DEVELOPMENT OF AN ALGORITHM FOR EVALUATING THE DOMINANT FACTORS THAT HAVE THE GREATEST IMPACT ON THE ENERGY INTENSITY OF PRODUCTS

[1]I U Rakhmonov*, [1]F A Hoshimov

[1]Tashkent state technical university, Tashkent.

Abstract. The article deals with the assessment of the dominant factors that have the greatest impact on the energy intensity of products. A method is proposed that makes it possible to single out the most significant factors from all the variety of factors affecting energy indicators and to give an appropriate assessment to each of them. Methods and mathematical models are proposed that can be used for current and medium-term forecasting of electricity consumption. The analysis is carried out and the energy characteristics of production units are developed, depending on the reduced and total volume of manufactured products. Using linear programming methods, the energy intensity coefficients of the entire range of products manufactured by individual production units are obtained.

Introduction

Development of an algorithm for evaluating the dominant factors that have the greatest impact on the energy intensity of products In addition to a study of the impact of each factor individually, it is also necessary to assess their cumulative impact, since in the production environment all factors are interconnected and operate simultaneously.

Main part

Dispersion, correlation and regression analysis methods are mainly used to solve multivariate statistical problems. Dispersion and correlation methods ensure the determination of the degree of influence of various factors on each other and on dependent variables, while regression methods make it possible to analytically present the nature of the regularities of the influence of these factors and to evaluate and forecast the results for the future period. The multiple regression equation is generally as follows:

$$\overset{\Lambda}{Y}_{x_1 \ldots x_2 \ldots x_n} = B_0 + B_1 X_1 + B_2 X_2 + \ldots + B_n X_n \qquad (1)$$

Where, B_0 – free member of the equation;
B_I– coefficient i=1÷n;
X_I – factors i=1÷n;
n – number of factors.

If the class of functions describing the phenomenon under study is not known, the type of relationship is determined empirically by selection, building a number of functions and statistical reliability using the multiple correlation coefficients, F - criterion, residual dispersion and relative approximation error [1-5].

In empirical function selection, the regression equation is initially represented by a Taylor segment. For practical tasks, it is usually limited to a second-degree polynomial:

$$\overset{\Lambda}{Y} = A_0 + \sum_{i=1}^{n} A_1 X_1 + \sum_{k=1}^{n} \sum_{i=1}^{n} A_k X_i X_k \qquad (2)$$

Where, A_0 – free member of the equation;
A_k, A_i – coefficient ;
X_k, X_i–factors.

The second-order terms X_k, X_i for $i \neq k$ are included in the regression equation (2) in order to take into account the so-called effects of the combined action of factors of arguments on the dependent variable (Y), and the terms X_k, X_i for $i = k$ (i.e. X_2) take into account the nonlinearity of the change in the dependence of the variable (Y) when changing the i-th argument [6-8].

The process of finding a significant segment of the Taylor series equation is carried out as follows.

First, all unknowns in the first degrees are included in the model and the resulting modules are assessed by the F-criterion with a relative approximation error. If the assessment of the F - criterion turns out to be insignificant, then the model includes the values of the unknowns in the second powers (paired products). The process of increasing the degree of the polynomial continues until the estimate of the equation becomes significant.

However, this method has significant disadvantages:

- the number of coefficients of the regression equation when using even a second degree polynomial (2) grows very quickly with an increase in the number of arguments;
- with a not very large sample size, the random arrangement of points in space can lead to the fact that some random (false) connections characterizing the effects of the relationship will also be significant;

* Corresponding author: ilider1987@yandex.com

- when studying the regression equation given in the form (2), a circumstance arises that complicates the application of the least squares method. In this regard, the equation (2) is preferred to be set in the following form:

$$\hat{Y} = A_0 + \sum_{i=1}^{n} A_i \bar{X}_i + \sum_{i=1}^{n}\sum_{k=1}^{n} A_{ik}(X_i - \bar{X}_i)(X_k - \bar{X}_k) \qquad (3)$$

Where, $\bar{X}_i, \bar{X}_k$ - average values *i*–th and и *k* – th parameters in the studied statistical population.

In addition, the complexity in integrating the results obtained causes computational difficulties, since the matrix of the system of normal equations becomes close to degenerate [9-11].

Research has established that of all the existing methods for solving the problem, it is advisable to use the regression step method of Efroimson.

This method is economical from a computational point of view and includes the advantageous aspects of the method of all possible regressions, the method of elimination, and also allows you to obtain an adequate mathematical model of the object under study, suitable for practical use.

The main idea of this method is to find regression with several variables in the form of a series of linear regression dependencies and to rebuild the correlation matrix step by step to the end. The method starts with a simple correlation matrix.

The variables (factors) are included in the equation in turn. The order of inclusion is determined using the partial correlation coefficient, as a measure of the importance of variables not yet included in the equation.

The variable X_i is selected (let's say it is x_i), which is most correlated with Y_i, and a linear (first-order) regression equation is found [12-14].

$$\hat{Y} = f(x_i)$$

Then the partial correlation coefficient xi ($i \neq 1$) is determined (taking into account the correction for x_i). Mathematically, this is equivalent to determining the correlation between:

a) residuals from regression $\hat{Y} = f(x_i)$;

b) residuals from another regression $X_i = f(x_i)$.

Then such a value x_k (let's say x_2) is selected, which possesses the above properties, and as a result of the calculation, the second regression equation is obtained $\hat{Y} = f(x_i, x_k)$. This process continues further.

At each stage, the variables included in the equation at the previous stages are additionally investigated.

A variable entered into an equation in a previous step may become redundant at a later time due to the relationship between this and other variables contained in the equation.

For verification, at each stage, the partial F_1 criteria for each variable of the equation are calculated and compared with the preselected F_T significance equation (the F value is the distribution according to the table). If $F_i < F_T$, then the i - th variable is excluded from the equation and the regression equation is recalculated taking into account the remaining variables. Otherwise, the equation remains unchanged. This relationship allows you to make a judgment about what contribution the i - th variable and the equation can make. A variable that makes a minor contribution is eliminated from the equation [15-16].

This process continues until minor variables are added and removed from the equation.

The construction of a mathematical model consists of the following stages.

1. A table with the dimension {n, m} is compiled from the values of the factors influencing the investigated energy indicator (e, W), where n is the number of measurements by factors, m is the number of factors determined by the method of expert assessments.

2. The available values of each factor are summed up:

$$S_m = \sum_{i=1}^{n} X_i \qquad (4)$$

$$S_m = \sum_{i=1}^{n} Y_i \qquad (5)$$

where , m - the number of factors;

n - the number of measurements by factors;

Y_i - investigated energy parameter (e, P, W).

1. The average value is determined by the following expression:

$$\bar{X}_r = S_m / n$$

$$\bar{Y}_r = S_{m+1} / n \qquad (6)$$

The standard of deviation is:

$$\sigma = \frac{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})}}{n} \qquad (7)$$

4. The adjusted sum of squares is calculated::

$$P_x = \sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)/n \qquad (8)$$

$$P_y = \sum_{i=1}^{n} Y_i^2 - (\sum_{i=1}^{n} Y_i)/n \qquad (9)$$

5. The unadjusted sum of the mixed works is:

$$S_i = \sum_{i=1}^{n} X_i Y_i \qquad (10)$$

6. A matrix of uncorrected sums of mixed products $X^{/} X$ is constructed, where $X^{/}$ is the transported matrix.

7. The correlation coefficients between each factor and the response are calculated:

$$r_{x_1,x_2} = \frac{\sum X_1 X_2}{\sqrt{(\sum X_2^1)(\sum X_2^2)}} \qquad (11)$$

$$r_{xy} = \frac{\sum\left(X_1 - \bar{X}\right)\left(Y_1 - \bar{Y}\right)}{\sqrt{\sum\left(X_1 - \bar{X}\right)^2 \sum\left(Y_1 - \bar{Y}\right)^2}} \qquad (12)$$

8. A matrix of correlation coefficients is constructed. The correlation matrix has dimension mxm, that is:

$$R=\begin{bmatrix} r_{11}, r_{12}, \ldots\ldots\ldots\ldots, r_{1m} \\ r_{21}, r_{22}, \ldots\ldots\ldots\ldots, r_{2m} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ r_{m1}, r_{m2}, \ldots\ldots\ldots\ldots, r_{mm} \end{bmatrix} \quad (13)$$

9. We expand the correlation matrix R as follows:

$$A=\begin{bmatrix} R(m\times m) & T'(m\times 1) & I(m\times m) \\ T(1\times m) & S(1\times 1) & O(1\times m) \\ -I(m\times m) & O(m\times 1) & O(m\times m) \end{bmatrix} \quad (14)$$

Where, $R(m\times m)$– correlation matrix for m - independent variables;

$T(1\times m)$– correlation vector for m - independent variables with deviation Y;

$T'(m\times 1)$– transformed vector T;

$S(1\times 1)$– the correlation coefficient of the response itself (1);

$I(m\times m)$– identity matrix;

$-I(m\times m)$– negative identity matrix. The elements of this matrix will be denoted A_{ij}.

10. To introduce a factor into the regression model. We calculate V_i which is determined by the formula:

$$V_i = r_{iy}\cdot r_{iy_1}/r = A_{n-1}\cdot A_{k-1}/A_n, \quad (15)$$

where $i=1\div k$.

11. The V_i with the maximum value is chosen, let's say it is V_i, then X_L is a factor to be considered [17-19].

12. The table is used to determine the standard FT test with a confidence level $P=0{,}95(n-1)$ and a degree of freedom (table F - distribution).

13. If X_L is the first variable included in the regression equation, then the value from the ANOVA table is determined as follows:

for general:

sum of squares $SS=r_{yy}^2=1$

degree of freedom $C_1=n-1$

for regression:

$$C_T=1; \qquad SS=r_{iy}^2$$

For the remainder $C_T=n-2; SS=r_{iy}^2$

$$M_S=\frac{SS}{(n-2)} \quad (16)$$

14. We calculate the sequential F - criterion:

$$F_T=\frac{M_{S.pez.}}{M_{S.ocm.}}, \quad (17)$$

where $M_{S.pez.}$ – the mean square due to the regression;

$M_{S.ocm.}$ – mean square due to the remainder.

15. If $F_T < F_{noc}$, then the XL factor is included in the regression.

At any stage of the procedure, the critical attitude for factor inclusion has a general formula:

$$\{\phi\cdot V_{\max}\}/d_{mm} - V_{\max}\cdot 3, \quad (18)$$

where d_{mm} – a table element that replaces S at any stage;

φ – the number of residual degrees of freedom after the inclusion of the factor.

16. Since the X_L factor is included in the regression, the correlation matrix must be transformed, for which the L-th row of the matrix is divided by and a second table is compiled, the elements of which are denoted by letters and they are determined as follows:

$$B_{L1}=\frac{A_{L1}}{A_{LL}},$$

$$B_{L2}=\frac{A_{L2}}{A_{LL}} \quad (19)$$

17. The rest of the table elements are calculated according to the following algorithm:

$$B_0 = A_n - \frac{A_n\cdot A_\eta}{A_{LL}}, \quad (20)$$

18. The proportion of the explained variation is determined:

$$D = M\cdot S_{pec}\cdot 100\%, \quad (21)$$

19. The standard deviation is:

$$S=\sqrt{M\cdot S_{ocm}}, \quad (22)$$

20. Standardized B– factor for X_L is:

$$B_1=\frac{\sum X_1Y_1-(\sum X_1)\left(\sum Y_1\right)/n}{\sum X_1^2-\left(\sum X_1\right)^2/n}, \quad (23)$$

21. B - coefficient for X_L in the original scale is determined by:

$$B_L = B_{L^m}\frac{S_m}{S_L}, \quad (24)$$

22. The standard error of the natural B– coefficient for X_L is:

$$S_{ош}=\frac{\sqrt{B_{дд}}}{\sum X_L^2}, \quad (25)$$

23. Using the partial correlation coefficients, as before, we choose as the next factor to include in the equation such a value X, which is characterized by the highest partial correlation coefficient with the response.

24. Determine for factors not included in the equation:

$$V_m=\frac{B_{im}\cdot B_{mi}}{B_{ij}}, \quad (26)$$

where $j\neq L, j=1\div M$

Determine the value of F - to include the factor X by the formula (3.30).
26. If F> F_{ij}, then the factor X_j is not included in the regression equation and the step method ends.
27. The free term of the equation at each step is calculated by the following formula:

$$B_0 = \overline{Y} - \sum_{t=1}^{k} B_t X_L, \qquad (27)$$

28. When F <F_n - th factor is included in the regression equation and repeated calculations are made starting from the 15th stage.
29. Since the factors X_j and X_L are included in the regression equation, particular F - criteria are determined. Partial F - criterion for X_L, in the presence of X_j in the model is determined by the formula:

$$F_{xL} = \frac{\phi(C_{Lm}+1)^2}{(C_{m+1},C_{m+1})(C_{m+2},C_{m+2})}, \qquad (28)$$

Where φ – residual number of degrees of freedom.
F – criterion for X_j in the presence of X_L is:

$$F_{xi} = \frac{\phi(C_{im}+1)^2}{(C_{m+1},C_{m+1})(C_{2m+1},C_{2m+1})}, \qquad (29)$$

30. For the factors L and j, the coefficients of variation, the deviation of the residuals, the standardized b - coefficients for X_j and X_L are determined according to the formulas specified in steps 17, 18, 19.
31. If F is a criterion for either the j-th or L-th factor is less than the critical one, that is, F_{xj} <F_{cr}, then the j-th factor is excluded from the regression equation.
32. Go to step 16.
33. When calculating V_{max}, the value of F is determined to include the factor. If some k-th factor has just been excluded from the regression dependence at the previous step and F for inclusion is more than F -critical, that is, $F_к$> $F_{кр}$, then the process of including and excluding factors in the model stops and the free term of the equation is determined by the formula (3.40).
34. The mathematical model of specific power consumption, depending on the factors obtained, can be expressed by the following equation:

$$e = B_0 \pm B_1 X_1 \pm .... \pm B_k X_k, \qquad (30)$$

where $B_1 - (i = 1 \div k)$, natural coefficients for included factors;
k – numbers of factors included in the model.
35. The adequacy of the resulting model is determined by substituting the values of the included factors into the regression equation:

$$Z = e - \grave{e}, \qquad (31)$$

where Z – остаток по модели;
e – raw data included for processing;
$\grave{e}$ – value after substitution of factor values into the model.
36. Determine the normal deviation using the following formula:

$$N = \frac{e - \grave{e}}{\phi} = \frac{Z}{\phi}, \qquad (32)$$

To assess the influence of a separate *i*-th factor on the specific power consumption included in the regression model:

$$e = B_o \pm \sum_{j}^{k} B_j X_j, \qquad (33)$$

with the constancy of other factors, i.e.

$$C = \sum_{j}^{k=1} B_j X_j = const,$$

where *k* – numbers of factors included.
In (33), we introduce the following transformation:

$$e = B_0 \pm C \pm B_i X_i, \qquad i = 1 \div k, \qquad (34)$$

When the i-th factor changes by β percentage, which is determined by the following formula:

$$\beta^0 = X_i \pm \frac{BX_i}{100}, \qquad (35)$$

Specific electricity consumption will change by β percent, which is determined by the following formula:

$$\beta' = \left(\frac{\beta^0 \beta_j + (B_0 \pm C)}{e} - 1 \right) \cdot 100\%, \qquad (36)$$

The given method makes it possible to single out from the whole variety of factors influencing energy indicators, the most significant ones and to give an appropriate assessment to each of them.
The above algorithm is unified for all the above calculation options and is included in the corresponding resulting calculations of the specific power consumption rates.

Conclusion

- The proposed methods and mathematical models can be used for current and medium-term forecasting of electricity consumption.
- The analysis has been performed and energy characteristics of production units have been developed depending on the given and total volume of products manufactured. Energy intensity coefficients for the entire range of products manufactured by individual production units were obtained using linear programming methods;
- the possibility of applying a mathematical method of experiment planning to calculate energy characteristics and identify technological factors that have a significant impact on energy indicators has been justified;
- a method has been proposed to assess the impact of the volume of yarn and spindle downtime on the energy performance of coiled, twisted, spinning and weaving production units

References

1. Rakhmonov, I.U., Niyozov, N.N. (2019) Optimization setting of steel-smelting industry in the issue of alloy steels *E3S Web Conf* **139** doi:10.1051/e3sconf/201913901077
2. Rakhmonov, I. U., Tovbaev, A.N., Nematov, L.A., Alibekova,T.Sh. (2020) Development of forecasted values of specific norms for the issues of produced products in industrial enterprises *Journal of Physics: Conference Series* **1515** doi:10.1088/1742-6596/1515/2/022050

3. Rakhmonov, I.U., Reymov, K.M., Najimova, A.M., Uzakov, B.T., Seytmuratov, B.T. (2019) Analysis and calculation of optimum parameters of electric arc furnace *Journal of Physics: Conference Series* **1399** doi:10.1088/1742-6596/1399/5/055048
4. Hoshimov, F.A., Bakhadirov, I.I., Erejepov, M., Djumamuratov, B. (2019) Development of method for normalizing electricity consumption *E3S Web Conf* **139** doi:10.1051/e3sconf/201913901074
5. Karimov R.Ch., Bobojanov M.K., Rasulov A.N., Usmanov E.G. E3S Web of Conferences, 139, 01039, (2019), doi.org/10.1051/e3sconf/201913901039;
6. Karimov R.Ch., Shamsiyev K., and others. IOP Conf. Series: Materials Science and Engineering, 883(1), 012142, (2020). doi:10.1088/1757-899X/883/1/012142;
7. E.G.Usmanov, A.N.Rasulov, M.K.Bobojanov, R.Ch.Karimov. E3S Web of Conferences 139, 01079 (2019), doi.org/10.1051/e3sconf/201913901079;
8. Karimov R.Ch., Shamsiyeva N. and others. IOP Conf. Series: Materials Science and Engineering, 883(1), 012120, (2020). doi:10.1088/1757-899X/883/1/012120
9. G.R.Rafikova, M.R.Ruzinazarov, S.K.Makhmutkhonov. E3S Web of Conferences, 139, 01075, (2019), https://doi.org/10.1051/e3sconf/201913901075
10. Khakimov,H.T., Shayumova,Z.M., Kurbanbaeva,Z.K ., Khusanov,B.M. Development of optimal modes and mathematical models of energy performance of electric steelmaking production//E3S Web of Conferences, 2019, 139, 01076
11. Khushnud Sapaev and Shukhrat Umarov. Two approaches for automating analysis of electromagnetic processes in non-linear circuits with valves. Web of Conferences https://doi.org/10.1051/e3sconf/20191390 E 3S 139 (2019) 10 1085 RSES 2019 8
12. Allayev, K.R., Fedorenko, G.M.,Postnikov, V.I.,Ostapchuk, L.B. Asynchronous generators as power system's natural dampers. 43rd International Conference on Large High Voltage Electric Systems 2010, CIGRE 20102010, 9p43rd International Conference on Large High Voltage Electric Systems 2010, CIGRE 2010; Paris; France; 22 August 2010.
13.Fazylov, Kh.F.,Allaev, K.R. Analysis of the operation of an electrical system during simultaneous operation of synchronous and asynchronous generators. Power engineering New York Volume 18, Issue 3, 1980, Pages 81-88.
14.Fazylov, Kh.F.,Allaev, K.R. Asynchronous turbogenerators with stator excitation and the prospects for their utilization. Power engineering New York Volume 23, Issue 2, 1985, Pages 7-13.
15. Fazylov, Kh.F.,Allaev, K.R. Calculation and experimental analysis of conditions of electrical power systems containing induction generators Power Engineering New York Volume 27, Issue 6, 1989, Pages 27-34.
16. R.Karimov, M.Bobojanov and others, IOP Conf. Series: Materials Science and Engineering, 2020, 883(1), 012120, doi:10.1088/1757-899X/883/1/012120
17. Usmanov, E.G., Rasulov, A.N., Bobojanov, M.K., Karimov, R.Ch. E3S Web of Conferences, 2019, 139, 01079, doi.org/10.1051/e3sconf/201913901079
18. R.Karimov, A.Rasulov, and others, IOP Conf. Series: Materials Science and Engineering, 2020, 883(1), 012142, doi:10.1088/1757-899X/883/1/012142
19. Karimov, R.Ch., Bobojanov, M.K., Rasulov, A.N., Usmanov, E.G. E3S Web of Conferences, 2019, 139, 01039, doi.org/10.1051/e3sconf/201913901039