

PAPER • OPEN ACCESS

Fault Detection of Wind Turbine Converters with Time Sequence Processing and Attention Model

To cite this article: Anqi Wang *et al* 2020 *J. Phys.: Conf. Ser.* **1659** 012059

View the [article online](#) for updates and enhancements.

You may also like

- [A self-supervised learning method for fault detection of wind turbines](#)
Shaodan Zhi and Haikuo Shen
- [Investigation of deep transfer learning for cross-turbine diagnosis of wind turbine faults](#)
Ping Xie, Xingmin Zhang, Guoqian Jiang et al.
- [Wind turbine gearbox fault prognosis using high-frequency SCADA data](#)
Ayush Verma, Donatella Zappalá, Shawn Sheng et al.

 The Electrochemical Society
Advancing solid state & electrochemical science & technology

UNITED THROUGH SCIENCE & TECHNOLOGY

248th ECS Meeting

Chicago, IL
October 12-16, 2025
Hilton Chicago



Science + Technology + YOU!

SUBMIT ABSTRACTS by March 28, 2025

[SUBMIT NOW](#)

Fault Detection of Wind Turbine Converters with Time Sequence Processing and Attention Model

Anqi Wang^{1,*}, Zheng Qian¹ and Bo Jing¹

¹School of Instrumentation and Optoelectronic Engineering, Beihang University, Haidian Dist., Beijing, China

*Email: wanganqi@buaa.edu.cn

Abstract. Fault detection methods of converters can improve the stability of wind turbines and reduce the maintenance cost. In order to achieve more accurate converter fault detection, this paper proposes an encoder-decoder structure based on the Long Short-Term Memory (LSTM) network. The LSTM network is considered to extract the sequence information of SCADA data to improve the accuracy of fault detection. Moreover, we introduce attention models into our method. Attention models can enable the network to selectively focus on the most useful variables and data during training and testing period. Specifically, we use the time attention model to calculate the importance of data at different times in the sequence, and the feature attention model to learn the importance of different variables. Experimental results on SCADA data show that our method achieves 7% -14% higher fault detection accuracy than other anomaly detection methods.

1. Introduction

In order to analyse and evaluate the health status of wind turbines (WTs) in time, the early fault detection method is used to monitor the health status of WTs, which can improve the reliability of WTs. However, according to the industry statistical results [1], as a kind of equipment with complex internal structure, the converter is one of the most vulnerable components in the WTs because it is often affected by high stress in the wind farm. In addition, the failure statistics of WTs show that [2], the total failure rate of converter in the drive system is higher than that of other components of WTs, so the failure detection of converter in wind farms is particularly important for the health of WTs.

Many approaches have been developed and reported for fault detection of converters. Available approaches can be generally classified into two categories, i.e., model-based methods and data-driven methods [3]. The model-based method is mainly based on the working principle and control theories to construct an explicit system dynamic model of the converter. [4] proposed a fault diagnostic technique of converter based on Park's vector approach. This method can visualize the fault mode of the converter by extracting the relevant electrical parameters and combining the three-phase electrical balance principle. Based on the Park's vector approach, the average current Park's vector approach for converter fault detection is proposed, which improves the accuracy compared with the previous method [5]. [6] proposed a model-based fault detection and identification (FDI) method for switching power converters using a model-based state estimator approach. In order to compare different FDI methods, [7-10] proposed model-based methods using combined observer and Kalman filter, up-down counter, Gaussian kernel support vector machine, and estimators, respectively. Besides, a multi-physics graphical model-based FDI was reported in [11]. Data-driven methods with the characteristics



of low cost and high feasibility, are currently receiving considerable attention. These methods use SCADA data for data mining to detect significant changes in the state parameters of the wind turbine. The SCADA data is mainly used for equipment fault diagnosis and health condition assessment, among which the most common vibration signals and oil monitoring signals are introduced in the literature [12-14]. Trending of SCADA parameters can reveal the development of a failure using historical data. [15] utilized SCADA data for a variety of wind turbine condition monitoring. One approach compared the temperature trends of different turbines in a particular wind farm. Different studies have shown that the changes in temperature have a certain degree of individual, which require the manual interpretation. It is not beneficial to describe the trend of data changes only by the numerical description. Therefore, there is a trend of using SCADA data for condition monitoring mentioned in [16], which is to use clustering algorithms to automate the classification of 'normal' and 'faulty' observations.

However, the existing research methods on fault detection of wind farm converters have the following problems: most researches on fault detection of converter focus on a certain time of operation, and seldom consider the information of time series. Sufficient quantities and types of data are difficult to obtain. The traditional method of converter fault detection lacks the research on the relationship between internal variables, but only considers the relationship between these variables and the operation state of wind turbine. Therefore, in order to overcome the major drawbacks of the previous fault detection methods for converters, a new method of converter fault detection is proposed in this paper. The proposed fault detection method consists of encoder-decoder network and attention model, both based on LSTM. This method not only makes full use of the time sequence information of the converter, but also only needs the data of the wind turbine in normal operation to identify the faults. In addition, in order to improve the accuracy of detection, an attention model is introduced into the proposed method. It aims to learn the importance weights of variables at the same time and data at different times in a sequence. We use the time attention model to calculate the importance of data at different times in the sequence, and the feature attention model to learn the importance of different variables for fault detection. Experimental results on SCADA data show that our method achieves 7%-14% higher fault detection accuracy than other anomaly detection methods.

2. Methodology

The encoder-decoder network based on LSTM is proposed in this paper for converter fault detection of WTs. Combined with attention model, it aims to learn the importance weights of variables at the same time and data at different times in a sequence, so as to improve the accuracy of fault detection. The complete structure is presented in Figure 1.

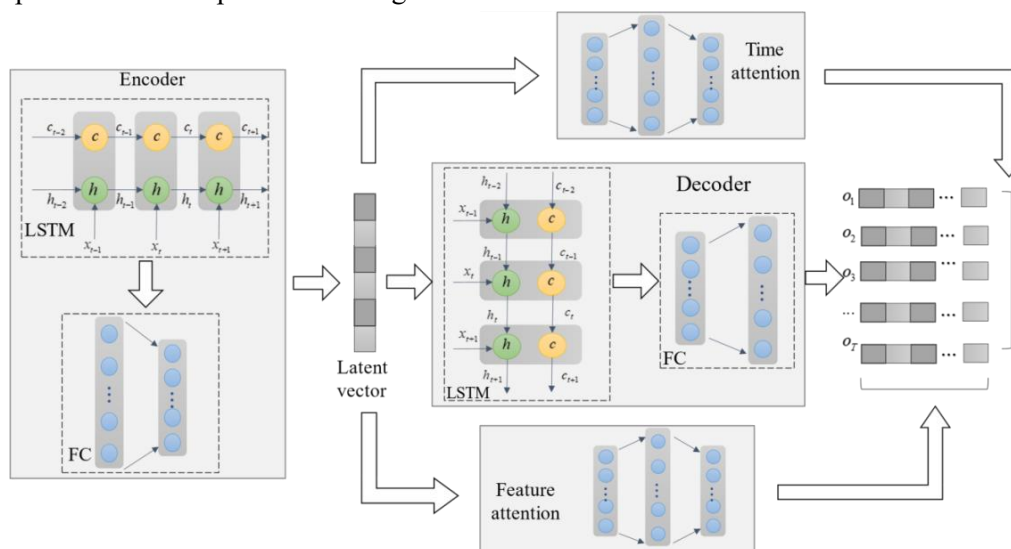


Figure 1. The overall structure of our method.

2.1 Networks Architecture

The encoder-decoder network is regarded as an effective method in the fault detection of converters of WTs. It can learn the distribution of normal data and speculate the abnormal degree of actual input to realize fault detection. However, the existing fault detection methods are usually based on MLP (multi-layer perceptron) for the design of encoder-decoder network, which is lack for the ability of learning sequential data. As the operation of the converters is a continuous process, the variation trend of the operation data in a certain period of time plays an important role in the advance detection of faults. Therefore, LSTM is used for the construction of the fault detection model of converters, which is composed of an encoder and a decoder. The detailed introduction is as follows:

2.1.1 The encoder network. The encoder consists of an LSTM network and a MLP. The LSTM can learn the information from the time sequences, and the MLP can be used for the accomplishment of encoding the information extracted by the LSTM for the next prediction. Assuming that the input time series data is represented by x_1, \dots, x_T and that the output of LSTM is represented by h_1, \dots, h_T the calculation process of LSTM can be represented by the following formula:

$$h_t = g(h_{t-1}, x_t) \quad (1)$$

In addition, we use an MLP to encode the output state of the LSTM at the last moment, and regard it as a hidden vector output by the encoder:

$$z = c_1 + V_1 * \text{ReLU}(c_2 + V_2 h_T) \quad (2)$$

Where c_1, c_2 and V_1, V_2 represent the bias and weight matrix respectively.

2.1.2 The decoder network. The decoder network is also composed of an LSTM module and a MLP module, but unlike the encoding network, the input of the decoding network is not a sequence of data, but the hidden vector z obtained by the encoder. The dimensions of the output and input data of the decoding network are the same, denoted by o_1, \dots, o_T . For the output at time t , the decoding network first calculates the hidden state (*hidden state*) h_t through the LSTM module.

$$h_t = g(h_{t-1}, z) \quad (3)$$

Afterwards, h_t is decoded by the MLP into an output o_t of the same dimension as the input data.

$$o_t = c_3 + V_3 * \text{ReLU}(c_4 + V_4 h_t) \quad (4)$$

Where c_3, c_4 and V_3, V_4 represent the bias and weight matrix respectively.

2.2 Attention Model

As an effective fault detection framework, the encoder-decoder network can learn the distribution of normal data, but during the training and testing of the encoder-decoder network, the impact of data at different times and different variables at the same time on the fault detection results is considered to be equal, without considering the dynamic impact of different variables and different data on the results. Therefore, we propose an attention model for converter fault detection, including time attention model and feature attention model, which are used to extract the importance of data at different times in the sequence and the importance of different variables to the results respectively.

Time attention model aims to learn the importance of data to the results at different times. The model takes the hidden vectors obtained by the encoding network as input and outputs the time attention vector A_t with the dimension of $1 \times T$:

$$A_t = \text{soft max}(c_t + v_t z) \quad (5)$$

Feature attention model is designed to learn the importance of different variables in the data at the same time for fault detection, so as to ensure that the network can always pay attention to the most effective variable for the result. The feature attention model takes the output hidden vector from the coding network as the input and output the feature attention vector A_f . The length of the vector is the same as the number of input variables.

$$A_f = \text{soft max}(c_f + v_f z) \quad (6)$$

2.3 Training Object

In order to enable the network model to learn the distribution of normal data and get a small residual error in the normal data and a large residual error in the fault data, we calculate the L1 loss function of the input data and the reconstructed data as a loss function during network training. In addition, we use the attention vector to weight-adjust the loss function from the time dimension and feature dimension, so that the network can focus more on typical data and typical variables. The specific form of the loss function is as follows:

$$L = \sum_t^T A_t^i \sum_j |x_i^j - o_i^j| A_f^j \quad (7)$$

Where A_t^i represents the i -th element of the time attention vector, x_i^j, o_i^j represents the j -th variable of the data at the i -th time of the input and output, and A_f^j represents the j -th element of the feature attention vector.

3. Experiments

3.1 Experimental Dataset

In order to train the model and test the effect of the model, we select the SCADA data of a wind farm in northwestern China. The wind farm contains 24 wind turbines. We select the SCADA data of wind turbines No. 1 and No. 2 in 2016 as our experimental dataset. According to the operation and maintenance records the wind turbine, we label the dataset as fault dataset and normal dataset. In addition, the normal data is divided into training data and test data according to the ratio of 7:3.

The SCADA data contains 170 operating variables, many of which are independent of converters. Therefore, we use the random forest method to pick out variables that are closely related to the state of converters. Random forest can be regarded as a simple but effective method for feature selection, by random sampling variables and features to train different classifiers. Random forest method can provide a metric of importance for each feature, and the metric can be regarded as the impact of different features on fault detection. We use the random forest method to select 22 variables for converter fault detection. After that, we normalize the data to $[-1, 1]$.

3.2 Comparison with other methods

To evaluate the effectiveness of our method, we compare it with other anomaly detection methods, including one class SVM and Isolate forest. All comparison methods are trained by the normal data from the training dataset. The results of the comparison are shown in Table 1. It can be seen from the comparison results that our method has achieved the best results in average detection accuracy. Although the Isolate forest and One class SVM methods can achieve good results in distinguishing normal data, they are far inferior to our methods in the detection of fault data, which shows that these methods have a high rate of false alarm. Figure 2 shows the residual variation of our method in the face of normal data and fault data. At this time, the wind turbine does not show any abnormality, but

our method can predict the failure of the converter about three days in advance, thereby reducing maintenance costs.

Table 1. Comparison results with other methods. The column "Normal" and "Fault" represents the accuracy of normal and fault data. The column "Average" represents the average accuracy.

Method	Accuracy		
	Normal	Fault	Average
Isolate Forest	100%	48.60%	74.3%
One Class SVM	90.76%	71.39%	81.08%
Our Method	86.47%	91.01%	88.74%

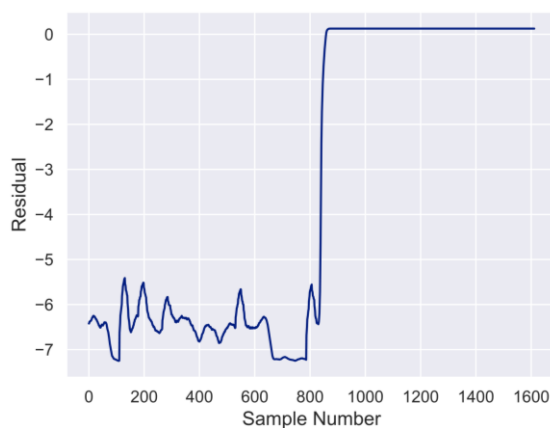


Figure 2. The Residual of our method. the wind turbine does not show any abnormality, but our method can predict the fault of the converter about three days in advance.

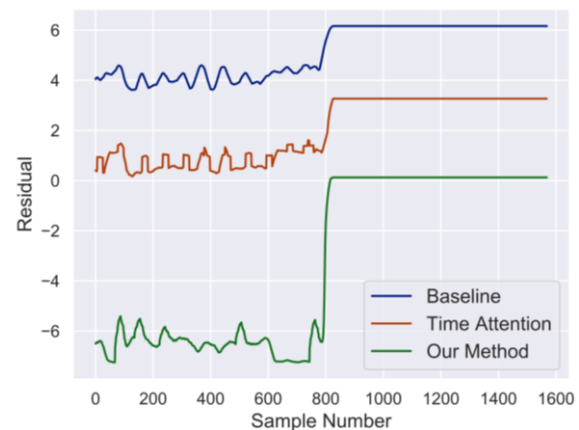


Figure 3. The ablation experiment. "Baseline" means no attention model. "Time Attention" means time attention only. "Our Method" means both attention models are used.

3.3 Ablation Analysis

In order to accurately evaluate the effect of the attention model, we design an ablation experiment, comparing the effects of using time attention and feature attention, using only time attention, and not using attention. The experimental results are shown in table 2. It can be seen that time attention and feature attention models have significantly improved the effect of fault detection. In addition, you can also see that the attention model has improved the accuracy of fault data detection more than normal data. This also shows that our method with attention can correctly focus on the most effective data and variables for fault detection, thereby improving the detection effect. Figure 3 shows the residual variation in the ablation experiment. It can be seen that our method with attention model can have a larger variation in the residual when faced with fault data.

Table 2. The ablation results of using time attention and feature attention, using only time attention, and not using attention.

Time Attention	Feature Attention	Accuracy		
		Normal	Fault	Average
×	×	85.01%	88.76%	86.89%
✓	×	85.48%	90.20%	87.84%
✓	✓	86.47%	91.01%	88.74%

4. Conclusion

In this paper, we propose a new fault detection model of wind turbine converter. We construct the encoding and decoding structure based on LSTM network, which can fully learn the timing information in the converter's SCADA data and improve the accuracy of fault detection. Moreover, we also introduce the attention model into our method. The attention model can enable the network to selectively focus on the most useful variables and data for fault detection during training and testing, thereby improving the accuracy of fault detection. Experimental results on the SCADA data show that our method can achieve better results in fault detection than other anomaly detection methods with a lower false alarm.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (No. 61573046) and Program for Changjiang Scholars and Innovative Research Team in University (No. IRT1203).

References

- [1] Qiao W, Lu D 2015 A Survey on Wind Turbine Condition Monitoring and Fault Diagnosis *IEEE Trans. Ind. Electron.* **62** 1-1
- [2] Zhao H, Li L 2017 Fault diagnosis of wind turbine bearing based on variational mode decomposition and Teager energy operator *IET Renew. Power Gen.* **11** 453-460
- [3] Zhao H, Liu H and Hu W 2018 Anomaly Detection and Fault Analysis of Wind Turbine Components Based on Deep Learning Network *Renew. Energy* **127** 825-834
- [4] Mendes A, Marques Cardoso A and Saraiva E 1998 Voltage source inverter fault diagnosis in variable speed AC drives, by Park's vector approach *Proc. Conf. on IEE* vol 456 (London: IET) pp 538-543
- [5] Mendes A, Marques Cardoso A 1999 Voltage source inverter fault diagnosis in variable speed AC drives, by the average current Park's vector approach. *Int. Conf. on IEEE* (New York: IEEE) pp 704-706
- [6] Poon, J, Jain P, Konstantakopoulos I, Spanos C, Panda S and Sanders S 2017 Model-Based Fault Detection and Identification for Switching Power Converters *IEEE Trans. Power Electron.* **32** 1419-30
- [7] Chen W, Ding S, Haghani A, Naik A, Khan A and Yin S 2011 Observer-based FDI Schemes for Wind Turbine Benchmark *Ifac Proc.* vol 44 (Amsterdam: Elsevier) pp 7073-78
- [8] Ozdemir A, Seiler P and Balas G 2011 Wind turbine fault detection using counter-based residual thresholding *Ifac Proc.* vol 44 (Amsterdam: Elsevier) pp 8289-94
- [9] Laouti N, Sheibat-Othman N and Othman S 2011 Support vector machines for fault detection in wind turbines *Ifac Proc.* vol 44 (Amsterdam: Elsevier) pp 7067-72
- [10] Zhang X, Zhang Q, Zhao S, Ferrari R, Polycarpou M and Parisini T 2011 Fault detection and isolation of the wind turbine benchmark: An estimation-based approach *Ifac Proc.* (Amsterdam: Elsevier) vol 44 pp 8295-300
- [11] Mojallal A and Lotfifard S 2018 Multi-physics graphical model based fault detection and isolation in wind turbines *IEEE Trans. Smart Grid* **9** 5599-612
- [12] Márquez F, Tobias A, Pérez J and Papaelias M 2012 Condition monitoring of wind turbines Techniques and methods *Renew. Energy* **46** 169-178
- [13] Qian H, Jiang D, Hong L and Ding Y 2008 Application of wavelet neural networks on vibration fault diagnosis for wind turbine gearbox *Int. Symp. on Advances in Neural Networks* vol 5264 (Berlin: Springer) pp 313-320
- [14] Zhang J 2011 Extraction of rolling bearing fault feature based on time-wavelet energy spectrum *Journal of Mechanical Engineering* **47** 44
- [15] Wilkinson M., Harman K., Van D, Herman K 2014 Comparison of methods for wind turbine condition monitoring with SCADA data *IET Renew. Power Gener.* **8** 390-397
- [16] Tautz J, Watson S 2017 Using SCADA data for wind turbine condition monitoring- a review. *Renew. Power Gener. IET* **11** 382-39