Mathematical model of the energy resource consumption process in the form of a random process with piecewise homogeneous components

Leonid Scherbak[a], Iaroslav Lytvynenko[b], Serhii Kharchenko[a], Oleg Nazarevych[b], Volodymyr Hotovych[b]

[a]*Institute of General Energy of the National Academy of Sciences of Ukraine, 172, Antonovycha Street, Kyiv, 03150, Ukraine*
[b]*Ternopil Ivan Puluj National Technical University, 56, Ruska Street, Ternopil, 46001, Ukraine*

Abstract

The paper is devoted to the development of a generalized mathematical model for the description of random processes of energy resource consumption (electricity, water, natural gas consumption). The model is presented in the form of a random process, which consists of the sum of components: a trend (non-oscillatory component), several piecewise homogeneous components of an oscillatory nature, and a stochastic component. To build this model, the results of statistical processing of implementations of the energy resource consumption processes of different scales and the nature of functioning of objects were used. At the same time, the decomposition method of Singular Spectrum Analysis (SSA) was used to break down realizations of random processes into components, and the Pelt method was used to select homogeneity time intervals within the general observation interval. Polynomials of the third order (trend component) and sinusoids (oscillating components) were used to approximate the components of the energy resource consumption processes at homogeneity intervals. The results of statistical processing show the adequacy of this model, and the selected process components are physically justified.

Keywords

Energy resource consumption process, random process, SSA-Caterpillar method, Pelt method, change points, trend, piecewise homogeneous components, stochastic component.

1. Introduction

The need for economy and optimal consumption of energy resources (electricity, water, natural gas consumption) determine the relevance of the tasks of analysis, control, diagnosis and monitoring of the relevant energy resource consumption processes. Today, such tasks are solved with the help of specialized automated systems. The basis for the functioning of these systems is algorithmic software, in particular, mathematical models of the energy resource consumption processes. At the same time, the correctness and efficiency of the functioning of automated systems depend on the adequacy and accuracy of mathematical models.

This paper is devoted to the development of a mathematical model of the energy resource consumption processes in the form of a random process with piecewise homogeneous components, which is a generalization of known models of resource consumption processes developed on the basis of the random processes theory.

ITTAP'2022: 2nd International Workshop on Information Technologies: Theoretical and Applied Problems, November 22–24, 2022, Ternopil, Ukraine

EMAIL: prof_Scherbak@ukr.net (L. Scherbak); iaroslav.lytvynenko@gmail.com (I. Lytvynenko); nanoavia@ukr.net (S. Kharchenko); taltek.te@gmail.com (O. Nazarevych), gotovych@gmail.com (V. Hotovych)
ORCID: 0000-0002-1536-4806 (L. Scherbak); 0000-0001-7311-4103 (I. Lytvynenko); 0000-0001-9808-7607 (S. Kharchenko); 0000-0002-8883-6157 (O. Nazarevych), 0000-0003-2143-6818 (V. Hotovych)

2. Analysis of recent research

A review of the literature [7-14] showed that the processes of energy resource consumption (electricity, water, natural gas consumption) by objects of different scales (house, settlement, district, region, etc.) have the following common features:

1) They are formed as a result of the combined action of many individual consumers of different capacities, the moments of turning on and off (resource consumption sessions), as well as the duration of such sessions, are random (stochastic nature of the process);
2) They are periodic, because the activity of a person and society, in general, is periodic, with different periods: hour, day, month, year, etc. (periodic nature of the process).

Therefore, it is quite justified to consider such processes as stochastic-periodic and, accordingly, to apply the theory of random processes for their modeling, as well as the apparatus of mathematical statistics for the practical study of their realizations.

In general, there are many scientific and technical publications dedicated to the construction and application of mathematical models based on random processes. For example, in [1] the use of the model of a periodic random process for a wide range of problems of stochastic process research is given. In [2], the model of a random process is presented as the sum of a stationary random process and the value of a mathematical expectation that changes under the influence of an indicator function. Also, there are well-known approaches associated with the use of periodically correlated random processes [3, 4], Markov processes [5, 6], etc.

In this paper, the set of random processes of energy consumption is considered as a subset of a large and numerous set of random processes.

Among the well-known publications, a significant number of publications are devoted to the problem of modeling energy consumption processes. In particular: in [7, 8], the mechanism of formation of the energy resource consumption process (specifically, natural gas consumption) as a result of the action of many consumers of different power and with random consumption sessions in time is substantiated, and it is also proposed to use a conditional linear periodic random process as a model; in [9] a model is proposed as the sum of a deterministic trend and a piecewise stochastic-periodic process, which takes into account the change in the mode of operation of the consumer network over a long observation interval; [10] proposed a model of the power consumption process in the form of a piecewise homogeneous periodic random process, which includes a set of homogeneous components (periodic with a period of 24 hours) of random processes as a vector periodic process; in [11] a mathematical model of the electric load of organizations at an interval of one day was developed and substantiated in the form of a piecewise homogeneous random process with uncorrelated values (white noise in the broad sense); in [12] a model of electricity consumption is proposed as a multi-component random process with moments of disorder (a sudden change in process dynamics), which consists of a trend, a piecewise-periodic random process, and a piecewise-stationary random process; in [13] a cyclostationary conditional linear random process is proposed as a model of the random process of water consumption; in [14], the model of the gas consumption process is presented in the form of the sum of the trend, cyclic random process and stochastic residual, which also takes into account the segment cyclical structure of the process.

In these publications, partial cases of the energy resource consumption processes were considered as the subsets of a large set of stochastically periodic random processes. This paper proposes a comprehensive generalized model that combines both deterministic and stochastic components of the energy resource consumption process.

3. Main part

In general, the research process of random processes of energy resource consumption can be described as shown in Fig. 1.

Here, under the formation of an energy resource, we mean a certain network (topology) of spatially separated consumers of such resources as electricity, water, and natural gas, and this network itself is called a stochastic spatio-temporal environment of consumption (city, village, district, region, country, organization, etc.).

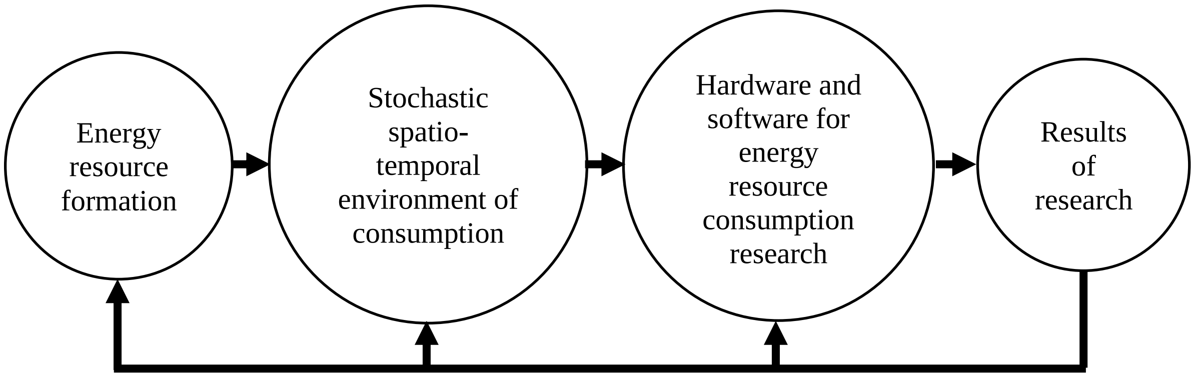

Figure 1: The general scheme of the formation of the process of energy resources consumption and their research

Taking into account the nature of energy resource consumption processes, and the results of considered studies [6 – 14], in general, the model of the energy resource consumption process can be represented by the expression (1):

$$\xi(\omega,t) = A(t) + B(t) + Z(\omega,t), \tag{1}$$

where $A(t)$ is a deterministic component (trend), $B(t)$ is a deterministic, periodic component with a certain period of oscillation, $Z(\omega,t)$ is a stochastic component.

The proposed mathematical model is an additive sum of components. As a rule, the trend component has the character of a smooth curve without sharp changes (oscillations). As a component $B(t)$, there can be several components with different periods and amplitudes of oscillations (the specific number of components is determined in the process of statistical processing based on the experience of the researcher). In this case, the oscillating component $B(t)$ will be the sum of these components and model (1) will take the following form:

$$\xi(\omega,t) = A(t) + \sum_{i=1}^{n} B_i(t) + Z(\omega,t). \tag{2}$$

The stochastic component of the model $Z(\omega,t)$ is a set of random factors that affect the object of research. In the literature dealing with the theoretical foundations of the SSA method [15], this component is often called the stochastic residual.

The obtained results of statistical processing, which make it possible to decompose the implementation of a random process into components and find change points (time moments of rapid or even sudden changes in the dynamics of the process), make it possible to specify the general model and present it in the form:

$$\begin{gathered}\xi(\omega,t) = \sum_{j=1}^{n} A_j(t) I(t,\Delta t_j) + \sum_{j=1}^{n} B_j(t) I(t,\Delta t_j) + \sum_{j=1}^{n} Z_j(\omega,t) I(t,\Delta t_j),\ \ \omega\in\Omega,\ \ t\in[0,T], \\ \bigcup_{j=1}^{n}\Delta t_j = [0,T],\ \ \Delta t_j = \begin{cases}\left[t_{j-1},t_j\right), & j=\overline{1,n-1} \\ \left[t_{j-1},t_j\right], & j=n\end{cases},\ \ \Delta t_j \cap \Delta t_i = \varnothing\ \ for\ i\neq j,\ \ \Delta t_j \neq \varnothing.\end{gathered} \tag{3}$$

Here, the interval of observation $t\in\left[0,T\right]$ is divided by change points (distortion points) t_j, $j=\overline{1,n}$ into n intervals. The belonging of a specific fragment of a process component to the j-th interval is given by the indicator function:

$$I(t,\Delta t_j) = \begin{cases}1, & t\in\Delta t_j \\ 0, & t\notin\Delta t_j\end{cases},\ \ j=\overline{1,n}. \tag{4}$$

At each of the n intervals, the trend and fluctuation components are described by some deterministic functions. At the same time, the trend and fluctuating components have a deterministic character on several observation subintervals, and together these observation intervals form the overall observation interval $t \in [0, T]$ of the implementation of a specific energy resource consumption process.

In essence, each separate subinterval is a time interval within which the topology of energy resource consumers functions in a somewhat stable mode and the cumulative effect of various factors on this topology (weather factors, length of day and night, economic factors, etc.) remains unchanged. A change in these factors leads to the transition of the network of energy resource consumers to a different mode of consumption.

To confirm the adequacy of the model (3), the work carried out statistical processing of several implementations of the energy resource consumption processes. This statistical processing consists of two stages:

1. Decomposition of the process implementation into individual components.
2. Change points detection and using them to divide process components into fragments that correspond to separate intervals of process homogeneity.

In this study, singular spectrum analysis (SSA), also known as the Caterpillar method [15, 16], was used to decompose realizations of energy consumption processes. Unlike other known methods (autoregression and integrated moving average (ARIMA) [17], group method of data handling (GMDH) [18], principal component method (PCA) [19], empirical mode decomposition method (EMD) [20], wavelet analysis [21], etc.) the SSA method is characterized by relative simplicity (it mainly consists in working with numerical matrices) and interactivity. The interactivity of the SSA method is that the researcher can choose the degree of detail of decomposition (the number of components that will be obtained as a result of decomposition) in an interactive mode. Some prior experience in using this method is required at this stage of statistical processing. Some recommendations for the practical application of the method are given in [15].

The SSA method takes as input the implementation of a random process obtained as a result of a measurement experiment with a certain accumulation step, as a time-ordered sequence of counts (time series). The purpose of the method is to decompose the time series into a set of components (trend, periodic oscillatory components, stochastic noise component), which in total give the original series. That is, the researcher a priori assumes the possibility of such decomposition as a consequence of the nature of some real phenomenon, which is represented by the registered implementation of a random process. At the same time, the reliability of such components can be explained by the fact that:

- a trend is a smooth curve of a non-oscillating type, which shows the general nature of the change in the dynamics of the process over time and is considered a certain constant component of the process;
- fluctuating components are present in the structure of the process due to the periodicity of the real phenomenon under investigation (a part of human society limited in space and time);
- the presence of a stochastic component in the structure of the process is caused by a large number of stochastic factors forming the process, as a rule, of low intensity. In essence, this component is the sum of these factors.

In the second stage of statistical processing, various statistical methods can be used to search for change points. They can be, for example, binary segmentation [22], the method of neighbouring segments [23], and the Brodsky-Darkhovsky method [24]. A more detailed analysis of various methods of finding points of disturbance (classification and comparison of the efficiency of their application in various fields: medicine, climate change, analysis of human activity, etc.) can be found in [25 – 27].

In this work, the Pelt method [28, 29] was used to find change points. This is an a posteriori method that belongs to the family of methods for finding the optimum of the likelihood function and is characterized by high computational efficiency. This method was chosen because of its interactivity, as it allows the researcher to specify a different number of change points to be found, as well as to specify the statistical characteristic behind which the distortion occurs (in particular, the mean and variance).

The change points detection process was based on a sudden change in the variance of the stochastic component obtained as a result of decomposition at the previous stage. After that, the points of distortion were generalized to other components of the energy resource consumption process. That is, it was considered that other components have the same distortion points.

The paper researches several implementations of energy resource consumption processes that characterize objects of different scales. In particular, statistical processing of data on electricity consumption (building № 1 of the Ivan Puluj National Technical University in 2016), natural gas consumption (Ternopil in 2009) and water consumption (100 apartment buildings in Ternopil in 2009) were carried out. Some of the obtained results are presented in fig. 2 – 7. In particular, fig. 2 shows the results of the decomposition of the time series of the electric energy consumption process. For clarity in fig. 2, a) graphs of the process itself and its trend component are combined.

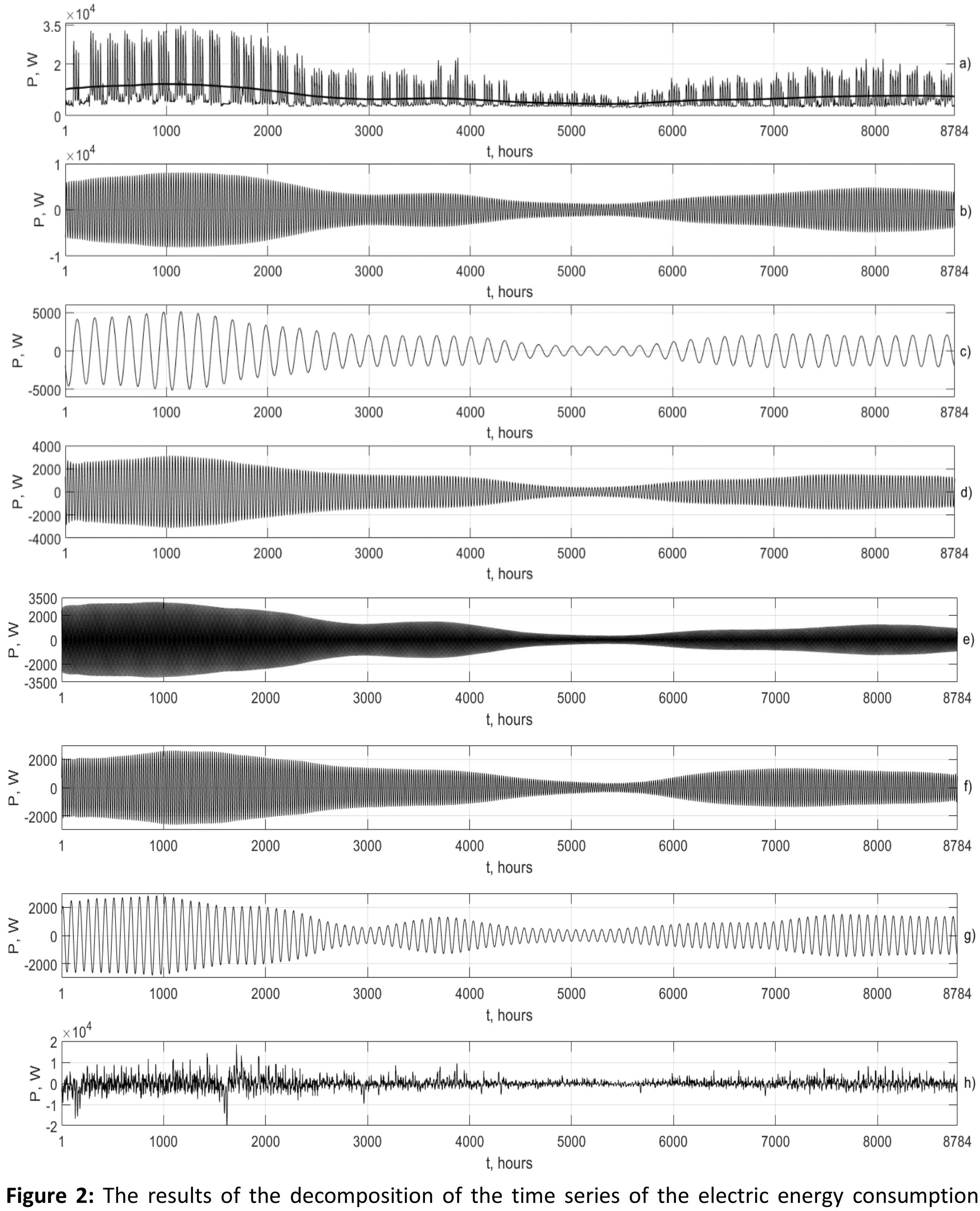

Figure 2: The results of the decomposition of the time series of the electric energy consumption process using the SSA method: a) implementation and trend (component №1 of decomposition results); b)-g) components №2-7 of the decomposition of an oscillatory nature; h) component №8 of decomposition results (stochastic “noise”)

Fig. 3 shows the results of dividing the power consumption process and its component into homogeneity intervals, which are represented by vertical lines on the graphs. Analogous results of statistical processing of implementations of natural gas consumption and water consumption processes are shown in Figs. 4 and 5, respectively.

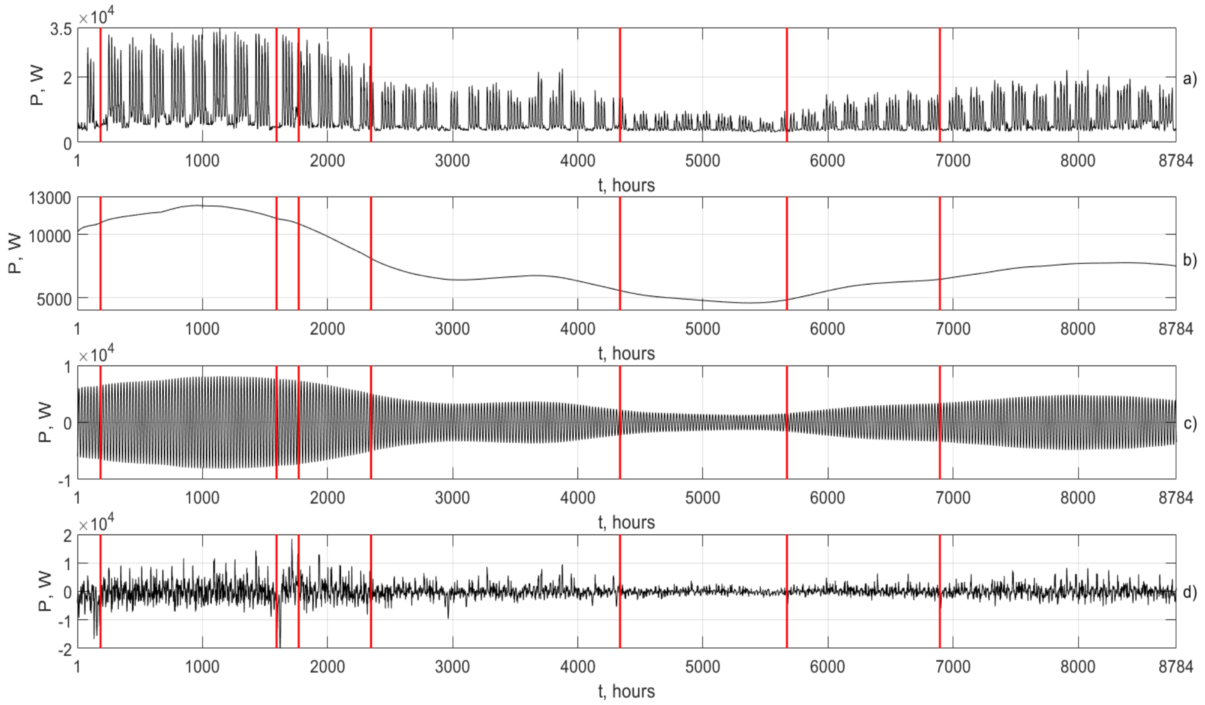

Figure 3: The division into homogeneity intervals: a) implementation of the electric energy consumption process; b) trend component; c) component with an oscillation period of one day; d) stochastic component

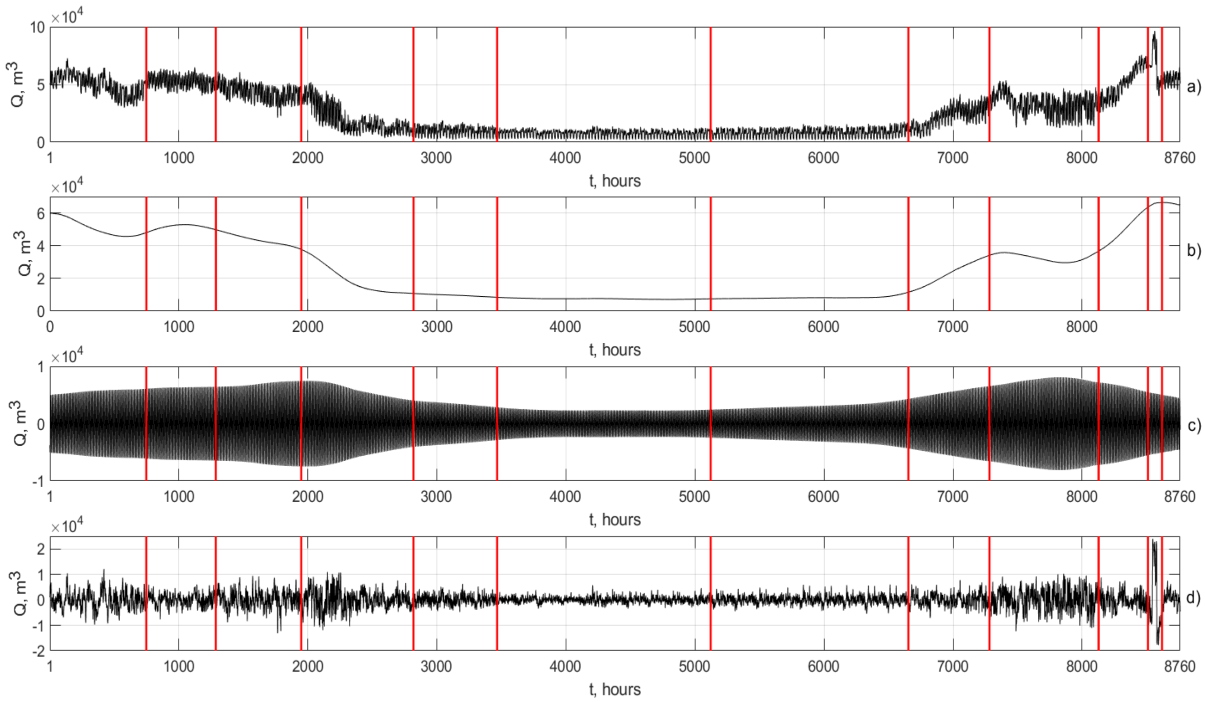

Figure 4: The division into homogeneity intervals: a) implementation of the natural gas consumption process; b) trend component; c) component with an oscillation period of one day; d) stochastic component

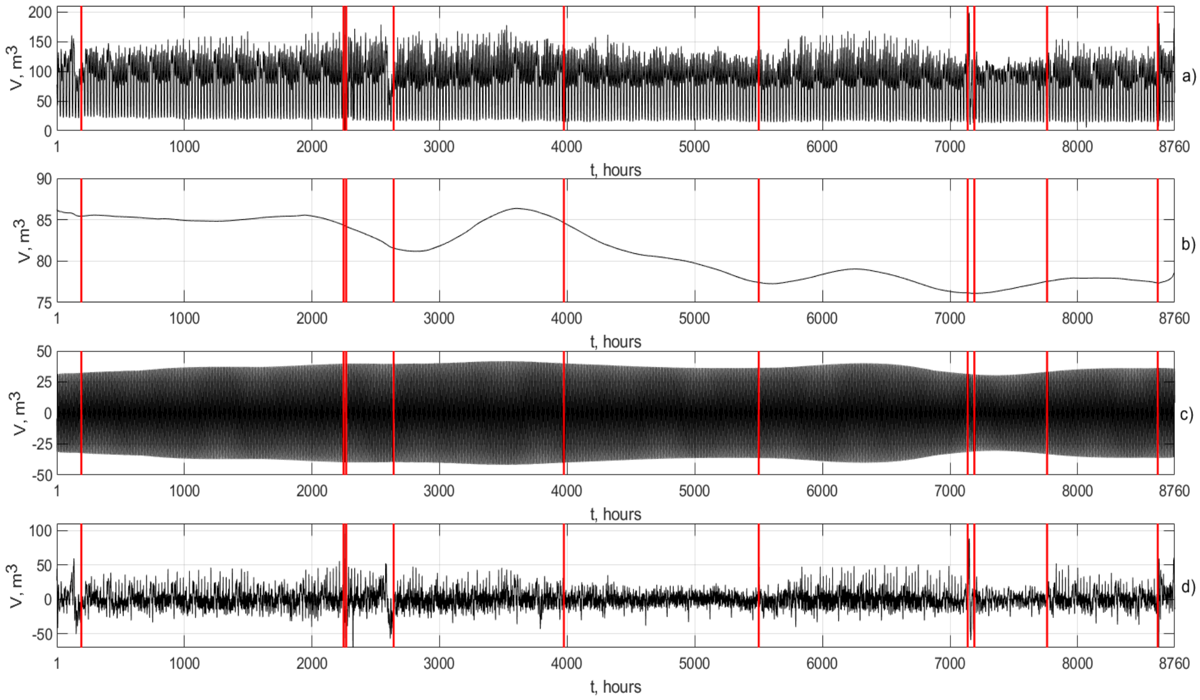

Figure 5: The division into homogeneity intervals: a) implementation of the water consumption process; b) trend component; c) component with an oscillation period of one day; d) stochastic component

Figs. 6 and 7 show some of the results of components approximation of the energy resource consumption processes. The approximation of the trend component was carried out using the polyfit and polyval functions (the method of least squares) of the Matlab software environment, while the Curve Fitting Toolbox package of the Matlab environment was used to approximate the fluctuating components (by finding the coefficients of the Fourier series of the approximating curve).

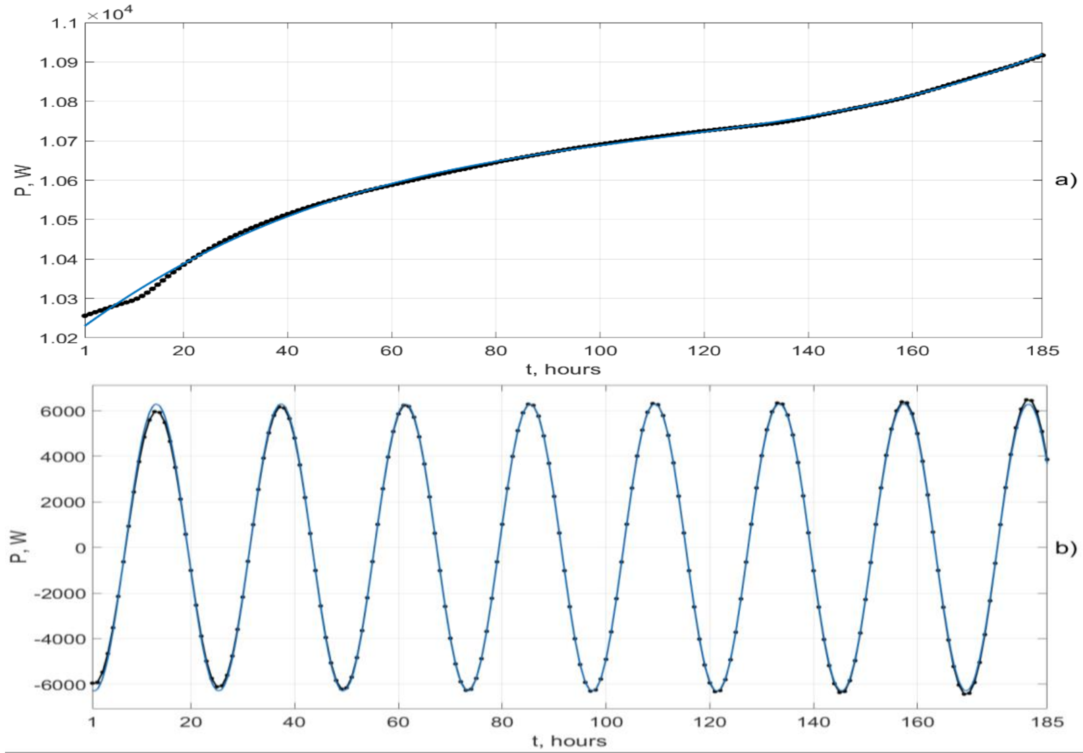

Figure 6: Approximation on the first interval of homogeneity of the components of the electric energy consumption process: a) trend component; b) component with an oscillation period of one day

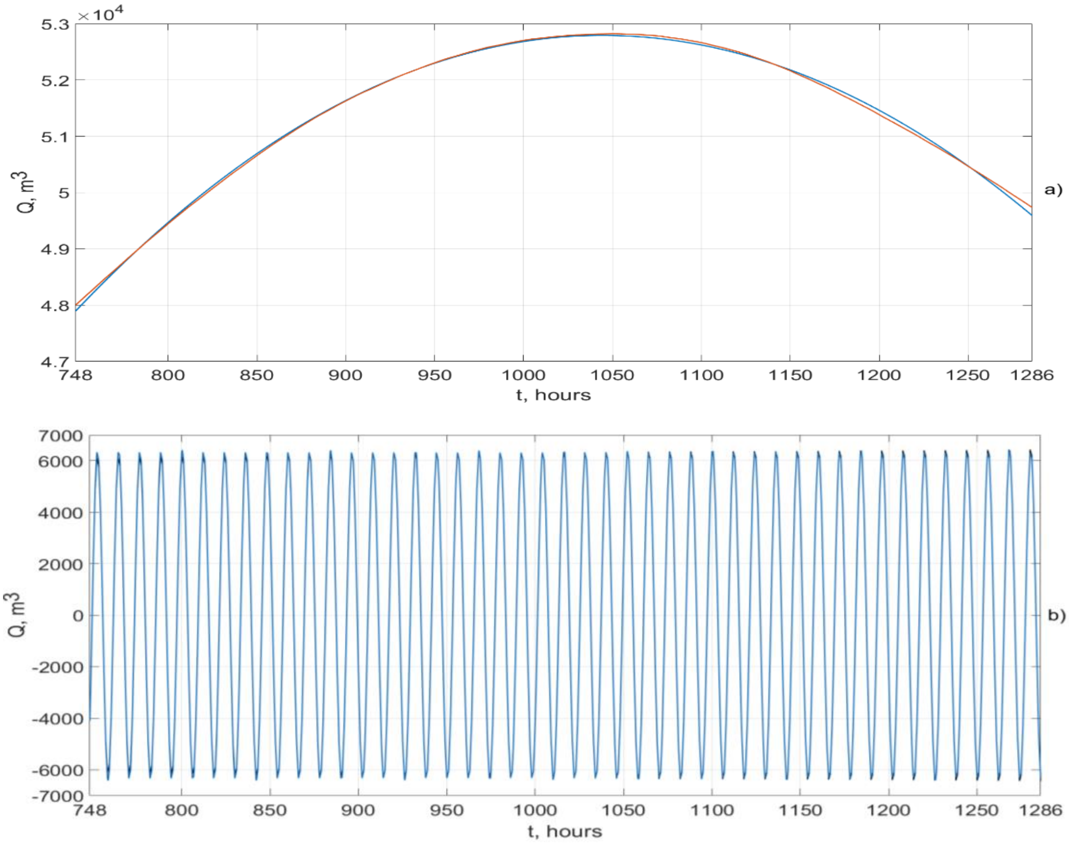

Figure 7: Approximation on the second interval of homogeneity of the components of the natural gas consumption process: a) trend component; b) component with an oscillation period of one day

In particular, for the electric energy consumption process on the first homogeneity interval, the trend component is approximated by a polynomial of the third degree $a_1(t) = 21612t^3 - 0,0722t^2 + 9,7314t + 1,0221,\ t \in [1,185]$ hours, and the oscillating component with an oscillation period of one day is approximated by a sinusoid $b_1(t) = -0,4316 - 5898\cos(0.2617t) - 2202\sin(0.2617t),\ t \in [1,185]$ hours.

For the water consumption process on the second homogeneity interval, the trend component is approximated by a polynomial of the third degree $a_2(t) = 1,7259 * 10^{-6} t^3 - 0,0607t^2 + 121,0737t - 9447,\ t \in [748,1286]$ hours, and the oscillating component with an oscillation period of one day is approximated by a sinusoid $b_2(t) = -0,5403 - 2596\cos(0.5235t) - 5873\sin(0.5235t),\ t \in [748,1286]$ hours.

A similar approximation can be carried out for all fluctuating components and the trend component of the decomposition of the energy resource consumption process implementation on all homogeneity intervals.

4. Discussion of obtained results

The physical interpretation of change points consists in the transition of the network of energy consumers from one established mode of consumption to another, for example, in the transition between seasons (spring, summer, etc.), which affect the consumer network. In this case, homogeneity intervals are time intervals within which the functioning of the network of energy resource consumers is unchanged (it is within a certain mode).

On the other hand, the components that we get as a result of the decomposition of the implementation of the energy resource consumption process are also interesting. The trend shows the general dynamics of the development of the process and is some generalized, integral sum of the functioning of all consumers who are included in the network of energy resource consumers. Each of the oscillating components has a constant period of oscillation and its presence as part of a random process is explained by the periodicity of human activity. For example, in the composition of all implementations of the energy resource consumption processes developed in this work, there is a component with an oscillation period of one day (component № 2 of decomposition), i.e., 24 counts for an accumulation interval of one hour. The stochastic component is the sum of all random factors that affect the users of the energy resource, and therefore it was used to search for change points during the statistical processing carried out in the paper.

Naturally, each network of energy resource consumers has its own individual characteristics. Therefore, the task of concretizing the model proposed in this paper (decomposition with a certain level of detail (the number of detected components), change points detection, approximation of deterministic components, etc.) is one of the tasks that are solved during the study of the corresponding random process as a characteristic of a specific network of energy consumers.

5. Conclusions

The paper proposes a model of the energy resource consumption process, which generalizes known stochastic models of energy resource consumption processes and combines both deterministic and additive components. This is consistent with the stochastically periodic nature of processes of this type. In order to confirm the adequacy of the proposed model, statistical processing of implementations of resource consumption processes was carried out, which characterize the topologies of consumers with different scales and modes of operation. The SSA and Pelt statistical methods were used for the corresponding processing, although it is also possible to use other, similar methods. The highlighted components of energy resource consumption processes are physically justified.

Prospects for further research include clarifying the nature of the stochastic component of the model (for example, the distribution on different intervals of homogeneity), which depend on the specific implementation of the energy resource consumption process being studied, as well as simulation modeling of both individual components of the model and the implementation of the energy resource consumption process as a sum of components as a whole.

6. References

[1] W. A. Gardner, A. Napolitano, L. Paura, Cyclostationarity: Half a century of research, Signal Processing, № 86 (2006), pp. 639–697. https://doi.org/10.1016/j.sigpro.2005.06.016.

[2] E. Rafajłowicz, M. Pawlak, A. Steland, "Nonparametric sequential change-point detection by a vertically trimmed box method, " IEEE Transactions on Information Theory, vol.56, pp.3621 - 3634, July 2010.

[3] I. Javorskyj, R. Yuzefovych, I. Kravets, I. Matsko, Methods of periodically correlated random processes and their generalizations. In: Cyclostationarity: Theory and Methods, 73–93 (2014).

[4] H. L. Hurd, A. G. Miamee, Periodically Correlated Random Sequences: Spectral Theory and Practice, Wiley, New York, 2006.

[5] A. R. Nematollahi, A. R. Soltani, Discrete time periodically correlated Markov processes, Probability and Mathematical Statistics. 2000, No. 20 (1), pp. 127-140.

[6] M. Pryimak, O. Matsiuk, O. Maevskyi, S. Proshyn, Models and methods of investigation for markov type queuing systems under conditions of stochastic periodicity and its application in energetics. Technical Electrodynamicsthis, 2014 (2), pp. 11–16.

[7] B. Marchenko, N. Mulyk, “Umovnyi liniinyi vypadkovyi protses yak matematychna model hazospozhyvannia” XIX Scientific conference of the Ternopil state technical university, 2005, Ternopil, p. 64. [in Ukrainian].

[8] B. Marchenko, N. Mulyk, M. Fryz, Obgruntuvannia matematychnoi modeli hazonavantazhen. Scientific Journal of the Ternopil state technical university 2005, No 2, pp. 138-143. [in Ukrainian].

[9] O. B. Nazarevych, Information technology for monitoring of city gas consumption (Information technologies), thesis for a Candidate degree, Ternopil Ivan Puluj National Technical University, Ternopil, Ukraine, 2015. [in Ukrainian].
[10] T. L. Scherbak, Information technology of diagnostics of dynamics of processes of electro-consumption in the regular and nonpermanent modes (Information technologies), thesis for a Candidate degree, National aviation university, Kyiv, Ukraine, 2010. [in Ukrainian].
[11] S. V. Marcenko, Mathematical modeling and measurement data processing methods in problems of monitoring electricity load (Mathematical modeling and computational methods), thesis for a Candidate degree, Ternopil Ivan Puluj National Technical University, Ternopil, Ukraine, 2011. [in Ukrainian].
[12] V. A. Gotovych, Mathematical modeling and statistical estimation of characteristics of the regular mode of organizations` power consumption (Mathematical modeling and computational methods), thesis for a Candidate degree, Ternopil Ivan Puluj National Technical University, Ternopil, Ukraine, 2019. [in Ukrainian].
[13] T. Mykhailovych and M. Fryz, "Model and Information Technology for Hourly Water Consumption Interval Forecasting," 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 2020, pp. 341-345, doi: 10.1109/TCSET49122.2020.235452.
[14] I. Lytvynenko, S. Lupenko, O. Nazarevych, G. Shymchuk and V. Hotovych, "Mathematical model of gas consumption process in the form of cyclic random process," 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT), 2021, pp. 232-235, doi: 10.1109/CSIT52700.2021.9648621.
[15] N. E. Golyandina, V. V. Nekrutkin, A. A. Zhigljavsky, Analysis of Time Series Structure: SSA and Related Techniques. Boca Raton: Chapman&Hall/CRC. 2000. 305 p.
[16] Groth, A., and M. Ghil, 2015: Monte Carlo Singular Spectrum Analysis (SSA) revisited: Detecting oscillator clusters in multivariate datasets, Journal of Climate, 28, pp. 7873-7893.
[17] Box, G. E. P & Jenkins, G. M (1976). Time Series Analysis: Forecasting and Control, Revised Edition, Holden Day.
[18] Ivakhnenko A. G. (1968) The group method of data of handling; a rival of the method of stochastic approximation. Soviet Auto Contr 13 (1): 43–55.
[19] Holland, Steven M. "Principal components analysis (PCA)." Department of Geology, University of Georgia, Athens, GA (2008): 30602-2501.
[20] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, Q. Shin, H. H. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The Empirical Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis‟, Proc. Royal Soc. London A, vol. 454, pp. 903-995, 1998.
[21] Lau, K.-M., and H.-Y. Weng, 1995: Climate signal detection using wavelet transform: How to make a time series sing. Bull. Amer. Meteor. Soc., 76, pp. 2391–2402.
[22] Scott, A. J., and Knott, M. (1974), "A Cluster Analysis Method for Grouping Means in the Analysis of Variance," Biometrics, 30, pp. 507–512.
[23] Auger, I. E., and Lawrence, C. E. (1989), "Algorithms for the Optimal Identification of Segment Neighborhoods," Bulletin of Mathematical Biology, 51, pp. 39–54.
[24] Brodsky B. E., Darkhovsky B. S. Nonparametric Methods in ChangePoint Problems. Dordrecht: Kluver Academic Publishings, 1993. 210 p.
[25] C. Truong, L. Oudre, N. Vayatis, A selective review of offline change point detection methods. Signal Processing. 167:107299, 2020. doi: https://doi.org/10.1016/j.sigpro.2019.107299.
[26] Aminikhanghanhi, S. & Cook, D. J., A survey of methods for time series change point detection. Knowl. Inf. Syst. 51, 339–367 (2017).
[27] C. Truong, L. Oudre, and N. Vayatis, A review of change point detection methods. arXiv preprint arXiv:1801.00718, 2018.
[28] Killick Rebecca, Fearnhead Paul, and Idris A. Eckley. Optimal detection of changepoints with a linear computational cost. Journal of the American Statistical Association. Vol. 107, No. 500, 2012, pp. 1590–1598. doi: http://dx.doi.org/10.1080/01621459.2012.737745.
[29] B. Jackson, J. D. Sargle, D. Barnes, S. Arabhi, A. Alt, P. Gioumousis, E. Gwin, P. Sangtrakulcharoen, L. Tan and T. T. Tsai, (2005), "An Algorithm for Optimal Partitioning of Data on an Interval", IEEE Signal Processing Letters, 12, pp. 105–108.