This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2016.2621135, IEEE Transactions on Smart Grid

Wind Turbine Blade Breakage Monitoring with Deep Autoencoders

Long Wang, *Student Member, IEEE*, Zijun Zhang, *Member, IEEE*, Jia Xu, and Ruihua Liu

***Abstract*--Monitoring wind turbine blade breakages based on SCADA data is investigated in this research. A preliminary data analysis is performed to demonstrate that existing SCADA features are unable to present irregular patterns prior to occurrences of blade breakages. A Deep Autoencoder (DA) model is introduced to derive an indicator of impending blade breakages, the reconstruction error (RE), from SCADA data. The DA model is a neural network of multiple hidden layers organized symmetrically. In training DA models, the restricted Boltzmann Machine (RBM) is applied to initialize weights and biases. The back-propagation method is subsequently employed to further optimize the network structure. Through examining SCADA data, we observe that the trend of RE will shift by the blade breakage. To effectively detect RE shifts through online monitoring, the exponentially weighted moving average (EWMA) control chart is deployed. The effectiveness of the proposed monitoring approach is validated by blade breakage cases collected from wind farms located in China. The computational results prove the capability of the proposed monitoring approach in identifying impending blade breakages.**

***Index Terms*--blade breakages, condition monitoring, deep autoencoders, statistical process control, wind turbine**

I. Introduction

COMPARED with traditional power generators, wind turbines are distributed large-scale power systems exposing to variable weather conditions and harsh environments. Hence, wind turbines suffer from frequent malfunctions and failures which induce significant costs. The development of advanced condition monitoring systems (CMS) which could continuously monitor wind turbine conditions and identify impending failures has attracted great interests of the wind energy industry. The condition monitoring and fault diagnosis of wind turbine components including the drive train [1], generator bearing [2], and gearbox [3] with physics-based and data-driven approaches have been reported. Rotor blades are major components of wind turbines and the frequency of their failures has increased due to the aging of wind farms [4]. Blade failures, such as blade breakages, can lead to catastrophic consequences including the significant capital loss [5], unscheduled downtime [6], and environmental hazards [7, 8]. Therefore, wind turbine blade monitoring methods are highly valuable and desire research attentions.

Previous studies focused on monitoring wind turbine blade conditions with signals measured by additionally installed sensors. Sutherland *et al.* [9] presented an early study of monitoring blade conditions based on stress distributions measured by 14 strain gauges. However, industrial application of the strain gauge was constrained due to its low durability and sensitivity. Optical fiber sensors and acoustic emission (AE) sensors were also utilized to detect blade deteriorations. Sorensen *et al.* [10] developed micro-bending optical fiber sensors for detecting cracks in adhesive joint of blades. Dutton [11] applied AE sensors to detect vibrations caused by blade damage propagations. Recent studies [12, 13] deployed accelerometers for monitoring the blade condition and examining their effectiveness through experiments. More advanced sensors including macro-fiber composite (MFC) sensors [14], Fiber Bragg grating (FBG) sensors [15] and scanning laser Doppler vibrometer (SLDV) sensors [16] have also been applied into studies of blade condition monitoring. Besides applications of advanced sensors, signal processing algorithms were introduced to advance monitoring functions. Based on FBG sensors, Lee *et al.* [15] developed a novel signal transforming algorithm to process the 3-blade moment signals for monitoring blade conditions. Karabayir *et al.* [17] proposed a signal processing model for Doppler radars and tested the model via simulations. Although the wind turbine blade condition monitoring methods based on extra sensors have been studied, their applications are limited due to following drawbacks. First, additional sensors increase the wind turbine complexity and induce extra costs in the wind turbine maintenance. Next, the sensor degradation impairs the signal accuracy and subsequently affects the reliability of condition monitoring methods. Moreover, the installation of sensors for monitoring blade conditions requires additional capital investments. Therefore, it is interesting and valuable to investigate nonintrusive monitoring approaches which do not request extra sensors.

Supervisory control and data acquisition (SCADA) systems have been widely mounted on commercial wind farms. The value of SCADA data in analyzing and monitoring conditions of wind turbine major components including the generator and gearbox has been reported in studies [18-20]. Monitoring blade health conditions with SCADA data has not been

Long Wang and Zijun Zhang are with the Department of Systems Engineering and Engineering Management, College of Science and Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: long.wang@my.cityu.edu.hk, zijzhang@cityu.edu.hk).

Jia Xu and Ruihua Liu are with the Centre of Wind Farm Data Analysis and Performance Optimization, China Longyuan Power Group Corporation Ltd., Beijing, China.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2016.2621135, IEEE Transactions on Smart Grid

sufficiently discussed. It is meaningful to investigate its feasibility and evaluate its effectiveness. Compared with previous studies [18-20], monitoring blade breakages is a completely different problem and desires new approaches due to following challenges: 1) SCADA data commonly do not contain measurements relative to wind turbine blade conditions, such as, blade vibrations, mechanical loads, and acoustic emissions; 2) available SCADA parameters do not directly reflect wind turbine blade conditions; 3) effective methods for identifying blade breakages through SCADA data are not available.

In this research, a deep neural network (DNN) based framework for monitoring wind turbine blade conditions and identifying its breakages based on SCADA data is proposed. Previous studies [21-24] reported that DNNs were capable to model higher complexities and extract useful representations from data through building multiple layers of nonlinear transformations. A class of DNNs, the deep autoencoder (DA), is applied to derive a signal, the reconstruction error (RE), from available SCADA data to present the blade breakage trends. SCADA data of four wind farms, Shandong, Anhui, Tianjin, and Ningxia, are applied in this study. Based on the SCADA data of Shandong and Anhui wind farms, computational analyses of three blade breakage cases (see Fig. 1) are performed to show that the RE pattern will shift prior to breakage occurrences. To monitor the variation of RE and detect its shift, the Exponentially Weighted Moving Average (EWMA) control chart is applied. A blind study is conducted based on SCADA data of Tianjin and Ningxia wind farms to validate the effectiveness of the proposed framework on identifying impending blade breakages. Prior knowledge of the blade breakages in the blind study including the location and time is not available and results show that blade breakages are all successfully identified before their occurrences.

Fig. 1. Blade breakages in Shandong, China

II. Preliminary Data Analysis

Preliminary data analyses are conducted to demonstrate difficulties of directly identifying two blade breakages based on SCADA parameters. Data utilized in preliminary analyses are collected from a commercial wind farm located in Shandong Province, China.

The targeted wind farm is composed of 33 wind turbines and two blade breakage cases were reported during the year of 2014. The first case occurred at turbine No. 9 on May 13th, 2014 and another case occurred at turbine No. 2 on Jun 27th, 2014. The replacement of the broken blade at each turbine consumed about one week. The 10-min SCADA data of 25 wind turbine conditional parameters listed in Table I are collected from Feb 1st, 2014 to Jun 27th, 2014. A comparative analysis is firstly conducted to examine the trends of conditional parameters among wind turbines with and without blade breakages. However, irregular patterns of conditional parameters before the blade breakage occurrence cannot be discovered. The power curve shape, the power production, and the frequency of experiencing wind gusts are next analyzed to seek indicators of blade breakages.

TABLE I
Description of Wind Turbine Conditional Parameters

Parameter	Notation	Parameter	Notation
Time	t	Wind direction	D_w
Wind speed	v_w	Wind angle	θ_w
Power	P	Blade angle	θ_b
Generator speed	v_g	Environment temperature	T_e
Rotor speed	v_r	Nacelle temperature	T_n
U1 voltage	U_1	Gearbox oil temperature	T_g
U2 voltage	U_2	Gearbox lubrication pressure	P_g
U3 voltage	U_3	U1 winding temperature	T_{u1}
U1 current	I_1	Gearbox shaft 1 temperature	T_{g1}
U2 current	I_2	Gearbox shaft 2 temperature	T_{g2}
U3 current	I_3	Generator bearing A temperature	T_A
Power factor	f_p	Generator bearing B temperature	T_B
Power grid frequency	f_g	Generating capacity	P_c

A. Power Curve and Power Production Analyses

The power curve analysis aims to examine the impact of impending blade breakages on power curve shapes. The Fig. 2 plots power curves of turbines No. 2 and No. 9 based on SCADA data recorded over the week prior to occurrences of blade breakages. As shown in Fig. 2, the shape of power curves of two turbines with impending blade breakages is regular and such result indicates that it is inefficient to detect impending blade breakages by simply checking power curve shapes.

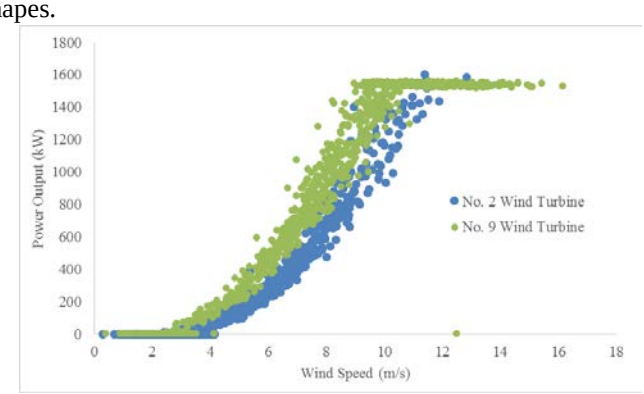


Fig. 2. Power curves of turbines No. 2 and No. 9

Power outputs of turbines No. 9 and No. 2 over the week before occurrences of blade breakages are compared with the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2016.2621135, IEEE Transactions on Smart Grid

average power output (APO) of all healthy turbines over same periods to examine irregularities caused by impending blade breakages. Results are summarized in Table II and we can observe that power outputs of turbines with and without impending blade breakages are similar. Thus, the power production cannot be considered as an efficient indicator for identifying impending blade breakages as well.

TABLE II
POWER PRODUCTION OF WIND TURBINES

Power output of turbine No. 9	APO of all turbines except No. 9	Power output of turbine No. 2	APO of all turbines except No. 2
164.73 MWh	160.18 MWh	35.55 MWh	33.03 MWh

B. Wind Gust Analysis

The frequent experience of wind gusts is a suspect factor inducing blade breakages [25]. Thus, wind speed data of all wind turbines are analyzed to figure out the average and maximal number of gusts as well as the maximal gust speed experienced by wind turbines. The procedure of detecting gusts follows the automated surface observing system (ASOS) algorithm [26]. During the period from May 12th 2014 to May 13th 2014, turbine No. 9 encountered 7 gusts which were similar to the maximal number of gusts experienced by other healthy wind turbines. However, its maximal gust speed (34.12 m/s) was less than that of healthy wind turbines (35.99 m/s). Meanwhile, during the period from June 26th 2014 to June 27th 2014, turbine No. 2 encountered smaller number of gusts and lower maximal gust speed compared with other healthy wind turbines. Therefore, the gust information cannot indicate impending blade breakages.

C. Summary and Discussions

Preliminary data analyses present that the straightforward application of existing SCADA parameters cannot effectively detect impending blade breakages. Meanwhile, power curves and power productions cannot reflect changes of blade conditions. The frequency of experiencing gusts cannot be regarded as a root cause of all blade breakages. The feasibility of utilizing SCADA data in blade condition monitoring requires further investigations and more sophisticated approaches need to be developed.

III. THE MONITORING METHOD

A DA based framework for monitoring blade breakages is presented in this section. The framework is composed of two parts, constructions of the DA model and monitoring chart. The DA model is built with all 25 parameters listed in Table I based on data of normal wind turbines. Based on the DA model, RE is derived from the comparison of a group of parameters to indicate the impending blade breakages. To monitor blade breakages with REs, a EWMA chart is constructed. A brief monitoring procedure is visualized in Fig. 3 and described in the following steps:

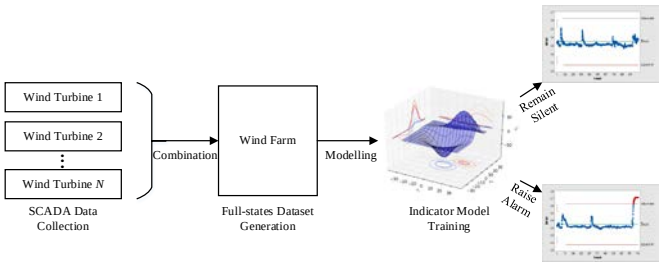


Fig. 3. Blade breakage monitoring framework

Step 1. Generate a training dataset: Collect SCADA data of all wind turbines in the wind farm and incorporate SCADA data of normal wind turbines into one dataset.

Step 2. Build a DA model: Apply all SCADA parameters to develop the DA model based on the training dataset.

Step 3. Compute REs: Compute the RE of wind turbine i, i = 1, 2, …, N, via the DA model. Wind turbines with impending blade breakages will have larger REs.

Step 4. Develop the EWMA control chart: Based on REs, a EWMA control chart is applied to estimate the upper and lower limits. If the RE of the wind turbine i exceeds the control limit, an alarm of the impending blade breakages will be activated.

In Step 2, the DA utilizes 25 parameters as inputs and reconstructs them via multiple layers of transformations. Compared with classical neural networks (NN), DAs do not specify a response parameter. During training process, the DA encodes and decodes the original data and learns the intrinsic structure of data in an unsupervised manner. RE is the square of Euclidean distance between the reconstructed inputs, $\hat{\mathbf{x}}$, and original inputs, $\mathbf{x}$, as expressed in (1).

$$\mathrm{RE} = \left\| \hat{\mathbf{x}} - \mathbf{x} \right\|_2^2 \tag{1}$$

In (1), $\| \cdot \|_2$ means the Euclidean norm.

A. The DA Model

The DA is a class of DNN which models relationships among all input parameters through constructing a symmetric network composed of the input layer, output layer, and multiple hidden layers. Fig. 4 presents the mechanism of developing the DA model including encoding and decoding processes. In the encoding process, the input $\mathbf{x}$ is firstly transformed to produce a set of features for further layer-wise transformations which develop the first half of the symmetric network. Restricted Boltzmann Machines (RBMs) are utilized as the mapping models in the transformation. Finally, code $\mathbf{y}$ is obtained through the encoding process. In the decoding process, the code $\mathbf{y}$ is iteratively transformed back to the original input space via RBMs and a reconstruction of $\mathbf{x}$, $\hat{\mathbf{x}}$, is generated. In order to minimize the difference between $\hat{\mathbf{x}}$ and $\mathbf{x}$, the model is treated as a feed-forward NN with multiple hidden layers trained by the back-propagation method [27]. Mathematical descriptions of the encoding and decoding processes are provided in (2) – (6).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2016.2621135, IEEE Transactions on Smart Grid

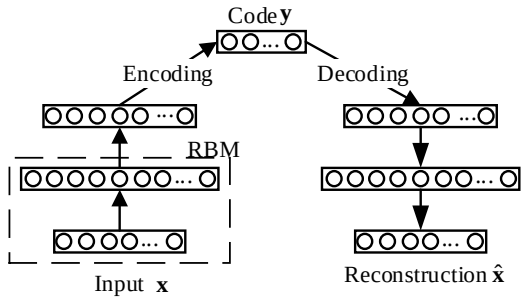


Fig. 4. Schematic diagram of Deep Autoencoder

According to Table 1, the input vector is

$$\mathbf{x} = [v_w, P, v_g, v_r, U_1, U_2, U_3, I_1, I_2, I_3, f_p, f_g, D_w, \theta_w, \theta_b, T_e, T_n, T_g, T_h, T_{u1}, T_{g1}, T_{g2}, T_A, T_B, P_c]^T \quad (2)$$

Next, **x** is transformed into a code **y**:

$$\mathbf{y} = f(\mathbf{x}) \quad (3)$$

where $f(\cdot)$ stands for the encoding process of **x**. The code, **y**, is mapped back to reconstruct **x** and produces $\hat{\mathbf{x}}$ The mapping follows $g(\cdot)$

$$\hat{\mathbf{x}} = g(\mathbf{y}) \quad (4)$$

where g(·) stands for the decoding process of **y**. The sigmoid function in (5) is employed as the activation function in the network.

$$s(x) = \frac{1}{1+e^{-x}} \quad (5)$$

The DA training process described in (6) aims to attain the optimal parameters which minimize the sum of REs.

$$\{\mathbf{W}_l, \mathbf{b}_l\} = \underset{\mathbf{W}_l, \mathbf{b}_l}{\operatorname{argmin}} \sum_{i=1}^{N} \frac{1}{2} \left\| \hat{\mathbf{x}}_i - \mathbf{x}_i \right\|^2, l = 0,1,\cdots,L \quad (6)$$

where $\mathbf{W}_l$ and $\mathbf{b}_l$ are weights and bias of the layer l, L is the number of layers in DA, and N is the number of samples in the dataset.

Training DNN models with the back-propagation algorithm might lead to local optima [28]. Therefore, a training process containing two phases, the pre-training and fine-turning, is introduced to better optimize the DA model. The layer-wise construction of RBMs is implemented in the pre-training phase to sequentially initialize weights and biases of each hidden layer. After training all hidden layers with RBMs, the network is next fine-tuned via the back-propagation method to obtain the globally optimized parameters.

1) Restricted Boltzmann Machines

RBMs are developed to learn initial weights and biases of hidden layers in the DA model. The RBM is a graphical model composed of hidden and visible units as shown in Fig. 5. In an RBM, connections between hidden and visible units are included while intra-connections among them are not considered. The m visible units, **V**, represent observable data and n hidden units, **H**, capture dependencies between observed variables. Two types of RBMs, binary RBM and Gaussian-Bernoulli RBM (GBRBM), are utilized in this study. GBRBM is employed to transform the original input into binary representations while binary RBMs are applied into the construction of hidden layers.

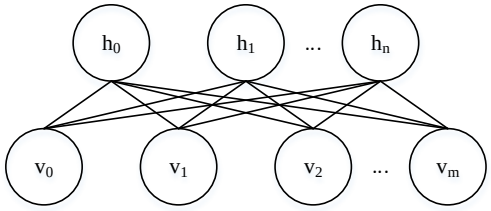


Fig. 5. Restricted Boltzmann Machine

In binary RBMs, (**V**, **H**) take values $(\mathbf{v}, \mathbf{h}) \in \{0,1\}^{m+n}$ and the joint distribution is given by the Gibbs distribution [29] in (7):

$$P(\mathbf{v}, \mathbf{h}) = \frac{1}{z} e^{-\mathbf{h}^T \mathbf{W} \mathbf{v} - \mathbf{c}^T \mathbf{h} - \mathbf{b}^T \mathbf{v}} \quad (7)$$

where z is the normalization constant, and **W** is the symmetric weights between **H** and **V** with bias **b** and **c,** separately. An energy function is then defined in (8):

$$E(\mathbf{v}, \mathbf{h}) = \mathbf{h}^T \mathbf{W} \mathbf{v} + \mathbf{c}^T \mathbf{h} + \mathbf{b}^T \mathbf{v} \quad (8)$$

Due to the graphical structure of RBMs, hidden units are independent given the state of the visible units and vice versa. Thus, conditional probabilities of hidden and visible units can be formulated as:

$$p(\mathbf{v} \mid \mathbf{h}) = \prod_{i \in \text{visible}} p(v_i \mid \mathbf{h}) \quad (9)$$

$$p(\mathbf{h} \mid \mathbf{v}) = \prod_{j \in \text{hidden}} p(h_j \mid \mathbf{v}) \quad (10)$$

where v_i and h_j are binary states of visible unit i and hidden unit j.

If a stochastic NN is utilized to interpret RBM, the conditional probability of a unit being one is calculated based on the output of a neutron activation function:

$$p(v_i = 1 \mid \mathbf{h}) = s(b_i + \mathbf{W}_i \mathbf{h}) \quad (11)$$

$$p(h_j = 1 \mid \mathbf{v}) = s(c_j + \mathbf{W}'_j \mathbf{v}) \quad (12)$$

Where $\mathbf{W}_i$ and $\mathbf{W}'_j$ are the weight vectors of visible unit i and hidden unit j, and b_i and c_j are biases. By minimizing the negative log-likelihood of the observed input data (minimizing the energy of RBM) with a gradient descent method, the update rule for **W** turns out to be

$$\Delta w_{ij} = \varepsilon(< v_i h_j >_{data} - < v_i h_j >_{\infty}) \quad (13)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2016.2621135, IEEE Transactions on Smart Grid

where ε is the learning rate, $<\cdot>_{data}$ denotes the expectation with respect to the data distribution and $<\cdot>_{\infty}$ denotes the expectation with respect to the model distribution. As the gradient is intractable, a contrastive divergence (CD) algorithm [30] is applied to estimate the gradient. In practice, the 1-iteration Gibbs sampling is often applied. Thus, the update rule is described in (14).

$$\Delta W_{ij} = \varepsilon(< v_i h_j >_{data} - < v_i h_j >_1) \tag{14}$$

This 1-step CD algorithm is described by the Pseudo Code, OneStepCD(). In OneStepCD(), *S* is the training dataset.

OneStepCD(S, ε)
{
  let $\Delta w_{ij} = \Delta b_j = \Delta c_i = 0$ *for* $i = 1,\ldots,n,\ j = 1,\ldots,m$
  for $\mathbf{v} \in S$:
    let $\mathbf{v}^{(0)} = \mathbf{v}$
    for $i \in [0,n]$:
      do sample $h_i^{(0)} \sim p(h_i \mid \mathbf{v}^{(0)})$
    end for
    for $j \in [0,m]$:
      do sample $v_j^{(1)} \sim p(v_j \mid \mathbf{h}^{(0)})$
    end for
    for $i \in [0,n],\ j \in [0,m]$:
      let $\Delta w_{ij} = \Delta w_{ij} + p(H_i = 1 \mid \mathbf{v}^{(0)}) \cdot v_j^{(0)} - p(H_i = 1 \mid \mathbf{v}^{(1)}) \cdot v_j^{(1)}$
      let $\Delta b_j = \Delta b_j + v_j^{(0)} - v_j^{(1)}$
      let $\Delta c_i = \Delta c_i + p(H_i = 1 \mid \mathbf{v}^{(0)}) - p(H_i = 1 \mid \mathbf{v}^{(1)})$
    end for
}

Since input parameters are continuous, a GBRBM is utilized to train the first hidden layer. The GBRBM has visible units with real values and binary hidden units. The energy function of GBRBM is then defined as

$$E(\mathbf{v},\mathbf{h}) = -\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij} h_j \frac{v_i}{\sigma_i} - \sum_{i=1}^{m} \frac{(v_i - b_i)^2}{2\sigma_i^2} - \sum_{j=1}^{n} c_j h_j \tag{15}$$

where σ_i is the standard deviation associated with Gaussian visible units. Similarly, conditional probabilities of visible and hidden units are

$$p(v_i = v \mid \mathbf{h}) = N(\mathbf{v} \mid b_i + \sum_{j \in \text{hidden}} h_j w_{ij}, \sigma_i^2) \tag{16}$$

$$p(h_j = 1 \mid \mathbf{v}) = sigm(c_j + \sum_{i \in \text{visible}} w_{ij} \frac{v_i}{\sigma_i^2}) \tag{17}$$

where $N(\cdot|\mu,\sigma^2)$ denotes the Gaussian probability density function with a mean μ and a standard deviation σ. The one-step CD learning is also utilized in the updating rule described in (18)

$$\Delta w_{ij} = \varepsilon(< \frac{1}{\sigma_i^2} v_i h_j >_{data} - < \frac{1}{\sigma_i^2} v_i h_j >_1) \tag{18}$$

Given the observed input data, a RBM is trained to obtain hidden representations of the data. Next, learnt representations can be viewed as new inputs and the subsequent RBM is trained. This procedure is repeated many times to initialize the weights and biases of each hidden layer.

2) Finalization of DA

The DA model in this study contains five hidden layers organized in a symmetric form. Four configurations of the number of units distributed in five hidden layers, Config1 = [100, 50, 10, 50, 100], Config2 = [300, 150, 10, 150, 300], Config3 = [500, 250, 10, 250, 500], and Config4 = [700, 350, 10, 350, 700] are considered in developing the DA model. The number of hidden units in the middle hidden layer, 10, is selected based on a principal component analysis (PCA) [31] considering a criterion that the summation of first 10 largest eigenvalues contributes 90% of the sum of all eigenvalues in the covariance matrix. A large number of hidden units in other layers are arbitrary options. Hidden layers are pre-trained by stacking RBMs developed with the 1-step CD algorithm. After the pre-training is complete, a linear layer is selected as the output layer to reconstruct original inputs based on outputs of the last hidden layer. Next, the whole DA is fine-tuned with a mini-batch gradient descent algorithm [32] and the batch size is set to 20 according to [33]. The mini-batch gradient descent update of DA parameters, $\mathbf{W}_l$ and $\mathbf{b}_l$, is described as

$$\begin{aligned} \mathbf{W}_l^{(t)} &\leftarrow \mathbf{W}_l^{(t-1)} - \alpha \frac{1}{20} \sum_{t'=20t+1}^{20(t+1)} \frac{\partial \left\| \mathbf{z}_{t'} - \hat{\mathbf{z}}_{t'} \right\|^2}{\partial \mathbf{W}_l} \\ \mathbf{b}_l^{(t)} &\leftarrow \mathbf{b}_l^{(t-1)} - \alpha \frac{1}{20} \sum_{t'=20t+1}^{20(t+1)} \frac{\partial \left\| \mathbf{z}_{t'} - \hat{\mathbf{z}}_{t'} \right\|^2}{\partial \mathbf{b}_l} \end{aligned} \tag{19}$$

where $\mathbf{z}_{t'}$ is a data point sampled at iteration t', $\hat{\mathbf{z}}_{t'}$ is the reconstruction of $\mathbf{z}_{t'}$ via DA, and α is the learning rate, which is set to 0.01 according to [33]. After training, DA models with two configurations, Config3 and Config4, result in similar least REs and the DA model with the configuration, Config3, is finally selected because of its simpler structure.

B. The RE Results

According to previously mentioned procedures, REs of wind turbines in the Shandong wind farm based on data from May 6th, 2014 to May 13th, 2014 are computed. REs of DA models of the turbine No. 9 and a randomly selected healthy wind turbine are compared in Fig. 6. It is observable that there exists a RE shift prior to the blade breakage at turbine No. 9 and such shift can be utilized to indicate the impending blade breakage. In addition, DA models without considering pre-training or temperature parameters as well as a shallow NN, three layers perceptron, are developed based on the same data separately. Their REs are demonstrated in Figs. 7, 8, and 9.

As shown in Figs. 7 and 8, the RE shift disappears and it indicates that pre-training and the consideration of temperature

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2016.2621135, IEEE Transactions on Smart Grid

parameters are important to identify the impending blade breakage based on the DA model. Moreover, according to Fig. 9, the shallow NN fails to extract meaningful patterns for identifying blade breakages from data.

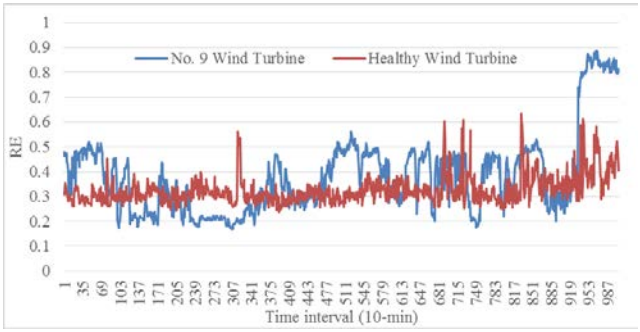


Fig. 6. REs based on turbines with and without impending blade breakages.

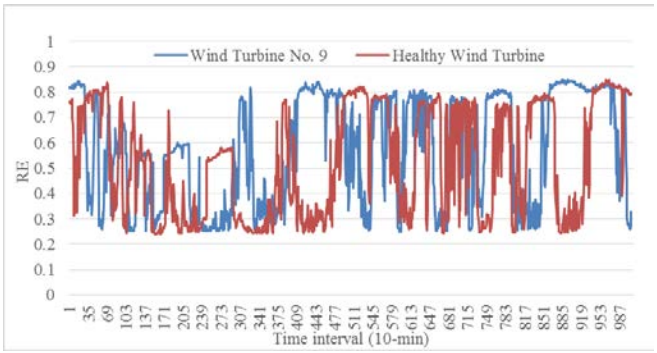


Fig. 7. REs of DA model of two turbines without pre-training.

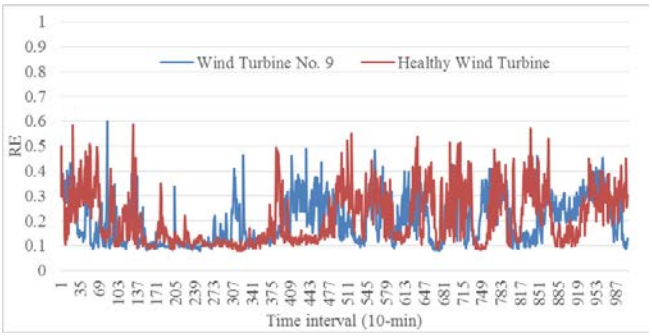


Fig. 8. REs of DA model of two turbines without temperature parameters.

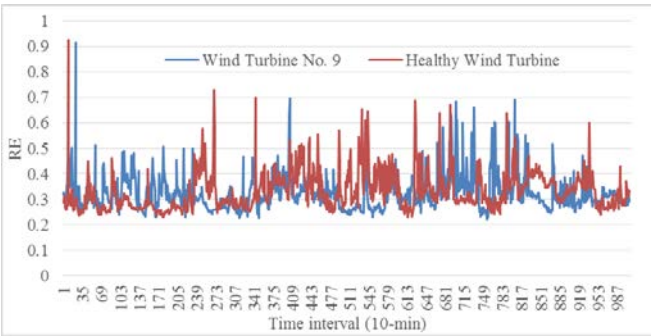


Fig. 9. REs of a shallow NN of two turbines.

C. EWMA Control Chart

The process of the blade breakage is gradual and the early detection is meaningful to decide suitable actions, which may prevent their occurrences. Since the DA model is trained with data of healthy wind turbines, the change of RE patterns can indicate impending blade breakages. As the shift of RE pattern is small before blade breakages (see Fig. 6), the EWMA chart designed for identifying small process shifts [34] is a suitable option for continuously monitoring the variation of the RE.

The EWMA, q_t, is computed as (20).

$$q_t = \lambda RE_t + (1-\lambda)q_{t-1} \tag{20}$$

where t is the time index, $\lambda \in (0, 1]$ is a weight of historical RE considered in constructing the EWMA, and q_0 is the mean of REs for the monitored wind turbine during a specific period. The λ is arbitrarily set to 0.2 in this study based on the recommendation from [34].

The mean and variance of q_t are next calculated according to (21).

$$\mu_{q_t} = \mu_{\mathrm{RE}}, \sigma_{q_t}^2 = \frac{\sigma_{\mathrm{RE}}^2}{n_s}\left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2t}\right] \tag{21}$$

where μ_{RE} is the mean of REs for all wind turbines within the same wind farm, σ_{RE} is the standard deviation of REs, and n_s is the sample size.

The upper and lower control limits of the EWMA chart in (22) and (23) are functions of t.

$$UCL(t) = \mu_{\mathrm{RE}} + L\sigma_{\mathrm{RE}}\sqrt{\frac{\lambda[1-(1-\lambda)^{2t}]}{(2-\lambda)n_s}} \tag{22}$$

$$LCL(t) = \mu_{\mathrm{RE}} - L\sigma_{\mathrm{RE}}\sqrt{\frac{\lambda[1-(1-\lambda)^{2t}]}{(2-\lambda)n_s}} \tag{23}$$

where L is usually set to 3 [35].

Besides recommended settings of L and λ, the grid search [36] can be applied to iteratively identify most suitable settings of L and λ based on the RE obtained in training.

IV. Case Studies

A computational study and a blind validation are presented in this section. In the computational study, the proposed framework is evaluated with 10-min SCADA data of two wind farms in Shandong and Anhui, China. REs of modeling wind turbines with and without impending blade breakages are computed. The EWMA chart is constructed to monitor REs and its capability of identifying wind turbines with impending blade breakages is evaluated. In the blind validation, the proposed monitoring approach is directly applied into data of two wind farms selected by the industrial partner. Two wind farms are located in Tianjin and Ningxia Provinces of China and the prior knowledge of their blade breakage cases is not provided. The identification accuracy is next examined by the industrial partner.

A. The Computational Study

Besides the dataset utilized in Section II, data of another wind farm composed of 33 wind turbines in Anhui, China is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2016.2621135, IEEE Transactions on Smart Grid

utilized in the computational study. A blade breakage case occurred at turbine No. 5 on August 2nd, 2015 is reported. The same SCADA system is installed in the Anhui wind farm and data of wind turbines are collected from May 1st, 2015 to August 31st, 2015.

Two DA models are developed based on SCADA data collected from Shandong and Anhui wind farms respectively. Parameters of healthy wind turbines described in Table I are utilized in training DA models. The training datasets from Shandong and Anhui wind farms include 399,639 and 321,280 data points separately. In training DA models, weights and biases are firstly initialized based on the layer-wised RBM pre-training. Next, weights and biases are updated via the mini-batch SGD to minimize the sum of REs. After training DA models, REs are computed according to (1). The EWMA control chart is constructed to monitor REs. To better visualize results, REs over the week prior to the blade breakage occurrence are plot in EWMA charts. End points in Figs. 11, 13, and 14 are 0:10 a.m. on August 2nd, 7:00 a.m. on May 13th, and 1:40 a.m. on June 27th, respectively. The UCL and LCL of EWMA charts are set according to (22) and (23). Selected EWMA charts are shown in Figs. 10-14.

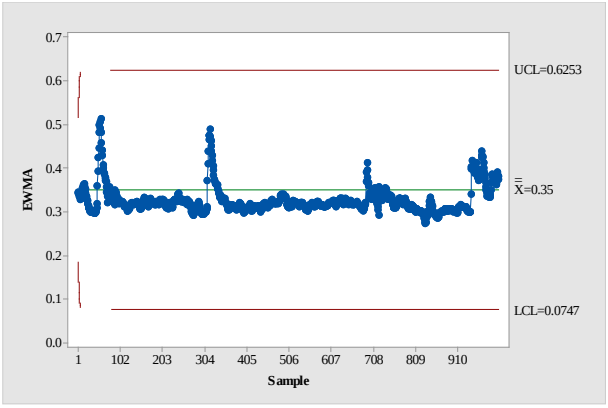


Fig. 10. The EWMA chart of turbine No. 3 in Anhui (July 26th to August 2rd, 2015)

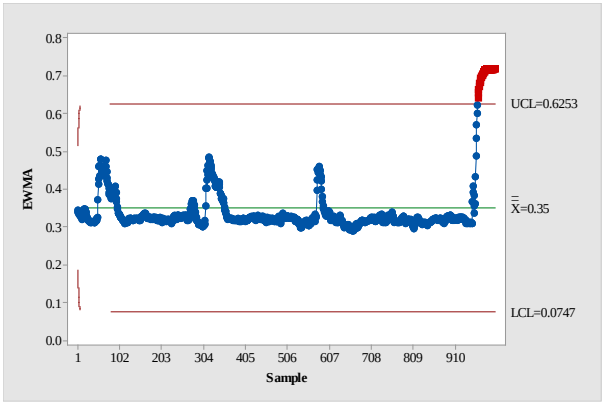


Fig. 11. The EWMA chart of turbine No. 5 in Anhui (July 26th to August 2rd, 2015)

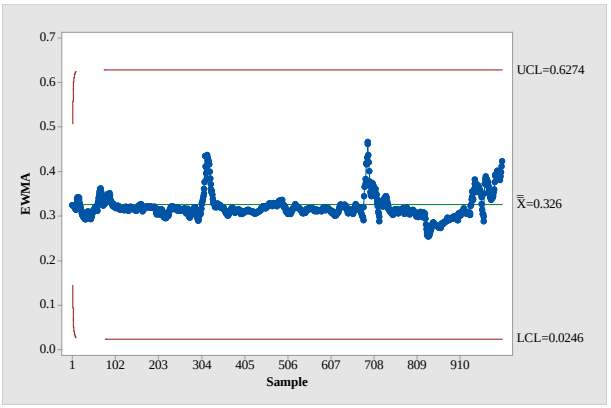


Fig. 12. The EWMA chart of turbine No. 15 in Shandong (May 6th to May 13th, 2014)

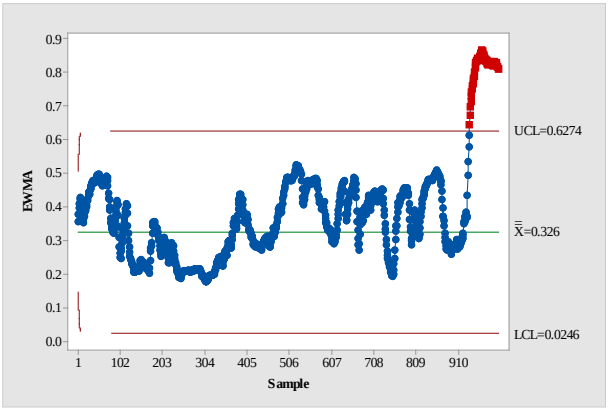


Fig. 13. The EWMA chart of turbine No. 9 in Shandong (May 6th to May 13th, 2014)

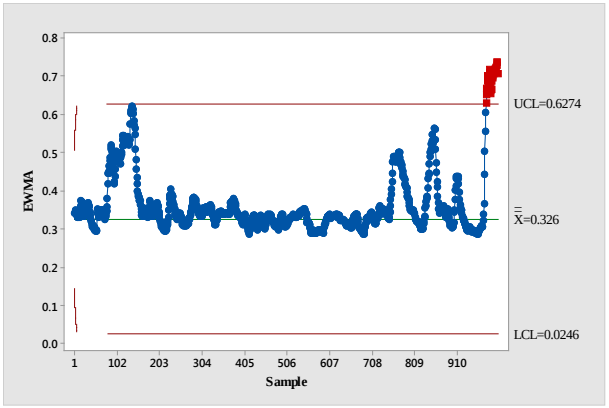


Fig. 14. The EWMA chart of turbine No. 2 in Shandong (June 20th to June 27th, 2014)

Fig. 10 plots the EWMA chart of a healthy wind turbine randomly selected from the Anhui wind farm while Fig. 11 plots that of a wind turbine with the impending blade breakage in the same wind farm. The EWMAs in Fig. 10 all fall within the control limits while outliers are observed as shown in Fig. 10. Moreover, the first outlier in Fig. 11 is detected about 8 hours prior to the blade breakage occurrence so that an early alarm can be issued. Figs. 12-14 present similar results of wind turbines in the Shandong wind farm. All EWMA charts of abnormal wind turbines can detect outliers prior to the occurrence of blade breakages (see Figs. 13 and 14).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2016.2621135, IEEE Transactions on Smart Grid

Based on computational results, we observe that all three blade breakage cases are successfully identified with the proposed monitoring framework. Meanwhile, impending blade breakages can be identified more than 6 hours ahead which offers sufficient time for wind farm operators to take reactions. Details of the identification of impending blade breakages are described in Table III.

TABLE III
IDENTIFICATIONS OF IMPENDING BLADE BREAKAGES

Wind farm	Turbines	Time prior to breakages
Shandong	No. 9	7 hours and 20 minutes
Shandong	No. 2	6 hours and 10 minutes
Anhui	No. 5	8 hours

B. *The Blind Validation*

SCADA data of two wind farms located in Ningxia and Tianjin, China provided by the industrial partner are applied to further validate the effectiveness of our proposed monitoring framework in this study. Datasets of the Ningxia and Tianjin wind farms contain SCADA data of 5 wind turbines and SCADA data of 22 wind turbines separately. Prior information of blade breakages in two wind farms is unavailable in the blind validation study. The monitoring results are firstly produced and next compared with the field report by the industrial partner. The SCADA data of Ningxia and Tianjin wind farms are collected from June 7th, 2015 to August 5th, 2015 and from May 1st, 2015 to July 25th, 2015, respectively.

Similarly, two DA models are developed based on SCADA data of two wind farms. As the information of wind turbines with blade breakages is unavailable, SCADA data of all wind turbines are considered as the training dataset in the validation. REs of all wind turbines are firstly computed for the whole period. Next, the EWMA chart is developed based on REs of each week for all wind turbines. After examining the EWMA charts of all wind turbines week by week, we discover that each wind farm has a wind turbine with outliers detected by the EWMA chart (see Figs. 15 and 16). The monitoring results are then compared with the field reports by the industrial partner to confirm their accuracies. According to the field report, wind turbines with blade breakages in two wind farms have been successfully identified and the accuracy is 100%.

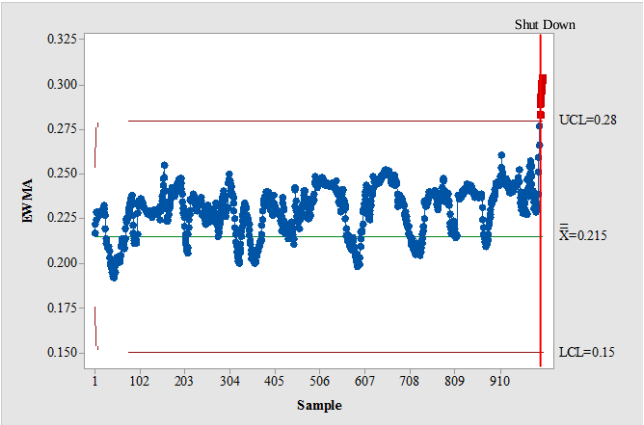


Fig. 15. The EWMA chart of turbine C05 in Ningxia (July 19th to July 25th, 2015)

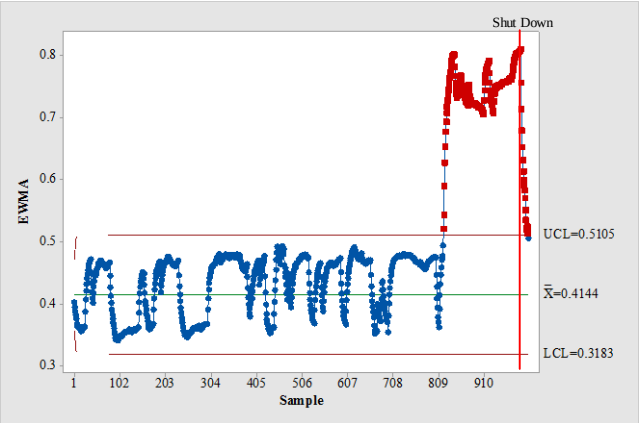


Fig. 16. The EWMA chart of turbine No. 70 in Tianjin (July 18th to July 24th, 2015)

As shown in Figs. 15 and 16, the proposed monitoring approach is able to identify impending blade breakages at least 2 hours prior to their occurrences (23:45 p.m. on July 25th and 19:57 p.m. on July 24th respectively). As SCADA data of all wind turbines are applied into the model development, wind turbines with impending blade breakages are identified slightly later relative to results of the computational study. However, the sweet spot is that false alarms are not observed in this blind validation which could reduce the unnecessary check of blade conditions.

Besides the breakage identification accuracy, the pre-training with RBMs benefits the training process of DA in terms of the computational time. A PC with a CORE I5 CPU and 8GB memory is employed in this study to perform the analysis and the training time of DA models with 5 hidden layers is less than 40-min. Once the DA model is developed, it can generate REs immediately. Thus, both the effectiveness and computational efficiency makes the proposed framework applicable in the industry. Although the proposed framework works well so far, it needs to be further validated by field tests in the future. To apply the proposed framework into a wind farm containing multiple categories of turbines in terms of the age and model, the wind farm dataset needs to be firstly separated by turbine categories and next the proposed framework can be constructed based on sub-datasets. To configure a more effective training dataset of DA models in practice, a validation dataset is firstly constructed by randomly sampling historical data. Next, a base training dataset will be gradually expanded until the REs based on validation data do not decrease. This strategy can better generalize the DA model.

V. CONCLUSION

This paper presented a data-driven framework for identifying impending wind turbine blade breakages based on SCADA data. As commercial SCADA systems did not directly measure blade conditions, the DA model was applied to derive an indicator of blade breakages, the RE, from SCADA data. The EWMA control chart was chosen to construct upper and lower criteria for identifying abnormal REs. SCADA data collected from four wind farms located in Shandong, Anhui, Ningxia, and Tianjin Provinces in China were utilized in this

study. The feasibility and effectiveness of the proposed monitoring approach was examined through the computational study and blind validation.

In the computational study, SCADA data of all healthy wind turbines were considered as a training dataset and the DA model was trained with a 1-step CD algorithm. Based on computed REs, EWMA control charts were developed for wind turbines. Impending blade breakages were successfully identified according to the computational results. In the blind validation, prior information of blade breakages was unavailable. SCADA data of all wind turbines were applied to develop DA models. EWMA charts were built for wind turbines based on REs of one week length. Through repeatedly examining EWMA charts weekly, wind turbines with impending blade breakages were successfully identified.

Results demonstrated that the proposed method was applicable to identify impending blade breakages based on SCADA data. Early identification could be achieved based on computational results and false alarms were not observed. In the future study, the proposed approach needs to be further validated with field tests.

VI. References

[1] M. Schlechtingen and I.F. Santos, "Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection," *Mech. Syst. Signal Pr.*, vol. 25, no. 5, pp. 1849-1875 2011.

[2] J. Xiang, S. Watson and Y. Liu, "Smart monitoring of wind turbines using neural networks," in *Sustainability in Energy and Buildings*, Brighton : Springer, 2009, pp. 1-8.

[3] Z. Zhang, A. Verma and A. Kusiak, "Fault analysis and condition monitoring of the wind turbine gearbox," *IEEE Trans. Energy Convers.*, vol. 27, no. 2, pp. 526-535 2012.

[4] J. Chou, C. Chiu, I. Huang and K. Chi, "Failure analysis of wind turbine blade under critical wind loads," *Eng. Failure Anal.*, vol. 27, pp. 99-118, 1 2013.

[5] K.M. Farinholt, S.G. Taylor, G. Park and C.M. Ammerman, "Full-scale fatigue tests of CX-100 wind turbine blades. Part I: testing," in *SPIE Smart Structures and Materials Nondestructive Evaluation and Health Monitoring*, 2012, pp. 83430P-83430P-8.

[6] S. Faulstich, B. Hahn and P.J. Tavner, "Wind turbine downtime and its importance for offshore deployment," *Wind Energy*, vol. 14, no. 3, pp. 327-337 2011.

[7] S. Butterfield, S. Sheng and F. Oyague, *Wind Energy's New Role in Supplying the World's Energy: What Role Will Structural Health Monitoring Play?*, National Renewable Energy Laboratory (NREL), Golden, CO., 2009.

[8] S. Uadiale, E. Urban, R. Carvel, D. Lange and G. Rein, "Overview of problems and solutions in fire protection engineering of wind turbines," *Fire Safety Science*, vol. 11, PP. 983-995 2014.

[9] H. Sutherland, A. Beattie, B. Hansche, W. Musial, J. Allread, J. Johnson and M. Summers, "The application of non-destructive techniques to the testing of a wind turbine blade," *Sandia National Laboratories, Albuquerque, NM* 1994.

[10] B.F. Sørensen, L. Lading, P. Sendrup, M. McGugan, C.P. Debel, O.J. Kristensen, G.C. Larsen, A.M. Hansen, J. Rheinländer and J. Rusborg, "Fundamentals for remote structural health monitoring of wind turbine blades-a preproject", Risoe National Laboratory for Sustainable Energy 2002.

[11] A. Dutton, "Thermoelastic stress measurement and acoustic emission monitoring in wind turbine blade testing," in *European Wind Energy Conference London*, 2004, pp. 22-25.

[12] M.W. Häckell and R. Rolfes, "Monitoring a 5MW offshore wind energy converter—Condition parameters and triangulation based extraction of modal parameters," *Mech. Syst. Signal Pr.*, vol. 40, no. 1, pp. 322-343 2013.

[13] I. Prowell, M. Veletzos, A. Elgamal and J. Restrepo, "Experimental and numerical seismic response of a 65 kW wind turbine," *J. Earthquake Eng.*, vol. 13, no. 8, pp. 1172-1190 2009.

[14] C. Pitchford, B.L. Grisso and D.J. Inman, "Impedance-based structural health monitoring of wind turbine blades," in *The 14th International Symposium on: Smart Structures and Materials & Nondestructive Evaluation and Health Monitoring*, 2007, pp. 65321I-65321I-11.

[15] J. Lee, J. Park, K. Oh, S. Ju and J. Lee, "Transformation algorithm of wind turbine blade moment signals for blade condition monitoring," *Renew. Energy*, vol. 79, pp. 209-218 2015.

[16] M. Ozbek, F. Meng and D.J. Rixen, "Challenges in testing and monitoring the in-operation vibration characteristics of wind turbines," *Mech. Syst. Signal Pr.*, vol. 41, no. 1, pp. 649-666 2013.

[17] O. Karabayir, S.M. Yucedag, A.F. Coskun, O.M. Yucedag, H.A. Serim and S. Kent, "Wind turbine signal modelling approach for pulse doppler radars and applications," *IET Radar Sonar Nav.*, vol. 9, no. 3, pp. 276-284 2015.

[18] A.B. Borchersen and M. Kinnaert, "Model-based fault detection for generator cooling system in wind turbines using SCADA data," *Wind Energy*, pp. n/a-n/a 2015.

[19] P. Bangalore and L.B. Tjernberg, "An artificial neural network approach for early fault detection of gearbox bearings," *IEEE Trans. Smart Grid* vol. 6, no. 2, pp. 980-987 2015.

[20] H. Long, L. Wang, Z. Zhang, Z. Song and J. Xu, "Data-driven wind turbine power generation performance monitoring," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6627-6635 2015.

[21] Y. Park and M. Kellis, "Deep learning for regulatory genomics," *Nat. Biotechnol.*, vol. 33, no. 8, pp. 825-826 2015.

[22] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland and G. Ostrovski, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529-533 2015.

[23] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444 2015.

[24] B. Alipanahi, A. Delong, M.T. Weirauch and B.J. Frey, "Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning," *Nat. Biotechnol.* 2015.

[25] T. Burton, D. Sharpe, N. Jenkins and E. Bossanyi, *Wind energy handbook*, John Wiley & Sons, 2001.

[26] M.D. Powell, "Wind measurement and archival under the automated surface observing system (ASOS): User concerns and opportunity for improvement," *Bull. Am. Meteorol. Soc.*, vol. 74, no. 4, pp. 615-623 1993.

[27] Williams, DE Rumelhart GE Hinton RJ and G. Hinton, "Learning representations by back-propagating errors," *Nature*, pp. 323,533-538 1986.

[28] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *International conference on artificial intelligence and statistics*, 2010, pp. 249-256.

[29] A. Fischer and C. Igel, "An introduction to restricted Boltzmann machines," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, Buenos Aires : Springer, 2012, pp. 14-36.

[30] M.A. Carreira-Perpinan and G.E. Hinton, "On contrastive divergence learning," in *Proceedings of the tenth international workshop on artificial intelligence and statistics*, 2005, pp. 33-40.

[31] I. Jolliffe, *Principal component analysis*, New York: John Wiley & Sons, 2002.

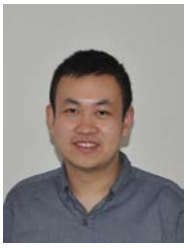
[32] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010*, Heidelberg: Springer, 2010, pp. 177-186.

[33] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural Networks: Tricks of the Trade*, Berlin: Springer, 2012, pp. 437-478.

[34] D.C. Montgomery, *Introduction to statistical quality control*, New York: John Wiley & Sons, 2007.

[35] S.S. Prabhu and G.C. Runger, "Designing a multivariate EWMA control chart," *J. Qual. Technol.*, vol. 29, no. 1, pp. 8 1997.

[36] P. Cortez, *Modern optimization with R*, New York: Springer, 2014.



Long Wang (S'16) received his M.Sc. degree in Computer Science with distinction from University College London, London, U.K., in 2014. He is currently pursuing the Ph.D. degree in the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong.

His research interests include wind turbine condition monitoring, electricity price forecasting, computer vision, and reinforcement learning.

Zijun Zhang (M'12) received his Ph.D. and M.S. degrees in Industrial Engineering from the University of Iowa, Iowa City, IA, USA, in 2012 and 2009, respectively, and B.Eng. degree in Systems Engineering and Engineering Management from the Chinese University of Hong Kong, Hong Kong, China, in 2008.

Currently, he is an Assistant Professor in the Department of Systems Engineering and Engineering Management at the City University of Hong Kong, Hong Kong, China. His research focuses on data mining and computational intelligence with applications in wind energy, HVAC and wastewater processing domains.

Jia Xu received his M.S. degree in Instrument Science and Technology at Xi'an Jiaotong University, China, in 2010.

Now, he is the head of the Centre of the Wind Farm Data Analysis and Performance Optimization at China Longyuan Power Group Co. Ltd., Beijing, China.

Ruihua Liu received the B.Eng. degree in Mechatronic Engineering and the M.S. degree in mechanical manufacturing and automation from the North China Electric Power University.

He is currently an Engineer at China Longyuan Power Group Co. Ltd., Beijing, China.