Regression model in the operation of wind turbines

Sylwester Borowski[*]*, Klaudiusz* Migawa, *Agnieszka* Sołtysiak, *Andrzej* Neubauer and *Anna* Mazurkiewicz

[1] Faculty of Mechanical Engineering, University of Science and Technology, Kaliskiego 7, 85-796, Bydgoszcz, Poland

Abstract. The manuscript presents linear regression models used to predict the amount of energy produced in a wind turbine. As part of the own research, data obtained from the SCADA program for the Enercon E-82 wind turbine was analyzed. It has been shown that it is possible to build a mathematical model to determine the amount of energy generated from the average wind speed. This method can be used to forecast the production volume and balance the energy system. This is important due to the increasing share of renewable energy sources. The application of the developed method in practice will facilitate and accelerate the implementation of the decision-making process in electricity generation systems, while reducing the risk of error. This model can also be used to make wind turbine repowering decisions.

1 Introduction

In many countries, including Poland, an increase in the share of renewable energy in the national energy balance is observed. However, renewable energy processors are unstable. This applies to wind and photovoltaic energy. For wind farms, this is due to the changing wind speed. The national energy system must be based on the knowledge of energy demand and the possibility of its production. Due to this nature of work, wind and solar farms are troublesome. In the case of high variability of electricity production, it is necessary to maintain other sources. These sources include coal, nuclear and hydro power plants. An "energy mix" is created based on electricity from renewable sources and fossil fuels. However, energy storage facilities are being built more and more often. Energy storage can be built in the form of hydropower plants, flywheels, or more and more often batteries. A system constructed in this way must ensure a high potential for reliability of electricity supply. Due to this, a very complex system is created. Its maintenance is becoming more and more expensive [1, 2].

One of the most important costs related to the maintenance of wind farms are its maintenance and repair activities. In order to use the production potential, activities related to the diagnosis of the technical condition are constantly carried out. All parameters that

* Corresponding author: sylwester.borowski@utp.edu.pl

may indicate technical condition are measured, saved and analyzed. Complex diagnostic systems are maintained and repairs are anticipated. This approach allows for technical readiness of wind turbines. Therefore, the phenomenon of replacing old models of wind turbines with new technical solutions is observed. The second reason for replacing old wind turbines is repowering. New turbines have a higher power rating or a higher efficiency. This phenomenon is beneficial for the investor and the energy system. A wind turbine with greater reliability is introduced into the system and the installed capacity increases. New turbines are built using the experience of designers, using new technologies and materials. For this reason, they can be larger and produce more energy. The second element that improves the operation of new wind turbines are energy production control systems. Production forecasting and wind resource modeling are also introduced to the system. The energy system operator now has greater possibilities to forecast energy supplies from wind turbines. The replacement of wind turbines with new models should be based on the information gathered from the old turbines currently in use. Better knowledge of wind behavior results in an optimal wind farm design [3].

Investors are aware that their wind turbines may be archaic. However, nowadays it is more and more difficult to find new places, especially on land, where new constructions can be installed. Therefore, old wind turbines that are worn out are disposed of. In their place, you can build new ones that bring better income. Many old wind turbines are disposed of before they are completely destroyed. Investors want new wind turbines for which the profits will be greater. Replacing the old turbine with a new generation turbine results in an increase in net energy production. The repowering process also includes the modernization of the installation with an increase in power and efficiency [4].

Replacing an old turbine with a new one is very costly. The investor should have the knowledge that helps him make a decision.

In this case, the amount of electricity produced by the old wind turbine is known. However, the amount of energy that could be produced at the same site by a new wind turbine is unknown. Therefore, before making a decision, one should also balance the amount of energy that can be produced by a new wind turbine. Relying on the production characteristics of a wind turbine and on the wind speed at the installation site is troublesome. A good solution is to use a mathematical model that helped to do this based on data obtained from the old turbine. It is then possible to easily compare the amount of energy produced by the old turbine and the "theoretical" energy from the new one [5].

Another place where such a model can be used is to use it to estimate the wear of the turbine blades. During the use of a wind turbine, its blades degrade. The resistance of such a blade increases and the amount of converted energy decreases. These resistances change the power curve of the turbine. Worn blades also increase the production of noise and vibration [6, 7, 8, 9].

The aim of the article is to present analytical research on the construction of a simple mathematical model that allows to forecast electricity production depending on the average wind strength. The model indicated in the article [10] was used in the manuscript. The model is built on the basis of historical data obtained from the Enercon SCADA software [11].

2 Methodology

The model built in the article [10] is based on data obtained from the Wicko wind farm. The plant consists of five Enercon E-82 wind turbines. As part of the research, historical data was obtained and subjected to statistical analysis. The main aim of the research was to check the developed regression model. The tested model makes it possible to determine the average daily amounts of electricity generated by wind turbines depending on the average

daily wind speed. The tested model is based on historical values obtained from the SCADA system in 2019. Therefore, it is appropriate for the analyzed wind turbine. This is due to the lack of analysis of other factors that may depend on the location of the wind turbine. These factors include roughness of the terrain, daytime temperature and other disturbing factors [12, 13, 14]. The check was made on the basis of data from previous years (2018 and 2017).

As part of the analyzes, linear regression models were developed. The influence of the average daily wind speed on the amount of energy generated during various time intervals was analyzed. Analyzes were performed for days, weeks and months.

The tested regression models were developed on the basis of data obtained from the Enercon SCADA system. The data presented the average daily wind speed [m s^{-1}] and the average daily electricity production [kWh] in the analyzed period. Wind speed is measured with a gauge placed on the nacelle of the wind turbine.

3 Analysis of the results

3.1 Linear regression model developed for individual months of 2019

The article [10] proposes a linear regression model. It was built on the basis of data obtained from the Enercon SCADA system. The model determines the dependence of the average daily amount of electricity produced on the average daily wind speed.

Formula (1) shows the linear regression function developed for individual months of 2019. The tables present the statistics of the developed regression model.

$$E_m = 4398.12 \ v_m - 14404.94 \tag{1}$$

were:

E_m – average daily energy [kWh],
v_m – average daily wind speed [m s^{-1}].

Table 1. Statistics of the linear regression model developed for individual months of 2019

Coefficient	Value of the coefficient 2018	Value of the coefficient 2017
Correlation coefficient R	0.9905	0.9874
R^2	0.9811	0.9750
Adjusted R^2	0.9792	0.9725
Standard error	574.3714	694.9133
Number of observations	12	12

Then, for the linear regression model represented by the formula (1), the values of the average daily amount of electricity produced were determined depending on the average daily wind speed. The obtained results are shown in Figures 1 and 2.
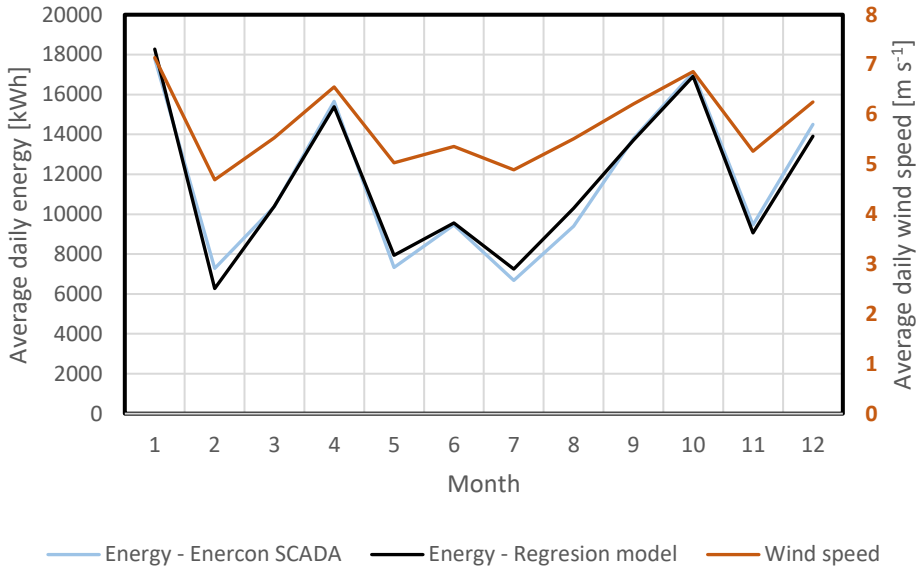
Fig. 1. Average monthly energy [kWh] generated by the Wicko wind farm in the following months of 2018 determined on the basis of real data (Enercon SCADA) and for the linear regression model
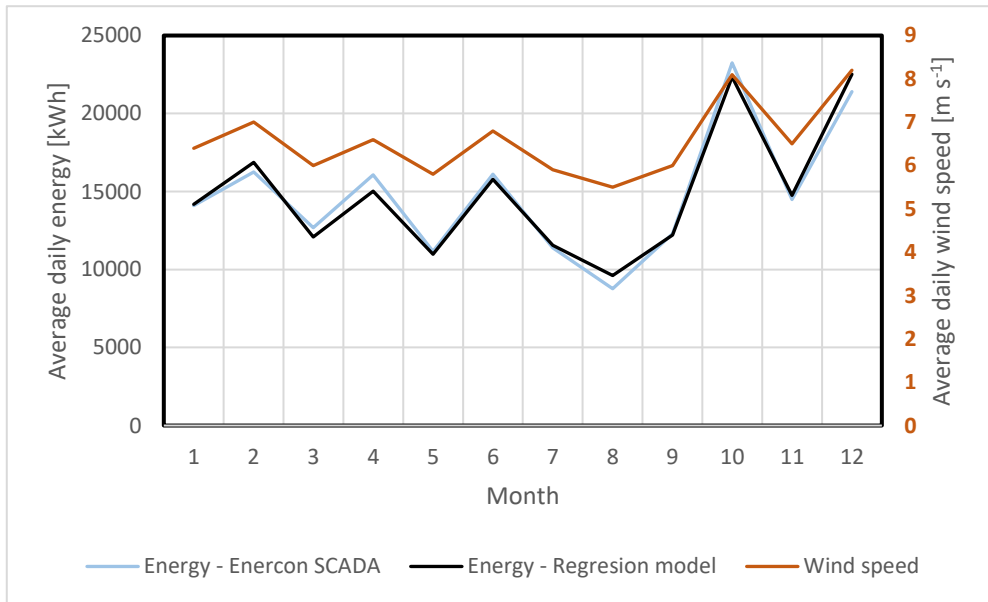

Fig. 2. Average monthly energy [kWh] generated by the Wicko wind farm in the following months of 2017 determined on the basis of real data (Enercon SCADA) and for the linear regression model

3.2 Linear regression model developed for each week of 2019

The article [10] proposes a linear regression model. It was built on the basis of data obtained from the Enercon SCADA system. The model determines the dependence of the average daily amount of electricity produced on the average daily wind speed.

Formula (2) presents the linear regression function developed for individual weeks of 2019. The tables present the statistics of the developed regression model.

$$E_{m} = 5022.27 \ v_{m} - 17960.18 \qquad (2)$$

were:

E_m – average daily energy [kWh],

v_m – average daily wind speed [m s^{-1}].

Table 2. Statistics of the linear regression model developed for individual weeks of 2019

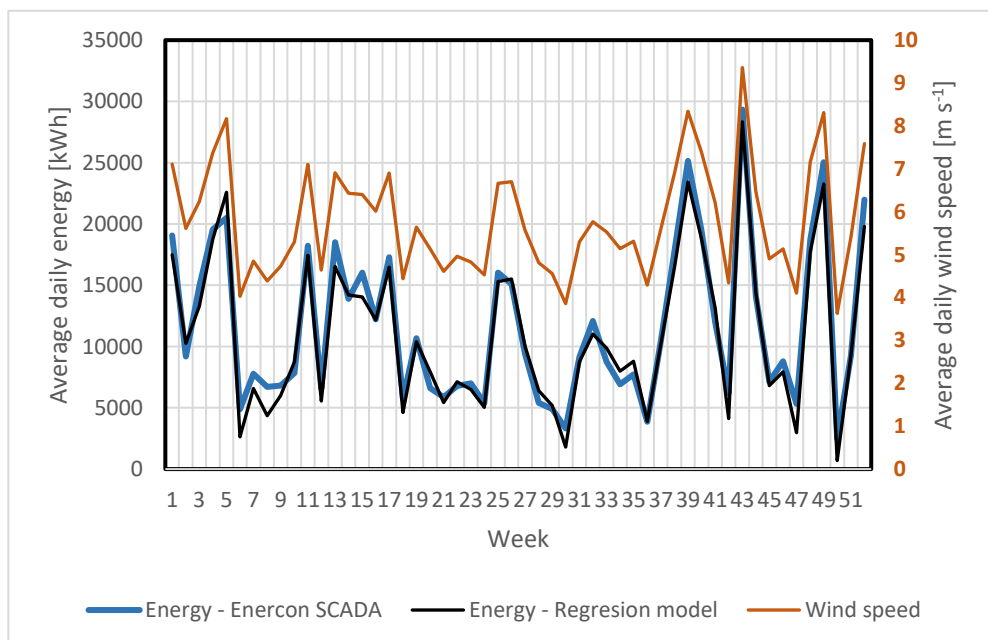| Coefficient	Value of the coefficient 2018	Value of the coefficient 2017
Correlation coefficient R	0.9848	0.9713
R^2	0.9699	0.9435
Adjusted R^2	0.9693	0.9423
Standard error	1128.6410	1813.2086
Number of observations	52	52

Fig. 3. Average monthly energy [kWh] generated by the Wicko wind farm in the following weeks of 2018 determined on the basis of real data (Enercon SCADA) and for the linear regression model

Fig. 4. Average monthly energy [kWh] generated by the Wicko wind farm in the following weeks of 2017 determined on the basis of real data (Enercon SCADA) and for the linear regression model
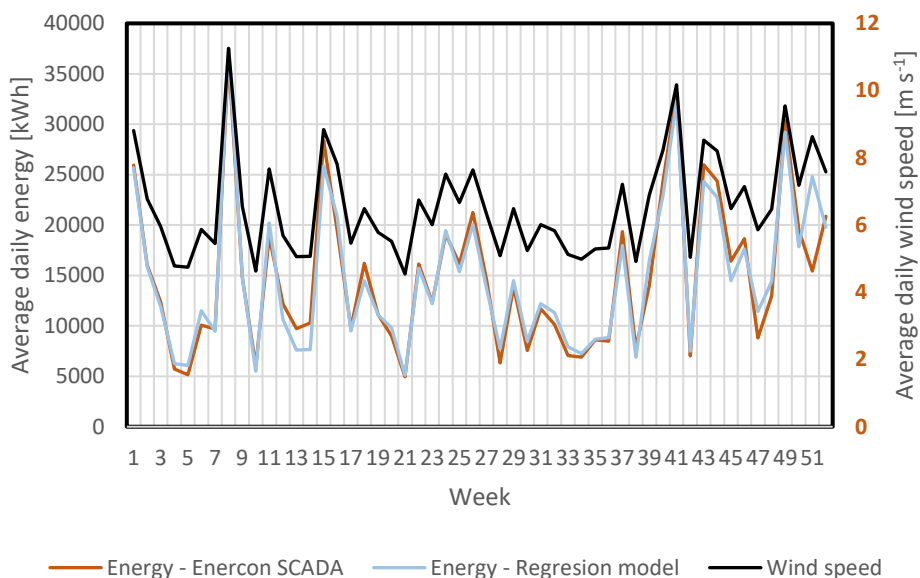
4 Summary

The manuscript presents the validation of the regression model developed for a specific wind turbine from the Wicko power plant. The developed regression models based on the data from 2019 show a good fit with the data obtained from 2017 and 2018. The R2 coefficient is in all cases above 0.92. For the week analysis, it is 0.97 and 0.94. For the analysis of months, even above 0.97 for both years. It means that the applied mode explains well the variability of electricity production due to the wind force.

The abbreviated research results presented above constitute the next stage of the research. Their goal is to develop a predictive method based on artificial intelligence solutions and expert knowledge. This method will be useful primarily for forecasting the volume of electricity production depending on the forecast wind strength. This will contribute to improving the perception of renewable energy sources in the bill of energy. The proposed method will be developed using predictive models that enable the analysis of time series (stationary and non-stationary), such as the ARMA and ARIMA models, as well as neural networks. Methods of this type are necessary to support the decision-making process in electricity generation systems. This reduces the risk of making a mistake (making wrong decisions). It is also possible to use the developed models to indicate the benefits of repowering.

References

1.　B. Parsons, M. Milligan, B. Zavadil, D. Brooks, B. Kirby, K. Dragoon, and J. Caldwell, Grid impacts of wind power: A summary of recent studies in the United States, in Proc. EWEC, Madrid, Spain, 2003.

2. S. Borowski, M. Szubartowski, L. Knopik, K. Migawa: Wind turbine condition monitoring system as a source of diagnostic information. In Procedings of the 17th International Conference Diagnostics of Machines and Vehicles (2018) https://doi.org/10.1051/matecconf/201818201015

3. M.I. Blanco, Renew. Sust. Energ. Rev. 13 (2009)

4. G. Szala, B. Ligaj, Fatigue failure and fracture mechanics, 69-76 (2012)

5. K. Migawa, L. Knopik, A. Sołtysiak P. Kolber, The method of risk assessment in transport system. Engineering Mechanics (2017)

6. M. Karwowska, J. Mikolajczak, S. Borowski, Z.J. Dolatowski, J. Marc-Pienkowska, W. Budzinski, Ann. Anim. Sci. 14, 441-451 (2014)

7. J. Marć-Pieńkowska, Z. Podkówka, S. Borowski. J. Cent. Eur. Agric. 20(4), 1055-1062 (2019)

8. M. Karwowska, J. Mikolajczak, Z.J. Dolatowski, S. Borowski, Ann. Anim. Sci. 15,1043-1054 (2015)

9. L. Knopik, K. Migawa, Maintenance and Reliability 21, 2 (2019)

10. S. Borowski, A. Sołtysiak, K. Migawa, A. Neubauer. MATEC Web of Conference, 332, 1-7 (2021)

11. L. Knopik, K. Migawa, A. Wdzieczny, Polish Maritime Research 23, 1 (2016)

12. L. Knopik, K. Migawa, P. Kolber: Statistical analysis of parameters of rail vehicles. Engineering Mechanics (2016)

13. K. Migawa, L. Knopik, A. Neubauer, D. Perczynski, Simulation of the model of technical object availability control. Engineering Mechanics (2017)

14. K. Migawa, L. Knopik, S. Wawrzyniak, Application of genetic algorithm to control the availability of technical systems. Engineering Mechanics (2016)