Agronomy Research **12**(2), 633–644, 2014

Cleaner pellet production – an energy consumption study using statistical analysis

H. Vigants[1,*], P. Uuemaa[2], I. Veidenbergs[1] and D. Blumberga[1]

[1]Department of Electrical Power Engineering, Riga Technical University, Kronvalda bulv.1, LV-1010 Riga, Latvia; [*]Correspondence: haralds.vigants@rtu.lv
[2]Department of Electrical Power Engineering, Tallinn University of Technology, Ehitajate tee 5, EE19086 Tallinn, Estonia

Abstract. This study investigates and analyses the methodology for introducing cleaner wood pellet production. A statistical model is developed for the energy consumption analysis. Efficiency indicators have been chosen which allow determining the impact on the production process. The developed model can be used in other similar type of industries. This study has processed large empirical data with statistical methods in order to establish the efficiency indicators. The modelled results enable to define the indicators which lead to higher efficiency and hence to the cleaner production.

Key words: Energy efficiency, energy optimization, improvement of the production process, industrial process.

INTRODUCTION

Production of wood pellets in the Baltic countries has been growing (Mola-Yudego et al., 2013) as new large and small scale facilities emerge. Cogeneration plants are added next to the most advanced production sites of wood pellets in order to improve the production efficiency and secure cleaner production (Anderson & Toffolo, 2013; Kohl et al., 2013). The concept of cleaner production comprises the aspects of resources consumption. Reducing the resources consumption is an essential tool for industrial production. Efficiency is the basis for analysing and developing guidelines for continuous improvement in the production process (Song et al., 2011). At pellet production facilities, these resources are biomass, which is the raw material for pellets, and fuel for energy production, electricity, heat, and water. The contribution of this study is to develop a model for efficient use of these resources.

The second section introduces the materials and methods; the 3rd section outlines the modelling study results and includes the discussion. The 4th section concludes the study and makes recommendations for future work.

MATERIALS AND METHODS

Production description

When wood pellets are produced, they have to comply with generally accepted quality standards. Only clean sawdust can be used for production. The sawdust may not contain any impurities, it has to be dry, free of any sand, abrasive particles and chemistry. The produced pellets must be mechanically robust, should not contain any small sawdust particles, and have to be free of any foreign objects.

Damp sawdust is the basic material for the production of wood pellets. After additional processing, cellulose fibre and technological wood chips can also be used for the production. Processing of wood chips takes place at a special cutting device. The system of the cutter consists of a range of electrical machines like chippers, peelers, conveyors, ventilators, hydro devices, etc.

The wood chips that are made of branches and bark can be used as fuel and are combusted for production of flue gases and for energy production in the cogeneration plant. The furnace system in the flue gas production consists of several electrical motors that power the equipment for ensuring operation of the furnace, i.e. heating, cooling, feeding of fuel, ventilators, valves, and hydro machinery. The heat from the cogeneration plant on the other hand is transferred to belt-type dryers. The biggest electricity consumers of belt dryers are ventilators, which suck heated air through sawdust.

Sawdust needs to be dried to obtain the required humidity level for the production of wood pellets. Dried sawdust material is dosed to a hammer mill where the sawdust is crushed. During the milling process, dried sawdust is turned into small particles and dust, a uniform substance is obtained. The milling system consists of several electrical motors that power the equipment to ensure milling of sawdust, i.e. heating, cooling, a dosing device, a ventilator, a worm-type transporter, a mill, and hydro machinery.

In the granulating device, there is a process during which the mix of sawdust is delivered into two pressing rolls and a rotating matrix. The process results in production of hot pellets with a diameter of 8 mm. The following process is cooling and screening, where the pellets become hard and are cleaned of any dust. After this process, the pellets are ready for storage and transportation. The granulating system consists of several electrical motors that power the equipment, i.e. conveyors, suction devices, coolers, worm-type transporters, a vibrant-sieve, dosing devices, a mixer, a central lubricating system, and presses. On Fig. 1 the operational scheme of the pellet production facility with indications to analysed data is shown.

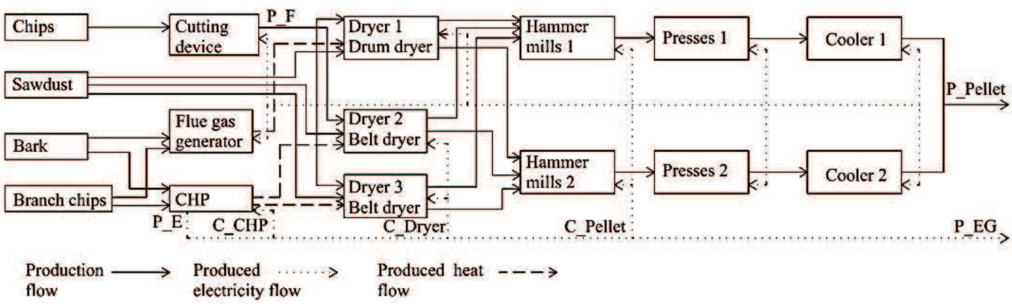

Figure 1. The operational scheme of the pellet production facility with the analysed indicators.

If power consumption at an industrial production facility is reduced, the share of generated electricity available for sale on the electricity market increases. Optimization of the production of the entire facility reduces the electricity bill of the pellet plant and increases the revenues from electricity sales.

This study will provide an analysis of power generation at a cogeneration plant, P_E. The target is to find the factors which have an impact on the power generation and therefore, once these factors have been defined, the next step is to optimize the production accordingly. Table 1 shows the list and explanations of the analysed data.

Table 1. Analysed data

Abbreviations	Explanation	Unit
C_CHP	Cogeneration plant power consumption	kWh
C_Dryer	Power consumption of the belt dryers	kWh
C_Pellet	Power consumption of the pellet production facility, excluding the belt-type dryers	kWh
E1 = C_DP	Power consumption of granulation	kWh
C_Total	Total power consumption	kWh
P_E	Produced electricity	kWh
P_EG	Produced electricity to grid	kWh
P_Pellet	Produced pellets	t
P_F	Sawdust crushed in the cutting device	m^3
C_DP/P_Pellet	Power consumption per produced ton of pellets	kWh t^{-1}
C_CHP/P_E	Power consumption per generated kWh	kWh kWh^{-1}

Analysis and data processing

Authors (Savola et al., 2007; Söderman & Ahtila, 2010) have used the modelling programme MINLP, MILP or simulation software (Mikita et al., 2012; Mobini et al., 2013). In this study, the empiric data was processed by applying statistical methods for data processing, correlation, and regression analysis. By means of a correlation analysis, the mutual link between two values and its strength are determined. The regression analysis is used for identifying the statistical importance of the multi-factor regression model and its coefficients (Blasnik, 1995).

The computer software STATGRAPHICSPlus was used for statistical processing of the data and development of the multi-factor empiric model. A similar model has been developed by other authors (Revina, 2002; Beloborodko et al., 2012).

In order to select the type of the regression equation, the linkage of the parameters by means of performing the correlation analysis is established by a single-factor linear model. The strength of the mutual link between independent and dependent random variables (correlation) can be assessed by means of a correlation coefficient. In case of a single-factor mathematic model, the Pearson's equation (1) is used for its estimation:

$$r = \frac{\sum_{i=1}^{m}(x_1 - x) \cdot (y_1 - y)}{(m-1) \cdot S_x \cdot S_y}, \qquad (1)$$

where: x_i and y_i are the independent variables and pairs of their corresponding dependent variables; x and $y$are the arithmetic values of independent and dependent variables; S_x and S_y are the variables of the selection dispersion.

Correlation coefficients were used for evaluating the accuracy of the mathematic models describing the strength of the correlation. It is assumed that a correlation is good if the correlation coefficient is above 0.8. It should be noted that in software for statistic processing of data, the squared correlation coefficient is usually calculated. When the value R^2 is multiplied by 100, the value that characterises the changes of dependent variables is described by the resulting empiric equation. For example, $R^2 = 0.9$ indicates that the relevant regression equation characterises 90% of the changes of the dependent random variables.

Correlation analysis of produced electricity

In this study, the purpose is to analyse the operation of the production facility and find the correlation between produced electricity *P*_E and the following parameters:

- Cogeneration plant power consumption *C*_CHP, kWh;
- Power consumption of the belt dryers *C*_Dryer, kWh;
- Power consumption of the pellet production facility *C*_Pellet, kWh;
- Produced pellets *P*_Pellet, t;
- Crushed sawdust *P*_F, m^3.

RESULTS

Only the graphs where correlation between the values of dependent variables and independent variables can be seen are presented below. The dependence of electricity generation on the auxiliary power consumption of the cogeneration plant *C*_CHP is presented in Fig. 2.

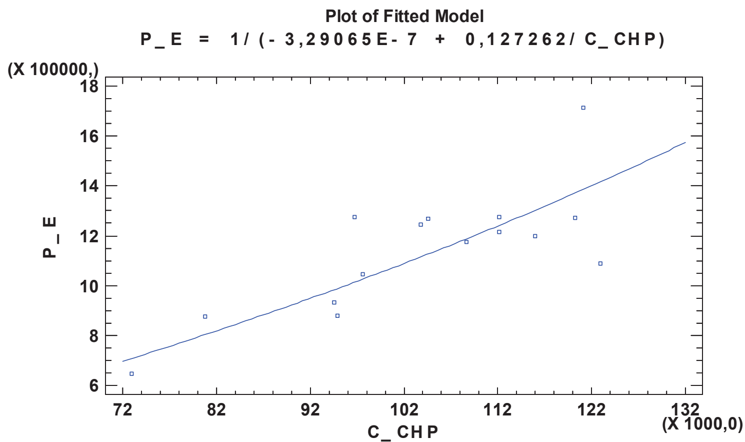

Figure 2. Produced electricity *P*_E depending of the CHP power consumption *C*_CHP.

In Fig. 2, it can be seen that there is a good mutual correlation between both variables. The value of the squared correlation coefficient as determined by the analysis is $R^2 = 0.75$ and the correlation coefficient $R = 0.87$. The relationship between the variables is non-linear and it is defined as follows, (2):

$$P\_E = 1/(-3.29065E\text{-}7 + 0.127262/C\_CHP), \tag{2}$$

The Eqn (2) explains 75% of the analysed changes in the data and it can be used for approximate calculations. 25% of generated electricity should be explained by the impact of other parameters.

The analysis of the data correlation shows that there is a certain correlation between the produced electricity *P*_E and the power consumption of the pellet production facility *C*_Pellet. The changes of the values are presented in Fig. 3.

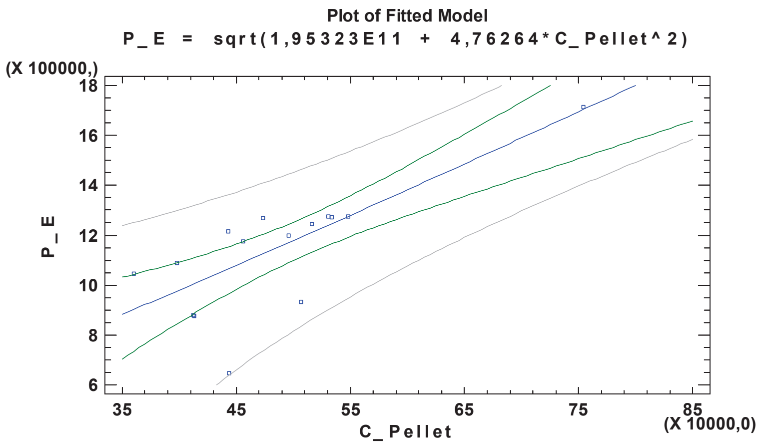

Figure 3. Produced electricity *P*_E depending on the power consumption of the pellet production facility *C*_Pellet.

The mutual correlation of the analysed variables is described by the value of the squared correlation coefficient $R^2 = 0.71$ and the correlation coefficient $R = 0.84$. The relationship between the variables is non-linear and it is defined as follows, (3):

$$P\_E = sqrt(1.95323E11 + 4.76264 \cdot C\_Pellet^2), \tag{3}$$

The mutual correlation of the variables is slightly lower than the correlation to the CHP electricity consumption. Eqn (3) explains 71% of the changes in the studied data. The impact of other parameters is higher, e.g. 29% of the electricity generation. The review of the correlation of other parameters demonstrates that there is no considerable correlation. Therefore, in further multi-factor regression analysis, the changes in the dependent variable of the produced electricity depend on two factors, i.e. the cogeneration plant power consumption *C*_CHP and the power consumption of the pellet production facility *C*_Pellet, Eqn (4):

$$P\_E = f(C\_CHP; C\_Pellet), \tag{4}$$

The performed correlation analysis of the data makes further regression analysis easier, as the set of factors that needs to be included in the multi-factor regression equation has been established.

Regression analysis of the data of power generation

The regression analysis is aimed at obtaining an empirical equation that would provide a quantitative description of the power generation depending on the indices that characterise the operation of the pellet production facility. These characteristics are statistically important and would serve as the basis for improving and evaluating the energy efficiency of the production facility. The regression analysis defines the accurate quantitative parameters of the change in random variables, i.e. explains the importance of the stochastic link by functional relationships.

The sequence of the regression analysis was as follows:

- the rule of the distribution of the dependent variable, i.e. the produced electricity *P*_E, was verified;
- the regression equation was established by applying the smallest square method;
- statistical analysis of the obtained results was performed.

The results of a regression analysis are correct if the rules for its application are complied with (Beloborodko et al., 2012). The number of rules is high and they cannot always be fully followed in practice. There are several main preconditions behind the application of the regression analysis. The use of the regression analysis of the data is correct if the normal distribution law is applicable to the dependent variable (produced electricity *P*_E). This requirement is not applicable to independent variables. The above means that the analysis starts with establishment of the distribution of dependent variables and the analysis may be continued if this distribution complies with the rule of the normal distribution. The results of verification of the rule of distribution are presented in Fig. 4. Normal distribution within logarithmic coordinates is graphically presented by a line. As can be seen in Fig. 4, the analysed data are placed close to the line in the graph. It means that the distribution is close to the rule of normal distribution and the application of the regression analysis is justified.

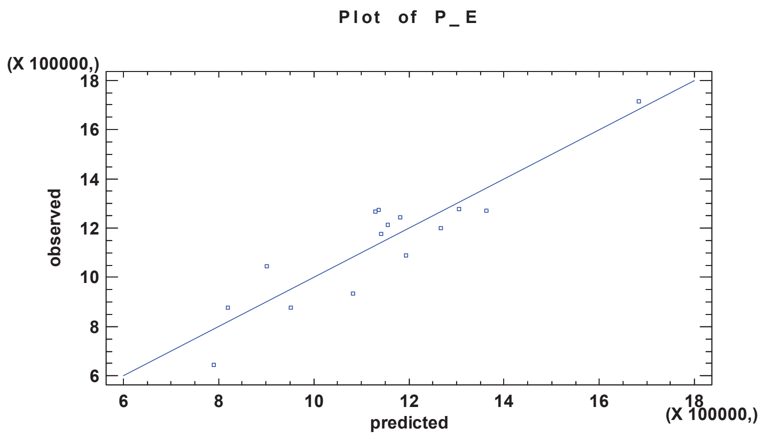

Figure 4. Distribution of the produced electricity *P*_E values.

When empirical models are developed in the form of the regression equation, several issues have to be solved. Whether the model comprises all the independent variables describing the analysed phenomenon and whether the model does not comprise unnecessary and non-essential variables, thus making the model too complicated. The answer to the above questions is provided by evaluation of the statistical importance of the variables contained in the model and the dispersion analysis of the model (Beloborodko et al., 2012).

The regression equation that is used by the author does not contain the effects of double and triple interaction of independent variables and is as follows in Eqn (5) (Beloborodko et al., 2012):

$$y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + \ldots + b \cdot x = b_0 + \sum_{i-1}^{n} b_1 \cdot x_1, \qquad (5)$$

where: y is the dependent variable; b_o is the free member of the regression; b_i... b_n are the regression coefficients and x_1...x_n are the independent variables.

The regression equation that corresponds to the Eqn (5) and was obtained as the result of statistical processing of the data and contains statistically important independent variables, as in Eqn (6):

$$P\_E = b_0 + b_1 \cdot C\_CHP + b_2 \cdot C\_Pellet \qquad (6)$$

The values of the coefficients of the regression equation and their statistical values are presented in Table 2.

Table 2. Coefficients of the regression equation and their values

Coefficients b_i	Values	t statistics	P value
Constant b_0	-5.27157	-2.36019	0.0360
Coefficient b_1	9.39524	4.27046	0.0011
Coefficient b_2	1.42205	4.10665	0.0015

In the data processing, the level of importance P = 0.1 was selected and it corresponds to the probability of credibility 0.90. For the purpose of evaluation of the importance of the coefficients b_0.....b_n of the regression Eqn (6), the t criterion with the Student's distribution with f freedom levels is applied in Eqn (7).

$$f = m - (n + 1), \qquad (7)$$

where m is the volume of the data which is the subject of the analysis, n is the number of independent variables in the regression equation.

The level of freedom is defined in Eqn (8):

$$f = m - (n + 1) = 12 - (2 + 1) = 9, \qquad (8)$$

The value of the t criterion corresponding to these values as taken from the tables of the Student's distribution is $t_{tab} = 1.9$. As can be seen from Table 2, the relationship (Blasnik, 1995) > t_{tab} is valid in all cases. It means that all the parameters are important and should be maintained in the equation.

The study has resulted in obtaining a regression equation that defines produced electricity depending on the data of the production facility, i.e. the cogeneration plant power consumption C_CHP and the power consumption of the pellet production facility C_Pellet in Eqn (9):

$$P\_\mathrm{E} = -5.27157 + 9.39524 \cdot C\_\mathrm{CHP} + 1.42205 \cdot C\_\mathrm{Pellet}, \tag{9}$$

The value of R^2 as determined as the result of statistical processing of the data which was established in the empirical model equals 0.83. It means that the established model (8) explains 83% of the change in the analysed data. The remaining 17% refer to independent variables that have not been included in the equation or defined in the study or the effect of their mutual interaction.

Evaluation of adequacy of the regression equation

Evaluation of the adequacy of the equation (9) is performed by means of the dispersion analysis by applying the Fisher's criterion F. For this purpose, the relationship between the dispersion of the dependent variable and the balance dispersion is analysed, Eqn (10):

$$F(f_1 f_2) = \frac{S_y^2(f_1)}{S_{atl}^2(f_2)} \tag{10}$$

where: $S^2_y(f_1)$ is y dispersion of the dependent variable and $S^2_{\mathrm{atl}}(f_2)$ is the balance dispersion.

The balance is defined as the difference between the dependent variable and the value calculated by means of the regression equation $y_i - y_{iapr}$.

The value as determined by means of the dispersion analysis performed by the software is F = 19.16. The obtained value is compared to the table value of the criterion, which is determined by applying the values of the freedom levels, Eqn (11):

$$f_1 = m - 1 = 12 - 1 = 11 \text{ and } f_2 = m - n = 12 - 2 = 10 \tag{11}$$

The table value of the Fisher's criterion is $F_{\mathrm{tab.}} = 2.9$. As can be seen, the relationship $F > F_{\mathrm{tab}}$ is valid and it means that the Eqn (9) is adequate and can be used for describing the analysed data within the framework of their change:

- For produced electricity from 0.65 to 1.72 GWh per month;
- For the CHP power consumption C_CHP from 0.07 to 0.12 GWh per month;
- For the power consumption of the pellet production facility C_Pellet from 0.36 to 0.75 GWh per month.

Verification of the rules of correct application of the correlation analysis

When following establishment of the regression equation, it is possible to perform verification of the rules of correct application of the regression analysis based upon a range of other indices. These are autocorrelation, multicollinearity and heteroscedasticity.

By means of application of the Durbin-Watson's (DW) test, in the course of statistical treatment of the data and the data analysis, the DW criterion has been established. Its value equals 2.3 and exceeds the marginal value of 1.4. This means that there is no considerable autocorrelation of the balance and assessments of the values by means of the smallest squared values method in the course of the analysis are not performed.

The verification has been performed by analysing the correlation matrix of the coefficients calculated by means of the regression equation, presented in Table 3.

Table 3. Correlation matrix of the coefficients of the regression equation

Coeff.	Const.	*C*_CHP	*C*_Pellet
Const.	1.0000	-0.7155	-0.3338
*C*_CHP	-0.7155	1.000	-0.4095
*C*_Pellet	-0.3338	-0.4095	1.000

The analysis of the correlation matrix of the coefficients of the regression equation demonstrates that there is no considerable correlation between the coefficients and independent variables. This is attested by the low values of the correlation coefficient in Table 2. The values presented in the Table are below 0.5 or close to this level, and this means that the evaluation of the coefficients of the regression equation is correct.

The verification of the heteroscedasticity has been performed by means of graphic analysis of the distribution of balances depending on the cogeneration plant power consumption *C*_CHP. If an increase of variations can be seen in graphs (the points form a triangle or a wedge), it means that there is heteroscedasticity. The distribution of the balances is presented in Fig. 5.

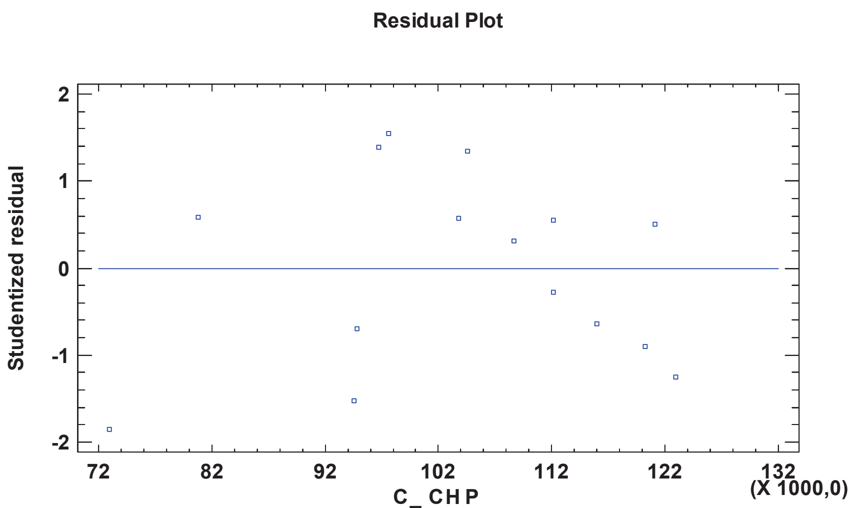

Figure 5. Distribution of balances depending on the cogeneration plant power consumption *C*_CHP.

In Fig. 5, it can be seen that there are no considerable changes in the distribution of balances depending on the cogeneration plant power consumption *C*_CHP. The values of the balances are similar along the whole range of changes in *C*_CHP. The distribution of the balances has been analysed based on other factors. In all cases, the conclusion has been that there is no heteroscedasticity and the standard error has been identified correctly.

One of the types of verification of the regression equation is related to the verification of the signs of its constituents and the fact that there is a logical explanation behind them. The identified changes in the equation from the physical essence perspective are described in the processes. In the regression equation for determining the produced electricity *P*_E (9), the signs of all the parameters are positive and an increase in their values causes an increase in produced electricity *P*_E. When the CHP power consumption *C*_CHP is increased, the produced electricity increases because, for a power plant to be able to generate more electricity, it has to consume more resources. The visible trends comply with the essence of the processes and there is a logical explanation behind them.

The question as to how complete is the correlation between the results calculated by means of the regression equation and the analysed data is among the basic questions regarding the use of empirical equations. It can only be stated in the case of satisfactory correlation that the model adequately describes the situation in practice and its use for simulating the situation is correct. For the purpose of verification of the adequacy of the empirical equation, the empirical and calculated data have been compared. The graphic presentation of the data comparison is in Fig. 6.

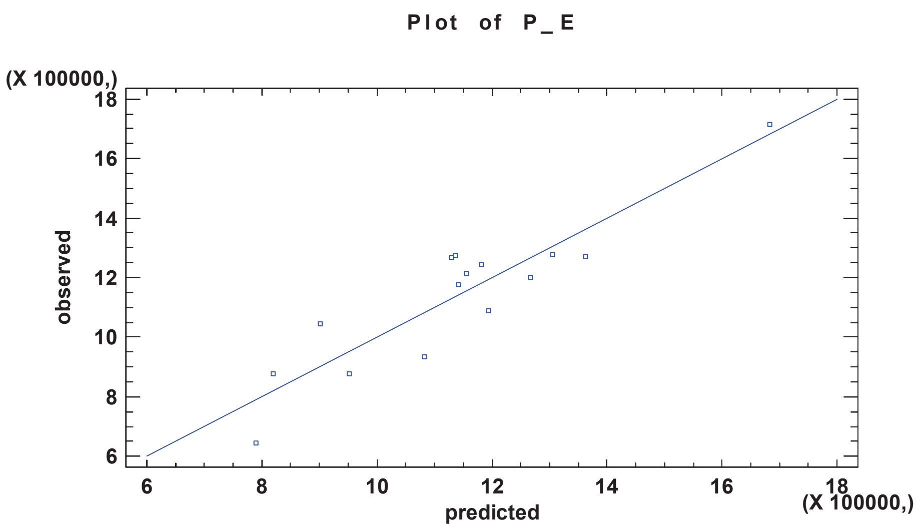

Figure 6. Comparison of the analysed and calculated data of produced electricity.

As can be seen from Fig. 6, there is a good correlation between both data sets. If the calculated value corresponded accurately to the surveyed data, the points would be located on the line in the figure. There is an increased dispersion of points at low values of the reduction of power generation.

CONCLUSION AND FUTURE WORK

Using the statistical processing and applying the methods of regression analysis, the most important factors describing the operation of the production equipment were identified. The relationship between the produced electricity and the parameters impacting this is defined by the regression equation which was obtained during the data processing. During the regression analysis, tests were performed at every stage regarding the correctness of the implemented steps.

According to the performed analysis, the electricity produced is determined by two statistically important parameters, cogeneration plant power consumption and power consumption of the pellet production facility. The adequacy of the equation was verified by applying the Fisher's criterion. The equation describes 83% of the changes in produced electricity.

This study has shown that there is a possibility to find a good equation which describes some independent values using variable values. This study shows that there is a possibility to make a model which describes all factory processes. This additionally enables to use this model for demand side management and hence improve the economic feasibility of the facility.

In future, this task should be studied further. More data must be gathered about another values, analysed and put in a model which can describe factory work.

ACKNOWLEDGEMENTS. This study is supported by the European Social Fund’s Doctoral Studies and Internationalization Program DoRa, Estonian Archimedes Foundation program.

REFERENCES

Anderson, J.O. & Toffolo, A. 2013. Improving energy efficiency of sawmill industrial sites by integration with pellet and CHP plants. *Applied Energy* **111**, 791–800.

Beloborodko, A., Timma, L., Žandeckis, A. & Romagnoli, F. 2012. The regression model for the evaluation of the quality parameters for pellets. *Agronomy Research* **10**(S1), 17–24.

Blasnik, M. 1995. The need for statistical analysis and reporting requirements: some suggestions for regression models. *Energy Program Evaluation Conf.*, Chicago, US, p. 14.

Kohl, T., Laukkanen, T., Järvinen, M. & Fogelholm, C.J. 2013. Energetic and environmental performance of three biomass upgrading processes integrated with a CHP plant. *Applied Energy* **107**, 124–134.

Mikita, V., Roots, J. & Olt, J. 2012. Simulation model of the combustion processes of diesel engine. *Agronomy Research* **10**(S1), 157–166.

Mola-Yudego, B., Selkimäki, M. & González-Olabarria, J.R. 2013. Spatial analysis of the wood pellet production for energy in Europe. *Renewable Energy* **63**, 76–83.

Mobini, M., Sowlati,, T. & Sokhansanj, S. 2013. A simulation model for the design and analysis of wood pellet supply chains. *Applied Energy* **111**, 1239–1249.

Savola, T., Tveit, T.M. & Fogelholm, C.J. 2007. A MINLP model including the pressure levels and multiperiods for CHP process optimisation. *Applied Thermal Engineering* **27**, 1857–1867.

Song, H., Dotzauer, E., Thorin, E. & Yan, J. 2011. Annual performance analysis and comparison of pellet production integrated with an existing combined heat and power plant. *Bioresource Technology* **102**, 6317–6325.

Söderman, J. & Ahtila, P. 2010. Optimisation model for integration of cooling and heating systems in large industrial plants. *Applied Thermal Engineering* **30**, 15–22.

Revina, I. 2002. Ekonometrija. *EuroFacultv*, Rīga, pp. 271.

Lehtikangas, P. 2000. Storage effects on pelletised sawdust, logging residues and bark. *Biomass and Bioenergy* **19**, 287–293.