

Selection Methodology of Energy Consumption Model Based on Data Envelopment Analysis

Vladimir Nakhodov (*Associate Professor, National Technical University of Ukraine “Kyiv Polytechnic Institute”*), Algirdas Baskys (*Professor, Vilnius Gediminas Technical University, Center for Physical Science and Technology*), Nils-Olav Skeie (*Associate Professor, Telemark University College, Norway*), Carlos F. Pfeiffer (*Professor, Telemark University College, Norway*), Ivanko Dmytro (*Ph.D. student, Vilnius Gediminas Technical University, National Technical University of Ukraine “Kyiv Polytechnic Institute”*)

Abstract – The energy efficiency monitoring methods in industry are based on statistical modeling of energy consumption. In the present paper, the widely used method of energy efficiency monitoring “Monitoring and Targeting systems” has been considered, highlighting one of the most important issues – selection of the proper mathematical model of energy consumption. The paper gives a list of different models that can be applied in the corresponding systems. The numbers of criteria that estimate certain characteristics of the mathematical model are represented. The traditional criteria of model adequacy and the “additional” criteria, which allow estimating the model characteristics more precisely, are proposed for choosing the mathematical model of energy consumption in “Monitoring and Targeting systems”. In order to provide the comparison of different models by several criteria simultaneously, an approach based on Data Envelopment Analysis is proposed. Such approach allows providing a more accurate and reliable energy efficiency monitoring.

Keywords – Energy consumption; Energy efficiency; Data models; Production management.

I. INTRODUCTION

Nowadays, the energy consumption in European countries is growing rapidly. This leads to a number of problems: increase in energy prices; increase in the negative environmental impact; loss of generating capacities; increase in the political influence of energy-exporting countries, etc.

It is known that the main energy consumer in the developed countries is industry [1]. A promising way to overcome the crisis phenomena in the energy sector is a widespread implementation of energy saving measures and technologies in industry [2]. Therefore, over the last few decades, various international and regional organizations have started investing additional money in the projects and funds intended for the promotion of energy efficiency in the industrial sector [3].

However, in order to obtain significant results in energy saving, it is not enough just implementing the energy saving measures [4]. To achieve this goal requires systematical monitoring of energy efficiency in enterprises [5].

One of the most reliable and accurate methods of energy efficiency monitoring is the use of the Monitoring and Targeting system (M&T system) [6]. This method allows providing the monitoring of energy efficiency at technological objects: work units, machines, technological processes, enterprise divisions, and other industrial objects.

M&T systems estimate the level of energy efficiency by comparing the actual energy consumption and the consumption obtained by a mathematical model. Such mathematical model reflects the influence of the main factors on energy consumption [7]. Appropriate factors characterize the conditions of the production process: output quantity, operation time of machines, pressure, temperature, etc.

Traditionally, in M&T systems, simplified mathematical models of energy consumption, such as simple linear regression or constant, are used [8]. In some cases it is proposed to use more complex, multivariable mathematical models with three or less independent variables.

Certain factors have nonlinear impact on energy consumption, which is necessary to take into account in mathematical models.

Inappropriate and simplified models can lead to a significant inaccuracy of energy efficiency monitoring. Therefore, the correct selection of a mathematical model is an important task in order to provide successful monitoring of energy efficiency.

The different methods of mathematical modeling allow calculating multivariable linear and nonlinear models. The software makes it possible to easily obtain different models and use them in energy efficiency monitoring. However, the selection methodology of energy consumption model in M&T systems, taking into account the different features of modeling, does not exist. This paper focuses on the development of the methodology of selecting an energy consumption model in M&T systems.

II. METHODS

Many publications are devoted to the application of various mathematical models for the monitoring of energy efficiency [6]–[9].

The most common in the use for calculating the linear and nonlinear mathematical models of energy consumption in M&T systems are the Method of Least Squares and Maximum Likelihood Estimation [10]–[12]. The Method of Least Squares is based on the minimization of the sum of the squares of the modeling errors [10]. Maximum Likelihood Estimation evaluates the unknown parameters of the model by maximizing the likelihood function [11].

A popular method of modeling in industrial processes is the Partial Least Squares (PLS) method, which is based on the

Principal Component Analysis. This method allows calculating the linear mathematical models in conditions when a small number of statistical data is available and the numbers of independent variables in the model are large [13]. An additional benefit of the PLS method is that it takes into account multicollinearity between the independent variables.

The paper [14] proposes to use time series in order to control energy consumption. Time series can be calculated by the following methods: Moving Average, Autoregressive Integrated Moving Average (ARIMA), Singular Spectrum Analysis (SSA), etc. The authors of this paper consider the application of these methods in M&T systems less effective than the application of regression models.

Some methods allow obtaining the mathematical model in the form of interval values. One of these methods is Simple Interval Estimation, which is presented in [15].

The method of calculating complex multivariable models is the Group Method of Data Handling (GMDH) [16]. It is based on the self-organization theory of mathematical models. GMDH allows obtaining the models with complex basic functions: polynomials, exponential, Fourier series, etc.

During the last few decades, the methods of mathematical modeling based on neural networks have become increasingly popular. Neural networks are the methods of simulation processes, which allow reflecting the complex relationships between the variables. The Generalized Regression Neural Network (GRNN) calculates the mathematical models of high complexity, which demonstrates good results in the conditions of the impact of nonlinear factors on energy consumption [17].

It is known that every industrial enterprise is unique, with its specific features of energy consumption, production conditions, different number of factors and their impact on energy consumption. Therefore, it is impossible find the concrete type of energy consumption model that would be suitable for any technological object.

The selection of an energy consumption model should be done individually for each technological object, considering its unique characteristics and conditions. A suitable model can be found by comparing different mathematical models obtained on the same set of data.

Such comparison should be implemented by the criteria which allow drawing conclusions about the adequacy of certain features of mathematical models (accuracy, complexity, stability, etc.) [11], [12].

Various publications describe a wide range of the "traditional" criteria of the adequacy of mathematical models (Table I). Numerical calculations show that using different criteria can lead to obtaining controversial results of selecting the most acceptable model. Therefore, the selection of the energy consumption model in M&T systems should be done by several different criteria simultaneously.

Table I shows that many criteria duplicate each other and evaluate the same features of a mathematical model. The use of "similar" criteria can lead to the selection of an inappropriate model of energy consumption. Also, it does not allow fully evaluating the characteristics of mathematical models.

Thus, the authors of the present paper considered the question of determining the set of the criteria of mathematical models adequacy which allow taking into account the main characteristics of models and do not duplicate each other.

Conventionally, the criteria shown in Table I can be divided into two groups by their contents.

The application of the criteria which belong to the first group is based on comparing the values obtained by calculation with some ideal, tabulated values. Such comparison allows making a conclusion about the suitability of the model by a certain feature. Appropriate criteria are Choy test, F -criteria, t -criteria, and DW -criteria. These criteria can be used for preselecting of mathematical models.

The second group includes all the remaining criteria. Different models can be compared by the values of these criteria in order to select the most suitable model. As noted, the use of criteria which duplicate each other is not recommended. Therefore, the authors proposed the minimum required set of criteria which allow avoiding controversial results and unnecessary calculations.

The criteria of the second group can be divided into three subgroups by the features of the models they consider:

1) the criteria which estimate the accuracy of mathematical models: ESS , D , SEE , R^2 , R^2_{Adj} , \bar{A} , MAE , $RSME$, $\Delta^2(B)$;

2) the criteria which estimate the accuracy and complexity of mathematical models: AIC , BIC ;

3) the criteria which estimate the stability of mathematical models, namely, feasibility of their application on new data: n .

The criteria within one subgroup should be checked on duplicating. For this purpose, the authors carried out numerical simulations where the closeness of the relationship between different criteria was evaluated by correlation analysis. The obtained results: ESS , D , SEE , \bar{A} , MAE , $RSME$, $\Delta^2(B)$ are similar criteria; R^2 , R^2_{Adj} is the modification of one criterion which also makes them similar; AIC , BIC are similar criteria. Thus, the similar criteria are duplicating each other and their simultaneous application is not advisable.

It is necessary to pay attention that the AIC and BIC criteria estimate the complexity and the accuracy of the model simultaneously. Therefore, in order to eliminate the duplication of the criteria from the subgroups 1 and 2, it is advisable to calculate one of them on the training sample, and the other one – on the testing sample. It will help selecting the most balanced model.

The criteria presented in Table I evaluate only the basic features of the mathematical models. In order to consider the models more precisely, it is appropriate to use some "additional" criteria which are taking into account other important features of the modeling. For example, the "additional" criteria can include the reasonable time of mathematical modeling, the financial expenses of modeling, etc.

Which "additional" criteria is necessary to use should be decided individually taking into account the characteristics of the technological objects where M&T systems have to be created.

To determine the numerical values of “additional” criteria, the authors propose to use the methods of expert survey and fuzzy logic.

The procedure of determining the “additional” criterion – the reasonable time of mathematical modeling – is shown in Table I.

TABLE I
WELL-KNOWN CRITERIA OF MATHEMATICAL MODEL ADEQUACY

No.	Title	Formula	Symbols
1	ESS (Explained Sum of Square)	$ESS = \sum_{k=1}^n (y_k - y_k^*)^2$	y_k, y_k^* – the actual and estimated values of the dependent variable
2	The residual variance	$D = \frac{1}{n-1} \sum_{k=1}^n (y_k - y_k^*)^2$	y_k, y_k^* – the actual and estimated values of the variable; n – the number of observations
3	Standard Error of the Estimate (SEE)	$SEE = \sqrt{\frac{ESS}{n-k-1}}$	ESS – explained Sum of Square; n – the number of observations; k – the number of independent variables
4	Coefficient of determination (R-square)	$R^2 = 1 - \frac{ESS}{TSS}$	ESS – explained Sum of Square; $TSS = \sum_{k=1}^n (y_k - \bar{y})^2$ – the sum of the squared deviations of the actual values of the dependent variable from its mean value
5	Adjusted coefficient of determination (Adjusted R-square)	$R_{Adj}^2 = 1 - \frac{ESS(n-1)}{TSS(m-1)}$	n – the number of observations; m – the number of independent variables
6	The average approximation error	$\bar{A} = \frac{1}{n} \sum_{k=1}^n \left \frac{y_k - y_k^*}{y_k} \right 100\%$	y_k, y_k^* – the actual and estimated values of the dependent variable; n – the number of observations
7	Mean Absolute Error (MAE)	$MAE = \frac{1}{n} \sum_{k=1}^n y_k - y_k^* $	y_k, y_k^* – the actual and estimated values of the dependent variable; n – the number of observations
8	Akaike information criterion	$AIC = \ln(\sigma^2) + \frac{2k}{n}$	σ^2 – residual variance modeling; n – the number of observations; k – the number of the parameters or independent variables of the model
9	Schwartz Criteria (BIC)	$BIC = \ln(\sigma^2) + \frac{k \ln n}{n}$	σ^2 – residual variance modeling; n – the number of observations; k – the number of the parameters or independent variables of the model
10	F-criteria Fisher	$F = \frac{R^2 / k}{R^2 / (n - k - 1)}$	R^2 – coefficient of determination; n – the number of observations; k – the number of the parameters or independent variables of the model
11	Root-mean-square deviation (RSME)	$RSME = \sqrt{\frac{ESS}{n-m}}$	ESS – explained Sum of Square; n – the number of observations; m – the number of the parameters or independent variables of the model
12	Student test (t-criteria)	$t_{ai} = \frac{a_i \sqrt{n-k-1}}{\sigma_{ai}}$	n – the number of observations; k – the number of the parameters or independent variables of the model; σ_{ai} – the standard deviation value of i -th parameter of the model; a_i – i -th parameter of the model
13	Durbin-Watson test (DW-criteria)	$DW = 2(1 - p)$	p – the correlation coefficient between squares residues modeling defined for training and checking samples.
14	Choy test	$F(T, m) = \frac{(RSS_{T+m} + RSS_T) / k}{RSS_T / (n - k - 1)}$	RSS_T and RSS_{T+m} – the residual sum of squares modeling, determined according to the sample, including the data for the period T and period $T + m$
15	Regularity criteria	$\Delta^2(B) = \frac{\sum_{k=1}^{N_B} (y_k - y_k^*)^2}{\sum_{k=1}^{N_B} (y_k)^2}$	y_k – the experimental values of the dependent variable belonging to checking sequence (B); y_k^* – the dependent value variable calculated on the relevant model; N_B – the number of experimental points in the checking sequence
16	Minimum bias criteria	$n = \frac{1}{R_1 + R_2} \sum_{k=1}^{R_1+R_2} (Z_k^* - Z_k^{**})^2$	R_1, R_2 – the size of the first and second sub-samples of data, respectively; Z_k^*, Z_k^{**} – the prediction value of the first and second models on all points samples

The time required for calculating the mathematical model of energy consumption is different for each specific technological object. It depends on the period of energy efficiency monitoring, the number of equipment covered by M&T systems, the qualification of the staff, etc.

The various methods of energy consumption modeling have different actual time expenses (T_a), required to determine the factors that have a significant impact on energy consumption, to handle statistical data, to calculate the model and use the procedure of energy efficiency monitoring. The values of T_a for each model can be obtained from the experiments or the previous experience of solving similar problems at other objects. In order to estimate reasonable time (T_r) for calculating the mathematical models for each specific technological object, it is advisable to use the knowledge and experience of the experts.

Usually, the expert assessments have a subjective character; therefore, the processing of expert survey can be done by the methods of fuzzy logic.

For this purpose, each expert should indicate by a vertical line on a special scale the approximate value of T_r , which according to his point of view is acceptable for this technological object. In Fig. 1, $T_{a.min}$ is the least amount of the time required for obtaining the model of energy consumption by using the considered modeling techniques. T_m is the period of energy efficiency monitoring at the appropriate technological object.

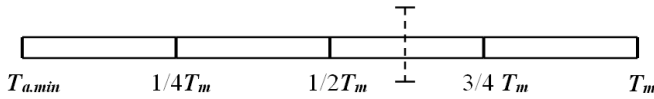


Fig. 1. The expert survey about the reasonable time for calculating the mathematical models of the energy consumption of appropriate technological equipment.

Fuzzy logic mathematical tools allow determining the level of the reliability of the subjective assessments of each expert (Fig. 2).

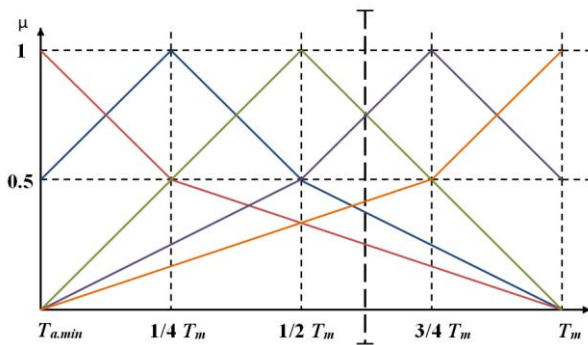


Fig. 2. Fuzzy parameter functions of the reasonable time for calculating the mathematical models of the energy consumption of appropriate technological equipment.

Different numbers of the levels of specified fuzzy parameter can be used in order to conduct an expert survey. The high numbers of appropriate levels increase the accuracy of the obtained results as well as make the expert survey more complex. Fig. 2 shows the five levels of the fuzzy parameter – the reasonable time of mathematical models calculation.

The number of the piecewise linear membership functions of fuzzy parameters, which is used for processing the results of expert survey, is equal to the number of the levels of fuzzy parameter.

The five membership function of the reasonable time of mathematical models calculation is represented by the following equations:

$$\mu_0(T) = \begin{cases} 1 - \frac{T}{2}, & T \in [0, 1], \\ \frac{2}{3} - \frac{T}{6}, & T \in (1, 4]; \end{cases} \quad (1)$$

$$\mu_{1/4 T_m}(T) = \begin{cases} \frac{1}{2} + \frac{T}{2}, & T \in [0, 1], \\ \frac{3}{2} - \frac{T}{2}, & T \in (1, 2], \\ 1 - \frac{T}{4}, & T \in (2, 4]; \end{cases} \quad (2)$$

$$\mu_{1/2 T_m}(T) = \begin{cases} \frac{T}{2}, & T \in [0, 2], \\ 2 - \frac{T}{2}, & T \in (2, 4]; \end{cases} \quad (3)$$

$$\mu_{3/4 T_m}(T) = \begin{cases} \frac{T}{4}, & T \in [0, 2], \\ \frac{T}{2} - \frac{1}{2}, & T \in (2, 3], \\ \frac{5}{2} - \frac{T}{2}, & T \in (3, 4]; \end{cases} \quad (4)$$

$$\mu_{T_m}(T) = \begin{cases} \frac{T}{6}, & T \in [0, 3], \\ \frac{T}{2} - 1, & T \in (3, 4]. \end{cases} \quad (5)$$

In order to estimate the opinion of each expert about the reasonable time of mathematical models calculation, from Fig. 2 or eq. (1)–(5), the levels of the conformity (μ_j^k) of expert assessment with one of the specified levels of fuzzy parameter can be found.

The average values of μ_j^k by the levels of fuzzy parameter can be found from the equation:

$$\mu_j = \frac{1}{N} \sum_{k=1}^N \mu_j^k,$$

where N – the total number of experts; j – the number of specified level values; k – the number of expert.

The values of μ_j allow finding the most probable value of the reasonable time of mathematical models calculation for energy efficiency monitoring. Table II shows that the most probable value of the reasonable time of mathematical models calculation will be under $3/4 T_m$; therefore, T_r is taken $3/4 T_m$.

The numerical value of the “additional” criterion – time required for calculating the mathematical model – can be obtained by the equation:

$$T_i = \frac{T_{a,i}}{T_r},$$

where i — the number of mathematical model.

For example, the period of monitoring is one hour. The time expenses for obtaining the mathematical model are 22 minutes and 30 seconds. Then

$$T = \frac{22.5}{3/4 \cdot 60} = 0.5.$$

TABLE II
CALCULATION OF THE MOST PROBABLE REASONABLE TIME
OF MATHEMATICAL MODEL CALCULATION

The level of subjective probabilities	Number of expert				μ_j
	No. 1	No. 2	No. 3	No. 4	
$T_{a.min}$	0.28	0.22	0.1	0.14	0.18
$1/4T_m$	0.38	0.34	0.13	0.23	0.27
$1/2T_m$	0.8	0.63	0.3	0.42	0.53
$3/4T_m$	0.8	0.9	0.8	0.92	0.85
T_m	0.42	0.45	0.8	0.6	0.56

The selection of the suitable model of energy consumption in M&T systems, based on several different criteria, is a multi-optimization problem. The main approach of solving this issue is comparing the weighted sum of the criteria of different models.

Usually, the weighted sum of all criteria is determined by the methods within Multi-Attribute Utility Theory. However, one significant drawback of the appropriate methods is the lack of clear recommendations about obtaining the weights for each of the criteria. These weights have a significant impact on the selection of the model.

The approach for selecting the suitable model of energy consumption by several criteria simultaneously proposed by the authors of this paper is based on the application of the Data Envelopment Analysis (DEA).

Generalized criteria of models suitability for each mathematical model DEA formed as maximizing functional with restrictions [18]:

$$f_n = \max_{u_i, v_j} \frac{u_1 Y_1 + u_2 Y_2 + \dots + u_k Y_k}{v_1 X_1 + v_2 X_2 + \dots + v_m Y_m} \quad (6)$$

$$\frac{u_1 Y_1 + u_2 Y_2 + \dots + u_n Y_n}{v_1 X_1 + v_2 X_2 + \dots + v_m Y_m} \leq 1, \quad (7)$$

$$n = 1, 2, \dots, N,$$

where Y_i — i -th criterion of the models, which characterizes the quality characteristics of the model; X_j — j -th criterion which characterizes the expenses for model calculation; n — the number of the model.

The system of the equations (6) and (7) is solved for each of mathematical models. The obtained solution of each of the n -models gives the generalized criteria of the models' suitability f_n , ranking on interval $[0, 1]$, and the corresponding weights $u_i = \{u_1, u_2, \dots, u_k\}$ and $v_j = \{v_1, v_2, \dots, v_m\}$.

It should be noted that some criteria for proper use of DEA require transformations because the lowest values of certain

criteria indicate a better model, and other criteria have opposite properties.

III. EXPERIMENTS

The selection of energy consumption model in M&T systems has been conducted on a bottling line where beverage is filled into bottles. The bottling line has the automated accounting system that allows obtaining information about electrical energy consumption and the amount of manufactured products. The main goods which are produced on the corresponding technological line are 5-beer, 8-cider and 11-kvass, where 5, 8, 11 are conditional numbers of the products in assortment of the enterprise. Statistical information about the amount of produced goods and the total electrical energy consumption are shown in Fig. 3.

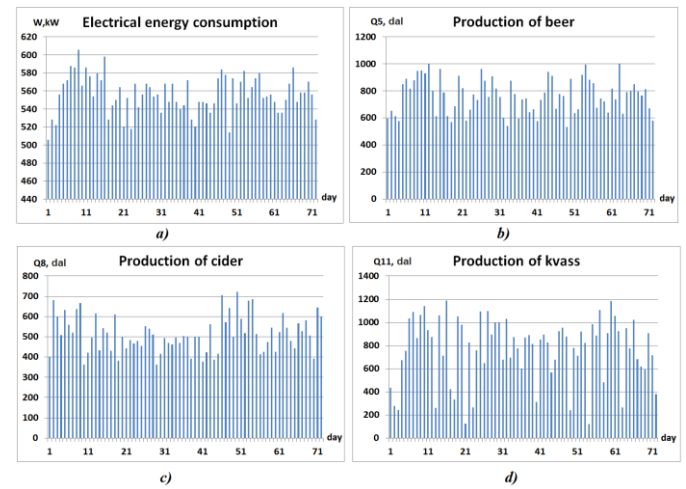


Fig. 3. Statistical data about the produced products and electrical energy consumption of a bottling line, where: a) electrical energy consumption; b) production of beer; c) production of cider; d) production of kvass.

Based on the statistical data presented in Fig. 3, different multivariable mathematical models of the energy consumption at the bottling line were calculated:

- the complex nonlinear model obtained by combining the “individual” models for different variants of the technological process:

$$P = \sum_{i=1}^N f(Q_i) = f(Q_5) + f(Q_8) + f(Q_{11}),$$

where Q_5, Q_8, Q_{11} — the amount of manufactured products;

- mathematical models of the energy consumption of the product number 5:

$$P_5 = 191.9 - 1.2 \cdot \cos(Q_5 \cdot 0.01) - 10.9 \cdot \sin(Q_5 \cdot 0.01) + 0.06 \cdot \cos(Q_5 \cdot 0.02) - 2.7 \cdot \sin(Q_5 \cdot 0.02) - 1.7 \cdot \cos(Q_5 \cdot 0.03) - 3.3 \cdot \sin(Q_5 \cdot 0.03)$$

(Mathematical models of energy consumption of the manufacturing of the products 8 and 11 have the similar shape to the model of product number 5);

- multivariable linear regression obtained by the Method of Least Squares:

$$P = 432.7 + 0.04 \cdot Q_5 + 0.1 \cdot Q_8 + 0.05 \cdot Q_{11};$$

- multivariable nonlinear polynomial regression:

$$P = 490.5 + 8.5 \cdot 10^5 \cdot Q_{11}^2 + 9.2 \cdot 10^5 \cdot Q_8 + 2.1 \cdot 10^{-14} \cdot Q_5^5 - 3.7 \cdot 10^{-8} \cdot Q_{11}^3;$$

- multivariable nonlinear regression obtained by the Group Method of Data Handling:

$$P = 415.8 + 8 \cdot 10^{-6} \cdot Q_{11}^2 + 4.42 \cdot \sqrt{Q_8} + 2.1 \cdot 10^{-14} \cdot Q_5^5 - 3 \cdot 10^8 \cdot Q_{11}^3;$$

- the model obtained by the Generalized Regression Neural Network.

The numerical values of the “traditional” criteria which estimate the adequacy of mathematical models is shown in Table III.

TABLE III
CALCULATION OF THE MOST PROBABLE REASONABLE TIME
OF MATHEMATICAL MODEL CALCULATION

Criterion	Model $P = P_5 + P_8 + P_{11}$	Multivariable linear regression	Multivariable nonlinear regression		GRNN
			Polynomial	GMDH	
D	362.66	203.914	217.51	215.284	201.02
R^2	0.842	0.731	0.73	0.741	0.855
BIC	6.323	5.985	5.899	5.959	5.75
n	146.821	134.179	134.97	134.926	100.6

The obtained-by-expert-survey “additional” criteria which characterize the time expenses of model calculation, the financial expenses, and the expenses of production resources are shown in Table IV.

TABLE IV
“ADDITIONAL” CRITERIA THAT ARE CONSIDERED
THE FEATURES OF THE MODELING

Criterion	Model $P = f(Q_5) + f(Q_8) + f(Q_{11})$	Multivariable linear regression	Multivariable nonlinear regression		GRNN
			Polynomial	GMDH	
Time expenses, T	0.709	0.5	0.59	0.62	0.34
Financial expenses, C	0.67	0.4	0.45	0.5	0.4
Expenses of production resources, K	0.56	0.4	0.34	0.32	0.45

According to DEA, the generalized criteria of models suitability can be considered as a set of N conditions of mathematical models with 5 inputs and 3 outputs, which reflects the “traditional” and “additional” criteria of model adequacy.

The input parameters in DEA will be taken as the inverse value of the “traditional” criteria, the outputs – as the inverse value of the “additional” criteria.

The inverse value of the criteria is taken in order the bigger value of the generalized criteria of models suitability indicates the better model.

Some criteria of model adequacy have been taken as reverse values in the equations (6) and (7). Because of this, the greater value of the generalized criterion obtained by DEA indicates the better model.

Thus, the optimization problem for calculating the generalized criteria of models suitability is written as:

$$f_n = \max_{u_i, v_j} \frac{u_1 \cdot 1/T + u_2 \cdot 1/C + u_3 \cdot 1/K}{v_1 \cdot D + v_2 \cdot 1/R^2 + v_3 \cdot AIC + v_4 \cdot n}, \quad (8)$$

$$\frac{u_1 \cdot 1/T + u_2 \cdot 1/C + u_3 \cdot 1/K}{v_1 \cdot D + v_2 \cdot 1/R^2 + v_3 \cdot AIC + v_4 \cdot n} \leq 1. \quad (9)$$

Solving the equations (8) and (9) for each mathematical model shows that the generalized criteria of models suitability is optimal in the case of applying the Generalized Regression Neural Network (Table V). The weights of the criteria that have been found from eq. (8) and (9) are shown in Table VI.

TABLE V
THE VALUES OF WEIGHT CALCULATED BY DEA

Weight						
$U1$	$U2$	$U3$	$V1$	$V2$	$V3$	$V4$
0.15	0.13	0.11	0.004	0.00002	0.0001	0.002

TABLE VI
GENERALIZED CRITERIA OF MODELS SUITABILITY OF EACH MATHEMATICAL
MODEL CALCULATED BY DEA

f_n	Model $P = f(Q_5) + f(Q_8) + f(Q_{11})$	Multivariable linear regression	Multivariable nonlinear regression		GRNN
			Polynomial	GMDH	
	0.34	0.83	0.76	0.75	1

Table VII shows that the most suitable mathematical model of energy consumption is the Generalized Regression Neural Network.

The application of Data Envelopment Analysis in order to select the mathematical model of energy consumption has a number of significant advantages:

- allows calculating the generalized criteria of models suitability which estimate the complex characteristics of the mathematical model;
- can simultaneously handle many inputs and outputs, and, besides, can be measured in different units;
- does not require preliminary determination of the weighting coefficients for solving the optimization problem;
- if necessary, allows taking into account the authority of experts concerning the importance of a certain criterion.

IV. CONCLUSION

The development of accurate energy efficiency monitoring methods is an important task for widespread implementation of energy saving projects in industry. The energy efficiency monitoring methods in industry are based on the statistical modeling of energy consumption. Inappropriate and simplified models can lead to a significant inaccuracy of energy efficiency monitoring. Therefore, the correct selection of mathematical model is an important task in order to provide successful monitoring of energy efficiency.

The mathematical models of energy consumption for the technological objects can be obtained by different methods: the Method of Least Squares, Maximum Likelihood Estimation, Group Method of Data Handling, Partial Least

Squares, Generalized Regression Neural Network, etc. The selection of an appropriate energy consumption model should be done individually for each technological object, considering its unique characteristics and conditions. Traditionally, for the estimation of the characteristics of a mathematical model, special criteria are used. These criteria allow estimating the accuracy, complexity stability of mathematical models.

For a more complete assessment of mathematical models besides the traditional criteria of model adequacy, it is proposed to apply the “additional” criteria that are linked to other features of the modeling (the time required for obtaining a mathematical model, the financial costs of modeling, etc.). To determine the numerical values of “additional” criteria, a procedure based on expert survey and fuzzy logic is proposed.

In the present paper, the selection of the mathematical model of energy consumption for energy efficiency monitoring is considered as a multi-optimization problem. In order to solve the problem of selecting the mathematical model, the Data Envelopment Analysis, which has a number of significant advantages over the method of the Multi-Attribute Utility Theory has been used. The proposed method allows improving the methods of energy efficiency monitoring.

REFERENCES

- [1] F. Kreith and D. Y. Goswami, Eds., *Energy management and conservation handbook* (Mechanical and Aerospace Engineering Series), CRC Press, Taylor & Francis Group, 2007.
- [2] A. Baskys, V. Nakhodov, D. Ivanko and C. Pfeiffer, “Calculation of Electrical Energy Balances of Production Systems Based on Probabilistic-Statistical Approach,” in *2015 IEEE 3rd Workshop on Advances in Inform., Electron. and Elect. Eng. (AIEEE)*, Riga, 2015, pp. 1–6. <https://doi.org/10.1109/aieee.2015.7367293>
- [3] European Commission – Fact Sheet. (2016, June 14). *The Investment Plan for Europe and Energy: making the Energy Union a reality*. [Online]. Available: http://europa.eu/rapid/press-release_MEMO-16-2195_en.htm
- [4] I. Beinarts, U. Grunde and A. Jakovics, “Distributed Multi-Sensor Real-Time Building Environmental Parameters Monitoring System with Remote Data Access,” *Elect., Control and Communication Eng.*, vol. 7, pp. 41–46, March 2014. Available: <https://doi.org/10.1515/ecce-2014-0022>
- [5] S. Hu, F. Liu, Y. He and T. Hu, “An on-line approach for energy efficiency monitoring of machine tools,” *J. of Cleaner Production*, vol. 27, pp. 133–140, May 2012. <https://doi.org/10.1016/j.jclepro.2012.01.013>
- [6] V. Nahodov, O. Borychenko and D. Ivanko, “Monitoring of energy efficiency in energy management systems,” *J. of KNUITD*, vol. 6, pp. 67–77, 2013. (in Ukrainian).
- [7] ETSU and Cheriton Technology Management Ltd. (1998). *Monitoring and Targeting in large companies. Good Practice Guide 112*. [Online]. Available: <http://www.cibse.org/getmedia/f0fa2ac9-c4bb-4aa1-9dfd-8e1858fbbd4c/GPG112-Monitoring-and-Targeting-in-Large-Companies.pdf.aspx>
- [8] P. Jones, Getting started with Monitoring&Targeting (M&T). *Fundamental Series*, no. 7, pp. 29–32, 2004.
- [9] Computer Based Monitoring And Targeting On A Hot Rolling Mill. Energy Efficiency Enquiries Bureau, ETSU, Harwell, Oxford shire, OX11. Best Practice Program. 1992. 26 p.
- [10] C. Dougherty, *Introduction to econometrics*, 2nd ed. Oxford University Press, 2002, Paperback, 424 p.
- [11] A. A. Afifi and V. Clark, *Computer-Aided Multivariate Analysis*, 3rd ed. Springer Science Business Media, 1996, 456 p.
- [12] G. A. F. Seber and C. J. Wild, *Nonlinear Regression* (Wiley Series in Probability and Statistics), New York: John Wiley & Sons, Inc., 1989. <https://doi.org/10.1002/0471725315>
- [13] K. H. Esbensen, *Multivariate Data Analysis – In Practice*, 5th ed. Oslo, Norway: CAMO Software, 2002.
- [14] M. A. Piette, S. Kinney and H. Friedman, “EMCS Time-Series Energy Data Analysis in a Large Government Office Building,” *Proc. of 9th National Conf. on Building Commissioning*, Cherry Hill, New Jersey, May 9–11, 2001, pp. 1–10.
- [15] O. Y. Rodionova, K. H. Esbensen and A. L. Pomerantsev, “Application of SIC (simple interval calculation) for object status classification and outlier detection–comparison with regression approach,” *J. of Chemometrics*, vol. 18, issue 9, pp. 402–413, 2004. <https://doi.org/10.1002/cem.885>
- [16] A. G. Ivakhnenko, G. A. Ivakhnenko, “The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH),” *Pattern Recognition and Image Analysis*, vol. 5, no. 4, pp. 527–535, 1995.
- [17] M. H. Beale, M. T. Hagan and H. B. Demuth. (2010). *Neural Network Toolbox™ 7 User’s Guide*, 1992–2011, The MathWorks, Inc.,
- [18] J. Zhu, Ed., *Data Envelopment Analysis: A Handbook of Models and Methods*, Yew York: Springer, 2015, 472 p.



Vladimir Nakhodov received the Ph.D. degree in energy systems and complex from the National Technical University of Ukraine “Kyiv Polytechnic Institute”. He is currently an Associate Professor at the Institute of Energy Saving and Energy Management of the National Technical University of Ukraine “Kyiv Polytechnic Institute”. His research interests include electrical engineering, the methods of energy efficiency monitoring, electricity tariffs, and energy management.
Address: Borschagivska 103, Kiev, Ukraine.
E-mail: nakhodov@ukr.net



Algirdas Baskys, Prof., Ph.D. is the Head of the Computer Engineering Department of Vilnius Gediminas Technical University, and the Head of Electronic Systems Laboratory of the Center for Physical Sciences and Technology, Vilnius, Lithuania.
His research interests include power electronics, control methods and controllers, and energy efficiency technologies.
Address: Naugarduko g. 41, Vilnius, Lithuania.
E-mail: algirdas.baskys@vgtu.lt



Nils-Olav Skeie received the Ph.D. degree in cybernetics from Telemark University College (TUC). Since 2008, he is an Associate Professor of Measurement Techniques, Software Design and System Design at the B.Sc. and M.Sc. levels. From 1985 to 2003, he gained his industrial experience in system design, both software and hardware development, mainly with monitoring systems within the aviation and maritime sectors.
His research interests include smart buildings, sensor networks, level measurements, and soft sensors.
Address: Gullbringvegen 36, Porsgrunn, Norway.
E-mail: Nils-Olav.Skeie@hit.no



Carlos F. Pfeiffer received the Ph.D. degree in process control from the University of Texas at Austin in 1999. Since 2011, he is a Professor of Systems and Control Engineering at Telemark University College, Porsgrunn, Norway. Up to 2011, he has been a Professor at ITESM and a Process Engineer in Motorola, Austin Tx.
His research interests include process modeling, optimization and control.
Address: Gullbringvegen 36, Porsgrunn, Norway.
E-mail: carlos.pfeiffer@hit.no



Ivanko Dmytro received the M.Sc. degree in electrical engineering from National Technical University of Ukraine “Kyiv Polytechnic Institute” in 2012. He is currently pursuing the Ph.D. degree with “Monitoring of energy efficiency in production systems based on probabilistic-statistical approach”. In 2014, he participated in the exchange postgraduate program “Eurasia” at Telemark University College, Porsgrunn, Norway. During 2015–2016, he was taking part in the exchange postgraduate program “Erasmus” at Vilnius Gediminas Technical University.

His research interests include monitoring of energy efficiency in production, calculation of electrical energy balances in production units, and energy management.

Address: Borschagivska 103, Kiev, Ukraine.

E-mail: dimo-vin@yandex.ua