Research on Fault Detection Method of Wind Turbine Generator Based on SCADA Data

Xin Wang¹,², *, Pengfei Guo¹,², Weijiang Liu¹,² and Chao Li³

¹Zhejiang Windey Co., Ltd, Hangzhou, China
²Key Laboratory of Wind Power Technology of Zhejiang Province, Hangzhou, China
³China Standard & Quality Control Research Institute, Ministry of Water Resource, Hangzhou, China

*Corresponding author e-mail: wangx@chinawindey.com

Abstract. In this paper, a data driven fault detection model of the generator based on improved nonlinear state estimation (INSET) method was established and the residuals of all the relevant parameters were predicted synchronously. According to the residual distribution characteristics the alarm rules were designed. In case studies, the fault was detected timely and exactly for avoiding serious accident. In addition, compared with other conventional algorithms, the results showed higher prediction accuracy and sensitivity and indicated the feasibility of the method.

1. Introduction

Most wind turbines operate in harsh outdoor natural environment and have more faults than conventional generators. At present wind turbine generator fault detection mainly rely on SCADA system but most of the faults didn’t trigger SCADA alarm until developed to serious. In recent years, the application of wind turbine operation data has been deepened gradually. Many scholars have carried out a lot of research in the field of wind turbine fault detection based on SCADA data [1-5]. In this paper, a generator fault detection model was established based on an improved non-linear state estimation (INSET) method. A process memory matrix which covered the whole normal working space of the generator was constructed by extracting characteristic variables properly. Reasonable alarm rules were designed according to the residual distribution characteristics of observation variables, and the trend of generator fault was monitored in real time to improve the sensitivity and accuracy of fault detection.

2. Selection of Characteristic Variables

Generator faults usually characterized by three forms: vibration, noise and temperature. When the bearing of generator is damaged, the winding is aged or the ventilation is not smooth, the temperature of generator will exceed the normal level [6]. In order to monitor the temperature change of the generator, 5 variables including generator front bearing temperature, generator back bearing temperature and stator 1/2/3 winding temperature were selected as the observation variables. The rotor speed, power, wind speed, ambient temperature and nacelle transverse vibration acceleration were in the top 6 average correlation coefficients with the five observation variables.
Considering the changes of temperature rise in most cases can better characterize the abnormal state of wind turbine, the following 2 variables were constructed as characteristic variables:

Front bearing temperature-rotor speed ratio: The ratio of front bearing temperature change to rotor speed change at adjacent sampling time;

Back bearing temperature-rotor speed ratio: The ratio of back bearing temperature change to rotor speed change at adjacent sampling time.

3. Construction of Fault Detection Model

3.1. Principle of the NSET Model

NSET was a non-parametric adaptive modeling method widely used [7]. It had been successfully applied in equipment monitoring and life prediction of electronic products [8-10]. A set of n input values at time \( i \) was defined as the observation vector \( X(i) \).

\[
X(i) = [x_1(i) \quad x_2(i) \quad \ldots \quad x_n(i)]^T
\]  

(1)

All observation vectors were chosen from historical data in normal state to compose matrix \( D \) which called memory matrix.

\[
D = \begin{bmatrix}
X(1) & X(2) & \ldots & X(m) \\
x_1(1) & x_1(2) & \ldots & x_1(m) \\
x_2(1) & x_2(2) & \ldots & x_2(m) \\
\vdots & \vdots & \vdots & \vdots \\
x_n(1) & x_n(2) & \ldots & x_n(m)
\end{bmatrix}_{n \times m}
\]

(2)

The input values was defined as vectors \( X_{obs} \) and output as \( X_{est} \). It assumed that the output value is a linear combination of all observation vectors in memory matrix \( D \).

\[
X_{est} = DW
\]

(3)

Where

\[
W = [w_1 \quad w_2 \quad \ldots \quad w_m]^T
\]

Then the residual of the input and output values was

\[
\epsilon = X_{est} - X_{obs}
\]

(4)

The \( W \) could be obtained by minimizing the residual.

\[
W = \left( D^T \cdot D \right)^{-1} \cdot \left( D^T \cdot X_{obs} \right)
\]

(5)

Euclidean distance between two vectors was directly proportional to the similarity.

\[
W = \left( D^T \otimes D \right)^{-1} \cdot \left( D^T \otimes X_{obs} \right)
\]

(6)

\[
X \otimes Y = \sqrt{\sum_{j=1}^{p} (x_j - y_j)^2}
\]

(7)

It can be obtained that the relationship between the output and the input value of the model.
$$X_{est} = D \cdot \left(D^T \otimes D\right)^{-1} \cdot \left(D^T \otimes X_{obs}\right)$$ (8)

3.2. Construction of Process Memory Matrix $D$

It was necessary to choose a reasonable and accurate train set which covered the whole workspace of the wind turbine in normal operation, so the data must have a long enough time span and power limitation and fault condition should be eliminated. Then all variables in the train set must be normalized to [0, 1].

In most papers the process memory matrix $D$ was generally constructed by equal step method [11]. In this paper $D$ was constructed by variable step based on the probability density distribution of each characteristic variable in the train set. According to probability density function based on characteristic variable, an equal ratio descent method was used to select the $\delta$ relying on the experience of model construction. The flow chart of construction matrix $D$ by variable step method was showed in Figure 1.

![Flow chart of construction matrix $D$ by variable step](image)

Figure 1. Flow chart of construction matrix $D$ by variable step

4. Validation of Prediction Effect

All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper. In this paper, an 115-2000MW wind turbine was selected to validate the model. A generator bearing fault occurred at 2018-4-5 11:20. Then the cooling circuit and bearings of the generator were checked and it was that the oil collect tank of the back-bearing is getting black and there is abnormal noise when turbine restarts. At last the turbine resumed operation after repair at 2018-4-6 8:02.

The historical operation data with 5 minutes sampling period from 2016-1-1 to 2017-12-31 were selected as train set. After eliminating the data of shutdown state, power limit and fault conditions, the remaining data amount was 168670. The $D$ was constructed as the process shown in Figure 1.

The data from 2018-1-24 to 2018-1-25 was randomly extracted, in which period the turbine operated normally all time. Multivariate linear regression (MLR), BP neural network (BPNN), NSET and the proposed INSET algorithm were used to established model. The prediction residual time series diagram of each algorithm was shown in Figure 2.
Table 1. Statistics of prediction results of different algorithms

| Algorithms | MSE   | $R^2$  |
|------------|-------|--------|
| MLR        | 0.1421 | 0.5433 |
| BPNN       | 0.1154 | 0.9115 |
| NSET       | 0.0951 | 0.9867 |
| INSET      | 0.0148 | 0.9982 |

The statistical results of mean square error MSE and correlation coefficient $R^2$ of each residual were shown in Table 1. It proved that the INSET algorithm had smaller prediction residual and better prediction effect than other conventional algorithms.

5. Alarm Rules and Fault Location
The model based on INSET algorithm could synchronously monitor all observation variables. The data for the period from 2018-3-20 to 2018-4-10 was selected as test set and predicted results were showed in Figure 3. There were obvious anomalies in the residuals of 5 observation variables before the SCADA system alarm time, and the residuals continued to grow larger until the back bearing was repaired on 2018-4-10. It proved the proposed model had a high sensitivity to generator bearing abnormality.
In this paper, the ‘threshold + duration’ mode was chosen for fault alarm. The residual thresholds of 5 observed variables were set as follows: front bearing temperature ≤ 2 °C, back bearing temperature ≤ 2 °C, stator 1/2/3 winding temperature ≤ 4 °C. The duration was set to 2 hours uniformly. When the residual exceeded the threshold and lasted more than 2 hours, an alarm was given out.

The original residuals were processed by moving average for marking alarm time based on threshold and duration. The alarm results were shown as Figure 3. The back bearing temperature residual exceeded alarm threshold at 2018-3-28 3:55, and it triggered alarm after 2 hours. The alarm time was 8 days and 5 hours earlier than the first alarm time of SCADA system. It was proved that the alarm mechanism can effectively detect the fault of generator system.

6. Conclusion
This paper established a data driven fault detection model of the generator based on INSET algorithm and completed model training with historical data. Two new variables were customized to improve the accuracy of detection. A variable step method taking full account of the distribution features of different variables was adopted to construct the process memory matrix $D$. Finally, the predicted residual is processed by moving average method, and a reasonable alarm mechanism was set for fault alarm and location. The actual verification results showed that the prediction residuals of INSET model is smaller than BP neural network and other algorithms. The proposed method could give alarm 8 days earlier than SCADA, which had great practical significance for the operation and maintenance work of wind farms.
Acknowledgments
This work was financially supported by Key Research and Development Program of Zhejiang Province (No. 2019C01050).

References
[1] Quiroga J, Cart ES D A, Edrington C S. Neural Network Based Fault Detection of PMSM Stator Winding Short Under Load Fluctuation. EPE-PEMC, 2008, pp. 793-798.
[2] Alahmar E, Benbouzid Meh, Amirat Y, Ben Elghalil S.E. DFIG Based Wind Turbine Fault Diagnosis Using a Specific Discrete Wavelet Transform. Proceedings of the 2008 International Conference on Electrical Machines, WuHan, China, 2009, pp. 17-20.
[3] Andrew Kusiak, Wenyan Li. The prediction and diagnosis of wind turbine faults. Renewable Energy, 36, pp. 16-23.
[4] P.Bangalore, M.Patriksson. Analysis of SCADA data for early fault detection, with application to the maintenance management of wind turbines. Renewable Energy, 115, pp. 521-532.
[5] Phong B. Dao, Wieslaw J. Staszewski, Tomasz Barszcz, Tadeusz Uhl. Condition monitoring and fault detection in wind turbines based on cointegration analysis of SCADA data. Renewable Energy, 116, pp. 107-122.
[6] Dong Mao. Research on the fault early warning methods of wind turbine generator under the big data. North China Electric Power University. Beijing, 2016.
[7] Gross K C, Singer R M, Wegerich S W. Application of a model-based fault detection system to nuclear plant signals. International Conference on Intelligent Systems Applications to Power Systems. Seoul, Korea, 1997.
[8] Bockhorst F K, Gross K C, Herzog J P. MSET modeling of crystal river-3 venturi flow meters. Proceedings of International Conference on Nuclear Engineering, San Diego, CA. 1998.
[9] Cheng S F, Pecht M G. Multivariate state estimation technique for remaining useful life prediction of electronic products. Proceedings of AAAI Fall Symp Artif. Intell. Prognostics (26-32), Arlington, VA. 2007.
[10] Cassidy K J, Gross K C, Malekpour A. Advanced pattern recognition for detection of complex software aging phenomena in online transaction processing servers. Proceedings of Dependable Systems and Networks, Washington, D.C., USA. 2002.
[11] Peng Guo, David Infield, Xiyun Yang. Wind Turbine Gearbox Condition Monitoring Using Temperature Trend Analysis. Proceedings of the CSEE, 31 (32), pp.129-136