LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-667088

Wind Power Curve Modeling in Simple and Complex Terrain

V. Bulaevskaya, S. Wharton, Z. Irons, G. Qualley

February 10, 2015

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Wind Power Curve Modeling in Simple and Complex Terrain Using Statistical Models

Vera Bulaevskaya[1], Sonia Wharton[1], Zackary Irons[2], and Grant Qualley[3]

[1]Lawrence Livermore National Laboratory, Livermore, CA
[2]Enel Green Power North America, Andover, MA
[3]Pentalum, Colleyville, TX

February 12, 2015

LLNL-TR-667088

Abstract

Our previous work on wind power curve modeling using statistical models focused on a location with a moderately complex terrain in the Altamont Pass region in northern California (CA). The work described here is the follow-up to that work, but at a location with a simple terrain in northern Oklahoma (OK). The goal of the present analysis was to determine the gain in predictive ability afforded by adding information beyond the hub-height wind speed, such as wind speeds at other heights, as well as other atmospheric variables, to the power prediction model at this new location and compare the results to those obtained at the CA site in the previous study. While we reach some of the same conclusions at both sites, many results reported for the CA site do not hold at the OK site. In particular, using the entire vertical profile of wind speeds improves the accuracy of wind power prediction relative to using the hub-height wind speed alone at both sites. However, in contrast to the CA site, the rotor equivalent wind speed (REWS) performs almost as well as the entire profile at the OK site. Another difference is that at the CA site, adding wind veer as a predictor significantly improved the power prediction accuracy. The same was true for that site when air density was added to the model separately instead of using the standard air density adjustment. At the OK site, these additional variables result in no significant benefit for the prediction accuracy.

1 Introduction

Our previous work on wind power curve modeling using statistical models focused on a location with a moderately complex terrain in the Altamont Pass region in northern California (CA). The goal of that work was to determine the gain in predictive ability afforded by adding information beyond the hub-height wind speed, such as wind speeds at other heights, as well as other atmospheric variables, to the power prediction model at that location. We found that considerable improvements in prediction accuracy can be achieved both through the addition of predictors other than the hub-height wind speed and the use of statistical models. In particular, we found that expanding the wind speed information from hub-height wind speed alone to a vertical profile of wind speed measurements at 12 heights reduced the root-mean-squared error (RMSE) by 2% of the rated power in any given 10-minute period. In addition, adding wind veer as an input to the power curve model yielded significant reductions in RMSE for any set of wind speed inputs. The same was true for using air density as a separate input instead of using the air density adjustment to wind speed recommended by the International Electrotechnical Commission (IEC) [1]. The study and its conclusions, as well as relevant prior work done by various authors, are all described in detail in [2]. In this report, we repeat the same analysis for a new location, in northern Oklahoma (OK), where the terrain is simple in contrast to the CA site. We compare the results obtained at the two sites.

2 Site and Data Description

2.1 California (CA) site

The wind farm studied in [2] is located in the Altamont Pass Wind Region in Northern California, approximately 70 km east of San Francisco. The exact location and name of the site is proprietary information, and we will refer to it as the California (CA) site. The farm consists of 38 IEC Class IIA horizontal-axis turbines that are located along the top of roughly parallel ridgelines (Fig. 1). The wind turbines are blade-pitch controlled, 1-MW Mitsubishi (MWT62-1000A). The specifications of the individual turbines are listed in Table 1, and the corresponding manufacturer's power curve (MPC) is shown in Fig. 2(a).

The terrain both upwind of and within the Altamont Pass Wind Farm is moderately complex. Locally, the hills and ridgelines have maximum heights of around 400 m although most are 200 m or less. The area is covered with short, uniform grassland of low canopy roughness. Wind speeds peak in

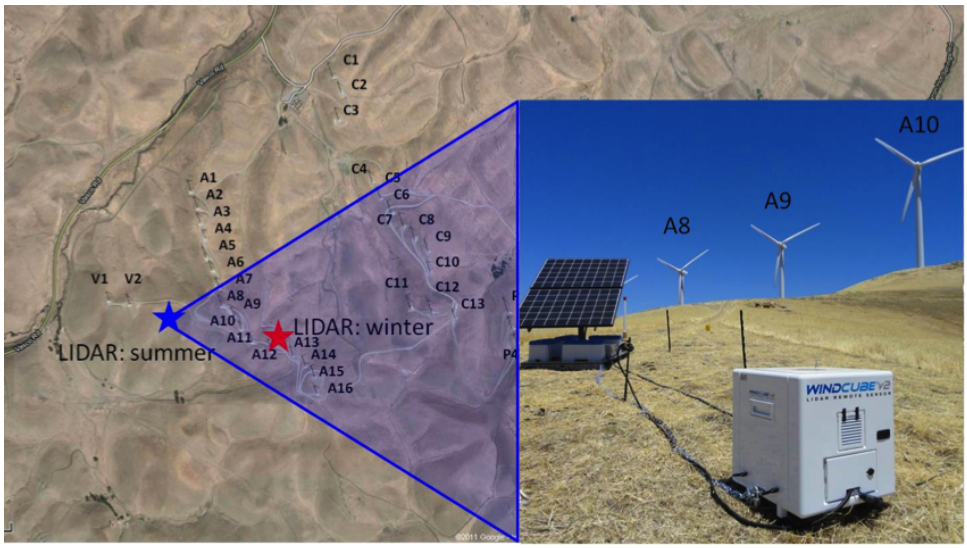

Figure 1: Satellite and photographic images of the 38-MW Altamont Pass wind farm showing the turbines, surrounding terrain and locations of the lidar campaigns including the summertime lidar deployment upwind of turbines A7–A10. The wind farm is located in the Cañada de los Vaqueros hills near the Altamont Pass in Contra Costa County, California.

(a) MWT62-1000A turbine (CA site)

(b) GE 1.68-82.5-60 Hz turbine (OK site)

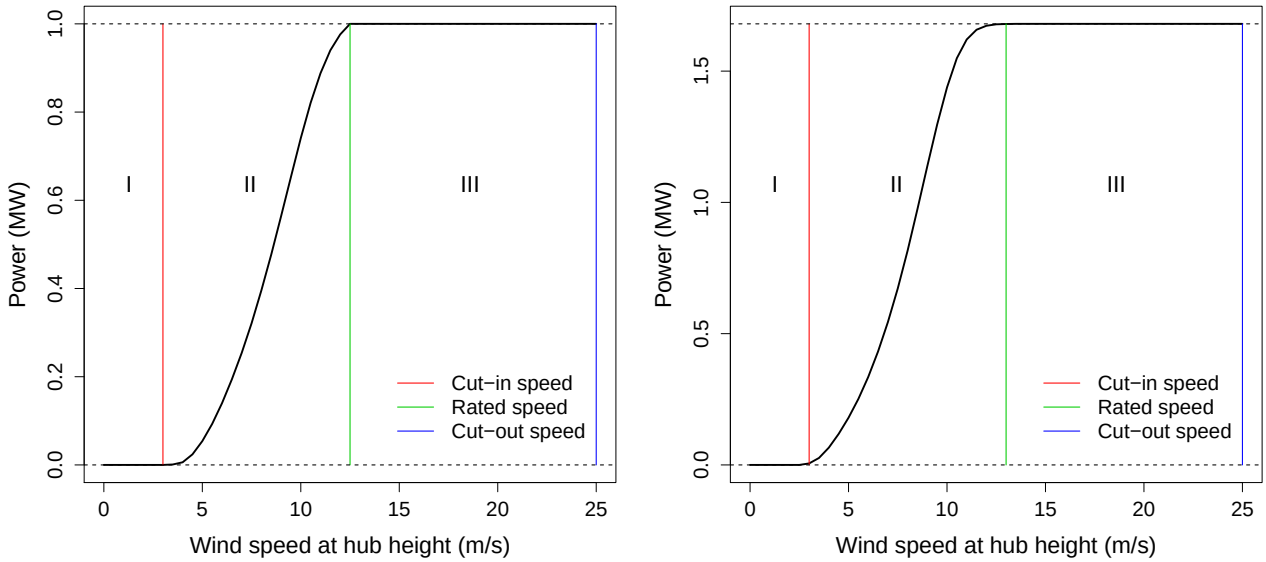

Figure 2: The MPC (power in MW vs. hub-height wind speed in m/s) for (a) the MWT62-1000A turbine [3] at the CA site and (b) the GE 1.68-82.5-60 Hz turbine [4] at the OK site, with the vertical lines indicating the cut-in, rated and cut-out speeds listed in Tables 1 and 2, respectively, and Roman numerals labeling the three regions of each power curve defined by these speeds.

Table 1: Specifications of the MWT62-1000A turbines used in the study [3].

Wind class	IEC Class IIA	Hub height	55 m
Cut-in speed	3 m/s	Rotor diameter	61.4 m
Rated speed	12.5 m/s	Blade length	29.5 m
Cut-out speed	25 m/s	Rated power	1 MW

the spring and summer afternoon hours due to strong, synoptically-driven sea breezes. During these months flow at the farm is influenced by three topographical features of decreasing scale: the California Coast Range which influences the regional upper air flow, a 9-km-long canyon just to the west of the park which alters the landscape flow, and the surrounding hills, ridges and small drainage valleys which change the local flow. Regional flow is controlled by a strong diurnal temperature contrast between the interior land and Pacific Ocean which synoptically induces a westerly flow of strong, cooler marine air.

In June through August of 2012, we deployed a pulsed Doppler light detection and ranging (lidar) instrument (Leosphere, Orsay, France) to obtain measurements of the free-stream, un-waked flow (also called inflow) as it approached a leading row of turbines (Fig. 1). The lidar provided nearly 2,000 hours of horizontal wind speed (m/s), wind direction (°), vertical wind speed (m/s) and turbulence intensity (%) measurements across the entire rotor disk. The lidar was programmed to measure from 40 m to 150 m at 12 levels spaced 10 m apart. Due to the moderately complex terrain and the placement of the turbines along the ridgeline, the lidar was deployed at an elevation lower than the turbines. This was done to ensure that the instrument was placed 2–4 rotor diameters upwind of the turbines as recommended by the International Electrotechnical Commission (IEC) 61400-12-1 report [1]. Note that complex terrain reduces the accuracy of lidar, and at this site the uncertainty in the wind speed measurements is estimated to be 3–5% based on studies of test sites with similar topography [5].

The turbines directly downwind of the lidar were A7, A8, A9, A10 and A11 depending on the wind direction at any given time (Fig. 3(a)). Turbines V1 and V2 were west of the lidar and could have waked the lidar measurements when the winds were from 270–280° (see Fig. 1). Fortunately, this wind direction was very infrequent. Instead, the strong prevailing wind direction (225–250°) meant that turbine A8 was most frequently downwind of the lidar. Careful analysis of the wind direction at each measurement height showed topographic-induced flows at heights closer to the ground. The lidar was placed at the edge of an abutment on the hill (Fig. 4(a)) and the wind direction measurements at 40 m to 70 m indicated that the flow was often redirected around the small hill feature. This caused significant wind veer (i.e., changes in wind direction with height) over these distances. Veer was significantly smaller at heights above 70 m.

The Supervisory Control and Data Acquisition (SCADA) system provided 10-minute averages of ground-level air density, hub-height wind speed and power output for all turbines, for the same 10-minute intervals as the lidar data. The SCADA data also contained other statistical summaries, including standard deviations, for these 10-minute intervals. Wind speed was measured with a nacelle-mounted cup anemometer (NRG #40, NRG Systems, Inc., Vermont, USA). The site also had one 30-m tall meteorological tower which took measurements of wind speed and direction at one height at the top of the tower. However, these measurements were not used here because the tower was not upwind of turbine A8 and there were concerns about the data quality of these measurements. As described in detail in [2], besides filtering the data to include only those points without any missing lidar and SCADA data, we also restricted our analysis to observations corresponding to 90-m wind speeds between 3.5 and 11.5 m/s (to exclude observations in regions I and III of the power curve) and 90-m wind direction in the range 225–250°.

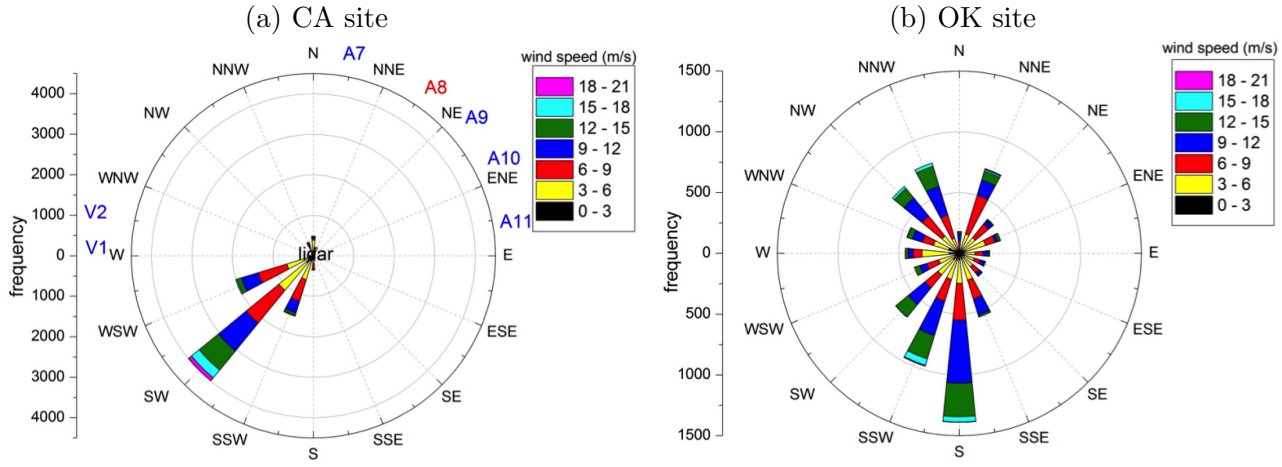

Figure 3: (a) Wind rose of 90-m wind measurements taken from the lidar during the campaign period (June–August 2012) and the relative locations of nearby turbines. The plot shows a high frequency of winds coming from the southwest (220–240°), implying that A8 (highlighted in red) is the turbine most often downwind of the lidar. (b) Wind rose of 80-m wind measurements obtained by the lidar during the campaign period (November 2013–January 2014). Westerly and easterly winds were rarely seen at the site. Winds were predominately from the north or south. The lidar measured free-stream, wake-free inflow measurements for turbine B06 during southerly winds.

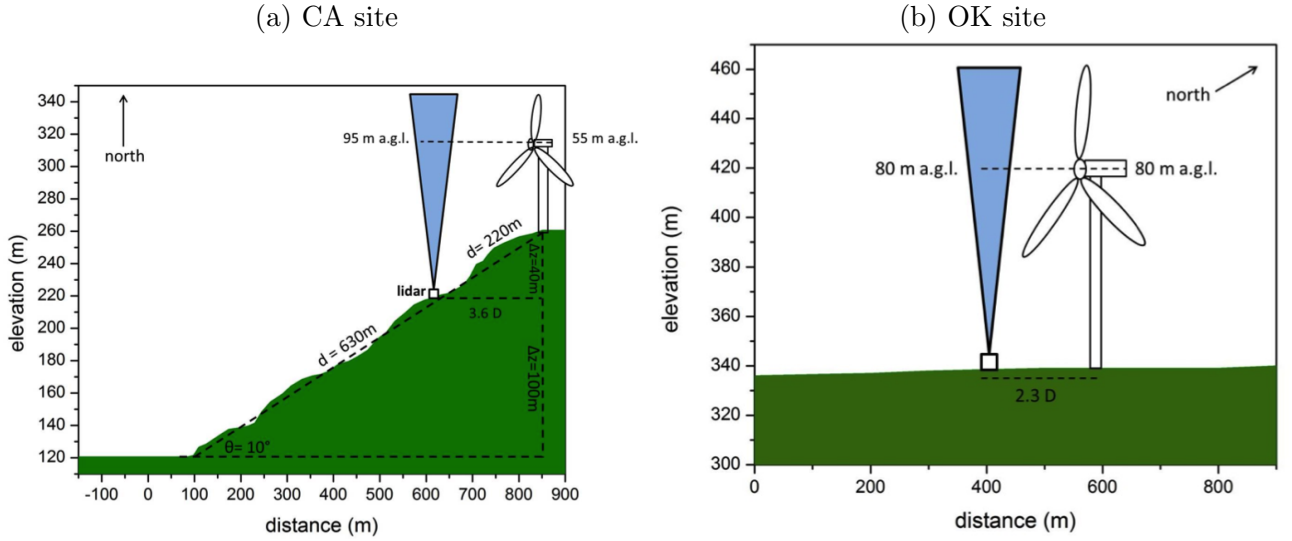

Figure 4: (a) Schematic of the hill profile and locations of the lidar and turbine A8 relative to one another, at the CA site. The lidar was located 3.6 rotor diameters (D) upwind of A8. (b) Schematic of the terrain profile and locations of the lidar and turbine B06 relative to one another, at the OK site. The lidar was located 2.3 rotor diameters (D) upwind of B06. Free-steam inflow measurements were easy to obtain at this site as the terrain influences were minimal.

2.2 Oklahoma (OK) site

The second wind farm we studied is located in north-central Oklahoma, approximately 120 km from Oklahoma City. The exact location and name of the site is proprietary information, and we will refer to it as the Oklahoma (OK) site. The farm consists of 140 horizontal-axis turbines that are spread across 25 km from west-to-east in rural Oklahoma. The wind turbines are blade-pitch controlled, GE 1.68 MW (1.68-82.5-60 Hz). The specifications of individual turbines are listed in Table 2, and the corresponding MPC is shown in Fig. 2(b).

Table 2: Specifications of the GE 1.68-82.5-60 Hz turbines used in the study [4].

Wind class	IEC Class IIB	Hub height	80 m
Cut-in speed	3 m/s	Rotor diameter	82.5 m
Rated speed	13 m/s	Blade length	40.3 m
Cut-out speed	25 m/s	Rated power	1.68 MW

The local and regional terrain is relatively simple; however, the synoptic meteorology is relatively complex and is influenced by mountain systems to the west and cold/warm frontal passages originating from the north or south. On atmospherically stable nights, a decoupled flow event, called a nocturnal low level jet (LLJ), often forms during the spring and summer months. The wind farm is located within a climatological maximum of U.S. LLJ occurrences [6]. LLJs produce maximum wind speeds frequently within the lowest 500 m of the atmosphere and can go down to heights found within the upper half of the rotor disk. Mechanically-generated turbulence is often observed below the maximum of the jet and may also penetrate heights found within the rotor disks. Wind shear and turbulence characteristics found at heights within the rotor disk are also strongly influenced by local surface heating and cooling.

A continuous-wave Doppler ZephIR 300 lidar (Zephir Ltd., Ledbury, UK) was deployed at the farm November 11, 2013–January 31, 2014. We focused on capturing inflow to a leading row turbine at the south end of the wind farm called, referred to as B06 (see Figs. 5 and 4(b)). This turbine was unwaked and downwind of the lidar and a 80-meter-tall meteorological tower when winds were southerly (135–225°). The lidar was programmed to measure at heights equivalent to the entire rotor disk (40–120 m). Exact heights were 40–130 m with measurements taken every 10 m. Measurements included 10-minute mean values of horizontal wind speed (m/s), wind direction (°), vertical wind speed (m/s), and turbulence intensity (%). Reported accuracy for the ZephIR in ideal conditions (e.g., flat, homogenous terrain) is less than 0.5% for wind speed and below 0.5° for wind direction.

The SCADA system provided 10-minute statistics (mean, maximum, minimum, standard deviation) of ground-level air density, hub-height wind speed, nacelle position, pitch angle, rotor speed, and power output for all turbines, for the same time intervals as the lidar data. Wind speed was measured with a nacelle-mounted cup anemometer in a similar fashion to the CA site.

The co-located lidar and 80-meter-tall meteorological tower measurements were first compared for quantifying instrument agreement in wind speed and direction at hub height (80 m). We found excellent agreement for both wind speed and direction (correlation of 0.99 for both). This gave us confidence in the lidar's accuracy. The 10-minute averages of lidar measurements at 80 m are thus our baseline in this work. Fig. 6 shows the distribution of the differences between the nacelle and the 80-meter lidar wind speed means. As can be seen from it, the wind speed measured at the nacelle tends to be lower than the 80-m lidar wind speed, with the median discrepancy of 0.75 m/s and the RMSE between the two measurements of 0.87 m/s. The consistently lower measurements are due to the flow obstruction at the nacelle caused by the turbine blades.

The wind direction at the site is predominantly from the north or south depending on season. During the winter the winds have a bimodal distribution between the north and south sectors (Fig. 3(b)). Southerly winds occurred approximately 45% of the time and put the first-row turbines directly down-

Figure 5: Satellite image showing the meteorological tower and first-row turbines, B04–B06 and C01. The lidar was co-located with the 80-m-tall met tower. The lidar and tower were 2.3 rotor diameters (D) (190 m) upstream of Turbine B06. The site's flat terrain is in distinct contrast to the moderately complex terrain found at the CA site.

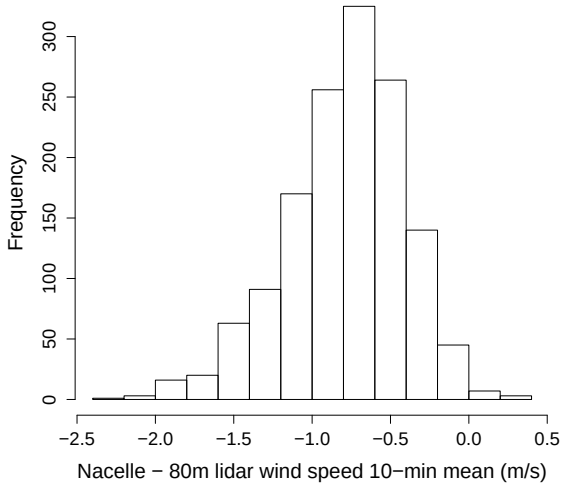

Figure 6: The distribution of the differences between the nacelle and the 80-meter lidar wind speed means (m/s), with a positive difference implying that the nacelle mean is the larger of the two (RMSE = 0.87 m/s, correlation = 0.98).

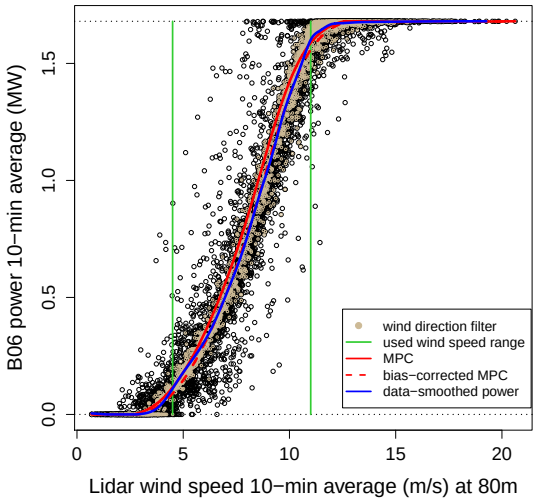

Figure 7: Ten-minute average power (MW) at turbine B06 vs. 10-minute average wind speed (m/s) measured by the lidar at 80 meters above ground (black circles). Also shown are the data filters described in Section 2, the predictions using the MPC and the bias-corrected MPC (see Section 6), and the data-smoothed power curve as a function of the wind speed mean at 80 m (see legend).

wind of the lidar. As a result, wind directions in the range 135°–225° were ideal for the lidar to capture inflow conditions seen by B06 (note that north corresponds to 0°), so the data were restricted to these wind directions. The data were further limited to the time periods without any missing SCADA data.

Moreover, as in [2], in this study we focused only on the 80-meter wind speeds that corresponded to region II, or the portion of the power curve where power is most sensitive to small changes in wind speed. For an MPC, this corresponds to the range between the cut-in and the rated hub-height wind speeds, i.e., between 3.5 m/s and 13 m/s in our study, as can be seen from Table 2 and Fig. 2(b). Since our data are 10-minute averages, we adjusted this range so that the observed 10-minute power averages are largely constant (very close to 0 or rated power) outside the chosen limits and are thus minimally contaminated by observations from regions I and III. This adjustment led us to restrict the data to 80-meter average speeds between 4.5 m/s and 11 m/s (see Fig. 2(b)). This and all the other filters discussed above resulted in 1404 data points, each representing a 10-minute period of lidar wind flow and power data.

3 Inputs to the Power Curve Model

3.1 Wind speed inputs

Since the study performed at the CA site is covered in detail in [2], for the remainder of this report we only discuss the OK study (with the exception of Section 6, where we compare and contrast the results obtained at the two sites).

Although the SCADA data provided wind speed measurements taken at the nacelle of the turbine, these measurements can deviate from the true inflow conditions because the cup anemometer is subject to flow distortion from the blades and the nacelle hub. Consequently, we considered lidar measurements of free-stream wind speed rather than the nacelle measurements as the baseline in this study. The appropriateness of using the former as the baseline is discussed in more detail in Section 6.1.

Recall from Section 2 that lidar wind speed averages measured at 80 meters were considered the baseline input in this study. The distribution of these measurements is shown in Fig. 8(a). As in [2], we applied the IEC-recommended air density adjustment [1] to the 80-meter lidar wind speed average and treated this adjusted wind speed as the baseline. The adjustment is given by

$$U_{d.adj} = U\left(\frac{\rho}{\rho_0}\right)^{1/3}, \tag{1}$$

where $U_{d.adj}$ and U are the density-adjusted and the original wind speeds (in m/s), respectively, while ρ and ρ_0 are the measured and the standard air densities (in kg/m^3), respectively ($\rho_0 = 1.225$ kg/m^3). The motivation for this adjustment is explained in more detail in [2].

First, we modeled the power output at B06 as a function of density-adjusted wind speed variables alone (using the adjustment in Eq. (1)). One way to include information about the entire vertical profile of wind speeds is to estimate the average inflow seen by the entire rotor disk of the turbine. This can be done using the rotor-equivalent wind speed (REWS), defined in [7]. As explained in more detail in [2], the REWS is simply the weighted average of the mean wind speeds in each height interval across the rotor disk, where the weight for each interval is equal to the area occupied by the interval as the fraction of the total area of the rotor disk.

Fig. 8(b) shows the distribution of the differences between the 10-minute average wind speed at 80 meters and the REWS. It reveals that these differences can be large, particularly when compared to those at the CA site (see Fig. 9(b)). For reference, the absolute difference exceeds 0.1 m/s for 25% of the observations at the OK site, compared to 6% at the OK site. This is because, as can be seen from Fig. 8(c), the magnitude of wind shear can be high at this site. Wind shear is defined as the value of α in the power-law profile of wind speed $U(z)$ as a function of height z: $U(z) = \beta z^{\alpha}$. By contrast, at the CA site examined in [2], the magnitudes of the differences between the 90-meter wind speed (the baseline wind speed input at that site) and the REWS, as well as the magnitudes of shear were much lower, as shown in Figs. 9(c) and 9(d), respectively.

As in [2], in addition to using the REWS as an input, we also explored using the entire profile of wind speed means in the model directly. In addition, we studied the effect of supplementing this profile with the profile of wind speed standard deviations, as these provide estimates of turbulence in the mean wind flow. The effects of turbulence on power and the motivation for including turbulence information in a power curve model are discussed in [2]. For reference, the distribution of one of the most commonly used measures of turbulence in the wind power industry, the turbulence intensity (TI), is shown in Fig. 8(d). It is defined as the ratio of the standard deviation to the mean of wind speeds in a 10-minute period at a given height (80 meters in the plot). Fig. 8(d) shows that for the majority of observations (71%), the TI is below 10%, in contrast to the CA site, where the equivalent fraction of observations is only 5.5%. This difference is due to the fact that the variability of the flow (as measured by the 10-minute standard deviation of the wind speeds) is smaller at the OK site than at the CA site.

As with the 80-meter wind speed mean, we applied the density adjustment in Eq. (1) to the entire vertical profile of means. In addition, an equivalent adjustment was applied to the profile of standard deviations:

$$SD_{d.adj}(U) = SD(U)\left(\frac{\rho}{1.225}\right)^{1/3}. \tag{2}$$

We explored the same five sets of wind speed variables as in [2]. These were all adjusted for air density and are henceforth collectively referred to as Group 1 (the parentheses after each input set contain the label used for that set in the plots in Section 6):

1. Wind speed mean at 80 meters (*ws 80m mean*),

2. Wind speed mean at 80 meters + standard deviation at 80 meters (*ws 80m mean + sd*),

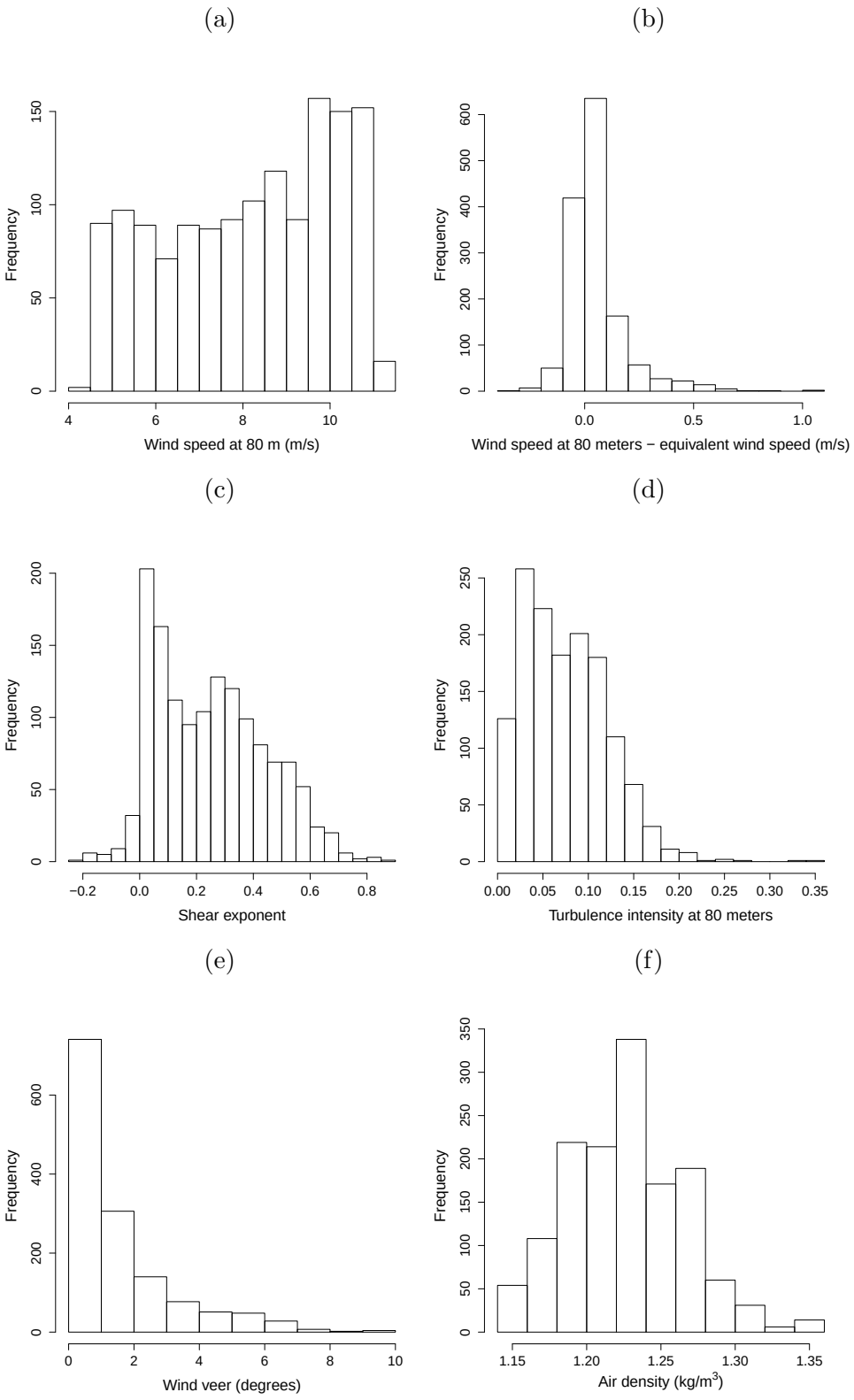

Figure 8: The observed frequency distributions of the following (based on 10-minute means) at the **OK site**: (a) 80-m wind speed (m/s), (b) difference between the 80-m wind speed and the REWS (m/s), (c) wind shear exponent α, (d) TI at 80 m, (e) wind veer (degrees), as defined in Section 3.2, and (f) ground-level air densities (kg/m^3).

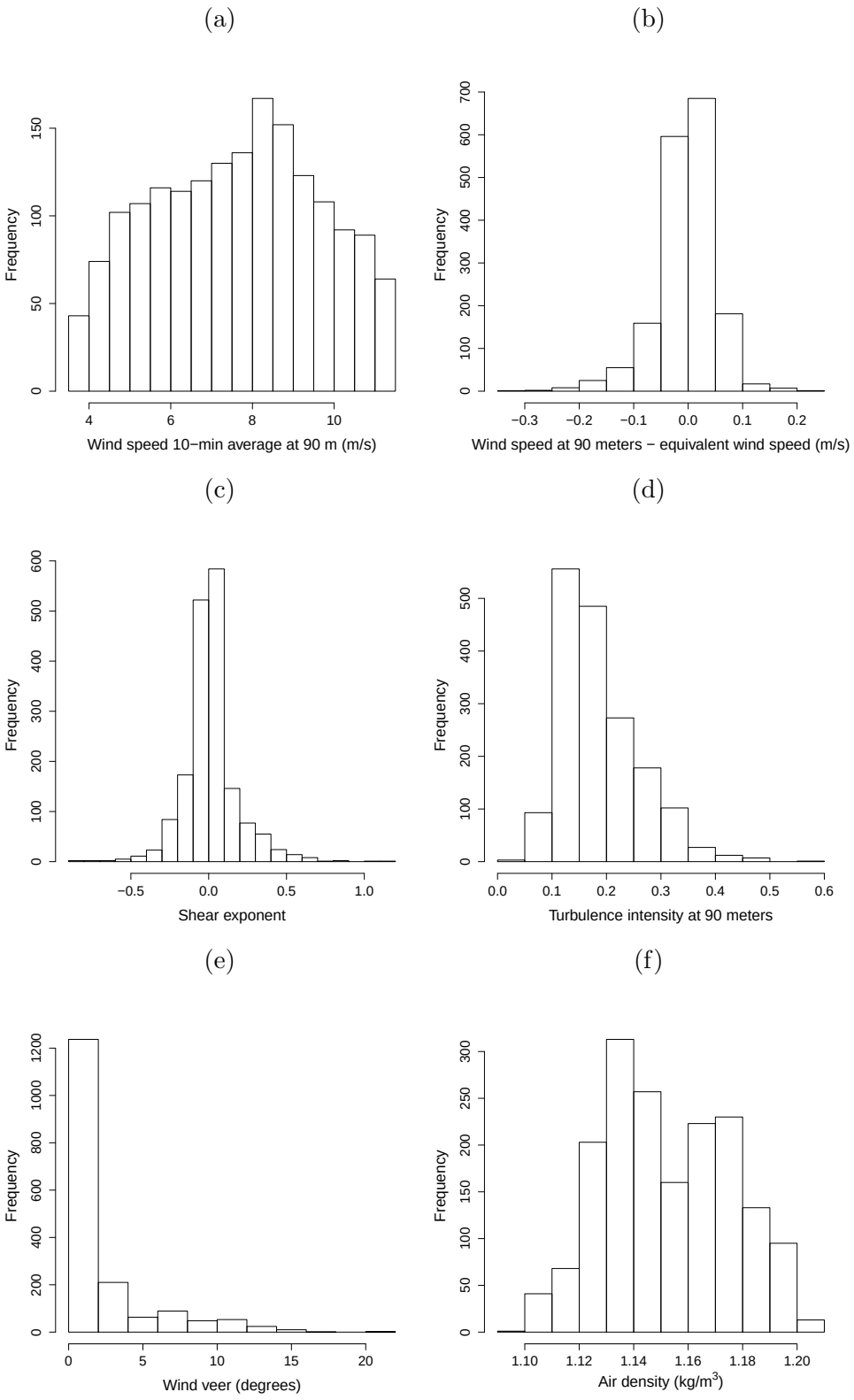

Figure 9: The observed frequency distributions of the following (based on 10-minute means) at the **CA site**: (a) 90-m wind speed (m/s), (b) difference between the 90-m wind speed and the REWS (m/s), (c) wind shear exponent α, (d) TI at 90 m, (e) wind veer (degrees), as defined in Section 3.2, and (f) ground-level air densities (kg/m^3).

3. Rotor-equivalent wind speed(*ws equiv*),

4. Entire profile of wind speed means at the 10 heights (40, 50, ..., 130 meters) (*ws all mean*) and

5. Entire profile of wind speed means + entire profile of wind speed standard deviations at the 10 heights (*ws all mean + sd*).

Note that the first set in the above list corresponds to our baseline method of modeling power as a function of the density-adjusted wind speed at the hub height alone. The other sets, on the other hand, include additional information, either in the form of the temporal variability in the wind speed (set 2) or wind speeds measured in other parts of the profile (sets 3 and 4), or both (set 5).

Furthermore, because the dimensions of the last two sets above are quite large (10 and 20 for sets 4 and 5, respectively) and the wind speeds at various heights are highly correlated (correlation values range between 0.66 and 0.99), these two sets were also subjected to a principal component analysis (PCA, see [8]) in order to reduce their dimensions, thereby greatly simplifying fitting of the statistical models discussed in Section 4. The top components that together accounted for at least 99% of the total variance were retained. Note that the aforementioned Group 1 of input sets includes the PCA-reduced versions of sets 4 and 5.

3.2 Additional atmospheric inputs

To assess the effect of wind veer in the OK data, we used the circular standard deviation (see [9]) of all wind direction averages measured by the lidar in the 50–80 meter height range as a way to quantify it. This range was used because wind direction averages at these heights had much higher circular correlations (see [9]) with the nacelle wind direction average than those at the other heights. Fig. 8(e) shows the frequency distribution of wind veer defined in this way. Relative to the CA site, wind veer at the OK site tended to be much smaller, as can be seen by comparing Figs. 8(e) and 9(e). Its effect on power at this site is thus also expected to be smaller than at the CA site. Nevertheless, we added it to each of the input sets in Group 1, thus creating another group of input sets, referred to as Group 2.

To account for the effect of air density on power, in addition to using the standard air density adjustment in Eq. (1), we also considered the approach of adding the 10-minute air density average at ground level as a separate input to the *unadjusted* versions of each of the input sets in Group 1, thus creating Group 3. The frequency distribution of air density values, shown in Fig. 8(f), reveals that the densities at this site are centered around the standard value of 1.225 kg/m^3. This is in stark contrast to the air density distribution at the CA site (shown in Fig. 9(f)), where all the density values were below the standard value. Because most of the density values in the present data are close to the standard value, we do not expect a large difference between the methods of adjusting for density in terms of the power prediction accuracy.

As in previous work, Group 4 of input sets was created by adding both wind veer and air density to the *unadjusted* wind speed input sets in Group 1. The four groups thus resulted in a total $4 \times 5 = 20$ input sets. The variables included in each group are summarized in Table 3.

4 Statistical Models

As in [2], we compared three statistical models–neural networks (NN), random forests (RF) and Gaussian process models (GPM)–in terms of their ability to predict power. Whiel Appendix A gives basic details of each model, these models and the motivation for using them for this purpose are discussed in greater detail in [2]. The reader is also referred to [10, 11], [12, 11], and [13, 14, 15] for a detailed treatment of NN, RF and GPM, respectively, for an even more detailed treatment of each model.

Table 3: Summary of the power model input variable combinations considered in our analysis. An "x" under a variable name indicates its inclusion in the group. Each group consists of 5 sets: each of the five is made up of the wind speed variables enumerated in Section 3.1 (referred to collectively as "Wind speed" in the table) and all of the variables marked with an "x". Note that in Groups 1 and 2 the air density adjustments in Eqs. (1) and (2) are applied to the wind speed variables, while no adjustment is applied to these variables in Groups 3 and 4.

Group	Input variables		
	Wind speed	Wind veer	Air density
1	x		
2	x	x	
3	x		x
4	x	x	x

5 Performance Study

To assess the performance of each of the sets of inputs and the three statistical models outlined above, the 1404 observations consisting of the lidar and SCADA data were randomly divided into two mutually exclusive sets of 702 points each, the first of which was used to train each of the models with each of the input sets in each group. For comparison, using the same training set, we also trained each of the statistical models with the 10-minute mean of the nacelle wind speed as an input.

Each combination of the trained model and input set was then used to make wind power predictions for the remaining 702 points in the data, and these predictions were compared to the observed wind power values for these points (henceforth referred to as the validation set). In addition, we obtained predictions of power using the MPC [4] (shown in Fig. 2(b)) using the same training and validation sets.

This experiment of partitioning the data, training the models with the training set and making predictions for the validation set was repeated a total of 30 times to obtain a representative set of results. For each experiment and model/input set combination, the RMSE of 10-minute average wind power predictions relative to the actual SCADA data was computed as a metric of predictive skill. The RMSE for the i^{th} experiment is given by

$$RMSE_i = \sqrt{\frac{\sum_{j=1}^{702}(P_{ij} - \hat{P}_{ij})^2}{702}}, \tag{3}$$

where P_{ij} and $\hat{P}_{ij}$ are the actual and predicted power value, respectively, for the j^{th} observation in i^{th} experiment's validation set with $i = 1, ..., 30$ and $j = 1, ..., 702$.

6 Results and Discussion

6.1 Wind speed inputs

We first consider the results for the input sets involving only the wind speed variables (Group 1). Fig. 10(a) shows the RMSEs, as fractions of the rated power (1.68 MW), for the 30 experiments described in Section 5 using lidar measurements of the wind speed as inputs to each of the statistical models. Also shown in the plot are the RMSEs for the predictions obtained using the bias-corrected MPC using the density-adjusted wind speed mean at 80 meters as an input. The bias correction is motivated by the fact that the original MPC tends to overpredict power, as can be seen from the solid red curve in Fig. 7. Such a consistent bias can be easily corrected by shifting the curve horizontally, to

(a) Lidar wind speed inputs (b) Nacelle wind speed inputs

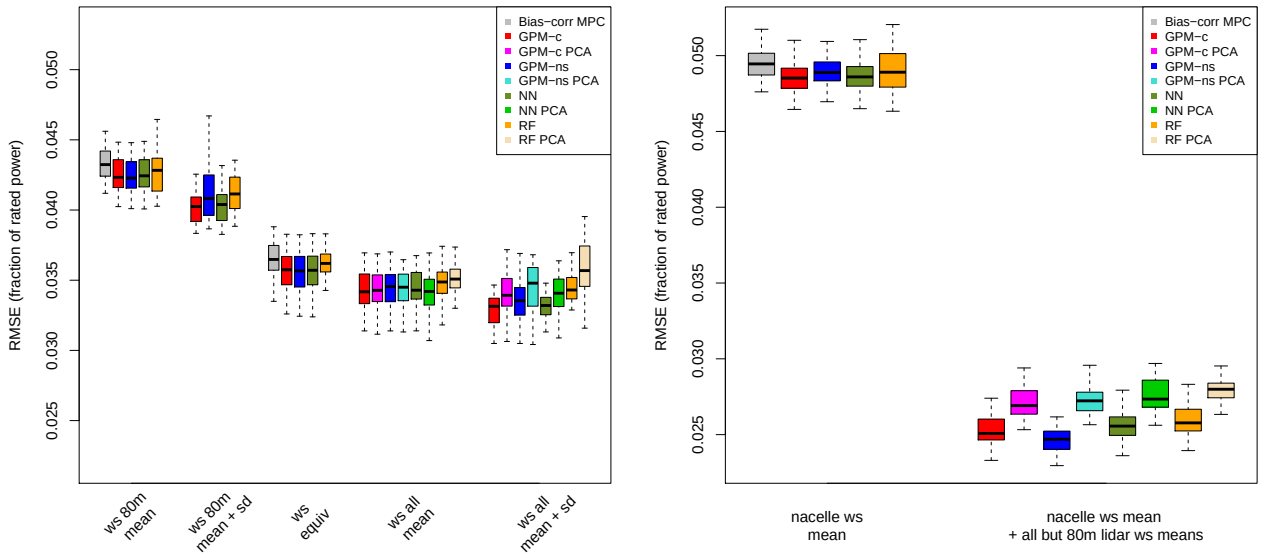

Figure 10: Boxplots of RMSE values (as fractions of rated power) at the **OK site** for the 30 experiments described in Section 5 using (a) lidar wind speed inputs (sets 1–5 in the list in Section 3.1) and (b) either nacelle only or a combination of nacelle and lidar wind speed inputs, as described in the text. The colors represent the statistical models or the MPC as indicated in the legend. The PCA-reduced versions of each model are denoted with "PCA" after the name of the model. Note that the scales on the y-axis are the same in the two plots.

the right in this case (in the case of consistent underprediction, as was the case in the CA data [2], the necessary shift is to the left). The value for the shift can be chosen by finding the value that minimizes the RMSE (or any other sensible measure of discrepancy) of the predictions relative to the data. We apply this correction in our study since we assume that any user of an MPC would do so given a sufficient amount of data, and we treat the bias-corrected MPC as the baseline. The bias-corrected MPC is shown in Fig. 7 (red dashed curve) and is close to the data-smoothed curve (in blue), indicating that the correction is reasonable.

The comparison of the RMSEs in Figs. 10(a) and 11(a) reveals that as fractions of the rated power, the RMSEs at the OK site are much smaller than those at the CA site for all input sets (note that since the rated power at the CA site is 1 MW, the absolute RMSE value at this site is equal to its value as the fraction of rated power). Even the worst-performing input set at the OK site (one consisting of the 80-meter wind speed mean alone) yields much smaller RMSEs (averaging at approximately 0.0425, or 4.25% of the rated power, across the models) than the best-performing input set (one consisting of the entire vertical profile of wind speed means and standard deviation) at the CA site (averages ranging from 0.073 to 0.08). This is not surprising since complex terrain is generally associated with more complex physical processes than simpler terrain, making wind power prediction at the former more difficult.

In terms of the relative standing of the input sets with respect to one another, some of the patterns observed for the CA site are reproduced in these data, while others diverge. As mentioned above, among the five input sets we considered, the set with just the wind speed mean at 80 meters (the 1st cluster of boxplots in Fig. 10(a) is the worst performer in terms of the RMSE, regardless of the model. This was the case in [2] for the 90-meter wind speed, which was the baseline in that study.

(a) Lidar wind speed inputs

(b) Nacelle wind speed inputs

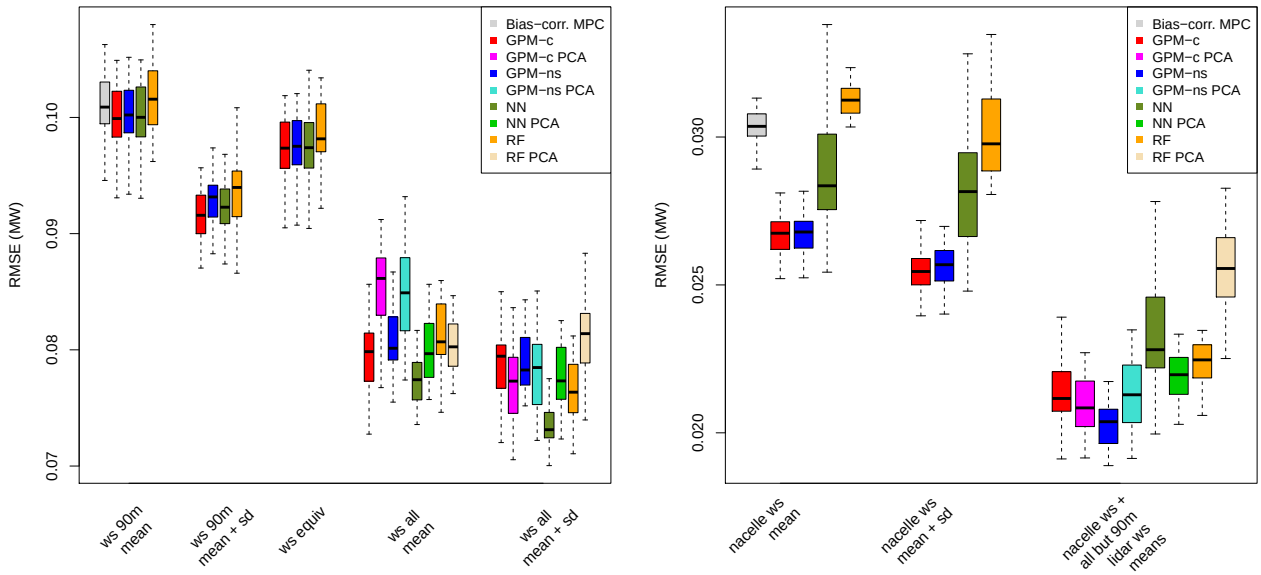

Figure 11: Same as Fig. 10, but for the **CA site**. Note that the scales on the y-axis are different in the two plots.

Moreover, as in [2], for this simplest input set, the bias-corrected MPC and the statistical models tend to perform equally in terms of prediction accuracy, suggesting that when only the 80-meter free-stream wind speed mean data are available, the main reason to prefer statistical models over the MPC is the ability of the former to produce uncertainty estimates associated with predictions.

The comparison of the first two clusters of boxplots in Fig. 10(a) reveals that adding the 10-minute standard deviation information at 80 meters did not reduce the RMSE in a practically or statistically significant way (average RMSE reduction was 0.002 as fraction of the rated power). This is in contrast to the results in the CA data, where adding the standard deviation led to a practically significant reduction in RMSE fraction of 0.007–0.008, as can be seen from Fig. 11. This is because, as discussed in Section 3.1, the 10-minute standard deviation of the wind speeds is smaller at the OK site than at the CA site, making the benefit of adding it to the model much more limited.

On the other hand, in contrast to the CA site, the REWS (3rd cluster of boxplots in Fig. 10(a)) performs significantly better than the wind speed mean at 80 meters, with the average RMSE difference of 0.007 as fraction of the rated power, across the models. This is due to the fact that, as discussed in Section 3 and shown in Fig. 8(c), there was a substantial amount of shear at the OK site, which is largely captured in the REWS. Indeed, the REWS achieves almost the same degree of accuracy as using the entire profile of wind speed means (4th cluster of boxplots in Fig. 10(a)), with the difference of 0.001 in average RMSE fraction between the two sets of inputs, across all models. Compared to the 80-meter wind speed alone, the entire profile of wind speed means thus reduces the RMSE by a little over 0.008, or 0.8% of the rated power. Although this is a statistically significant difference, this is a much lower reduction than at the CA site, where the difference between these two inputs' RMSEs was 2%, as shown in Fig. 11. The reason for the much smaller relative impact of using the entire wind speed profile on the RMSE of power predictions at the OK site is the same as alluded to earlier: due to the simpler terrain at the site, even using the hub height speed alone yields highly accurate predictions compared to the complex terrain site such as the CA farm, so the benefit of adding other information is relatively small at the OK site.

Adding the standard deviation profile to the mean profile (5th vs. 4th cluster of boxplots in Fig. 10(a)) results in very little benefit (RMSE reduction of less than 0.001 as fraction of rated power, or 0.1%), as in the case of the CA site. This is most likely because the mean profile largely captures the turbulence information contained in the standard deviation profile. In fact, for PCA-reduced models (every other boxplot in the 4th and 5th cluster of boxplots in the figure), this addition sometimes decreases the model accuracy. When using the entire profile of wind speed means only, the PCA reduction does not lead to a deterioration in predictive skill, however, as was the case in the CA data.

Similarly to the CA site, the statistical models tend to perform more or less equally at the OK site (the largest difference between a pair of models for the same input site is below 0.0001). The fact that GPM performs no worse than the other two models is important because, as discussed earlier, it has the advantage of providing uncertainty estimates in a much more natural manner than NN and RF. Given this advantage and comparable performance to that of NN and RF, GPM may be the preferred model in this case.

We now consider the nacelle measurement of the wind speed as an input to the power model. The 1st cluster of boxplots in Fig. 10(b) shows the performance of the bias-corrected MPC and the statistical models using the nacelle wind speed mean as an input. In contrast to our findings at the CA site, in the case of the OK site, the nacelle wind speed performs slightly worse than the 80-m lidar wind speed in terms of the RMSE, with the average difference in RMSE of 0.006 for all models as a fraction of the rated power. This validates our choice of 80-meter lidar wind speed as the baseline (rather than the nacelle wind speed) and confirms our hypothesis in [2] that in flat terrain lidar measurements provide more accurate wind speed data than in complex terrain. While the nacelle wind speed performs worse than the lidar wind speeds, combining it with the lidar wind speeds at other heights (2nd cluster of boxplots in Fig. 10(b)) results in the most accurate prediction of all the wind input sets, as was the case for the CA site.

In [2], we compared the power prediction accuracy of the wind speed inputs described above with that resulting from using wind shear exponent α (defined in Section 3.1) and turbulence intensity (TI), at the CA site. We made the same comparison for the OK site. The results are summarized in Fig. 12, which shows boxplots of RMSEs for the same experiments as described above, as well as either shear or the TI added to some of the wind input sets described in Section 3.1. Only the GPM-c and NN models are shown since the other two models produced very similar RMSE distributions.

While adding shear to the 80-meter wind speed mean (2nd vs. 1st cluster of boxplots) leads to a small improvement in RMSEs for all statistical models, a much more pronounced reduction is achieved by using either the REWS or the entire profile of wind speed means (5th and 6th cluster of boxplots, respectively). This is not surprising since wind shear, which reflects the shape of the wind's vertical profile, only partially captures the information found in the entire vertical profile of wind speed means. Moreover, as the plot shows, even using only the 80-meter wind speed mean and standard deviation (4th cluster of boxplots) leads to slightly better results than using 80-meter wind speed and wind shear. As expected, adding shear to a full profile of wind speed means (7th vs. 6th cluster of boxplots) does not lead to an improvement in RMSE because the information captured in the former is already conveyed by the latter. These results are very similar to those in [2], with the exception of a really good relative performance of the REWS.

As in [2], adding TI at 80 meters to the 80-meter wind speed mean (3rd cluster of boxplots) leads to the same reduction in RMSE as adding the standard deviation at 80 meters (4th cluster of boxplots). The same is true if an entire profile of TIs is added, when compared to adding an entire profile of standard deviations (9th vs. 8th cluster of boxplots in Fig. 12).

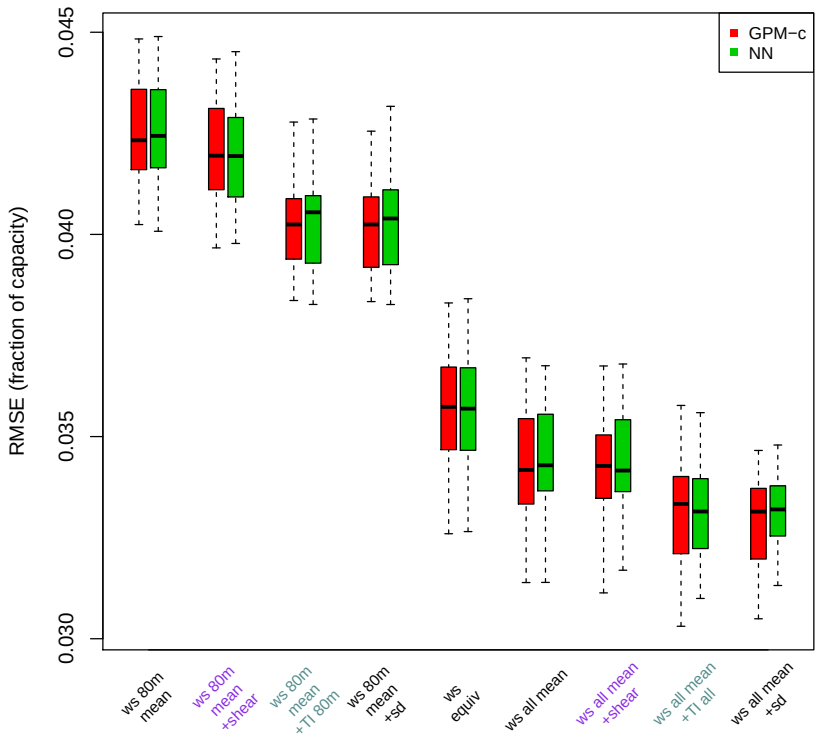

Figure 12: Boxplots of RMSE values (MW) for 30 experiments at the **OK site** using lidar wind speed inputs (sets 1–5 in the enumerated list in Section 3.1, labeled in black on the x-axis) and sets 1 and 4 combined with either shear exponent (labeled in purple on the x-axis) or TI (labeled in teal on the x-axis). When combined with set 1, only the TI at 80 meters was used, whereas TI values at all 10 heights were used when combined with set 4. Only the GPM-c (red) and NN (green) model results are shown for ease of presentation.

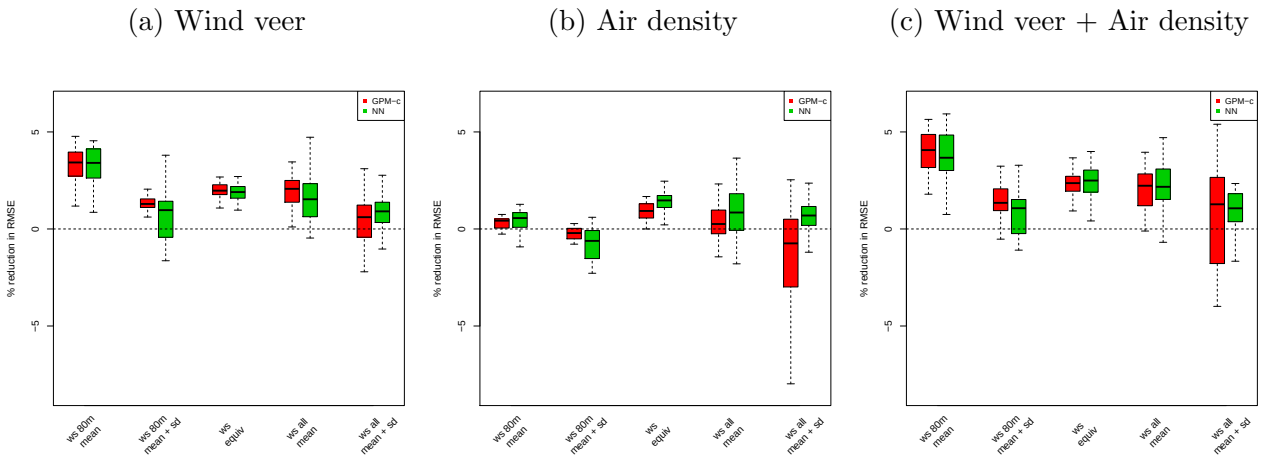

Figure 13: Boxplots of percent reductions in RMSE (as defined in Section 6.2) for the 30 experiments, resulting from using variables in Groups 2–4 of Table 3 as inputs relative to using the density-adjusted wind speed inputs of Group 1 (shown for the GPM-c and NN models only). Percent reductions resulting from adding (a) **wind veer** to the density-adjusted wind speed inputs (Group 2 vs. 1), (b) **air density** separately to the *unadjusted* wind speed inputs (Group 3 vs. 1), (c) **wind veer and air density** to the *unadjusted* wind speed inputs (Group 4 vs. 1). The dashed horizontal line represents the RMSE reduction of 0. Note that the scale on the y-axis is the same for all 3 plots.

6.2 Additional atmospheric variables

Here we examine the results for groups of variables that include inputs other than the wind speed information, specifically, wind veer, air density added separately to the unadjusted wind speeds and the combination of the two (Groups 2–4 outlined in Section 3.2 and listed in Table 3). GPM-c and NN tended to be the best-performing models in the results shown above and in all the other groups of variables, so in this section we limit our discussion only to these two models. Furthermore, the PCA-reduced versions of all the larger wind profiles (sets 4 and 5 in the enumerated list in Section 3.1) tended to perform virtually the same as the original profiles for all the groups, so we only show the results for the former when discussing the larger sets.

Fig. 13(a) shows the percent reduction in RMSE due to adding wind veer to the model already containing density-adjusted wind speed inputs (Group 2 in Table 3). This is the difference in RMSE for the models with and without wind veer as a percentage of the RMSE of the latter model, with a positive reduction implying higher RMSE for the latter model. In contrast to the CA site, where adding wind veer resulted in median RMSE percent reductions of 11–26% (across input sets and models) relative to the original RMSEs, in the case of the OK site, this addition yielded very little improvement for all input sets, with median RMSE percent reductions ranging from 0.6% to 3.4%. This difference in the results between the two sites is not surprising since, as shown in Figs. 8(e) and 9(e) and mentioned in Section 3.2, the amount of wind veer is much smaller at the OK site than at the CA site.

As can be seen from Fig. 13(b), while for some input sets there is minimal reduction in RMSE due to adding density separately to the model rather than using the IEC-recommended density adjustment in Eq. (1), if the results are averaged over all the input sets and models, the two ways of accounting for density perform approximately equally. In addition, the positive reductions are not significantly different from 0. This is in contrast to the CA site, where including density as a separate input yielded significant improvements in RMSE (12–28%) relative to using the IEC adjustment. This difference between the two sites is also expected since, as shown in Fig. 8(f) and mentioned in Section 3.2, the density values at the OK site center around the standard density. The closer the observed density

value is to the standard density, the less important it is to account for density, so the method of doing so becomes more and more inconsequential.

The plot in Fig. 13(c) shows the improvement resulting from having both wind veer and air density in the model when these are added to the unadjusted wind speed inputs (Group 4 in Table 3), relative to using only the air density-adjusted wind speed inputs. The combined effect of adding wind veer and air density is approximately the same as the effect of adding wind veer because, as mentioned above, there is no benefit from adding air density separately relative to using the IEC air density adjustment.

6.3 Caveats

Many of the caveats in our CA site data analysis (discussed in [2]) apply to the OK data, as well. The following is a summary of the shared limitations between the two sites' analyses:

1. The study was limited to a period of 10 weeks.
2. The study was limited to just one turbine rather than an entire wind park.
3. The analysis did not address numerical weather prediction model or lidar measurement error.
4. The analysis was limited to region II of the power curve.
5. The analysis ignored the effects of the turbine and only focused on the atmospheric variables.
6. The use of RMSE as a performance metric has limitations, and other metrics, such as the ability to accurately detect various parameters of ramps in power should be examined.
7. Ten-minute averages may be too coarse to capture certain meteorological phenomena.

It should be noted, however, that one of the main purposes of this study was to begin to address an additional caveat mentioned in [2]. Namely, in that work we noted that our analysis was limited to one location. The study in this report provides information about an additional location and terrain type and is thus a good starting point to obtaining a cross-site comparison and gaining a wider understanding of key determinants of wind power production.

7 Summary and conclusions

This report presented the results of the power curve analysis for the OK site, which has a flat terrain, and compared them to the previous findings for the CA site, which has a moderately complex terrain. Many of the conclusions are the same for the two sites, but there are several important differences, primarily due to the difference in the terrain at the two sites.

The following is a list of similarities in the findings at the two sites:

1. The statistical models performed equally well relative to one another in terms of power prediction accuracy.
2. Using the entire vertical profile of wind speeds improved the accuracy of power predictions relative to using the hub-height wind speed alone (however, the degree of improvement differed at the two sites, as summarized below).
3. Shear exponent was inferior to both the standard deviation of hub-height wind speeds and the entire vertical profile of wind speed means as a predictor of wind power. At the OK site, it was also inferior to the rotor-equivalent wind speed (REWS).

4. Using the hub-height turbulence intensity in conjunction with the hub-height wind speed mean led to the same prediction accuracy as using the hub-height standard deviation together with the hub-height wind speed mean.

The differences between the results at the two sites were as follows:

1. Prediction accuracy was much higher at the OK site than at the CA site due to the much simpler terrain of the former. As a result, the worst-performing input predictor set (hub-height wind speed mean alone) at the OK site resulted in greater accuracy (as fraction of the rated power) than the best-performing predictor set (the entire vertical profile of wind speed means) at the CA site.

2. The degree of improvement afforded by using the entire vertical profile of wind speed means relative to using the hub-heigh wind speed alone was much greater at the CA site than at the OK site. This is because even the simplest predictor sets at the OK site resulted in very high accuracy relative to the CA site (for the reason stated above), so the benefit from adding more predictors was marginal at the former.

3. At the OK site, using the REWS resulted in significant improvement in power prediction accuracy relative to the hub-height wind speed. This was not the case at the CA site, where the two yielded very similar prediction accuracy.

4. Adding wind veer to the predictor set resulted in a significant improvement in power prediction accuracy at the CA site, but not at the OK site. This is because the amount of wind veer was very small at the latter.

5. Adding air density separately to the predictor set instead of adjusting the wind speed for air density using the IEC-recommended adjustment resulted in a significant improvement in power prediction accuracy at the CA site, but not at the OK site. This is because air density values tended to be close to standard at the OK site, making the air density adjustment less relevant there than at the CA site.

Table 4 summarizes the improvements in RMSE resulting from using various input sets relative to the hub-height wind speed mean for each of the two sites. All of the results presented here must be further validated at more locations with different terrain and turbine types, for longer time periods and other seasons.

Table 4: Median RMSE reductions (**as fractions of rated power or percent change**, as indicated) relative to using the hub-height wind speed (HH WS) mean as an input to the power curve model at each of the two sites (sd = standard deviation). Note that a negative value means that the RMSE increased.

Input(s)	OK site	CA site
HH WS mean + sd	0.002	0.007
REWS	0.007	0.003
Profile WS mean	0.008	0.02
Nacelle WS mean	-0.006	0.06–0.09
Wind veer	3.4%	27%
Air density added *separately* (wrt IEC adjustment)	Insignificant difference	28%

Acknowledgments

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under the Contract DE-AC52-07NA27344 and funded by the LLNL Energy Program. The authors are extremely grateful for Infigen's and Enel's partnership, including data sharing, intellectual input and field campaign assistance. The authors would also like to thank Jennifer Newman (University of Oklahoma) for assistance in lidar deployments and data collection and Bryan Maddock (Infigen) and James Clark (Enel) for site access and field deployment assistance.

A Additional details on the statistical models

A.1 Neural network

A neural network (NN, also known as an artificial neural network, or ANN) is a non-linear regression model whose mathematical structure mimics the behavior of biological neural networks. In particular, if using a single-layer network, the response Y, such as power output in our case, is modeled as a linear combination of basis functions ϕ_j, $j = 1, ..., J$, each of which is a non-linear function of predictors $\boldsymbol{x}$ (such as those discussed in Section 3):

$$Y(\boldsymbol{x}) = \sum_{j=1}^{J} w_j \phi_j(\boldsymbol{x}). \tag{4}$$

We used the default sigmoid function for the basis functions ϕ_j, whereas the weights w_j and the number of basis functions J were both determined using cross-validation. The library *nnet* [16] in the R package [17] was used to fit the NN model. See [10, 11] for a more thorough treatment of NNs.

A.2 Random forest

Like NN, a random forest (RF) is a regression model. This method constructs an ensemble of regression trees, and the individual trees' predictions are averaged to obtain the final prediction. Different trees are constructed by first randomly sampling the training data and then randomly sampling the predictors at each node of the tree. The library *randomForest* [18] in the R package [17] was used to fit the RF model. Defaults of various parameters required by the model were used, except for the number of predictors that are randomly sampled at each split, which was tuned to minimize the out-of-bag error. For a more thorough treatment of RFs, see [11, 12].

A.3 Gaussian process model

A general formulation of the stationary Gaussian process model (GPM) is as follows:

$$Y(\boldsymbol{x}) = \boldsymbol{g^T}(\boldsymbol{x})\boldsymbol{\beta} + \boldsymbol{Z(x)}, \tag{5}$$

where Y is the output, $\boldsymbol{x}$ is the vector of predictors, $\boldsymbol{g}$ is a user-defined function, $\boldsymbol{\beta}$ is the vector of the coefficients associated with each predictor, and $\boldsymbol{Z(x)}$ is a stationary error process, modeled as a multivariate Gaussian random variable with 0 mean vector and a user-specified covariance, given by

$$Cov(\boldsymbol{x}_i, \boldsymbol{x}_j) = K(||\boldsymbol{x}_i - \boldsymbol{x}_j||). \tag{6}$$

The covariance between any two output values is thus modeled as a user-specified function K of the Euclidean distance between their corresponding predictor vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, $||\boldsymbol{x}_i - \boldsymbol{x}_j||$.

The above model specification for the output value $Y(\boldsymbol{x})$ implies that the joint distribution of the predicted value $Y(\boldsymbol{x_{new}})$ at a new set of input values $\boldsymbol{x_{new}}$ and the observed training set values $y(\boldsymbol{x_1}), ..., y(\boldsymbol{x_n})$ (in keeping with the standard practice in GPM literature, we denote the predictions and the observed values with the large-case and small-case letters, respectively) is multivariate Gaussian, while the predictive distribution of $Y(\boldsymbol{x_{new}})$ (i.e., its distribution conditional on the observed response data) is univariate Gaussian. This result is the foundation of GPM prediction inference. See [13, 14, 15] for a detailed treatment of GPM.

The Gaussian random variable is defined on the entire real number line, whereas power output is bounded below by 0 MW and above by the rated power (1.68 MW in the case of GE 1.68-82.5-60 Hz

wind turbine considered in this work). Consequently, the power values are first transformed using a logit transformation, as follows:

$$Y(\boldsymbol{x}) = log\left(\frac{P^*(\boldsymbol{x})}{1 - P^*(\boldsymbol{x})}\right) \tag{7}$$

where $P(\boldsymbol{x})$ is the power value normalized to the rated power and

$$P^*(\boldsymbol{x}) = P(\boldsymbol{x}) \cdot (1 - 2c) + c, \tag{8}$$

with c a positive constant close to 0. Including c ensures that the logit in Eq. (7) is defined for normalized power values of 0 and 1 since Eq. (8) results in $P^*(\boldsymbol{x})$ of c or $1-c$ when power $P(\boldsymbol{x})$ is 0 or 1, respectively. The value of c is chosen so that $Y(\boldsymbol{x})$ in Eq. (7) is as close as possible to the Gaussian distribution.

The GPM is then used to obtain the parameters of the predictive Gaussian distribution of $Y(\boldsymbol{x_{new}})$, which in turn leads to the logit-normal [19] as the predictive distribution of $P^*(\boldsymbol{x_{new}})$, i.e., the conditional distribution of $P^*(\boldsymbol{x_{new}})$ given the observed power values. From Eq. (8) it follows that

$$P(\boldsymbol{x}) = \frac{P^*(\boldsymbol{x}) - c}{1 - 2c}, \tag{9}$$

so the predictive distribution of $P(\boldsymbol{x_{new}})$ is a logit-normal shifted by c and scaled by $1-2c$. The point prediction of power is then the mean of the predictive distribution of $P(\boldsymbol{x_{new}})$ since this mean is the minimum mean squared prediction error predictor of $P(\boldsymbol{x_{new}})$. It is obtained by numerical integration using the *logitnorm* package [20] in R and applying the transformation on the right-hand side of Eq. (9). The associated prediction interval is obtained by computing the appropriate quantiles of the predictive Gaussian distribution for $Y(\boldsymbol{x_{new}})$ and applying the logistic transformation (the inverse of the logit) to them, followed by the transformation in Eq. (9).

The following covariance function was used for the GPM:

$$\begin{aligned} &Cov(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sigma^2 \left[\exp\left(-\sum_{k=1}^{K} \frac{|x_{ki} - x_{kj}|^{\nu}}{d_k}\right) + q\delta_{ij}\right], \\ &\text{where} \\ &d_k = \text{range parameter for the } k^{\text{th}} \text{ predictor}, \\ &q = \text{nugget and} \\ &\delta_{ij} = \text{Kroeneker delta} \end{aligned} \tag{10}$$

(note that this specification is adapted from [21]). The range (d_k, for each $k = 1, ..., K$, with K denoting the total number of predictors), the shape (ν), the nugget (q), and the variance (σ^2) parameters were estimated using maximum likelihood estimation. This and other operations needed to implement the GPM were done using routine functions available in the R package [17].

Finally, we considered two forms of the mean function $\boldsymbol{g^T}(\boldsymbol{x})\boldsymbol{\beta}$ in Eq. (5), resulting from setting $\boldsymbol{g}$ to a constant or using a natural spline of order 3 (see [22, 23]) to model it. In case of the former, all the predictors were used to model only the error process. In the latter case, we first separated the predictor set $\boldsymbol{x}$ into two subsets: one used as inputs to the function $\boldsymbol{g}$, denoted with $\boldsymbol{x_m}$, and the other used to model the error process $\boldsymbol{Z}$, denoted with $\boldsymbol{x_e}$, as indicated in Table 5 (note that some predictors were in both sets). We then modeled power output Y as a function of the input vector $\boldsymbol{x} = (\boldsymbol{x_m}, \boldsymbol{x_e})$, as follows:

$$Y(\boldsymbol{x}) = Y(\boldsymbol{x_m}, \boldsymbol{x_e}) = \boldsymbol{g^T}(\boldsymbol{x_m})\boldsymbol{\beta} + \boldsymbol{Z}(\boldsymbol{x_e}),$$

with the covariance function as in Eq. (10).

Table 5: Subsets of inputs used to model the mean function and the error process (denoted $\boldsymbol{x_m}$ and $\boldsymbol{x_e}$, respectively) for each of the five input sets enumerated in Section 3 when a natural spline of order 3 was used for the function $\boldsymbol{g}$. The input labels are as in Section 3.1.

Input set	$\boldsymbol{x_m}$	$\boldsymbol{x_e}$
1	ws 80m mean	ws 80m mean
2	ws 80m mean	ws 80m sd
3	ws equiv	ws equiv
4	ws 80m mean	ws all mean
5	ws 80m mean	ws all mean + sd

References

[1] International Electrotechnical Commission (IEC) International Standard 61400-12-1: Wind turbine generator systems - Wind turbine power performance testing, 1998.

[2] V. Bulaevskaya, S. Wharton, A. Clifton, G. Qualley, and W. Miller. Wind power curve modeling in complex terrain using statistical models. *Journal of Renewable and Sustainable Energy*, 7 103103, 2015.

[3] Mitsubishi wind turbine generator specifications. `https://www.mhi-global.com/products/pdf/mwt1000a.pdf` (last accessed on 11/8/14).

[4] Enel Group. Personal communication, 2014.

[5] R. Krishnamurthy and M. Boquet. Case studies of WINDCUBE measurement uncertainty for complex terrain using FLOW complexity recognition (FCR). Poster presentation at the EWEA conference in Barcelona, Spain, 2014.

[6] J. Song, K. Liao, R. L. Coulter, and B. M. Lesht. Climatology of the low-level jet at the Southern Great Plains Atmospheric Boundary Layer Experiments site. *Journal of Applied Meteorology*, 44:1593–1606, 2005.

[7] R. Wagner, I. Antoniou, S. M. Pedersen, M. S. Courtney, and H. E. Jorgensen. The influence of the wind speed profile on wind turbine performance measurements. *Wind Energy*, 12:348–362, 2009.

[8] I. T. Jolliffe. *Principal Component Analysis*. Springer, 2002.

[9] J. R. Jammalamadaka and A. SenGupta. *Topics in circular statistics*. World Scientific Press, 2001.

[10] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2007.

[11] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer, 2009.

[12] L. Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

[13] J. Sacks, W. J. Welch, T. J Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statistical Science*, 4:409–423, 1989.

[14] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2005.

[15] T. J. Santner, B. J. Williams, and W. I. Notz. *The Design and Analysis of Computer Experiments*. Springer, 2003.

[16] W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer, New York, fourth edition, 2002. ISBN 0-387-95457-0.

[17] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. `http://www.R-project.org` (last accessed on 11/8/14).

[18] A. Liaw and M. Wiener. Classification and regression by randomforest. *R News*, 2(3):18–22, 2002. `http://CRAN.R-project.org/doc/Rnews/` (last accessed on 11/8/14).

[19] R. Mead. A generalised logit-normal distribution. *Biometrics*, 21:721–732, 1965.

[20] T. Wutzler. Package 'logitnorm': Functions for the logitnormal distribution. `http://cran.r-project.org/web/packages/logitnorm/logitnorm.pdf` (last accessed on 11/8/14).

[21] R. B. Gramacy and H. K. H. Lee. Adaptive design and analysis of supercomputer experiments. *Technometrics*, 51:130–145, 2009.

[22] T. Hastie. *Generalized Additive Models*. Chapman and Hall/CRC, 1990.

[23] J. M. Chambers and T. J. Hastie. *Statistical Models in S*. Wadsworth and Brooks/Cole, 1992.