Fault Prognosis of Wind Turbine Generator Using SCADA Data

Yingying Zhao[†], Dongsheng Li[‡], Ao Dong[†], Jiajia Lin[†], Dahai Kang[¶], Li Shang[†]
[†] Tongji University, Shanghai 201804 P.R.China
[‡] IBM Research - China, Shanghai 201203 P.R. China
[¶] Concord New Energy Group Limited - China, Beijing 100048 P.R. China

***Abstract*—Accurate prognosis of wind turbine generator failures is essential for reducing operation and maintenance costs in wind farms. Existing methods rely on expensive, purpose-built condition monitoring systems to conduct diagnosis and prognosis of wind turbine generator failures. In this paper, we present a prognosis method to predict the remaining useful life (RUL) of generators, which requires no additional hardware support beyond widely adopted SCADA system. This work first introduces a notion, Anomaly Operation Index (AOI), to quantitatively measure wind turbine performance degradation in runtime. It then presents a data-driven wind turbine anomaly detection method and a time series analysis method to predict the wind turbine generator RUL. Experimental study on real-world wind farm data demonstrates that the proposed methods can achieve accurate prediction of wind turbine generator RUL and provide sufficient lead time for scheduling maintenance and repair.**

I. Introduction

With the rapid growth of wind power installation capacity, wind farms are faced with the ever-increasing challenges of high operation and maintenance (O&M) cost [1], [2]. To minimize O&M cost, wind turbine prognosis becomes increasingly important to ensure sufficient lead time for scheduling repair or acquiring replacement components before failures occur [3]. The failure mechanisms of wind turbines have been extensively studied in the past. Generator is one of the most expensive components in wind turbines, and the failures of which response for long downtime and high maintenance costs. Recent studies have shown that the major source of failures has centered on generators [4]. Generator-related failures bring significant downtime of wind turbines [5], which is one of the highest downtimes caused among all wind turbine failures. To this end, the goal of this work is to develop effective prognostic methods for wind turbine generator to reduce O&M cost.

Recent research work has tackled the prognosis problem for wind turbine generator. High cost is the primary limiting factor to existing solutions, as expensive, purpose-built condition monitoring system (CMS) is required [5], [6], [7]. For instance, Yang et al. [6] developed a fault diagnosis method for the main shaft in the generator sub-system by analyzing vibration signals provided by CMS [6]. Due to the high installation and maintenance cost, to date, the adoption rate of CMS has been low in existing wind farms. As a result, adopting CMS-based generator prognosis method is challenging. Compared to CMS, the Supervisory Control And Data Acquisition (SCADA) system is considered as a standard installation equipment and has been widely installed in existing wind farms. SCADA monitors the run-time operation condition of wind turbine, such as temperature, speed, and power [6]. Such information may be leveraged to support generator prognosis. On the other hand, compared to CMS, the information collected by SCADA is limited, making SCADA-based failure prognosis challenging. For instance, generator vibration information, a critical measure for generator mechanical failure detection, is not available in SCADA. Recently, researchers have started investigating SCADA-based prognosis methods for wind turbine gearbox [8], [9], [10]. However, there is few fault prognostic work focusing on generator fault prediction.

This paper aims to tackle the remaining useful life (RUL) prediction problem for wind turbine generator using SCADA data. We first introduce a notion, Anomaly Operating Index (AOI), to quantitatively measure wind turbine historical performance and predict wind turbine future performance based on a statistical model. Next, we propose a data-driven anomaly detection technique to capture the $anomaly$ data change in SCADA system associated with generator fault, i.e., capture the anomaly of runtime AOI. We then propose a prognosis method to predict the RUL of wind turbine generator. This method consists of an autoregressive integrated moving average (ARIMA) based statistical model to conduct online prognostic, and a time series analysis-based RUL estimation method to provide accurate RUL prediction. Experimental study using real world wind farm data (33 wind turbines over 17 months) demonstrates that the proposed methods can perform accurate prognosis of wind turbine generator RUL with sufficient lead time for maintenance and repairing. In summary, this work makes the following contributions.

1. A new notion, namely AOI, is introduced to quantitatively measure wind turbine performance degradation in runtime.
2. A data-driven analysis method is proposed to support anomaly detection, which adopts a clustering method (DB-Scan) to distinguish $anomaly$ data and $normal$ data from unlabeled historical SCADA data, and a classification method (SVM) to classify $anomaly\ AOI$ and $normal\ AOI$ in runtime.
3. An ARIMA model is adopted to analyze realtime AOIs, which can estimate wind turbine RUL and predict future AOI of wind turbines.
4. Evaluation on a wind farm with 33 wind turbines demonstrates that the proposed method can provide sufficient lead time to wind farm operators to schedule maintenance before generator failures occur.

The rest of this paper is organized as follows. Section II surveys the related work. Section III discusses the challenges in

978-1-5090-3270-9/16/$31.00 ©2016 IEEE

SCADA-based generator fault prognosis. Section IV presents the proposed anomaly detection method and prognostic model. Experimental results are presented in Section V. We conclude this work in Section VI.

II. Related Work

Existing wind turbine generator fault prognosis methods can be categorized into two classes: model-based approaches and data-driven approaches [3], [11].

Model-based approaches typically use either physical specific or explicit mathematical model to reason the causality of the machine inherent operational principle. The residuals between the model outputs and actual signals are used to suggest that a failure is present or not [12]. There are various model-based approaches focusing on generators failure diagnosis and prognosis [13]. A reliable and accurate model is the key in model-based approaches. The proposed approaches are beneficial for detecting specific generator components fault. However, one main limitation of the model-based approach is that it may not be feasible for complex systems [12]. Wind turbines principle operation is an uninterrupted process and failures may be caused by a series of components co-operation. However, the wind turbines co-operation is neglected in the current fault prediction methods. On the other hand, these model-based approaches typically require machine features that can only be collected by high frequency sensors or advanced interpreter techniques, which cannot be supported by the SCADA system.

The data-driven method utilizes statistical approaches and artificial intelligence (AI) approaches, which learn models directly from the data. Support vector machine (SVM) [3] and hybrid SVM-Bayesian network (BN) [14] are used to realize an optimized classification of machine states. Many approaches have used artificial neural networks (ANNs) [10], [15] to model the system, such as dynamic wavelet neural networks (DWNN) [16] and recurrent neural networks (RNNs) [17]. These approaches work by representing the linear or nonlinear correlation among different features to determine the association between system inputs and outputs. The residuals between the estimated values and the actual features values are used to indicate the *anomaly* change so as to realize an incipient fault prognosis. To develop an accurate model, it requires large amount of data and numerous examples [10], [15]. Till today, there is few method targeting generator fault prognosis via SCADA data.

III. Challenges

The primary challenge to tackle generators fault prognosis has always been the choice of a meaningful and actionable data analysis method to capture the hidden relationship among low frequency signals collected by the SCADA system so as to recognize wind turbines anomaly behaviors. Compared with purpose-built CMS that collects high frequency data such as acceleration and acoustic emission, SCADA system is not initially designed for condition monitoring purpose so that it collects limited and low-frequency signals, e.g. power output and temperature. Therefore, it is challenging to capture incipient faults from the raw SCADA data without effective data analysis methods.

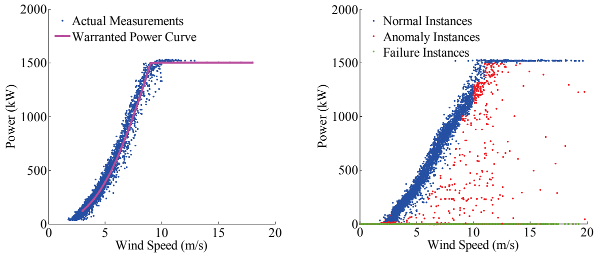

Fig. 1. Power curves of a wind turbine in normal condition (left) and a turbine with an anomaly condition and an occurred failure (right).

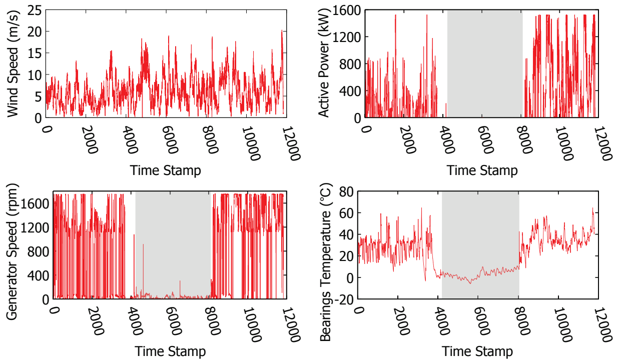

Fig. 2. Examples of data analyzing challenges on time series basis.

Nevertheless, SCADA system provides information which can potentially reveal wind turbines health condition. In wind farms, power curve has been widely used to diagnose wind turbine faults. More specifically, the correlation between wind speed and active power output collected by SCADA is compared against the power specification provided by the equipment manufacturer. Under-performing turbines can then be identified. Figure 1 shows a wind turbine that operates properly and one that experiences a fault. On the other hand, using power curve alone, it is difficult to identify the true cause, or develop fault progression trajectory and perform fault prognosis. Therefore, other features, collected by SCADA, associated with the generator fault process must be included in the modeling method, e.g., temperature profile and speed information of individual generator components. Indeed, Qiu et al. [18] showed that generator faults can change heat transfer and result in temperature changes. Yang et al. [19] suggested that mis-correlation between generator speed and generator active power will indicate a fault in generator.

Noise imposes further challenges to SCADA-based generator fault prognosis. Figure 2 shows exemplary time series data collected by SCADA, highlighting the data analysis challenges due to high data noise. In this figure, the gray zone shows the occurrence of a generator failure. The generator data collected by SCADA during the fault developing phase (before the gray zone) or the normal operation after repair and maintenance (after the gray zone) exhibits high fluctuation, making fault prognosis analysis challenging.

IV. Prognostics Approach

This section first formulates the problem. Then, two layers of the proposed prognostic, i.e., data layer and decision support layer, are presented in detail. The data layer consists of the following key steps: 1) data preprocessing; 2) feature selection from SCADA system; and 3) feature reduction. The decision support layer consists of three key methods: 1) a clustering-based method to distinguish $normal$ data and $anomaly$ data from unlabeled historical SCADA data; 2) a classification method, which is trained based on the previous clustering results, to classify $normal$ data and $anomaly$ data in runtime; 3) an ARIMA-based method to estimate runtime RUL of wind turbine generator.

A. Problem Formulation

The proposed work captures the $anomaly$ change based on a simple intuition that $normal$ data instances exhibit similar characteristics and $anomaly$ data instances exhibit different characteristics from the $normal$ instances. To this end, cluster analysis can be adopted to classify instances into different categories on the feature similarity. In addition, in a wind farm, most of the generators function properly most of the time, and failures rarely occur. Therefore, in this study, clustering-based anomaly detection technique is applied under the following assumptions:

1. Normal data instances belong to the largest and densest cluster.
2. Anomaly instances either belong to the small or sparse clusters, or do not belong to any clusters.

Therefore, to develop an accurate model for a specific fault, we can measure the change of runtime $anomaly$ instances compared to the total instances to illustrate wind turbines performance change. Higher fraction of $anomaly$ instances means lower wind turbines performance. Thus, by analyzing the runtime fraction of $anomaly$ instances, we can develop the growth trajectory of an incipient fault.

B. The Data Layer

Preprocessing. The data collected by a SCADA system usually contains errors caused by sensors and malfunctions of the data collection system. These errors may be missing values, out-of-range values, dirty values, and etc. In this work, these errors are filtered out prior to data analysis using data preprocessing.

Feature Selection. Since not all SCADA data are related to generator faults, we first need to identify the subset of the features collected by SCADA system reflect generator operation performance. More specifically, the features that allow us to analyze generator faults fall into three categories: the condition parameters, the health parameters, and the performance parameters. The condition parameters contain wind speed and ambient temperatures, which can determine the power output of wind turbines. The health parameters contain main bearing temperature, low speed shaft temperature, high speed shaft temperature, and gearbox oil temperature, which can help analyze the health condition of wind turbines. The performance parameters contain rotor speed, generator speed and active power, which measure wind turbines operational performance.

Feature Reduction. The above features are selected based on insights from domain experts, but some of them may not be strongly correlated with generator faults. Therefore, to maximize the clustering quality, principle component analysis (PCA) technique proposed by Wold et al. [20] is adopted to find the optimal subset of features from all relevant features. Since PCA-based feature reduction is not the key contribution of this work, the details can be referred to [20].

C. Clustering-based Anomaly Detection

The optimal subset of features are adopted by density-based spatial clustering of applications with noise (DBScan) algorithm to distinguish normal data and anomaly data in the historical data from the SCADA systems. Note that, various cluster algorithms can be adopted here to perform anomaly detection, and we choose DBScan algorithm due to the following reasons:

- DBScan is based on density estimation. According to the previous two assumptions, we can find the largest and densest cluster for adequate time series SCADA data;
- the cluster shape can be arbitrary in DBScan. This is suitable for the detection purpose, so that we do not need to predefine the parameters for clusterings, e.g., number of clusters or sizes of each cluster.

To obtain the optimal clustering results, we adopt an internal validation measure [21], S_Dbw, to assist model selection, i.e., to select the best clustering results from a variety of different runs. This measure has two criteria:

- compactness $Dens\_bw$, which measures how closely related the objects in a cluster are; and
- separation $Scat$, which measures how distinct a cluster is from other clusters. S_Dbw is defined as follows:

$$S\_Dbw{=}Scat(c){+}Dens\_bw(c) \tag{1}$$

where c is the number of the clusters. $Scat(c)$ and $Dens\_bw(c)$ are defined in Eq. (2) and Eq. (3), respectively.

$$Scat(c){=}\tfrac{1}{c}\textstyle\sum_{i=1}^{c}\tfrac{\|\sigma(C_i)\|}{\|\sigma(D)\|} \tag{2}$$

$$Dens\_bw(c){=}\tfrac{1}{c(c-1)}\textstyle\sum_{i,j=1,j\neq i}^{c}\tfrac{dens(u_{ij})}{max\{dens(v_i),dens(v_j)\}} \tag{3}$$

where $\sigma(C_i)$ is variance of the i-th cluster C_i, $\sigma(D)$ is the variance of dataset D, v_i, v_j are centers of cluster c_i, c_j, respectively, and u_{ij} is the middle point of the line segment formed by v_i and v_j. $dens(u)$ is defined as follows:

$$dens(u){=}\textstyle\sum_{l=1}^{n_{ij}}f(z_l,u) \tag{4}$$

where n_{ij} is the number of tuples that belong to the clusters c_i and c_j. $f(x,u)$ is 1 when the distance between x and u is larger than the average standard deviation of clusters, and 0 otherwise.

The clusters from DBScan fall into two categories: normal set (NS) and anomaly set (AS). NS is the largest and densest cluster, which consists of the normal instances. And AS denotes all the other clusters and noise points, which consist of the anomaly instances. Based on the clustering results, Anomaly Operation Index ($AOI(s_q)$) is proposed to measure a wind turbine's performance based on its historical data sequence s_q:

$$AOI(s_q){=}1{-}m_{qi}/m_q \tag{5}$$

where $m_q = |NS \cup AS|$ and $m_{qi} = |NS|$. $AOI(s_q)$ measures the proportion of anomaly instances to total instances for a test sequence s_q. Therefore, AOI could be used to indicate the wind turbine performance for a time period. Larger AOI indicates higher possibility of wind turbines performance degradation in a given time period. In addition, we can also predict wind turbine's future performance by predicting its future AOI. Therefore, contiguous AOIs tendency corresponding to time series data should be considered to describe wind turbine performance trajectory. Since wind turbine performance degradation is a stochastic and accumulated process, a specific $AOI(s_q)$ for a sequence s_q may exhibit different degree of fluctuation according to the dimensions of a test sequence, so that it cannot describe wind turbine performance degradation accurately. Thus, we use Piecewise Aggregate Approximation (PAA) method to reduce the dimension of the time series s aiming to smooth the fluctuation of $AOI(s_q)$.

A sequence $s = \{s_1, \ldots, s_n\}$ of dimension n can be represented in a w-dimensional space $\bar{s} = \{\bar{s}_1, \ldots, \bar{s}_w\}$ using PAA representation. The i^{th} element of $\bar{s}$ is calculated by the following equation:

$$\bar{s}_i = \frac{1}{\alpha} \sum_{j=\alpha \cdot (i-1)+1}^{\alpha \cdot i} s_j \tag{6}$$

where $\alpha = \lceil n/w \rceil$ is the width of smoothing window. Larger α indicates longer accumulation process, so that we can choosing appropriate α to help us accurately measure the performance of different kinds of wind turbines.

D. ARIMA-based Prognostic

Contiguous AOIs can be described as time series data. Besides, seasonal features and correlation features, e.g., periodic variation that recur within wind turbines different status and future trend should also be analyzed together with AOI. To this end, an ARIMA model [22] is adopted here, which can deal with time series data and forecast the future data points by analyzing the seasonal data and correlation data. ARIMA model can be generally described as follows:

$$\varphi_p \cdot (1-B)^d \cdot AOI_t = \theta_0 + \theta_q \cdot (B) \cdot \epsilon_t \tag{7}$$

where AOI_t ($t \in \{1, \ldots, n\}$) is the anomaly operation index for a sequence s, ϵ_t is white noise sequence, B is backward shift operator, φ_p is autoregressive operator, θ_q is moving average operator, p, d, q are the order of the autoregressive, the order of difference, the order of the moving-average, respectively. ARIMA model can be simply described as: ARIMA (p, d, q). We can use ARIMA model to understand the trends of AOI so as to forecast its future tendency. The three iterative steps of using ARIMA model to achieve wind turbine AOI prognostic are as follows:

1. Model Identification. Stationarity is a necessary condition in building an ARIMA model. At first, difference data to make data stationary on mean and variance. The difference process can be described as: $\bigtriangledown^d AOI_t = AOI_t - AOI_{t-1}$. Then, identification of best fit ARIMA model. Box et al. [23] proposed to use autocorrelation function and the partial autocorrelation function to identify patterns in the above stationary data.
2. Parameter Estimation. This step is an optimization procedure, with intending of minimizing the overall errors.
3. Diagnostic Checking. Check the tentative model is adequate or not.

The above three steps will be iterated until the model is adequate. Then, AOIs will be forecasted using the best fit ARIMA model.

E. RUL Prediction

After establishing an ARIMA model to analyze future AOIs of wind turbines, wind turbine generator RUL prediction can be performed. Two parameters are adopted to predict generator RUL in wind turbine in this work: 1) boundary value β ranging from 1 (normal) to 0 (anomaly); and 2) historical remaining useful life τ_i at a certain time i. The RUL prediction of wind turbine generator is described in Eq. (8).

$$RUL = \sum_{i=1}^{w} x_i \cdot \tau_i \tag{8}$$

where x_i is 1 under the condition of $\bar{s}_i \geqslant \beta$, and 0 otherwise. $\bar{s}$ is the smoothed history sequence. Obtaining the value of β can be viewed as a classification process, and many different kinds of classification methods can be used here. Support vector machine (SVM) [24], due to its excellent generalization ability, is a popular binary classification method in the machine learning community. In this work, we employ SVM to learn the boundary (optimal β) between normal instances and anomaly instances.

V. Experiments and Results

This section evaluates the proposed generator fault prognosis work using real-world wind farm data.

A. Dataset Description and Evaluation Metrics

In order to evaluate and test the proposed prognosis approach, SCADA data collected from a wind farm (33 wind turbines) located in north west China for a period of 17 months is used. More specifically, three major generator-related failures occurred on different wind turbines are identified and used to evaluate the overall prognosis performance. In addition, a case study is given to illustrate the process of model training and test. Since ten-minute sampling interval has been widely adopted in existing SCADA-based fault prognosis literature, we re-sampled the data using the same ten-minute interval.

To evaluate the accuracy of RUL prediction, The prediction accuracy for a certain prediction step t is calculated using Eq. (9).

$$ACC = 1 - \frac{|\tau_t^* - \tau_t|}{\tau_t} \tag{9}$$

where τ_t^* is estimated RUL at t-th day ahead and τ_t is actual RUL at t-th day ahead. In addition, the mean relative error (MRE) for predicting generator RUL with n days ahead is defined using Eq. (10).

$$MRE = \sum_{i=1}^{n} \frac{|\tau_i^* - \tau_i|}{n \cdot \tau_i} \tag{10}$$

where τ_i^* is estimated RUL at i-th day ahead and τ_i is actual RUL at i-th day ahead.

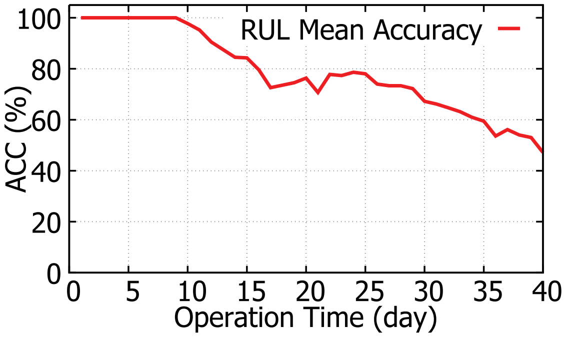

Fig. 3. Average prediction accuracy of RUL for three wind turbines.

TABLE I. PREDICTION STEPS FOR ANOMALY PROGNOSTICS LEVEL.

anomaly level	prediction accuracy (%)	prediction step (day)
suggestion	[50, 60]	39
trival	(60, 70]	34
minor	(70, 80]	29
major	(80, 90]	16
crash	(90,100]	12

B. Overall Performance

The average prediction accuracy of RUL for the three wind turbines experiencing generator failures during the testing period is provided in Figure 3. As shown in Figure 3, there is a general trend that the average prediction accuracy improves as the required repair and maintenance lead time decreases. Although there is a slightly fluctuation cased by stochastic of wind turbines operation in the curve, the overall trend is clear. More specifically, the average prediction accuracy is higher than 90% if prediction if the required repair and maintenance leadtime is 12 days ahead.

In practice, sufficient lead time (10-30 days) to schedule maintenance activities before a generator failure occurs will be sufficient. Generally, if failure occurs in a wind farm, two actions will be taken: 1) repair components or 2) replace components. Time spent on repairing components is uncertain, which is based on the failure types. If there is a need of replacing a component, such as replacing generator, purchase of the component usually takes about 20 days or 30 days, and installation and debug may take 1 or 2 days.

According to different prediction accuracy and lead time for plan maintenance, we divided anomaly levels into five categories: suggestion, trivial, minor, major and crash as shown in Table I. Thus, we can provide operators with different anomaly levels to warn a potential failure in the near future, it will be more flexible and efficient for operators to plan maintenance activities.

C. A Case Study

1) Fault Detection Based on AOI: Figure 4 shows how daily AOI values vary for wind turbine 28 in the wind farm over a period of 180 days, in which an incipient failure of generator is successfully prognosed. As illustrated in this case study, a failure occurred from the 117-th day to the 144-th day, which is colored as red in Figure 4. When the failure occurs, AOI values are almost all 1, which reflects the fact that the wind turbine stopped operating over that period of time.

In Figure 4, ARIMA model is adopted to predict the daily AOIs, and the difference order d is set to 1, which means that daily AOIs are stationary after first order difference. If we exclude the AOI values during the time when failure occurs, the rest of AOI values can be divided into two categories: 1) normal AOI values (the grey parts in Figure 4) and 2) anomaly AOI values before failure occurs (the yellow part in Figure 4). Moreover, the patterns of the two kinds of AOI values are quite different, which means that we can distinguish the two kinds of AOI values by their means and variances.

For the case study of wind turbine No. 28 in Figure 4, the gray range denotes wind turbine operations in normal conditions and the yellow range denotes wind turbine performance degradation. In this case study, wind turbines performance degradation has lasted for about 44 days until the failure occurs, and similar phenomenon is observed on other wind turbines. Therefore, we use 44 days ahead prognoses as a baseline to predict a failure.

2) Fault Prognostics: According to the analysis in section IV-A, wind turbine performance degradation is a stochastic process and this process will last for a certain time period before failure occurs. In the above case study, we choose 44 days as a baseline to predict an impending failure. In addition, the width of time smooth window α and the boundary from normal to anomaly β will affect the prediction accuracy, so we analyze how α and β values will affect the prediction accuracy in this section.

Impact on α. Theoretically, larger α will lead to lower MRE, because AOIs will exhibit lower fluctuation, and it is easier for us to distinguish the normal AOIs and anomaly AOIs. As shown in Figure 5, when the prediction is made 44 days ahead, with same β values, we can see that MRE increases with α decreases and the lowest MRE occurs with the largest α. In fact, however, we need lower α to implement accurate prediction because higher α means higher prediction interval and higher possibility of missing some fault events. Actually, when β is larger than 0.4 and α is equal or larger than 2, the MRE is of little difference. We could use $\alpha = 2$ to conduct the later experiments.

Impact on β. β stands for the boundary from normal to anomaly and ranges from 0 to 1. Theoretically, the closer to the boundary, the lower the MRE. Therefore, if β increases from 0 to 1, the MRE will decrease at first. But once β exceeds a critical value, MRE will increase as β increases. That is, there exists an optimum β to achieve a minimum MRE. In this work, the optimal β is calculated by SVM classification method. As shown in Figure 5, the optimal β is around 0.41 and we will use $\beta = 0.41$ in the following experiments.

3) RUL Prediction: In the prediction of the wind turbine generator RUL, the optimal α, β and prediction steps are kept the same for both model training and testing. Figure 6 shows the result of comparison between the actual RUL and estimated RUL on training set, the mean relative error of which is approximately 0.27.

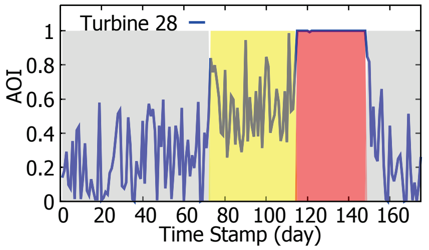

Fig. 4. A case study: failure diagnosis and prognosis for wind turbine 28.

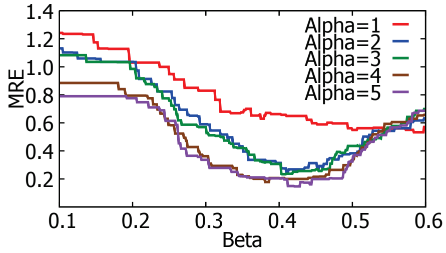

Fig. 5. Performance trends of the proposed method with different parameters.

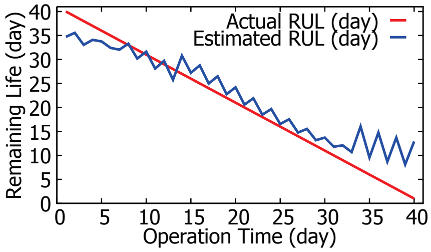

Fig. 6. Comparison of actual RUL and estimated RUL for wind turbine 28.

VI. Conclusions

This paper proposes a wind turbine generator fault prognostic approach based on anomaly detection technique. Through prior analysis of historical failure data, a notion AOI is introduced to evaluate the wind turbine historical performance, which is then used for performance prediction. By analyzing the existing data mining and machine learning techniques, DBScan algorithm is adopted to conduct the clustering to obtain $normal$ data and $anomaly$ data from unlabeled historical wind farm data, and SVM is adopted to distinguish $normal$ data and $anomaly$ data in runtime. The proposed approach is evaluated using wind turbine generator SCADA data collected from a wind farm. The experimental results indicate that the proposed method can provide effective estimation for wind turbine generator RUL via SCADA system. More importantly, the experimental results also show that the proposed method can provide wind turbine operators with sufficient lead time to schedule a repair or replace plan before generator failure occurs, which will reduce O&M cost to a large extend.

Acknowledgment

This work was supported in part by National Natural Science Foundation of China under Grant No. 61233016 and the International Exchange Program for Graduate Students, Tongji University.

References

[1] D. S. Li, F. L. Lu, Q. Lv, and L. Shang, “Lifetime cost optimized wind power control using hybrid energy storage system”, *in North American Power Symposium (NAPS)*, 2013, pp. 1-6.

[2] Y. Qiu, Y. Feng, P. Tavner, P. Richardson, G. Erdos, and B. Chen, “Wind turbine SCADA alarm analysis for improving reliability”, *Wind Energy*, vol. 15, no. 8, pp. 951-966, 2012.

[3] H. E. Kim, A. C. C. Tan, J. Mathew, and B. K. Choi, “Bearing fault prognosis based on health state probability estimation”, *Expert Systems with Applications*, vol. 39, no. 5, pp. 5200-5213, 2012.

[4] B. Lu, Y. Li, X. Wu, and Z. Yang, “A review of recent advances in wind turbine condition monitoring and fault diagnosis”, *in Power Electronics and Machines in Wind Applications (PEMWA)*, 2009, pp. 1-7.

[5] J., Ribrant, and L. Bertling, “Survey of failures in wind power systems with focus on Swedish wind power plants during 1997-2005”, *in Power Engineering Society General Meeting*, 2007, pp.1-8.

[6] W. Yang, P. J. Tavner, C. J. Crabtree, Y. Feng, and Y. Qiu, “Wind turbine condition monitoring: technical and commercial challenges”, *Wind Energy*, vol. 17, no. 5, pp. 673-693, 2014.

[7] I. P. Girsang, J. S. Dhupia, E. Muljadi, M. Singh, and L. Y. Pao, “Gearbox and drivetrain models to study dynamic effects of modern wind turbines”, *IEEE Transactions on Industry Applications*, vol. 50, no. 6, pp. 3777-3786, 2014.

[8] K. Kim, G. Parthasarathy, O. Uluyol, W. Foslien, S. Sheng, and P. Fleming, “Use of SCADA data for failure detection in wind turbines”, *in American Society of Mechanical Engineers (ASME) 2011 5th International Conference on Energy Sustainability*, 2011, pp. 2071-2079.

[9] A. Kusiak, and W. Li, “The prediction and diagnosis of wind turbine faults”, *Renewable Energy*, vol. 36, no. 1, pp. 16-23, 2011.

[10] A. Zaher, S. D. J. McArthur, D. G. Infield, and Y. Patel, “Online wind turbine fault detection through automated SCADA data analysis”, *Wind Energy*, vol. 12, no. 6, pp. 574-593, 2009.

[11] M. Schwabacher, “A survey of data-driven prognostics”, *in Proceedings of the AIAA Infotech@ Aerospace Conference*, 2005, pp. 1-5.

[12] A. K. S. Jardine, D. Lin, and D. Banjevic, “A review on machinery diagnostics and prognostics implementing condition-based maintenance”, *Mechanical systems and signal processing*, vol. 20, no. 7, pp. 1483-1510, 2006.

[13] K. Rothenhagen, and F. W. Fuchs, “Doubly fed induction generator model-based sensor fault detection and control loop reconfiguration”, *IEEE Transactions on Industrial Electronics*, vol. 56, no. 10, pp. 4229-4238, 2009.

[14] R. Ramesh, M. A. Mannan, A. N. Poo, and C. Lucas, “Thermal error measurement and modeling in machine tools. Part II. Hybrid Bayesian Network-support vector machine model”, *International Journal of Machine Tools and Manufacture*, vol. 43, no. 4, pp. 405-419, 2003.

[15] M. Schlechtingen, and I. F. Santos, “Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection”, *Mechanical systems and signal processing*, vol. 25, no. 5, pp. 849-1875, 2011.

[16] P. Wang, and G. Vachtsevanos, “Fault prognostics using dynamic wavelet neural networks”, *AI EDAM*, vol. 15, no. 4, pp. 349-365, 2001.

[17] R. C. M. Yam, P. W. Tse, L. Li, and P. Tu, “Intelligent predictive decision support system for condition-based maintenance”, *The International Journal of Advanced Manufacturing Technology*, vol. 17, no. 5, pp. 383-391, 2001.

[18] Y. Qiu, Y. Feng, J. Sun, W. Zhang, and D. Infield, “Applying thermophysics for wind turbine drivetrain fault diagnosis using SCADA data”, *IET Renewable Power Generation*, 2016.

[19] W. Yang, R. Court, and J. Jiang, “Wind turbine condition monitoring by the approach of SCADA data analysis”, *Renewable Energy*, vol. 53, pp. 365-376, 2013.

[20] S. Wold, K. Esbensen, and P. Geladi. “Principal component analysis”, *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37-52, 1987.

[21] M. Halkidi, and M. Vazirgiannis, “Clustering validity assessment: Finding the optimal partitioning of a data set”, *in Proceedings IEEE International Conference on Data Mining (ICDM)*, 2001, pp. 187-194.

[22] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen. “ARIMA-based time series model of stochastic wind power generation”, *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 667-676, 2010.

[23] G. E. P. Box, and G. Jenkins. “Time Series Analysis, Forecasting and Control”, Holden-Day, San Francisco, CA, 1970.

[24] C. J. Burges. “A tutorial on support vector machines for pattern recognition”, *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121-167, 1998.