

Condition Monitoring of Wind Power System With Nonparametric Regression Analysis

Nattavut Yampikulsakul, Eunshin Byon, *Member, IEEE*, Shuai Huang, Shuangwen Sheng, *Member, IEEE*, and Mingdi You

Abstract—Condition monitoring helps reduce the operations and maintenance costs by providing information about the physical condition of wind power systems. This study proposes to use a statistical method for effective condition monitoring. The turbine operation is significantly affected by external weather conditions. We model the wind turbine response as a function of weather variables, using a nonparametric regression method named least squares support vector regression. In practice, online condition monitoring of wind power systems often relies on datasets contaminated with outliers. This study proposes to use a weighted version of least squares support vector regression that provides a formal procedure for removing the outlier effects. We determine the decision boundaries to distinguish faulty conditions from normal conditions by examining the variations in the operational responses that are significantly affected by external weather. The results show that the proposed method effectively detects anomalies.

Index Terms—Control chart, fault diagnosis, statistical process control, support vector regression (SVR), wind energy.

I. INTRODUCTION

A U.S. Department of Energy (DOE) scenario published in 2008 projects that wind will supply 20% of U.S. electricity by 2030 [1]. Integrating this substantial volume of energy into the power grid will require significant advances in the wind power system management. Today, premature failure rates, restoration times, and operation and maintenance (O&M) costs of wind power are higher than desirable [1], [2]. Recently, condition monitoring systems have been introduced for controlling and reducing the O&M costs [3], [4]. However, stochastic weather conditions, sensor noise, and inherent variability of turbine components lead to miss-detection of faults and frequent false alarms.

This study examines condition monitoring of wind turbines by investigating the dependence of turbine responses on external

weather conditions. Based on a set of sophisticated statistical tools, we investigate the variations due to external environmental effects, and discriminate them from those caused by faults. Using historical data, we construct a baseline (reference) regression curve to represent the turbine output as a function of the weather variables. We theorize that significant changes from the baseline model will indicate a possible fault or anomaly.

Our proposed methodology is based on a nonparametric regression technique for characterizing the relationship between the turbine response and weather conditions. Wind turbine operation involves complicated aerodynamics, structural dynamics, and controls, and thus, it is difficult to define the shape of a response as a function of input weather conditions *a priori*. In addition, the turbine structure is subject to nonlinear stresses, so the structural responses show nonlinear patterns over a wide range of weather conditions. We employ a support vector machine (SVM), a nonparametric statistical method, to capture the complexity of turbine response, which does not require any predefined function structure. To solve various reliability problems, SVM has been utilized, including reliability evaluations of composite power systems [5] and fault localization in the transmission lines [6]. There are two types of SVMs: the support vector classifier can be used for problems where the output variable takes a finite number of values, and the support vector regression (SVR) is used for problems where the output variable is continuous. This study uses the SVR [7] because the turbine outputs, e.g., the power generation and bending moments from strain sensors, are continuous. Specifically, least squares SVR (LS-SVR) is employed due to its computationally attractive features.

Modeling the baseline curve requires uncontaminated data collected from a faultless condition because the curve should characterize the normal behavior of the response. However, wind turbines operate in the presence of large uncertainties, and many outliers are likely to be generated, particularly during transient operations such as start-up and shut-down. The presence of outliers has undesirable effects on modeling and monitoring of functional data, as discussed in [8], including inaccurate specification of the curve and incorrect estimation of the output variability. In practice, identifying these outliers often relies on experience or ad hoc procedures. The studies in [9] and [10] point out the necessities of using faultless data in finding the nominal behavior of the turbine outputs, but they do not provide formal procedures for removing the outliers. Therefore, the challenge is to extract the uncontaminated data collected under normal conditions in order to generate a robust reference curve. To this end, Suykens *et al.* [7] suggest using

Manuscript received December 18, 2012; revised August 21, 2013; accepted September 19, 2013. Date of publication January 16, 2014; date of current version May 15, 2014. Paper no. TEC-00672-2012.

N. Yampikulsakul, E. Byon, and M. You are with the Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, USA (e-mail: nattavut@umich.edu; ebyon@umich.edu; mingdiyou@umich.edu).

S. Huang is with the department of Industrial and Management Systems Engineering, University of South Florida, Tampa, FL 33620, USA (e-mail: shuaihuang@usf.edu).

S. Sheng is with the National Renewable Energy Laboratory, Golden, CO 80401, USA (e-mail: Shuangwen.Sheng@nrel.gov).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TEC.2013.2295301

a weighted version of LS-SVR, and it has been successfully employed in many applications [11]. In this study, we present a computationally efficient method to solve the weighted LS-SVR by taking advantage of an incomplete Cholesky matrix approximation [12], [13].

We also discuss how to set the decision boundaries, i.e., the upper and lower control limits, to determine the abnormal conditions. Detecting abnormalities falls into the general category of statistical process control (SPC). In the traditional SPC analysis, fixed control limits are used for fault monitoring. However, wind turbine responses and their variability are different across a range of input vectors. Consequently, usage of fixed control limits could skew the monitoring decision. We investigate the heterogeneous variability of the response, and by combining the results from the weighted LS-SVR model and SPC theory, we determine the decision boundaries that enable us to detect changes in the functional relationship. Two case studies using field data from real operations suggest that the proposed method has strong fault detection capability.

The remainder of the paper is organized as follows. We review related studies in Section II. Section III describes the dataset used in the case studies. Sections IV and V discuss how to construct the baseline regression model and the control charts, respectively. Sections VI and VII present case studies. Finally, we conclude the paper in Section VIII.

II. LITERATURE REVIEW

In general, the literature on condition monitoring techniques for wind power systems can be divided into two streams. The first (and the focus of this study) looks at finding the relationship between weather conditions and turbine response. For example, Caselitz and Giehardt [9] analyze the power characteristics for offshore turbines to provide information about the rotor condition. They categorize wind speeds with a bin width of 0.5 m/s. At each bin, they compute the average of the active power outputs collected under faultless rotor condition. The decision boundaries to detect abnormality are heuristically determined based on past experience. In the monitoring phase, an alarm triggers when a pre-determined number of subsequent measurements exceed the alarm limit. Due to its simplicity, this so-called “binning approach” represents current industrial practice. In this approach, separate alarm limits are used at each bin without considering the relationship between the power generation and weather variables.

Unlike the binning method in [9], Kusiak *et al.* [10] build a power curve as a function of weather variables, using various data mining techniques. Based on the Shewhart type control chart, they set static decision boundaries regardless of input weather conditions. Similarly, Stephen *et al.* [14] show the relationship between power generation and wind speed by deriving a joint density function using a Copula model. Uluyol *et al.* [15] also analyze the power curve by considering site-specific operational regimes, and use the residual distribution for component diagnostics and prognostics.

The second stream focuses on signal processing using data from supervisory control and data acquisition (SCADA) systems

TABLE I
SPECIFICATIONS OF WIND TURBINE USED IN THIS STUDY

Wind turbine model	NEG-Micon/2750
Hub height (m)	80
Rotor diameter (m)	92
Cut-in wind speed (m/s)	4
Cut-out wind speed (m/s)	25
Rated wind speed (m/s)	14
Nominal power (kW)	2,750
Control system	Pitch control
Location	Alborg, Denmark
Terrain	Coastal

and sensors installed for condition monitoring. Many of these studies investigate the series of turbine outputs based on Fourier transforms, wavelet analysis, cepstrum analysis, or time domain analysis, in order to extract damage-sensitive features [16], [17]. These studies perform output-only analysis without explicitly considering the effects of operating conditions on the variability in the responses. Because external conditions indirectly affect the subsystems inside the turbine nacelle after they interact with other components, such studies could be useful for monitoring some drive-train components where the relationship between weather conditions and sensor signals is harder to characterize.

III. DATASET

We use two datasets collected from real land-based turbines in this study.

A. Dataset 1

The first dataset, provided by Risø-DTU (Technical University of Denmark) [18], is used to illustrate the proposed method in Sections IV and V. Table I summarizes the turbine’s specifications in the first dataset. We consider power generation and blade bending moment as the response variables. The data is recorded at 25 HZ. In our analysis, we use a ten-minute average of the measurements. From the total of 3 010 ten-minute average records measured from November 2005 to January 2006, we use the first 2 500 records for constructing the baseline regression curves in the training phase and the remaining records for demonstrating our proposed method’s monitoring performance.

Fig. 1 shows the two scatter plots of the 3 010 records between the ten-minute average of each response variable and ten-minute average wind speed. The plots demonstrate the weather effects on the turbine responses. We observe nonlinear patterns between each response and wind speeds as well as the large variations and outliers in both scatter plots.

B. Dataset 2

In order to validate the presented approach, we also use a dataset provided by DOE’s National Renewable Energy Laboratory (NREL). The data were collected at 100 HZ from the two-blade Controls Advanced Research Turbine 2 (CART2) in Colorado whose rated power is 600 kW. This dataset includes measurements when the gear failure inside the turbine’s gearbox occurred.

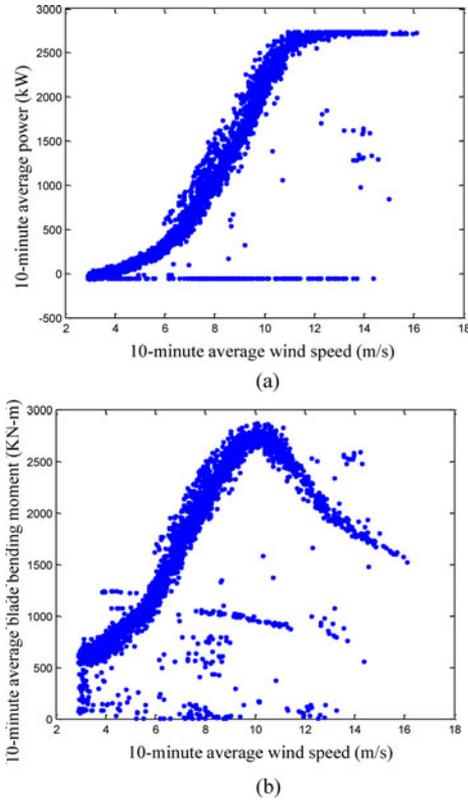


Fig. 1. Scatter plots between ten-minute average turbine responses and ten-minute average wind speed. (a) Power generation. (b) Blade bending moment in flapwise direction.

IV. NONPARAMETRIC REGRESSION WITH WEIGHTED LS-SVR

As mentioned, our primary objective is to determine the control limits for monitoring the turbine condition. We view the turbine response as a functional random variable whose values are affected by the weather variables. This section identifies the relationship between the weather condition and turbine responses, and construct a robust regression model when the historical dataset include outliers.

A. Model

We briefly present the weighted LS-SVR described by Suykens *et al.* [7], which can be used to learn the functional relationship between the weather variables (denoted as \mathbf{x}) and the turbine response variable (denoted as y). Specifically, the functional relationship can be written as $y = f(\mathbf{x}) + e_i$, where e_i is a random noise. From the SCADA system or sensors, we collect a historical dataset, $S = \{(y_i, \mathbf{x}_i), i = 1, \dots, M\}$, which consists of a set of realizations of this function. The dataset, S , provides the basis for learning the unknown function, $f(\cdot)$. Here, \mathbf{x}_i is the vector of explanatory variables (weather variables) for the i^{th} record in S , $y_i \in \mathbb{R}$ is the response variable (e.g., power generation and blade bending moment), and M is the size of the training dataset.

However, it is challenging to learn $f(\cdot)$ as the functional form since $f(\cdot)$ is unknown and possibly nonlinear. One approach is to use a nonparametric regression model to approximate $f(\cdot)$.

Specifically, this nonparametric regression model can be obtained by applying a nonlinear mapping, ϕ , which maps the input vector, \mathbf{x}_i , to a high-dimensional feature space, as:

$$y(\mathbf{x}_i) = b + w^T \phi(\mathbf{x}_i) + e_i, \quad i = 1, \dots, M \quad (1)$$

where b and w are the regression parameters.

Note that the regression function in (1) should represent the turbine output under the faultless condition. When the training dataset includes abnormal samples collected under faulty conditions such as sensor faults, measurement errors or subsystem damage (e.g., the data depicted in Fig. 1 demonstrates the presence of abnormal data), we should remove them from the dataset, S . Identifying abnormal samples is not straightforward in many cases, and generally ad hoc rules are applied. The weighted LS-SVR [7] provides a formal and automated procedure where the function is estimated by solving the following optimization problem:

$$\min \quad \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum v_i e_i^2 \quad (2)$$

s.t.

$$y(\mathbf{x}_i) = w^T \phi(\mathbf{x}_i) + b + e_i, \quad i = 1, \dots, M \quad (3)$$

where γ is called the regularization parameter which seeks an optimal balance between model fit and model complexity, and v_i is the weight on the i^{th} measurement. The choice of the weight, v_i , in (2) determines the contribution of each observation to the final function estimation. Suykens *et al.* [7] suggest giving minimal weights for the observations with large errors (i.e., outliers), and propose the following rule [19], [20]:

$$v_i = \begin{cases} 1 & \text{if } e_i/\hat{s} \leq c_1 \\ \frac{c_2 - |e_i/\hat{s}|}{c_2 - c_1} & \text{if } c_1 \leq e_i/\hat{s} \leq c_2 \\ 10^{-4} & \text{if } e_i/\hat{s} > c_2. \end{cases} \quad (4)$$

Here, \hat{s} is an estimate of the standard deviation of e_i , which is given by $IQR/(2 \times 0.6745)$. The interquartile range, IQR , is the difference between the 25th and 75th percentile. The typical choice for c_1 and c_2 are 2.5 and 3, respectively.

To obtain a robust function estimation, we apply an iterative procedure [7] as follows: in the first step, unweighted LS-SVR is applied to the dataset, S , i.e., solving (2)–(3) with equal weights (e.g., $v_i = 1, \forall i$). Next, the weight for each observation is determined using the rules in (4) based on the errors, $e_i, i = 1, \dots, M$. Then, the weighted LS-SVR is applied to the data with the associated weights. The procedure repeats until the change of weights in consecutive iterations is negligible. This iterative procedure enables us to construct a robust regression function with uncontaminated data by excluding the outlier effects.

B. Solving Iterative Weighted LS-SVR

The computational procedure for solving the weighted LS-SVR in (2)–(3) is equivalent to solving the following linear

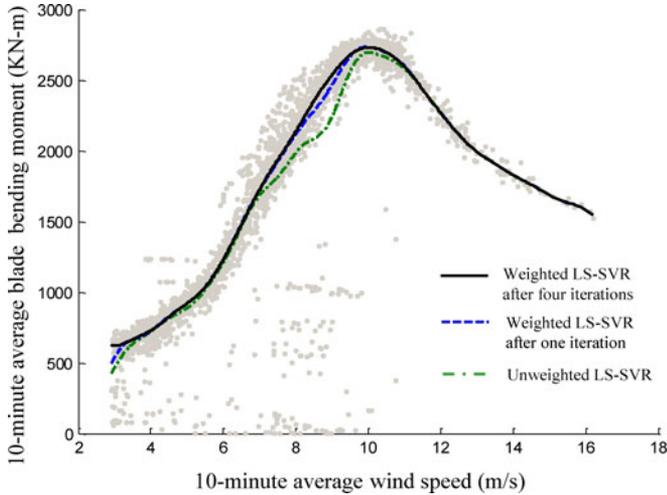


Fig. 2. Comparison of the estimated baselines from LS-SVR and weighted LS-SVR for blade bending moment.

system (see appendix A or [7] for details):

$$\begin{bmatrix} 0 & 1_M^T \\ 1_M & \mathbf{K} + V_\gamma \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix}. \quad (5)$$

Here, $Y = [y(\mathbf{x}_1), \dots, y(\mathbf{x}_M)]^T$. 1_M is an $(M \times 1)$ vector whose components are 1, and $\hat{\alpha} = [\hat{\alpha}_1, \dots, \hat{\alpha}_M]^T$ is a vector of Lagrangian multipliers. \mathbf{K} is an $(M \times M)$ kernel matrix whose $(i, j)^{th}$ component is $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$. $k(\cdot, \cdot)$ is called a kernel function. There are several kernel functions available, including the radial basis, polynomial, and spline functions. V_γ is an $(M \times M)$ diagonal matrix whose $(i, i)^{th}$ component is $\frac{1}{\gamma v_i}$, i.e., $V_\gamma = \text{diag}\{\frac{1}{\gamma v_1}, \dots, \frac{1}{\gamma v_M}\}$.

By solving (5), we obtain [21]

$$\hat{\alpha} = (\mathbf{K} + V_\gamma)^{-1} (Y - 1_M \hat{b}), \quad (6)$$

$$\hat{b} = \frac{1_M^T (\mathbf{K} + V_\gamma)^{-1} Y}{1_M^T (\mathbf{K} + V_\gamma)^{-1} 1_M}, \quad (7)$$

where $1_M = [1, \dots, 1]^T$.

Then, the estimated regression function at an input vector, \mathbf{x} , is given by

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^M \hat{\alpha}_i k(\mathbf{x}, \mathbf{x}_i) + \hat{b} \quad (8)$$

and the error is $e_i = e(\mathbf{x}_i) = y(\mathbf{x}_i) - \hat{y}(\mathbf{x}_i)$, $i = 1, \dots, M$. Following the rules in (4), these errors determine the weights, v_i 's, in the next iteration.

Suykens *et al.* [7] claim that after solving the unweighted LS-SVR, a single additional weighted LS-SVR iteration is generally sufficient for excluding the outlier effects. In this study, we perform iterations until the change of v_i 's in consecutive iterations is negligible. Specifically, we run iterations until $\max(|v_i^{\text{prev}} - v_i^{\text{current}}|, i = 1, \dots, M) < 0.5$, where v_i^{prev} and v_i^{current} are the weights of the i^{th} observation at the previous and current iterations, respectively. Fig. 2 compares the regression curve for the blade bending moment from the weighted

LS-SVR and from the unweighted LS-SVR using the 2500 records in the training dataset described in Section III-A. Note that the regression line from the unweighted LS-SVR is affected by many outliers in the lower part and shifts downwards. On the contrary, an additional iteration with weighted LS-SVR removes the influence of the outliers significantly, and after four to five iterations, we obtain a smooth curve that characterizes the normal behavior in the blade response and power generation.

In our implementation, Gaussian Kernel function is employed. The regularization parameter, γ , in (2) and hyperparameter in Gaussian Kernel function are optimized by using cross-validation techniques [13]. When a data size is large, finding optimal parameters and building the baseline regression function through the aforementioned iterative procedure could be computationally challenging. The incomplete Cholesky matrix approximation method can reduce the computational burden when it is used in conjunction with the matrix inverse lemma (see appendix B for the detail).

C. Estimation of Heterogeneous Variance

In order to determine decision boundaries, we need to estimate the variance of output, y , to capture the variability of turbine responses under the normal conditions. As can be seen in Fig. 1, the variability of a turbine's response manifests differently under different wind speeds, called heteroscedasticity. For instance, the variability of blade bending moments tends to be small at high wind speeds. We estimate the heterogeneous variance of the output, y , by smoothing the squared errors, following the procedure in [21]. Specifically, we apply the LS-SVR to the dataset $S' = \{(e_i^2, \mathbf{x}_i), i = 1, \dots, M\}$, where $e_i = y(\mathbf{x}_i) - \hat{y}(\mathbf{x}_i)$ is the i^{th} error from the final regression line. The detailed procedure is given in appendix C.

D. Comparison with Linear Models

We compare the proposed nonlinear regression with a linear regression in constructing a baseline curve. For a fair comparison, we employ an iterative reweighted least square (IRLS) robust regression that is used for removing the outlier effects in the linear regression analysis [22]. IRLS iteratively finds regression coefficients that minimize a weighted least squares criterion, and at each iteration, the weight of each observation is determined by the error. We can see that the overall idea of IRLS is similar to the iterative weighted LS-SVR. The primary difference is that to estimate a regression function, IRLS needs to pre-specify a linear function type, whereas LS-SVR does not have this limitation. In this comparison, we use polynomial functions with a bisquare weighting.

Fig. 3(a) shows the estimated regression line from IRLS with the sixth order polynomial function using the blade bending moment data. The polynomial function seems to provide a reasonably good fit. However, a closer look indicates that the fit is less accurate at wind speeds over 12 m/s. In order to see whether the polynomial function can generate a good regression curve over a full range of wind speeds, we need data for the entire operating conditions. However, the dataset lacks measurements under high wind speeds. Therefore, we simulate data for a high

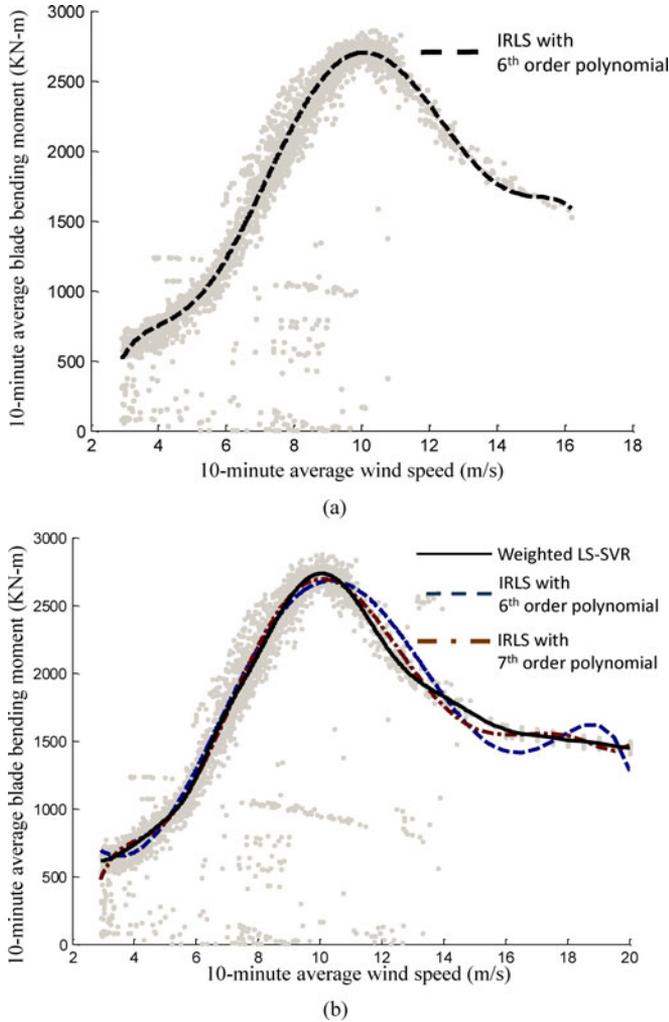


Fig. 3. Comparison of weighted LS-SVR and linear regressions for blade bending moment. (a) Estimated baseline with sixth order polynomial regression using the original training dataset. (b) Estimated baselines using the augmented dataset.

wind regime. We note that the blade bending moment in a flap-wise direction slowly declines at high wind speeds due to a pitch control to protect blades. Based on this understanding, we simulate blade bending moments in a gradually decreasing pattern at 16–20 m/s. Then, we augment the simulated data to the original training dataset, and apply both IRLS with polynomial functions and the weighted LS-SVR to the augmented dataset. Fig. 3(b) suggests that the sixth order polynomial function does not provide a good fit at high wind speeds. The seventh order polynomial function provides an improved fit, but the estimation at 14–16 m/s deviates from the observations. With the augmented dataset, the weighted LS-SVR still generates a better regression curve that captures the nominal pattern of blade bending moments.

In addition, linear regression does a poor job of estimating the heterogeneous variance when the variance presents complicated patterns. When the variance of outputs shows specific patterns, we can estimate the variance by regressing the squared residuals on carefully selected explanatory variables [22]. This procedure

again requires a predefined function (e.g., assuming squared residuals linearly increase in x). Overall, linear model-based approaches fail to capture the fine local features and the changes in variability over a range of wind conditions.

V. REGRESSION CONTROL CHART

The regression line obtained in the previous sections serves as a baseline function in setting the control limits that separate in-control (i.e., normal) conditions from out-of-control (i.e., abnormal) conditions. The resulting control chart is called a “regression control chart” or “functional data control chart” [23]. We propose two methods to define the control limits.

A. Residual Control Chart

A residual is the difference between the observed output and the model’s predicted output, i.e., $r(\mathbf{x}) = y(\mathbf{x}) - \hat{y}(\mathbf{x})$ (note that $r(\mathbf{x})$ is essentially the same as $e(\mathbf{x})$; here, we use $r(\mathbf{x})$, instead of $e(\mathbf{x})$, to distinguish the error in the *monitoring* phase from the error in the training phase). A large absolute value of residuals will indicate a possible abnormal situation. A well-fitting regression curve provides a good prediction, $\hat{y}(\mathbf{x})$, for a future response, $y(\mathbf{x})$, under normal conditions, so $|r(\mathbf{x})|$ is small. On the contrary, when a regression model is not properly selected, large $|r(\mathbf{x})|$ ’s often will be generated during normal operations, causing frequent false alarms.

To set the limit values to distinguish normal and abnormal conditions, we need to find the distribution of the residual. Using a similar technique as in [21], we note that the predicted output at an input vector, \mathbf{x} , can be expressed as a linear combination of the past outputs in the training dataset:

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^M \hat{\alpha}_i k(\mathbf{x}, \mathbf{x}_i) + \hat{b} \quad (9)$$

$$= \mathbf{K}_x^T \hat{\alpha} + \hat{b}, \quad (10)$$

$$= \mathbf{K}_x^T Z \left(Y - \frac{1}{M} \frac{1_M^T Z Y}{1_M^T Z 1_M} \right) + \frac{1}{M} \frac{1_M^T Z Y}{1_M^T Z 1_M} \quad (11)$$

$$= \left[\mathbf{K}_x^T Z \left(I - \frac{1}{M} \frac{1_M^T Z}{1_M^T Z 1_M} \right) + \frac{1}{M} \frac{1_M^T Z}{1_M^T Z 1_M} \right] Y \quad (12)$$

$$= L(\mathbf{x})^T Y \quad (13)$$

where $\mathbf{K}_x = [k(\mathbf{x}, \mathbf{x}_1), \dots, k(\mathbf{x}, \mathbf{x}_M)]^T$ and $Z = (\mathbf{K} + V_\gamma)^{-1}$. $L(\mathbf{x}) = \left[\mathbf{K}_x^T Z \left(I - \frac{1}{M} \frac{1_M^T Z}{1_M^T Z 1_M} \right) + \frac{1}{M} \frac{1_M^T Z}{1_M^T Z 1_M} \right]^T$ is called a linear smoother [21].

As pointed out in [21], $\hat{y}(\mathbf{x})$ is a biased estimator of the mean response, $E(y)$, where $E(\cdot)$ denotes an expectation. That is, $E[\hat{y}(\mathbf{x})] \neq E[y(\mathbf{x})]$, but $E[\hat{y}(\mathbf{x})] = E[y(\mathbf{x})] + \text{bias}(\mathbf{x})$. Brabanter *et al.* [21] suggest estimating the bias with $\text{bias}(\mathbf{x}) = \sum_{i=1}^M l_i(\mathbf{x}) \hat{y}(\mathbf{x}_i) - \hat{y}(\mathbf{x})$, where $l_i(\mathbf{x})$ is the i^{th} component of $L(\mathbf{x})$ in (13). Therefore, the bias-corrected estimator of the mean response, $E[y(\mathbf{x})]$, is given by

$$\hat{y}_c(\mathbf{x}) = \hat{y}(\mathbf{x}) - \widehat{\text{bias}}(\mathbf{x}) = 2\hat{y}(\mathbf{x}) - \sum_{i=1}^M l_i(\mathbf{x}) \hat{y}(\mathbf{x}_i). \quad (14)$$

Using this, we define the following bias-corrected residual whose expectation is approximately zero:

$$r_c(\mathbf{x}) = y(\mathbf{x}) - \hat{y}_c(\mathbf{x}). \quad (15)$$

To find the variance of $r_c(\mathbf{x})$, we assume that the bias-corrected predicted output, $\hat{y}_c(\mathbf{x})$, which is obtained from the historical dataset, is independent of the future output, y . We estimate the variance of $r_c(\mathbf{x})$ with $\hat{\sigma}^2(\mathbf{x}) + \widehat{Var}(\hat{y}_c(\mathbf{x}))$, where $\hat{\sigma}^2(\mathbf{x})$ is the estimated variance of the output, y , at an input vector, \mathbf{x} . $\widehat{Var}(\hat{y}_c(\mathbf{x}))$ is the estimated variance of $\hat{y}_c(\mathbf{x})$, given by $\widehat{Var}(\hat{y}_c(\mathbf{x})) = \sum_{i=1}^M l_i(\mathbf{x})^2 \hat{\sigma}^2(\mathbf{x}_i)$ (see the Appendix for more details).

To monitor the process, we use the average of multiple residuals. Suppose we observe N_r observations, $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_{N_r}, y_{N_r})$. We assume that the outputs are independent of one another. Applying the central limit theorem, the average of N_r residuals, $r_c(\mathbf{x}_1), \dots, r_c(\mathbf{x}_{N_r})$, will asymptotically follow Normal distribution under certain regularity conditions [21], [24]:

$$\bar{r}_c = \frac{1}{N_r} \sum_{i=1}^{N_r} r_c(\mathbf{x}_i) \quad (16)$$

$$\sim N \left(0, \frac{1}{N_r^2} \sum_{i=1}^{N_r} \left(\hat{\sigma}^2(\mathbf{x}_i) + \widehat{Var}(\hat{y}_c(\mathbf{x}_i)) \right) \right). \quad (17)$$

Therefore, given the significance level, α (or, type I error, false alarm rate), we define the upper and lower control limits (UCL and LCL) for the average residual, \bar{r}_c , as follows:

$$\text{UCL} = +z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{N_r^2} \sum_{i=1}^{N_r} \left(\hat{\sigma}^2(\mathbf{x}_i) + \widehat{Var}(\hat{y}_c(\mathbf{x}_i)) \right)}, \quad (18)$$

$$\text{LCL} = -z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{N_r^2} \sum_{i=1}^{N_r} \left(\hat{\sigma}^2(\mathbf{x}_i) + \widehat{Var}(\hat{y}_c(\mathbf{x}_i)) \right)} \quad (19)$$

where $z_{1-\frac{\alpha}{2}}$ is the $100(1 - \frac{\alpha}{2})$ quantile of the standard normal distribution. The typical choice of $z_{1-\frac{\alpha}{2}}$ is 3, corresponding to $\alpha = 0.0027$. This type of chart, which is called a three-sigma control chart, is the standard control chart in many industries. Here, $\alpha = 0.0027$ implies that 0.27% false alarm rate is allowed in each monitoring of \bar{r}_c .

When the average residual, \bar{r}_c , falls within the control limits, we assume that the system is performing normally. The capability of detecting abnormal conditions will depend on the monitoring size, N_r . See Sections VI and VII for a comparison of the performance of control charts with different monitoring sizes.

Note that the control limits in (18) and (19) differ from those in the typical control charts in SPC. In general, Shewhart type control charts decide the control limits using the historical dataset in the training phase, and the same, fixed control limits are used in the monitoring phase. However, recall again that the prediction accuracy and variability of the turbine output will vary across weather conditions; thus, the control limits in (18) and (19) reflect the heterogeneity in the output variance. As a

result, the control limits are determined in the monitoring phase, depending on the input weather conditions.

B. Response Control Chart

The second approach is to specify the control limits for the output, y . The techniques used here follow Brabanter *et al.* [21] who derive prediction intervals for the LS-SVR model. We extend their techniques to the weighted LS-SVR with the incomplete Cholesky matrix approximation method, and use the prediction interval for determining the control limits.

Recall that $\hat{y}_c(\mathbf{x})$ in (14) is an approximated, unbiased estimator of the output, $y(\mathbf{x})$, and that the variance of $y(\mathbf{x}) - \hat{y}_c(\mathbf{x})$ is estimated by $\hat{\sigma}^2(\mathbf{x}) + \widehat{Var}(\hat{y}_c(\mathbf{x}))$. We use these results to obtain the approximate $100(1 - \beta)$ pointwise interval for a new output, y , at an input vector, \mathbf{x} , as follows:

$$\hat{y}_c(\mathbf{x}) \pm z_{1-\frac{\beta}{2}} \sqrt{\hat{\sigma}^2(\mathbf{x}) + \widehat{Var}(\hat{y}_c(\mathbf{x}))}. \quad (20)$$

The interval in (20) implies that an individual response approximately falls into the interval with $100(1 - \beta)\%$ probability. Suppose that we observe multiple outputs and use them together to decide whether or not the system is in-control. Assuming independence among turbine outputs, the probability that all of N_y observations are within the interval is $(1 - \beta)^{N_y}$. Setting this probability equal to $(1 - \alpha)$ gives $\beta = 1 - (1 - \alpha)^{1/N_y}$ where α is the significance level. This interval is called a simultaneous prediction interval [21]. We use this interval as control limits. In this case, an alarm triggers when any response among the N_y observations is outside this boundary. Here, when the response takes nonnegative values, the lower control limit for the response is set to be $\max\{\hat{y}_c(\mathbf{x}) \pm z_{1-\frac{\beta}{2}} \sqrt{\hat{\sigma}^2(\mathbf{x}) + \widehat{Var}(\hat{y}_c(\mathbf{x}))}, 0\}$.

Note that the interval in (20) provides the uncertainty (or, variability) of the turbine output [25]. Similar to the residual control chart, the control limits vary, depending on the input vector, and they are determined in the monitoring phase.

VI. CASE STUDY 1

We apply the proposed method to the dataset of 3 010 records described in Section III. We use 2 500 records to build our baseline regression line and the remaining 510 records to test the monitoring performance. This section discusses the monitoring results from the residual control chart and response control chart for each response variable using the 510 monitoring records. Fig. 4 depicts the scatter plots in the monitoring dataset (in the plot, we mark the observation numbers of the records that seem to be abnormal).

A. Residual Control Chart

We use three-sigma control charts to monitor the average residuals. To examine the effects of different monitoring size N_r , we consider two cases; one with $N_r = 30$ and the other with $N_r = 100$. Figs. 5–7 depict the control charts for power generation and blade bending moment. The middle points in each chart are the average residuals, \bar{r}_c . For example, the first

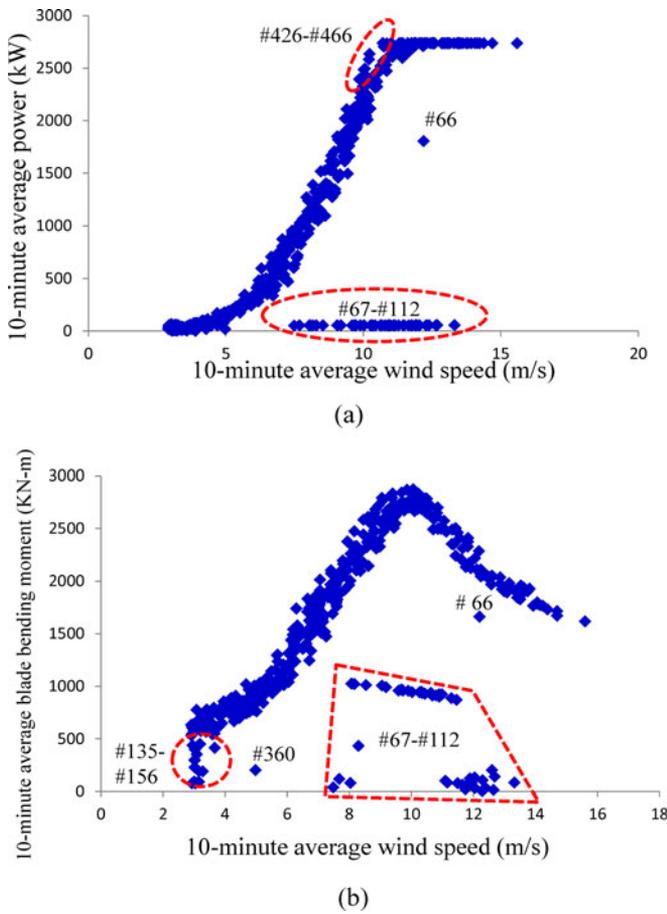


Fig. 4. Scatter plots in the monitoring dataset. (a) Power generation. (b) Blade bending moment.

point in Fig. 5(a) is the average of the first 30 residuals. The points on the upper and lower lines in each figure describe the upper and lower control limits (that is, UCL and LCL in (18)–(19)), respectively.

In Fig. 4(a), the abnormal data are observed in Nos. 66–112 records in the monitoring dataset. These abnormal responses are detected in the third and fourth points in Fig. 5(a) with $N_r = 30$ and the first and second points in Fig. 5(b) with $N_r = 100$. In Nos. 426–466 records, consistently higher power above the baseline curve [see Fig. 8(c)] has been observed around 10 m/s. These observations make the 15th and 5th average residuals in Fig. 5(a) and (b) lie above their upper control limits, respectively. Fig. 6, an enlarged version of Fig. 5(a), shows that the control limits are not fixed, but vary, depending on the input values since the variances of the responses depend on the input vector.

Similarly, Fig. 7 demonstrates that the abnormal observations are detected for the blade bending moment. Both control charts with $N_r = 30$ and $N_r = 100$ identify the abnormal observations in Nos. 66–112 and 135–156. Also, due to the consistently higher moments in Nos. 361–390 over the baseline curve [see Fig. 9(c)], the 13th average residual is out of upper control limit in Fig. 7(a). However, these observations are not detected in the residual chart with $N_r = 100$ in Fig. 7(b) since the majority

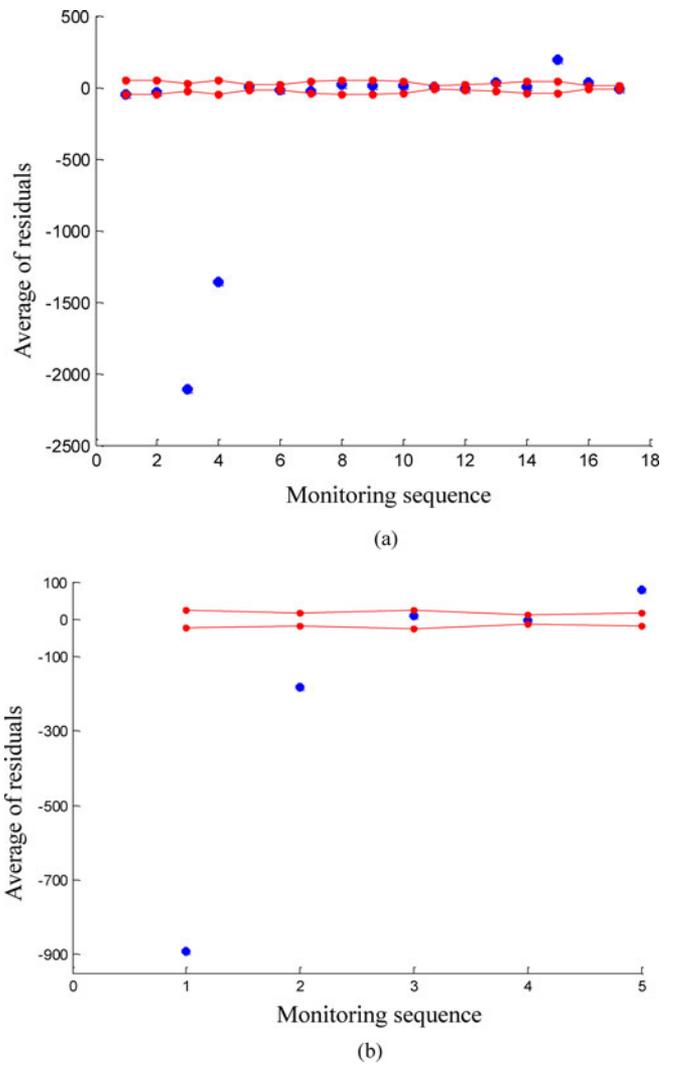


Fig. 5. Residual control chart for power generation. (a) $N_r = 30$. (b) $N_r = 100$.

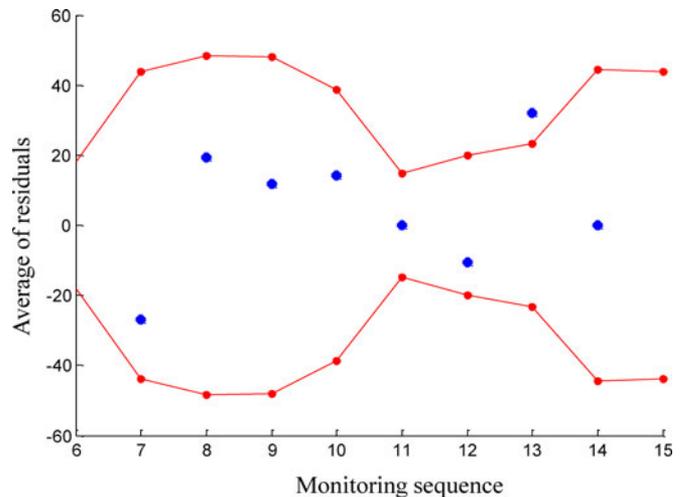


Fig. 6. Enlarged residual control chart for power generation with $N_r = 30$.

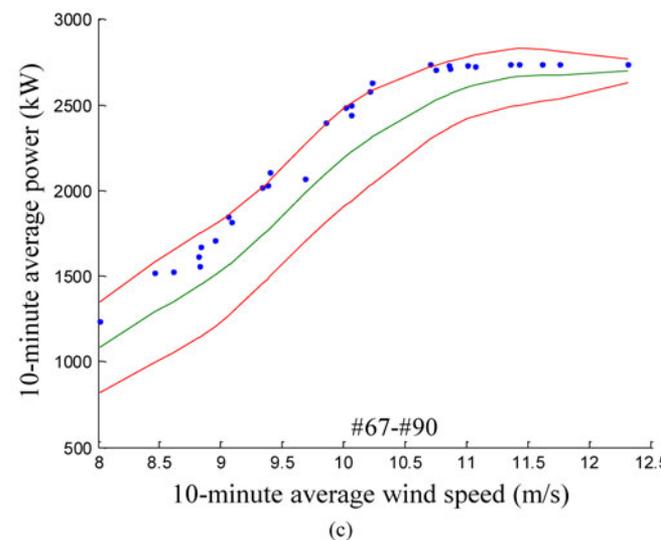
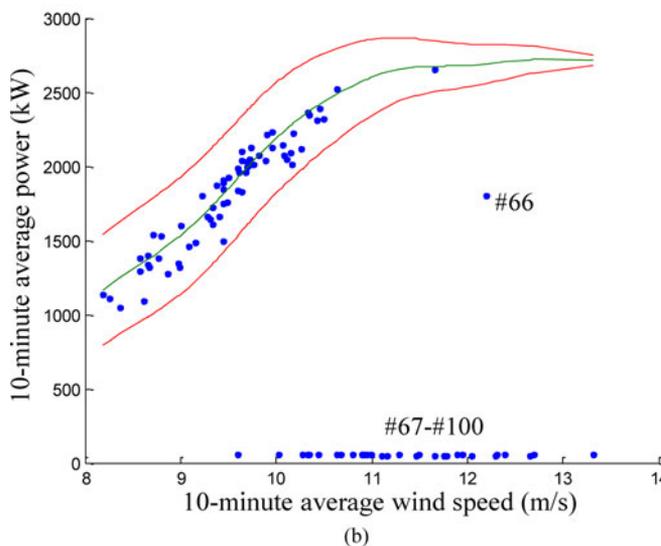
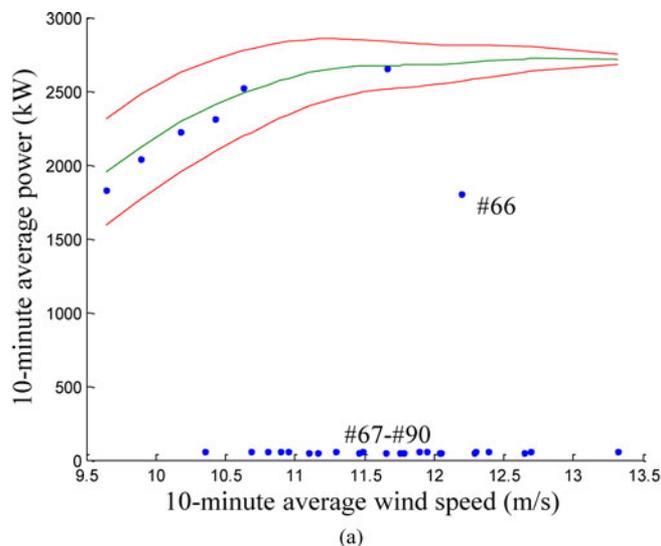
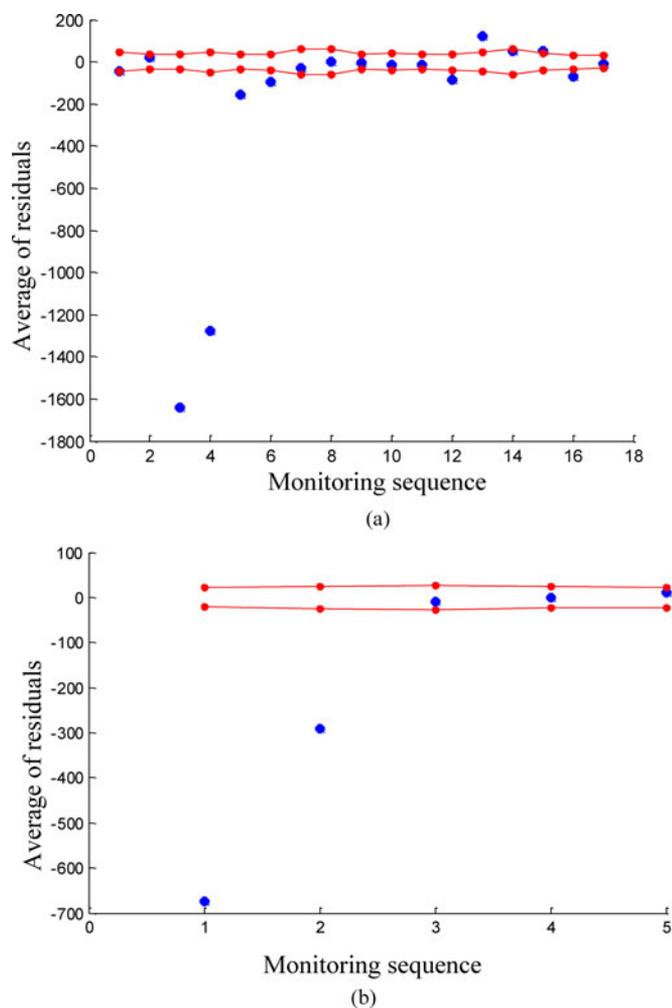


Fig. 7. Residual control chart for blade bending moment. (a) $N_r = 30$. (b) $N_r = 100$.

of the observations in Nos. 301–400 are in normal condition and accordingly, the effects of a small number of abnormal observations are averaged out. We can see that usage of a small sample size, N_r , leads to high detection power, but we should also note that it could cause frequent false alarms. As such, a careful decision would be necessary, considering the tradeoff between the miss-detection and false alarm rates.

B. Response Control Chart

To generate the response control charts, we use $\alpha = 0.0027$. Fig. 8 shows control charts for power generation in the monitoring dataset. As discussed earlier, the width of the control limit varies, depending on the input weather condition, reflecting that the resulting control limits accommodate the inherent variability of the response. For example, Fig. 8 shows that the control limits are narrow at high wind speeds due to a small variability in power generation [see Fig. 1(a)]. The anomalies shown in Fig. 4(a) are detected in the response control charts. Recall

Fig. 8. Response control charts for power generation. (a) Nos. 61–90 observations. (b) Nos. 1–100 observations. (c) Nos. 421–450 observations.

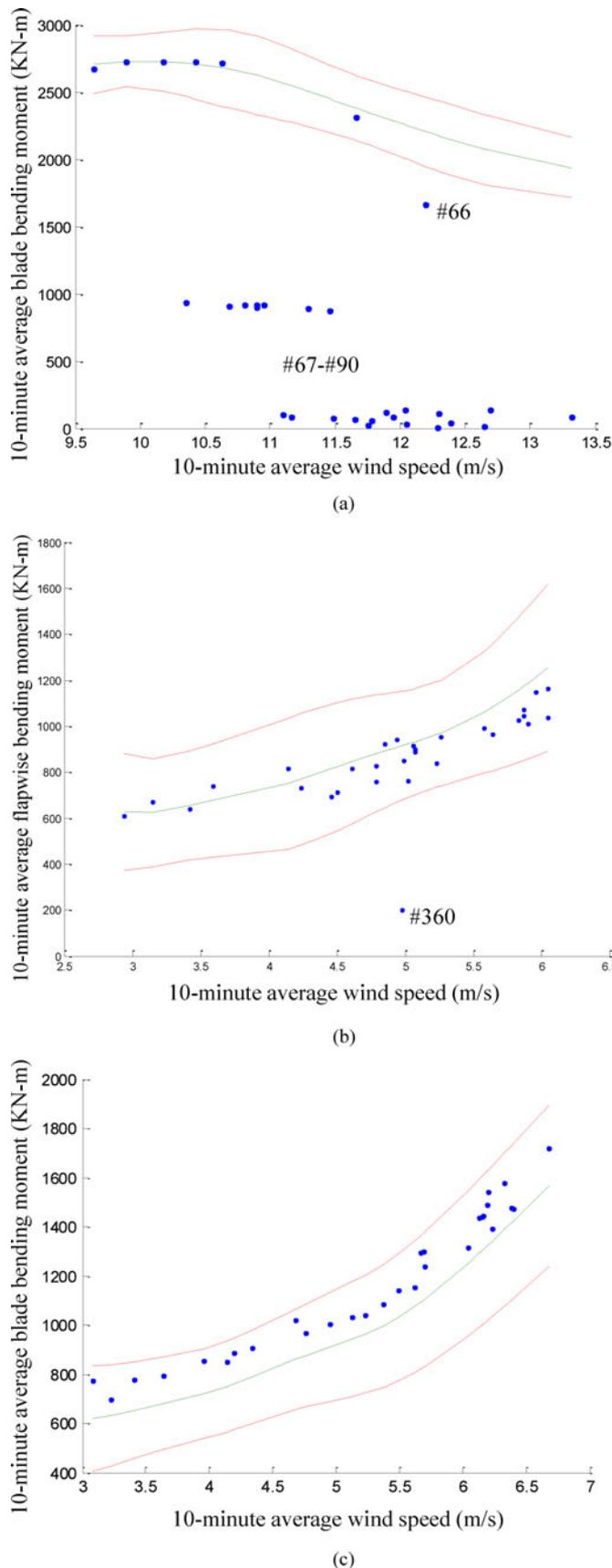


Fig. 9. Response control charts for blade bending moment. (a) Nos. 61–90 observations. (b) Nos. 331–360 observations. (c) Nos. 361–390 observations.

that in the residual control chart, the detection capability of abnormal observations is sensitive to the selection of the sample size, N_r . In the response control chart, the choice of N_y does not considerably influence the detection power. With $N_y = 100$, the abnormal outputs are also well identified. Fig. 8(c) shows that all of the power outputs in Nos. 421–450 are above the baseline, and some of them are beyond the control limits.

Similarly, Fig. 9 illustrates that the anomalies in blade bending moments are detected in the response control charts. Response No. 360, which was not detected in the residual control charts in Fig. 7, is identified as an abnormal case in Fig. 9(b). In Fig. 9(c), even though the bending moments are between the upper and lower control limits, all of them lie above the baseline, indicating a possible fault.

C. Implication of Abnormality

Abnormal outputs are likely to be generated in transient operations. However, since Nos. 66–112 records correspond to eight hours of operation, we believe these abnormal data are collected during nontransient operations.

Modern turbines employ control mechanisms such as pitch and torque controls to regulate turbine operations [26]–[28]. In Region 2 (from cut-in to rated wind speeds), the goal of the control operation is to capture as much power from the wind as possible, whereas in Region 3 (from rated to cut-out wind speeds), a variable-speed turbine sheds additional power and maintains a rated power by pitching its blades [27]. In Figs. 8–9, where Nos. 66–112 (and some observations among Nos. 421–450) are identified as abnormal, these observations occur at around the rated wind speed. One possibility is that such anomalies occur when the controllers switch their operations from Region 2 to Region 3, indicating unstable control operations.

VII. CASE STUDY 2

We validate the presented approach using the dataset collected from the CART2 turbine at NREL. When some damage progresses in a gearbox, the response from the high-speed shaft (HSS), which connects a gearbox and a generator, can change. We consider the ratio of HSS torque to HSS RPM (hereafter, we call it HSS ratio) as a response variable, and investigate the relationship between the HSS ratio and wind speed using the one-minute average of measurements. Fig. 10 shows that when the gearbox was damaged, the HSS ratio was lower at low to mid wind speed range (i.e., 3–10 m/s) compared to the ratio during normal operations.

We build the baseline curve using the data collected during normal operations and monitor the abnormal data with the regression control charts. The residual charts in Fig. 11 present most data points outside the lower control limits, indicating the damage progression. The average residual inside the decision boundaries in Fig. 11 is obtained at mid to high wind speeds where the HSS ratios show a similar pattern to that during normal operations. The response control charts in Fig. 12 also identify the abnormal HSS ratios. As seen in Section VI, the width of decision boundaries vary in both types of control charts due to

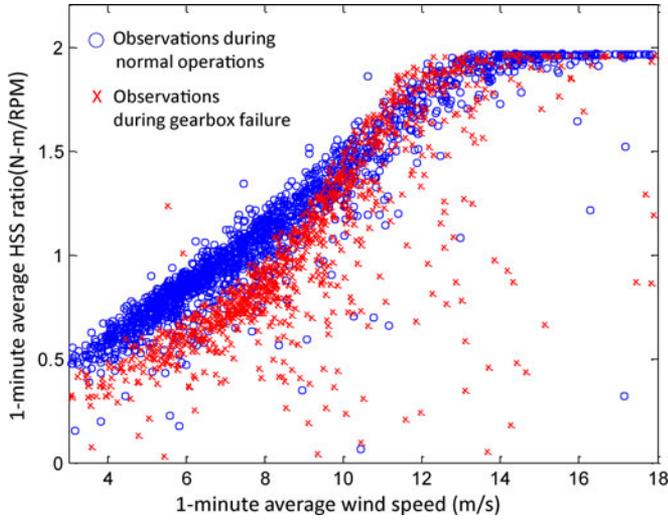


Fig. 10. Observations from the CART2 gearbox.

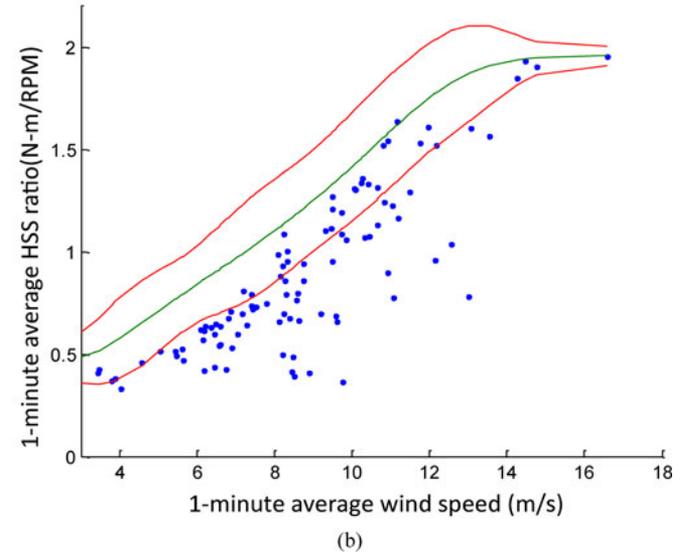
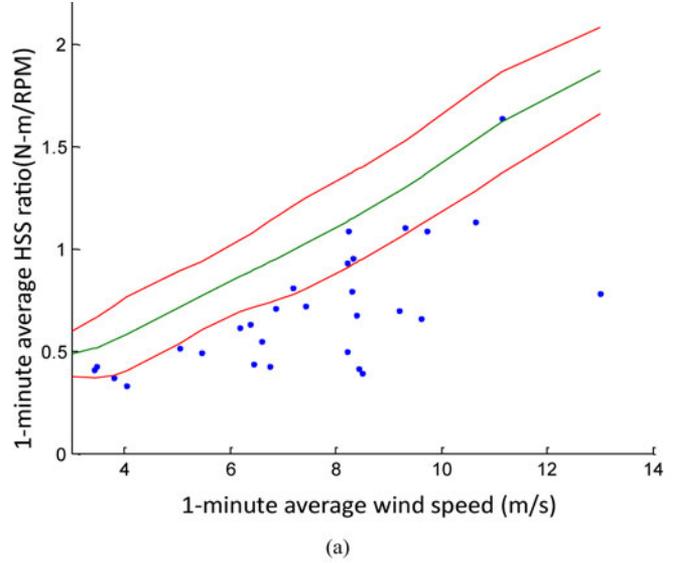


Fig. 12. Response control charts for HSS ratio. (a) $N_y = 30$ (low-to-mid wind speed). (b) $N_y = 100$.

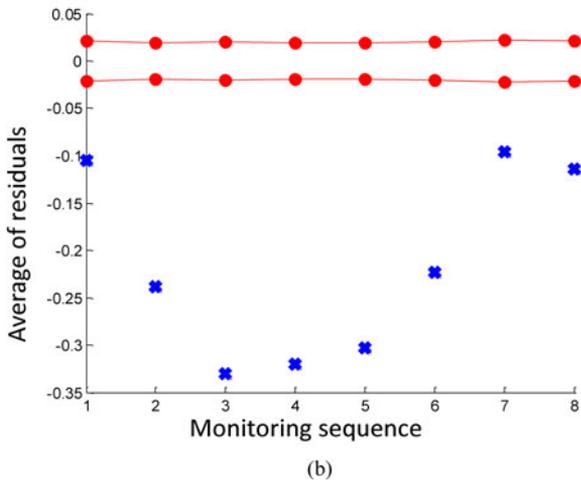
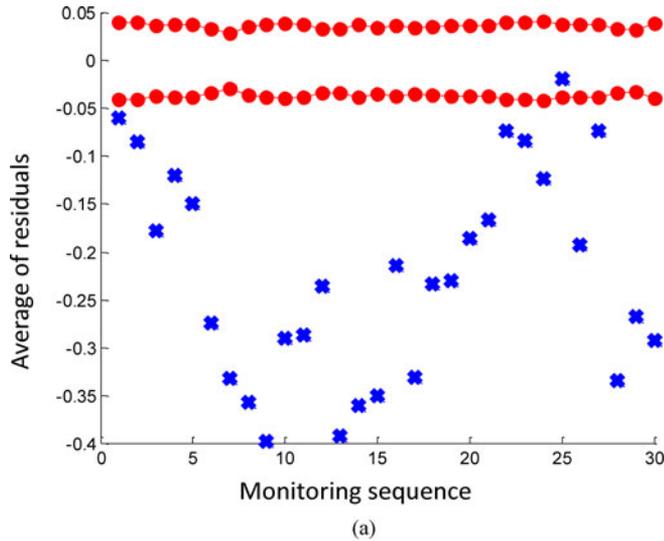


Fig. 11. Residual control charts for HSS ratio. (a) $N_r = 30$. (b) $N_r = 100$.

the heterogeneous variability of the HSS ratios over a range of wind speeds.

VIII. CONCLUSION

This study proposes a regression model-based online condition monitoring procedure for wind power systems. We employ a nonparametric regression model to characterize the baseline relationship between turbine responses and weather variables. To remove the effects of outliers, we propose to use a weighted LS-SVR that finds the nominal baseline curve. We construct two control charts by considering the variability of turbine response over a range of wind speeds. Unlike traditional control charts, the control limits in our method vary with the input weather conditions. Consequently, the decision boundaries are determined *on the fly* in the monitoring phase.

Real case studies show that the proposed method is an effective tool to monitor the power generation and the components of turbines. Our analysis demonstrates that the response control chart has a strong capability for identifying abnormal signals individually without delayed decision. As such, the information from the response control charts can be used for short-term monitoring (e.g., control system monitoring). One caveat is that the use of response control charts in damage detection can cause frequent false alarms. Although the residual control chart with a large monitoring size may fail to capture a small number of abnormality incidents, it still provides a robust monitoring capability with fewer false alarms. Anomalies consistently observed for long-term periods in the residual control charts would indicate subsystem damage (as seen in Case Study 2). Also, the residual control charts with the small monitoring size could be used for both short- and long-term monitoring.

Future research will extend the proposed method to incorporate the correlation of multiple responses, rather than analyzing individual responses, in order to improve detection capability. We believe that this research extension will make condition monitoring of wind turbines even more effective.

APPENDIX

A. Derivation of the solution for weighted LS-SVR

We solve the optimization problem in (2) and (3) by defining the Lagrangian multipliers, $\hat{\alpha}_i$'s:

$$Q = \frac{1}{2}w^T w + \frac{1}{2}\gamma \sum v_i e_i^2 - \sum_{i=1}^M \hat{\alpha}_i (\omega^T \phi(\mathbf{x}_i) + b + e_i - y_i).$$

By taking the partial derivatives of Q with respect to ω , b , e_i , and $\hat{\alpha}_i$, $i = 1, \dots, m$, and letting the derivatives be zeros, we can obtain the linear system in (5).

B. Incomplete Cholesky Matrix Approximation

The major computational burden in applying the weighted LS-SVR to a large dataset (i.e., M is large) results from the need to compute the inverse of $(\mathbf{K} + V_\gamma)$ in (6)–(7), which requires $O(M^3)$ computational efforts. The incomplete Cholesky decomposition allows us to approximate the original $M \times M$ kernel matrix, \mathbf{K} , as a low rank matrix, $\mathbf{K} \approx \mathbf{G}\mathbf{G}^T$, where \mathbf{G} is the $M \times N$ matrix with $N \ll M$ [12], [13]. Then, we can apply the matrix Inverse Lemma, also known as the *Sherman–Morrison–Woodbury* formula [13], [29], to find the approximation of $(\mathbf{K} + V_\gamma)^{-1}$, as follows:

$$(\mathbf{K} + V_\gamma)^{-1} \approx (\mathbf{G}\mathbf{G}^T + V_\gamma)^{-1} \quad (21)$$

$$= V_\gamma^{-1} - V_\gamma^{-1} \mathbf{G} (I_N + \mathbf{G}^T V_\gamma^{-1} \mathbf{G})^{-1} \mathbf{G}^T V_\gamma^{-1} \quad (22)$$

where I_N is an $N \times N$ identity matrix.

Let Z denote the $(\mathbf{K} + V_\gamma)^{-1}$ matrix. That is, $Z \approx V_\gamma^{-1} - V_\gamma^{-1} \mathbf{G} (I_N + \mathbf{G}^T V_\gamma^{-1} \mathbf{G})^{-1} \mathbf{G}^T V_\gamma^{-1}$. By replacing $(\mathbf{K} + V_\gamma)^{-1}$ in (6) and (7) by Z , $\hat{\alpha}$, and \hat{b} can be approximated as

follows:

$$\hat{\alpha} \approx Z(Y - 1_M b) \quad (23)$$

$$\hat{b} \approx \frac{1_M^T Z Y}{1_M^T Z 1_M}. \quad (24)$$

Then, we use $\hat{\alpha}$ and \hat{b} in (23) and (24) to find the regression function, $\hat{y} = \sum_{i=1}^M \hat{\alpha}_i k(\mathbf{x}, \mathbf{x}_i) + \hat{b}$.

We use the cross-validation technique to find the optimal regularization parameter, γ . This requires to solve the optimization problem in (2)–(3) with different values of γ . Then, to get a stable regression function, we solve the problem with associated weights from (4) iteratively. Note that \mathbf{G} in (22) is the same, regardless of the choice of γ , but only V_γ changes. Therefore, Z can be obtained with less computational burden, unlike the original procedure which requires us to compute the inverse of the whole matrix $(\mathbf{K} + V_\gamma)$.

C. Variance Derivation

As discussed earlier, to estimate the variance of the output, y , we apply the LS-SVR to the dataset, $S' = \{(e_i^2, \mathbf{x}_i), i = 1, \dots, M\}$. Then, we can get a linear smoother, $L'(\mathbf{x})$, in this new dataset, S' , similar to $L(\mathbf{x})$ in (13). We denote $d(A)$ as the column vector whose i^{th} row is the $(i, i)^{th}$ element of a square matrix, A . By noting that $\mathbf{E}[L'(\mathbf{x})^T d(ee^T)] = \sigma^2(\mathbf{x})(1 + L'(\mathbf{x})d(LL^T - L - L^T))$ [21], the variance estimator at a point, \mathbf{x} , is given by

$$\hat{\sigma}^2(\mathbf{x}) = \frac{L'(\mathbf{x})^T d(ee^T)}{1 + L'(\mathbf{x})^T d(LL^T - L - L^T)}.$$

Here, $e = [e_1, \dots, e_M]^T$ is the vector of errors in the training dataset, and L is an $(M \times M)$ matrix whose i^{th} row is $L(\mathbf{x}_i)^T$, $i = 1, \dots, M$.

With this estimator, the variance of the predicted output, $\hat{y}(\mathbf{x})$, is

$$\begin{aligned} Var(\hat{y}(\mathbf{x})) &= Var(L(\mathbf{x})^T Y) \\ &= Var\left(\sum_{i=1}^M l_i(\mathbf{x})y(\mathbf{x}_i)\right) = \sum_{i=1}^M l_i(\mathbf{x})^2 \sigma^2(\mathbf{x}_i). \end{aligned}$$

Assuming the variance of the bias-corrected estimator, $\hat{y}_c(\mathbf{x})$, is same as $Var(\hat{y}(\mathbf{x}))$, we estimate the variance of $\hat{y}_c(\mathbf{x})$ by $\sum_{i=1}^M l_i(\mathbf{x})^2 \hat{\sigma}^2(\mathbf{x}_i)$ [21].

REFERENCES

- [1] U.S. Department of Energy. (2008). "20 percent wind energy by 2030: Increasing wind energy's contribution to U.S. electricity supply," Washington DC, Tech. Rep., [Online]. Available: <http://www.nrel.gov/docs/fy08osti/41869.pdf>
- [2] North American Electric Reliability Corporation (NERC). (2009). "2009 long-term reliability assessment," U.S. Department of Energy, Princeton, NJ, Tech. Rep., [Online]. Available: <http://www.nerc.com/files/2009-LTRA.pdf>
- [3] J. Nilsson and L. Bertling, "Maintenance management of wind power systems using condition monitoring systems-life cycle cost analysis for two case studies," *IEEE Trans. Energy Convers.*, vol. 22, no. 1, pp. 223–229, Mar. 2007.
- [4] Z. Zhang, A. Verma, and A. Kusiak, "Fault analysis of the wind turbine gearbox," *IEEE Trans. Energy Convers.*, vol. 27, no. 2, pp. 526–535, Jun. 2012.

- [5] N. Pindoriya, P. Jirutitijaroen, D. Srinivasan, and C. Singh, "Composite reliability evaluation using Monte Carlo simulation and least squares support vector classifier," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2483–2490, Nov. 2011.
- [6] R. Salat and S. Osowski, "Accurate fault location in the power transmission line using support vector machine approach," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 979–986, May 2004.
- [7] J. Suykens, J. D. Brabanter, L. Lukas, and J. Vandewalle, "Weighted least squares support vector machines: Robustness and sparse approximation," *Neurocomputing*, vol. 48, no. 3, pp. 85–105, 2002.
- [8] Y. Ding, L. Zeng, and S. Zhou, "Phase I analysis for monitoring nonlinear profiles in manufacturing processes," *J. Quality Technol.*, vol. 38, no. 3, pp. 199–216, 2006.
- [9] P. Caselitz and J. Giebhardt, "Rotor condition monitoring for improved operational safety of offshore wind energy converters," *J. Sol. Energy Eng.*, vol. 127, no. 2, pp. 253–261, 2005.
- [10] A. Kusiak, H. Zheng, and Z. Song, "On-line monitoring of power curves," *Renewable Energy*, vol. 34, no. 6, pp. 1487–1493, 2009.
- [11] H. Tang, S. Xue, R. Chen, and T. Sato, "Online weighted LS-SVM for hysteretic structural system identification," *Eng. Struct.*, vol. 28, no. 3, pp. 1728–1735, 2006.
- [12] S. Fine and K. Scheinberg, "Efficient SVM training using low-rank kernel representation," *J. Mach. Learn. Res.*, vol. 2, pp. 243–264, 2001.
- [13] S. An, W. Liu, and S. Venkatesh, "Fast cross-validation algorithms for least squares support vector machine and kernel ridge regression," *Pattern Recognit.*, vol. 40, no. 8, pp. 2154–2162, 2007.
- [14] B. Stephen, S. Galloway, D. McMillan, D. Hill, and D. Infield, "A copula model of wind turbine performance," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 965–966, May 2011.
- [15] O. Uluyol, G. Parthasarathy, W. Foslien, and K. Kim, "Power curve analytic for wind turbine performance monitoring and prognostics," in *Proc. Annu. Conf. Prognost. Health Manag. Soc.*, Montreal, PQ, Canada, 2011, pp. 1–7.
- [16] W. Yang, P. Tavner, C. Crabtree, and M. Wilkinson, "Cost-effective condition monitoring for wind turbines," *IEEE Trans. Ind. Electron.*, vol. 57, no. 1, pp. 263–271, Jan. 2010.
- [17] Z. Hameed, Y. Hong, Y. Cho, S. Ahn, and C. Song, "Condition monitoring and fault detection of wind turbines and related algorithms: A review," *Renewable Sustainable Energy Rev.*, vol. 13, pp. 1–39, 2009.
- [18] WindData, [Online]. Available: <http://www.winddata.com>, Accessed Jul. 2010.
- [19] H. A. David, "Early sample measures of variability," *Statist. Sci.*, vol. 13, no. 4, pp. 368–377, 1998.
- [20] R. Fletcher, *Practical Methods of Optimization*. New York, NY, USA: Wiley, 1987.
- [21] K. D. Brabanter, J. D. Brabanter, J. A. K. Suykens, and B. D. Moor, "Approximate confidence and prediction intervals for least square support vector regression," *IEEE Trans. Neural Netw.*, vol. 22, no. 1, pp. 110–120, Jan. 2011.
- [22] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, *Applied Linear Statistical Models*, 5th ed. New York, NY, USA: McGraw-Hill/Irwin, 2004.
- [23] L. Shu, K. Tsui, and F. Tsung, "Regression control charts," in *Encyclopedia Statistics in Quality and Reliability*. New York, NY, USA: Wiley, 2008.
- [24] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Pacific Grove, CA, USA: Duxbury, 2002.
- [25] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1845–1856, Nov. 2010.
- [26] D. Gautam, L. Goel, R. Ayyanar, V. Vittal, and T. Harbour, "Control strategy to mitigate the impact of reduced inertia due to doubly fed induction generators on large power systems," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 214–224, Feb. 2011.
- [27] L. Y. Pao and K. E. Johnson, "Control of wind turbines," *IEEE Contr. Syst. Mag.*, vol. 31, no. 2, pp. 44–62, Apr. 2011.
- [28] S. S. Smater and A. D. Dominguez-Garcia, "A framework for reliability and performance assessment of wind energy conversion systems," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2235–2245, Nov. 2011.
- [29] M. A. Woodbury, "Inverting modified matrices," Statistical Research Group, Princeton University, Princeton, NJ, USA, Tech. Rep., 1950.



Nattavut Yampikulsakul received the B.S. degree in computer science with a minor in economics, and the M.S. degree in operations research from Columbia University, New York, NY, USA, in 2008 and 2010, respectively. He is currently working toward the Ph.D. degree in the Department of Industrial & Operations Engineering at the University of Michigan, Ann Arbor, MI, USA.

His research interest includes the fault diagnosis of wind power systems and modeling, analysis and predictions of nonstationary processes.



Eunshin Byon (S'09–M'12) received the B.S. and the M.S. degrees in industrial and systems engineering from the Korea Advanced Institute of Science and Technology, Daejeon, Korea, and the Ph.D. degree in industrial and systems engineering from Texas A&M University, College Station, TX, USA. She is currently an Assistant Professor in the Department of Industrial and Operations Engineering at the University of Michigan, Ann Arbor, MI, USA.

Her research interests include optimizing operations and management of wind power systems, data analytics, quality and reliability engineering, and discrete-event simulation. She is a Member of the IEEE, IIE, and INFORMS.



Shuai Huang received the B.S. degree in statistics from the University of Science and Technology of China, Anhui, China, in 2007, and the Ph.D. degree in industrial engineering from the Arizona State University, Phoenix, AZ, USA, in 2012.

He is currently an Assistant Professor in Industrial and Management Systems Engineering at University of South Florida, Tampa, FL, USA. His research interest is statistical learning and data mining with applications in healthcare and manufacturing.



Dr. Shuangwen Sheng (S'03–M'08) received the B.S. and M.S. degrees both in electrical engineering and the Ph.D. degree in mechanical engineering.

He is currently a Senior Engineer with the National Renewable Energy Laboratory (NREL), Golden, CO, USA. Shawn is currently leading wind turbine condition monitoring, gearbox reliability database, and wind plant operation and maintenance research at NREL. He has a broad range of experiences, including mechanical and electrical system modeling and analysis, sensing and sensor placement, signal processing, defect classification and level evaluation, prognosis, and multiscale modeling and control.



Mingdi You received the B.S. degree in mathematics from Southeast University, Nanjing, China, in 2011, and the M.S. degree in Industrial & Operations Engineering from the University of Michigan, Ann Arbor, MI, USA, in 2013. He is currently working toward the Ph.D. degree in the Department of Industrial & Operations Engineering at the University of Michigan.

His research interest includes applied statistics and spatio-temporal data analytics and reliability engineering with applications on wind energy.