Renewable Energy 113 (2017) 1589–1597

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Quantifying the effect of vortex generator installation on wind power production: An academia-industry case study

Hoon Hwangbo [a], Yu Ding [a,*], Oliver Eisele [b], Guido Weinzierl [b], Ulrich Lang [b], Georgios Pechlivanoglou [b,**]

[a] Department of Industrial & Systems Engineering, Texas A&M University, College Station, TX, 77843, USA
[b] SMART BLADE® GmbH, Zuppingerstraße 14, D-88213, Ravensburg, Germany

ARTICLE INFO

Article history:
Received 28 January 2017
Received in revised form
23 May 2017
Accepted 3 July 2017
Available online 5 July 2017

Keywords:
Machine learning
Power curve
Power versus power
Statistical analysis
Vortex generator

ABSTRACT

This paper presents an academia-industry joint study concerning effective methods to estimate and quantify the effect of vortex generator installation on wind power production. This problem has presented a challenge to the wind industry, because (a) vortex generator installation may lead to a moderate 1–5% extra power production, but this level of improvement is difficult to be accurately detected; and (b) it is equally difficult to validate the estimated effect of vortex generator installation because a controlled experiment is practically impossible to conduct to provide a credible baseline. An academic institute and a wind technology company team up to tackle this challenge. The two teams develop their own version of quantification methods, which are profoundly different. The academic method uses 10-min data and makes use of both power and environmental data, whereas the company method uses high-frequency data via primarily a direct power comparison approach that relies less on the environmental data. When applying the respective methods to two inland wind farms, each of which presents four pairs of turbines, the quantification results from the two methods are surprisingly consistent. We believe the consistent outcome presents a strong case of cross validation, testifying to the respective method's capability and credibility.

1. Introduction

The wind industry has long been aware of the vortex generator (VG) technology and the potential benefit that a VG may bring to wind power production [1–3]. One challenging question remaining elusive to the wind industry is how much benefit, if any at all, VG installation brings to operators under the commercially operating conditions. The focus of this paper is not to answer that question directly, but to present a pair of methods which we believe are in the best position to answer the question.

There are two major difficulties in tackling this challenging problem. The first is how to model and estimate the effect of VG installation using turbine operational data. The second is how to validate the estimated effect.

Although the precise magnitude of the benefit from VG installation is unknown, the general feeling in the industry is that it would be moderate in scale resulting in 1–5% extra wind energy production under the same wind and environmental conditions. Detecting this moderate improvement in the turbine operational data, with the presence of large amounts of noise, is not a trivial task. The International Electrotechnical Commission (IEC)'s 61400-12 standard procedure for power performance measurements [4,5] is probably the most widely used approach in the wind industry for estimating and quantifying a turbine's performance before and after VG installation. The IEC standard method is, however, ineffective in this endeavor, which has been noticed by industrial practitioners and documented in previous studies [6,7]. IEC admits that "*Depending on site conditions and climate, the uncertainty may amount to several percent*" [4]. Based on other empirical studies, for example [8], one can expect a typical measurement uncertainty of 3–5% in flat terrain and 4–8% in complex terrain.

The IEC method's ineffectiveness is rooted in its lack of control of the influence of environmental factors other than wind speed.

* Corresponding author.
** Corresponding author.
E-mail addresses: hhwangbo@tamu.edu (H. Hwangbo), yuding@tamu.edu (Y. Ding), o.eisele@smart-blade.com (O. Eisele), g.weinzierl@smart-blade.com (G. Weinzierl), u.lang@smart-blade.com (U. Lang), g.pechli@smart-blade.com (G. Pechlivanoglou).

http://dx.doi.org/10.1016/j.renene.2017.07.009
0960-1481/

Considering the stochastic nature of the energy production system, one cannot evaluate the performance of wind turbine generators simply by comparing their power output, as the input conditions change all the time. At the very minimum, the influence of wind speed must be controlled while comparing power output, suggesting the use of power curve in turbine performance evaluation. That is in fact what the IEC method is based on. In addition to wind speed, however, other environmental factors such as wind direction, air density, humidity, turbulence intensity, and wind shear all potentially affect wind power production. The IEC method does take air density into account through the use of air density adjusted wind speed and controls the wind direction by filtering for clean wind sectors. But the study in [9] shows that, while wind direction and air density are indeed important factors, IEC's approach in accounting for their influence is not optimal. The complexity involved in controlling for the influence of multiple environmental factors contributes primarily to the first difficulty mentioned above.

Similar complexity also contributes to the second difficulty, namely the validation difficulty. In order to validate the estimated VG effect, one ought to know the ground truth of the actual effect. For that purpose, one would ideally conduct a controlled experiment, in which all environmental conditions are set the same before and after a VG is installed, so that the difference in power outputs before and after the installation signifies the VG effect. The problem is that such a controlled experiment is impractical and will probably never be feasible, considering the sheer physical size of commercial wind turbine generators. Researchers could conduct small-scale experiments in a wind tunnel, but the amount of uncertainty generated from extrapolation of the small-scale wind tunnel test to commercial operations makes such results much less credible to use. Even for field tests, the early results, such as the up to 25% improvement in power output claimed in [1], are now believed overly optimistic.

In this study, an academic institute, Texas A&M University (TAMU), and a wind technology company, SMART BLADE® GmbH, team up to tackle this challenge. The two teams develop their own version of estimation and quantification methods. The two methods adopt different mechanisms to control for the influence of the environmental factors. TAMU's approach, labeled as Kernel PLUS, is a machine learning model-based method, which relies on a multi-dimensional power curve model [9] to account for all environmental variables measured in a commercial operation setting. SMART BLADE's approach, known as power-vs-power, is a pure data-driven method. It considers both a control turbine with no VG installation and the test turbine on which VGs are installed and then uses the control turbine as a baseline to neutralize the influence of the environmental factors. The Kernel PLUS and the power-vs-power approaches are profoundly different. The Kernel PLUS uses 10-min data and relies on both power and environmental data, whereas the power-vs-power method uses high-frequency data and does not involve the environmental data in the power difference calculation (although wind direction and speed data are needed in pre- and post-processing). When the respective methods are applied to two wind farms, each of which presents four pairs of turbines, the quantification results of the two methods are surprisingly consistent. We believe the consistent outcome presents a strong case of cross validation, testifying to the respective method's capability and credibility.

In the remaining parts of the paper, we first elaborate the details of the power-vs-power method in Section 2. Part of the Kernel PLUS method has already been published in [6]. Therefore, we present in Section 3 the method's general idea without repeating the detail. In the same section, we also explain a new step added to the Kernel PLUS method that may be needed in some circumstances. In Section 4, we present the case study showing the outcomes of the respective methods. Finally, we conclude the paper in Section 5 and discuss the pros and cons of the two methods.

2. SMART BLADE's approach: the power-vs-power method

SMART BLADE develops its own VG technology and has installed it on more than 2000 wind turbines worldwide. Their power-vs-power method has been developed and used since 2013. Their calculation was already verified by DNV-GL in two projects and the method itself is in the process of certification.

As we argued in Section 1, one key element in an effective estimation of the VG effect is how it controls for the influence of the time-varying environmental factors. The basic idea behind the power-vs-power approach is to use a pair of wind turbines that are located very closely in space. It can be reasonably assumed and subsequently proven that the pair of turbines are subject to comparable, or even identical (to some extent), wind and environmental conditions. The power-vs-power approach uses one of the turbines as the control turbine, which does not undergo any modification, while treating the other as the test turbine, or VG turbine, which has VGs installed at a certain time. By using the power output from the control turbine as a baseline, which naturally incorporates the change in wind and environmental conditions, the difference of power outputs between the control and test turbines should exhibit a change point if there is a genuine effect caused by the VG installation. Computing the change before and after the VG installation in the power difference presumably quantifies the VG effect.

Let us define some notations and terminology first. For the quantification purpose, there are two time periods corresponding to before and after the VG installation, referred to as PRE and POST, respectively. Denote y as the power output of a turbine. Denote $\mathbf{x} := (x_1, \dots, x_p)$ by the set of p environmental variables measured on a wind farm, by V the wind speed, D the wind direction, and ρ the air density. If we let V, D, and ρ be the first three elements in $\mathbf{x}$, then $x_1 = V$, $x_2 = D$, and $x_3 = \rho$. We use a subscript to indicate which turbine and a superscript to indicate which period a variable is associated with. For instance, y_{Crtl}^{PRE} means the power output from the control turbine in the time period before VG installation and V_{VG}^{POST} means the wind speed associated with the VG turbine in the time period after VG installation. When the power or wind data are binned, we use N to denote the number of bins.

The power-vs-power approach entails the following five main steps:

1. Determine the valid wind sectors and eliminate the wind and power measurements taken under wake conditions. Also apply all other data filters (Status.Flag, Yaw.Error, etc.).
2. Apply a power density normalization, namely, normalize the wind power output through $y \times \frac{\rho}{\rho_0}$, where ρ_0 is the sea-level dry air density. Air density is calcualted by $\rho = \frac{P}{R \cdot T}$ for a given air temperature, T, expressed in Kelvin, and air pressure, P, expressed in Newtons/m^2, where $R = 287$ (Joule)(kg)$^{-1}$ (Kelvin)$^{-1}$ is the gas constant [10]. Use the density-normalized power in the subsequent analysis.
3. If necessary, verify whether there is any other source of variation significantly affecting the power difference between the PRE and POST periods. If such a source of variation is identified, further reduce the dataset so that its effect is controlled for.
4. Compute the bin-wise power difference, namely, calculate the PRE and POST power production difference of the VG turbine, relative to that of the control turbine, for each of the power output bins.

5. Compute the power difference produced by the VG installation over the whole power output spectrum.

Step 1 is needed because of the assumption that, when two turbines are close enough in space, it is likely that the wind and environmental conditions they are subject to are comparable. This assumption is generally reasonable and verifiable, except for the situation when one turbine is in the wake of the other one. Step 1 is to identify the wake free conditions, also known as the valid wind sectors, in the turbine operational data set and then only use the wake free data in the subsequent analysis. The specific steps are:

- For each time stamp, compute the wind speed ratio $\frac{V_{VG}}{V_{Ctrl}}$ or the power output ratio $\frac{y_{VG}}{y_{Ctrl}}$.
- Bin the ratios based on a wind direction signal measured by a nearby met mast or the control turbine wind direction derived by yaw position and wind vane signal.
- Plot a graphical representation of the bin-wise boxplots. Identify the coherent regions (for a minimum 30-degree width) showing wind direction independence of all considered signal ratios.

Fig. 1 shows an example using the power output ratio of two turbines. The two regions in which the ratio obviously departs from a ratio of 1 suggest the presence of a strong wake effect; these regions shall be filtered out in the subsequent analysis.

Step 2 of the power-vs-power approach performs an air density normalization. The thought behind this is similar to that of using the density-normalized wind speed, as recommended by IEC [4], and the motivation for doing so is to account for the air density effect. In the case of the power-vs-power approach, no wind speed signals are involved in the power comparison step. Therefore, density normalization must be accomplished by direct normalization of the power values for the below rated region.

Step 3 is another step designed to verify and uphold the assumption that both turbines must see the same conditions and must operate similarly. In general, the assumption is reasonably valid for relatively flat terrain. However, if there is an obvious source of variation, e.g., due to severely uneven terrain, the variation should be controlled for. Consider the following example. At different altitudes caused by the uneven terrain, the pair of turbines will face inflow wind of different speeds. The difference in wind speed is due to the speed up/hill effect. As such, the power difference of both turbines varies by time, although both turbines do operate in wake free conditions. This requires an additional categorization next to the wind direction filter to split the data into sets of equal conditions. In cases where different controller modes can be identified (such as pitch mode/variable speed mode) the data set can be separately analysed for each controller mode. Data points in which both turbines show different controller modes at the same time are rejected. Afterwards, the individual results can be merged by means of a probability weighted average.

After completion of the pre-processing steps that filter, clean, and normalize the data, Step 4 of the power-vs-power approach is to compute the bin-wise power difference between the two turbines. Specifically,

- Take the high frequency power output data of the *control* turbine and partition the data into N bins by using a bin width of, say, 100 kilowatts (kW). The bin width can be adjusted for other applications, but, for megawatts capacity turbines, 100 kW appears to be a reasonable default number.
- For each bin, calculate the median of the power difference between the VG turbine and the control turbine. We understand that IEC instructs to use the mean, however, the power difference is skewed. As a result, the median is a much better statistical measure to represent the average bin-wise power performance than the mean.
- Conduct the above two steps for the PRE and POST periods individually. Denote the resulting power differences by $\Delta\bar{y}_i^{\mathrm{PRE}}$ and $\Delta\bar{y}_i^{\mathrm{POST}}$, respectively, for $i = 1, \ldots, N$.
- Conduct a PRE bin comparison between the control and test turbine to verify the performance similarities between the pair of turbines, thereby proving the initial assumptions.
- Calculate the bin-wise power difference as $\Delta\bar{y}_i = \Delta\bar{y}_i^{\mathrm{POST}} - \Delta\bar{y}_i^{\mathrm{PRE}}$, for $i = 1, \ldots, N$.

Finally, Step 5 of the power-vs-power approach combines all the bin-wise power differences by using the weights derived from the power distribution over a given year; the resulting metric serves as the estimate of the VG effect. The detailed procedure is:

- Following the standard IEC procedure [4,5], compute a power curve and the site representative probability distribution of wind speed using the measurements taken from the control

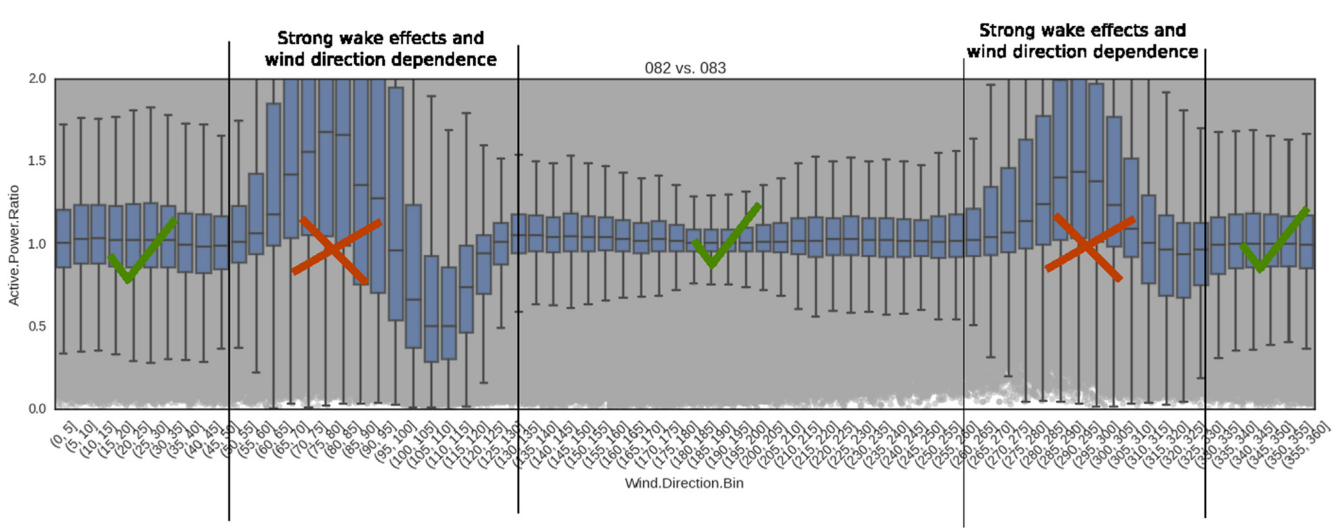

Fig. 1. Identify the wake free data using the power output ratio.

turbine. Alternatively, one can use the OEM certified reference power curve.

- Using the power curve, find the specific wind speeds $V_{i,left}$ and $V_{i,right}$ that corresponds to the lower and upper bound of the i-th power bin, respectively. Convert the wind speed distribution into a power distribution through

$$Prob(y_i) = F\left(V_{i,right}\right) - F\left(V_{i,left}\right), \tag{1}$$

where y_i is the midpoint of the i-th power bin, $F(\cdot)$ is the cumulative distribution function of wind speed, and $Prob(y_i)$ is the probability of the i-th power bin or, intuitively, the relative occurrence frequency of that particular power bin in the period of evaluation (i.e., a given year).

- Estimate the overall VG effect as

$$\Delta_{VG}[\%] = \frac{\sum_{i=1}^{N} \Delta\bar{y}_i \cdot Prob(y_i)}{\sum_{i=1}^{N} y_i \cdot Prob(y_i)} \times 100\%. \tag{2}$$

We want to note that ideas similar to the power-vs-power approach have been mentioned previously, for instance, in Refs. [11,12]. Sometimes, the ideas were called the side-by-side comparison, referring to a pair of turbines standing side by side in physical vicinity. As commented in Ref. [6], "*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.*" Despite the similarity in ideas, the difference is in the details. None of the previous works used high frequency power data (i.e., the original historian data), the use of which we believe is crucial to the success of a power-vs-power, or side-by-side, approach. For the power-vs-power (or generally, side-by-side) approach, the VG effect is varied at different power regimes of the turbine. Although a strong VG effect is expected in the mid-power-regime, regimes with no VG effect at all do exist, for instance, in the pitch region. Hence, in order to identify the true characteristics of the VG effect, a proper resolution of the power regime of the turbine is required. Because of the unsteady nature of the inflow, the turbine operates at different operational regimes during a certain period of time. This means that a proper resolution of the operational regime of the turbine can be achieved only with a proper temporal resolution. The procedure outlined in Ref. [11] is not purely data driven because certain regression models were fit, making their approach different from the power-vs-power approach delineated here. The procedure outlined in Ref. [12] lacks sufficient details with which to be compared (as [12] was a poster publication).

3. TAMU's approach: the Kernel PLUS method

TAMU's Kernel PLUS method takes a different approach to control for the influence of environmental factors. The basic idea is to establish a multi-dimensional power curve model that incorporates the effect of all environmental variables measured on a wind farm. The current version of the Kernel PLUS method uses the 10-min data instead of the high frequency data.

Typical environmental variables measured include wind speeds at the hub height as well as those above or below hub height, wind directions at hub height and sometimes at other heights, temperature, air pressure, and humidity. As shown previously, the air density is computed by using the temperature and air pressure measurements. The turbulence intensity is computed by using the hub height wind speed measurements. The wind shear is computed by using wind speed measurements at two different heights (typically, one of them is the hub height speed).

Overall, there are six elements, but likely more, in the input vector $\mathbf{x}$, which are the hub height wind speed V, hub height wind direction D, air density ρ, humidity H, turbulence intensity S, and wind shear W. As such, $\mathbf{x} = (V, D, \rho, H, S, W)$. The multi-dimensional power curve model will find a functional relationship $f(\cdot)$ that links the inputs to the power output y, i.e.,

$$y = f(\mathbf{x}) = f(V, D, \rho, H, S, W) + \varepsilon, \tag{3}$$

where ε is the residual noise in the data that cannot be fully explained by the model. Once a model in the form of equation (3) is established by using the data of the PRE period, the model is essentially the mathematical surrogate of the turbine's physical reality under the wind and environmental conditions in the PRE period. When this model fitted with PRE data is applied to the wind and environmental conditions under the POST period, namely, with V, D, ρ, H, S, and W measured in the POST period, it is equivalent to running the old, pre-upgrade turbine under the new wind and environmental conditions. Then, the difference between the model output (pre-upgrade turbine under new conditions) and the actual power measurement (VG upgraded turbine under the same new conditions) exhibits the VG effect.

TAMU's Kernel PLUS method has been published in [6]. The specific mechanism used to model and learn $f(\cdot)$ from a set of data is based on the Kernel regression method [13,14], one of the methods broadly used in the machine learning field [15]. The method in [6] tailored a special model structure for wind applications and also included a self-calibration step to alleviate any potential bias introduced by the use of the Kernel regression. The resulting method is different from a standard version of the Kernel regression, and was therefore given the name "Kernel PLUS."

Considering that the full details of Kernel PLUS are available in [6], we will not repeat them here again. We do want to note a blind study reported in [7], in which Kernel PLUS was applied to three sets of turbine operational data without any prior knowledge of how many, and if any, which, turbines had undergone VG installation. The Kernel PLUS method was able to identify the right VG turbine in that blind study.

While working on this joint study, we realized that one more issue needs to be discussed, which has not yet been addressed in the Kernel PLUS as published in [6]. The issue concerns the use of wind speed. For the set of wind and environmental measurements discussed above, the wind measurements may be from either a nearby mast or the nacelle, whereas the other measurements are from the mast. The wind measurement, if from nacelle, is in the wake of the rotor. Installation of VGs adjusts wind flow separation behind the rotor, so much so that for the same free inflow wind in front of the rotor, the wind speed measurements taken by the nacelle anemometer before and after the VG installation are most likely different. This difference could introduce a degree of inaccuracy if left unaddressed. For this reason, we propose to add a wind speed correction step prior to the use of the Kernel PLUS method for the circumstances where the wind speed measurements are taken from a nacelle anemometer. The wind speed correction step is explained below.

We understand that IEC 61400-12-2 [5] deals with nacelle measurements through a nacelle transfer function (NTF), which is the relation between the free inflow wind speed and that measured at the nacelle anemometer. Typically a NTF can be obtained by comparing the nacelle measurements with that on a nearby mast or with a nacelle mounted LIDAR. Some operators establish a NTF for a VG turbine, so that the wind speed correction is not needed. That is what happened in our previous studies, as our industrial partner provided us the adjusted wind speed through the use of a NTF, even for their VG turbines. However, if the NTF was not established, then a wind speed correction, as described below, is needed.

The idea for wind speed correction is to make use of the wind speed measurements from the control turbine. Because the control turbine does not undergo the VG installation, the difference between its nacelle anemometer measurements and their free stream counterparts should stay the same in the PRE and POST periods, all other conditions being equal. In light of this thought, we would like to be able to assume the following relationship:

$$V_{VG}^{\text{POST}} - V_{VG}^{\text{PRE}} = V_{Ctrl}^{\text{POST}} - V_{Ctrl}^{\text{PRE}} + C, \tag{4}$$

where the offset (i.e., the correction value), C, represents the impact of VG installation on the nacelle wind speed measurements. The equivalent nacelle wind speed on the VG turbine after the VG installation is supposed to be:

$$\tilde{V}_{VG}^{\text{POST}} = V_{VG}^{\text{POST}} - C. \tag{5}$$

In order to assume the relationship in (4), the underlying assumption is that the difference between $V_{VG}^{\text{POST}} - V_{VG}^{\text{PRE}}$ and $V_{Ctrl}^{\text{POST}} - V_{Ctrl}^{\text{PRE}}$, or equivalently, the difference between $V_{VG}^{POST} - V_{Ctrl}^{POST}$ and $V_{VG}^{\text{PRE}} - V_{Ctrl}^{\text{PRE}}$, is the consequence of the VG installation. This assumption is, however, not always true. If the VG turbine is in the wake of the control turbine during the PRE period but not in the wake during the POST period, then, even without VG installation, there is a difference between $V_{VG}^{\text{POST}} - V_{Ctrl}^{\text{POST}}$ and $V_{VG}^{\text{PRE}} - V_{Ctrl}^{\text{PRE}}$. To alleviate this problem, we first match the wind directions observed during the PRE and POST periods. By doing so, one turbine is exposed to the other turbine's wake at a similar frequency in both periods, rendering the wind speed differences comparable on average. In the following, we explain i) how to align the probability densities of wind direction and ii) how to adjust the wind speed for the VG turbine after VG installation.

3.1. Aligning probability densities of wind direction

Matching the wind direction is done in a probabilistic sense. That is to say, we match the probability density function of the wind directions in the PRE period with that in the POST period.

Let D_i^{PRE} for $i = 1, \ldots, n^{\text{PRE}}$ and D_j^{POST} for $j = 1, \ldots, n^{\text{POST}}$ denote the wind direction measurements during the PRE and POST periods, respectively, where n denotes the number of data points with the respective superscript indicating the period it corresponds to. To align the probability densities for the two periods, a subset will be taken from each period's dataset, so that the subsets have a comparable probability density in terms of wind directions. This can be achieved by matching individual observations in the POST period with an observation in the PRE period based on their observed wind directions.

Provided a D_j^{POST}, a dissimilarity score is calculated by measuring how different it is from D_i^{PRE}, for each and every $i = 1, \ldots, n^{\text{PRE}}$. We propose to use the dissimilarity score, S_i, defined below:

$$S_i = \frac{min\left\{\left|D_j^{\text{POST}} - D_i^{\text{PRE}}\right|, 360 - \left|D_j^{\text{POST}} - D_i^{\text{PRE}}\right|\right\}}{D_j^{\text{POST}}}. \tag{6}$$

The dissimilarity score measures the relative difference between the two wind direction values. Because wind direction values are circular in nature, the above formula makes sure that the largest degree difference between two wind directions are 180° (namely, opposite directions). Then, the wind direction in the PRE period with the smallest S_i is chosen as a candidate match for the jth wind direction observation in the POST period. To avoid matching observations with an unsatisfactorily large dissimilarity score, the candidate match is compared with a prescribed threshold, α; only when the dissimilarity score is smaller than α, a match is declared. Specifically, equation (7) below decides which wind direction in the PRE period matches the jth wind direction observation in the POST period:

$$i^*(j) = \underset{i \in \{1,\ldots,n^{\text{PRE}}\}}{\operatorname{argmin}} \{S_i : S_i \le \alpha\}. \tag{7}$$

The threshold α can be determined as a fraction of the coefficient of variation, i.e., $\alpha = c \cdot (\sigma_D^{\text{POST}} / \overline{D}^{\text{POST}})$ where $\overline{D}^{\text{POST}}$ and σ_D^{POST} are the mean and the standard deviation of $\{D_j^{\text{POST}}, j = 1, \ldots, n^{\text{POST}}\}$, respectively. Based on our experience, a $c = 0.25$ provides a reasonable match.

Once $i^*(j) \in \{1, \ldots, n^{\text{PRE}}\}$ is chosen, the corresponding data point in the PRE dataset is removed and will not be considered as a candidate match to another observation in the POST period. Let $\mathscr{J}$ define the set of the observations in the POST period that have a match and $\mathscr{I}$ define the set of the matched observations in the PRE period, i.e., the set of all $i^*(j)$'s. Then, $\mathscr{I}$ and $\mathscr{J}$ are the respective subset of the PRE and POST period; only the observations in the two subsets are used to calculate the offset value C in the following section.

3.2. Calculating the wind speed offset C

In addition to the influence of wind direction, the offset C also depends on the wind speed value, meaning that the offset for wind speed, say of 7 m/s, could be different from that for wind speed of 10 m/s. Because of this, we first bin the wind speed and then calculate the bin-wise offset values to account for this inhomogeneity across the wind speed spectrum. As the control turbine is free of the VG effect rendering its wind speed measurements more comparable between the PRE and POST periods, the reference wind speed in our action of binning is that of the control turbine.

Denote m as the number of bins, with $b = 1, \ldots, m$ as the index for binning, and $\mathscr{I}_b \subset \mathscr{I}$ and $\mathscr{J}_b \subset \mathscr{J}$ the subsets of data belonging to the b-th bin of the respective PRE and POST datasets. Let the number of observations in $\mathscr{I}_b$ and $\mathscr{J}_b$ be n_b^{PRE} and n_b^{POST}, $b = 1, \ldots, m$, respectively.

If equation (4) is true, then it implies $C = [V_{VG}^{\text{POST}} - V_{Ctrl}^{\text{POST}}] - [V_{VG}^{\text{PRE}} - V_{Ctrl}^{\text{PRE}}]$. As such, the bin-wise offset C_b can be estimated by using the sample averages as follows:

$$C_b = \frac{1}{n_b^{\text{POST}}} \sum_{j \in \mathscr{J}_b} \left(V_{VG,j}^{\text{POST}} - V_{Ctrl,j}^{\text{POST}}\right) - \frac{1}{n_b^{\text{PRE}}} \sum_{i \in \mathscr{I}_b} \left(V_{VG,i}^{\text{PRE}} - V_{Ctrl,i}^{\text{PRE}}\right). \tag{8}$$

Once C_b's, $b = 1, \ldots, m$, are estimated, they are used to adjust the wind speed measurements taken from the VG turbine during the POST period, namely that for $b = 1, \ldots, m$,

$$\tilde{V}_{VG,j}^{\text{POST}} = V_{VG,j}^{\text{POST}} - C_b, \quad \text{if } j \text{ belongs to bin } b. \tag{9}$$

In summary, the revised Kernel PLUS method entails the following step:

1. Match the probability distribution of the wind direction in the PRE and POST periods.

2. Calculate the bin-wise wind speed offsets, C_b's.
3. Adjust the wind speed of the VG turbine in the POST period.
4. Apply the Kernel PLUS method as in [6].

The Kernel PLUS method can be individually applied to the VG turbine data and the control turbine data, producing two percentage improvement values, $\Delta_{VG}[\%]$ and $\Delta_{Ctrl}[\%]$, respectively. Since the power-vs-power approach reports the VG turbine improvement relative to the control turbine, we recommend using the difference, $\Delta_{VG}[\%] - \Delta_{Ctrl}[\%]$, as the final Kernel PLUS's estimation of the VG effect to make it consistent with that of the power-vs-power approach.

4. Analysis outcome

In this case study, we have turbine pairs from two different wind farms. Both wind farms are inland but of different terrain complexity. Each farm presents four pairs of wind turbines, and each pair comprises a control turbine and a VG turbine co-located on a wind farm. The historian data was collected in high temporal resolution ($\sim 0.01 - 1$ Hz) with no averaging applied; this is the high frequency data referred to earlier in this report. The 10-min data is produced from the historian data. The power-vs-power approach uses the high frequency data, whereas the Kernel PLUS method uses the 10-min data. Periods that are known to be under curtailment were manually excluded prior to the analysis. The two teams conducted their studies independently. The data was initially provided by the respective wind farm operator to SMART BLADE, who conducted their analysis first. Then, SMART BLADE provided the data to the TAMU team for them to conduct the analysis using the Kernel PLUS method. At that stage, SMART BLADE withheld their own analysis results, meaning that the TAMU team conducted its analysis without knowing SMART BLADE's estimates of the VG effect.

4.1. Wind farm #1

The layout of the four turbine pairs on the first wind farm is illustrated in Fig. 2. The wind farm is not on flat terrain but that of medium complexity. The turbines on the farm belong to the general 2 MW turbine class. The VG installation took place in a summer month of 2014, but it was conducted on different days for each of the four VG turbines. There are six months of turbine data, including wind speed and wind power, in the PRE period and 13 months of the data in the POST period. As mentioned before, several of the environmental measurements, such as air density and humidity, were taken from the mast. Mast data is available for almost the same period of time; approximately six months in the PRE period and 13 months in the POST period. Missing data is common in all data sets and in both periods. Other details of the datasets and turbines are withheld due to the confidentiality agreement in place.

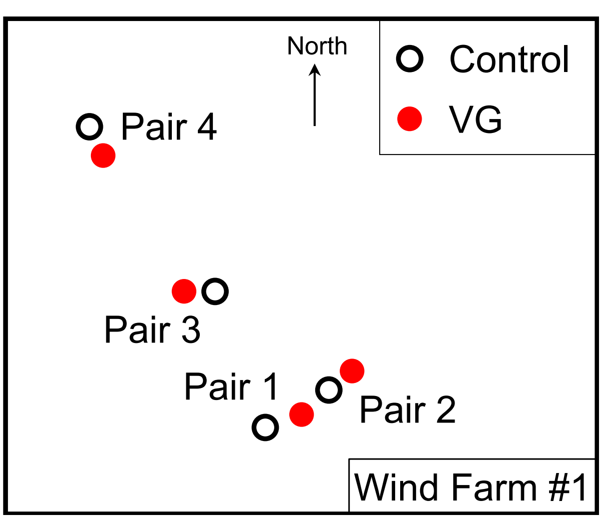

Fig. 2. Layout of the four turbine pairs on wind farm #1. The distance among the turbines are not scaled precisely, but their relative positions, as well as their locations on the farm, reflect the reality. The between-turbine distances are expressed as multiples of turbine rotor diameter, d, as follows: Pair 1, $7d$; Pair 2, $5.5d$; Pair 3, $3d$; and Pair 4, $4.5d$. The met mast is directly north of all turbine pairs. Its distance to the turbine pairs are: Pair 1 & Pair 2, 11 kilometers (km); Pair 3, 8.8 km; and Pair 4, 6 km.

Because the nacelle wind speed is used in the Kernel PLUS method, the wind speed correction procedure presented in Section 3.2 is used. The estimated VG effect on the four pairs of turbines is presented in Fig. 3. Uncertainty quantification is conducted via the bootstrap resampling method [15], and as such, 90% confidence intervals are added in the plot on top of the respective mean estimates. Understandably, the two sets of estimates are not exactly the same, but they are reasonably consistent, especially in terms of the relative significance of the VG effect on a specific turbine. The difference between the two sets of estimates are well within the margin of error, and the overall difference between the two methods, averaged over the four pairs of turbines, is about 0.86%, with the Kernel PLUS slightly overestimating relative to the power-vs-power approach.

4.2. Wind farm #2

The layout of the four turbine pairs on the second wind farm is illustrated in Fig. 4. The wind farm is in a coastal area and on relatively flat terrain. The turbines on the second farm belong to the general 2 MW turbine class. The VG installation took place in December of 2015, but it was also conducted on different days for each of the four VG turbines. The duration of the common period where both the turbine data and mast data are available is 3.5 months in the PRE period and one month in the POST period. In this analysis, because the mast is close to the turbines, we use the wind speed measurements from the mast. Of course, the rest of the environmental measurements were taken from the mast as well. The humidity was not measured on site. We therefore use the average of the humidity measurements taken from two weather stations, one located at 10 km north of the wind farm and the other at 10 km east of the farm. Missing data is also common in all data sets and in both periods. Other details of the datasets and turbines are withheld due to the confidentiality agreement in place.

The estimated VG effect on the four pairs of turbines is presented in Fig. 5, in which the confidence intervals are computed via a bootstrap resampling procedure. Again, we see consistent outcomes from the two methods: the overall difference between the two methods, averaged over the four pairs of turbines, is about 0.15%, with the Kernel PLUS still slightly overestimating relative to the power-vs-power approach.

5. Discussion

We would like to stress the merit of this academia-industry joint exercise: it presents a pair of methods to tackle a challenging, yet critically important question in the wind industry. Despite the profound difference in the underlying mechanism design and data usage, the two methods produce consistent results on two wind farm case studies, each of which presents four turbine pairs. We want to emphasize that the raw data has a considerable amount of noise. Yet, our two respective methods differ, on average, 0.86% and

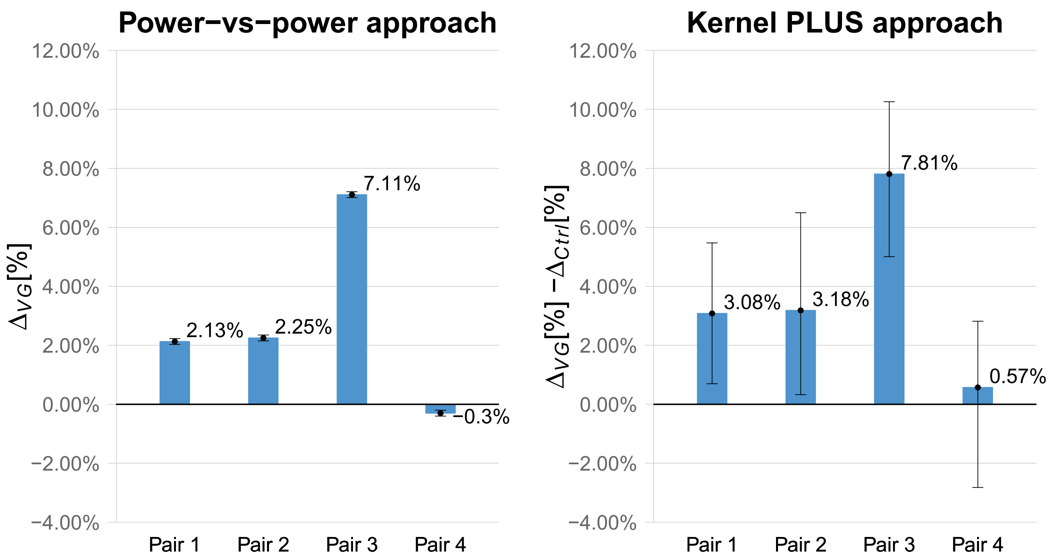

Fig. 3. Estimates of the VG effects, together with the respective 90% confidence intervals, on the four pairs of turbines on wind farm #1.

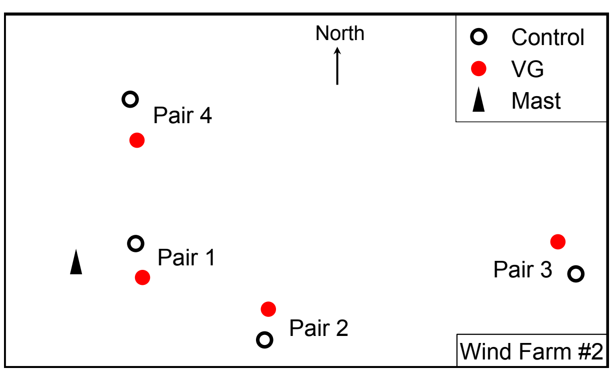

Fig. 4. Layout of the four turbine pairs on wind farm #2. The between-turbine distances are: Pair 1, 3*d*; Pair 2, 3.3*d*; Pair 3, 3.3*d*; and Pair 4, 3.7*d*. The met mast's distance to the turbine pairs are: Pair 1, 0.2 km, Pair 2, 1.3 km; Pair 3, 3.6 km; and Pair 4, 1.3 km.

0.15%, respectively, and no individual pair has a difference greater than 1%. We hope that this consistency presents a degree of clarity and credibility on the estimated VG effect and helps address the validation difficulty in the general study of estimating the VG effect in the field.

While bringing up the side-by-side approach in [6], the academia authors of this report commented previously that "*It would be interesting and valuable to conduct a carefully devised comparison study to determine which method [Kernel PLUS versus side-by-side (or power-vs-power)] is more effective and robust in practice.*" We suppose that the case study presented here fulfills that wish with the conclusion being, when carefully designed, that the two methods could be both effective and produce comparable results.

We call attention to the pros and cons of both methods. The power-vs-power method is a data-driven method. The procedure is

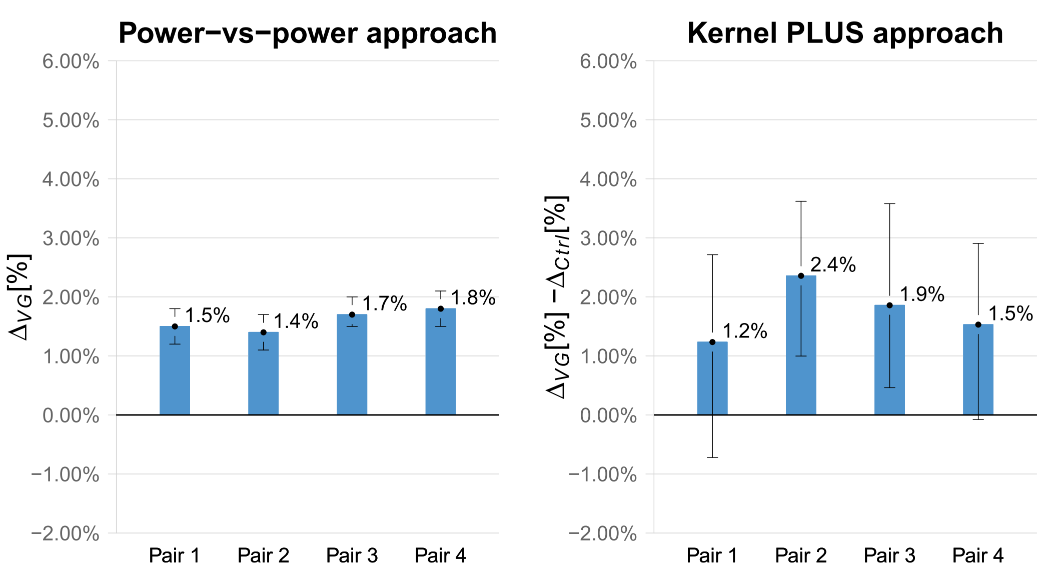

Fig. 5. Estimates of the VG effects, together with the respective 90% confidence intervals, on the four pairs of turbines on wind farm #2.

simple and relies on fewer assumptions. We can see why practitioners could be in favor of this approach. By using the high frequency data and having a larger sample size, the power-vs-power approach produces VG estimates with tighter confidence intervals. The main assumption for the power-vs-power approach is that the control and VG turbines are subject to comparable wind and environmental conditions, which is ensured by using the data in the valid wind sector and checking wind turbine operations. On the other hand, the power-vs-power approach can only be used for a pair of turbines. For small wind farms on which there are too few turbines and no two of them form a reasonable pair due to, among other reasons, the relatively large distance between them, the power-vs-power approach will no longer be applicable.

The Kernel PLUS method can possibly be applied to a single turbine when a control turbine does not exist. This explains why the Kernel PLUS passed the blind test [7] in which no prior knowledge of control and VG turbines was given and no turbine pairs were provided. Yet, it is always beneficial to have a control turbine, whenever possible, as an additional reference. The premise of Kernel PLUS is that it controls for the influence of the environmental factors through the learning of $f(\cdot)$, but the number of inputs currently included in $f(\cdot)$ may not be comprehensive enough. It is, of course, possible that measurements of certain environmental factors are not available on a wind farm or, for the time being, people may not realize the importance of other environmental factors. Still, the studies conducted so far (here and previously) show that using the current six input variables produces a remarkably competent model. The method itself, however, is certainly open to accommodating other input variables as they become available in operations.

For a wind farm without a mast or with a malfunctioning mast, many of the environmental measurements may be unavailable. This will make the Kernel PLUS inapplicable. The power-vs-power approach can still be used without a mast. The approach does need wind measurements, as wind direction is used in Step 1, and wind speed (through the use of power curve) is used in Step 5, but the wind measurements could be obtained at a nacelle instead of at a mast.

The Kernel PLUS method makes a connection with the IEC standard procedure. Briefly speaking, the Kernel PLUS method can be considered as a generalization of the IEC procedure. As explained in [7], the Kernel regression uses a smoothly curved window to produce a weighted average of all the data points falling into that window, whereas the IEC standard procedure can be viewed as using a step-function window to produce an equally weighted average (a step function gives equal weights to all data points in the window); please refer to Fig. 5 in [7] for an illustration. In addition to the window type difference, Kernel PLUS extends the input dimensionality from one to multiple inputs. Combining these two aspects, the IEC standard procedure can be considered as a one-dimensional special case of the Kernel PLUS method using a step-function window. Understanding this connection makes it easier to appreciate why Kernel PLUS can do better where the IEC procedure falls short.

One more observation we would like to discuss is that, while the general understanding of the VG effect is an extra 1–5% power production, we do see a greater than 7% improvement on Pair #3 of Farm 1, but at the same time, a near 0% effect on Pair #4 of Farm 1. As noted before, Wind Farm 1 is a medium complexity site that makes the wind inflow conditions complicated. Our studies indicate that the VG effect tends to be greater when the wind inflow condition is more turbulent on a complex terrain. Because the IEC recommends using the clean wind sector data that are less turbulent, this partially explains why the IEC method, for VG effect quantification, usually produces a smaller estimate of the effect.

The existence of this variation also suggests the importance of testing on more than one pair of turbines to get a general sense of the VG benefit through a site specific average, which, in this case, is about 2.80% based on the power-vs-power approach and 3.66% based on the Kernel PLUS. The range of site-averaged VG effects is consistent with the current understanding in industrial practice. On the other hand, the performance of the second site, which is flat and at which wind inflow conditions are simpler and measured with higher confidence, the VG effects fall into a much narrower range, with the site average at 1.60% based on the power-vs-power approach and 1.75% based on the Kernel PLUS method. The difference between the two methods tends to be greater when the terrain is more complex and this tendency is more accurately reflected in the farm-level averages rather than in the difference between an individual turbine pair.

Another reason for using site-specific averages for decision making is because we do anticipate that the difference between the two methods on some individual turbines may be greater than those on others. But, the averaged difference from a few turbine pairs on the same farm is more stable. The use of the farm-level average irons out potential biases and reduces variability. The decision for wind farm owners/operators is not whether to install VGs on a particular turbine, but rather, whether to install VGs on the tens or even hundreds of turbines on their wind farm. For that purpose, the site-specific average is a more indicative metric.

Acknowledgement

Hwangbo and Ding's research was partially supported by the U.S. National Science Foundation under grant no. CMMI-1300560.

References

[1] S. Øye, The effect of vortex generators on the performance of the ELKRAFT 1000 kw turbine, in: Aerodynamics of Wind Turbines: 9th IEA Symposium, December 11-12, Stockholm, Sweden, 1995, pp. 9–14.
[2] C.M. Velte, M.O.L. Hansen, K.E. Meyer, P. Fuglsang, Evaluation of the performance of vortex generators on the DU 91-W2-250 profile using stereoscopic PIV, in: Proceedings of the 12th World Multi-conference on Systemics, Cybernetics and Informatics (WMSCI 2008), vol. 2, 2008, pp. 263–267. June 29-July 2, Orlando, Florida, USA.
[3] H. Mueller-Vahl, G. Pechlivanoglou, C. Nayeri, C. Paschereit, Vortex generators for wind turbine blades: a combined wind tunnel and wind turbine parametric study, in: Proceedings of ASME Turbo Expo 2015: Turbine Technical Conference and Exposition, June 11-15, Copenhagen, Denmark, 2012, pp. 899–914.
[4] IEC 61400-12-1, Wind Turbines - Part 12-1: Power Performance Measurements of Electricity Producing Wind Turbines, International Electrotechnical Commission, Geneva, Switzerland, 2005.
[5] IEC 61400-12-2, Wind Turbines - Part 12-2: Power Performance of Electricity Producing Wind Turbines Based on Nacelle Anemometry, International Electrotechnical Commission, Geneva, Switzerland, 2013.
[6] G. Lee, Y. Ding, L. Xie, M.G. Genton, Kernel plus method for quantifying wind turbine upgrades, Wind Energy 18 (2015) 1207–1219.
[7] Y. Ding, J. Tang, J.Z. Huang, Data analytics methods for wind energy applications, in: Proceedings of ASME Turbo Expo 2015: Turbine Technical Conference and Exposition, June 15–19, Montreal, Canada, 2015 pp. GT2015–43286.
[8] M. Young, Power curve measurement experiences, and new approaches (a workshop presentation), in: Proceedings of the EWEA Technology Wrokshop: Resource Assessment 2013, June 25–26, Dublin, Ireland, 2013 available at: http://www.ewea.org/events/workshops/wp-content/uploads/2013/06/EWEA-RA2013-Dublin-3-3-Mark-Young-DNV-Kema.pdf.
[9] G. Lee, Y. Ding, M.G. Genton, L. Xie, Power curve estimation with multivariate environmental variables for inland and offshore wind farms, J. Am. Stat. Assoc. 110 (2015) 56–67.
[10] O. Uluyol, G. Parthasarathy, W. Foslien, K. Kim, Power curve analytic for wind turbine performance monitoring and prognostics, Annu. Conf. Progn. Health Manag. Soc. 2 (2011). Publication Control Number 049.
[11] A. Albers, Relative and integral wind turbine power performance evaluation, in: Proceedings of the 2012 European Wind Energy Conference & Exhibition, November 22-25, London, UK, 2004.
[12] I. Antoniou, R. Downey, H. Paulsen, P. Enevoldsen, A. Piperas, L. Jensen, Improved power performance of upgraded turbine rotors: a full scale

verification study (a poster presentation), in: Proceedings of the 2012 European Wind Energy Conference & Exhibition, April 16–19, Copenhagen, Denmark, 2012.
[13] E. Nadaraya, On estimating regression, Theory Probab. Appl. 9 (1964) 141–142, http://dx.doi.org/10.1137/1109020.
[14] G. Watson, Smooth regression analysis, Sankhyā Indian J. Stat., Ser. A 26 (1964) 359–372, http://dx.doi.org/10.2307/25049340.
[15] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction, second ed.ction, Springer, New York, 2009, http://dx.doi.org/10.1007/978-0-387-84858-7.