Parameter Anomaly Identification Model About Supervisory Control And Data Acquisition

Guo Ping¹,*
¹Liaoning Jianzhu Vocational College, Liao Yang City, Liao Ning Province, 111000, China
*Corresponding author’s e-mail: guoping3360@126.com

Abstract. With the rapid progress of sensor technology, state monitoring of wind turbines is more comprehensive, the amount and type of the sensor used will be more, the state parameters of wind turbines will be more. According to the status indicators to parameter index of target parameters selector model is established in this paper, and the number of input parameters are reduced on the basis of the guarantee accuracy. The two sub-models constitute the abnormal identification model of the working condition index. Finally, the model is analyzed by an example to verify the accuracy and effectiveness of the model. The research in this paper on abnormal discrimination of working condition parameters and evaluation methods of wind turbines is an important premise and basis for making scientific and reasonable operation and maintenance decisions of wind farms, and it has important academic significance.

1. Introduction
Supervisory Control And Data Acquisition (SCADA) system is an online monitoring and Control system for wind turbines, which records and stores the operating Data of the turbines. Online monitoring information of SCADA system is the first-hand data to display the working conditions of wind turbines, which contains a large number of valuable information, and provides data support and possibility for abnormal analysis of working conditions parameters and evaluation of operating conditions of wind turbines [1-2]. Therefore, the research on abnormal discrimination of wind turbine working condition parameters and evaluation methods is an important prerequisite and basis for wind farm to make scientific and reasonable operation and maintenance decisions. It has important application value and academic significance to ensure the safety and improve the economy of wind farm[3].

2. SCADA system of wind turbine

2.1. An overview of the SCADA system
SCADA system which the full name is Data Acquisition And Monitoring Control System, and it is a commonly used wind farm monitoring and Control system[4]. When the monitored parameter index value is greater than the set threshold value, the system will issue a warning, and the work manager will carry out maintenance according to the relevant warning. Monitoring index parameters can be divided into two kinds, one is discrete quantity, the other is continuous quantity. The parameters to be monitored include:

1) Speed indicator. It mainly refers to the speed of the generator and the spindle. According to the size of the speed, the unit will take off the grid, connect to the grid and run the overspeed protection command.
2) Temperature monitoring index variables. The control system controls and protects the unit according to the temperature value. The temperature sensing equipment is distributed in the generator, gear box, outside the engine room and other positions. The temperature index parameters measured by the sensor are transmitted to the control system.

3) The wind monitoring. It includes wind speed and direction monitoring.

2.2. Data standardization processing about SCADA

SCADA system contains a variety of state parameters. Different state parameters have different measurement units, and the numerical range of different state parameters is very different, so correlation analysis cannot be conducted directly. Therefore, it is necessary to standardize the initial SCADA system data. The method of SCADA system state parameter standardization is as follows.

Assuming that the initial data matrix is $X = (x_{ij})_{n \times p}$, the normalized transformation of the initial data is carried out by formula (1).

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{D(x_j)}}$$

Where, $z_{ij}$ is an element of the standard transformation data matrix $Z = (z_{ij})_{n \times p}$; $D(x_j)$ is the variance before the transformation; $\bar{x}_j$ is the average, it is calculated by formula (2).

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$$

The variance after the normalization of state parameters is calculated by formula (3), and the result is 1.

$$D(z_j) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{x_{ij} - \bar{x}_j}{\sqrt{D(x_j)}} \right)^2 = \frac{D(x_j)}{D(x_j)} = 1$$

The covariance $u_{ij}$ is calculated by formula (4).

$$u_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} (z_{ik} - \bar{z}_i)(z_{kj} - \bar{z}_j)$$

Formula (5) can be obtained.

$$u_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} z_{ik}z_{kj} = \frac{1}{n-1} \sum_{k=1}^{n} \frac{x_{ik} - \bar{x}_i}{\sqrt{D(x_i)}} \frac{x_{kj} - \bar{x}_j}{\sqrt{D(x_j)}} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}}$$

Where, $s_{ij}$ is formula(6).

$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} (k_{ik} - \bar{k}_i)(x_{ij} - \bar{x}_j)$$

Correlation coefficient after data standardization is $r_{z_iz_j}$ by formula (7).

$$r_{z_iz_j} = \frac{u_{ij}}{\sqrt{u_{ii}u_{jj}}} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}} = r_{x_ix_j}$$

Formula(7) shows that the correlation between parameters and indicators has not been changed after normalization.
3. Operating condition parameter index selection sub-model

3.1. Forecast model input indicators
Wind turbine working condition parameters are divided into three categories, which are unit component working condition indicators, grid indicators and environmental indicators.

1) Working condition index of unit components. It includes two kinds: temperature index and vibration index.
2) The grid index. The output state of unit power can be reflected by two indexes: active power and reactive power.
3) Environmental indicators. There are two main environmental indicators, namely wind speed index and environmental temperature index.

3.2. Working condition parameter index selection sub-model
In the parameter modeling of wind turbine working condition index, the proper input index parameters should be selected first, and the relationship between the input index and the target parameter index should be determined. Therefore, a sub-model of working condition index selection based on BPNN is proposed.

1) BPNN based working condition parameter index selection model
   Determining the number of hidden layers of BPNN(Back-propagation neural network), the number of nodes of each layer and the sample data are the main process of BPNN prediction model construction. In wind power parameter prediction, the use of three-layer structure of BPNN can get more accurate results. Therefore, the BPNN parameter of single hidden layer is selected to select the prediction model.
   Methods is shown in Formula (8-10).

\[
W(t+1) = W(t) - \eta d(k) \tag{8}
\]
\[
d(k) = -[\hat{H}(k)]^{-1} g(k) \tag{9}
\]
\[
\hat{H}(k) = H(k) + \beta Q(k) \tag{10}
\]
   The calculation of the number of hidden layer nodes in the BPNN model is shown in formula (11).
\[
l = \sqrt{m + n + a} \tag{11}
\]
Ten thousand records were randomly selected as samples from the normal operation data of No. 15 unit of the wind farm, among which each record contained 20 unit working condition indicators, 9000 records were used as training samples, and 1000 records were used as test samples. Taking the temperature of generator bearing B as an example, the input index set was selected for the forecast model of working condition index parameters. The temperature of generator bearing B is the output index of the BPNN model, so the remaining 19 working condition indexes are the input indexes to be selected for the BPNN model. According to the above formula, the number of hidden layer nodes should be 5 ~ 14, then there will be 10 BPNN models, each of which will be trained 10 times, and the average root mean square error of the 10 times will be obtained. RMSE(Root mean square error) calculation is shown in formula (12).
\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{r}_i - r_i)} \tag{12}
\]

4. Abnormal discriminant sub-model of working condition index

4.1. Combined prediction model of performance indicators
Wind turbine bearing B temperature parameters were taken as the target parameters, and a linear combination prediction model was established based on BPNN, RBFNN (Radial Basis Function Neural
Network) and LS-SVM (Least Squares Support Vector Machines) models to minimize the sum of the squares of prediction errors. The specific process is as follows:

1) Ten thousand data were randomly selected from the one-year normal operation data of Unit 15 of the wind farm and divided into training data samples and test data samples according to the ratio of 9:1. According to the obtained data samples, the single and combined models are established.

2) Choose models with high precision. Further, the three models constructed in the first step were repeatedly trained and tested, and then the models with higher accuracy and accuracy were selected from these repeated methods.

3) N test samples are randomly selected from 10000 data samples, and each test sample contains 1000 data. The test samples were used to establish a single item model for testing, and the corresponding residual sequence of each test sample was recorded respectively. The weight distribution of each single model was obtained by using the non-negative weight optimal combination prediction iterative algorithm, and the combined model was finally obtained.

4.2. Anomaly analysis of state indicators
The main process of abnormal identification of state indicators is as follows: in the first step, we need to divide the working conditions or states of wind turbines and their subsystems into two kinds, namely, the normal and safe operating state and the abnormal operating state. The second step is to build the model of working condition index under normal state with the operation data of the unit in normal operation. The third step is to predict each state index with the help of the model established in the second step, and then calculate the residuals between them according to the predicted values and actual values of the prediction model. Generally, the smaller the residuals, the less serious the anomaly will be, and vice versa.

5. Verification and Analysis

5.1. Comparative analysis of index selection methods.
Genetic algorithm combined with Partical Least Squares Regression (GAPLS) is widely used in variable selection and index selection.

The temperature parameter prediction model of generator bearing B is taken as an example for testing. First, the SCADA system data of wind farm on March 16, 2016 solstice on April 17, 2016 is selected, and then the model established by GAPLS index selection method and the index selection method in this paper is used for testing. The test results are shown in Figure 1 and Figure 2.

![Figure 1. parameter selection base on text](image1)
![Figure 2. parameter selection base on GAPLS](image2)

From Fig.1 and 2 it is shown, using the parameter index selection method in this paper, the model residual is mainly located in the interval [-6℃,5℃] and distributed centrally, mainly around 0℃, and the model error is small. After calculation, RMSE of the data in Fig.1 is 1.8203℃, and RMSE of the
data in Fig. 2 is 2.4319°C. Therefore, the prediction accuracy of the bearing B temperature model established based on the index selection method is higher. It shows that the parameter index selection method in this paper can effectively determine the input index set of the prediction model and simplify the structure of the parameter model while ensuring the accuracy of the model.

5.2. Validation of parameter anomaly identification method
On July 31, 2016, No. 15 wind turbine failed, mainly caused by overheating of generator bearing B. In this section, the monitoring data before failure is selected to study the changes of the working condition indicators of wind turbines, and the data of SCADA system from May 3, 2016 to July 31, 2016 are selected for research and analysis. The temperature parameter analysis results of bearing B of Unit 15 are shown in Fig. 3 and Fig. 4.

Figure 3 shows the variation trend of RMSE with a calculation period of one day. Before 70 days, the amplitude of RMSE was almost all less than 4°C, and the wind turbine was in normal working state. Then, at 75 days, RMSE rapidly increased to 6.5°C, and then decreased, until it suddenly increased 1 or 2 days before the failure. The RMSE at around 75 days is significantly different from the RMSE at 1 or 2 days prior to the failure compared to the RMSE at normal operation of the unit.

Under the normal operation of the wind turbine, the residual temperature distribution of the generator bearing B will not change dramatically, that is, the entropy of the residual is basically stable and will not change too much. Figure 4 shows the change trend of the entropy value of the residual with a calculation period of 1 day. From the beginning to the 75th day, the entropy value was less than 2.5, and then increased rapidly and then decreased, until it suddenly increased before the failure. The temperature index of generator bearing B is abnormal, the predicted value of the index model deviates from the actual value to a certain extent, and the residual distribution changes appear chaotic and disorderly, which is completely inconsistent with the stable and orderly condition in normal operation. Therefore, we can see that the information entropy can reflect the abnormal state of the working condition index.

6. Conclusion
In this paper, a sub-model of performance index selection is established, and the input index set of the model is determined by judging the degree of influence on the target index. When there is a large correlation of the input index with the target index, the overall correlation index can be used to simplify the selection process of input parameters. When the correlation with the target indicator is not different, use the work condition indicator selection method in this chapter to determine the input indicator. Least Squares Support Vector Machines algorithm, Back Propagation Neural Network algorithm and Radial Basis Function Neural Network algorithm are studied. Then, the combined prediction model based on BPNN, RBFNN and LS-SVM was established. Combined with the information entropy of prediction residual, the discriminant submodel of parameter anomaly was established. The prediction error of the working condition index is reduced, and the abnormal discriminant sub-model of the working condition index is established. Finally, an example is given to verify the accuracy and effectiveness of the model.
Acknowledgment
This work is supported by Science Research Project of Education Department of Liaoning Province (JYT-L202005).

References
[1] Yang D, Li H, Hu Y, et al. Vibration Condition Monitoring System for Wind Turbine Bearings Based on Noise Suppression with Multi-point Data Fusion[J]. Renewable Energy, 2016, 92: 104-116.
[2] Tang J, Soua S, Mares C, et al. An Experimental Study of Acoustic Emission Methodology for in Service Condition Monitoring of Wind Turbine Blades[J]. Renewable Energy, 2016, 99: 170-179.
[3] Ata R. Artificial Neural Networks Applications in Wind Energy Systems: A Review[J]. Renewable Sustainable Energy Review, 2015, 49: 534-562.
[4] Schlechtingen M, Santos I F. Wind Turbine Condition Monitoring Based on SCADA Data Using Normal Behavior Models. Part 2: Application Examples[J]. Applied Soft Computer Journal, 2014, 14: 447-460.