Wind Energy

WIND ENERGY
Wind Energy 2015; **18**:1207–1219
Published online 21 April 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/we.1755

RESEARCH ARTICLE

A kernel plus method for quantifying wind turbine performance upgrades

Giwhyun Lee[1], Yu Ding[2,3], Le Xie[3] and Marc G. Genton[4]

[1] Korea Army Academy, Yeongcheon, Korea
[2] Department of Industrial and Systems Engineering, Texas A&M University, College Station, Texas 77843, USA
[3] Department of Electrical and Computer Engineering, Texas A&M University, College Station, Texas 77843, USA
[4] CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia

ABSTRACT

Power curves are commonly estimated using the binning method recommended by the International Electrotechnical Commission, which primarily incorporates wind speed information. When such power curves are used to quantify a turbine's upgrade, the results may not be accurate because many other environmental factors in addition to wind speed, such as temperature, air pressure, turbulence intensity, wind shear and humidity, all potentially affect the turbine's power output. Wind industry practitioners are aware of the need to filter out effects from environmental conditions. Toward that objective, we developed a kernel plus method that allows incorporation of multivariate environmental factors in a power curve model, thereby controlling the effects from environmental factors while comparing power outputs. We demonstrate that the kernel plus method can serve as a useful tool for quantifying a turbine's upgrade because it is sensitive to small and moderate changes caused by certain turbine upgrades. Although we demonstrate the utility of the kernel plus method in this specific application, the resulting method is a general, multivariate model that can connect other physical factors, as long as their measurements are available, with a turbine's power output, which may allow us to explore new physical properties associated with wind turbine performance. Copyright © 2014 John Wiley & Sons, Ltd.

KEYWORDS
Binning method; non-parametric methods; kernel estimation; power curve; wind turbine

Correspondence
Yu Ding, Room 4016 ETB, Department of Industrial and Systems Engineering, Texas A&M University, College Station, Texas 77843-3131, USA.
E-mail: yuding@iemail.tamu.edu

Received 17 June 2013; Revised 29 January 2014; Accepted 27 March 2014

1. INTRODUCTION

Wind turbine generators are arguably one of the most crucial elements in the wind energy infrastructure. Turbine performance assessment plays an important role in wind turbine maintenance, equipment procurement and wind energy planning. Over time, a wind turbine naturally degrades, losing efficiency in power generation. To maintain the production efficiency of a wind turbine, practitioners sometimes perform a retrofit to an existing turbine known as an upgrade. An upgrade adds new components to a turbine and/or adjusts existing ones.

Upgrading is costly. For instance, one type of upgrade we present in the following is a vortex generator installation. Past studies, for instance,[1] claimed that having vortex generators could improve the lift characteristics of the blades and thus benefit wind energy production. Installing vortex generators, however, requires retrofitting of the wind turbine blades, and doing so incurs material and labor costs and halts energy production during installation. Wind operators often wonder whether or not the performance of a wind turbine is improved enough to justify the cost of upgrading.

Practitioners recognize that wind power output from a turbine is driven by wind input. It therefore makes little sense to compare only the difference in wind power production before and after an upgrade. Even if there exists a difference, this output-only difference would not reveal if the difference comes from the upgrading of the turbine or merely from the occurrence of a strong wind after the upgrade. A more sensible method, which is also a common practice in the wind industry, is through an analysis of the power curve associated with the turbine. A power curve measures the relationship

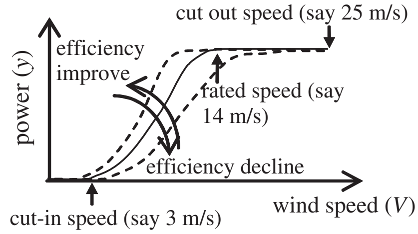

Figure 1. An example of a power curve and the associated power production efficiency.

between the power output of a turbine and the wind speed, as illustrated in Figure 1. A change in the position and slope of a turbine's power curve, especially in the part between the cut-in wind speed and rated speed, indicates a change in energy production efficiency. In theory, it is possible to estimate the power curve associated with a turbine before and after an upgrade and compare the power curves to determine if the production efficiency has improved.

The current industrial practice of estimating power curves from measurements relies on the binning method as recommended by the International Electrotechnical Commission.[2] The basic idea of the binning method is to discretize the domain of wind speed into a finite number of bins using a bin width of, for example, 0.5 m s^{-1}. Then, the value to be used for representing the power output for a given bin is simply the sample average of all the data points falling within that specific bin. The details of the binning method and a number of other variants are reviewed in Section 2.

A shortcoming of the binning method is that it produces an average power curve for a turbine using only the wind speed. Besides the wind speed, however, many other environmental factors, such as temperature, air pressure, turbulence intensity, wind shear and humidity, all potentially affect the power output of a turbine. When there is a change in the power curve, it is still difficult to discern the source of the change. A power curve (or response) determined by a turbine's own aerodynamics, after the influence from the major environmental factors is controlled for, is needed. When the environmental factors are not controlled for, the power output data on a power curve plot scatter widely around the average curve, indicating a large amount of uncertainty.

On the other hand, the meteorological mast (MAST) on each wind farm includes a wide array of sensors that measure more than just wind speed. Other environmental variables measured may include wind direction, D, temperature, T, air pressure, P, and humidity, H. Based on wind speed measurements, it is also possible to calculate turbulence intensity, I (equal to the standard deviation (SD) of short-duration wind speeds divided by the average wind speed of the same duration) and wind shear, S (using wind speeds measured at different heights). Ongoing innovations are also introducing spatial measurement capabilities through LIght Detection And Ranging (LiDAR) or SOnic Detection And Ranging (SODAR) technologies that can potentially measure and represent the spatial structure of the wind field in front of a turbine. Our conjecture is that incorporating extra environmental variables (also known as the weather covariates) or spatial measurements in a power curve model gives us a tool for controlling for variations from the environmental variables, thereby reducing the uncertainty in subsequent analysis.

There are inherent difficulties to extend the binning method to incorporate the extra environmental variables; we provide more detailed arguments in the next section. To tackle these problems, we developed a non-parametric, multivariate method, called *kernel plus*, that can incorporate many environmental variables.

In the current application, we use environmental measurement data (as well as power output data) from a commercial wind farm as is. The set of data does not include LiDAR spatial measurements but it does include several meteorological measurements, such as temperature and air pressure. Using the proposed method, we conduct two case studies related to turbine upgrades, demonstrating the superiority of the kernel plus method in detecting small-to-moderate changes resulting from certain turbine upgrades. We also wish to point out that the utility of the proposed kernel plus method is broader than merely the application of turbine upgrade quantification; it can be used to facilitate the exploration for and understanding of how any newly measured physical factor affects a turbine's power production ability.

The remaining parts of this paper are discussed in the following order. Section 2 reviews the methods and practices of turbine performance assessment. Section 3 presents the details of the kernel plus method as well as a procedure for quantifying a turbine's upgrade. We present the case studies in Section 4 and conclude the paper in Section 5.

2. LITERATURE REVIEW

For turbine performance assessment, a sensible approach is to use a power output to account for environmental effects. This can be achieved by using the *power residual*, denoted by R, which is the difference between the observed power output, y, and the power output estimated through a power curve model, $\hat{y}$, as in the following:

Wind Energy 2015; **18**:1207–1219 © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/we

$$R = y - \hat{y}. \tag{1}$$

Of course, whether or not the power residual captures an endogenous change to a turbine depends on how $\hat{y}$ is obtained. Please note that the power curve model $\hat{y}$ is represented by a functional surface across all environmental variables and can be used to study unmodified turbines as well as modified ones.

As we mentioned in Section 1, the most popular tool used by practitioners for estimating a power curve from measurements is the binning method, computed by

$$\hat{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{i,j} \tag{2}$$

where $y_{i,j}$ is the power output of the j^{th} data point in bin i, and n_i is the number of data points in bin i. In this plain version of the binning method, all environmental variables, other than the wind speed, are ignored.

The commonly used physical relationship in wind power generation[3,4] states that

$$y = \frac{1}{2}\cdot C_p \cdot \rho \cdot \pi r^2 \cdot V^3 \tag{3}$$

where r is the radius of the rotor, ρ is the air density and C_p is the so-called power coefficient. Motivated by this physical relation, practitioners recognize the need to include air density as a factor in calculating the power output and do so through a formula known as the air density correction.[2] If V is the raw wind speed, the air density correction adjusts the wind speed based on the measured air density, ρ, namely,

$$V' = V\left(\frac{\rho}{\rho_0}\right)^{\frac{1}{3}} \tag{4}$$

where ρ_0 is the dry air density at the sea level ($= 1.225$ kg m^{-3}) per international atmosphere standard. The binning method with the air density correction uses this corrected wind speed, V', and the power output, y, to establish a power curve. In the subsequent analysis, as well as in Section 4, by 'binning method', we mean this air density corrected version unless otherwise noted.

Other existing methods of fitting a power curve are similar to the binning method in the sense that wind speed is used as the sole explanatory variable, although the specific techniques used for curve fitting are quite different.[5–9] Wan *et al.*[10] considered more than just wind speed; they binned both wind speed and direction and used a neural network model using both wind speed and air density as inputs. In the end, Wan *et al.*[10] concluded that the binning method with air density correction produced the best power curve-fitting outcome among all the alternatives they investigated.

Despite the availability of other environmental measurements and their potential impact on power curve estimation, most current methods still make use of wind speed only. The power residuals calculated by the current methods reveal a non-random pattern and high variance in the residuals; such residual characteristics hinder us from detecting small-to-moderate changes in turbine performance. The need to develop power curve methods with multivariate dependencies has been noted.[11] However, doing so through an extension of the binning method is not easy. The binning method has its own merit, not only that it is a simple procedure and very easy to understand but also a non-parametric method that relies on few assumptions. It is therefore robust. Its major limitation lies in its rigid compartmentalization of data and its separate use of data for each individual bin rather than borrowing strength from other bins. A binning method is effective when used on single dimensional data but runs into the 'curse of dimensionality' and loses its effectiveness when used for handling multivariate data.

To see this, consider the following. Suppose that we have a whole year worth of data arranged in increments of 10 min, translating to 52,560 data points, and we now try to bin the data against, say, five input variables, and each of them uses 20 bins. Then, we end up with more than three million bins. On average, every 60 bins share one data point, meaning that we are going to see a lot of empty bins with no data at all. Such data scarcity will cause inaccuracy or higher uncertainty in subsequent data analysis and statistical estimation. Researchers may argue that we can use fewer bins. But reducing the number of bins does not fundamentally reduce the curse of dimensionality. Moreover, with fewer bins, the binning method loses the resolution that is needed for producing a sufficiently smooth curve. The choice of bin width is by itself a problem often under debate among both academic researchers and practitioners.

3. METHODOLOGY

We present our proposed methodology in this section. We first present some background information on the conventional kernel regression method in Section 3.1 and explain why the existing kernel method does not address the challenges

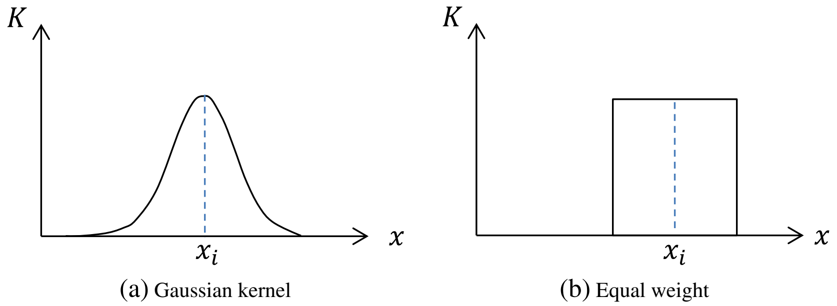

Figure 2. The kernel function in a kernel regression.

identified earlier. Then, we present our new method, followed by a summary procedure to show how the method can be used for the purpose of assessing changes in turbine performance after an upgrade.

3.1. The conventional kernel method

Consider the task of regression using a dataset, $\{x_i, y_i)\}_{i=1}^N$, where x_i is the covariate (or explanatory variable) and N is the data amount. For a power curve as shown in Figure 1, x is the wind speed, V. Kernel regression is a non-parametric, localized regression method.[12] The basic idea is to make an estimation, $\hat{y}$, at x, using observation pairs, (x_i, y_i), close to x. This localization is achieved by associating a weighting function with each data point x_i. This weighting function is a smoothly decaying function, symmetric with respect to x_i and also integrable to 1, such that the magnitude of $\hat{y}$ is consistent with that of the original data, y. When all the conditions on the weighting function are satisfied, it is called a kernel function and is represented by $K_\lambda(x, x_i)$, where λ is known as the bandwidth of the function, indicating how fast the function decays from its peak toward zero and hence controlling the smoothness of the regression function. One of the popular kernel functions is the Gaussian kernel function, taking the form

$$K_\lambda(x, x_i) = \frac{1}{\sqrt{2\pi\lambda^2}} \exp\left(-\frac{||x - x_i||^2}{2\lambda^2}\right). \tag{5}$$

Figure 2(a) presents a Gaussian kernel function. Using this kernel function, a kernel regression estimator,[13,14] also known as the Nadaraya–Watson estimator, can be expressed as

$$\hat{y}(x) = \sum_{i=1}^{N} w_i(x) \cdot y_i, \text{ where } \quad w_i(x) = \frac{K_\lambda(x, x_i)}{\sum_{i=1}^{N} K_\lambda(x, x_i)}. \tag{6}$$

In essence, a kernel regression method is a weighted average of all the data points with the weighting coefficient, $w_i(x)$, associated with the data point, x_i. The weighting coefficient, $w_i(x)$, decays toward zero, as the distance increases between x_i and x, the location where the estimation is made. Even though the Gaussian kernel in the theory gives every data point a positive weight, the weight is practically zero when x_i is sufficiently farther, say 3 λ away from x. For this reason, a kernel regression is considered a localized regression method.

Compared with the simple one-dimensional kernel regression, the binning method bears a certain similarity. The binning method can even be considered as a special kernel method but it uses a step function as the weighting function, giving a constant weight to any location, x, within the step function window of data point, x_i; see Figure 2(b). The binning method is not really a kernel method for another important reason: in the kernel regression, the estimation location, x, moves continuously along the abscissa axis, producing a continuous, smooth curve, whereas the corresponding x in the binning method jumps from one bin to the next one, so that the resulting power curve from the binning method, if magnified enough, is discretized. However, the difference between the two methods is not great, as long as the covariate remains a single variable like the wind speed. With a one-dimensional covariate, changing the binning method to a kernel regression produces just a slightly different power curve, assuming that a fine enough resolution has been used to bin the wind speed.

Kernel regression can be extended to multivariate cases. Suppose that multiple covariates are represented by the vector $\mathbf{x}$ of dimension q, namely, $\mathbf{x} = (x_1, x_2, \ldots, x_q)$. A multivariate kernel regression still bears the form as in Equation (6), but the kernel function used therein should be a multivariate kernel function and is often composed of a product kernel that is a multiplication of univariate kernel functions, such as

Wind Energy 2015; **18**:1207–1219 © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/we

$$K_{\boldsymbol{\lambda}}(\mathbf{x}, \mathbf{x}_i) = K_{\lambda_1}(x_1, x_{i,1}) \cdot K_{\lambda_2}(x_2, x_{i,2}) \cdots K_{\lambda_q}(x_q, x_{i,q}) \tag{7}$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_q)$ is the set of bandwidths associated with each covariate, and $x_{i,j}$ is the jth element of $\mathbf{x}_i$. The multivariate kernel regression suffers from the curse of dimensionality as well—when q is large, the estimation quality deteriorates because the data become sparse in a high-dimensional space.

3.2. The kernel plus method

We propose a kernel regression procedure, taking advantage of the sound statistical properties of conventional kernel methods while addressing their shortcomings. There are two primary shortcomings we address in this section: (i) the curse of dimensionality issue as mentioned earlier and (ii) the bias associated with a Nadaraya–Watson kernel estimator when it is used in a new dataset. We call the resulting method the *kernel plus* method. To address the curse of dimensionality issue, kernel plus chooses a hybrid structure that includes multiplicative kernel functions in an additive model. A pure additive model is scalable and can easily include many explanatory variables or covariates without running into the dimensionality problem. The downside is that such a model fails to capture the interaction among the covariates. This is a crucial drawback for additive models, as the physical law in Equation (3) governing wind energy generation tells us that the important covariates, such as wind speed and air density, interact in a nonlinear fashion when both are connected with the power output. We have in fact attempted pure additive models, and they are not effective as expected.

On the other hand, a pure multiplicative model, while capable of capturing the complicated nonlinear relationship among covariates, becomes less practical when its product kernel has more than three elements. A natural compromise is to use a multiplicative–additive hybrid: the multiplicative kernel has only three element kernels to capture the important covariate interaction, and the additive model structure ensures the scalability of the resulting model. Observing again the governing physics in Equation (3), we believe that wind speed, V, and wind direction, D, are the two most important covariates and should be included in all the multiplicative kernels. The third element in a trivariate multiplicative kernel is another weather-relevant covariate. The final kernel plus model includes $q-2$ trivariate multiplicative kernels, each of which takes a different covariate as its third element.

We designate the first two elements of $\mathbf{x}$ always as $x_1 = V$ and $x_2 = D$ and introduce a new symbol, $\mathbf{x}^j$, such that $\mathbf{x}^j := (x_1, x_2, x_j), j = 3, \ldots, q$. With this notation, our hybrid kernel estimator takes the form of

$$\hat{y}^{HK}(\mathbf{x}) = \frac{1}{(q-2)} \left[\hat{y}(\mathbf{x}^3) + \ldots + \hat{y}(\mathbf{x}^q) \right] \tag{8}$$

where $\hat{y}(\mathbf{x}^j)$ is the Nadaraya–Watson kernel estimator in Equation (6) with its kernel function, a multiplication of three univariate kernel functions, $K_{\lambda_1}(x_1, x_{i,1})$, $K_{\lambda_2}(x_2, x_{i,2})$ and $K_{\lambda_j}(x_j, x_{i,j})$, $j = 3, \ldots, q$. When additional covariates become available, we plan to add extra additive terms, each of which has the same structure as the current terms, namely, a trivariate multiplicative kernel having inputs of V, D and a new covariate. To address the bias issue of the Nadaraya–Watson kernel estimator, we propose a self-calibration procedure. The existence of bias based on finite samples is a common problem in statistical prediction; see Figure 7.2 of Hastie *et al.*[12] for a full illustration of biases and variances involved in statistical prediction.

Before we present the procedure, we define a few terms. Our method is a data-driven method, and it involves three datasets to quantify a wind turbine upgrade: a training dataset of historical observations of $(\mathbf{x}, y)$ pairs and two test datasets before and after an upgrade; see Figure 3. The training dataset is referred to as 'Data0' and is used to fit a power curve model. This model is the mathematical surrogate of a physical turbine, and it is supposed to represent the 'old turbine' before upgrade. Data0 should be collected from a reasonable duration of a turbine's operation, for instance, 1 year, such that the seasonal weather effects are represented in the data. The two test sets, referred to as 'Data1' and 'Data2', respectively, are collected for the same length of duration before and after the upgrade. They are used to detect and quantify the upgrade. Their corresponding data duration can be much shorter than that of Data0; a few weeks to a few months may be sufficient.

The self-calibration procedure is done primarily by using subsets of the training data in Data0. To select a calibration set of data that has similar weather conditions to those in Data1 and Data2, we define a distance measure, which is in spirit similar to the Mahalanobis distance,[15] i.e., a weighted distance between two multidimensional points based on the corresponding covariance matrix. Our defined distance measure is indeed a weighted distance, but instead of using the covariance matrix, we employ a diagonal matrix whose diagonal elements are from the bandwidth vector $\boldsymbol{\lambda}$. Specifically, we denote this diagonal matrix as Λ, so that $\Lambda_{i,i} = \lambda_i$ and $\Lambda_{i,j} = 0 \quad \forall i \neq j$. The resulting distance measure between a training data point, $\mathbf{x}_i \in$ Data0, and a testing data point, $\mathbf{x}_j$, in either Data1 or Data2 is

$$D(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T \Lambda^{-1} (\mathbf{x}_i - \mathbf{x}_j)}. \tag{9}$$

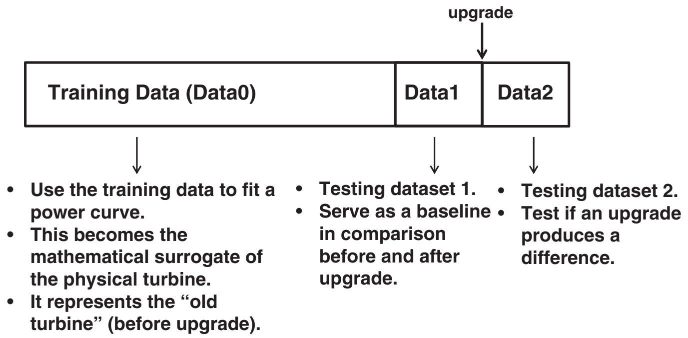

Figure 3. Data set partitions used in quantifying an upgrade.

The reason we choose this distance measure to define the similarity between two **x**'s is as follows. A simple Euclidean distance does not reflect the similarity between the **x**'s well because different elements in **x** have different physical units, leading to different value ranges. To define a sensible similarity measure, a key issue is to weigh different elements in **x** consistently with their relative importance pertinent to the power output. The original Mahalanobis distance does not serve this purpose because the squared distance associated with an input variable is weighted by the inverse of its variance. In a power curve model, wind speed is arguably the most important variable, yet it has a large variance. Because of this large variance, using the Mahalanobis distance will in fact diminish the importance of wind speed relative to other variables that have a smaller variance. Our choice of using the kernel bandwidth parameters as the weighting coefficients in Λ is consistent with our goal of weighting each element according to its relative importance, because the bandwidth parameters are selected based on how sensitive the power output is to a unit change in the corresponding input variable. If an input variable has a small bandwidth, it means that the power output could produce an appreciable difference with a small change in the corresponding input, suggesting that this variable is relatively important. On the other hand, a large bandwidth indicates a less important input variable.

For any testing data point $\mathbf{x}_j$, we can choose a calibration point $\mathbf{x}_i^{\text{cal}}$ from Data0, which has the minimum $D(\mathbf{x}_i^{\text{cal}}, \mathbf{x}_j)$ distance with the testing data point. Then, we employ the self calibration procedure as follows:

1. For $\mathbf{x}_i^{\text{cal}} \in$ Data0, compute $\hat{y}^{HK}\left(\mathbf{x}_i^{\text{cal}}\right)$
2. Compute the calibration value $R_j^{\text{cal}} = y\left(\mathbf{x}_i^{\text{cal}}\right) - \hat{y}^{HK}\left(\mathbf{x}_i^{\text{cal}}\right)$
3. For any testing data point $\mathbf{x}_j$, the final power estimate from the kernel plus method is $\hat{y}^{KP}(\mathbf{x}_j) = \hat{y}^{HK}(\mathbf{x}_j) + R_j^{\text{cal}}$.

3.3. Procedure for quantifying an upgrade

Figure 4 outlines the procedure for detecting and quantifying an upgrade using the kernel plus method and the three sets of data as described in the previous subsection.

To start, we establish a power curve model using the kernel plus method. This step includes the use of both the hybrid kernel and the self-calibration. Then, this power curve model, representing the 'old' turbine, is used to make a prediction/estimation of power output under a new weather profile **x** in either Data1 or Data2, and the result is denoted as $\hat{y}(\mathbf{x}^{(1)})$ and $\hat{y}(\mathbf{x}^{(2)})$, respectively. Consequently, the corresponding power residuals can be computed. Had a turbine undergone an upgrade, we would expect the residuals before and after the upgrade to be different. To detect a potential difference in the residuals, it is necessary to invoke a statistical test because the power output data are noisy; specifically, a Student's *t*-test[16] is used. Suppose that N_1 and N_2 are the number of data points in Data1 and Data2, respectively. The statistical test procedure is as follows:

- Compute the residuals before and after an upgrade, i.e., for Data1, $R(\mathbf{x}^{(1)}) := y(\mathbf{x}^{(1)}) - \hat{y}(\mathbf{x}^{(1)})$, and for Data2, $R(\mathbf{x}^{(2)}) := y(\mathbf{x}^{(2)}) - \hat{y}(\mathbf{x}^{(2)})$;
- Compute the two sample means and the corresponding SDs by using the following formula:

$$\bar{R}_k = \frac{\sum_{j=1}^{N_k} R(\mathbf{x}_j^{(k)})}{N_k} \text{ and } S_k = \sqrt{\frac{\sum_{j=1}^{N_k}\left(R\left(\mathbf{x}_j^{(k)}\right) - \bar{R}_k\right)^2}{N_k - 1}} \qquad \text{for} \quad k = 1, 2$$

where N_1 and N_2 are the number of observations in Data1 and Data2, respectively.

Wind Energy 2015; **18**:1207–1219 © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/we

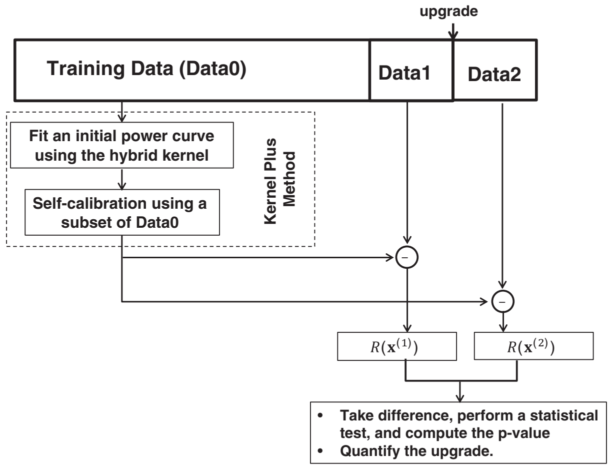

Figure 4. The procedure of quantifying a turbine upgrade using the kernel plus method.

- Then, the pooled estimate of SD σ_R is calculated as

$$\sigma_R = \sqrt{\frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}}.$$

- The t statistic is calculated as

$$t = \frac{\bar{R}_2 - \bar{R}_1}{\sigma_R \cdot \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}.$$

- Finally, we calculate the p-value of the t statistic. A p-value is a probability, taking values between 0 and 1. The smaller the p-value, the more significant the difference.

The aforementioned procedure is devised to confirm any detectable difference resulting from an upgrade. The output is binary, either the upgrade produces a statistically significant difference in a turbine's performance or it does not.

Another practical question is that if the aforementioned t-test does indicate a significant difference, how much a difference in terms of power generation does the upgrade produce? To answer this question, we define a quantifier as follows:

$$\text{DIFF}(\mathbf{x}) = \frac{\sum_{\mathbf{x} \in \mathcal{D}_{test}} (y(\mathbf{x}) - \hat{y}(\mathbf{x}))}{\sum_{\mathbf{x} \in \mathcal{D}_{test}} y(\mathbf{x})} \times 100\% \tag{10}$$

where $\mathcal{D}_{test}$ is a test dataset and can be either Data1 or Data2, so that $\mathbf{x}$ can be either $\mathbf{x}^{(1)}$ or $\mathbf{x}^{(2)}$ accordingly. Similar to the residual analysis described earlier, comparing DIFF($\mathbf{x}^{(2)}$) with DIFF($\mathbf{x}^{(1)}$) quantifies the benefit of the upgrade. Note further that DIFF $=$ DIFF($\mathbf{x}^{(2)}$) $-$ DIFF($\mathbf{x}^{(1)}$) provides the difference in power generation between the old turbine and the new turbine when the effects of the weather profile are controlled for. Obviously, the bigger the DIFF, the more wind energy an upgraded turbine produces under the same weather profile.

4. CASE STUDY

4.1. Background and datasets

We study two types of upgrades in this section: one is a vortex generator installation and the other is an artificial pitch angle adjustment. In the literature, both upgrades were believed to provide benefit in improving wind energy production.[1,17] We have wind farm operational data from a turbine that underwent vortex generator installation. The reason we include the artificial pitch angle case is to facilitate the understanding of a method's sensitivity to a known perturbation. What we do in this artificial case is to take another turbine from the same wind farm and its current operational data and synthesize the after-upgrade data according to the understood characteristic of pitch angle adjustment.

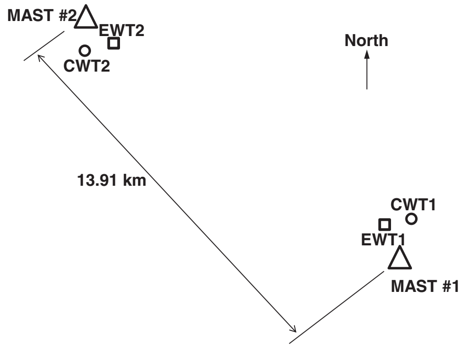

Figure 5. Wind farm layout. Triangles: meteorological masts; squares: experimental turbines that had an upgrade; circles: turbines in the control group that did not have the upgrade.

Table I. Specifications of the wind turbines in the case study.

Hub height (m)	80
Rotor diameter (m)	About 80
Cut-in wind speed (m s^{-1})	3.5
Cut-out wind speed (m s^{-1})	20
Rated wind speed (m s^{-1})	Around 13
Rated power (MW)	1.5 - 2.0
Location	In land, USA

As such, we select two groups of wind turbines from an inland wind farm, each consisting of a MAST and two wind turbine pairs, one is an experimental wind turbine (EWT) that had an upgrade, and the other is a control wind turbine (CWT) in the vicinity of the EWT but did not have the upgrade. In the wind farm layout as shown in Figure 5, the distance between a mast and a turbine is not scaled accurately but their relative positions, as well as their positions relative to the mast, reflect the reality.

The turbine characteristics are presented in Table I, in which an approximation rather than the accurate value is given to certain characteristics to protect the turbine manufacturers' and wind farms' identities. In both pairs, the EWT and CWT are practically identical turbines and were installed and put into service at the same time. We selected them because we believe that all other differences associated with a turbine pair are controlled for as practically as possible. This wind farm is on a reasonably flat terrain, so the elevation differences between the EWT and CWT pairs are about 2–3 m.

The power output, y, is measured on individual turbines, whereas the environmental variables in $\mathbf{x}$ (i.e., the weather covariates) are measured by sensors at the nearby mast. The sensor equipment at the mast is compliant with the International Electrotechnical Commission standard.[2] The pairs of $(\mathbf{x}, y)$ constitute the data points in the respective datasets. For this particular wind farm, the following environmental variables are measured: wind speed, wind direction, air pressure and temperature. We convert the air pressure and temperature measurements into air density, ρ, following a standard formula.[7] Those environmental measurements, as well as the power output data, are arranged in 10-min block and with an average of 10-min data recording. We understand that such treatment is consistent with industry practice.

Using wind speed measurements, we can also compute the turbulence intensity, I, and the vertical below-hub wind shear, S. The vertical wind shear, S, uses two wind speed measurements, one at the hub height and the other at a height below the hub height. There are no wind measurements that allow us to compute different kinds of wind shear other than this vertical one. Humidity could be measured, but for the datasets we have at hand, humidity measurements are not available. Taking all the aforementioned discussions into account, we have five covariates in $\mathbf{x}$, i.e., $\mathbf{x} = (V, D, \rho, I, S)$. When using this $\mathbf{x}$ to fit a kernel plus model, we find that including the vertical wind shear, S, does not make the final model any better. In the end, we choose to leave out wind shear, S. The final weather covariate vector is therefore $\mathbf{x} = (V, D, \rho, I)$ and $q = 4$.

Turbine EWT1 underwent a vortex generator installation. The vortex generators are on the suction side of the blade and are located radially on the inboard (root) end of the blade. For the group of MAST # 1, EWT1 and CWT1, we have 13 months of actual operational data in Data0 and 2 weeks of operational data in both Data1 and Data2.

For the group that include MAST # 2, EWT2 and CWT2, we have 14 months of data in Data0 and 2 weeks of data in both Data1 and Data2. Turbine EWT2 is used to mimic the effect from a pitch angle adjustment. Since it does not actually undergo such an upgrade, we generate the after-upgrade data, i.e., Data2, by modifying its existing data. Our treatment is based on the study[17] that simulated the power output based on different pitch angles. Figure 6 illustrates the results from

 Wind Energy 2015; **18**:1207–1219 © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/we

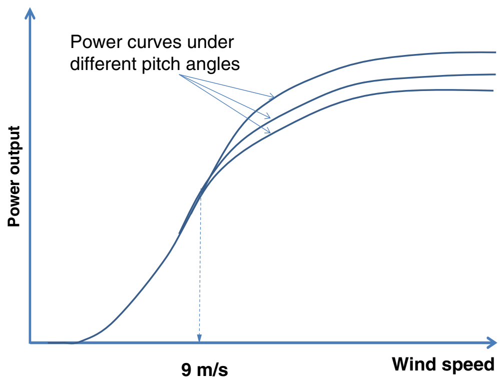

Figure 6. Illustration of power output results based on different pitch angles. This is a simplified adaptation from a figure in Wang *et al.*[17] The original work simulated three attack angles, 5.5°, 6.0° and 6.5°, under the same designed tip speed ratio of 7 and designed wind speed of 8.0 m s^{-1}.

Table II. Bandwidth parameters used in the kernel plus method.

Turbine	λ_V	λ_D	λ_ρ	λ_I
EWT1	0.2163	3.4057	0.0019	0.0043
CWT1	0.2377	3.2234	0.0021	0.0050
EWT2	0.2907	3.6651	0.0018	0.0054
CWT2	0.2575	4.1766	0.0014	0.0050

EWT, experimental wind turbine; CWT, control wind turbine.

Wang *et al.*,[17] concerning the rotor power performance corresponding to different pitch angels, where we notice that the pitch angle adjustment increases the power performance primarily for wind speeds above 9 m s^{-1}. Accordingly, to mimic the behavior of a turbine that underwent a pitch angle adjustment, we multiply the original power outputs (i.e., y values) in Data2 of EWT2 by a factor of 1.05 for those corresponding to wind speeds above 9 m s^{-1}. We did not make any other changes to the $\mathbf{x}$ vector nor the y value in Data0 and Data1 of EWT2. No changes were made to any datasets associated with CWT2.

4.2. Upgrade detection

We first apply both the kernel plus method and the binning method to the two groups of turbines and conduct the residual analysis. The use of the kernel plus method follows the logic flow shown in Figure 4. When using the binning method, we simply replace the kernel plus method, i.e., the dashed line rectangle in Figure 4, with the binning method (the version with air density adjustment). To determine the bandwidth parameters in the kernel plus method, we follow a standard procedure in the statistical literature.[18] Table II presents the bandwidth parameters used for the data associated with the four turbines.

For an upgraded turbine, a method is supposed to produce a large t-statistic (in its absolute value), which will further lead to a small p-value that signifies the difference between the residuals, whereas for a turbine without upgrade, a small t statistic, or equivalently, a large p-value is expected. In statistical inferences, the commonly used threshold of a p-value to indicate significance is 0.05, which is what we use in this analysis.

Table III presents the outcomes from the residual analysis of both groups of turbines. The table's message is clear: the kernel plus method has significant outcomes consistent with the upgrade action, whereas the binning method does not. Although the binning method has significant outcomes on EWT2, it indicates the change in CWT2 as equally significant.

We believe that the erratic outcome of the binning method is attributable to the still large amount of uncertainty unaccounted for in its residuals, leading to big biases and/or great variability. To intuitively understand the outcomes of the statistical tests, we present in Figure 7 the residual plots using data from EWT1 when applying the binning and the kernel plus method. The residuals of the binning method follow an obvious pattern (leading to bias) and have a large dispersion, suggesting a poor model fit and large uncertainty, whereas the residuals of the kernel plus method have a considerably smaller dispersion and exhibit a random pattern, indicating an adequate model fit and reduced uncertainty. The residual plots using other turbines' data lead to the same insight and are thus omitted.

Table III. Comparing the kernel plus and binning methods on their ability to detect a turbine upgrade.

	Binning		Kernel plus	
Turbine	*t*-statistic	*p*-value	*t*-statistic	*p*-value
EWT1	-0.46	0.65	2.24	0.025
CWT1	-2.54	0.01	-0.14	0.89
EWT2	5.09	3.89×10^{-7}	3.18	0.002
CWT2	4.51	6.84×10^{-6}	-1.71	0.09

EWT, experimental wind turbine; CWT, control wind turbine.

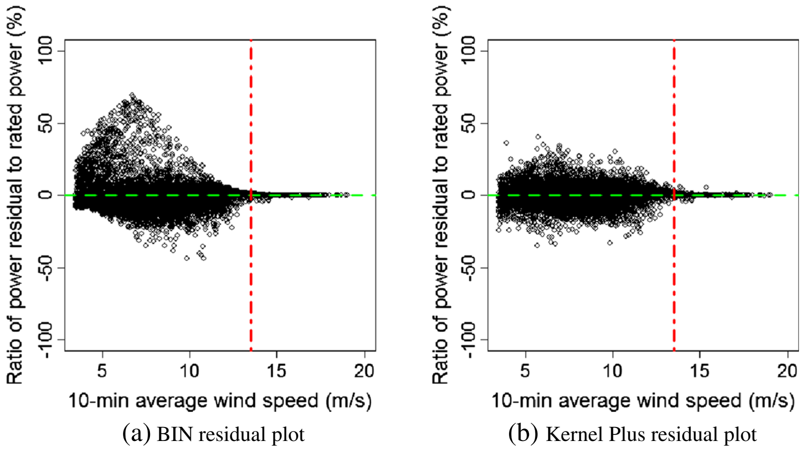

Figure 7. Residual plots after the binning method (left panel) or the kernel plus method (right panel) is applied. The vertical dashed line indicates the rated wind speed.

Table IV. Comparing the kernel plus and binning methods on their ability to quantify a turbine's upgrade.

Turbine	DIFF from binning	DIFF from kernel plus
EWT1	−0.43% (0.55%)	1.68% (0.73%)
CWT1	−2.92% (1.20%)	−0.13% (0.82%)
EWT2	5.98% (0.70%)	2.79% (0.59%)
CWT2	3.92% (0.98%)	−1.45% (0.75%)

EWT, experimental wind turbine; CWT, control wind turbine.

4.3. Upgrade quantification

We compute the DIFF metric as defined in Section 3.3. Once again, both the binning method and the kernel plus method are used. The original dataset is bootstrapped 10 times, producing 10 replications of the quantification analysis. Both the average improvement and its SD (the quantity in parenthesis), reported as percentages, are presented in Table IV.

The outcomes from the kernel plus method are consistent with the upgrade action. When there are both a vortex generator installation and a pitch control adjustment, EWT1 and EWT2 can produce, respectively, 1.68% and 2.79% more wind energy after the upgrade than before the upgrade under the same weather profile. For EWT1, the range of the percentage improvement recorded among the 10 replication is $[0.82\%, 3.15\%]$. For EWT2, the range is $[1.88\%, 3.88\%]$. The 95% confidence interval (CI) of the improvement can be roughly estimated by using the average $\pm 2\times$ SD. It is obvious that when using the kernel plus method, the 95% CI of the improvements in EWT1 and EWT2 are both greater than zero, whereas those of CWT1 and CWT2 both contain zero, consistent with the detection action.

For the case of pitch angle adjustment, since we synthesize the after-upgrade data, we can calculate the actual power increment rate in the data. Recall that we multiplied the original power outputs by a factor of 1.05 but only for the power output corresponding to wind speed of 9 m s^{-1} or above. If we compute the power output difference before and after the adjustment for the entire wind spectrum from the cut-in to the cut-out wind speeds, the power increment rate is 2.80%, which is rather close to the average estimated by the kernel plus method (2.79%).

Wind Energy 2015; **18**:1207–1219 © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/we

On the other hand, the binning method does not appear to be effective in quantifying the benefit of the upgrade. For the first group, the binning method indicates that EWT1 and CWT1 perform, on average, worse after the upgrade than before the upgrade. When considering its 95% CI, EWT1 does not produce a statistically significant improvement, because its 95% CI contains zero. For the second group, the binning method appears to indicate that EWT2 and CWT2 performs much better after the upgrade. However, the percentage of improvement associated with EWT2 is more than two times larger than the actual power production increase, leaving the credibility of the binning-based quantification in doubt.

4.4. Power curves

When we use the kernel plus method, what we fit is no longer a one-dimensional power curve; it is a multi-dimensional power response surface that is difficult to visualize. Practitioners in the wind industry, however, believe that

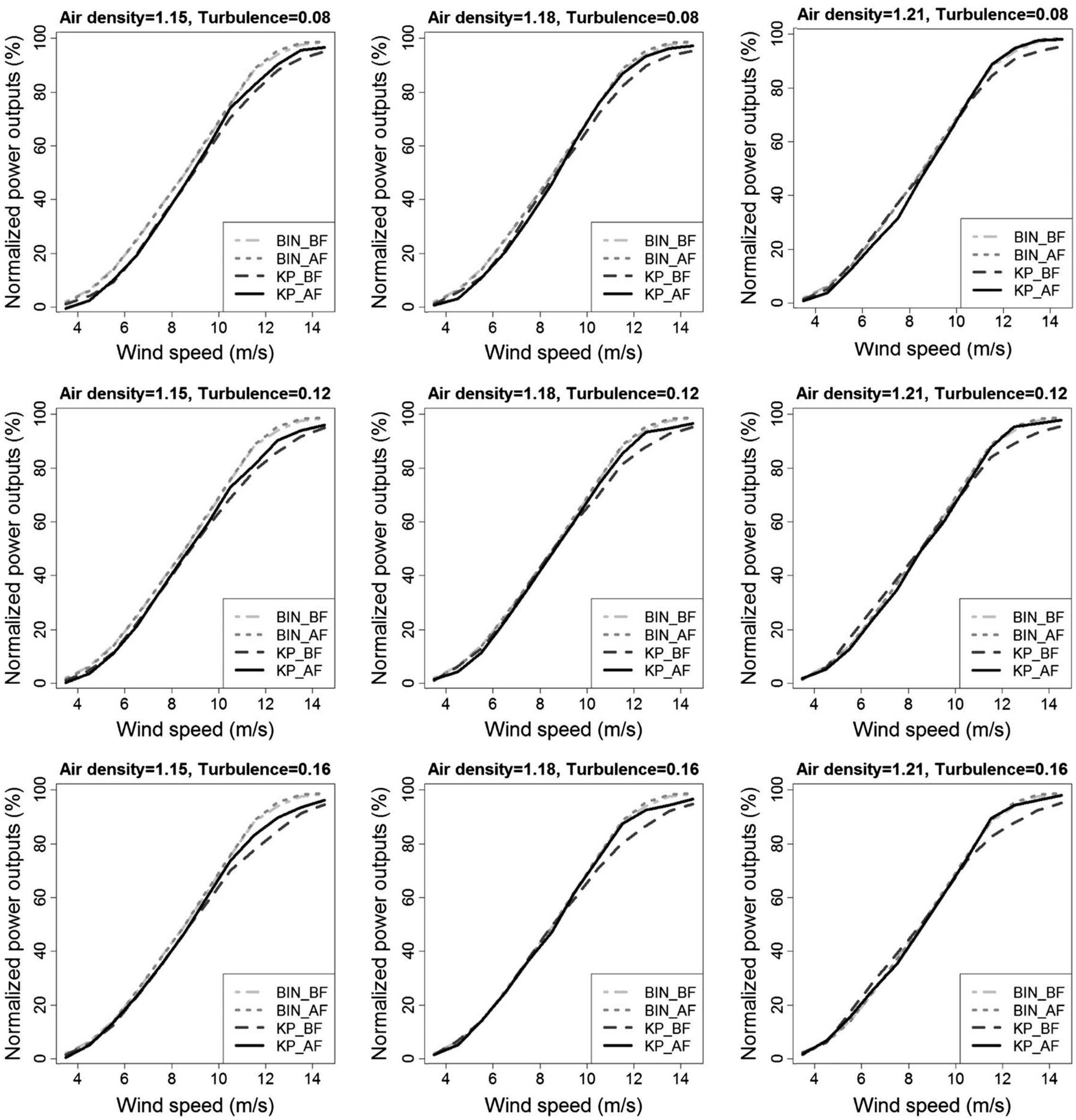

Figure 8. Power curves conditioned on air density and turbulence intensity and averaged across wind directions. The power outputs are normalized by the turbine's rated power. BIN_BF: binning method before upgrade; BIN_AF: binning method after upgrade; KP_BF: kernel plus method before upgrade; KP_AF: kernel plus method after upgrade.

one-dimensional power curves often provide useful insights on where the improvement comes from. In order to produce a one-dimensional power curve using the kernel plus method, covariates other than the wind speed need to be either conditioned on a constant value or averaged among all the possible values. Our current power curve model incorporates four weather covariates (V, D, ρ, and I). We choose to produce a series of power curves under different combinations of ρ and I but averaged over all possible values of wind directions. We choose three settings for both ρ and I, such as $\rho = (1.15, 1.18, 1.21)$ and $I = (0.08, 0.12, 0.16)$; altogether, there are nine combinations. We use the data from EWT1 to produce the power curves. The same can be done for other turbines, but to save space, we present only the results for EWT1.

In Figure 8, we include the power curves produced by both the kernel plus method and the binning method. In using the kernel plus method, there are observable differences between the power curves before and after the upgrade in several subplots. The difference is usually the most pronounced around the rated wind speed. By comparison, the binning method produces power curves with no visually detectable difference. This result is consistent with the message of the previous two subsections. We want to caution readers that our purpose here is not to advocate a specific type of turbine upgrade or retrofit option but to advocate a new method for turbine performance assessment. When the binning method indicates that there is no difference in a turbine's power output before and after installing vortex generators, a question remains: is there really no benefit to have a vortex generator for this particular turbine, or is it possible that the method used, namely, the binning method, is incapable of detecting small to moderate changes due to the method's inherent limitations? We believe that our case study demonstrates that the latter possibility is plausible, and as an alternative, the kernel plus method is a better tool for practitioners to use in assessing turbine performance.

5. CONCLUDING REMARKS

We reported a kernel plus method in this paper, extending the conventional kernel method in two ways: (i) our method uses an additive–multiplicative hybrid model structure that allows the resulting method to capture important nonlinear interactions when modeling a power response and to remain scalable at the same time, and (ii) it includes a self-calibration component to alleviate the bias associated with kernel estimators. This kernel plus method serves as a useful tool for producing power curves in which the influence of environmental factors can be controlled for, such that the energy production efficiency of a turbine can be better quantified. By using the kernel plus method to quantify an upgrade taken by a turbine and comparing the outcomes with those using the binning method, we found that the kernel plus method is consistent with the action of upgrading, whereas the findings from the binning method appear random, due to the large, uncontrolled uncertainty associated with the method. We do acknowledge that in the field of the wind energy, a big uncertainty is the quality of wind speed measurement itself; our reported method does not mitigate this uncertainty.

On the methodology front, the kernel plus method is admittedly a data-driven method, which by itself does not produce new physics. But we believe that our method is a good facilitator and a capable tool in helping explore new physical discoveries about wind turbine performance. To understand how multiple physical factors may have impacts on turbine performance, a multivariate method must be in place to connect the measured physical factors with the power output. We hereby present the kernel plus method as a possible solution for this broader objective. On the application side, we use the kernel plus method to quantify a turbine's upgrade. We believe that the same method is equally applicable to the detection and quantification of turbine degradation. Another area of application is in wind power prediction. Since the kernel plus method produces an improved power curve, it could help to produce more accurate power predictions, given a set of measured weather covariates beyond the wind speed.

Finally, one referee pointed out a different school of thought on conducting turbine performance assessment. As published by Albers and Antoniou *et al.*,[19,20] the difference in power between two side-by-side turbines is measured in a timeline including data before and after the upgrade. The correlation of the wind power generated by the side-by-side turbines may remove the uncertainty of environmental measurements, leaving only the effect resulting from the turbine upgrade. It would be interesting and valuable to conduct a carefully devised comparison study to determine which method is more effective and robust in practice.

ACKNOWLEDGEMENTS

Lee and Ding were partially supported by grants from NSF (CMMI-0926803 and CMMI-1300236) and King Abdullah University of Science and Technology (KUS-CI-016-04). Xie was partially supported by a grant from NSF (ECCS 1150944).

REFERENCES

1. Øye S. The effect of vortex generators on the performance of the ELKRAFT 1000 kw turbine. *Aerodynamics of Wind Turbines: 9th IEA Symposium*, Stockholm, Sweden, 1995; 9–14. ISSN:0590-8809.
2. IEC. *IEC 61400-12-1 Ed 1, Wind Turbines - Part 12-1: Power performance Measurements of Electricity Producing Wind Turbines*. International Electrotechnical Commission: Geneva, Switzerland, 2005.
3. Söder L, Ackermann T. *Wind Power in Power Systems*. John Wiley & Sons: Chichester, UK, 2005. DOI: 10.1002/0470012684.
4. Belghazi O, Cherkaoui M. Pitch angle control for variable speed wind turbines using genetic algorithm controller. *Journal of Theoretical and Applied Information Technology* 2012; **39**: 5–10.
5. Yan Y, Osadciw LA, Benson G, White E. Inverse data transformation for change detection in wind turbine diagnostics. *Proceedings of the 22nd IEEE Canadian Conference on Electrical and Computer Engineering*, St. John's, Newfoundland, Canada, 2009; 944–949. DOI: 10.1109/CCECE.2009.5090267.
6. Kusiak A, Zheng H, Song Z. On-line monitoring of power curves. *Renewable Energy* 2009; **34**: 1487–1493. DOI: 10.1016/j.renene.2008.10.022.
7. Uluyol O, Parthasarathy G, Foslien W, Kim K. Power curve analytic for wind turbine performance monitoring and prognostics. *Proceedings of Annual Conference of the Prognostics and Health Management Society*, Vol. 2, Montreal, Canada, 2011; 049:1–8.
8. Osadciw LA, Yan Y, Ye X, Benson G, White E. Wind turbine diagnostics based on power curve using particle swarm optimization. In *Wind Power Systems: Applications of Computational Intelligence*, Wang L, Singh C, Kusiak A (eds). Springer-Verlag: Berlin, 2010; 151–165. DOI: 10.1007/978-3-642-13250-6_6.
9. Hayes BP, Ilie I, Porpodas A, Djokic SZ, Chicco G. Equivalent power curve model of a wind farm based on field measurement data. *Proceedings of IEEE PowerTech*, Trondheim, Norway, 2011; 1–7. DOI: 10.1109/PTC.2011.6019318.
10. Wan Y, Ela E, Orwig K, Development of an equivalent wind plant power curve. *Technical Report* NREL/CP-550-48146, National Renewable Energy Laboratory, 2010. Available: http://www.nrel.gov/docs/fy10osti/48146.pdf. (Accessed 12 April 2014)
11. Stephen B, Galloway SJ, McMillan D, Hill CD, Infield DG. A copula model of wind turbine performance. *IEEE Transactions on Power Systems* 2011; **26**: 965–966. DOI: 10.1109/TPWRS.2010.2073550.
12. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer: New York, 2009. DOI: 10.1007/978-0-387-84858-7.
13. Nadaraya E. On estimating regression. *Theory of Probability and Its Applications* 1964; **9**: 141–142. DOI: 10.1137/1109020.
14. Watson G. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A* 1964; **26**: 359–372. DOI: 10.2307/25049340.
15. Mahalanobis PC. On the generalised distance in statistics. *Proceedings of the National Institute of Science of India* 1936; **2**: 49–55.
16. Snedecor GW, Cochran WG. *Statistical Methods* 8th ed. Iowa State University Press: Iowa City, Iowa, USA, 1989.
17. Wang L, Tang X, Liu X. Blade design optimization for fixed-pitch fixed-speed wind turbines. *International Scholarly Research Network Renewable Energy* 2012; **682859**: 1–8. DOI: 10.5402/2012/682859.
18. Ruppert D, Sheather SJ, Wand MP. An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Assosication* 1995; **90**: 1257–1270.
19. Albers A. Relative and integral wind turbine power performance evaluation. *Proceedings of the 2012 European Wind Energy Conference & Exhibition*. November 22-25, London, UK, 2004.
20. Antoniou I, Downey R, Paulsen H, Enevoldsen P, Piperas A, Jensen L. Improved power performance of upgraded turbine rotors: a full scale verification study. (a poster presentation). *Proceedings of the 2012 European Wind Energy Conference & Exhibition*. April 16-19, Copenhagen, Denmark, 2012.