS series, subtracting the deterministic component of the series leaves the series non-stationary due to the presence of a stochastic trend.

**3.1. Review of methods for primary testing of time series.** Mostly, methods for checking time series for stationarity can be divided into three groups:

- graphic;
- statistical tests;
- procedures for estimating the statistical characteristics of a series.

**3.1.1. Graphic methods.** As methods of *graphical analysis*, it can be distinguished both the simplest ones as analysis of the graph of a time series

and the graph of moving statistics; and the analysis of (partial) autocorrelation functions. It is clear that the first two “naive” methods are practically not applicable to real data series, since for them it is not clearly determined whether there is noise in the data or a trend change has occurred (as for example for the series shown in Figure 1).

#### ***Autocorrelation Function Analysis***

The more mathematically correct graphical method is analysis of time series correlograms. The degree of closeness of linear relationship between sequences of observations  $X_t = \{x_{t_1}, x_{t_2}, \dots, x_{t_n}\}$  and  $X_{t+k} = \{x_{t_1+k}, x_{t_2+k}, \dots, x_{t_n+k}\}$  of a stationary series (with lag  $k$ ) can be determined using the correlation coefficient

$$\rho(k) = \frac{E[(X_t - a)(X_{t+k} - a)]}{\sigma^2}, \quad (1)$$

where  $E[X_t] = E[X_{t+k}] = a$  is the expectation of the time series  $X_t$ ,  $\sigma$  is the standard deviation of the series  $X_t$ ,  $X_{t+k}$ . Since the coefficient  $\rho(k)$  measures the correlation between members of the same series, it is called the autocorrelation coefficient, and the dependence of  $\rho(k)$  on the lag value  $k$  is called the autocorrelation function (ACF). In time series analysis, the autocorrelation function shows the degree of linear statistical relationship between the values of a time series. If any long-term external factors act on a series, this leads to the appearance of trends (tendencies) and a cyclical component in the series, which the ACF allows to detect. Based on the appearance of the correlogram, one can, for example, draw the following conclusions. If the expressed maximum of the correlogram turns out to be for lag  $l = k$ , then the time series contains a cyclic component with period  $k$ . If the correlogram has a maximum at  $l = 1$ , then the series contains only a linear trend. If the correlogram does not have pronounced maxima, then the series does not contain a trend or a cyclic component, and the random component dominates in it. This may also indicate that the series contains a strong nonlinear trend that cannot be expressed by a linear correlation coefficient. Also, if the correlogram represents small changes around zero, that means then the series is random.

Also, together with the autocorrelation function, the autocorrelation function (PACF) is considered. For a time series  $X_t$ , partial autocorrelation with lag  $k$  is the autocorrelation between  $X_t$  and  $X_{t+k}$  obtained after removing the influence of any correlation with shorter lags  $1, 2, \dots, k - 1$ . Note that the autocorrelation of observation  $x_i$  and the observation at the previous time step  $x_{i-1}$  includes direct correlation and indirect correlation. These indirect correlations are correlations of observations of linear functions. The partial autocorrelation function attempts to remove these indirect correlations. Partial autocorrelation of time series is used to find periodicities in time series and find the order of the autoregressive model of the series. Partial autocorrelation with lag  $k$  can be found according to the Durbin-Levinson algorithm

[21]

$$\phi(k, k) = \frac{\rho(k) - \sum_{i=1}^{k-1} \phi(k-1, i)\rho(k-i)}{1 - \sum_{i=1}^{k-1} \phi(k-1, i)\rho(i)}, \quad (2)$$

where  $\phi(k, i) = \phi(k, i) = \phi(k-1, i) - \phi(k, k)\phi(k-1, k-i)$   $1 \leq i \leq k-1$  and  $\rho(k)$  is the value of the autocorrelation function with lag  $k$ .

It should be noted that ACF and PACF analysis is most often used in time series forecasting using ARIMA model, where the lag at which the regression coefficients become insignificant determines the order of the model, see, for example, the results of modeling in Hydrological Processes [22]. The analysis of corelograms is a powerful and simple tool based on the constancy of the autocovariance, mean and variance of a stationary process, but for a clearly DS-stationary series, the method can give an erroneous result due to the analysis of only linear dependencies.

**3.1.2. Statistical Tests.** Note that the main method of checking a time series for stationarity remains statistical tests. The tests are based on the assumption that the modeled process is represented as an autoregressive process, for which the characteristic equation can be easily written (for details, see [23]). If a given characteristic equation has a unit root, then this indicates a strong correlation between neighboring terms of the series, that is, non-stationarity. There are several unit root tests and most of them have common implementations in Python, R and other statistical packages. Let's highlight some of them. *Augmented Dickey-Fuller* test checks the stationarity around a trend under the assumption that time series is presented as  $p$ -order autocorrelation process [24]. The modern implementation of *Phillip-Perron test (PP)* [25] makes the same under the assumption that the residuals cannot be distributed normally. In the *KPSS test*, the time series under study is written as the sum of a trend, a random walk process [26]. There are also tests that evaluate a series for stationarity under the assumption that it has one or more break points (trend changes). These include the *Zivot-Andrews test (ZA)* (for a single structural break point, see in [27]) and the *CUSUM test* (for multiple ones, see original works [28] and its improvements [29]). Note, that CUSUM (CUmulative SUM) is, first of all, a procedure for searching for points of characteristic change. The original procedure, proposed in 1975 [30], was focused on checking whether changes occur in the statistical mean of the series. The first work that took into account the joint change in the mean and variance of a series (that is, dividing the series into presumably stationary subsequences) was [31]. Now, there are many modifications (see [32]–[34] and other) of the method, including for series with a long memory [35]; real-time detection [36] and with the combination with other methods [37]. Today, CUSUM is able to detect transitions between a presumably stationary (without checking the constancy of the autocovariance) and non-stationary segment, but it cannot guarantee the accuracy of this identification.

The popularity of the designated tests in the Google Scholar system for the partial year 2024 is distributed as follows: almost 5000 works using the

Dickey-Fuller test, just over 3000 works mentioning the CUSUM test, about 2000, 1500 and 300 for the PP, KPSS and ZA tests respectively. At the same time, in most cases, works related to the CUSUM test are research related to the study of the procedure and its modifications. The remaining tests have been developed and researched for more than 20 years and works related to them mention them as a tool used in making predictions in real-world areas. Thus, when solving real problems, the Dickey-Fuller test remains the most used, and its usefulness is confirmed by works comparing classical tests with each other, where the ADF test shows good results [38]. Here, the ZA and CUSUM methods will not be used to analyze the data under study, since it was previously assumed that the data does not have structural changes. Describe the remaining tests in more detail and show what their difference is.

#### ***Augmented Dickey-Fuller (ADF) Test***

In statistics, ADF procedure [24] tests the null hypothesis that a sample of time series has an unit root. The alternative hypothesis differs depending on which version of the test is used, but it is usually stationarity or stationarity around trend. In fact, the null hypothesis of the Dickey-Fuller test is that the series belongs to the DS class, and the alternative – to the TS class. The process of constructing the extended Dickey-Fuller test begins by considering an autoregressive process of order  $p$ :

$$\Delta x_t = f(t) + \alpha x_{t-1} + \theta_1 \Delta x_{t-2} + \dots + \theta_p \Delta x_{t-p} + \varepsilon_t, \quad (3)$$

where  $f(t)$  is the expected trend (usually, constant or linear);  $p$  is the number of lags;  $\varepsilon_t$  are random variables having a normal distribution with zero mean; and  $\Delta x_{t-i} = x_{t-i} - x_{t-i-1}$ ,  $\forall i = 0, \dots, p$ . The equation (3) is estimated using the least squares method, after which the statistical significance of the coefficient  $\alpha$  is checked. The null hypothesis  $H_0 : \alpha = 0$  corresponds to the situation of a non-stationary time series. Alternative hypothesis  $H_1$ : the coefficient  $\alpha$  is significant and negative. Calculated statistics

$$\tau = \frac{\bar{b}}{SE_{\bar{b}}} \quad (4)$$

has a distribution provided by the Dickey-Fuller table (see [39]). A formalized procedure for using the Dickey-Fuller criteria with a sequential check of the possibility of reducing a statistical model is given in the work [40]. As mentioned above, the Dickey-Fuller test is most often used by researchers.

#### ***Phillips-Perron (PP) test***

This criterion, proposed in [25], tests the hypothesis  $H_0 : \alpha = 0$  within the framework of a statistical model

$$\Delta x_t = f(t) + \alpha x_{t-1} + \varepsilon_t, \quad (5)$$

where, as in the Dickey-Fuller criterion,  $f(t)$  can be taken equal to zero. However, in contrast to the Dickey-Fuller criterion, random components  $\varepsilon_t$  with zero mathematical expectations can be autocorrelated (with a fairly

rapid decrease in the autocorrelation function), have different variances (heteroscedasticity) and are not necessarily selected from the normal distribution. Thus, in contrast to the criterion Dickey–Fuller, a wider class of time series is allowed for consideration. A more detailed description of the Phillips-Perron test is presented in [41].

***Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test***

The main difference between the KPSS test [26] and those discussed above is the permutation of hypotheses. In the KPSS test, the null hypothesis states that the time series is stationary, against the alternative that there is non-stationarity (belonging to the DS class). It is also necessary to note significant methodological differences. The KPSS test approach assumes that the time series  $X_t$  tested for stationarity with respect to trend can be decomposed into the sum of deterministic trend  $\beta t$ , random walk  $r_t$  and stationary error  $\varepsilon_t$ :

$$x_t = \beta t + r_t + \varepsilon_t, \quad r_t = r_{t-1} + u_t, \tag{6}$$

where  $u_t$  is a normally distributed process with zero mean and variance  $\sigma^2$  ( $u_t \sim N(0, \sigma^2)$ ).

From the equation above, it follows that the null hypothesis  $H_0$  that  $X_t$  is stationary is equivalent to the hypothesis  $\sigma^2 = 0$ , which implies that  $r_t = r_0$  for all  $t$  ( $r_0$  is a constant). Similarly, the alternative hypothesis  $H_1$  about nonstationarity is equivalent to the hypothesis  $\sigma^2 \neq 0$ . To test the hypothesis  $H_0$  against the alternative  $H_1$  the authors of the KPSS test obtained one-sided Lagrange multiplier test statistics [42]. They also calculate its asymptotic distribution and model the asymptotic critical values.

**3.2. Procedures for estimating the statistical characteristics of a series.** The last group of methods is based on determining stationarity and direct assessment of the statistical characteristics of the series. This may include *Cochrane procedure (variance ratio test)* [41, 43], which tests the random walk hypothesis against stationary alternatives, by exploiting the fact that the variance of random walk increments is linear in all sampling intervals. For real data, the representation of variance ratio is performed:

$$VR_k = \frac{V_k}{V_1}, \quad \text{where } V_k = \frac{N}{N-k+1} \frac{1}{k} D(X_t - X_{t-k}). \tag{7}$$

Here  $D(X_t)$  is the sample variance of  $X_t$  time series and  $N$  is number of series  $X_t$  elements. If  $X_t$  is random walk, then  $VR_k = 1, \forall k$  and if  $X_t$  is stationary process over constant or linear trend, then  $VR_k \rightarrow 0$  for  $k \rightarrow \infty$ .

Some of the other methods for testing time series for (non)stationarity involve its spectral decomposition. These methods have been developed since the seventies of the last century and new ideas and modifications are still being proposed. Pioneering work in this area [44] proposed a method for comparing a set of evolutionary spectra of a series assessed at different points in time. This approach was further developed, for example, in the works of [45, 46]. Methods for comparatively time-varying spectral density

estimation with its stationary approximation are also being developed [47]; and approaches based on Fourier transform and wavelets [48]–[51].

Methods for assessing the stationarity of time series based on frequency analysis and spectrum analysis are now being widely developed and make it possible to isolate changes in the series through periodicity analysis [23].

#### 4 Discussion of problems associated with applying some well-known methods to real data

Using the example of the data presented in Section 2, it will be shown that using a single tool to understand the behavior of a system can lead to serious errors. In the Table 1  $p$ -values of stationarity tests performed on the studied data are presented. Values that indicate presumptive non-stationarity of the selected series are highlighted in bold. As can be seen from the table, the set of tests does not allow us to draw an unambiguous conclusion about the part of the data that is most interesting for research. This is original data from field measurements of PM2.5 concentrations and their logarithmic representation.

TABLE 1. P-values for ADF, PP, KPSS tests in comparison with a significance level equaled 0.05

time series	tests taking into account possible constant or linear trends					
	ADF	ADF	PP	PP	KPSS	KPSS
	(const)	(linear)	(const)	(linear)	(const)	(linear)
original data	0.00	0.00	0.00	0.00	<b>0.025</b>	<b>0.000</b>
daily original data	0.00	0.00	0.00	0.00	0.434	0.086
logarithmic data	0.00	0.00	0.00	0.00	<b>0.039</b>	<b>0.000</b>
daily logarithmic data	0.00	0.00	0.00	0.00	0.556	0.140

Note that, it would seem, the most popular tests for stationarity (ADF and PP) counted the series as stationary; but really they only indicate that the series does not have *one* unit root. At the same time, the KPSS test regarding daily data also indicates that it is impossible to confidently draw a conclusion about their stationarity. Thus, in this case, statistical testing is not a comprehensive tool in analyzing the stationarity of time series.

Now turn to the analysis of autocorrelation functions. Figure 3 shows correlograms (ACF) of the data used for analysis. Here, the  $x$  axis is the lag value  $k$ , and the  $y$  axis is the value of the correlation coefficient, calculated using the formula (1). The area highlighted in blue indicates that the coefficients falling within this area are not significant with a probability of 95%. Thus, the corelogram of the original data shows their obvious non-stationarity and complex seasonality. Note that the period over which the oscillations occur is approximately about 70 measurements, that is, about a day in the format of the available data. This is confirmed by the fact that the use of

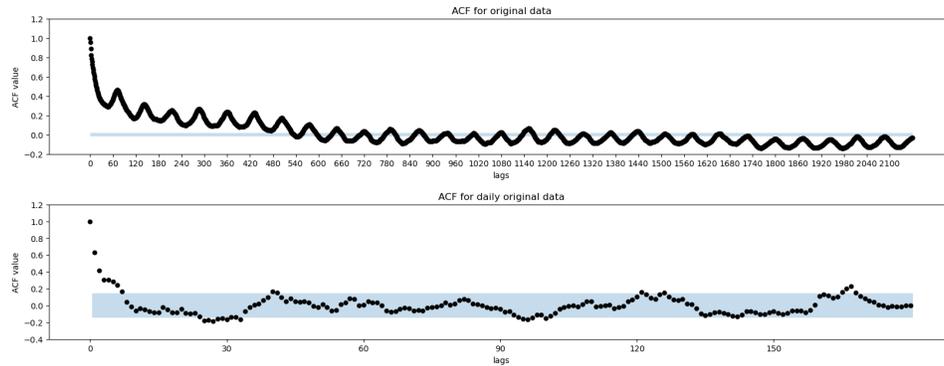


FIG. 2. ACF graphics for “original data” (top) and “daily original data” (bottom)

daily averaging makes it possible to eliminate such frequent fluctuations. The ACF of “daily original data” decreases rapidly, which may indicate both stationarity and the presence of a nonlinear trend, since the first few ACF values are significant, while the decrease in ACF also occurs non-linearly. Note that the figure does not show corellograms of the “logarithmic data”, since they do not differ from the corresponding graph for the “original data”.

Figure 3 shows the PACF for the data set under study. Analysis of partial

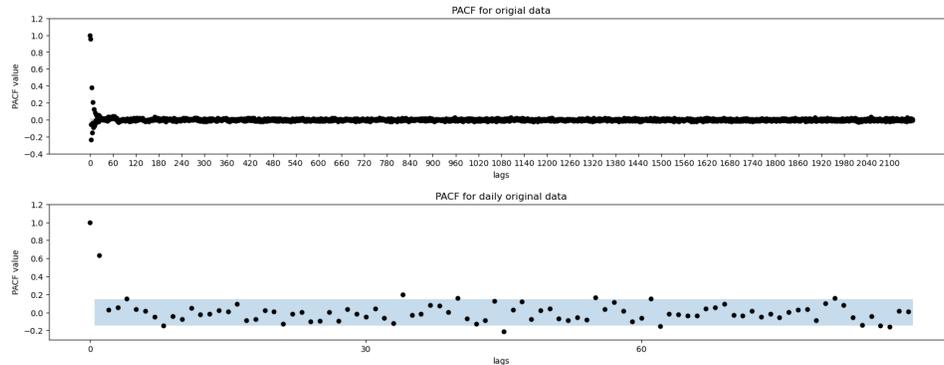


FIG. 3. PACF graphics for “original data” (top) and “daily original data” (bottom)

autocorrelation functions generally confirms the conclusions drawn from the ACF analysis.

Figure 4 shows the values of  $VR_k$  as a function of lag  $k$  for the Cochrane procedure described in Section 3.2. Note an interesting feature. The ratio of variances for the “original data” and the “logarithmic original data” behaves like white noise or the sum of white noise and linear trend, which is no different from the procedure under consideration. At the same time, their daily averaging, on the contrary, shows the non-stationarity of the process

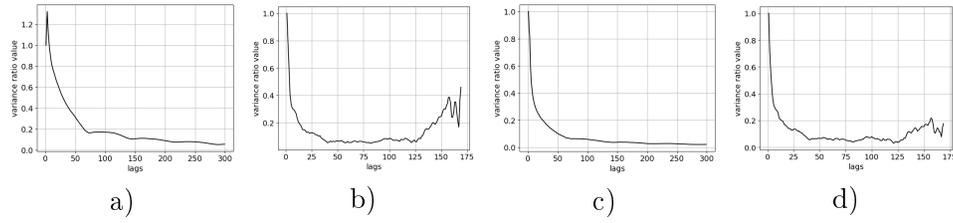


FIG. 4. Variance ratio  $VR_k$  values depending on lag  $k$  for a) “original data”, b) “daily original data”; c) “logarithmic data” and d) “daily logarithmic data”

to a greater (for “daily original data”) and lesser (for “daily logarithmic data”) degree. Most likely, this behavior is due to the measurement characteristics of the sensors from which the data is collected. Thus, optical sensors at low concentrations of PM2.5 can make significant errors (see [10]). Daily smoothing solves this problem and then the method can capture the non-stationary nature of the process itself.

To fully understand the structure of the time series used, it will be also constructed their periodograms. Periodogram analysis is one of the representatives of the methods of spectral analysis of a series. A periodogram is a function of frequency that shows an assessment of the spectral density of a time series. Periodograms are not used directly to assess the stationarity of a series, but they do provide insight into how different it is from white noise. Periodogram analysis also allows you to visually assess the presence of a trend and seasonality in data. Here high amplitudes in the low frequency region indicate the presence of red (i.e., correlated) noise in the data, which indirectly confirms the hypothesis of erroneous sensor measurements at low pollutant concentrations. More details about the analysis of periodograms and the spectrum of time series can be found in the work [52]. Figure 5 shows the periodograms of the analyzed data.

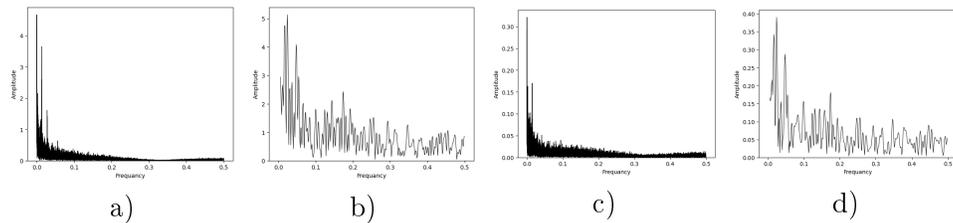


FIG. 5. Periodograms for a) “original data”, b) “daily original data”; c) “logarithmic data” and d) “daily logarithmic data”

## 5 Procedure for analyzing the mean and variance of time series

Thus, there are no simple and accurate algorithms for classifying time series as stationary. Available visual analysis is limited by sample size and researcher choice and is not mathematically rigorous. Statistical tests do not respond well to complex series. Note once again that checking the series for stationarity is closely related to the problems of understanding the structure of the series. This is also included in the definition of weak stationarity: constancy of the main statistical characteristics of the series (its mean, dispersion and autocovariance). In this section, a simple procedure will be proposed to assess both the change in the mean and variance of a series and its difference from the behavior of white noise.

**5.1. Procedure description.** The idea of the proposed method is quite simple and is based on the fact that if a series is stationary, then the distributions of the it's sample mean and sample variance of subsequences of large enough sizes are close to symmetrical. Prove this assumption. Assume that some stationary process is defined. Let

$$X = \{x_1, x_2, \dots, x_N\}, \quad (8)$$

where  $N$  is large enough to reflect all the properties of the statistical population, be a sample of sequential values from this process. Thus,  $X$  is a series with constant mathematical expectation  $\mu = \mu_X$  and constant variance  $\sigma^2 = \sigma_X^2$ , and autocovariance depends only on the lag value.

Suppose there is a sample  $\bar{X}_m$  such that

$$\bar{X}_m = \{\bar{X}_{i,m}\}, \text{ where } \bar{X}_{i,m} = \frac{x_i + x_{i+1} + \dots + x_{i+m}}{m}, \quad (9)$$

and  $i$  can take values from 1 to  $N - m$ ;  $|\bar{X}_m| = M$ . From the law of large numbers in Khinchin's form it follows that  $\bar{X}_{i,m} \rightarrow \mu_X$  in probability, when  $n \rightarrow \infty$ .

Now show that the sample third central moment tends to 0 in probability for the sample  $Y = \{\bar{X}_m\}$ , where  $m$  can take several values. For simplicity, show the assertion if  $Y$  consists of 2 samples:  $\bar{X}_m$  and  $\bar{X}_l$ . For a larger number of them, the proof will be similar.

Firstly, the sample mean of  $Y = \{\bar{X}_m, \bar{X}_l\}$ :

$$\bar{Y} = \frac{1}{M+L} \left( \sum_{k=1}^M \bar{X}_{i_k,m} + \sum_{k=1}^L \bar{X}_{i_k,l} \right) \rightarrow \mu_X \quad (10)$$

in probability, when  $m, l \rightarrow \infty$ . Here  $L = |\bar{X}_l|$ . Second, the sample third central moment is

$$M_3^Y = \frac{1}{M+L} \left( \sum_{k=1}^M (\bar{X}_{i_k,m} - \bar{Y})^3 + \sum_{k=1}^L (\bar{X}_{i_k,l} - \bar{Y})^3 \right) \rightarrow 0 \quad (11)$$

in probability, when  $m, l \rightarrow \infty$ , due to (10) and  $\bar{X}_{i,m} \rightarrow \mu_X$ .

Thus,  $M_3^Y$  is close to zero for the sample of sample means of subsequence of different sizes and the variance  $\sigma_Y^2 < \infty$  for sample  $Y$  due to stationarity of process described by  $X$ .

Carry out a similar proof for the sample variances of individual subsequences of the series  $X$ . For this, consider the series  $Z = \{x_1^2 - \mu_X, x_2^2 - \mu_X, \dots, x_N^2 - \mu_X\}$ . The sample mean  $\bar{Z}$  is equal to  $\sigma_X^2$ . Note that here it is not key whether the series  $Z$  is stationary; the assumption of stationarity of  $X$  is enough. Then repeat the steps (9),(10),(11) and the corresponding conclusions. Note that here for  $\bar{Z}_m = \{\bar{Z}_{i,m}\}$ :

$$\bar{Z}_{i,m} = \frac{1}{m} \sum_{j=i}^{i+m} (x_j^2 - \mu_X) \quad (12)$$

is the sample variance of the subsequence of size  $m$  of stationary series  $X$ . Thus,  $M_3^Y$  tends to zero in probability, where  $Y$  now is  $Y = \{\bar{Z}_m\}$ , where  $m$  can take several values. The variance  $\sigma_Y^2 < \infty$  due to the representation of (12) and stationarity of process  $X$ .

Now, describe the procedure. For time series  $X_t$ :

1. choose the subsequence of constants  $\Omega = \{\omega_1, \omega_2, \dots, \omega_M\}$  such that  $\max \Omega \leq N/2$ ;
2.  $\forall \omega_i$  get the set subsequence of time series of length  $\omega_i: \{x_{k\omega_i}, \dots, x_{k\omega_i + \omega_i}\}$   $k = 0, \dots, [N/\omega_i] - 1$ ;
3. generate samples  $Means = \{\bar{X}_{ik}\}$ ,  $Vars = \{\sigma^2(X_{i,k})\}$ , where  $\bar{X}_{i,k}$  is sample mean for subsequence  $\{x_{k\omega_i}, \dots, x_{k\omega_i + \omega_i}\}$ , and  $\sigma^2(X_{i,k})$  is sample variance for the same subsequence;
4. shuffle samples  $Means$  and  $Vars$  and divide them into two samples  $Means_1$ ,  $Means_2$  and  $Vars_1$ ,  $Vars_2$  of similar sizes (or not differing by more than one);
5. build scatterplots for  $Means_1$ ,  $Means_2$ ,  $Vars_1$ ,  $Vars_2$  and histograms for  $Means$  and  $Vars$  and draw conclusions about the time series in comparison with the reference behavior of stationary process.

Note that the proposed procedure is not enough for making a conclusion of stationarity of time series, because here we research only the mean and variance of time series and can be used for the analysis of the behavior of it. Here the visual analysis of scatterplots is an additional instrument.

Now demonstrate how the procedure works on synthetically generated data.

**5.2. Result of procedure application on the synthetic data.** First, generate a time series that is white noise. Put

$$X_t = \varepsilon, \quad (13)$$

where  $\varepsilon \sim N(0, 1)$  is a normally distributed variable with zero mean and unit standard deviation. The size of sample is equal to 10000, that in comparison of size of studied real data. Figure 6 shows the result of the algorithm for

the time series represented by (13) (the data were previously normalized on the  $[0, 1]$  scale for convenience of visualization).

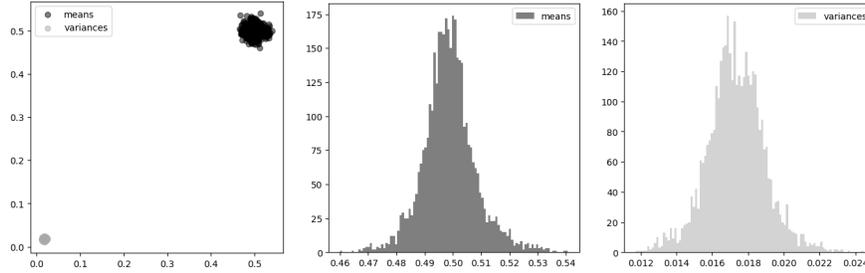


FIG. 6. The result of procedure application on white noise data represented by (13). On the left – scatter plot of *Means* and *Vars*; in the middle – histogram for *Means*; on the right – histogram for *Vars*

Now consider a time series represented by an additive model: the sum of white noise and trend.

$$X_t = at + \varepsilon, \tag{14}$$

$$X_t = at^2 + \varepsilon, \tag{15}$$

$$X_t = \exp(t^2) + \varepsilon. \tag{16}$$

Figure 7 shows the result of the proposed algorithm for the time series represented by (14)–(16) i.e. data with trends. As can be seen from the Figure 7, the presence of a trend significantly changes the average values obtained using the procedure, while the variance determined by white noise obviously does not change its behavior (except the (16) due to exponent). The presence of seasonal (periodic) changes, such as, for example, for a series given by the following dependence

$$X_t = 2 \sin(\pi t) + \varepsilon \tag{17}$$

changes the scatterplot and histogram as it is shown in Figure 8. As a final example, consider heteroscedastic data, i.e. heterogeneous observations, expressed in the variable variance of the random error of the regression model. The results of the procedure for such data are presented in Figure 9. Such data can be determined according to the law (13), where  $\varepsilon$  is a random variable distributed according to the normal law  $N(0, \sigma(t))$ , where  $\sigma(t) = t + \theta$ ,  $\theta$  is a uniformly distributed random variable. As can be seen from the results obtained, here the non-constancy of the series is observed precisely in the dispersion, while the distribution of means remains symmetric.

Thus, the proposed procedure allows to draw some conclusions about the behavior of the time series through the analysis of scatterplots and histograms of means and variances.

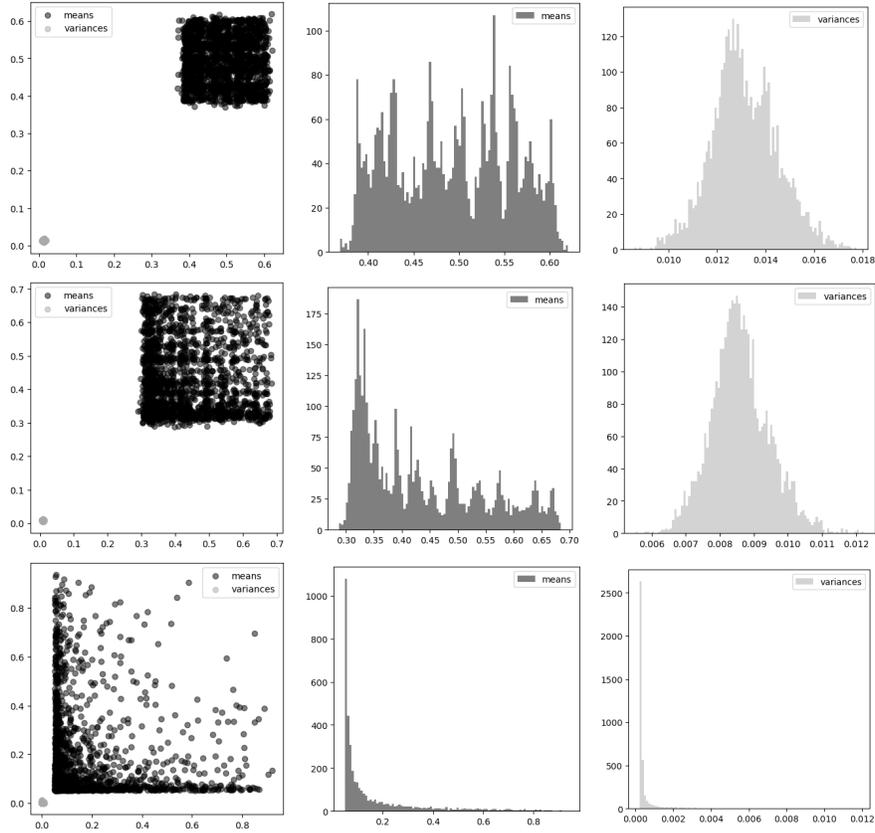


FIG. 7. The result of procedure application on white noise with trend data presented by equations (14) (first line), (15) (second line), (16) (third line). On the left – scatter plot of *Means* and *Vars*; in the middle – histogram for *Means*; on the right – histogram for *Vars*

**5.3. Result of procedure application on the studied data.** Now the proposed procedure will be applied to the studied data on PM2.5 concentration (Figure 10). The data was previously normalized within  $[0; 1]$ .

From the analysis of the obtained results, the following conclusion can be drawn. Histograms of means and variances for the “original data” are asymmetric and have a long tail to the right (see Figure 10, first line). At the same time, the trend present in the data is not expressed polynomially, which can indicate the multiple presence of imbalances in the time series. The same can be said about the “daily original data” (see Figure 10, second line). Note that after logarithm applying the histograms of *Means* and *Vars* become more symmetrical, but still have a tail to the right (see Figure 10, third line). For “daily logarithmic data” it can be negligible for further analysis.

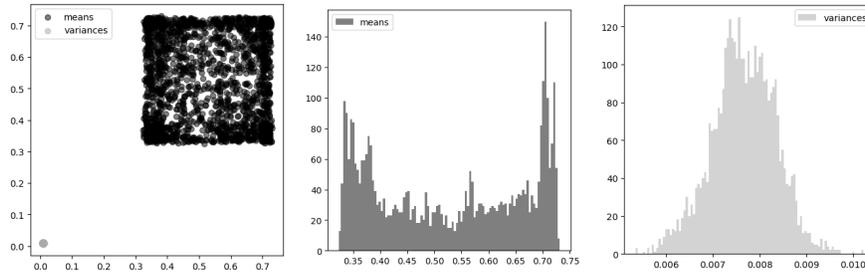


FIG. 8. The result of procedure application on white noise and seasonality data presented by (17). On the left – scatter plot of *Means* and *Vars*; in the middle – histogram for *Means*; on the right – histogram for *Vars*

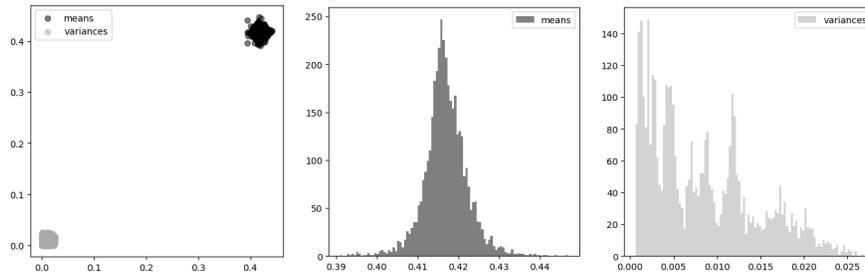


FIG. 9. Result of procedure application on heteroscedastic data. On the left – scatter plot of *Means* and *Vars*; in the middle – histogram for *Means*; on the right – histogram for *Vars*

## 6 Conclusions and discussion

The paper is devoted mostly to a review of methods that can be used to assess the stationarity of a time series. It is shown that there is no single criterion suitable for complex series of real data, and even such a primary analysis procedure can cause difficulties. Here attention was drawn to the correctness of using the most popular tools in solving such problems. It was shown that for time series of real measurements, different statistical tests, which are most used in practice, due to the peculiarities of constructing hypotheses, can give different results. Another equally frequently used tool, ACF and PACF analysis, shows good results for autocorrelated series that do not exhibit heteroscedasticity.

At the same time, stationarity has a fairly simple criterion in its definition – constancy of the mean, variance and autocovariance over time. So, the methods of primary analysis of these components can give us the propositions of time series behavior. Here the periodograms analysis can be referred as an

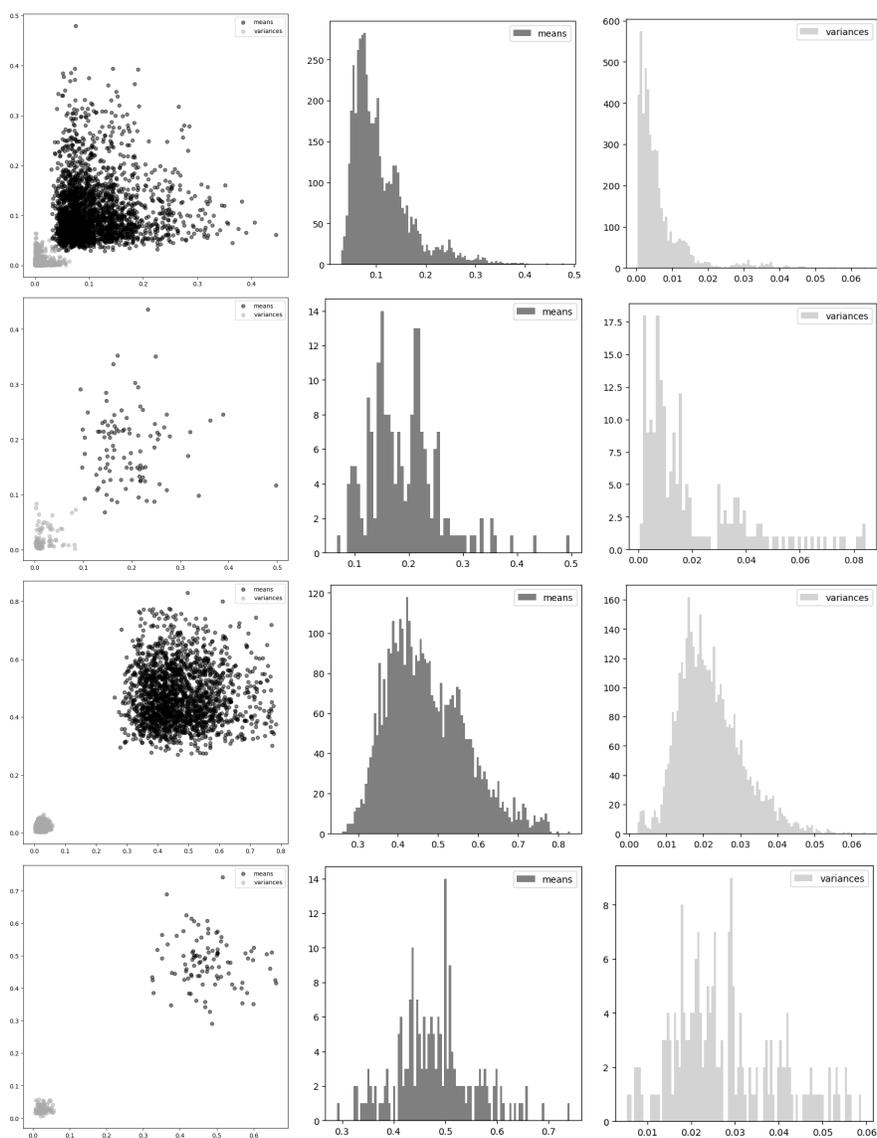


FIG. 10. The result of procedure application on studied data. First line: results for “original data”. Second line: results for “daily original data”. Third line: results for “logarithmic data”. Fourth line: results for “daily logarithmic data”. On the left – scatter plot of *Means* and *Vars*; in the middle – histogram for *Means*; on the right – histogram for *Vars*

instrument of analysis of differences data under research from white noise. There are also tests based on stationary definition, such as the Cochrane procedure (for variance). But this method does not allow detecting and

distinguishing poorly defined trends and trends more complex than linear. Here it is proposed a procedure for testing a series for changing in mean and variance of its subsequents. Synthetic data shows the operation of the procedure. Thus, the presence of a trend in the data violates the symmetry and unimodality of the distribution for a sample of means values; seasonality also leads to multimodality. And the heteroscedasticity of the source data violates the variance histogram. The proposed procedure was applied to measurement data of PM<sub>2.5</sub> concentrations in Krasnoyarsk, that allows to supplement knowledge about the data being studied. Once again, note that the procedure tests only the mean and variance of the time series, while the concept of weak stationarity also includes knowledge of autocovariance. Thus, the procedure cannot be used in isolation from the analysis of the ACF.

## References

- [1] C. Faloutsos, et al. *Forecasting big time series: old and new*, Proceedings of the VLDB Endowment, **11**:12 (2018), 2102–2105.
- [2] Z. Liu, et al. *Forecast methods for time series data: a survey*, Ieee Access, **9** (2021), 91896–91912.
- [3] P. Wang, et al., *A novel hybrid-Garch model based on ARIMA and SVM for PM<sub>2.5</sub> concentrations forecasting*, Atmospheric Pollution Research, **8**:5 (2017), 850–860.
- [4] P.J. Brockwell, R.A. Davis, *Introduction to time series and forecasting*, Springer New York, 2002.
- [5] H. Liu, Z. Duan, C. Chen, *A hybrid framework for forecasting PM<sub>2.5</sub> concentrations using multi-step deterministic and probabilistic strategy*, Air Quality, Atmosphere & Health, **12** (2019), 785–795.
- [6] S. Aminikhanghahi, D.J. Cook, *A survey of methods for time series change point detection*, Knowledge and Information Systems, **51** (2017), 339–367.
- [7] H. Wold, *A Study in the Analysis of Stationary Time Series*, Almqvist, Uppsala, 1954.
- [8] J. Wang, S. Ogawa, *Effects of meteorological conditions on PM<sub>2.5</sub> concentrations in Nagasaki, Japan*, International Journal of Environmental Research and Public Health, **12** (2015), 9089–9101.
- [9] N. Zhao, et al., *Day-of-week and seasonal patterns of PM<sub>2.5</sub> concentrations over the United States: Time-series analyses using the Prophet procedure*, Atmospheric Environment, **192** (2018), 116–127.
- [10] *E-BAM particulate monitor operation manual*, available online: 09.07.2024.
- [11] *Environmental Technology Verification Report*, 2001.
- [12] M. Russell, et al., *Daily, seasonal, and spatial trends in PM<sub>2.5</sub> mass and composition in southeast Texas special issue of aerosol science and technology on findings from the fine particulate matter supersites program*, Aerosol Science and Technology, **38** (2004), 14–26.
- [13] X. Zhao, *Seasonal and diurnal variations of ambient PM<sub>2.5</sub> concentration in urban and rural environments in Beijing*, Atmospheric Environment, **43** (2009), 2893–2900.
- [14] B. Zhai, et.al, *Relevance analysis on the variety characteristics of PM<sub>2.5</sub> concentrations in Beijing, China*, Sustainability, **10** (2018), 3228.
- [15] D. Oztuna, A.H. Elhan, E. Tuccar, *Investigation of four different normality tests in terms of type 1 error rate and power under different distributions*, Turkish Journal of Medical Sciences, **36** (2006), 171–176.

- [16] N.M. Razali, et al., *Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests*, Journal of Statistical Modeling and Analytics, **2** (2011), 21–33.
- [17] S. Keskin, *Comparison of several univariate normality tests regarding type I error rate and power of the test in simulation based small samples*, Journal of Applied Science Research, **2** (2006), 296–300.
- [18] B. Yazici, S. Yolacan, *A comparison of various tests of normality*, Journal of Statistical Computation and Simulation, **77** (2007), 175–183.
- [19] T. Thadewald, H. Buning, *Jarque-Bera test and its competitors for testing normality—a power comparison*, Journal of Applied Statistics, **34** (2007), 87–105.
- [20] M.B. Priestley, *Current developments in time series modelling*, Journal of Econometrics, **37**:1 (1988), 67–86
- [21] J. Durbin, *The Fitting of Time-Series Models*, Review of the International Statistical Institute, **28** (1960), 233–244.
- [22] V. Singh, et al, *Hydrological Processes Modelling and Data Analysis*. Water Science and Technology Library, in: Springer, 2024.
- [23] F. Dama, C. Sinoquet, *Time series analysis and modeling to forecast: A survey*, ArXiv:2104.00164. (2021).
- [24] D.A. Dickey, *Stationarity issues in time series models*, SAS Users Group International, **30** (2015), 192.
- [25] P.C.B. Phillips, P. Perron, *Testing for a unit root in time series regression*, Biometrika, **75** (1988), 335–346.
- [26] D. Kwiatkowski, *Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?*, Journal of Econometrics, **54** (1992), 159–178.
- [27] E. Zivot, D.W.K. Andrews, *Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis*, Journal of Business & Economic Statistics. **20** (2002) 25–44.
- [28] Z. Zhou, *Heteroscedasticity and autocorrelation robust structural change detection*, Journal of the American Statistical Association . **108** (2013) 726–740.
- [29] H. Dette, W. Wu, Z. Zhou, *Change point analysis of second order characteristics in non-stationary time series*, ArXiv:1503.08610 (2015).
- [30] E. Page, *Continuous inspection schemes*, Biometrika, **41** (1954), 100–115.
- [31] S. Lee, J. Ha, O. Na, S. Na, *The cusum test for parameter change in time series models*, Scandinavian Journal of Statistics, **30** (2003), 781–796.
- [32] K. Kim, et al., *Unsupervised change point detection and trend prediction for financial time-series using a new cusum-based approach*, IEEE Access, **10** (2022), 34690–34705.
- [33] O.A. Vanli, R. Giroux, *An adaptive cumulative sum method for monitoring integer-valued time-series data*, Quality Engineering, **34** (2022), 230–247.
- [34] L. Xie, G.V. Moustakides, Y. Xie, *Window-limited CUSUM for sequential change detection*, IEEE Transactions on Information Theory, **69**:9 (2023), 5990–6005.
- [35] K. Wenger, C. Leschinski, P. Sibbertsen, *Change-in-mean tests in long-memory time series: a review of recent developments*, AStA Advances in Statistical Analysis, **103** (2019), 237–256.
- [36] Y. Huang, et al., *Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis*, IEEE Systems Journal, **10** (2014), 532–543.
- [37] S. Lee, S. Lee, M. Moon, *Hybrid change point detection for time series via support vector regression and CUSUM method*, Applied Soft Computing, **89** (2020), 106101.
- [38] J.K. Afriyie, *Evaluating the performance of unit root tests in single time series processes*, Mathematics and Statistics, **8** (2020), 656–664.
- [39] *Augument Dickey-Fuller test tables*, available online: 09.07.2024.
- [40] J.J. Dolado, et al., *Cointegration and unit roots*, Journal of Economic Surveys, **4** (1990), 249–273.

- [41] V.P. Nosko, *Econometrics: Introduction to regression analysis of time series*, Moscow, 2002.
- [42] C.R. Rao, *Score test: historical review and recent developments*, *Advances in Ranking and Selection, Multiple Comparisons, and Reliability: Methodology and Applications*, **2** (2005), 3–20.
- [43] A. Charles, O. Darne, *Variance-ratio tests of random walk: an overview*, *Journal of Economic Surveys*, **23** (2009), 503–527.
- [44] M. Priestley, T. Rao, *A test for non-stationarity of time-series*, *Journal of the Royal Statistical Society Series B*, **31** (1969), 140–149.
- [45] E. Paparoditis, *Validating stationarity assumptions in time series analysis by rolling local periodograms*, *Journal of the American Statistical Association*, **105** (2010), 839–851.
- [46] S.K. Schmidt, et al., *An asymptotic test for constancy of the variance under short-range dependence*, *The Annals of Statistics*, **49** (2021), 3460–3481.
- [47] R. Puchstein, P. Preuß, *Testing for stationarity in multivariate locally stationary processes*, *Journal of Time Series Analysis*, **37** (2016), 3–29.
- [48] S. Basu, S. Subba Rao, *Graphical models for nonstationary time series*, *The Annals of Statistics*, **51** (2013), 1453–1483.
- [49] V.M. Panaretos, S. Tavakoli, *Fourier analysis of stationary time series in function space*, *The Annals of Statistics*, **41:2** (2013), 568–603.
- [50] S. Bandyopadhyay, C. Jentsch, S. Subba Rao, *A Spectral Domain Test for Stationarity of Spatio-Temporal Data*, *Journal of Time Series Analysis*, **38:2** (2017), 326–351.
- [51] M. Basta, *Wavelet-Based Test for Time Series Non-Stationarity*, *Statistika: Statistics & Economy Journal*, **95** (2015)
- [52] A. Diez-Noguera, *Methods for serial analysis of long time series in the study of biological rhythms*, *Journal of Circadian Rhythms*, **11** (2015), 1–19.

VIKTORIYA SERGEEVNA PETRAKOVA  
 INSTITUTE OF COMPUTATIONAL MODELLING,  
 ST. ACADEMGORODOK, 50/44,  
 660036, KRASNOYARSK, RUSSIA,  
 SOBOLEV INSTITUTE OF MATHEMATICS,  
 PR. KOPTYUGA, 4,  
 630090, NOVOSIBIRSK, RUSSIA  
 Email address: [vika-svetlakova@yandex.ru](mailto:vika-svetlakova@yandex.ru)