Review Article

AN IMPROVED PRE-FORECASTING ANALYSIS OF ELECTRICAL LOADS OF PUMPING STATION

Anatolii V. Voloshko[*,1], Yaroslav S. Bederak[2], Oleksandr A. Kozlovskyi[3]

[1] National Technical University of Ukraine «Ihor Sikorskyi Kyiv Polytechnic Institute»,
115, Borschahivs'kaStr, Kiev, 03056, Ukraine.

[2] PJSC «AZOT», 72, Pershotravneva Str., Cherkassy, 18014, Ukraine.

[3] Central Ukrainian National Technical University, 8, Universytetskyi Avenue, Kropyvnytskyi, 25006, Ukraine.

Abstract

Relevance of research. In order to reduce energy losses, an accurate and timely forecast of the amount of consumed electricity is necessary. Accurate forecasting of electrical loads of industrial enterprises and their divisions (productions, workshops, departments etc.) allows planning of normal operating conditions, concluding contracts for the electricity supply with the electricity supply company under more favorable conditions, and improving the electricity quality, which ultimately affects the final cost of the products produced by an enterprise. So far, more than 150 forecasting methods of electrical loads have been developed. Usually, the most convenient one is selected based on the forecaster experience by creating and analyzing several forecasting models in order to identify the best. Therefore, in order to simplify the forecasting procedure, it is necessary to develop the methodology for forecasting analysis. This methodology should enable canceling forecasting algorithms that will create lower quality forecasts. **The main objective** is to develop the methodology for making a forecasting analysis of power consumption on the example of a pumping station of an enterprise with a continuous cycle of work to increase the efficiency of energy consumption and implementation of energy-saving measures. **Objects of research**: the process of forecasting electrical loads of a pumping station of the enterprise with a continuous cycle of work. **Methods of research**: fundamental principles of the theory of electrical engineering, statistical methods for power consumption forecasting, the method for detecting the trend of radio signals, and fractal analysis of time series. **Research results.** The methodology for forecasting analysis of power consumption, which makes it possible to apply the most appropriate methods to forecast the operational power consumption, is developed. For the first time, the radio signal trend detection method is applied to identify the trend of electrical loads. The variation ranges of the fractal parameters of time series of electrical loads are established depending on the variation coefficient of the time series for different periods of time. The Brown method of exponential smoothing that is used to forecast the electrical loads, in the case of identifying the smoothing constant α is in the beyond set ($1 < \alpha < 2$), is further improved. The regularity of changes in the fractal parameters of time series of power consumption of a pumping station with an increase in the time series duration and their field of application are explained.

Keywords: Electrical load, time series, pre-forecasting analysis, fractal analysis, statistical analysis.

1. Introduction

In the conditions of fierce competition between manufacturers of the same type products on the world markets, energy-saving and rational use of electric energy by industrial enterprises become one of the most significant tasks and priority direction of the energy strategies of states. The reduction in energy losses is not possible without an accurate and timely forecast of the amount of electricity to be consumed. Forecasting electrical loads provide the

source of information necessary to plan normal operation conditions, correlate more favorable power consumption conditions with the power system, and ensure the quality of electricity, which ultimately affects the final cost of products produced by the enterprises [1].

At present, there is a large number of methods for electrical loads (EL) forecast. The main aim is solving the control problems and choosing optimal operation conditions of electric power systems. However, the issues of forecasting power consumption by industrial enterprises and their divisions are not paid much attention. In [2], it was noted that all forecasting methods could be divided into classical methods (analytical construction of a mathematical model)), statistical methods (regression, autoregression without and with a moving average (ARMA model) [3], with exponential smoothing [4]), and intelligent methods (expert systems, artificial neural networks, genetic algorithms, the method of group accounting of arguments, etc.) [5].

However, real-time series, such as the implementation of a random process, are difficult to forecast. Typically, the methods that provide a quantitative forecast are applied in the statistical analysis of time series. On the other hand, in time series in which the hypothesis of a certain trend existence is not confirmed, the application of such methods is meaningless. In this case, methods for detecting the general behavior tendency of the time series are applied. These methods make it possible to conduct a qualitative analysis of the time series [6].

Using the methods based on the theory of nonlinear dynamical systems and fractal theory provides an idea of the change in properties and behavior of various processes and objects under external physical influences [7].

The task of fractal analysis of any time series is to detect the presence of long-term memory in it, calculate the value of the Hurst exponent, and identify the trend resistance. Knowledge of the fractal characteristics of the considered time series allows the analyst to evaluate the prospects of a reliable forecast of the time series.

The analysis of daily, weekly, monthly, and annual graphs of electric load showed the presence of fractal properties and short-term and long-term memory. Thus, the fractal analysis, which determines the dependence of future values on retrospective information, achieving effective forecasting and control of power consumption [8], is employed in energy load forecasting.

In [9], the dependence of the quality of forecasting power consumption for one or two integration periods on the fractal dimension D was revealed. It was shown that the closer D was to the value of one (i.e., the stronger the persistence was), the smaller the forecast error was, which be used in practice for operational forecasting of electric load.

Besides, the analyses conducted in [10] proved that the forecasting method based on the theory of fractals could provide a higher forecasting quality than the neural networks-based method.

In [11], the EL forecasting with a large number was conducted using different methods, and it was noted that each of the methods had its own advantages and disadvantages. Therefore, for effective operational forecasting, not single, but several methods should be used simultaneously, realizing various hybrid forecasting systems [12].

In general, each forecasting type requires individual data processing, and most methods that were developed to forecast electrical loads are not universal. Therefore, for a specific consumer of electricity, an individual forecast is necessary. Thus, in order to reduce the number of possible forecasting models that adequately describe the time series of electrical loads, it is necessary to conduct a forecasting analysis, where even at the stage of choosing a forecasting method is possible to cancel the forecasting algorithms that will lead to lower quality forecasts. Carrying out such an analysis will ensure choosing some of the most effective forecasting methods for a particular consumer under specific conditions, which will both reduce the forecast time and improve forecast quality.

The most commonly used forecasting analysis of time series in medicine and economics [6, 13, 14] considers only the use of fractals to assess the persistence of time series. However, in order to conduct a forecasting analysis of time series of electrical loads, it is necessary to apply statistical methods for identifying abnormal data and testing autocorrelation along with the theory of fractals.

Currently, the methods based on the statistical criteria of Fisher and Student are used to detect the trend of electrical loads. Since the application of these criteria is labor-intensive and difficult for practitioners in industrial enterprises, it is advisable to apply a less complicated method to detect the trend of time series of electrical loads.

To the best of authors' knowledge, the law of changing fractal parameters of time series of power consumption with the time series duration has not

been studied yet. Therefore, this research allows obtaining detailed information about the behavior of time series of electric loads.

Thus, the improvement of the forecasting method of electrical loads is an important task.

Research methods

The probability models created on the basis of statistical analysis have been widely used to study the graphs of electrical loads of groups of electrical consumers. At the same time, these models have been idealized by introducing the assumption that the distribution of a random component of the electrical load follows the normal law. However, it is known that the distribution of electrical load graphs does not always follow Gaussian or normal distribution. Therefore, in the analysis of electrical loads, it is necessary to use nonlinear dynamics methods [15, 16] for a better understanding of processes and increase the accuracy of their forecasting.

The development of the methodology for a pre-forecasting analysis is to select from a large number of procedures of preliminary, fractal and statistical analyses, which would provide the best operational forecasting. The following sequence of the pre-forecasting analyses of energy efficiency for its subsequent operational forecasting is proposed:
1. Preliminary analysis.
    1.1 Plot the EL time series.
    1.2 Identify anomalous observations.
    1.3 Identify anomalous observations.
    1.4 Check the availability of the time series trend.
2. Fractal analysis [17].
    2.1 Calculate the fractal parameters of the time series of electrical loads.
    2.2 Perform autocorrelation testing.
3. Assessment of the results.
    3.1 Process the results.
    3.2 Choose an optimal method for operational forecasting.

The first stage of forecasting analysis of the time series of the electrical load is a preliminary analysis, which intendeds to check the available data for compliance regarding data comparability and its completeness, homogeneity, and stability. Next, the main dynamic characteristics of the time series under study are determined, i.e., the trend checking and autocorrelation coefficients.

Identification of anomalous observations (i.e., gross errors) is a mandatory procedure in the preliminary data analysis, since their presence distorts the time series processing results. Various methods have been developed for the diagnosis of blunders, such as the Irvine method [18]. After checking the series regarding the presence of gross errors, it is necessary to check the time series regarding the completeness and, in the absence of certain data, fill them by mathematical methods [19].

Methods based on Fisher's F-test and Student's t-test [20] can be used to verify the presence of a trend in the EL time series, but these methods are both laborious and time-consuming. Therefore, in order to detect the trend on these time series, it would be the most efficient to apply the method developed in [21]. This method consists of calculating the relative distance between three data series, called the direct (Direct List), reverse (Reverse List) and measuring (Measurement List) data series, which are formed from the original data series of volume N. The first data series (Direct List) is obtained by sorting the levels of the original data in ascending order, and the reverse one (Reverse List) is obtained by the sorting in descending order. The Measurement List is the source data series.

The distance between the specified lists, respectively, d_1, d_2, and d, is determined by the sum of the absolute differences between the forward r_d series, inverse r_r series, and measured r_m value series, which is expressed as:

$$d_1 = \Sigma \left| r_d - r_m \right|;$$
$$d = \Sigma \left| r_d - r_r \right|;$$
$$d_2 = \Sigma \left| r_r - r_m \right|.$$

The ordered rows (Direct List and Reverse List) form, respectively, the left and right boundary sequences, between which the levels of analyzed series are located. Quantitatively, the trend is estimated by the value of M_s, which represents the relative distance between the analyzed signal and its ordered boundary standards, and it is defined by:

$$M_s = \frac{d_1 - d_2}{d}. \tag{1}$$

The parameter M_s calculated according to (1), for the analyzed time series $\{Y_t\}$, can has one of three values: when $M_s \approx 0$, then $\{Y_t\}$ does not have trend; when $M_s \approx -1$, then $\{Y_t\}$ have a monotonously increasing trend with the regularity of the left border;

and when $M_s \approx 1$, then $\{Y_t\}$ has a monotonously decreasing trend with the regularity of the right.

Estimation of the irregularity degree of time series is performed using the coefficient of variability K [22].

The coefficient of variability K is proportional to the relative rate of change in values of the time series, and therefore can be used to classify time series according to the nature of their change. Thus, the values of $K = 1$ and $K = 2$ divide the entire set of time series into three classes as follows:

- time series with the predominant deterministic component (mostly regular) ($0 < K < 1$);
- time series with a high degree of irregularity (mostly random) ($1 < K < 2$);
- noise sequences ($K \geq 2$).

The second stage of the forecasting analysis of the EL time series consists of determining the fractal parameters of the time series of electrical loads and calculating the nth order r_n autocorrelation coefficients.

In order to determine the fractal dimension of time series using the number of observations of n, the method of normalized scope developed by P. Hurst [23, 24] is used. The Hurst exponent H is associated with the fractal dimension D of the time series such that $D = 2 - H$.

V_k- and V_z-methods were developed to isolate the class of oscillatory processes, assess their periodicity, and classify stationary random processes according to their distribution law. The main parameters of these methods are as follows [27]:

- the ratio of the accumulated deviation magnitude to the range of the series is $V_k = R/R_p$;
- the ratio of the accumulated deviation magnitude to the range of increments is $V_z = R/R_{inc}$, where R_{inc} denotes the increment scale, i.e., the difference between the maximum and minimum increments.

It is useful to evaluate a sample by the fractal parameters (R/R), (R/S) [26], H, C, D, NF [25], V_k, V_z, and the values of variation CV, form factors K_f, and fill K_{fil} of the daily graph.

The Hurst exponent H and fractal dimension D are used to estimate the dynamics of the EL time series, and the parameter (R/R) is intended to determine the type of process. Namely, the virtual volume NF, and parameters V_k, and V_z define the distribution type.

Dependence $(R/R) = f(n)$ is described by a theoretical model $(R/R) = A + Bn^{S'}$, where A, B, S' denote model coefficients, of which B and S' are classifying coefficients [27].

The whole area of the existence of fractal (R/R)-functions can be divided into the trend (T), stationary (S'), and oscillatory (O) areas. When, $0 < S' < 1$, then the process is oscillatory, and when $S' > 1$, then the process has a trend [27]. In the area of T, there are the processes with a clearly defined trend component; in the area of S, there are the stationary processes that do not have a trend, and they can be classified by parameter B, which characterizes the distribution law; lastly, the area «O» includes the oscillatory processes.

The generalized correlation coefficient C is used to determine the degree of long-term correlation between the past and future increments. Therefore, in the forecasting analysis, it is necessary to calculate the values of r_1 and r_n.

The third stage of the EL forecasting analysis is to assess the obtained results and to select the optimal method for operational forecasting.

Research results and discussion

The electricity meter was installed at the 10-kV input of the substation of the pumping station of the energy-intensive chemical enterprise for 30 days in 2018 for the purpose of collecting the active power values at a 15-minute and 60-minute averaging intervals. The time series $\{P_{1,t}\}$ and $\{P_{2,t}\}$ contained 2880 and 720 elements, respectively, as shown in Fig. 1.

The variability coefficients of the time series $\{P_{1,t}\}$ and of $\{P_{2,t}\}$ were $K = 0,041$ and $K = 0,038$, respectively (these are the average value for 30 days). Therefore, the studied time series of the electrical load were the series with the predominant deterministic component.

The measures of the relative distance for the monthly and daily time series of EL were calculated, respectively, and their trends were expressed implicitly.

The results of the preliminary analysis of the time series $\{P_{1,t}\}$ are summarized in Table 1. The load dips of the pumping station were examined in terms of accuracy due to short-term outages. The analysis of the results presented in Table 1 showed that:
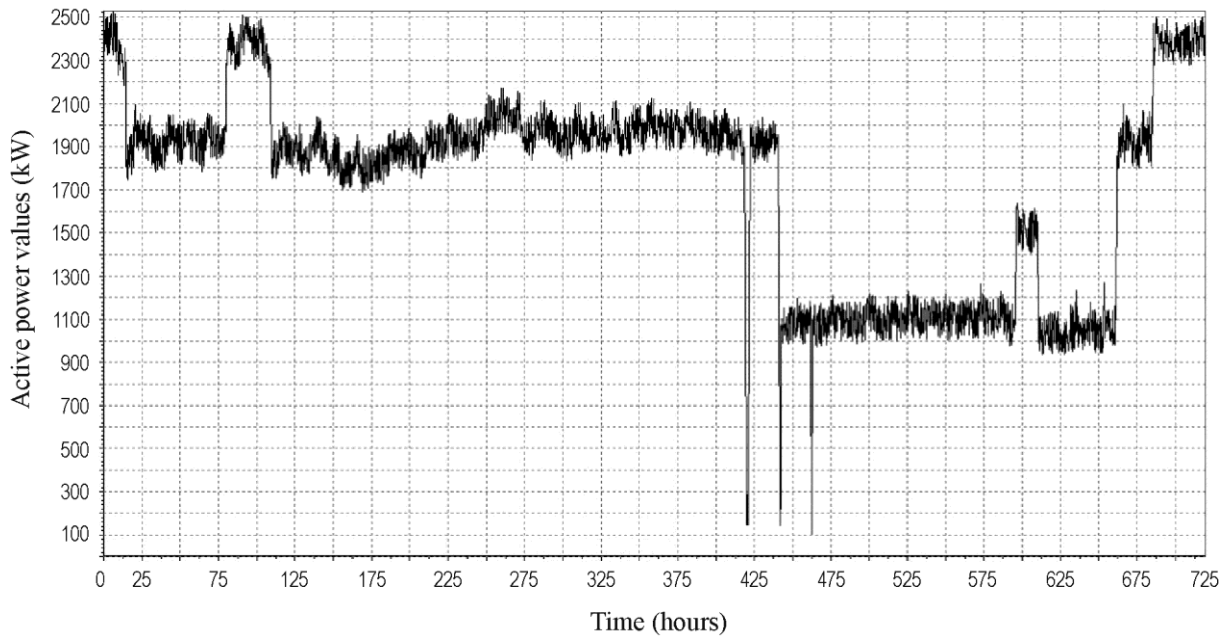
Fig. 1. Dynamics of power consumption of the pumping station (the time series $\{P_{1.t}\}$)

- an increase in the length of the time series led to an increase in the values of parameters (R/R), (R/S), V_k and V_z;
- the considered EL time series had a long-term memory effect and persistence;
- the value of parameter $NF = 28.6$ corresponded to the Rayleigh distribution [1].

Table 1. Calculation results of the fractal parameters, coefficient of variation at the input of the 10-kV pumping station for the time series $\{P_{1.t}\}$

N	R/R	R/S	H	C	D	NF	V_k	V_z	CV
72	124	27	0,92	0,78	1,08	12,8	7,5	10,7	9,6
96	567	44	0,98	0,93	1,02	10,4	13,5	19,3	11
120	1135	56	0,98	0,96	1,02	10,8	17,1	24,4	11,1
144	1370	66	0,98	0,95	1,02	10,7	20,2	28,8	11,4
168	1473	75	0,97	0,93	1,03	11,2	22,4	32	11,3
192	1516	82	0,97	0,91	1,03	12,4	23,4	33,4	10,8
216	1564	89	0,96	0,89	1,04	13,2	24,6	35	10,5
240	1591	96	0,95	0,87	1,05	14,3	25,3	36	10,2
264	1625	101	0,95	0,86	1,05	15	26,1	37,3	10
288	1728	109	0,94	0,85	1,06	15,9	27,2	38,7	9,7
480	8966	182	0,95	0,87	1,05	11,4	53,7	62,9	11,1
1440	10592	333	0,88	0,7	1,12	24,4	67,3	84,5	8,6
2880	111536	1194	0,97	0,93	1,03	28,6	223,1	214,9	26,3

The diagram of the dependence of the parameter (R/R) on the length N of time series $\{P_{1.t}\}$ in the double logarithmic scale is shown in Fig. 2, where it can be seen that the power consumption had a trend $(S > 1)$, which was because $(R/R) = BN^S$, where $B = 0.569$ and $S = 1.475$.
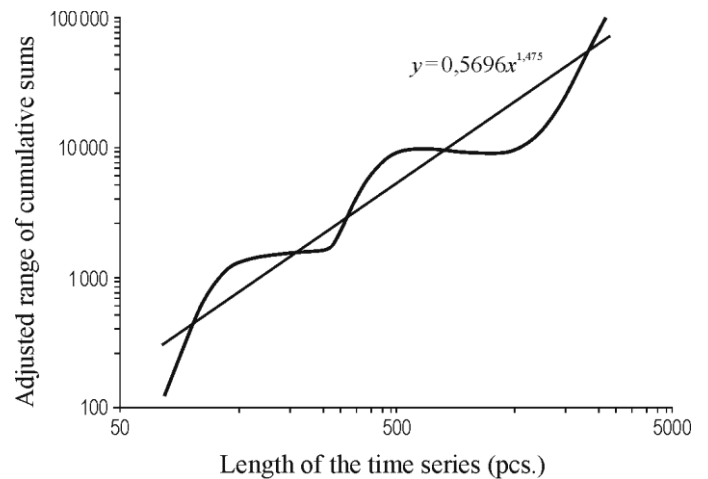

Fig. 2. The dependence of the parameter (R/R) on the length N of the time series $\{P_{1.t}\}$

The calculated fractal parameters and values of the coefficient of variation CV, autocorrelation coefficients of the first-order r_1, form K_f, and filling K_{fil} of a 15-minute load curve for each day of the first decade of the month are provided in Table 3.

Changes in the fractal parameters of the investigated time series $\{P_{1.t}\}$, which are presented in Table 2, did do not affect the coefficients of the electrical load graph and the value of the first-order autocorrelation coefficient. In the time series that have a strong non-linear tendency, such as a second-order parabola or exponent, the autocorrelation coefficient of the levels of the original series may approach to zero, as can be observed in the data collected on the first, sixth and seventh days (refer to Table 2).

24

Table 2. The calculation results of fractal parameters, coefficients of variation, filling, and forms of the time series of the 15-minute values of power consumption for each interval of 10 days

No of order	R/R	R/S	H	C	D	NF	V_k	V_z	CV	K_{fil}	K_f	r_1
1	*567*	*44*	*0,98*	*0.93*	*1.02*	*10.4*	*13.5*	*19.3*	*11*	*0.87*	*1.03*	*0*
2	164	26	0.84	0.6	1.16	14.8	6.7	5.1	3.6	0.94	1.01	0.10
3	25	11	0.62	0.18	1.38	16.2	2.7	1.8	3.4	0.94	1.1	0.10
4	*644*	*43*	*0.97*	*0.93*	*1.03*	*8.5*	*14.7*	*19.6*	*10.2*	*0.89*	*1.02*	*0.17*
5	*808*	*45*	*0.98*	*0.96*	*1.02*	*8.8*	*15.3*	*27.4*	*11.6*	*0.87*	*1.03*	*0.3*
6	86	21	0.79	0.5	1.21	18.6	5.0	3.8	3.3	0.93	1.01	0
7	117	23	0.8	0.52	1.2	15.4	5.7	3.9	3.7	0.93	1.01	0
8	124	24	0.82	0.56	1.18	17.1	5.8	4.8	3.8	0.93	1.01	0.37
9	72	20	0.77	0.46	1.23	20.5	4.4	3.4	3.2	0.93	1.01	0.28
10	84	18	0.74	0.39	1.26	14.2	4.6	2.7	3.1	0.95	1.01	0.29

Table 3. The results of the regression analysis of fractal parameters

No of order	Regression equation	R^2	R^2_{adj}
1	$\hat{y} = 0.649 + 0.020x_1$	0.887	0.873
2	$\hat{y} = 0.728 + 0.011x_2$	0.790	0.751
3	$\hat{y} = 0.672 + 0.028x_3$	0.756	0.726
4	$\hat{y} = 0.733 + 0.0004x_4$	0.799	0.774
5	$\hat{y} = 0.59 + 0.043x_1 - 0.01x_2$	0.918	0.894
6	$\hat{y} = 0.622 + 0.049x_1 - 0.004x_2 - 0.025x_3$	0.937	0.905
7	$\hat{y} = 0.486 + 0.090x_1 + 0.042x_2 - 0.038x_3 - 0.002x_4$	0.998	0.993

The calculation results showed that there were two main types of daily series of power consumption of the pumping station. In the first one, CV was 3–4 %. This CV corresponded to the following parameters: (R/R) up to 200, (R/S) up to 30, H smaller than 0.9, and V_k and V_z smaller than 10. In the second type of the series, which is represented by the italic text in Table 2, CV higher than 10%, which corresponded to the following parameters: (R/R) more than 500, (R/S) more than 40, H larger than 0.95, and V_k and V_z larger than 10. Thus, it is advisable to classify the studied time series into groups.

The analysis of fractal parameters showed that the Hurst exponent H depended on the parameters V_k, V_z, CV, and (R/R). The Hurst exponent H was considered as a dependent variable, and the regressors were V_k (factor x_1), V_z (factor x_2), CV (factor x_3), and (R/R) (factor x_4). The coefficient of determination of R^2 defined the quality of the regres-

sion function. Adding new variables x_2-x_4 to the regression function $\hat{y} = b_0 + \sum_{i=1}^{n} b_i x_i$ increases the normalized coefficient R^2_{adj} (i.e., the adjusted coefficient of determination). The calculation results of regression analysis parameters are summarized in Table 3.

The calculation of the fractal parameters of the time series of active power allows getting more detailed information about the dynamics of the power consumption process and EL time series trend, establishing the distribution of time series of power consumption, and thus choosing the most suitable method to forecast the electrical load. Therefore, it is necessary to simultaneously conduct a fractal analysis of the time series of active, reactive, and total power to determine the relationships and dependencies between them.

When the regression equation is based on the factors $V_k(x_1)$, $V_z(x_2)$, $CV_z(x_3)$, $R/R(x_4)$, then 99.8 % (coefficient of determination $R^2 = 0.998$) of the change in the trait can be explained by the variation

of regressors, and the remaining 0.2 % can be explained by other reasons (refer to Table 3).

The values of the autocorrelation coefficients were calculated for each day. Their values of the first day are given in Table 4. When the first-order autocorrelation coefficient turned out to be the highest, the series under study contained only a tendency. When the autocorrelation coefficient was of the order of r, the time series contained the cyclic oscillations with a periodicity at r times. Thus, the time series given in Table 4 has four local maxima, at 2, 8, 15, and 23 hours. Also, the highest was the 7-order autocorrelation coefficient, which means that the time series contained cyclic oscillations with a frequency of 7 hours. When none of the autocorrelation coefficients is significant, then a series does not contain trends and cyclical fluctuations and has a random structure.

Table 4. The calculation results of autocorrelation coefficients of daily EL time series of the first day, consisting of 24 hourly values

The order of the autocorrelation coefficient					
1	2	3	4	5	6
1.0000	0.0017	-0.4700	-0.0825	0.0819	-0.1906
0.0017	1.0000	-0.0623	-0.4559	-0.0847	0.1800
-0.4700	-0.0623	1.00000	-0.0293	-0.4712	0.0110
-0.0825	-0.4559	-0.0293	1.0000	-0.0298	-0.5587
0.0819	-0.0847	-0.4712	-0.0298	1.0000	-0.5984
-0.1906	0.1800	0.0110	-0.5587	-0.03208	1.0000

Thus, according to the results of pre-forecasting data analysis, the considered time series were persistent, with an implicitly expressed trend, and had a random structure. Therefore, in this case, it would be recommended to use the simplest and most reliable methods of operational forecasting for operational prediction, such as methods of exponential smoothing of Brown and Holt [29].

According to the results of the pre-forecasting analysis if the time series of power consumption does not have persistence, then for operational forecasting can be done by recommend regression methods (in the presence of numerical values of parameters, on which power consumption depends), methods for constructing a mathematical model, neural networks-based methods.

The hourly data of power consumption at the first input of the pumping station on the first day of the interval of ten days is given in Table 5. The value of the smoothing constant α was calculated by Brown's simple exponential smoothing using the 15-minute

daily samples and one hour every ten days [29]. In [30], it was proposed to expand the limit of the smoothing constant α when forecasting the economic processes for this constant was not between the classical limits $0 < \alpha < 1$, but within the interval $0 < \alpha < 2$; when $1 < \alpha < 2$, then a model is called the out-of-limit set. In the interval $0 < \alpha < 2$, it is possible to determine the smoothing constant for a number of EL. Previously, there has been no evidence that the value of the constant smoothing α of the time series of power consumption might be in the interval $1 < \alpha < 2$. In this paper, an example of such a time series is given for the first time, and it is provided in Table 5.

Table 5. The values of hourly power consumption of the pumping station for one day

Interval number	1	2	3	4	5	6	7	8
P (kW)	2420.0	2491.8	2346.0	2438.0	2377.6	2387.7	2488.8	2495.2
Interval number	9	10	11	12	13	14	15	16
P (kW)	2397.9	2327.0	2501.8	2435.9	2334.9	2427.1	2473.5	2398.2
Interval number	17	18	19	20	21	22	23	24
P (kW)	2337.6	2396.5	2503.2	2513.0	2425.2	2461.5	2351.0	2354.2

On the first and third days, the value of α of the daily samples for one hour was 1.1. This denotes a situation where the previous value could not be used as a good estimate of the forecasted value of the process being modeled. In this case, either the process of power consumption went beyond the framework of simple dynamics, and it developed a certain tendency in development, or the process was on the verge between evolutionary and chaotic dynamics, and its mathematical description was impossible to be obtained using any of the existing models.

For a time series of power consumption that consisted of 96 values per day, the smoothing constant for each of 10 days of the first 10-day interval was in the range 0.5–0.9. When the value of α is close to unity, then, in the forecast, the effects of only the latest observations should be taken into account. On the other hand, when the value of α is close to zero, then the weights by which the levels of the time series are weighted decrease slowly, i.e., all (or almost all) past observations are taken into account in the forecasting process.

The mean absolute percentage error (*MAPE*) of power consumption prediction of daily samples for ten days is given in Table 6. Their maximum value are of 0.4 and 3.4 % for the Holt method and the

Brown method, respectively. The presented results prove that forecasting errors for all the time series of ELs were almost constant over a day and did not depend on changes in the values of fractal parameters, and due to high persistence (Hurst exponent of each row was not less than 0.6) high-quality forecasting was achieved.

Table 6. The *MAPE* (%) of power consumption prediction of daily samples for ten days

Forecasting method	Day number									
	1	2	3	4	5	6	7	8	9	10
Brown	3.3	3.0	3.1	2.5	2.9	2.6	3.0	2.9	2.7	2.5
Holt	0.4	0.4	0.4	0.3	0.3	0.3	0.4	0.4	0.3	0.3

The obtained results are typical for all pumping stations for water supply to continuous petrochemical and metallurgical production. The result represented in this work can be used not only for pumping stations but also for workshops and productions with a continuous work cycle.

The results presented in this work show that it is advisable to calculate the Hurst exponent not only for certain time intervals but also for each period of time, starting with two periods of counter integration and ending with the end of a time series consisting of *n* integration intervals. Such a calculation of the Hurst exponent can reveal the change in the dynamics of the time series of power consumption.

The presented methodology for the forecasting analysis of the EL time series is easy to formalize and automate. This methodology includes simple methods for determining the trend of time series of power consumption, classifying time series by the nature of their change, and identifying the persistence or antipersistency, type (distribution law) of a random process, and stability of the production process. The proposed pre-forecasting analysis procedures are sufficient for the purposeful selection of the most suitable method for operational forecasting of electrical loads of large and medium-sized industrial enterprises, which is, in this case, the Holt's method.

Conclusion

The main contributions of this work can be summarized as follows:

1. It is proved that using the trend detection method of radio signals to identify the trend of time series of electrical loads simplifies calculations and does not require knowledge of statistical methods.

2. It is established that there are several types of daily power consumption of the pumping station, and each type has its own distribution law. Namely, in the first type, the coefficient of variation was 3–4 %, and the Hurst exponent was less than 0.9. The second type of series corresponded to the following parameters: the Hurst exponent was larger than 0.95, and the coefficient of variation was higher than 10 %. Thus, the information on the type of diurnal row allows choosing the most appropriate forecasting method for each day.

3. It is proved that the time series of power consumption of the pumping station contained cyclical fluctuations.

4. For the first time, it was revealed that the value of the smoothing constant α of the time series of electric loads of a pumping station during the forecasting could be within the beyond-limit set, i.e., $1 < \alpha < 2$. In this case, either the process of power consumption went beyond the framework of simple dynamics, and it developed a certain tendency in development, or the process was on the verge between evolutionary and chaotic dynamics, and its mathematical description could not be obtained by any method.

5. A methodology for conducting a forecasting analysis of power consumption, which allows choosing an optimal method for on-line forecasting of electrical loads of industrial enterprises, and thus reduce the forecast error, was introduced.

Nomenclature

ARMA – autoregressive moving average model;
EL – electrical loads;
MAPE – mean absolute percentage error;
A – coefficient of (R/R)-model;
B, S' – classifying coefficients (R/R)-model;

b_0, b_i – regression coefficients;
C – generalized correlation coefficient;
CV – coefficient of variation;
D – fractal dimension;
d – distance between data series of the Direct List and the Reverse List;
d_1 – distance between data series of the Direct List and the Measurement List;

d_2 — distance between data series of the Reverse List and the Measurement List;

H — Hurst exponent;

K — coefficient of variability;

K_f — coefficient of form;

K_{fil} — coefficient of filling;

M_s — quantification of a trend;

n — number of factors;

N — length of time series;

NF — virtual volume;

r_1 — first-order autocorrelation coefficient;

R^2 — coefficient of determination;

R_{inc} — increment scale;

r_d — counts of forward series;

r_m — counts of measured value series;

r_r — counts of inverse series;

T — trend area of existence of fractal (R/R)-functions;

S — stationary area of existence of fractal (R/R)-functions;

O — oscillatory area of existence of fractal (R/R)-functions;

V_k — fractal parameter;

V_z — fractal parameter;

R_{adj}^2 — adjusted coefficient of determination;

(R/S) — adjusted range of cumulative sums;

$\{P_t\}$ — time series;

x_i — model factor;

$\hat{y}$ — regression function;

Greek symbol

α — smoothing constant.

References

[1] Lezhniuk P.D., Shullie Yu.A. Operatyvne prohnozuvannia elektrychnykh navantazhen system elektrospozhyvannia z vrakhuvanniam yikh fraktalnykh vlastyvostei. Vinnytsia, VNTU, 2015, 104p.

[2] Alfares H.K., Nazeeruddin M. Electric load forecasting: literature survey and classification of methods. *International Journal of Systems Science*, 2002, vol. 33, no. 1, pp. 23–34. doi: 10.1080/00207720110067421.

[3] Almeshaiei E., Soltan H. A methodology for Electric Power Load Forecasting. *Alexandria Engineering Journal*, 2011, vol. 50, no. 2, pp. 137–144. doi: 10.1016/j.aej.2011.01.015.

[4] Baliyan A., Gaurav K., Kumar Mishra S. A review of short term load forecasting using artificial neural network models. *Procedia Computer Science*, 2015, vol. 48, no. C, pp. 121–125. doi: 10.1016/j.procs.2015.04.160.

[5] Abdurahmanov A.M., Volodin M.V., Zybin E.Yu., Ryabchenko V.N. Forecasting methods in electricity distribution networks (review). *Russian Internet Journal of Electrical Engineering*, 2016, vol. 3, no. 1, pp. 3–23. doi: 10.24892/RIJEE/20160101.

[6] Antonova I., Chikina N. Application of nonlinear dynamics methods in prepropriate analysis of the morbidity indicators in Ukraine by various skin pathology. *Norwegian Journal of development of the International Science*, 2018, vol. 1, no. 14, pp. 48–52.

[7] Kirillov A.K, Kazimirov A.N. The fractal dimension of time series of biological processes and objects. Bulletin of Donetsk National University. Series A: Natural Sciences, 2013, no. 2, pp. 149–155.

[8] Lezhniuk P., Bondarchuk A., Shullie Iu. Investigation and implementation of the fractal properties of electric load on civilian objects in order to efficiently forecast and control electrical consumption. *Eastern-European Journal of Enterprise Technologies*, 2019, vol. 3, no. 8-99, pp. 6–12. doi: 10.15587/1729-4061.2019.168182.

[9] Voloshko A., Bederak Y. Fractal analysis of electrical loads of chemical plants. *Power Engineering: Economics, Technique, Ecology*, 2016, no. 4 (46), pp. 53–59.

[10] Hongsheng Su. Short-Term Load Forecasting Method Based on Fractal Theory. *Wseas Transactions on Circuits and Systems*, 2012, vol. 11, no. 6, pp. 169–181.

[11] Panyukova D., Nurpeissova G. Review of modern approaches for power load forecasting. Proceeding of the 15th International Scientific Conference Information Technologies and Management, 2017, pp. 69–71.

[12] Voloshko A., Bederak Ya., Lutchin T. Operational forecasting of power consumption at enterprises with a continuous cycle of work. *Electronic Modeling*, 2015, vol. 37, no. 6, pp. 111–118.

[13] Melikhov E.V. Ebzeeva N.S. Identification of long-term memory of time series of the sale of homogeneous goods. Proceedings of the 1th International forum "Actual problem of modern science", 2005, pp. 53–56.

[14] Tambieva D. Nonlinear dynamics methods in economic time series investigations. *Izvestia: Herzen University Journal of Humanities & Sciences*, 2008, no. 71, pp. 192–200.

[15] Shulle Yu.A. Analysis of graphs of electrical loads of industrial enterprises by method of rescaled range (R/S-analysis). Visnyk of Vinnytsia Polytechnic Institute, 2015, no. 2, pp. 50–56.

[16] Cui H.R., Wu R.M. Research characteristic of power load based on fractal theory. *Journal of Electric Power*, 2012, no. 3, pp. 181–185.

[17] Jian-Kai L., Cattani C., Wan-Qing S. Power load prediction based on fractal theory. *Advances in Mathematical Physics*, 2015, vol.2015, Article number 827238. doi: 10.1155/2015/827238.

[18] Irwin J. O. On a criterion for the rejection of outlying observations. *Biometrika*, 1925, vol. 17, no. 3–4, pp. 238–250.

[19] Kaminskyi R., Kunanets N., Pasichnyk V., Rzheuskyi A., Khudyi A. Recovery Gaps in Experimental Data. Proceeding of the COLINS, 2018, pp. 108–118.

[20] Box G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M. *Time Series Analysis: Forecasting and Control*. Hoboken, John Wiley and Sons Inc., 2015, 712 p.

[21] Gorshenkov A.A., Klikushin Yu.N., Kobenko V.Y. Algorithm of detection, measurement and classification of signals trends. *Digital Signal Processing*, 2012, no. 2, pp. 46–49.

[22] Klikushin Yu.N. The quantitative assessment of the properties "regularity-chaotic state of signals". *Journal of Radio Electronics*, 2006, no. 4. Available at: http://jre.cplire.ru/jre/oct06/1/text.html.

[23] Hurst H.E. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 1951, vol. 116, Article number 2447.

[24] Kurskoi Yu.S. Fractal analysis of electrical measurements results. *Energy saving. Power engineering. Energy audit*, 2014, no. 1 (119), pp. 52–57.

[25] Klikushin Yu.N. The fractal scale for distribution's measurement. *Journal of Radio Electronics*, 2000, no. 3. Available at: http://jre.cplire.ru/koi/mar00/2/text.html.

[26] Klikushin Yu.N. The fractal classification method for complex signals. *Journal of Radio Electronics*, 2000, no. 4, Available at: http://jre.cplire.ru/jre/apr00/1/text.html.

[27] Kobenko V.Iu. Methods for processes identifying on fractal scales. Cand. tech. sci. diss. Omsk, 2000. 151 p.

[28] Bederak Ya.S. On substantiation of selection of economic and mathematical methods for the assessment of energy efficiency of production facilities. *Electrical engineering & electromechanics*, 2017, no. 1, pp. 67–72. doi: 10.20998/2074-272X.2017.1.11.

[29] Stetsenko I.V., Bederak Ya.S. Restoration and operational forecasting of the electricity consumption of enterprises with continuous cycle of operating using Holt's method. *Electronic Modeling*, 2015, vol. 37, no. 4, pp. 119–126.

[30] Svetun'kov S.G., Svetun'kov I.S. Methods of socio-economic forecasting. St. Petersburg, SPbGUEF Publ., 2010. 103 p.