Short-Term Reliability Prediction of Key Components of Wind Turbine Based on SCADA Data

Ketian Liu\textsuperscript{1*}, Jun Zhang\textsuperscript{2} and Feng Su\textsuperscript{3}

\textsuperscript{1}School of Electric Power Engineering, Nanjing Institute of Technology, Nanjing, China
\textsuperscript{2}State Grid Energy Research Institute Co., LTD, Beijing, China
\textsuperscript{3}State Grid Corporation of China, Beijing, China

*Corresponding author e-mail: kt_liu@126.com

Abstract. In this paper, the Principal Component Analysis (PCA) method combined with the Radial Basis Function (RBF) neural network is used to establish a short-term reliability prediction model for wind turbines based on the SCADA data. The PCA method is used to reduce the dimensionality of the SCADA data and extract the principal components as the input data of the RBF neural network. The RBF neural network is used to predict the running state of key components of the wind turbine. Finally, a short-term reliability prediction model of wind turbines based on PCA-RBF is established. With the real wind farm SCADA data, the short-term reliability of wind turbine gearbox is predicted. The result shows that the short-term reliability prediction model can better reflect the reliability of key components and provide reference for the operation and maintenance of wind turbines.

1. Introduction

With the rapid development of wind power industry, a series of problems are produced recent years. What is remarkable are that the fault early warning and reliability assessment techniques for wind turbines have not been significantly improved resulting in the frequent failures of wind turbines. The penetration of wind energy into power systems is steadily increasing and highlights the importance of operations, maintenance, and specifically the role of condition monitoring [1]. At the same time, the geographical locations and operation characteristics determine that the reliability of wind turbines is significantly different from that of the general industrial similar system and make the maintenance difficult [2]. How to use the historical and real-time monitoring data to establish a reasonable reliability prediction model is the key to ensure the stable operation of the wind turbines.

The wind farm has been installed with SCADA system at the beginning of operation. The use of SCADA data for wind turbine fault prediction and reliability analysis does not require additional hardware investment. So how to use SCADA data to develop wind turbine fault diagnosis and reliability analysis have become a research hotspot. Wind turbine fault prediction is mainly used to rationally arrange wind turbine maintenance. The three mainly methods for wind turbine fault and reliability analysis with SCADA data are signal trending, artificial intelligence (ANN) and physical model [3-5]. The signal trending method determines the faulty components of the wind turbine by comparing the SCADA data of the normal wind turbine with the faulty wind turbine during the same period. In [6], the wind turbine blade aging was tested and the wind turbine maintenance plan was optimized. With the development of artificial intelligence technology and data processing technology,
mathematical methods such as Markov prediction method, fuzzy neural network prediction method and renewal process are applied to the research of preventive maintenance modeling of wind turbines [7-9]. In order to improve the operating efficiency of wind turbines, the Markov model of component downtime and faults can be established at different stages of the life cycle of the wind turbine [10]. Literature [11] proposes a new method for fault diagnosis of wind turbines based on the co-integration analysis of the SCADA data. Literature [12] uses statistical algorithms to analyze the real-time operating states of various components of the wind turbines and predicts the overall reliability of wind power equipment. Due to the complexity of the working environment of wind turbines, the fault characteristics and reliability of wind turbines are significantly different in different regions [13].

The above research process are mostly based on theoretical mechanism, and the practical applications are difficult. Based on the wind turbine SCADA data, this paper uses PCA-RBF neural network to establish a short-term fault prediction model and reliability assessment model for key components of wind turbines by statistical methods.

2. PCA-RBF based fault prediction of key components of wind turbines

2.1. PCA algorithm principle

The idea of the PCA algorithm is to construct new variables formed by linear combination of original variables, so that the new variables reflect the information of the original variables as much as possible without being related to each other. The PCA method maps the original high-dimensional input variables to low dimensional ones to achieve the purpose of reducing dimensionality and reducing data quantity. The PCA algorithm can avoid the single indicator not reflecting the global data information and the information redundancy caused by a large number of indicators, and remove the correlation between the different parameter data of the original data.

Assume that the standardized sample set is \( X \), then the autocorrelation matrix \( R \) is calculated as follow,

\[
R = X^T X / (N - 1)
\]

where, \( N \) is the amount of samples.

The eigenvalues of autocorrelation matrix \( R \) are \( \lambda_1 \geq \lambda_2 \cdots \geq \lambda_m \), and the eigenvectors of every eigenvalue are \( \mu_1, \mu_2, \cdots, \mu_m \). Then, the sample matrix variance contribution \( \eta_i \) of \( \lambda_i \) is,

\[
\eta_i = \frac{\lambda_i}{\sum \lambda_i} \times 100\%
\]

The quantity of principal components depends on the cumulative variance contribution. Generally, when the cumulative variance contribution rate is greater than 75%~95%, the corresponding first \( p \) principal components contain most of the information that the original variables can provide. Then, the quantity of principal components is \( p \). The cumulative contribution of the first \( p \) eigenvalues arranged from large to small is,

\[
\eta_c(p) = \sum_{i=1}^{p} \eta_i
\]

The eigenvector corresponding to the first \( p \) eigenvalues constitute the reduced-dimensional coordinate system \( U_{m\times p} \):

\[
U_{m\times p} = \begin{bmatrix} \mu_1, & \mu_2, & \cdots, & \mu_p \end{bmatrix}
\]

Then the principal component \( Z_{N\times p} \) corresponding to the standardized new sample \( X^*_{N\times m} \) is

\[
Z_{N\times p} = X^*_{N\times m} \times U_{m\times p}
\]

2.2. PCA-RBF based fault prediction model

RBF neural network is widely used as an artificial neural network model. It is a neural network formed by introducing Radial Basis Function into neural networks [14]. Its remarkable features are simple structure, fast convergence, and the ability to approximate any nonlinear function. In this paper,
the PCA is used to reduce the dimension of the original sample data, and the PC of the original data is used as the input of the RBF neural network for the purpose of the RBF neural network training and wind turbine fault prediction. The PCA-RBF based wind turbine component failure prediction model is shown in Figure 2.

Select Gauss function as the activation function of PCA-RBF prediction model. The activation function

\[ R_{(x, c)} = \exp\left(-\frac{1}{2\sigma^2} \|X_p - c\|^2\right) \]  

(6)

Where, \(\|\|\) is the Euclidean norm, \(c\) is the center of Gauss function. \(\sigma\) is the variance.

The output of the PCA-RBF prediction model is weighted and summed by the output values of the hidden layer. The prediction results of PCA-RBF prediction model are

\[ y_j = \sum_{i=1}^{h} w_{ij} \exp\left(-\frac{1}{2\sigma^2} \|X_p - c_i\|^2\right) \]  

(7)

Where, \(X_p\) is the \(p\)th input sample, \(p=1, 2, \cdots, P\) and \(P\) is the total number of input samples. \(c_i\) is the center of the hidden layer structural point. \(w_{ij}\) is the connection weight value between hidden layer structural point \(i\) and output layer node \(j\) with \(i=1, 2, \cdots, h\) and \(j=1, 2, \cdots, n\). \(y_j\) is the output of sample.

![Figure 1. The PCA-RBF based wind turbine component fault prediction model](image)

3. PCA-RBF based reliability prediction of key components of wind turbines

The short-term reliability assessment not only needs to know the relevant online operating parameters of the wind turbine itself, but also needs to reflect the environmental data of its operating state. The wind farm SCADA system can monitor the operating state parameters of the wind turbine in real time and provide the data basis for short-term reliability assessment.

3.1. Short-term operational reliability calculation

Equipment monitoring data and short-term operational reliability can be approximated as follows,

\[ R = f(x_1, x_2, \cdots, x_n) \]  

(8)

Where, \(R\) is the current reliability of the equipment. \(f\) is a functional relationship between different parameters and short-term operational reliability. \(x_i\) is the \(i\)th monitoring parameter.

If the equipment monitoring value is within the optimal value or within the optimal range, the equipment’s short-term reliability reaches the maximum value \(R_{max}\). When the monitoring parameter exceeds the failure threshold, the current short-term reliability reaches the minimum value \(R_{min}\). When the monitoring parameter is between the optimal value and the threshold, the short-term operational reliability gradually decreases with the deviation from the optimal value. For the convenience of
analysis, the maximum value of reliability is 1 and the minimum value is 0, which does not distinguish between positive and negative. Therefore, the range of reliability reduction is [0, 1].

If a device has only a single monitoring parameter, the operating deviation is expressed as,
\[
D_r = \begin{cases} 
\frac{x_b - x}{x_b - x_{\text{min}}} & x_b > x > x_{\text{min}} \\
\frac{x - x_b}{x_{\text{max}} - x_b} & x_{\text{max}} > x > x_b 
\end{cases}
\]
(9)
Where, \(D_r\) is the operation deviation. \(x_b\) is the optimum value for the monitoring parameters. \(x_{\text{max}}\) and \(x_{\text{min}}\) are the maximum failure value and the minimum failure value of the monitoring parameters respectively. \(x\) is the current monitoring parameter.

The decrease in equipment reliability \(R_r\) is expressed as the maximum reliability minus the equipment operation deviation.
\[
R_r = R_{\text{max}} - D_r
\]
(10)
Where, \(R_r\) is the reliability of the device at the current time. \(R_{\text{max}}\) is the best reliability value, \(R_{\text{max}}=1\) in this paper.

In general, the optimum value of wind turbine monitoring parameters is affected by environmental factors. At different wind speeds, the optimum value of the generator temperature is also different. Some wind turbine SCADA systems will give the optimum monitoring parameters’ value \(x_b\). It can also determine the optimal operating value based on experience combined with the working environment, expert experience, fault mechanism and other factors. The upper and lower limits for important parameters are recorded in the SCADA system.

For wind turbines, many parameters often have the best monitoring range \([x_{\text{bmin}}, x_{\text{bmax}}]\). If the device is operated within this range, its status is considered to be optimal and its reliability is the highest. Then the wind turbine operation deviation is,
\[
D_r = \begin{cases} 
1 & x < x_{\text{min}} \\
\frac{x_{\text{bmax}} - x}{x_{\text{bmax}} - x_{\text{bmin}}} & x_{\text{bmin}} \leq x \leq x_{\text{bmax}} \\
0 & x_{\text{bmin}} \leq x \leq x_{\text{bmin}} \\
\frac{x_{\text{bmax}} - x}{x_{\text{max}} - x_{\text{bmax}}} & x_{\text{bmin}} \leq x \leq x_{\text{max}} \\
1 & x > x_{\text{max}}
\end{cases}
\]
(11)
It can be seen from the above equation that the range of operation deviation is [0, 1], which indicates the extent to which the actual operating state deviates from the optimal operating state.

The actual wind turbine has multiple monitoring parameters and the parameters are also related to each other. The PCA method introduced above is used to extract the principal components of the parameters and eliminate the interaction between the parameters. The weight of each principal component is determined using the entropy method. The wind turbine operation deviation can be obtained by calculating the deviation of the principal component.
\[
D_r = \sum_{i=1}^{P} \alpha_i D_{ci}
\]
(12)
Where, \(D_{ci}\) is the operation deviation of the \(i^{th}\) principal component, \(i=1,2,\ldots, P\). \(\alpha_i\) is the weight coefficient of the \(i^{th}\) principal component after the entropy method.

### 3.2. Division of short-term reliability evaluation levels

The short term reliability assessment of the equipment includes: evaluating the current short-term reliability level of the equipment according to the equipment operation reliability deviation; analyzing the operation status of the equipment in the future according to the short-term reliability prediction, and proposing corresponding measures. Referring to the reliability rating standard of general
electromechanical equipment, the reliability evaluation level of key components of wind turbines is divided into 4 levels. The value of different levels of reliability obtained by statistical analysis needs to fully consider the equipment failure mechanism and combine the historical data failure conditions, as shown in Table 1.

Table 1. Short-term reliability level division

| Reliability level | Threshold | Description                                           |
|-------------------|-----------|-------------------------------------------------------|
| 1                 | 0.2       | Very low reliability, impending exceptions            |
| 2                 | 0.5       | Low reliability and functional degradation            |
| 3                 | 0.7       | The reliability is slightly lower and the function is basically normal |
| 4                 | 0.9       | Good reliability and normal function                  |

4. Case Study

The paper only considers the impact of wind speed on the reliability of wind turbine gearboxes. To predict the reliability of a wind turbine gearbox. The upper and lower limits of the failure temperature threshold are 0 °C and 70 °C in this paper. The wind turbine SCADA data at certain moments those need to predict the reliability are shown in Table 2.

Table 2. Gearbox monitoring data for reliability prediction

| Moment | Active power | Ambient temperature | Wind speed m/s | Bearing temperature °C | Lubricating oil temperature °C |
|--------|--------------|---------------------|----------------|------------------------|--------------------------------|
| 1      | 1117.3       | 13.1                | 11.16          | 48.7                   | 53.4                           |
| 2      | 1111.69      | 13.4                | 11.11          | 49.5                   | 55.1                           |
| 3      | 1269.74      | 13.6                | 10.04          | 51                     | 56                             |
| 4      | 1349.25      | 13.6                | 11.17          | 51.3                   | 57.7                           |
| 5      | 1295.85      | 13.9                | 9.94           | 52.6                   | 58.7                           |
| 6      | 1205.84      | 13.9                | 12.16          | 52.6                   | 59.4                           |
| 7      | 1224.87      | 14                  | 10.16          | 52.7                   | 58.8                           |
| 8      | 1244.86      | 14                  | 11.47          | 52.6                   | 59.4                           |
| 9      | 1243.16      | 14.2                | 11.63          | 52.6                   | 59.9                           |
| 10     | 1724.38      | 14                  | 11.27          | 52.7                   | 60.2                           |
| 11     | 1178.28      | 14.2                | 11.35          | 52.9                   | 60.2                           |
| 12     | 1292.43      | 14.3                | 11.48          | 52.3                   | 60.8                           |
| 13     | 0            | 14.3                | 12.01          | 49.6                   | 59                             |
| 14     | 0            | 14.3                | 11.45          | 50.9                   | 58.1                           |
| 15     | 0            | 14.2                | 11.43          | 51.8                   | 57.6                           |

Use the normal operating wind turbine SCADA data to train the PCA-RBF model, and predict the wind turbine bearing temperature and lubricant oil temperature at each moment in Table 2. The gearbox temperature deviation value is calculated using Equation 13, where α is taken as 0.5. The reliability of the gearbox at each moment in Table 2 was calculated by Equation 11, and the calculation results are shown in Table 3 and Figure 2.

![Figure 2](image-url) Wind turbine gearbox reliability trend at all moments
### Table 3. Gearbox reliability predictions at all moments

| moment | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|
| Reliability | 0.86 | 0.96 | 0.79 | 0.64 | 0.48 | 0.43 | 0.46 | 0.43 |

| moment | 9   | 10  | 11  | 12  | 13  | 14  | 15  |
|--------|-----|-----|-----|-----|-----|-----|-----|
| Reliability | 0.39 | 0.36 | 0.35 | 0.35 | 0.67 | 0.64 | 0.61 |

5. Conclusion

In this paper, the short-term reliability prediction model of wind turbines is established by combining PCA with RBF neural network. The PCA method is used to reduce the dimension and extract the principal components of the wind turbine SCADA data, which effectively improves the prediction speed of the RBF. The PCA-RBF based short-term reliability prediction model established in this paper can better predict the operational reliability of key components of wind turbines and reflect the potential failures of wind turbines. The research results in this paper can provide some help for the operation and maintenance of wind turbines.

6. Acknowledgments

This work was financially supported by science and technology project of State Grid Corporation of China (WBS:521104170016).

References

[1] Pandit R K, Infield D, SCADA-based wind turbine anomaly detection using Gaussian process models for wind turbine condition monitoring purposes, IET Renewable Power Generation, 12 (2018) 1249-1255.

[2] Johnson K, Wingerden J V, Balas M J, et al, Special issue on Past, present and future modeling and control of wind turbines, Mechatronics, 21(2011) 633.

[3] Wilkinson, M, Comparison of methods for wind turbine condition monitoring with SCADA data, IET Renewable Power Generation, 8 (2014): 390-397.

[4] Wang K, Sharma V S, Zhang Z, SCADA data based condition monitoring of wind turbines, Advances in Manufacturing, 2 (2014) 61-69.

[5] Tautz-Weinert J, Watson S J, Using SCADA data for wind turbine condition monitoring – a review, IET Renewable Power Generation, 11 (2017) 382-394.

[6] Yang W, Jiang J, Wind turbine condition monitoring and reliability analysis by SCADA information, 2011 Second International Conference on Mechanic Automation and Control Engineering, Hohhot, 2011, 1872-1875.

[7] Dong M, He D, Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis, European Journal of Operational Research, 178(2007) 858-878.

[8] Dong M, He D, A segmental hidden semi-Markov model based diagnostics and prognostics framework and methodology, Mechanical Systems and Signal Processing, 21 (2007) 2248-2266.

[9] Chaturvedi D, Sinha A, Malik O, Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network, International Journal of Electrical Power & Energy Systems, 67 (2015) 230-237.

[10] Ossai C I, Boswell B, Davies I J, A Markovian approach for modelling the effects of maintenance on downtime and failure risk of wind turbine components, Renewable Energy, 96 (2016) 775-783.

[11] Dao P B, Staszewski W J, Barszcz T, et al, Condition monitoring and fault detection in wind turbines based on cointegration analysis of SCADA data, Renewable Energy, 116 (2018) 107-122.

[12] Kusiak A, Li W, The prediction and diagnosis of wind turbine faults, Renewable Energy, 36 (2011) 16-23.

[13] Johnson K, Wingerden J W V, Balas M J, et al, Special issue on “Past, present and future modeling and control of wind turbines”, Mechatronics, 21 (2011) 633.

[14] Chen X, Shen W, Cao Z, et al, Robust sliding mode observer using RBF neural network for lithium-ion battery state of charge estimation in electric vehicles, 2014 9th IEEE Conference on Industrial Electronics and Applications. IEEE, 2014, 42-47.