

SESSION E: THERMAL TECHNIQUE AND ENVIRONMENTAL PROTECTION

Analysis of the influence of the rotation speed of the DFIG wind turbine generator on the temperature of the generator

Gojko Krunic^{1*}, Srđan Vasković², Srećko Krile³, Aleksandra Ijačić¹

¹ University of East Sarajevo, Faculty of Production and Management, Trebinje, Bosnia and Herzegovina

² University of East Sarajevo, Faculty of Mechanical Engineering, East Sarajevo, Bosnia and Herzegovina

³ Rpglobal projekti d.o.o., Zagreb, Croatia

ARTICLE INFO

* Correspondence: gojko.krunic@fpm.ues.rs.ba

DOI: <https://doi.org/10.46793/ET26.E02K>

PAGES: E11 – E20

ABSTRACT

The paper analyzes the influence of the rotation speed of a DFIG (Doubly-Fed Induction Generator) wind turbine generator on the temperature in the operating mode and in the idle state, based on one-year data from the SCADA (Supervisory Control and Data Acquisition) system. The aim is to quantitatively determine the intensity and statistical significance of this relationship. The methodology includes data filtering, linear regression and correlation analysis, with the calculation of the Pearson coefficient (r) and the coefficient of determination (R^2). The accuracy of the model was assessed by RMSE and MAE, while statistical validation was carried out by ANOVA analysis and residual analysis. The results show a moderately strong positive correlation between rotation speed and temperature ($r = 0.7034$), with $R^2 = 0.4949$ indicating that speed changes explain about 49.5% of the variability. In idle state, the temperature depends mainly on external factors (ambient temperature, wind, solar radiation, humidity and surrounding objects). The proposed model enables reliable temperature prediction in the operating mode and can be used for operation optimization, monitoring and predictive maintenance of wind turbines.

KEYWORDS

Generator rotation speed, Wind turbine generator, Generator temperature, Linear regression, Residuals

1. INTRODUCTION

Wind energy is one of several renewable energy sources used in the production of electricity due to rising electricity consumption and environmental pollution [1]. Advances in wind energy technology have significantly decreased the cost of producing electricity from this renewable source. Nowadays, the generation of energy from wind sources plays a crucial role to increasing the green energy [2]. WPPs have become more popular in recent decades as they are cleaner, greener, and more sustainable [3].

The most prevalent wind power generator is the variable-speed wound-rotor doubly fed induction generator (DFIG) [4]. The DFIG technology has the advantage that it extracts maximum energy from the wind for low wind speeds by optimizing the turbine speed, also by minimizing the mechanical stresses on the turbine during sudden gusts of wind [5]. Doubly Fed Induction Generators (DFIGs) are equipped with a rotor winding that is powered through a power electronic converter. This configuration enables efficient control of both the generator's rotational speed and output power [6]. The stator of DFIG is directly connected to the grid, while the rotor is connected to the grid through a back-to-back power electronic converter, which is composed of a rotor-side converter (RSC) and a grid-side converter (GSC) [7]. A gearbox couples the rotor and generator of DFIG to manage the speed range differences between rotor and generator speed [8]. Bidirectionally powered induction generator of wind turbine is shown in Figure 1 [9].

Thermal analysis of a generator is a complex problem where a number of heat exchanges are involved. Since performance and efficiency of a generator depends on the thermal condition of its windings, understanding the power losses and the thermal mechanism within a DFIG are of great importance to the industry [10]. The rotation speed of a DFIG generator is a fundamental factor that directly affects the thermal load and functional reliability of the generator. DFIG systems are

designed to operate at variable rotation speeds, which allows for maximum utilization of the kinetic energy of the wind and improves the efficiency of electricity generation. Changes in rotation speed directly affect the stator and rotor currents, causing changes in copper and iron losses, which leads to an increase in the generator temperature.

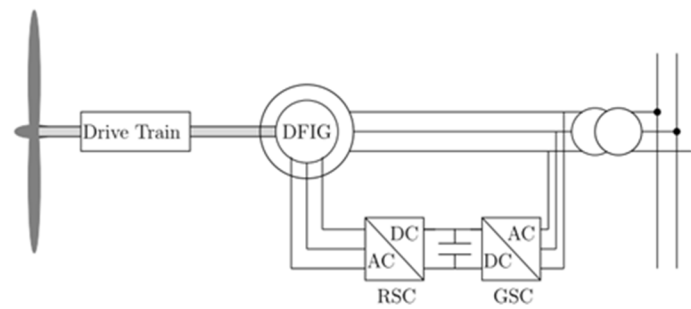


Figure 1: Bidirectionally powered induction wind turbine generator

Power losses of the generator and the temperature of stator winding increase with wind speed. As the induction generator torque is subject to the rotor current and stator voltage, therefore any unbalance of the stator voltage will cause an unstable rotor current in order to maintain an appropriate torque value rather than to maintain a constant power output. Therefore, an increase of stator voltage will cause a slight increase of stator current, and thus an increase of winding temperature [10]. The increase in power dissipation and changes in thermal conductivity lead to an increase in the temperature of the stator winding, and therefore of the entire generator.

The contribution of the work is reflected in the quantification of the influence of rotation speed on the temperature load of the generator based on measurements at one representative point, with a separate analysis of the operating mode and the idle state. The work applied linear regression analysis to determine the dependence between rotation speed and temperature, using Pearson's correlation coefficient, regression function parameters (slope and intercept), coefficient of determination, as well as statistical indicators of model accuracy and significance, including RMSE, MAE, p -value, t -statistics and residual analysis.

However, it is important to emphasize that the temperature analysis in this paper is based on measurements at a single selected point of the generator, which is considered representative for assessing the overall thermal state of the system. Although in real conditions the temperature of the generator is monitored at multiple locations, including the stator and rotor windings, as well as other critical parts, this approach implies a certain level of simplification. Accordingly, the obtained results describe local temperature characteristics and do not encompass the complete spatial distribution of temperature within the generator. The aforementioned limitation should be kept in mind when interpreting the results, while future research may be directed towards the application of multi-point measurements and more advanced thermal models.

The paper is organized as follows: after the introduction, the mathematical foundations of linear regression are presented in the second chapter, the analysis of the influence of the rotation speed on the generator temperature is given in the third chapter, while the conclusions and an overview of the literature used are presented at the end of the paper.

2. MATHEMATICAL BASIS OF LINEAR REGRESSION

The analysis of the influence of the rotation speed of a wind turbine DFIG generator on the generator temperature is based on the application of appropriate statistical methods that allow quantitative description and measurement of the degree of correlation between the observed quantities. In this context, simple linear regression represents the basic model for examining the linear relationship between the rotation speed as an independent variable and the generator temperature as a dependent variable.

In order to enable a detailed and quantitative analysis of the observed dependence, the mathematical foundations of the Pearson correlation coefficient (r), the coefficient of determination (R^2), as well as the parameters of the linear regression function, including the slope and intercept of the regression line, are presented below. In addition, the basic equation of the simple linear regression model is given, which allows for the description and prediction of the change in generator temperature depending on the change in rotation speed.

Pearson's correlation coefficient (r) is an indicator of the linear relationship between two quantitative variables. Its value reflects the direction and strength of the relationship between the variables being analyzed. In other words, it shows how much a change in one variable affects a change in the other, either in the same (positive correlation) or opposite (negative correlation) direction.

Based on the value of the Pearson correlation coefficient, the direction and strength of the linear relationship between the observed variables can be determined.

The direction and strength of the linear association of the variables depending on the value of the Pearson correlation coefficient are shown in Table 1 [11].

Table 1: The direction and strength of the linear association of the variables depending on the value of the Pearson correlation coefficient

The value of the Pearson correlation coefficient	Direction and strength of linear correlation of variables
$r = -1$	functional negative correlation
$r = 1$	functional positive correlation
$-1 < r \leq -0,8$	strong negative correlation
$0,8 \leq r < 1$	strong positive correlation
$-0,8 < r \leq -0,5$	medium strong negative correlation
$0,5 \leq r < 0,8$	medium strong positive correlation
$-0,5 < r < 0$	weak negative correlation
$0 < r < 0,5$	weak positive correlation
$r = 0$	there is no correlation

A scatter plot shows the relationship between two numerical variables using dots. A scatter diagram showing data with zero, complete positive, incomplete positive, complete negative, and incomplete negative correlations is shown in Figure 2.

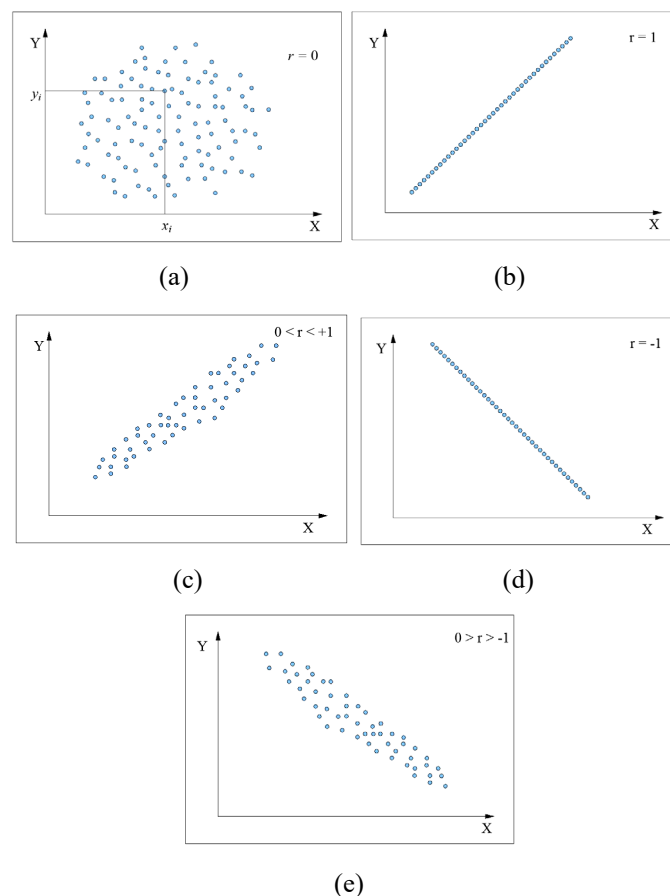


Figure 2: Scatter plot showing data with: (a) zero correlation, (b) complete positive correlation, (c) incomplete positive correlation, (d) complete negative correlation, and (e) incomplete negative correlation

The value of Pearson's correlation coefficient ranges from -1 to $+1$. A value of $r = 0$ indicates the absence of a linear relationship between variables, which does not necessarily mean that there is no relationship between them, but only that the relationship is not linear. A value of $r = +1$ indicates a complete positive linear relationship between the variables, while $r = -1$ indicates a complete negative linear relationship between the variables.

A complete positive linear relationship means that an increase in the value of one variable always follows an increase in the value of the other, that is, that the variables move in the same direction. Incomplete positive correlation exists when two variables tend to increase together, but the relationship is not perfectly linear. In this case, the increase in one variable usually follows the increase in the other, but the points in the scatter plot do not lie exactly on a straight line.

A perfect negative linear relationship means that an increase in the value of one variable is accompanied by a simultaneous decrease in the value of the other. An incomplete negative correlation exists when an increase in one variable often accompanies a decrease in the value of the other, but the relationship is not perfectly linear and the points in the scatterplot

do not lie exactly on a straight line, meaning that the variables change in opposite directions. The closer the value of the Pearson coefficient is to the values 1 or -1 , the stronger the linear correlation.

Pearson's correlation coefficient is represented by the following equation [12]:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (1)$$

where:

- x_i, y_i - individual values of the observed variables related to the rotation speed and temperature of the wind turbine generator.
- $\bar{x}, \bar{y}$ - arithmetic means of data sets related to the rotation speed and temperature of the wind turbine generator, respectively.
- n - sample size.

The coefficient of determination (R^2) in linear regression indicates how well the model explains changes in the dependent variable (wind turbine generator temperature) by the independent variable (wind turbine generator rotation speed). The coefficient of determination (R^2) is defined as the square of the Pearson correlation coefficient, i.e. $R^2 = r^2$.

The coefficient of determination (R^2) in linear regression is represented by the following equation:

$$R^2 = r^2 = \left(\frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \right)^2 \quad (2)$$

The value of the coefficient of determination is in the interval from 0 to 1, i.e. $0 \leq R^2 \leq 1$. The value of $R^2 = 0$ means that the linear model does not explain changes in the dependent variable at all, while $R^2 = 1$ indicates that the model perfectly explains all changes in the dependent variable. A higher value of the coefficient of determination indicates a better ability of the regression model to describe changes in the dependent variable.

Equation (2) represents the basis for analyzing the quality of a linear regression model, as it allows quantifying the degree and reliability of the linear dependence between two variables.

The slope b in a linear regression model indicates how much the dependent variable changes when the independent variable increases by one unit.

The slope in linear regression is represented by the following equation [13]:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3)$$

where:

- x_i - individual values of the rotation speed of the wind turbine generator.
- y_i - individual temperature values of wind turbine generators.
- $\bar{x}$ - the arithmetic mean of all values of the rotation speed of the wind turbine generator.
- $\bar{y}$ - the arithmetic mean of all the temperature values of the wind turbine generator.

The slope shows the direction and strength of the linear relationship between the independent and dependent variable. If the slope is positive ($b > 0$), it means that the value of the dependent variable increases as the value of the independent variable increases. In this case, the diagram shows a straight line going up from left to right.

If the slope is negative ($b < 0$), the value of the dependent variable decreases as the independent variable increases, which is shown in the diagram by a straight line going downward from left to right. In the case where the slope is zero ($b = 0$), there is no linear trend between the independent and dependent variables. The line is horizontal, indicating that changes in the independent variable have no effect on the values of the dependent variable.

The intercept in a linear regression model indicates the value of the dependent variable when the independent variable is zero. In the linear regression model, the intercept is represented by the following equation [13]:

$$b_0 = \bar{y} - b_1 \bar{x} \quad (4)$$

The intercept indicates the point where the regression line intersects the y-axis. A simple linear regression model is represented by the following equation [13]:

$$E(Y|x) = \beta_0 + \beta_1 x \quad (5)$$

where:

- $E(Y|x)$ - predicted (expected) value of the dependent variable Y for a given level of the variable x .
- β_0 - the intercept of the regression line.
- β_1 - the slope of the regression line.

A simple linear regression model is an analytical method that describes and quantifies the degree of linear relationship between an observed dependent and independent variable. A simple linear regression model is used in this paper, since the influence of one independent variable (generator rotation speed) on a dependent variable (generator temperature) is analyzed. The purpose is to determine the appropriate regression line that best depicts changes in the dependent variable in relation to the independent variable.

3. RESULTS AND ANALYSIS OF THE DEPENDENCE OF DFIG GENERATOR TEMPERATURE ON ROTATION SPEED

This chapter analyzes the influence of wind turbine generator rotation speed on generator temperature. The data used in the analysis were collected from one wind turbine of the Rudine wind farm. The generator temperature data were taken from the SCADA system and refer to one measurement point, i.e. one built-in temperature sensor, while at the same time the rotation speed data is included. The original dataset contains 52,510 time series records, which represents a high measurement resolution, but at the same time a significant processing load in standard software tools, including Microsoft Excel.

Downsampling is subsetting data such that the timestamps occur at a lower frequency than in the original time series [14]. In order to efficiently process data while preserving relevant information, a downsampling method was applied. From the original data set, systematic sampling was performed by selecting every fourth sample, reducing the total number of data records to 13,128.

This approach is a common technique in time series analysis, which reduces the size of the data set while preserving the basic characteristics of the observed process, thereby ensuring the representativeness of the reduced set. In this way, the data set used retains the key trends in the changes in the observed variables and allows for reliable statistical analysis of the relationship between the rotation speed as an independent variable and the generator temperature as a dependent variable. The relationship between the observed variables was examined using a linear regression model, which enabled quantification of the impact of changes in rotation speed on changes in generator temperature. The dependence of temperature on rotation speed was analyzed using a scatter diagram with a linear trend line, while the statistical characterization of temperature distribution was performed using a box-plot diagram for different generator operating modes.

3.1. Generator Temperature in Operating Mode

This subsection analyzes the influence of the rotation speed of the DFIG generator of a wind turbine on the temperature of the generator, as well as the temperature distribution, which illustrates the thermal behavior in the operating mode.

The influence of the rotation speed of the DFIG wind turbine generator on the temperature of the generator is shown in Figure 3.

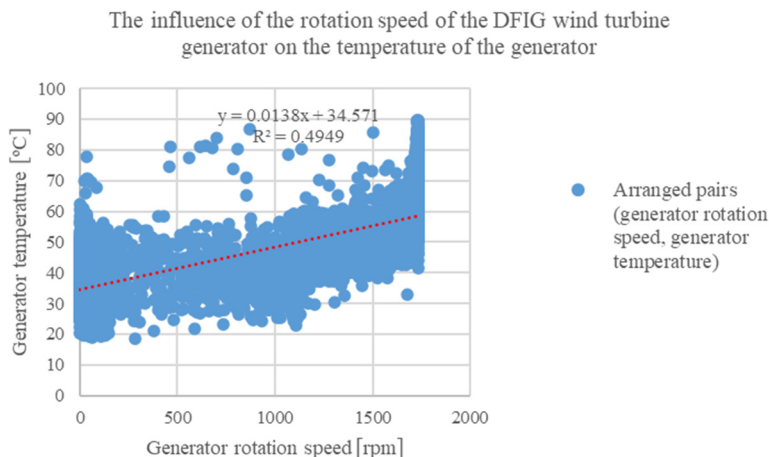


Figure 3: The influence of the rotation speed of the DFIG wind turbine generator on the temperature of the generator

Figure 3 shows a graphical representation of the dependence of the generator temperature on the rotation speed of the DFIG generator in operating mode. The blue dots represent ordered pairs of measured values, i.e. the corresponding values of the rotation speed and the generator temperature.

The dependence is approximated by a linear regression model, which confirms the obtained regression lines. However, the scattering of experimental points around the regression line indicates deviations from ideal linear behavior, which can be attributed to the influence of additional operating and ambient factors (generator load, cooling conditions, ambient temperature, etc.).

It is observed that at lower rotation speeds the temperatures are lower and relatively stable, while with increasing speed there is a slight but continuous increase in temperature. At higher rotation speeds there is more pronounced heating and greater dispersion of measured values, which indicates an increased sensitivity of the thermal regime of the generator in that range.

An approximately linear dependence with moderate deviations is observed, where the rotation speed is a significant factor in determining the thermal state of the generator, but not the only one.

Figure 3 shows a linear regression line that illustrates the trend of generator temperature versus rotation speed. The equation of the regression line (red dashed line) is given by the following expression:

$$y = 0,0138 \cdot x + 34,571 \quad (6)$$

where:

- y - temperature of the wind turbine generator (°C).
- x - rotation speed of the wind turbine generator (rpm).

The slope of the regression line is $b_1 = 0.0138$, which indicates that for each increase in the generator rotation speed by one revolution per minute, the generator temperature increases by approximately 0.0138 °C on average.

The intercept of the regression line is $b_0 = 34.571$ and corresponds to the estimated generator temperature at zero rotation speed (0 rpm). This value serves as a reference point and allows for the estimation of the initial temperature of the generator when not in operation.

The Pearson correlation coefficient is $r = 0.7034$, which indicates a medium-strong positive correlation between the rotation speed of the wind turbine generator and its temperature. With increasing rotation speed, a tendency for the generator temperature to increase is observed. This result confirms the existence of a significant linear dependence between the observed variables.

The value of the coefficient of determination $R^2 = 0.4949$ indicates a moderate level of statistical correlation between the rotation speed and the generator temperature. The results show that the rotation speed explains approximately 49.5% of the temperature variability, while the remaining part of the variation (about 50.5%) depends on other factors, such as cooling quality, generator load, ambient temperature, etc. Therefore, the rotation speed can be seen as a significant, but not dominant parameter in modeling the thermal behavior of the generator under operating conditions.

The generator temperature distributions, which illustrate the thermal behavior in the operating mode ($n > 0$ rpm) depending on the rotation speed, are shown using box-plot diagrams in Figure 4.

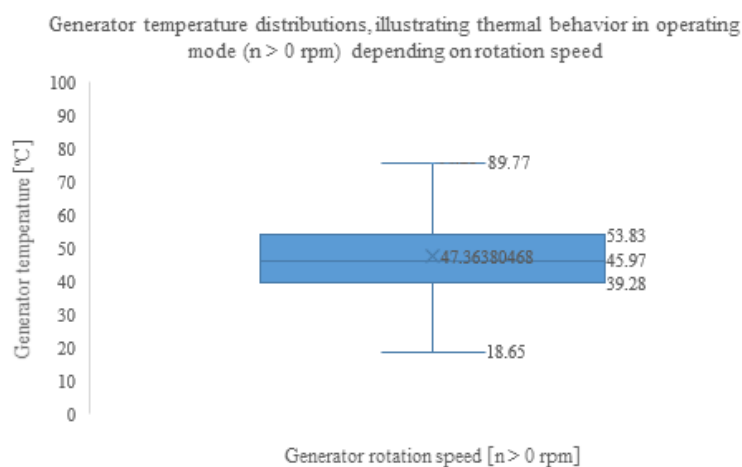


Figure 4: Generator temperature distributions illustrating thermal behavior in operating mode

The first quartile (Q1) is 39.28 °C and represents the value below which 25% of the measured temperatures are located. The median temperature is 45.97 °C and represents the central, i.e. representative, value of the thermal state of the generator in operating mode. The third quartile (Q3) is 53.83 °C and defines the value below which 75% of the measured temperatures are located, and represents the upper limit of the interquartile range. Based on the displayed parameters, it

can be concluded that most of the measured temperatures are in the interval from 39.28 °C to 53.83 °C. The arithmetic mean of the temperature is 47.36 °C and represents the average value of the measured temperatures. The minimum and maximum temperature values are 18.65 °C and 89.77 °C, respectively.

3.2. Generator Idle Temperature

The generator temperature distributions, which illustrate the thermal behavior in the idle mode ($n = 0$ rpm), are shown using a box-plot diagram in Figure 5.

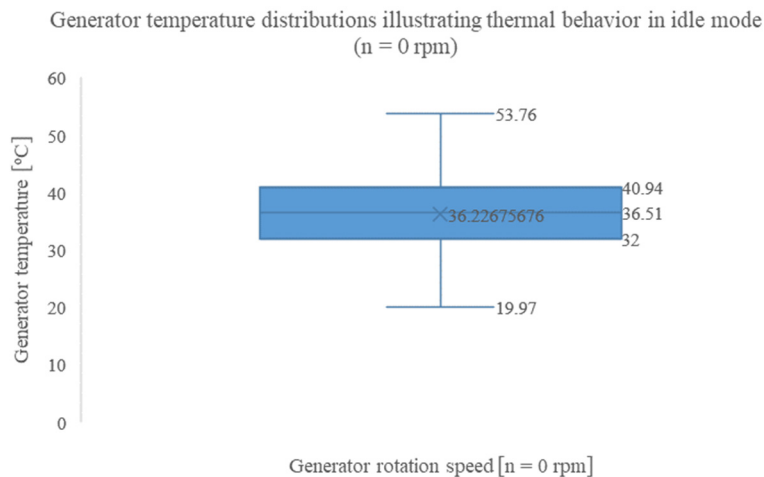


Figure 5: Generator temperature distributions illustrating thermal behavior in idle mode

The first quartile is 32 °C and indicates the value below which 25% of the measured temperatures are located. The median temperature is 36.51 °C and represents the central value of the thermal state of the generator in idle mode. The third quartile is 40.94 °C and indicates the value below which 75% of the measured temperatures are located, representing the upper limit of the interquartile range.

Based on the parameters shown, it can be concluded that the temperatures in the idle mode are relatively stable, with moderate dispersion and no pronounced extreme deviations. The arithmetic mean temperature is 36.22 °C and represents the average value of the measured temperatures, while the minimum and maximum values are 19.97 °C and 53.76 °C, respectively, which defines the overall range of temperature variation.

The analysis of the generator temperature distribution in the operating mode and the idle mode provides insight into the basic statistical parameters and differences between the observed data sets. The minimum generator temperature in the operating mode is 18.65 °C and is lower than the temperature in the idle mode, which is 19.97 °C. This phenomenon can be explained by the influence of ambient conditions, time asynchrony of the measurements and the measurement uncertainty of the sensor, where the difference of 1.32 °C belongs to the range of expected experimental variability.

In the idle mode, the generator temperatures are characterized by lower values and greater stability, with less dispersion and a more uniform distribution of measurement data. In contrast, in the operating mode, the temperature increases with increasing rotation speed, accompanied by more pronounced variability and a wider range of values, which indicates a more complex thermal behavior of the generator due to the action of additional influencing factors.

4. MODEL AND RESIDUAL ANALYSIS

This chapter presents the statistical validation of the developed model, as well as the analysis of residuals with the aim of assessing its adequacy.

4.1. Statistical Validation of the Model

ANOVA detects the presence or absence of a global effect of the independent variable on the dependent variable [15]. The ANOVA (Analysis of Variance) table of the regression model includes degrees of freedom (df), sum of squares (SS), mean squares (MS), F -statistics, and the corresponding p -value of the F -test (Significance F). The listed elements are systematically arranged according to the principal components of the model, including regression, residual (error), and total variance. ANOVA within the regression model was performed using the Data Analysis – Regression tool in [Microsoft Excel], and the results are shown in Table 2.

Table 2: ANOVA regression model

	df	SS	MS	F	Significance F
Regression	1	2429939429	2429939429	12027.17717	0
Residual	12275	2480008907	202037.3855		
Total	12276	4909948336			

In order to analyze the effect of the rotation speed of the wind turbine DFIG generator on its temperature, the linear regression method was applied, while the statistical significance of the model was assessed using the ANOVA approach. ANOVA analysis allows the decomposition of the total variability of the dependent variable into the part explained by the regression model and the part attributed to residual (unexplained) variation. The degrees of freedom (df) in the ANOVA table represent the number of independent pieces of information used to estimate the generator temperature variability, with the total degrees of freedom divided into the component associated with the regression model that includes the wind turbine DFIG generator rotation speed and the residual error component.

The resulting F -value indicates a high ratio between explained and unexplained variation, with the generator temperature variation explained by the DFIG generator rotation speed being significantly higher than the residual variation. The accompanying p -value of the F -test (Significance $F \approx 0$) shows that the effect of the DFIG generator rotation speed on the generator temperature is statistically highly significant, i.e. the probability that the observed relationship between these two variables arose by chance is negligibly small. This confirms that the rotation speed has a real and measurable effect on the generator temperature changes in the observed model. Based on the above results, it can be concluded that the linear regression model is statistically highly significant, i.e. that the rotation speed of the DFIG generator has a statistically significant effect on the generator temperature. The results confirm a statistically significant linear dependence between the observed variables.

The model shows high statistical significance; however, its validity is limited by the assumption of linearity of the relationship between the observed variables. Furthermore, the model does not consider additional physical influences, such as load changes, cooling system configuration and ambient conditions, which have a significant impact on the thermal processes in the generator. Based on the results obtained, ANOVA analysis confirms that the linear regression model adequately describes the statistical relationship between rotation speed and generator temperature, with the influence of rotation speed being statistically significant and dominant in explaining temperature variability. The high value of the t -statistic (109.67) and the negligibly small p -value (≈ 0) confirm that the influence of rotation speed on generator temperature is statistically highly significant, with the probability that the observed effect is the result of chance being practically negligible.

The RMSE and MAE values, which in this case are 8.28 and have identical numerical values, represent the basic indicators of the accuracy and quality of the prediction of the linear regression model that describes the influence of the rotation speed of the DFIG generator on the generator temperature. RMSE represents the root mean square error and is used as a measure that gives greater importance to larger deviations, which is why it is more sensitive to the presence of larger errors in the model prediction. MAE represents the mean absolute error, or the average absolute difference between the actual and predicted temperature values, where the obtained value shows that the model deviates on average by about 8.28 temperature units from the measured values, without taking into account the direction of the deviation. The identity of the RMSE and MAE values indicates that the model errors are relatively uniform, i.e. there are no pronounced extreme deviations that would lead to a significant increase in RMSE relative to MAE. This is an important indicator of model stability, as it indicates that the linear regression model consistently describes the relationship between rotation speed and generator temperature, without the presence of significant local deviations.

In the context of DFIG wind turbine generators, the obtained error values indicate that the model approximates the thermal behavior of the system with relatively high accuracy. The average deviation of about 8.28 temperature units can be viewed as a systematic error of the model, which is probably a consequence of neglected physical factors, such as load changes, cooling regime, wind turbulence and ambient conditions. Nevertheless, the low and stable error value indicates that the linear regression model has good predictive ability and is adequate to describe the general trend of the dependence of temperature on the generator rotation speed.

4.2. Residual Analysis

In order to verify the adequacy of the developed model for estimating the temperature of the DFIG generator, a residual analysis was performed. The residuals represent the difference between the measured and model-predicted generator temperature values, and provide insight into possible model deviations, the presence of systematic errors, and data variability. The residual diagram as a function of the generator rotation speed is shown in Figure 6.

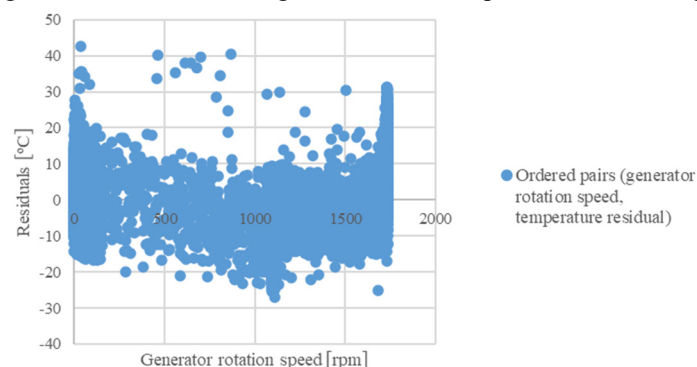


Figure 6: Residual diagram as a function of generator rotation speed

Figure 6 shows a diagram of residuals as a function of generator rotation speed, which allows a visual assessment of the existence of a dependence between these quantities. The blue points on the diagram represent ordered pairs (generator rotation speed, temperature residual), with each point corresponding to one measurement sample. Based on the diagram of residuals, it can be seen that the residuals are approximately randomly distributed around the zero value, without a clearly expressed systematic pattern, which indicates that the assumption of linearity of the model is satisfied and there is no significant systematic deviation.

However, there is a variation in the dispersion of the residuals along the generator rotation speed range, which suggests a deviation from the assumption of homoscedasticity. Also, the observed range of residuals and the presence of extreme values indicate limitations of the model in terms of the accuracy of temperature prediction in certain generator operating modes.

The observed changes in the variability of the residuals suggest that additional factors, such as load, cooling conditions and operating modes, can have a significant impact on the generator temperature, and that their inclusion in the model could contribute to improving its accuracy. Based on the overall analysis, the model shows a satisfactory level of adequacy, however, the observed deviations indicate the need for its further improvement by applying more complex model approaches, in order to increase the reliability and accuracy of the prediction.

Based on the analysis of the residuals, it can be concluded that the assumption of linearity of the model is satisfied, while the observed indications of impaired homoscedasticity indicate certain limitations of the observed model.

5. CONCLUSION

In this paper, a quantitative analysis of the dependence of the temperature of a wind turbine DFIG generator on its rotation speed was conducted, using linear regression analysis and statistical processing of SCADA data. The results of the correlation analysis confirm the existence of a medium-strong positive linear dependence between the observed quantities ($r = 0.7034$), while the coefficient of determination ($R^2 = 0.4949$) shows that the rotation speed explains almost half of the total variability of the generator temperature.

The results of the statistical validation of the model based on ANOVA analysis confirm its high statistical significance, with the F -test and t -statistics indicating a significant influence of the rotation speed on the generator temperature. The results of the error analysis (RMSE and MAE = 8.28) indicate a uniform and stable pattern of model deviations, without the presence of pronounced extreme values, which further confirms the high reliability and consistency of its predictions.

The findings of the residual analysis indicate their random distribution around the zero value, which confirms the adequacy of the linearity assumption of the model. However, the identified variability of the residual variance indicates a deviation from the homoscedasticity assumption, which indicates the presence of additional factors that affect the thermal behavior of the generator, such as load, cooling conditions and ambient influences.

Although the developed linear regression model confirms a statistically significant relationship between the rotation speed and the temperature of the DFIG generator, it is necessary to point out certain limitations of the applied approach. The developed model is based on a single independent variable, while other relevant factors, such as generator load, cooling regime and ambient conditions, are not included in the analysis. In addition, the temperature analysis is based on measurements at a single representative point, which limits the possibility of observing the spatial distribution of the thermal field inside the generator.

Although the linearity assumption is statistically confirmed on the observed data set, it may lead to a decrease in the accuracy of the model in conditions of nonlinear system behavior at extreme operating modes. The observed heteroskedasticity indicates that the application of more robust models could improve the reliability of the analysis.

Based on the obtained results, it can be concluded that the linear regression model adequately describes the basic dependence between the rotation speed and the temperature of the DFIG generator under operating conditions, and represents a reliable and efficient approach for assessing the thermal behavior of the system. However, the observed limitations in explaining the overall variability, as well as the presence of heteroscedasticity, indicate the need to improve the model by including additional input variables or applying more advanced nonlinear and multivariate approaches, in order to improve the accuracy and reliability of the generator temperature prediction.

In future research, it is recommended to expand the set of input variables, as well as the application of advanced multivariate regression models and machine learning methods, including neural networks and ensemble methods, in order to improve the accuracy, robustness and generalization ability of the model for generator temperature prediction under different operating conditions. Also, the application of multipoint measurements would enable a more detailed analysis of the spatial distribution of temperature within the DFIG system, which would improve the understanding of local thermal gradients and their impact on the generator operating characteristics and reliability.

REFERENCES

- [1] C. P. Chioncel, E. Spunei and G.O. Tirian, “Wind Turbines Around CutIn Speed: Startup Optimization and Behavior Analysis Reported to MPP,” *Appl. Sci.*, Vol. 15, p. 3026, --<https://doi.org/10.3390/app15063026>, (2025).
- [2] A. Dannier, E. Fedele, I. Spina and G. Brando, “DoublyFed Induction Generator (DFIG) in Connected or Weak Grids for TurbineBased Wind Energy Conversion System,” *Energies*, Vol. 15, p. 6402, --<https://doi.org/10.3390/en15176402>, (2022).
- [3] M. Arsal, T. Kamel, H. Asad and A. Khan, “A Systematic Review of Cyber Risk Analysis Approaches for Wind Power Plants,” *Energies*, Vol. 19, p. 677, <https://doi.org/10.3390/en19030677>, (2026)
- [4] M. Shuaibu, A. S. Abubakar and A. F. Shehu, “Techniques for ensuring fault ridethrough capability of grid connected DFIGbased wind turbine systems: A review,” *Niger. J. Technol. Dev.*, Vol. 18, pp. 39–46, --<https://doi.org/10.63746/njtd.v18i1.591>, (2021).
- [5] R. Sharma, V. S. Chauhan and H. Shah, “Doubly Fed Induction Generator for Wind Energy System,” *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, Vol. 6(5), pp. 3208–3213, <https://doi.org/10.15662/IJAREEIE>. 2017.0605004, (2017).
- [6] P. Prajzendanc and C. Kreischer, “A Review of New Technologies in the Design and Application of Wind Turbine Generators,” *Energies*, Vol. 18, p. 4082, <https://doi.org/10.3390/en18154082>, (2025).
- [7] K. Ni, Y. Hu, Y. Liu and C. Gan, “Performance Analysis of a Four Switch Three Phase Grid Side Converter with Modulation Simplification in a Doubly Fed Induction Generator Based Wind Turbine (DFIG WT) with Different External Disturbances,” *Energies*, Vol. 10, p. 706, <https://doi.org/10.3390/en10050706>, (2017)
- [8] H. A. Behabtu, T. Coosemans, M. Berecibar, K. A. Fante, A. A. Kebede, J. Van Mierlo and M. Messagie, “Performance Evaluation of Grid Connected Wind Turbine Generators,” *Energies*, Vol. 14, p. 6807, <https://doi.org/10.3390/en14206807>, (2021).
- [9] R. K. Behara and A. K. Saha, “Optimised Neural Network Model for Wind Turbine DFIG Converter Fault Diagnosis,” *Energies*, Vol. 18, p. 3409, <https://doi.org/10.3390/en18133409>, (2025).
- [10] Y. Qiu, W. Zhang, M. Cao, Y. Feng, and D. Infield, “An Electro Thermal Analysis of a Variable Speed Doubly Fed Induction Generator in a Wind Turbine,” *Energies*, Vol. 8(5), pp. 3386–3402, <https://doi.org/10.3390/en8053386>, (2015).
- [11] M. Biljan-August, S. Pivac and A. Štambuk, *Uporaba statistike u ekonomiji*, 2. izd., Rijeka: Ekonomski fakultet Sveučilišta u Rijeci, 2009.
- [12] A. Tabak, *Pearsonov korelacijski koeficijent*, Diplomski rad, Sveučilište J. J. Strossmayera u Osijeku, Odjel za matematiku, Diplomski studij matematike – smjer Financijska matematika i statistika, Osijek, Hrvatska, 2018.
- [13] H. J. Seltman, *Experimental Design and Analysis*, 2018. Available: <http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf>, accessed: 27.04.2026.
- [14] A. Nielsen, “Practical Time Series Analysis: Prediction with Statistics and Machine Learning” O’Reilly Media, Inc., Sebastopol, CA, USA, 2020. <https://www.oreilly.com/>
- [15] J. Juarros-Basterretxea, G. Aonso-Diego, Á. Postigo, P. Montes-Álvarez, Á. Menéndez-Aller i E. García-Cueto, “Post-Hoc Tests in One-Way ANOVA: The Case for Normal Distribution,” *Methodology*, Vol. 20(2), pp. 84–99, <https://doi.org/10.5964/meth.11721>, (2024).