energies

Article

Condition Assessment and Analysis of Bearing of Doubly Fed Wind Turbines Using Machine Learning Technique

Aiman Abbas Mahar [1], Nayyar Hussain Mirjat [2,*], Bhawani S. Chowdhry [3], Laveet Kumar [4], Quynh T. Tran [5] and Gaetano Zizzo [6,*]

[1] Department of Electronic Engineering, Mehran University of Engineering & Technology, Jamshoro 76062, Pakistan
[2] Department of Electrical Engineering, Mehran University of Engineering & Technology, Jamshoro 76062, Pakistan
[3] NCRA-HHRCMS Lab, Mehran University of Engineering and Technology, Jamshoro 76062, Pakistan
[4] Department of Mechanical Engineering, Mehran University of Engineering & Technology, Jamshoro 76062, Pakistan
[5] Institute of Energy Science, Vietnam Academy of Science and Technology, Hanoi 10000-04, Vietnam
[6] Department of Engineering, University of Palermo, 90128 Palermo, Italy
* Correspondence: nayyar.hussain@faculty.muet.edu.pk (N.H.M.); gaetano.zizzo@unipa.it (G.Z.)

Abstract: Condition monitoring of wind turbines is progressively increasing to maintain the continuity of clean energy supply to power grids. This issue is of great importance since it prevents wind turbines from failing and overheating, as most wind turbines with doubly fed induction generators (DFIG) are overheated due to faults in generator bearings. Bearing fault detection has become a main topic targeting the optimum operation, unscheduled downtime, and maintenance cost of turbine generators. Wind turbines are equipped with condition monitoring devices. However, effective and reliable fault detection still faces significant difficulties. As the majority of health monitoring techniques are primarily focused on a single operating condition, they are unable to effectively determine the health condition of turbines, which results in unwanted downtimes. New and reliable strategies for data analysis were incorporated into this research, given the large amount and variety of data. The development of a new model of the temperature of the DFIG bearing versus wind speed to identify false alarms is the key innovation of this work. This research aims to analyze the parameters for condition monitoring of DFIG bearings using SCADA data for k-means clustering training. The variables of k are obtained by the elbow method that revealed three classes of k (k = 0, 1, and 2). Box plot visualization is used to quantify data points. The average rotation speed and average temperature measurement of the DFIG bearings are found to be primary indicators to characterize normal or irregular operating conditions. In order to evaluate the performance of the clustering model, an analysis of the assessment indices is also executed. The ultimate goal of the study is to be able to use SCADA-recorded data to provide advance warning of failures or performance issues.

Keywords: bearings; condition monitoring; DFIG; K means; SCADA data; wind turbine

check for updates

Citation: Mahar, A.A.; Mirjat, N.H.; Chowdhry, B.S.; Kumar, L.; Tran, Q.T.; Zizzo, G. Condition Assessment and Analysis of Bearing of Doubly Fed Wind Turbines Using Machine Learning Technique. *Energies* **2023**, *16*, 2367. https://doi.org/10.3390/en16052367

Academic Editors: Davide Astolfi and Abu-Siada Ahmed

Received: 2 January 2023
Revised: 6 February 2023
Accepted: 27 February 2023
Published: 1 March 2023

1. Introduction

In recent years, renewable energy sources have attracted considerable interest on a global scale, becoming a viable option due to technological development, cost reduction, and increasing demand, especially in developing countries [1,2]. Wind energy conversion is one of the most promising renewable energy technologies that has developed rapidly in recent years and provides a substantial share of electricity in an increasing number of countries.

The cumulative global installed capacity of wind energy is about 93.6 GW of the additional wind generating capacity installed globally in 2021, as shown in Figure 1.

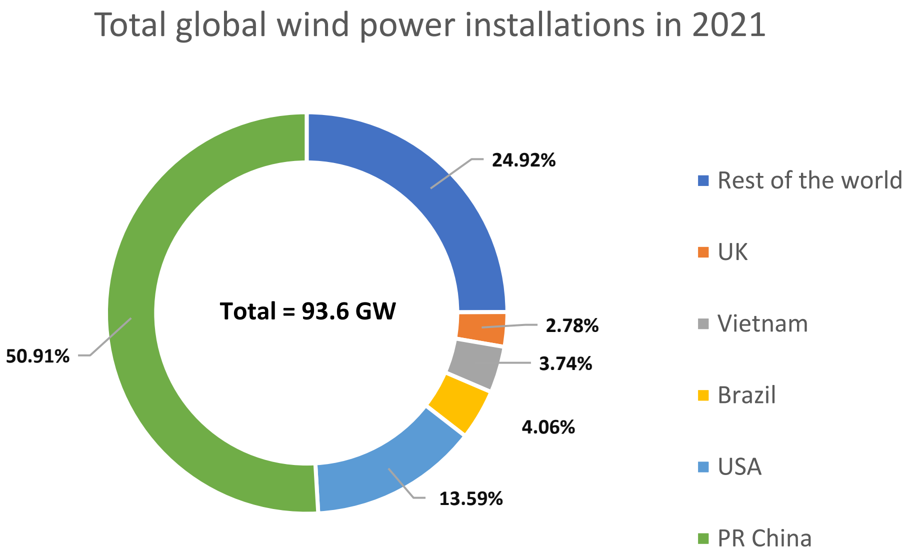


Figure 1. Global wind energy installation.

China, the U.S., Brazil, Vietnam, and the United Kingdom were the top five global markets for wind power installations in 2021 [3].

According to the report in [4], during 2021–2022, twelve wind power projects in Pakistan, with a cumulative installed capacity of 610 MW, achieved and started the supply of electricity to the national grid. As evidenced in Figure 2, currently, Pakistan has around 4% share of wind energy among other resources [5] and has a deployment of renewable energy throughout the land, particularly in two provinces, Sindh and Baluchistan [6], and aims to achieve 30% of its electricity generation from renewables by 2030 [7].

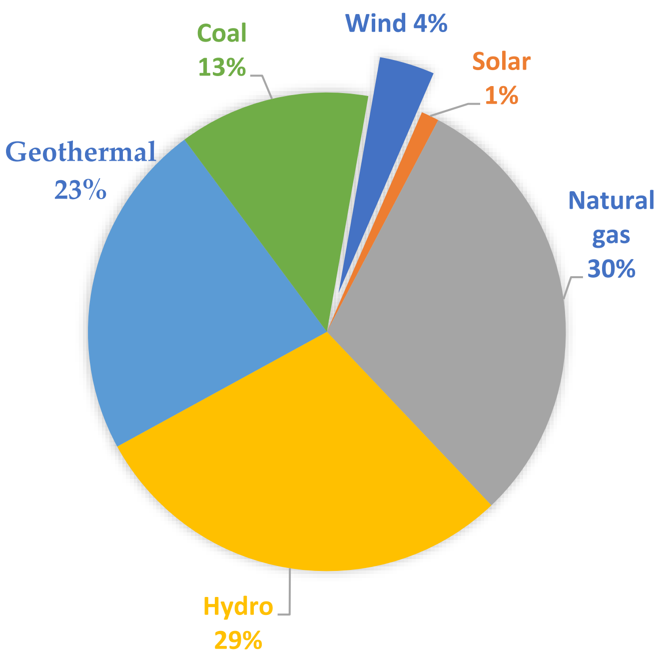


Figure 2. Pakistan's energy matrix.

The generator of a wind turbine is one of the most failure-prone assemblies due to the variable loads [8]. Continuous operations in all environmental conditions contribute to failures of wind turbine components, assemblies, and systems. Bearing failures account for more than 40% of the overall wind turbine generator failures leading to unexpected energy losses [9]. In Figure 3, a fire in a wind turbine is represented. According to [10], the turbine may have failed because of the generator's high bearing temperature due to a shutdown that led the generator to set on fire. The high temperature, combined with the leakage of oil from the rotating manifold nearby, probably triggered the fire.

Figure 3. Extend of damage due to wind turbine failure during October 2022 at Sapphire Wind Power Plant at Jhimpir.

Hence, a solution for effective condition monitoring of generator bearings and early identification of failure symptoms is needed. From actual maintenance practices of wind turbine generators, it was found that there is a high nonlinear relationship between the turbine fault and relevant factors [11]. Therefore, without proper monitoring, replacement of the damaged bearing will not solve the problem and will cause damage again. In order to avert such failures, it is critical to address the main causes of these failures in order to prevent them and create possible solutions to decrease the occurrence of damage in components. The development of condition monitoring techniques for rotating machinery, particularly the bearings, has received much attention during the past few decades. Since bearing failures results in prolonged downtime and wind turbine systems operate in adverse conditions with widely variable speeds, loads, and temperatures, they appear to be a solid option for model-based condition monitoring system [12,13]. However, cutting tools are the final executive component in the machining process and come into close contact with the product. This causes them to wear out quickly, which in turn impacts the workpiece's surface quality. According to [14], tool wear and damage are the primary factors causing the failure of the machining process. The resulting downtime accounts for 7–20% of the total downtime of the machining process, and the cost of tool and tool change accounts for 3–12% of the total machining cost.

A lack of information precisely describing primary bearing breakdowns in terms of frequency and damage types, as well as the most common bearing modeling and analysis setups, divided into dynamic and quasi-static categories, are presented in [15]. The study examines wind turbines' dynamic reliability under various control strategies and external conditions. The survival signature and fault tree analysis (FTA) were used to examine the system reliability level of wind turbine drive trains under various wind conditions in Ref. [16]. To maintain maximum power output, a doubly fed induction generator (DFIG) was chosen since it includes more than 50% of all major onshore wind power plants worldwide [17]. Doubly fed wind turbines are vulnerable to various types of generator losses; these failures lead to excess vibrations that might damage other components, such as bearing failures, which produce non-stationary vibrations [18,19]. Because the consequences of a failure are catastrophic for both business and customers, it is crucial to prevent the error more precisely [18]. Due to the longtime operation of the wind turbine in poor conditions, such as heavy loads, corrosion or failure in gearboxes can occur, such as misalignment, looseness, and contamination of the bearings [20]. Contaminants include dust in the air, dirt in the bearing, and any abrasive substance that gets into the bearing. Studies conducted on the reliability of wind turbines show that the drivetrain system accounts for about 20% of overall problems and accounts for nearly 30% of wind turbine outages when using DFIG [21].

With the use of data from the supervisory control and data acquisition system and the idea of energy saving, this study enhances a condition monitoring method for wind turbine main bearings. The objective of the current research was to employ model parameters as health indicators. The method was applied to find main bearing degradations over a two-year period in a wind farm with more than 100 WTGs. The method was evaluated because of the history of bearing failures that were known [22]. The two different types of datasets were used to test the newly created trigonometric entropy measure based on variational mode decomposition (VMD). One comes from the Centre for Intelligent Maintenance Systems, and the other from the XJTU-SY Bearing Databases. The suggested method has the ability to raise the alert about the beginnings of faults relatively at an early stage [23].

One of the research studies suggests an artificial neural network (ANN)-based defect detection approach for wind turbine main bearings based on current SCADA data. In this study, SCADA data from the Nord-Trndelag Elektrisitetsverk-owned Hundhammerfjellet wind farm were used. Turbine rear bearing temperature, a main shaft rear bearing characteristic in the SCDA data, provides an indication of how hot the bearings are operating and hence provides the opportunity to identify rear bearing overheating [24]. Another study was developed in such a way that 10 min average operating (SCADA) data for a group of 14 wind turbines are available, and a subset of 10 of those 14 turbines undergo monthly grease checks. The suggested method is completely hybrid and intended to combine data-driven and physics-informed layers in deep neural networks. The bearing damage of a wind turbine using recurrent neural networks was studied. It was specifically suggested that grease damage increments through a multi-layer perceptron and bearing fatigue damage through equations frequently employed in bearing reliability design [25]. Therefore, this research focuses on analyzing these faulty bearings in DFIG. The conditioning monitoring-based system was suggested to detect faulty bearings, and a machine learning approach was used for the detection of intensity/type of fault in DFIG. In this paper, system identification methods were applied to SCADA data to develop a condition monitoring model, which can be used to predict the generator's indices that affect the bearings in terms of wind speed, generator temperature, generator rotation speed, etc. The SCADA data were analyzed to assess wind turbine operating conditions using the developed model that effectively identifies potential failures or breakdowns.

The need for condition monitoring of wind turbines is growing as the size and location of modern wind turbines make their technical availability essential. Due to the limited accessibility of some remote-controlled wind farms on mountain and offshore wind farms, unexpected failures, especially of large and important components, might result in unnecessary delay and cost. Therefore, the goal of our research was to obtain real-time SCADA data for acquisition while minimizing downtime through condition monitoring. Many methods employing these data for early failure detection have recently been developed since CM using SCADA data is a potentially low-cost solution requiring no additional sensors. From all the parameters, it was found that the prediction of the temperature variation trend is crucial in order to provide an overheating warning and detect an optimization problem. In order to solve the optimization problem of the proposed model, the corresponding objective function was derived in a more tractable form, and an alternative update algorithm is presented, which is based on the identification of new concepts in unlabeled data. The method is used to achieve improved predictive performance in terms of improved predictive accuracy.

The rest of the paper is structured as follows: Section 2 presents the research methodology describing the data-based approach and reporting the research flow of the study. Section 3 provides the results and analysis of the study. Techniques and methods (K-means clustering approach, elbow method, boxplot visualization) are discussed and applied to the SCADA data where faults are known to have occurred, and conclusions and future recommendations are drawn.

2. Methodology

Wind turbine condition monitoring can be used to improve safety or to lower the cost of the existing level of safety, anticipating or detecting emerging major faults [26]. The study proposes a methodology based on wind turbine condition monitoring that, through a supervisory, control, and data acquisition SCADA system, records data from an inertial measurement unit (IMU) sensor mounted on the DFIG's bearings that detects system signals. The SCADA system provides historical signals, fault information, environmental condition parameters, and operational factors related to the DFIGs and their equipment in wind turbines. The collected data are then used by proposing a new method of bearing failure detection in the machines.

The method used in this study to compare the bearing status of the DFIG in the SCADA system is characterized by its feasibility and cost-effectiveness, and its flowchart is shown in Figure 4. Table 1 lists the key parameters of interest.

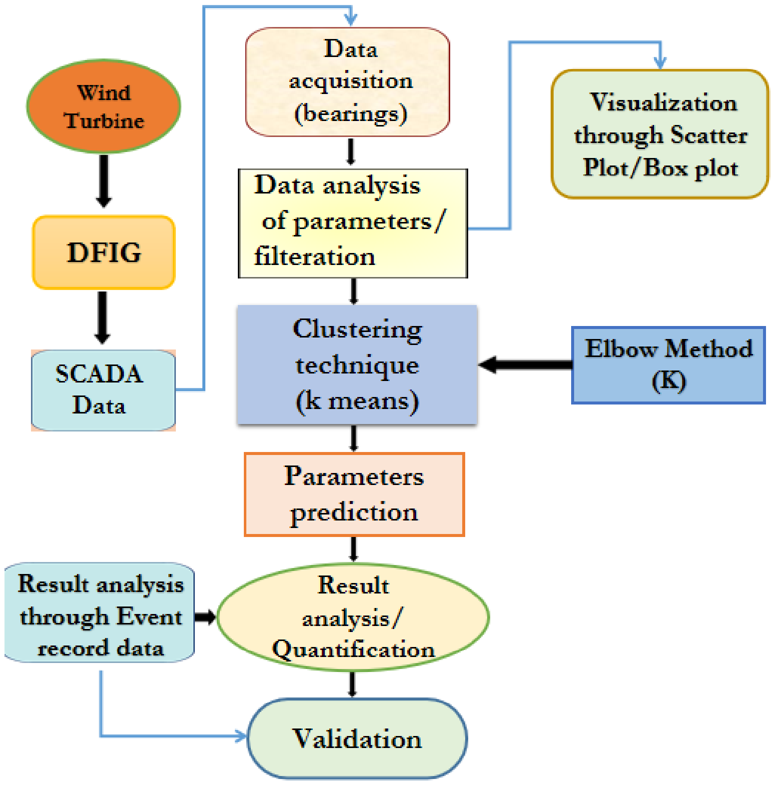


Figure 4. Research Methodology Flow chart.

Unlike supervised learning, where human-labeled data are necessary, unsupervised learning uses unlabeled data and provides an insight into the probable classes present in the dataset, which is then used to broadly classify the data. The main benefit of unsupervised learning lies in the fact that it does not require any human-labeled data, which saves much manual effort that is otherwise used to label every data point in a particular dataset individually.

Table 1. Assessment indices of DFIG bearings.

S.no.	SCADA Parameters	Unit
1	Average Active Power	(kW)
2	Average Wind Speed	(m/s)
3	Average Reactive Power	(kVar)
4	Wind Turbine Energy yield	(kWh)
5	Average Ambient Temp	(°C)
6	Average Gen Rotation Speed	(rpm)
7	Average Maximum generator temp	(°C)

Advanced condition monitoring systems are used by wind turbines, which are complex systems, to assess their state of health. Due to its high failure rate and downtime, the generator is one of the most important components. In wind farms, SCADA are used for real-time condition monitoring and control. New and reliable strategies for data analysis are needed because of the large amount and variety of the data. The development of a new model of the temperature of the DFIG bearing versus wind speed to identify false alarms was the key innovation of this work. A box plot was utilized to depict the distribution of the data, and a data partitioning strategy was used.

The model for determining failures in DFIG bearings is based on machine learning (ML) techniques for classification and parameter analysis for early diagnosis and predictive maintenance. The K-means clustering technique is used for the evaluation and condition analysis of DFIG bearings, and the elbow method is used to determine the K value for the classes in accordance with the clustering technique. Reducing the size and complexity of the dataset is the primary goal of the clustering technique. Compared to the original data, clustered groups of points occupy much less storage space and are easier to control, which is why this method is proposed. Once the clusters were determined, scatterplot and boxplot visualization was used for the statistical visualization of the predicted data. In the last step, validation of the predictive performance of the results was analyzed with the latest data obtained.

K-means Clustering

The unsupervised learning approach known as K-means clustering divides the unlabeled dataset into many clusters. The K-means algorithm was employed for the condition assessment of the DFIG bearings. The K-means clustering technique was applied to design the prediction model for the assessment indices that are based on DFIG bearings, such as temperature, rotation speed, and wind speed. The K-means model was used to identify the clusters in the available dataset. Here, the value of K (1, 2, 3, 4, 5) was determined.

The model is trained through the machine learning algorithm using python programming in Jupyter in Figure 5.

The steps in the K-means clustering algorithm are the following:

Let X = {x_1, x_2, x_3, . . . , x_n} be the set of data points and V = {v_1, v_2, . . . , v_n} be the set of centers:

(1) Select "c" cluster centers;
(2) Determine the Euclidean distance between the cluster centers and each data point;
(3) Assign the cluster with the shortest distance from all the other cluster centers;
(4) Recalculate the new cluster center using the following:

$$v_i = (1/c_i) \text{ sum } (j = 1, c_i)\ (x_i)$$

where "c_i" denotes the no. of data points in the ith cluster;

(5) Measure again the separation between each data point and the newly discovered cluster centers;
(6) Stop if no data point was moved; otherwise, go back to step (3).

For the progression, keep changing the value until the optimal clusters can be achieved.

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{ci} \left(\|x_i - v_j\| \right)^2$$

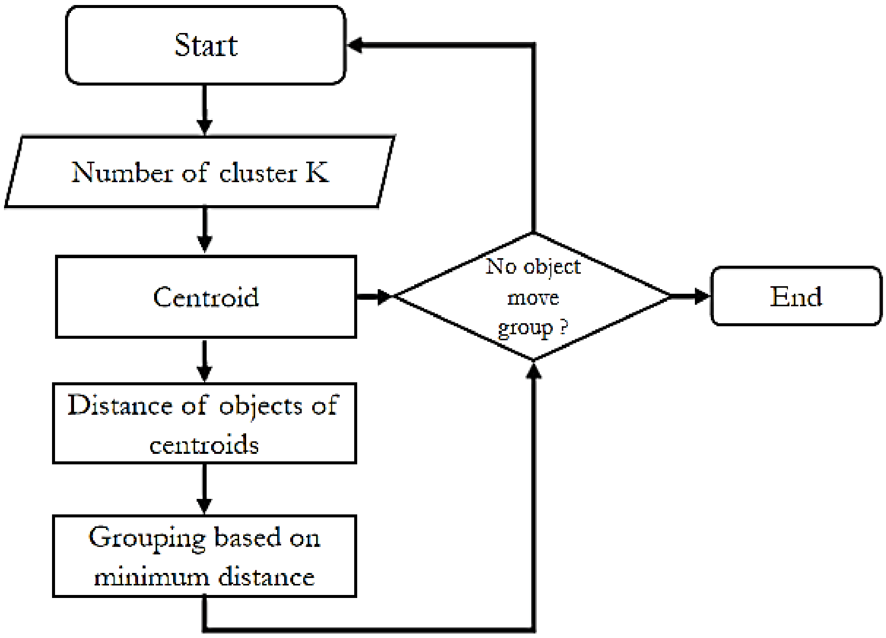


Figure 5. Flowchart of K-means technique.

The Euclidean distance between x_i and v_j is given as "$\|x_i - v_j\|$".

Prediction errors in the SCADA data of the model follow a normal distribution. Thus, it can be calculated based on the probability density function of normal distribution.

Elbow method for optimizing K value:

The elbow method is an effective method for cluster optimization and is used for clustering analysis since it is simple to implement and yields useful results. The elbow technique is a visual way to verify the consistency of the optimal number of clusters by analyzing the difference in the sum of square errors (SSE) of each cluster. The most severe difference creating the angle of the elbow indicates the optimum cluster number [27].

In this method, the optimal value of k using the elbow method is found by calculating the sum of square error (SSE) to evaluate K-means clustering using the elbow criterion. The idea of the elbow criteria approach is performed to select the k (number of clusters) at which the SSE decreases significantly. The elbow rule's fundamental concept is to employ a square of the distance between each cluster's centroid and sample points to generate a range of K values. As a performance measure, the sum of squared errors (SSE) is employed. SSE is calculated by iterating over the K-value. Smaller numbers represent more converging clusters. SSE displays a sharp reduction when the number of clusters is adjusted to be close to the number of actual clusters. When there are more clusters than there are actual clusters, SSE still decreases, but it does so more slowly. It is used to choose the best clusters.

Data Visualization

Large amounts of numerical data can be displayed using graphs, which can be used to demonstrate the relationships between the numerical values of various variables and to derive quantitative relationships between them. One of the most powerful and popular methods for visual data analysis is the scatter plot [28]. Different data points are positioned between an x- and y-axis to display these data. Each of these data points appears to be "scattered" across the graph. By using the generalized scatter plot technique, big datasets can be completely represented in the figure without any overlap. The primary concept is to provide the analyst with the flexibility to adjust the quantity of overlap and distortion to produce the optimal view. The model offers the option to smoothly zoom between the conventional and generalized scatter plots to enable effective usage. An optimization function considers overlap and visualization distortion.

Performance Analysis System

SCADA was utilized to create the system for analyzing model performance. Once a csv file is loaded with the dataset comprising the measurements, the model builds its predictions. Then, predictions are made using the seven parameters listed in Table 1. In the same sequence as the data received from the input file, the predictions are made and written to a csv file. By comparing the predicted label with the actual label, the generated csv files can be examined to determine how well the model performed. The software is used to analyze and visualize the performance of the model implemented in Jupyter Notebook using Python programming.

3. Results and Analysis

3.1. Case Study

In this study, SCADA data were collected and analyzed from an onshore wind farm with 66 wind turbines with DFIG generators with rated power of 1.5 MW each, located in the Jhimpir wind corridor. The data were collected in the form of excel sheets at 10 min intervals at a bearing speed of 1800 rpm. The database covers one month of operation, from January 2021 to February 2021. Up to 300 datasets were recorded for analysis and used to train the model.

3.2. Data Visualization

After data pre-processing, visualization of the data according to its correlation with all filtered parameters can be observed in Figure 5. Once the visualization of the data is acquired, the condition changes in the DFIG parameters summarized in Table 1 can be easily determined.

In Figure 6, a good correlation is observed between the average generator rotation speed and the average generator maximum temperature. Therefore, with the change in speed, the temperature also increases.

In order to analyze the trend and cycles, studies of the parameters with regard to time were devised, as shown in Figure 7. Based on the number of notices and downtime hours, Figure 6 displays the monthly report of the evaluation indices for a certain DFIG. One-month duration is displayed on the x-axis, and the left y-axis displays the ranges and downtime hours. The two temperature parameters exhibit nonlinear and progressive variation tendencies, and the wind speed varies significantly.

It is evident from the figure that as the wind speed increases, the generator's rotational speed and power generation also increase accordingly. The average temperature of the turbine also increases due to an increase in turbine speed. It can also be observed that the effect of ambient temperature on generator temperature is rather low compared to the effect of rotation speed. The turbine starts generation at 5 m/s and tries to maintain its speed at 17 m/s for maximum power generation. As wind speed reduces to less than 5 m/s, the turbine cannot maintain generation and reduces the generator speed to zero, as can be seen from the figure.

3.3. K-Means Clustering with Elbow Method

The K-means algorithm partitions the collected data into k clusters, in which each point belongs to the cluster with the smallest distance. K determines the default clusters to be generated during the process, with K = 2 creating two clusters and K = 3 creating three clusters. This allows for the analysis of a close connection between generator rotation speed and generator temperature. The technique is applied by using the elbow method algorithm, which provides a quick and intuitive response.

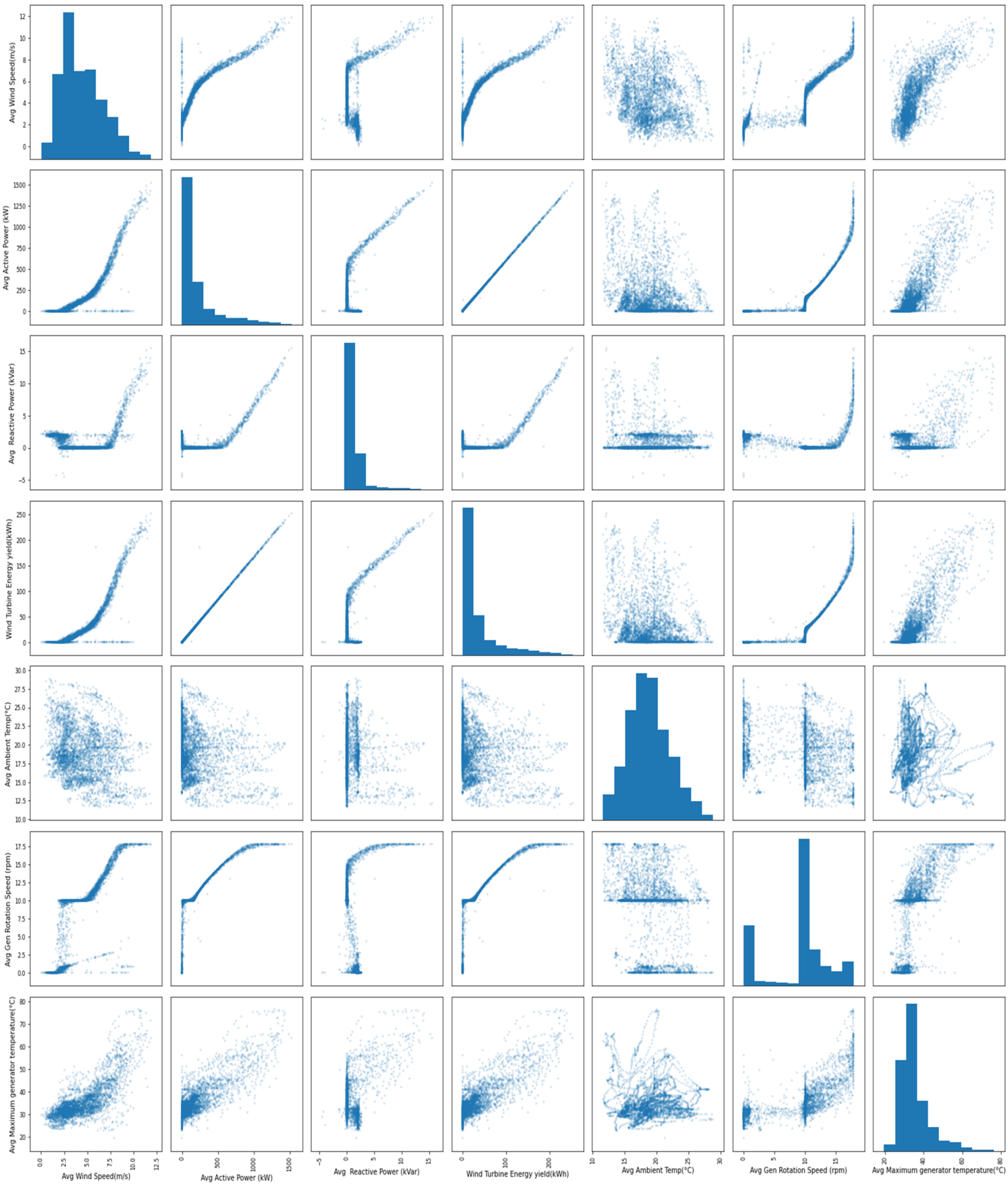


Figure 6. Scatterplot of units divided into clusters obtained according to its correlated parameters.

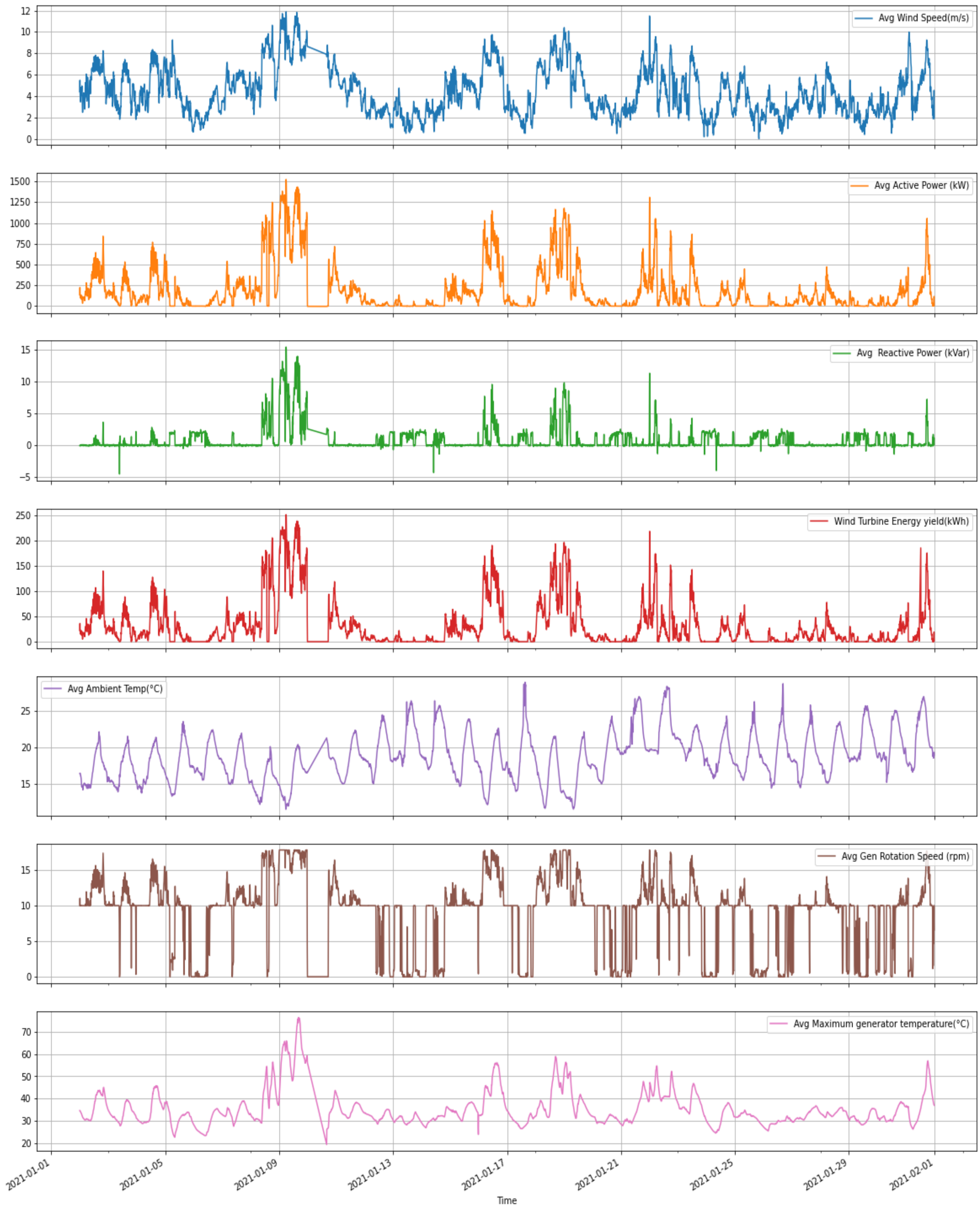


Figure 7. Time series graph of parameters for a given timeline.

In the elbow graph in Figure 8, it is seen that there is a sharp decrease in the SSE until the third cluster. Therefore, the optimal value of k = 3 is obtained.

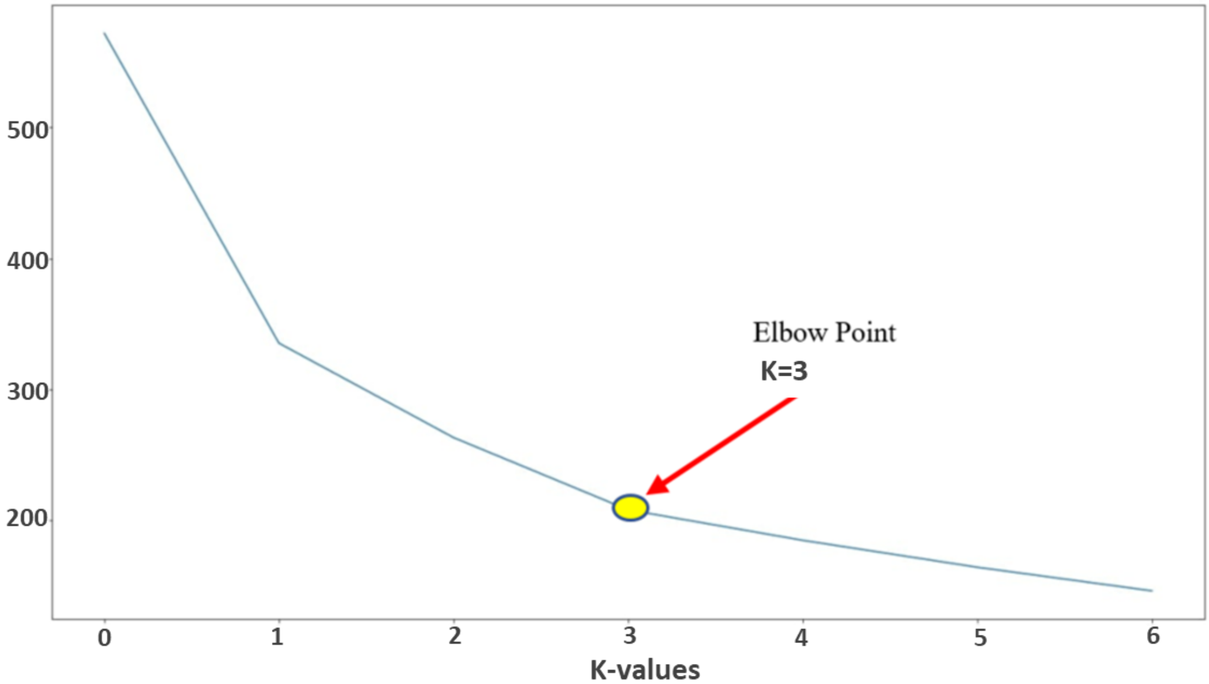


Figure 8. Determination of the best cluster using the elbow method: sum of squared errors (SSE) versus K values.

The centroid of a cluster and each individual observation allocated to that cluster are separated by the total within-cluster variation. The following centroids are chosen using the original max–min criterion, which involves picking the point that is farthest from its nearest centroid. When the cluster sizes are seriously out of balance, the max–min approach is extremely helpful in preventing the worst-case behavior of the random centroids [29].

The clusters are more clearly defined and confined the closer together these distances are by adjusting the k-value to 3 and observing the clusters again with respect to all parameters. The grouping of data points does not give a good correlation between all parameters except the average generator temperature and the average generator rotation speed, which are clearly grouped in Figure 9. However, a linear relationship (orange line) can be observed between the temperature and the rotation speed of the DFIG bearings.

The red dots showing the mean of each cluster's points, orange line shows the linearity and clusters are referred to as classes K = 0, 1, and 2, (3 colors = 3 clusters or k = 3) in Figure 10.

3.4. Box Plot Visualization

The relationship function between the assessment indices and their influencing factors was established using a boxplot. The condition assessment index of DFIG bearings was generated using a boxplot representation of k-means clustering algorithms. The boxplot visualization of the k-means cluster analysis groups individuals as average maximum generator temperature (°C) and average gen rotation speed (rpm).

The partition of units in clusters is:

Cluster 1: 0, for low temperature, low speed;
Cluster 2: 1, for medium temperature, moderate speed;
Cluster 3: 2, for severe temperature, high speed;

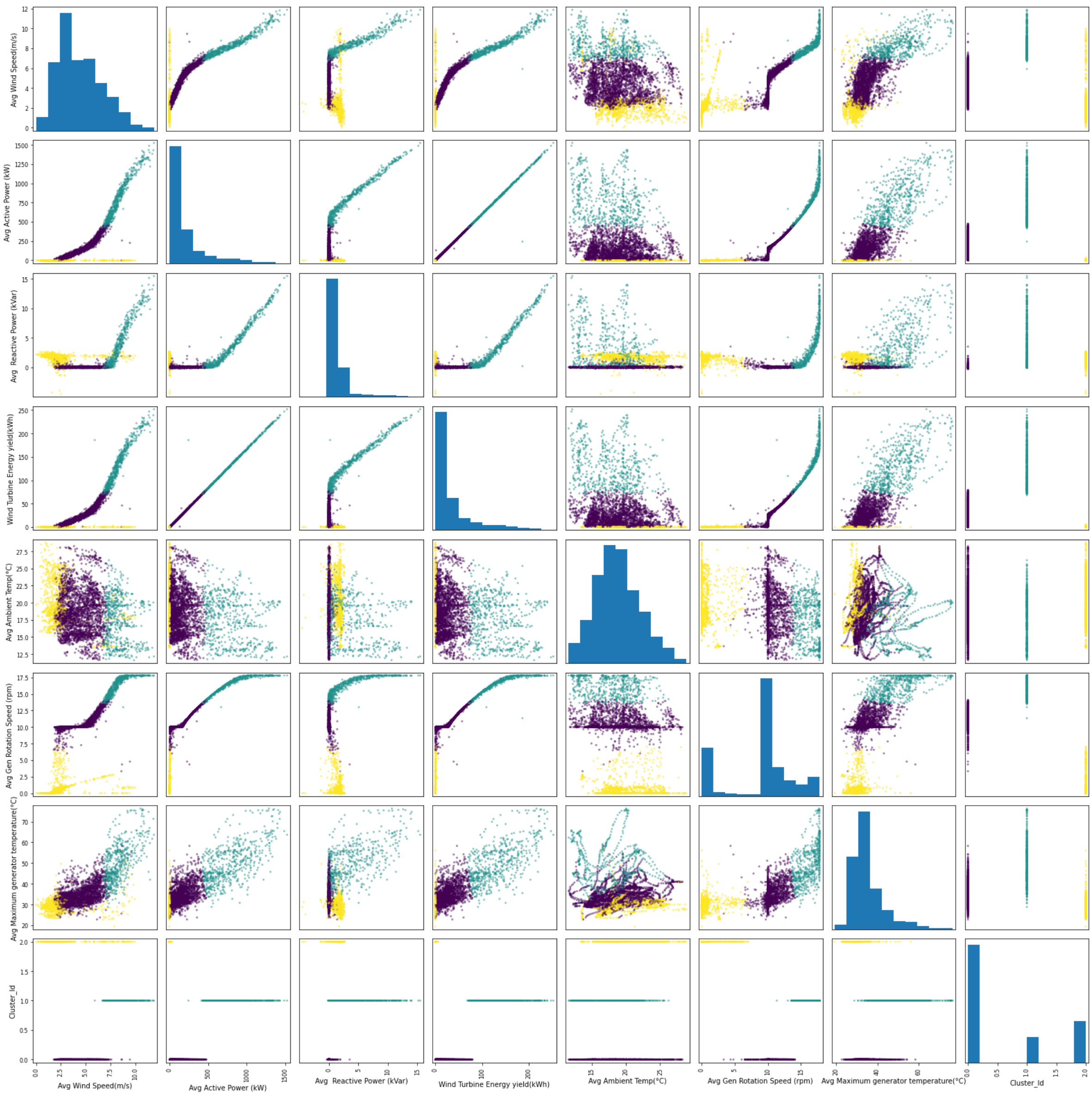


Figure 9. Data visualization according to its correlated parameters.

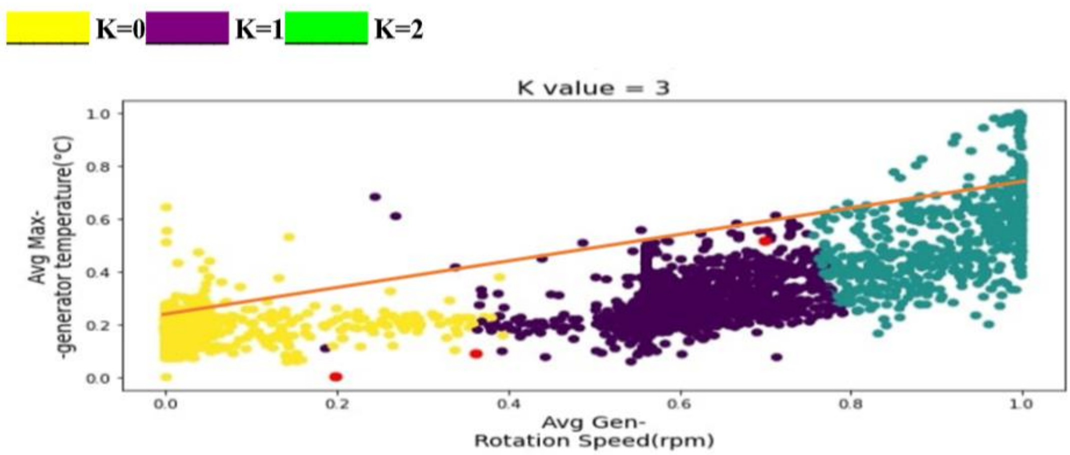
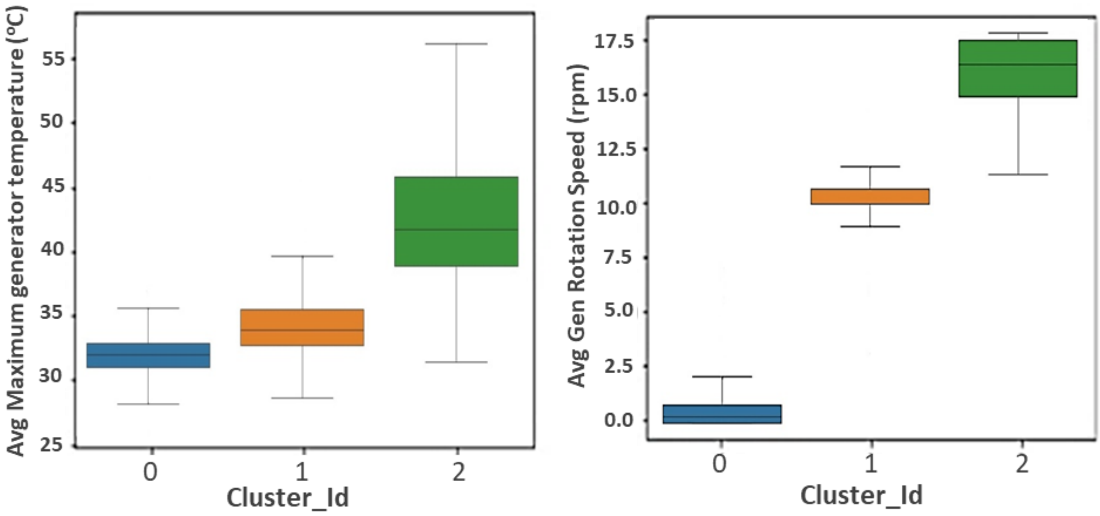


Figure 10. Data visualization of generator temperature and rotation speed according to K-means clustering.

In Figure 11, the relationship between rotation speed and temperature is analyzed. As the rotation increases, the temperature also increases. Therefore, three classes are extracted after the clustering method.



(**a**) (**b**)

Figure 11. Boxplot variables: (**a**) rotation speed vs. cluster id; (**b**) temperature vs. cluster id.

3.5. Validation of Results

Validation of the predictive performance of the results with the latest data was obtained. After processing the model again for getting qualitative performance, SCADA operational data were recorded continuously at 10 min intervals. This can take the form of the average, minimum, maximum, or standard deviation of live values recorded by the controller in the previous 10 min period. Signals such as the turbine power output, wind speed, temperatures of various components, electrical signals, and environmental conditions such as anemometer-measured wind speed and ambient temperature were recorded. The flow chart of validation is shown in Figure 12.

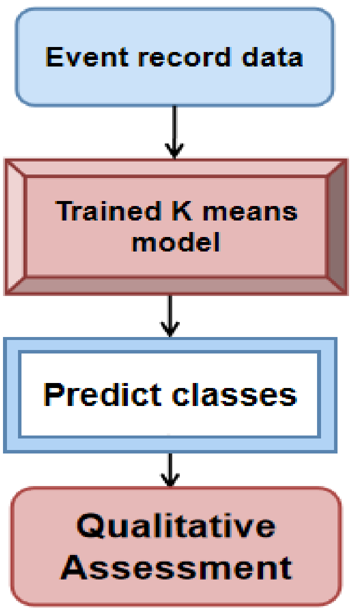


Figure 12. Flow chart of validation.

The comparison of average generator rotation speed and average generator temperature in Figure 13 shows their performance with their pros and cons in a particular scenario. The suggested evaluation approach can effectively forecast the change in operating circumstances prior to fault occurrences and can provide early warning of developing faults in DFIG bearings, according to the validation results.

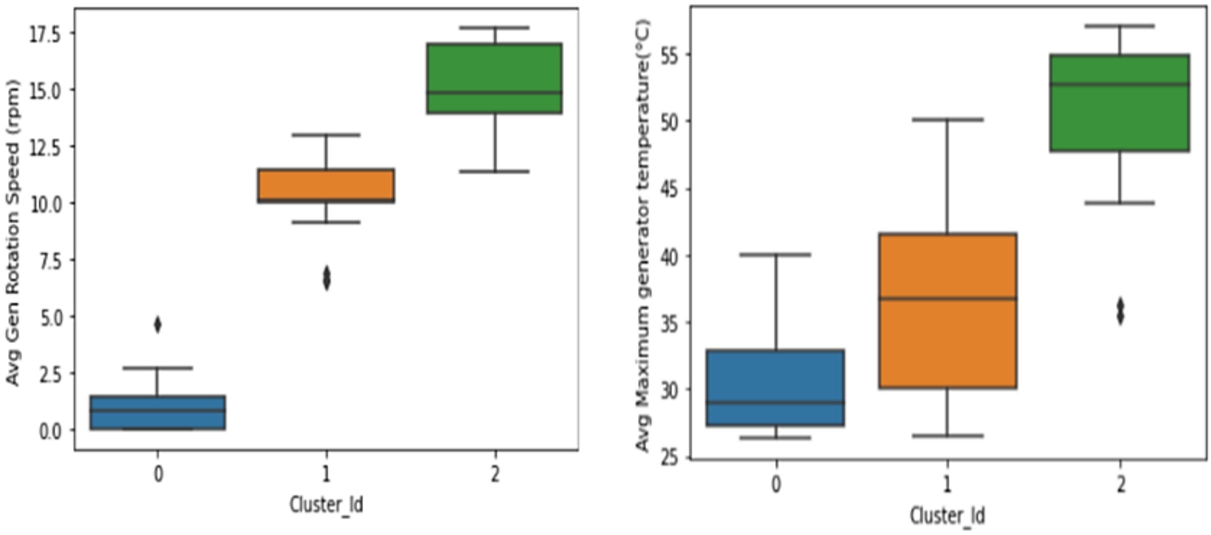


Figure 13. Validation of predicted results.

Validation of this research shows the change in speed and temperature. If the temperature exceeds greater than the range of 25 °C to 75 °C, it may fault the DFIG. After applying the machine learning technique, it was found that the rotation of the generator changes its speed, the high temperature could also be exceeded, and it can affect DFIG and could damage the bearings. Therefore, this research provides early fault detection of DFIG bearings.

4. Conclusions, Social Impact, and Recommendations

The approach was applied in a field study using one month of SCADA data from a wind farm consisting of 66 1.5 MW turbines in order to develop an effective condition monitoring system for early diagnosis and prognoses the conditions of the wind turbine's drive train to investigate bearing failures in DFIG. The box plot shows the visual graph of the affected parameters. In order to achieve forecasting with high accuracy, this paper proposes a novel model for the bearings in DFIG and a machine learning method for predictive maintenance. Therefore, forecasting the trend of temperature change is critical for overheating warnings. In order to evaluate the performance of the clustering model, experimental analysis was carried out. By combining the condition parameters in a scatter matrix, the linear elbow technique revealed the relationships between temperature and related variables. The results of comparative studies and early fault diagnosis show that the proposed method has better performance for temperature forecasting and average rotation speed of the main bearing of large-scale DFIG bearings. If these data can be used to identify potential failures or breakdowns, then this may well prove to be a very cost-effective means of condition monitoring; the SCADA data parameters prove to be a cost-effective method of condition monitoring that can be used for potential faults or malfunctions. The method will enable reducing downtime and monetary losses due to maintenance and replacement of various wind turbine components. The study provides an adaptable but reliable framework for the early identification of developing wind turbine damage, reducing wind turbine outages, and raising wind turbine dependability and revenue through operational improvement. Forecasting the trend of temperature change is essential for issuing overheating alerts. The outcomes of comparative studies and early fault diagnosis demonstrate that the proposed method has improved performance for temperature and rotation speed for forecasting the bearing of DFIG bearings. This research suggests a novel model for the bearings in DFIG and a machine learning method for preventive maintenance to achieve forecasts with high accuracy.

The sustainability of wind energy generation is ensured by using the approach suggested in this paper. Additionally, wind turbines may lessen the amount of power produced using fossil fuels, which lowers overall air pollution and carbon dioxide emissions. The potential for wind turbines to negatively impact wild animals through collisions as well as indirectly through noise pollution, habitat loss, and decreased survival or reproduction is a major issue for the business. Moreover, this study could be carried out to analyze the viability of this method for more than seven parameters of DFIG wind turbines in the future. With some changes, our project can also be used for health/condition monitoring of other equipment such as wind turbine gearboxes, medical sectors, helicopters, cars, etc.

Author Contributions: Conceptualization, A.A.M., N.H.M., B.S.C., L.K. and Q.T.T.; validation, A.A.M., N.H.M., B.S.C., L.K. and G.Z.; formal analysis, A.A.M., N.H.M., B.S.C., L.K. and G.Z.; writing—original draft preparation, A.A.M., N.H.M., B.S.C. and L.K.; writing—review and editing, A.A.M., N.H.M., B.S.C., L.K., Q.T.T. and G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The first author of the article has received funding/stipend from ICT endowment of Mehran University of Engineering and Technology, Jamshoro; Pakistan.

Acknowledgments: This research work is carried out at Mehran University of Engineering and Technology and is supported by NCRA established at the IICT building, Mehran UET, Jamshoro. Gaetano Zizzo is supported by the Research Fund for the Italian Electrical System under the Contract Agreement "Accordo di Programma 2022-2024" between ENEA and the Ministry of Ecological Transition.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Soomro, M.A.; Memon, Z.A.; Kumar, M.; Baloch, M.H. Wind energy integration: Dynamic modeling and control of DFIG based on super twisting fractional order terminal sliding mode controller. *Energy Rep.* **2021**, *7*, 6031–6043. [CrossRef]
2. Khatri, S.A.; Harijan, K.; Uqaili, M.A.; Shah, S.F.; Mirjat, N.H.; Kumar, L. A Logistic Modelling Analysis for Wind Energy Potential Assessment and Forecasting its Diffusion in Pakistan. *Front. Energy Res.* **2022**, *10*, 860092. [CrossRef]
3. *Global Wind Report 2022*; Global Wind Energy Council: Brussels, Belgium, 2022.
4. Annual Report 2021-22. Available online: https://www.nepra.org.pk/publications/Annual%20Reports/Annual%20Report%202021-22.pdf (accessed on 6 February 2023).
5. Turi, J.A.; Rosak-Szyrocka, J.; Mansoor, M.; Asif, H.; Nazir, A.; Balsalobre-Lorente, D. Assessing Wind Energy Projects Potential in Pakistan: Challenges and Way Forward. *Energies* **2022**, *15*, 9014. [CrossRef]
6. Asghar, R.; Ullah, Z.; Azeem, B.; Aslam, S.; Hashmi, M.H.; Rasool, E.; Shaker, B.; Anwar, M.J.; Mustafa, K. Wind Energy Potential in Pakistan: A Feasibility Study in Sindh Province. *Energies* **2022**, *15*, 8333. [CrossRef]
7. Solangi, Y.A.; Tan, Q.; Khan, M.W.A.; Mirjat, N.H.; Ahmed, I. The selection of wind power project location in the Southeastern Corridor of Pakistan: A factor analysis, AHP, and fuzzy-TOPSIS application. *Energies* **2018**, *11*, 1940. [CrossRef]
8. Kusiak, A.; Verma, A. Analyzing bearing faults in wind turbines: A data-mining approach. *Renew Energy* **2012**, *48*, 110–116. [CrossRef]
9. Tavner, P.J.; Greenwood, D.M.; Whittle, M.W.G.; Gindele, R.; Faulstich, S.; Hahn, B. Study of weather and location effects on wind turbine failure rates. *Wind Energy* **2013**, *16*, 175–187. [CrossRef]
10. Islam, M.M.; Sohag, K.; Alam, M.M. Mineral import demand and clean energy transitions in the top mineral-importing countries. *Resour. Policy* **2022**, *78*, 102893. [CrossRef]
11. Zhang, D.; Yuan, J.; Zhu, J.; Ji, Q.; Zhang, X.; Liu, H. Fault Diagnosis Strategy for Wind Turbine Generator Based on the Gaussian Process Metamodel. *Math. Probl. Eng.* **2020**, *2020*, 4295093. [CrossRef]
12. Garlick, W.G.; Dixon, R.; Watson, S.J.; Garlick, W.G.; Dixon, R.; Watson, S.J. A Model-Based Approach to Wind Turbine Condition Monitoring Using SCADA A Model-Based Approach to Wind Turbine Condition Monitoring Using SCADA Data Data A Model-based Approach to Wind Turbine Condition Monitoring Using SCADA Data. Available online: https://hdl.handle.net/2134/5488 (accessed on 6 February 2023).
13. Elasha, F.; Shanbr, S.; Li, X.; Mba, D. Prognosis of a wind turbine gearbox bearing using supervised machine learning. *Sensors* **2019**, *19*, 3092. [CrossRef] [PubMed]

14. Zhou, Y.; Zhi, G.; Chen, W.; Qian, Q.; He, D.; Sun, B.; Sun, W. A new tool wear condition monitoring method based on deep learning under small samples. *Measurement* **2022**, *189*, 110622. [CrossRef]
15. Hart, E.; Clarke, B.; Nicholas, G.; Kazemi Amiri, A.; Stirling, J.; Carroll, J.; Dwyer-Joyce, R.; McDonald, A.; Long, H. A review of wind turbine main bearings: Design, operation, modelling, damage mechanisms and fault detection. *Wind. Energy Sci.* **2020**, *5*, 105–124. [CrossRef]
16. Li, Y.; Zhu, C.; Chen, X.; Tan, J. Fatigue reliability analysis of wind turbine drivetrain considering strength degradation and load sharing using survival signature and FTA. *Energies* **2020**, *13*, 2108. [CrossRef]
17. Soomro, M.A.; Memon, Z.A.; Kumar, M. PWM Based VSC for Power Quality Assessment of Grid Integrated DFIG-WECS. *Int. J. Integr. Eng.* **2020**, *12*, 239–252. [CrossRef]
18. Liu, W.Y.; Tang, B.P.; Han, J.G.; Lu, X.N.; Hu, N.N.; He, Z.Z. The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review. *Renew. Sustain. Energy Rev.* **2015**, *44*, 466–472. [CrossRef]
19. Han, W.; Kim, J.; Kim, B. Effects of contamination and erosion at the leading edge of blade tip airfoils on the annual energy production of wind turbines. *Renew Energy* **2018**, *115*, 817–823. [CrossRef]
20. Qian, P.; Ma, X.; Zhang, D. Estimating health condition of the wind turbine drivetrain system. *Energies* **2017**, *10*, 1583. [CrossRef]
21. Gu, H.; Liu, W.; Gao, Q.; Zhang, Y. A review on wind turbines gearbox fault diagnosis methods. *J. Vibroeng.* **2021**, *23*, 26–43. [CrossRef]
22. de Oliveira-Filho, A.M.; Cambron, P.; Tahan, A. Condition Monitoring of Wind Turbine Main Bearing Using SCADA Data and Informed by the Principle of Energy Conservation. In Proceedings of the 2022 Prognostics and Health Management Conference, PHM-London 2022, London, UK, 27–29 May 2022; pp. 276–282. [CrossRef]
23. Kumar, A.; Gandhi, C.P.; Vashishtha, G.; Kundu, P.; Tang, H.; Glowacz, A.; Shukla, R.K.; Xiang, J. VMD based trigonometric entropy measure: A simple and effective tool for dynamic degradation monitoring of rolling element bearing. *Meas. Sci. Technol.* **2022**, *33*, 014005. [CrossRef]
24. Zhang, Z.Y.; Wang, K.S. Wind turbine fault detection based on SCADA data analysis using ANN. *Adv. Manuf.* **2014**, *2*, 70–78. [CrossRef]
25. Yucesan, Y.A.; Viana, F.A.C. A Physics-informed Neural Network for Wind Turbine Main Bearing Fatigue. *Int. J. Progn. Health Manag.* **2020**, *11*. [CrossRef]
26. Entezami, M. Wind Turbine Condition Monitoring System Wayside Acoustic Monitoring for Train Bearings View Project EC FP7-NIMO View Project. 2010. Available online: https://www.researchgate.net/publication/242568720 (accessed on 6 February 2023).
27. Umargono, E.; Suseno, J.E.; Gunawan, S. *K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based-On Mean and Median*; Atlantis Press: Paris, France, 2020; pp. 234–240. [CrossRef]
28. Keim, D.A.; Hao, M.C.; Dayal, U.; Janetzko, H.; Bak, P. Generalized scatter plots. *Inf. Vis.* **2010**, *9*, 301–311. [CrossRef]
29. Fränti, P.; Sieranoja, S. How much can k-means be improved by using better initialization and repeats? *Pattern Recognit* **2019**, *93*, 95–112. [CrossRef]