ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 10, Issue 2, February 2023

Forecasting Methods in Energy

Nazirova X.Z, Nazirova O.Z., Zayniddinova M.A.

Tashkent State Technical University named after Islam Karimov, Tashkent, Uzbekistan

ABSTRACT: The article discusses the issues of forecasting power consumption based on statistical methods that provide for the possibility of using an unlimited number of different input parameters, and the function of the influence of the input parameter not the output result can be of any complexity-linear, non-stationary.

I.INTRODUCTION

The main task of managing the EPS modes is to meet the needs of economic sectors and the population with high-quality electric energy and heat in the short and long term at minimal cost, ensuring the energy efficiency of their production and consumption.

In the context of scientific and technological progress and improvement of the energy system of the state, forecasting becomes one of the decisive scientific factors in the formation of the strategy and tactics of energy development.

Modern conditions require further improvement of the methodology and methods of energy forecasts. The higher the level of forecasting the development of energy processes, the more effective the planning and management of these processes in society [1].

The effectiveness of decisions to develop infrastructure and improve energy security is determined by the objectivity and validity of information obtained from forecasts. Planning for the development of energy systems in the system of territorial planning includes, as a basic element, the solution of economic issues of the development of the territory as a whole and its energy complex, in particular.

As practice shows, in the processes of territorial planning, the entire available arsenal of forecasting methods can be widely used. The development of informatics and computer technology creates the possibility of expanding the circle and improving the forecasting methods used.

The entire set of forecasting methods, including in the energy sector, can be divided into three main classes:

statistical (extrapolation of trends), based on the dissemination of conclusions made as a result of studying the state of the object of forecasting in the past and present, to its state in the future;

expert assessments based on statistical processing of expert opinions about the object of forecasting;

modeling, which consists in building a model of the object under study in order to analyze its behavior in the future.

II.GEO SCATTERED TYPE BIG DATA IN APPLICATION

Statistical forecasting methods are the most common and are based on the construction and analysis of time series that characterize the behavior of the system or forecasting object in the past and present, and extrapolation of the resulting trends for the forecast period [2].

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 10, Issue 2, February 2023

Statistical data serve as the initial information for the forecast. Statistical dependencies are approximated by known analytical functions - regression equations. The nature of the regression equation is determined on the basis of general ideas about the type of relationship or by direct enumeration of known functions:

Linear:

$$y = a + bt; \quad (1)$$

parabolas of the second degree:

$$y = a + bt + c \quad (2)$$

simple exponential (exponential):

$$y = abt; \quad (3)$$

modified exponential:

$$y = a - bc; \quad (4)$$

$$y = at; \quad (5)$$

where a, b and c are the coefficients of the regression equations.

Coefficients a, b and c in (1-5) can be obtained using the least squares method.

Energy forecasting is the process of developing energy forecasts based on scientific methods for understanding phenomena in the energy sector and using the entire set of methods, means and methods of forecasting [3].

- Analytical and computational forecasts are based on the analysis of the dynamics of available statistical indicators and various calculations carried out using: extrapolation, interpolation, processing of the dynamics of statistical series of indicators, correlation dependencies, comparison of indicators of the predicted object with other observations.

- Planning and settlement forecasts, the most typical type of which are forecasts for the implementation of the plan. The basis of planning and calculation forecasts, as a rule, is made up of various variants of initiating events.

- Forecasting using statistical methods. It is based, as a rule, on reliable initial data, the existing dynamics of the development of a certain area of research, the material level of production, resources, etc., identifying quantitative trends in the further development of the predicted object by applying statistical methods: extrapolation with its various modifications, interpolation, exponential analysis, etc.

- Forecasts, carried out mainly on the basis of expert assessments of specialists, including with the help of various types of questionnaires and their further processing, mainly with the help of mathematical statistics.

- Forecasts, the development of which is carried out simultaneously by statistical and expert methods (combined) using model engineering and the use of feedback to eliminate possible errors in the judgments of experts.

The use of forecasts of various types, their combination to create adequate judgments on the development of the energy industry in the future depends on the tasks that are set when developing a particular forecast, and to a certain extent determines the methods used to develop forecasts [5].

The most common groups of formalized forecasting methods include two subgroups: extrapolation and modeling. The first subgroup contains the methods of least squares, exponential smoothing, moving averages, adaptive smoothing. The second includes structural, network, matrix and simulation modeling.

Extrapolation is a method of scientific research, which is based on the dissemination of past and present trends, patterns, relationships to the future development of the forecasting object. In the narrow sense of the word, extrapolation is finding, from a series of data, a function of its other values that are outside this series. When evaluating the parameters of dependencies, the most common are the least squares method, the method of exponential smoothing of time series, the moving average method, and others.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 10, Issue 2, February 2023

To estimate the coefficients, the least squares method (LSM) is used more often than others. The essence of the LSM is to find the coefficients of the trend model that minimize its deviation from the original time series. The disadvantage of this method is that the trend model is rigidly fixed, and this makes it possible to use it only for short lead times, i.e. in short-term forecasting [10].

The method of exponential smoothing of time series is a modification of the least squares method for the analysis of time series, it is the averaging of the values of the time series. It can be performed according to different methods: as arithmetic or geometric means, by an even or odd number of points. The smoothing procedure is aimed at minimizing random deviations of the points of the series from some smooth curve of the assumed trend of the process.

This method allows you to evaluate the parameters of the model that describes the trend that was formed at the end of the base period and does not just extrapolate the current dependencies into the future, but adapts, adapts to conditions that change over time. The exponential smoothing method is used for short- and medium-term forecasting.

Adaptive methods include Brown, Holt and Holt-Winters methods. It is based on the model of a recursive harmonic process proposed by J.Yul.

Adaptive methods make it possible to take into account the different information value of the levels of the time series, the degree of "obsolescence" of the data. This property of adaptive methods is their essential advantage for predicting power consumption.

The initial construction of the predictive model is made on the basis of the first few observations of the object, a forecast is made, which is compared with the actual data. Based on the results of the forecast, the model is corrected, then a forecast is made for the next observations, and so on until all observations are exhausted. Adaptation is carried out iteratively with the receipt of each new actual point of the series. The model constantly "absorbs" new information, adapts to it and therefore reflects the development trend that exists at the moment.

Estimation of coefficients of the adaptive model is usually carried out on the basis of the recursive method. Thus, it is possible to obtain an accurate forecast for a larger interval than, for example, using the exponential smoothing method [15]. However, this is true only for very long time series. Currently, there is no methodology that calculates a sufficient amount of initial information for such models, which greatly complicates the process of their implementation.

Extrapolation is based on a number of assumptions: the development of the phenomenon can be reasonably characterized by a smooth trajectory-trend; the basic conditions that determine the trend of development in the past do not undergo significant changes in the future.

Thus, extrapolation in one form or another is widely used in predicting energy development trends. When forming forecasts using extrapolation, as a rule, they proceed from the trends in changes in certain quantitative characteristics of an object, which are statistically composed.

Estimated functional systemic and structural characteristics are extrapolated. Extrapolation methods are one of the most common and most developed among the totality of forecasting methods.

For the analysis, modeling and forecasting of regime parameters and characteristics of the EPS, it is advisable to apply the methods of direct construction of models based on observational data (statistics). Such methods make it possible to reveal implicit cause-and-effect relationships and patterns hidden in retrospective data and present them in an explicit form of mathematical models.

One of the most common methods for obtaining electricity consumption forecasts is the construction of a multiple regression model based on the least squares method.

With the help of correlation analysis, the presence of statistically significant relationships between variables is revealed and the degree of their tightness is assessed. Paired correlation coefficients characterize the relationship between two selected variables against the background of interaction with other indicators and are the most common indicators of the closeness of the relationship in statistical data analysis.

Then they proceed to the mathematical description of a specific type of dependency using regression analysis. To do this, a class of functions is selected that connects the effective indicator Y and the arguments X1, X2, X3, ... Xk,

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 10, Issue 2, February 2023

the most informative arguments are selected, estimates of unknown values of the parameters of the link equation are calculated, and the accuracy of the resulting equation is analyzed [24].

Correlation and regression methods take into account the influence of a wide range of input parameters (meteorological factors, the formation of morning and evening maxima, etc.) on the output forecast data of power consumption. The construction and evaluation of the regression model is carried out in Excel using the regression module of the data analysis package.

The simplest version of the regression model is linear regression. It is based on the assumption that there is a discrete external factor X(t) that affects the process Z(t), while the relationship between the process and the external factor is linear, but in practice the process Z(t) is influenced by a number of discrete external factors. factors X1(t),…,XS(t), such a model is called multiple. The disadvantage of this model is that in order to calculate the forecast of the process Z(t), it is necessary to know the future values of all factors X1(t),…,XS(t), which is almost impossible.

Correlation and regression methods predict the behavior of a variable based on the temporal relationship between it and another variable, which can be expressed as a statistical relationship called regression or correlation. In other words, these methods make it possible to establish the dependence of a change in one variable in the event of a change in another by a certain amount.

Regression analysis explores the relationship of a dependent variable from other independent ones, it is used if there is a connection between the predicted process and any factors that affect it. Linear regression is usually used for long term forecasts, but can also be used for shorter term forecasts. In the process of work, it is natural that new information appears, which requires correction of the data of the predictive model. Such an adjustment is a very complex and time-consuming procedure that requires a large amount of calculations.

This leads to the fact that a dynamic change in the characteristics of the predictive model is often impossible. Sometimes, the requirements of the predictive model for the initial information for real observations are not feasible, and therefore the forecast is inaccurate and ineffective.

When constructing a statistical relationship based on regression and correlation analysis, a statistical sample is used to evaluate the coefficients in the regression equation, and the choice of the type of function and informative features is carried out by the researcher himself.

An important stage of forecasting is the assessment of the adequacy of the model. When developing a forecasting model, the initial information is divided into two parts, one of which covers earlier data, and the other - more recent. Using the data of the first group, the parameters of the forecast model are evaluated, and the data of the second group are considered as the actual data of the predicted indicator. After developing a forecasting model, the resulting forecast is compared with the data of the second group to assess the adequacy of the developed model, i.e. compliance with actual statistics. In addition, when several forecasting models are developed, it is not always obvious which one is better or more accurate.

Various criteria are used to assess the accuracy of models.

When working with regression models, the verification procedure is adopted by performing the following steps:

the first stage is to check the statistical significance of the obtained coefficients of the regression equation;

the second stage is to check the overall quality of the constructed regression equation;

the third stage is the verification of data properties, the feasibility of which was planned when evaluating the equation.

In order to assess the significance of the regression equation as a whole, Fisher's test is used.

To determine the adequacy of the obtained regression equation for the process under study, one of the following options is possible:

1. The constructed model based on the Fisher F-test is recognized as adequate and all regression coefficients are significant. Such a model is considered effective, and it can be used to make management decisions and create forecasts.

2. The constructed model according to the Fisher F-criterion is generally adequate, but some of the coefficients are not significant. The model is suitable for making some decisions, but such a model cannot be used for making forecasts.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 10, Issue 2, February 2023

3. The constructed model according to Fisher's F-criterion is generally adequate, but all regression coefficients are not significant. Such a model is recognized as completely inadequate and it is sent for additional research. Based on such a model, it is impossible to make decisions and make forecasts.

Fisher's criterion is preceded by analysis of variance. Analysis of variance in mathematical statistics can be used as an independent tool for statistical analysis. In econometrics, analysis of variance is more often used as an additional tool for studying the quality of the constructed regression model. The principle of the main idea of the analysis of variance: the total sum of the squared deviations of the variable (y) from the mean value is decomposed into two components: "explainable" and "unexplained"

$$\sum(y-\bar{y})^2=\sum\left(\hat{y}_x-\bar{y}\right)^2+\sum\left(y-\hat{y}_x\right)^2 \quad (6)$$

where is the average value of a series of calculated values. The ratio of the explainable part of the variance (variable y) to the total variance is called the coefficient of determination and is used to assess the quality of the resulting regression equation. The ratio between the explainable and unexplained parts of the total variance can be calculated using an alternative formula:

$$R^2=\frac{\sum(\tilde{y}_x-\bar{y})^2}{\sum(y_i-\bar{y})^2}=1-\frac{\sum(y_i-\tilde{y}_x)^2}{\sum(y_i-\bar{y})^2} \quad (7)$$

The coefficient of determination R2 can take values in the range from zero to one $0\leq R2\leq 1$. With the help of the coefficient of determination R2, it is determined what part of the variance of the resultant result (y) could be explained by the regression equation. The higher R2, the better the regression equation describes the original data and the greater part of the variance of the resulting feature (y) explained the regression equation. If, as a result of the study, there is no relationship between (y) and (x), the coefficient of determination R2 will be close to zero. Thus, the coefficient of determination R2 can be used to assess the quality of the resulting regression equation. The use of the coefficient of determination R2 to assess the quality of the model has a drawback, if it is necessary to include a new factor (even an insignificant one) in the model, then this automatically increases the value of R2.

The analysis shows that none of the existing methods can give sufficient accuracy of forecasts for 20-25 years. The extrapolation method used in forecasting does not give accurate results for a long-term forecast, since this method proceeds from the past and present and thereby accumulates an error. This method gives positive results in the short term forecasting of certain objects - for 5-7 years.

The method of group accounting for arguments (GMCA) differs from those considered in that, using the idea of heuristic self-organization of a small sample of experimental data, it allows choosing the type of the approximating function and the arguments included in it.

The essence of the method lies in the gradual selection - enumeration of prediction equations that meet the conditions of the required accuracy of the results given experimental data. The choice comes from a set of candidate models according to a given criterion [31]. In other words, there is a directed selection according to certain criteria - forecasting accuracy, the greatest stability of the resulting model, etc.

To build a mathematical model of optimal complexity using GMDH, the initial experimental sample is divided into two sequences - learning (A) and testing (B). The learning sequence is used in a conventional regression analysis to optimize the estimates of the coefficients of an equation using a minimum standard error test. The check sequence serves to select the number of terms and construct the regression equation by minimizing the selection criteria.

The GMDH is based on a scheme according to which the step-by-step selection of mathematical models of processes is carried out, which, as a rule, leads to the choice of the optimal one that best describes the process under consideration. The algorithm has a probabilistic character, i.e. the probability of obtaining a better solution increases with the number of selections. In this case, the choice of the support functions of the method is decisive.

The following are mainly used as support functions in the formation of GMDH algorithms:

Various power polynomials of the form:

$$y=a_0+\sum_{i=1}^{n} aix^t \; ; \quad (8)$$

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 10, Issue 2, February 2023

Harmonic or logistic:

$$y= a_o+\sum_{i=1}^{n}(\frac{ai}{1}+\exp(-xi))\,; \quad (9)$$

Exponential:

$$y= \sum_{i=0}^{n} aie^{\beta it}\ ; \quad (10)$$

where the nature of the problem being solved is determined.

Other, combined types of functions, logical networks, fuzzy Zadeh sets, etc. are also used.

Conclusions. The importance and relevance of improving forecasting methods in the conditions of market relations in the energy sector is increasing due to the need to improve the accuracy of planning EPS modes, both in short-term and long-term planning, and improve the efficiency of the electrical system as a whole.

1. Application for short-term and long-term forecasting of the parameters of the mode and characteristics of the EPS of methods and principles of self-organization, closely adjacent to the EPS of Uzbekistan. In this regard, the use of the method of group accounting of arguments (GMCA) is especially effective, the basis of which is the use of natural selection - the selection of the optimal and the only mathematical model of the process under study from a variety of candidate models.

2. Comparison of GMDH with other models using regression methods, in particular, with the least squares method (LSM), shows its high computational efficiency, which is expressed in an increase in the accuracy of both short-term and long-term forecasting.

REFERENCES

1. Allaev K.R. Modern energy and prospects for its development. –T.: Fan va texnologiyalar nashriyot-matbaa uyi, 2021.- 952 p.
2. Allaev K.R. Power industry of Uzbekistan and the world.-T.: Fan va texnologiya, 2009.- 464 p.
3. Tikhonov E.E. Forecasting methods in market conditions: a textbook.- Nevinnomyssk, 2006.-221p.
4. Birman E.G. Comparative analysis of forecasting methods//NTI. Series 2 - 1986. - No. 1. – P. 11–16.
5. Ivakhnenko A.G. Long-term forecasting and control of complex systems.- Kyiv: Technique, 1975.-311 p.
6. Bogdanov V.A., Kochkarev Mathematical model of operational forecasting of active loads of power systems.// Electric stations, 1974. No. 4. S.22-24.
7. Lowdy D., Maxwell A. Factor analysis as a statistical method.-M.: Mir, 1967.
8. Ivakhnenko A.G., Kroto G.I., Cheberkus V.I. Multi-row algorithm for self-organization of long-term forecasts.// Avtomatika. 1980. No. 4. pp. 28-47.
9. Korovin G.B., Malyshev E.A. Forecasting the development of regional energy systems // Economics of the region. 2011. No. 2 (26). pp. 184–188.
10. Malyshev E.A., Sokol-Nomokonov E.N., Sokol-Nomokonov V.E. Planning and forecasting the development of regional energy systems: the formation of vertically integrated territorial-branch complexes for planning and managing the development of regional energy systems: monograph: Chita: ChitGU, 2010. 133,S.
11. Romanova O.V. Conditions and factors of structural modernization of the regional industrial system // Economics of the region. 2011. No. 2 (26). pp.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 10, Issue 2, February 2023

AUTHOR'S BIOGRAPHY

№	FULL NAME PLACE OF WORK, POSITION, ACADEMIC DEGREE AND RANK	PHOTO
1.	**Nazirova Khilola Zakhidzhanovna,** Senior lecturer of the department "Energy efficiency and energyaudit", Tashkent state technical university	
2.	**Nazirova Ozoda Zakhid Qizi,** Senior lecturer of the department "Foreign Languages" Tashkent state technical university	
3	**Zayniddinova Madina Ahmad Qizi,** Student "Energy efficiency and energyaudit", Tashkent state technical university