https://doi.org/10.33271/nvngu/2023-4/096

N. S. Dreshpak*,
orcid.org/0000-0002-4453-1378,
O. S. Dreshpak,
orcid.org/0000-0003-1019-4382

Dnipro University of Technology, Dnipro, Ukraine
* Corresponding author e-mail: dreshpak.n.s@nmu.one

PARAMETERIZATION OF THE STATISTICAL MODEL FOR ELECTRICAL ENERGY EFFICIENCY CONTROL

Purpose. Justification of a structural construction and parameters of a regression model for the normalization of specific energy consumption when controlling the production process energy efficiency.

Methodology. Analysis of the peculiarities of energy efficiency control of the production process in conditions of frequent and significant changes in specific energy consumption, followed by the determination of the structure and parameters of the regression model.

Findings. Based on the presence of frequent and significant changes in the energy efficiency control of the production process, the reasonableness of normalizing the specific energy consumption by using the regression model with a variable structure is substantiated. The actual daily specific energy consumption indicators, obtained during the month to control energy consumption efficiency and build the regression model of the variable structure, are used. The limited possibilities for the formation of voluminous statistical samples with homogeneous data, and the complexity and laboriousness of measuring a significant number of influence parameters make it necessary to reduce the number of explanatory variables of the regression model. The feasibility of using the value of the output volume, as a comprehensive characteristic of the level of energy consumption, is proven. The acceptability of the application of linear and non-linear univariate regression dependencies is determined. The nonlinear model, as a result of reducing the linear model of energy consumption to a nonlinear form characteristic of the values of its specific consumption, is obtained.

Originality. For the first time, the use of the regression model of the variable structure for the normalization of specific energy consumption in conditions of frequent and significant changes in the energy efficiency of the production process, which helps to increase the accuracy of their determination, is proposed. The need to reduce the number of explanatory variables of the regression model is proven. The expediency of using linear or non-linear one-factor regression dependencies in the given conditions of energy efficiency control, which helps to simplify the procedure of registering the initial data for their construction, is confirmed.

Practical value. The scientific results of the performed studies allow for taking into account the peculiarities of the production conditions when determining the structure and parameters of the regression model for normalizing the specific energy consumption. This contributes to increasing the accuracy and energy efficiency control of the production process.

Keywords: *energy efficiency control, specific energy consumption, regression dependence, model parameterization, normalization methods*

Introduction. The energy intensity of Ukraine's gross domestic product remains high and significantly exceeds the indicators of developed countries. Despite the fact that energy saving is a priority direction of state policy, there is a significant delay in the implementation of state programs and insufficient activity of industrial enterprises in implementing energy-saving measures, which affects the competitiveness of domestic products and creates financial problems for intra-economic activity. The law "On Energy Efficiency" adopted by the Verkhovna Rada of Ukraine in 2021 focuses on ensuring the "energy efficiency" of production processes [1]. It is planned to strengthen controlling actions by the state in the field of energy resource use, significantly reducing the amount of energy consumption [2]. The law of Ukraine provides for the introduction of energy management systems at production enterprises among the priority directions for solving the formulated tasks. In 2020, the standard of Ukraine DSTU ISO 50001:2020 (ISO 50001:2018, IDT) – Energy management. Requirements and instructions for use systems – came into force. This standard corresponds to the international standard ISO 50001:2018 – Energy management system – Requirements with guidance for use. The standard provides for implementing a certain list of actions that contribute to the determination of the enterprise's energy policy, tasks, and specific plans aimed at increasing the energy efficiency of production processes. The important functions of the system are control and management of energy efficiency, which are carried out on the basis of the concept of continuous improvement of the situation [3, 4].

The methodological component of control systems existing at enterprises was created mainly during the period of the planned economy and is not sufficiently effective in market conditions. In new conditions, the enterprise must constantly adapt to changes in the environment and ensure the competitiveness of products [5, 6]. Existing changes affect the structure and regimes of the energy supply of the production process, causing intense unpredictable changes in their indicators. Most of the methods of normalizing the energy efficiency indicator turned out to be unacceptable for use in conditions of frequent and significant changes in its values due to the insufficient validation of the results. There is a need to improve them, taking into account the existing promising solutions in this area. As shown in [7], it is necessary to ensure an increase in the reliability of control, informativeness, and efficiency, which will contribute to the creation of a meaningful theoretical basis for making balanced management decisions and allow reducing the specific energy costs for the production of products.

The electrical component of the power supply of an industrial enterprise deserves special attention. This is due to the wide scope of application and the relatively high cost of electricity. There are features of distribution and accounting and features of its consumption that differ from other types of energy. This determines the need to apply individual approaches to the implementation of power efficiency control systems.

Literature review. In [8], the expediency of using the statistical method of normalizing specific energy consumption is shown. The advantages of the method include the relatively low complexity of the calculation, the insignificant error of forecasting the indicator due to the use of experimental data, and the possibility of assessing the impact of individual factors on the overall result of specific energy consumption. The method requires obtaining reliable information about energy consumption by the control object at certain time intervals determined by the periodicity of control. In power supply systems, this can be ensured by an automatic energy meter reading (AMR) system [9, 10]. The normalization process also involves the creation of a mathematical model for calculating the predicted values of specific energy consumption. Various approaches to the development of statistical models are known. So, for example, in [8], it is proposed to identify the parameters of a statistical series determined by the values of specific energy consumption W at certain intervals. If the law of the distribution of the elements of this series is normal, which is a characteristic of the same conditions of the technological process, then the estimate of the average value of the specific energy consumption is determined as $\widehat{W}$ and the mean square deviation value as δ. Value $\widehat{W}$ is used as normalized and the degree of acceptable deviations of the indicator is associated with the values δ. It should be noted that the possibilities of using statistical series consisting exclusively of specific energy consumption values W in certain periods of time are limited. In this case, forecasting and, accordingly normalization of the value of W is possible for a limited number of future steps (usually one or two steps) (Ayvazyan, 2010). It is obvious that there is a need to take into account other parameters of the technological process, which characterize the conditions of its flow. If such conditions change significantly, then (Ayvazyan, 2010), it is proposed to determine the statistical dependence $\widehat{W}$ from a number of factors (production volume, processing speed, etc.)

$$\widehat{W} = f(X^{(1)}, X^{(2)} \ldots X^{(p)}), \quad (1)$$

where p is the number of factors X taken into account. Dependence (1) can be linear or non-linear. The nonlinearity of the function complicates the procedure for determining the error of calculating the value $\widehat{W}$, and limits the possibilities of applying matrix calculations, the use of which is important for the implementation of operational control associated with a significant amount of calculations in a short period of time. Therefore, in statistical modeling, the desire of the researcher to apply a linear model is manifested. So, for example, in the article [8], it is proposed to apply a linear multifactorial regression model when modeling the specific consumption of electricity by an industrial enterprise. This approach to modeling is, as a rule, traditional. Meanwhile, as shown in [8], there are modes of operation of the energy supply system where the manifestation of the nonlinearity of the function is significant. This primarily applies to cases where a significant part of the facility's energy consumption does not depend on the volume of production.

The system for controlling the efficiency of electric energy consumption, which is to be developed, needs to reduce the error of determining the normalized values. Therefore, the choice of the type of statistical dependence for their calculation is a priority task with the need to justify the decision made. It is known (Ayvazyan, 2010) that there is no clear theoretical basis for solving the problem of choosing the optimal type of statistical dependence. Recommendations for making such a choice are usually taken into account in the literature. As stated (Ayvazyan, 2010), the art of building a regression model is to combine the brevity of the parameterization with its sufficient adequacy in the problem being solved.

Unsolved aspects of the problem. The task of the research, the results of which are presented in the article, is to substantiate the approaches to the parameterization of the statistical model, taking into account the modern market conditions of production of products, which are often characterized by rapid (usually daily) changes in the specific indicators of energy consumption, as well as the peculiarities of the processes of accounting for the electricity consumption of industrial enterprises control of the energy efficiency of the production process.

Statement of the basic material of the research. A statistical model built based on the results of experimental studies in the conditions of a specific product has a limited scope of application related to the peculiarities of the influence of these conditions on the parameters of the production process. At the same time, the existing connection of the model with experimental data contributes to obtaining reliable modeling results directly related to the practice of production. Despite the presence of shortcomings, statistical models have received wide practical application, and their substantive part is continuously improved. The limited scope of application of specific models is compensated by the development of general recommendations regarding their structural construction (Ayvazyan, 2010). The choice of the type of function and explanatory (exogenous) variables should be aimed at minimizing the forecast error. A priori information about the meaningful essence of the addiction under investigation, about physical, economic, sociological, etc., should be used as much as possible. dependencies In addition, a preliminary analysis of the "geometric" structure of the source data is proposed. This reproduces the nature of the relationships existing between the model variables to apply statistical data processing methods that ensure the minimization of the error of predicted values. In regression models, the structure of functional dependence between variables should not be overcomplicated. It should be remembered that the values of the coefficients for their explanatory variables are obtained on the basis of one sample and they may not be acceptable for the next sample (Ayvazyan, 2010). The recommended sequence of choosing a mathematical model is from a simple option (linear model) to a more complex one (if necessary) (Ayvazyan, 2010). When choosing a model, it is advisable to reach a compromise between the complexity of its structure and the accuracy of the calculation of the predicted values. The limited sample size with a significant number of explanatory variables reduces the statistical reliability of the assessment.

The above recommendations and conclusions should be considered sufficiently justified. They are of a general nature, and their definition is due to the long-term positive experience of using statistical models in various spheres of social activity. The orientation of modern enterprises to economic activity in the conditions of a market economy requires increasing the efficiency, reliability, and informativeness of control systems. It is obvious that taking into account the existing production conditions should be reflected in the structural construction of statistical models. It is necessary to determine the general features of the functioning of these systems and, taking them into account, to develop general approaches to the formation of the structure of the statistical model. Let us focus on considering this task.

In the process of controlling the efficiency of electricity consumption, the dynamics of changes in the energy intensity of the process are analyzed. In the competitive conditions of industrial production, a stable production volume is not always realized. Often, the volume changes significantly within short periods of time (for example, within a day). Given that the value W, as a rule, depends on the value of the volume of production in a certain time interval, there is a corresponding fluctuation in the energy intensity of the process. The level of electricity consumption is determined by a number of influencing factors, the degree of influence of which can be different and often changes in the process of energy efficiency control. Changing the operating modes of technological equipment is often associated with adjusting the duration of the electric load or changing its intensity (equipment load level). The introduction of new equipment or its removal from the technological process usually occurs in short periods of time and is accompanied by a change in the system structure energy consumption. Therefore, structural and regime changes in the power supply system of the object of control are often of a rapid nature, caused by the introduction of new technological solutions into production. This is primarily characteristic of enterprises operating in a market economy.

Quantitative characteristics of the power consumption of the object of control depend, for example, on the characteristics of materials and semi-finished products used for manufacturing products, and external factors (environmental temperature, air humidity, and others). Note that some of the listed factors are controllable, while others are uncontrollable or poorly controllable. This distribution of influence factors is important for the implementation of the function of energy efficiency control since the management of production with the aim of reducing the energy intensity of the process involves a change in primarily controlled factors. Uncontrollable and uncontrollable factors, as a rule, cannot be changed. Their values are determined by conditions external to the company's activities. It should be noted that the division of influence factors into managed and unmanaged ones is conditional. It characterizes the degree of acceptability of the application of various influencing factors to manage the energy efficiency of the production process. A change in production conditions can contribute to the transfer of an uncontrollable factor to the rank of a controlled factor and vice versa. The use of controlled influencing factors to solve energy-saving tasks in production is the task of the enterprise team. In the control process, it is important to identify the influence of these factors on specific energy consumption indicators and to separate their effect from the influence of uncontrollable factors. This will contribute to the allocation of that part of the change in the energy efficiency of the process, which is due to the action of the service personnel. The use of this feature of the energy-efficient control of the production process will increase its informativeness by focusing attention on the assessment of the actions of the service personnel, comparing their effectiveness with the influence of uncontrollable factors.

The need to improve production processes often requires changing the structure of the facility's power supply system. The replacement of the composition and number of electrical energy receivers, and the introduction of new operational modes of operation of electrical equipment can take place quickly enough, which is a significant advantage of electrical systems, which necessitates the need to increase the efficiency of energy consumption control and increase the frequency of experimental data registration. It is also important that structural and regime changes in the power supply system can occur every time according to a new, previously unknown scenario, which affects the homogeneity of experimental data. This narrows the possibilities of applying previous experience to forecast electricity consumption in the future. Therefore, it is advisable to forecast the indicators based on the results of experimental data obtained in the past close to the moment of the forecast. This will bring the predicted situation closer to the previously observed situation and, accordingly, increase the accuracy of the forecast.

For effective control of electricity consumption, it is advisable to link the periodicity of control actions with the dynamics of the production process. Taking into account the fact that significant fluctuations in power consumption levels are quite possible in short periods of time, one should focus on reducing the duration of pauses in control actions. This will make it possible to promptly respond to the situation that has developed at the object of control, and to implement measures for its improvement in a timely manner. The question arises as to what measure of control efficiency is acceptable. The technical side of the answer to the question is that the value of electricity consumption is often recorded by the AMR system at certain time intervals. Usually, values of 30-minute observations, and values for a working shift, day, month, and year are recorded. It is possible to set an arbitrary value for the duration of the control period in the system. This is important in the case when the production of a batch of products takes place in predetermined periods with a clear registration of the beginning of the process and its completion. Strict requirements regarding the synchronization of indicators of electricity consumption and production output require careful analysis of the production process, consideration of its individual components, and the sequence of their implementation over time. The broad possibilities of accounting carried out by AMR allow for considering the task of determining the periodicity of energy efficiency control in relation to

$$Y_t = Q_0 + Q_1 X_t^{(1)} + \ldots + Q_p X_t^{(p)} + \delta_t, \quad (2)$$

where Y_t is the value of the explained (source) variable in t – time or space measure; $X_t^{(1)} \ldots X_t^{(p)}$ – the value of explanatory variables (influence factors); $Q_0 \ldots Q_p$ – coefficients to be determined (p – the number of explanatory variables); δ_t – a random component characterizing the difference between the actual Y_t and obtained from the linear model of the estimation of the average values of the explained variable $\widehat{Y}_t$.

Value $\widehat{Y}_t$ is calculated by dependence

$$\widehat{Y}_t = Q_0 + Q_1 X_t^{(1)} + \ldots + Q_p X_t^{(p)}. \quad (3)$$

In order to exclude a systematic error in the estimation of the value Y_t by approximate value $\widehat{Y}_t$ consider that the average value of the random component δ_t is equal to zero $(E_{\delta_t} = 0)$ (Ayvazyan, 2010). To use dependence (2) in the problem of normalization of the specific consumption of electrical energy, it is necessary to ensure the coincidence of the time cycles of measuring the values of the explained and explanatory variables with the determined moments of energy efficiency control. We will focus on daily monitoring. The results of measuring the variables are used both to calculate the parameters of the linear model (3) and to compare the actual values Y_t with normalized. It is planned to obtain experimental data in systems for controlling the efficiency of electric energy consumption by recording the values of variables at certain time intervals, that is, the option of spatial measurements at different control objects is not considered. A significant advantage of forming dependencies (2 and 3) in a linear form is that there is a possibility of applying matrix calculation methods to determine coefficients of regression dependence, errors of their estimation, and correlations between variables. This greatly simplifies the execution of the calculation procedure and facilitates control efficiency. The list of explanatory variables in equation (3) when performing model parameterization should be such that the calculated value $\widehat{Y}_t$ slightly differed from the values Y_t. Minimizing the difference δ_t is possible when the list contains variables whose influence on the explained change is significant. Based on the fact that the value of the specific consumption of electricity is subject to control, that is, the consumption related to the number of products produced by the company, entering the value of the volume of production into the right part of the equation (3) X in a controlled period of time is reasonable. Value X is a "target" parameter that defines the purpose of the production process. Another of its features is that the quantitative characteristics of production output are directly related to the energy costs of the full cycle of technologically related operations aimed at obtaining the final product. From this point of view, indicator X is a comprehensive description of the costs associated with the existing technological process. The value of these costs in the overall balance of energy consumption is, as a rule, decisive. Expanding the list of explanatory variables contributes to the reduction in the calculation error of the predicted values of electricity consumption. But such an expansion also requires an increase in the number of measurements to ensure the statistical reliability of the result, which is problematic from the point of view of data homogeneity in conditions where structural and regime changes occur in the electricity consumption system. If there are significant changes in the object of control, there is also incomplete determination in the list of influential explanatory variables, as this list can change. Regression models describing the energy supply process of such objects have a variable structure. The focus of research on the control of the efficiency of electricity consumption in a system with structural and regime changes requires the use of models of this type. It should be added that within each of the interconnected technological operations, there are indicators that affect the overall level of energy consumption, but its influence on this level is manifested not only due to their characteristic properties but also taking into account the quantitative indicator – the volume of output products *X*. Thanks to this, the energy consumption of each technological operation is proportional to the value of the entered complex characteristic *X*, which must be taken into account when parameterizing the model. Usually, in linear regression equations, such a relationship between explanatory variables is shown in the form of a product of these variables, which is fixed by their interaction [11, 12]. It is obvious that in the case when the values of a number of parameters of the technological process cannot be measured in the control process, the manifestation of the peculiarities of their change can be reflected by a change in the coefficient at the value *X*. This is an acceptable option when applying regression equations of a variable structure, which provides for a change in the values of the coefficients in the case of unknowns (Ayvazyan, 2010).

Taking into account the conditions of operation of the control system to be analyzed, we come to the conclusion about the feasibility of entering the value of the volume of production into the regression model *X* as the main explanatory variable that determines the purpose of the production process and corresponds to the level of the complex character of energy consumption. One should focus on limiting the number of additional explanatory variables, without excluding the option of completely rejecting their use. The last option is a good solution from the point of view of simplifying the procedure of obtaining experimental data for building a model, as it involves periodic registration of energy consumption and production volumes. AMRs provide registration of levels of electricity consumption. As for the volumes of output, it should be emphasized that enterprises pay special attention to the timely registration of quantitative indicators of products at various stages of the technological process of production. Synchronous registration of these two indicators contributes to increasing the reliability of the initial data for the construction of a univariate regression dependence. With a significant number of explanatory variables, additional difficulties are possible, associated with the need to ensure the synchronicity of their measurements. Differences are possible both in terms of data fixation and in the values of variables due to their random nature.

The orientation of the research on the use of linear regression models with the aim of practical implementation of their advantages determines the feasibility of developing an approach that allows using the linear regression dependence by means of its preliminary transformation for the subsequent representation of the nonlinear relationship of specific energy consumption with influencing factors. Let us consider the essence of the proposed approach. We will assume that the error of registering the values of the explanatory variables in the process of their measurement is insignificant. It is obvious that this fully applies to the quantitative indicators of production during the period of energy efficiency control. A significant number of existing production processes are characterized by the accumulation of energy costs *Y*, proportional to the volume of production *X* [8]. Therefore, this relationship between the parameters is appropriate for *Y* and *X* present in the form of a linear one-factor regression dependence. Then we will get an estimate of the average value of electricity consumption by the technological object from the equation

$$\widehat{Y} = G_0 + G_1 X, \tag{4}$$

where G_0, G_1 are coefficients of the regression equation.

Since the value *X* is not random, then the estimate of the average value of specific energy consumption $\widehat{W} = \widehat{Y} / X$ is defined by dependence

$$\widehat{W} = G_0 / X + G_1. \tag{5}$$

The peculiarity of the obtained regression equation (5) is that the dependence takes into account the nonlinearity of the relationship $\widehat{W}$ with value *X*, which is characteristic of a significant number of production processes. The presence of direct proportional dependence between parameters $\widehat{W}$ and $\widehat{Y}$ provides a simple calculation procedure $\widehat{W}$ by values $\widehat{Y}$. The determination of the nonlinear dependence (5) through the auxiliary linear dependence (4) is considered a reasonable decision due to the fact that the use of this approach allows the use of the linear function (4) in the calculations of the normalized values of specific energy consumption, using the above-mentioned advantages of the linear function and ensuring its subsequent transformation into nonlinear equation (5). Additional advantages of the formulated approach include:

- dependence (4) allows you to estimate the average value of the energy consumed by the control object. The parameter $\widehat{Y}$ is often used to predict the level of energy consumption;
- the confidence intervals of the predicted values of energy consumption of the linear dependence (4) can be used when constructing the confidence intervals of the dependence (5), which is important when assessing the reliability of the control.

In addition to non-linear dependence $\widehat{W}(X)$ it may be acceptable to use a linear relationship

$$\widehat{W}_* = Q_0 + Q_1 X. \tag{6}$$

The accuracy of calculation of average values *Y*(*X*), *W*(*X*), $W_*(X)$ by grade values $\widehat{Y}(X)$, $\widehat{W}(X)$, $\widehat{W}_*(X)$ (dependencies (4–6)) is determined by constructing confidence intervals

$$\begin{gathered} \widehat{Y}(X) - \Delta_y(X) \le Y(X) \le \widehat{Y}(X) + \Delta_y(X); \\ \widehat{W}(X) - \Delta_W(X) \le W(X) \le \widehat{W}(X) + \Delta_W(X); \\ \widehat{W}_*(X) - \Delta_{W_*}(X) \le W_*(X) \le \widehat{W}_*(X) + \Delta_{W_*}(X), \end{gathered} \tag{7}$$

where $\Delta_y(X)$, $\Delta_W(X)$, $\Delta_{W_*}(X)$ are maximum calculation errors.

Taking into account the dependence obtained in (Ayvazyan, 2010) for linear regression, we determine the limits of the confidence intervals $Y_{\frac{\max}{min}}(X) W_{*\frac{\max}{min}}(X)$ for *Y*(*X*), $W_*(X)$

$$Y_{\frac{\max}{\min}}(X) = \widehat{Y}(X) \pm t_{(1-P_D)/2} \cdot \sigma_y \sqrt{\frac{1}{n} + \frac{(X - X_0)^2}{\sum_{i=1}^{n}(X_i - X_0)^2}}; \tag{8}$$

$$W_{*\frac{\max}{\min}}(X) = \widehat{W}_*(X) \pm t_{(1-P_D)/2} \cdot \sigma_{*W} \sqrt{\frac{1}{n} + \frac{(X - X_0)^2}{\sum_{i=1}^{n}(X - X_0)^2}},$$

where σ_y, σ_{*W} are mean square deviations; *n* – the number of experimental data sample elements, $t_{(1-P_D)/2}$ – quantiles of *t* – distribution (Student's distribution) of *n* – 2 degrees of freedom, P_D – confidence probability.

The upper and lower limits of the confidence interval for a non-linear dependence *W*(*X*) are obtained taking into account (5, 7 and 8)

$$\begin{aligned} W_{\frac{\max}{\min}}(X) &= \widehat{W}(X) \pm \Delta_W(X) = Y_{\frac{\max}{\min}}(X) / X = \\ &= \widehat{Y}(X) / X \pm \Delta_y(X) / X. \end{aligned} \tag{9}$$

From (8, 9) we get

$$\Delta_W(X) = \frac{\Delta_y(X)}{X} = t_{(1-P_D)/2} \cdot \sigma_y / X \cdot \sqrt{\frac{1}{n} + \frac{(X - X_0)^2}{\sum_{i=1}^{n}(X_i - X_0)^2}}. \tag{10}$$

Dependencies (8, 10) allow determining the maximum calculation errors $Y(X)$, $W(X)$, $W_*(X)$ according to the values of their grades $Y(X)$, $W(X)$, $W_*(X)$. Considering the dependence of errors on X, it is advisable to compare them for value $X = X_0$. When

$$\Delta_y(X_0) = t_{(1-P_D)/2} \cdot \sigma_y \Big/ \sqrt{n};$$
$$\Delta_W(X_0) = t_{(1-P_D)/2} \cdot \sigma_y \Big/ (X_0\sqrt{n});$$
$$\Delta_{W_*}(X_0) = t_{(1-P_D)/2} \cdot \sigma_{*W} \Big/ \sqrt{n}. \quad (11)$$

Relative values of maximum errors are

$$\delta_y = \Delta_y(X_0) \Big/ \widehat{Y}(X_0);$$
$$\delta_W = \Delta_W(X_0) \Big/ \widehat{W}(X_0); \quad (12)$$
$$\delta_{W_*} = \Delta_{W_*}(X_0) \Big/ \widehat{W}_*(X_0).$$

We will determine the sequence of implementation of the proposed approach on a specific example. Monthly indicators of electricity consumption by a coal mine Y_i and corresponding volumes of production X_i ($i = 1 \ldots n$, $n = 12$) are given in the Table. In addition, the Table also contains values of specific energy consumption $W_i = Y_i/X_i$. It is necessary to make a selection of the regression dependence, considering options (5) or (6), to estimate the normalized value of the specific consumption of electrical energy.

To deepen the understanding of the applied approach, consider the correlation fields of variables Y_i and X_i, W_i and X_i, as well as dependencies $\widehat{Y}(X)$, $\widehat{W}(X)$, $\widehat{W}_*(X)$ in the options considered. We will apply the matrix method for calculating coefficients G_0, G_1 and Q_0, Q_1 regression equations (4, 6) using the method of least squares. We will get $G_0 = 48.2$, $G_1 = 9.1$, $Q_0 = 11.1$, $Q_1 = -0.021$. The Figure illustrates correlation fields of variables as well as regression dependencies. The Figure shows that the obtained regression dependences correspond to the correlation fields of the considered variables. The location of points W_i and X_i in the Figure, *b* caused the use of a segment of non-linear dependence $\widehat{W}(X)$, close to linear. Joint consideration of the correlation field of variables and possible options for applying regression dependencies helps to make a reasonable choice of the equation, taking into account possible regime changes of the technological process (primarily the volume of production) in the future. Changing the value X_i can be used in the process of energy efficiency management if this solution is acceptable taking into account the existing restrictions.

Let us determine the statistical characteristics of the regression dependencies obtained in the given example. The calculations were performed at the following parameter values: $P_D = 0.95$, $n = 12$, $t(1 - P_D)/2 = 2.228$, $X_0 = 51$ thousand tons. For dependence (4) we obtained: coefficient of determination $R^2 = 0.95$, $\Delta_y(X_0) = 10.25$ thousand kWh, $\delta_y = 2$ %. For formula (5): $R^2 = 0.18$, $\Delta_W(X_0) = 0.2$ kWh/t, $\delta_W = 2$ %. For formula (6): $R^2 = 0.155$, $\Delta_{W_*}(X_0) = 0.233$ kWh/t, $\delta_{W_*} = 2.3$ %. It can be seen that when using dependencies (5, 6) as a result of the calculation, close values of statistical characteristics were obtained. The obtained result does not prove the advantage of using one of the considered dependencies. It is important that the calculation results confirm the possibility of obtaining energy consumption indicators acceptable for the practice of normalizing their determination errors even with a small amount of experimental data ($n = 12$). This is mostly ordered by the fact that the error value is determined by the mean value of the random variable. Under these conditions, it is advisable to perform the necessary calculations with the following evaluation of the obtained result.

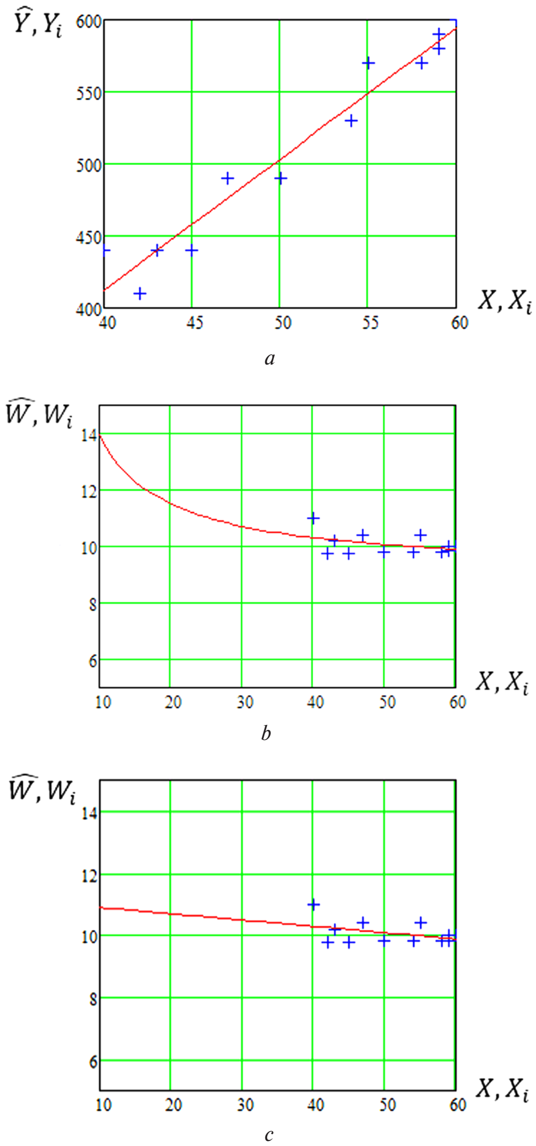

Fig. Correlation fields of variables and regression dependencies:

a – correlation field Y_i and X_i, regression dependence (4); b – correlation field W_i and X_i, regression dependence (5); c – correlation field W_i and X_i, regression dependence (6)

Low values R^2 for dependencies (5, 6) indicate the presence of explanatory variables not taken into account in the models, which determine the change in the levels of specific

Table

Statistical sampling of the mine's monthly coal production X_i (thousand tons), corresponding energy consumption Y_i (thousand kWh), specific costs W_i (kW h/t)

Month	January	February	March	April	May	June	July	August	September	October	November	December
X_i	60	40	50	55	47	43	54	45	42	59	58	59
Y_i	600	440	490	570	490	440	530	440	410	590	570	580
W_i	10	11	9.8	10.4	10.4	10.2	9.8	9.77	9.76	10	9.82	9.83

energy consumption, independent of the volume of production. Taking into account the fact that the normalization of energy consumption is a periodic process related to the control and management of energy use, the experimental data for constructing regression dependencies change periodically. The policy of energy saving at enterprises consists in strengthening the connection of energy consumption with the volume of production. Therefore, with the successful implementation of this policy, an increase in values should be expected over time R_2 in the future. Periodic calculations of values R_2 during the performance of the normalization operation make it possible to monitor the dynamics of their changes and can serve as a criterion for evaluating the effectiveness of the implementation of the energy-saving policy at the enterprise.

The parameterization of the linear model must take into account the conditions for obtaining the initial data for the calculation of the model parameters. All components of the vector of explanatory variables X_t $(X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(p)})$ must be measured with sufficient accuracy at specified time intervals t, which are determined by the periodicity of energy efficiency control. In the conditions of operational control, the collection of a significant amount of data related to the measurement of the components of the vector X_t, is a difficult task, since the value $X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(p)}$, , as a rule, are not registered in an automated mode (unlike Y_t, the registration of the values of which is carried out by AMR). With daily monitoring of energy efficiency, the task of manual information collection becomes somewhat more complicated. Due to the special importance of the production volume X to assess the efficiency of industrial enterprise, registration X is usually performed.

An attempt to abandon the mandatory measurement of the values of explanatory variables is outlined in (Ayvazyan, 2010). The essence of the approach is that action of some qualitative factors on explanatory change Y_t is considered rather than quantitative factors related, for example, to the seasonality of works. This action does not change the type of regression dependence. The values of the regression coefficients are subject to change Q_0, $Q_1, \ldots, Q_p$ (3). To account for this influence, dummy variables are introduced into the regression equation Z, which can take the value 1 or 0 depending on the presence or absence of influence. It is known that the statistical reliability of the estimation of the explanatory variable according to such equations increases as the ratio "$n/p + 1$" increases [1, 2]. An increase in the number of "n" measurements can be made by obtaining information from several control objects of the same type (spatial information collection option). Since the control of the efficiency of electricity consumption is planned to be carried out within the framework of one management object, which determines the collection of information in certain time intervals (the option of collecting information in time), the significant increase in the number of measurements "n" that meet the condition of minor structural or regime changes on objects of management, it is practically impossible, based on the peculiarities of control. Hence the expediency of reducing the number of "p" factors in the existing conditions of experimental data collection is clear. At the same time, the introduction of dummy variables helps to increase the value of "p". Often, the introduction of new explanatory variables into the regression model requires taking into account their interaction. Pair interaction is usually taken into account as a product of the values of these factors [11]. It is clear that their introduction into the regression equation also increases the value of "p" and the number of coefficients to be determined. Since structural and mode changes in the operation of the power supply system are a common phenomenon, it is necessary to update the data used to construct the regression dependence [7, 8]. Daily measurements of values X_i, Y_i allow using them both for performing the energy efficiency control procedure (both the value of the actual daily electricity consumption and the volume of production) and for constructing the regression function, which will be used in the following stages of control [7]. The duration of the set is important X_i, Y_i to determine a new function depending on the dynamics of structural and mode changes occurring in the power supply system. Focusing on sufficiently frequent changes, as well as on the expediency of increasing the number of measurements, it is proposed to collect statistical data for constructing a regression dependence within a month. Then, with daily energy efficiency control, the value of "n" is equal to the number of working days in a month. A monthly term for the formation of an updated regression model that corresponds to the changes that occurred at the object of control is an acceptable option from the point of view of the possibility of comparing the characteristics of the existing and updated models, clarifying the dynamics of their changes, and evaluating the effectiveness of management decisions in the reporting monthly period. The conducted analysis shows the expediency of reducing the number of explanatory variables "p". Given that value X is often sufficient to characterize the level of total energy consumption associated with production, it is proposed to form regression models to control energy efficiency in the form of one-factor linear dependencies. Such are the models (4, 6) considered above.

The proposal to reduce the number of explanatory variables also has negative consequences, as some influencing factors will not be taken into account, which will ensure an increase in the random component δ_t in formula (2). The absence of additional explanatory variables will lead to a significant variation of the coefficients Q_0, Q_1 and G_0, G_1 regression equations (4, 6) when they are built based on the results of different samples, that is, the situation comes down to considering models with a variable structure (Ayvazyan, 2010). A possible variant of the formation of such models is the distribution of experimental data of the initial sample into samples with homogeneous data, that is, data that ensure the relative stability of the coefficients of the regression equation. It is obvious that the proposed collection of information in a monthly period must meet the condition of homogeneity of the data used to construct the regression dependence. The planned periodic update of the regression model to be used is accompanied by a change in the coefficients of equations (4, 6). This ensures the reduction of the random component δ_t, making it possible to increase the accuracy of estimating the predicted value of electric energy consumption.

Limiting the number of explanatory variables in the regression model narrows the informativeness of the control process, as there is no quantitative assessment of the impact of unaccounted factors. Consideration in models (5, 6) of the target factor X solves the task of a quantitative assessment of its impact. Some of the factors that were not reflected in the regression model can be classified as controlled. It is obvious that the assessment of their impact is also essential for solving management tasks. It is possible to make such an assessment thanks to the daily energy consumption control proposed for implementation. The control procedure, which is carried out often enough, allows you to attribute recorded changes in daily energy consumption (changes relative to normalized values) to a specific list of changes that occurred during the reporting day in the structure or modes of energy consumption of the object. Registration of these changes by service personnel will allow for determining both the list of factors not taken into account in the model and assessing the degree of their influence. Such an assessment is especially important for determining the activity of service personnel in the implementation of energy-saving tasks.

The recommendations set out in the article should not be considered as an alternative list of necessary actions. The authors see rational grains in the proposed parameterization option, which can be fully or partially used in a specific situation.

Conclusions.

1. Focusing on the presence of frequent and significant changes in the energy efficiency of the production process of

the control object, it is proposed to use regression models with a variable structure to normalize the specific energy consumption, where each subsequent sample of homogeneous data provides a change in the regression coefficients. Daily actual energy consumption indicators registered during the month are used both to control the efficiency of energy use and build a regression model of the variable structure.

2. The limited possibilities for forming voluminous statistical samples with homogeneous data, and the complexity, and laboriousness of measuring a significant number of influence parameters make it necessary to reduce the number of explanatory variables of the regression model. It is proposed to focus on the use in the model of the value of the output volume as a comprehensive characteristic of the level of energy consumption.

3. Special conditions for controlling the efficiency of electricity consumption by an industrial enterprise determine the acceptability of applying linear and non-linear one-factor regression dependencies in the problem of normalizing specific energy consumption. The nonlinear model is proposed to be defined as a result of reducing the linear model of energy consumption to a nonlinear form characteristic of the values of its specific consumption.

4. Acceptable values of errors in the calculation of normalized energy consumption indicators using univariate regression dependencies can be obtained even with a limited amount of experimental data, which indicates the expediency of performing calculations in these conditions with the subsequent assessment of the accuracy of the obtained results. The calculation method is described.

5. It is proposed to compensate for the decrease in the informativeness of the energy efficiency control process due to the decrease in the number of explanatory variables of the regression model by attributing existing changes in daily electricity consumption to changes that occurred during the day in the structure or modes of energy supply of the object.

References.

1. Verkhovna Rada of Ukraine (2021). *About energy efficiency. 1818-IX§ section I Article 5.* Retrieved from https://zakon.rada.gov.ua/laws/show/1818-20#Text.

2. Government portal (2016). *Ukraine has adopted national standards for energy audit and energy management in accordance with European norms*. Retrieved from https://www.kmu.gov.ua/news/249113427.

3. *ISO 50001:2018, IDT. Energy Management Systems – Requirements with Guidance for Use – Guidelines* (2018). Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:50001:ed-2:v1:en.

4. *Implementation of the standard of energy management systems in the industry of Ukraine: UNIDO/GEF Project* (2015). Retrieved from http://www.ukriee.org.ua/uk/proekt/meta-proekta.

5. *ISO 50001 Benefits for Manufacturers* (2019). Retrieved from https://www.plantengineering.com/articles/iso-50001-benefits-for-manufacturers/.

6. Palekhova, L., & Simon, S. (2016). Competitive advantages through the implementation of international energy management standards. *Bulletin of the Dnieper State Academy of Construction and Architecture, 3*, 42-51.

7. Dreshpak, N.S. (2020). Energy efficiency control systems of production processes and ways of their improvement. *Electrical Engineering and Power Engineering, 1*, 40-48. https://doi.org/10.15588/1607-6761-2020-1-5.

8. Dreshpak, N.S., Dreshpak, O.S., & Vypanasenko, S.I. (2021). Specific standards of energy consumption in the problem of controlling its use efficiency *Electrical Engineering and Power Engineering, 3*, 31-39. https://doi.org/10.15588/1607-6761-2021-3-3.

9. Shulle, Yu.A., & Rogozyanskmiy, I.S. (2016). The use of AMR to increase the efficiency of energy use at industrial enterprises. *Information Technologies and Computer Engineering, 1*, 59-63.

10. Verkhovna Rada of Ukraine (2018). *Code of commercial accounting of electric energy. No 311§, Section I 1.2.* Retrieved from https://zakon.rada.gov.ua/laws/show/v0311874-18#Text.

11. Dreshpak, O.S., Dreshpak, N.S., & Vypanasenko, S.I. (2022). *Technology of Raw Materials Enrichment of Inhomogeneous Carbonate Deposits and Evaluation of its Energy Efficiency: Multi-authored: monograph,* (pp. 194-214). Romania: UNIVERSI-TAS Publishing. https://doi.org/10.31713/m1107.

12. Chung, S., Park, Y., & Cheong, T. (2020). A mathematical programming approach for integrated multiple linear regression subset selection and validation. *Pattern Recognition, 108*, 1-25. https://doi.org/10.1016/j.patcog.2020.107565.

Параметризація статистичної моделі контролю ефективності споживання електричної енергії

*Н. С. Дрешпак**, *О. С. Дрешпак*

Національний технічний університет «Дніпровська політехніка», м. Дніпро, Україна

*Автор-кореспондент e-mail: dreshpak.n.s@nmu.one

Мета. Обґрунтування структурної побудови та параметрів регресійної моделі для унормування питомих витрат енергії при виконанні контролю енергоефективності виробничого процесу.

Методика. Аналіз особливостей контролю енергоефективності виробничого процесу в умовах частих і суттєвих змін питомих витрат енергії з наступним визначенням структури та параметрів регресійної моделі.

Результати. Виходячи з наявності частих і суттєвих змін енергоефективності виробничого процесу об'єкта контролю, обґрунтована доцільність нормування питомих витрат енергії шляхом використання регресійної моделі зі змінною структурою. Щоденні фактичні показники питомого енергоспоживання, отримані протягом місяця, використані як для контролю ефективності використання енергії, так і для побудови регресійної моделі змінної структури. Обмеженість можливостей для формування об'ємних статистичних вибірок з однорідними даними, складність і трудомісткість вимірювань значної кількості параметрів впливу зумовлюють необхідність скорочення кількості пояснюючих змінних регресійної моделі. Доведена доцільність використання в моделі значення обсягу випуску продукції як комплексної характеристики рівня витрат енергії. Визначена прийнятність застосування лінійної та нелінійної однофакторних регресійних залежностей. Нелінійна модель отримана в результаті приведення лінійної моделі витрат енергії до нелінійного виду, характерного для значень її питомих витрат.

Наукова новизна. Уперше запропоноване використання регресійної моделі змінної структури для унормування питомих витрат енергії в умовах частих і суттєвих змін енергоефективності виробничого процесу, що сприяє підвищенню точності їх визначення. Доведена необхідність скорочення кількості пояснюючих змінних регресійної моделі. Підтверджена доцільність застосування в даних умовах контролю енергоефективності лінійної або нелінійної однофакторних регресійних залежностей, що сприяє спрощенню процедури реєстрації вихідних даних для їх побудови.

Практична значимість. Наукові результати виконаних досліджень дозволяють ураховувати особливості умов виробництва продукції при визначенні структури та параметрів регресійної моделі для унормування питомих витрат енергії. Це сприяє підвищенню точності та оперативності контролю енергоефективності виробничого процесу.

Ключові слова: *контроль енергоефективності, питомі витрати енергії, регресійна залежність, параметризація моделі, методи унормування*

The manuscript was submitted 21.12.22.