LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-677602

Wind Power Curve Modeling Using Statistical Models: An Investigation of Atmospheric Input Variables at a Flat and Complex Terrain Wind Farm

S. Wharton, V. Bulaevskaya, Z. Irons, G. Qualley, J. F. Newman, W. O. Miller

September 28, 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 Using Statistical Models: An Investigation of Atmospheric Input Variables at a Flat and Complex Terrain Wind Farm

Sonia Wharton[1], Vera Bulaevskaya[1], Zack Irons[2], Grant Qualley[3,^], Jennifer Newman[4,*] and Wayne Miller[1]

[1]Lawrence Livermore National Laboratory, Livermore, California

[2]Enel Green Power North America, Andover, Massachusetts

[3]Infigen Energy, Dallas, Texas

[4]University of Oklahoma, Norman, Oklahoma

[^]Current affiliation: Pentalum Technologies, Dallas, Texas

[*]Current affiliation: National Renewable Energy Laboratory, Golden, Colorado

LLNL-TR-677602

Prepared for the Wind and Water Power Technologies Office, Office of Energy Efficiency and Renewable Energy, Department of Energy by the Lawrence Livermore National Laboratory

September 30, 2015

1. Introduction

As the United States sets increasingly ambitious wind penetration goals, an ability to produce an accurate wind power forecast is becoming more and more critical. To be valuable, a forecasting tool should not only produce accurate point forecasts of power but also correctly quantify the uncertainty associated with these point predictions. Statistical analysis and modeling has the potential to be a powerful tool for designing a forecasting system with both of these features.

The goal of our FY15 project was to explore the use of statistical models and high-resolution atmospheric input data to develop more accurate prediction models for turbine power generation. We modeled power for two operational wind farms in two regions of the country. The first site is a 235 MW wind farm in Northern Oklahoma with 140 GE 1.68 turbines. Our second site is a 38 MW wind farm in the Altamont Pass Region of Northern California with 38 Mitsubishi 1 MW turbines. The farms are very different in topography, climatology, and turbine technology; however, both occupy high wind resource areas in the U.S. and are representative of typical wind farms found in their respective areas.

This work was motivated by a number of earlier studies which found discrepancies between nacelle hub-height wind speed and expected power generation (Motta et al. 2005, Sumner and Masson 2006, van den Berg 2008, Antoniou et al. 2009, Wagner et al. 2011, Vanderwende and Lundquist 2012). This study also follows a legacy of DOE funded research at LLNL. In FY09-11 the DOE Wind Program funded Drs. S. Wharton and J. Lundquist to explore the relationship between power generation and atmospheric stability for a wind farm in the West Coast, U.S.A. In that study we found that power generation was significantly related to atmospheric stability through the parameters turbulence intensity and wind shear; these two factors explained significant discrepancies in power output for a given nacelle wind speed (Wharton and Lundquist 2012).

Here we expand on prior work by examining a larger set of atmospheric variables and the use of three statistical models. The suite of atmospheric variables was taken with vertical profiling laser detection and ranging (lidar) instrumentation. Lidars were deployed to provide high resolution information about inflow conditions. Atmospheric variables were tested for their ability to predict power generation through the use of statistical models. With the collected data sets and model results we were able to answer the following questions:

1. Do high resolution measurements of wind shear, veer, and turbulence intensity from lidar add a significant improvement to power prediction versus relying on the nacelle hub-height measurements of wind speed and turbulence intensity? If so, which of the lidar measurements are most important for predicting power? Are the important variables the same at both wind farms or do they differ?

2. Can a single lidar add a predictive advantage for the entire wind farm or are the benefits limited to the closest turbine?

3. Which of the statistical models produce the most accurate prediction? Do statistical models provide a predictive advantage over using the manufacturer's power curve (MPC)?

2. Overview of statistical models

The Power Curve Working Group is in the process of evaluating the definition of wind speed to be used in manufacturer power curve (MPC) based models (PCWG 2015). Variables examined thus far include hub-height wind speed, rotor diameter equivalent wind speed, and turbulence and air density renormalized wind speed. While these variables expand the power curve modeling technique beyond the most frequent input, hub-height wind speed, it is by far not a comprehensive list of available input variables. Basic power curve models (e.g., MPC) cannot accommodate input sets that include multiple variables motivating the use of statistical models to predict power for the more complex input sets (e.g., lidar measurements). Moreover, even in the absence of the additional variables, a statistical model has two advantages. One advantage is that it does not assume a particular relationship between the inputs and output, as these relationships are learned by the model from the data. A second advantage is that statistical models provide an estimate of the uncertainty associated with the predictions.

We compare three statistical models in terms of their ability to predict power generation. These include neural networks (NN), random forests (RF), and Gaussian process models (GPM). For more details on each of these models see Bishop (2007), Hastie et al. (2009), Breiman (2001), Sacks et al. (1989), Rasmussen and Williams (2005), Santner et al. (2003), and Gramacy and Lee (2008). The statistical software R was used to train all three models. Each of the models requires the user to specific certain parameters. As discussed in Bulaevskaya et al. (2015) for some of these parameters we used well-established default values or functional forms, while others were determined using cross-validation or maximum likelihood estimate. Two different forms of the mean function for the GPM were used: a constant and a natural spline. These will be referred to as GPM-c and GPM-ns, respectively.

While NN and RF can be very powerful prediction tools and have been used previously in the wind energy field (e.g., Clifton et al. 2013, Fonte and Quadrado 2005), the GPM provides a much more natural framework for modeling prediction uncertainty. The capability to produce accurate uncertainty estimates associated with point predictions is a highly desirable feature in a wind power forecasting tool. GPM has not to our knowledge been used before in wind power modeling. Specific details about how we ran the models are found in Bulaevskaya et al. 2015.

3. Wind farm descriptions and field campaign details

3.1 Oklahoma – simple terrain wind farm

The first wind farm studied is located in north central Oklahoma, approximately 120 km from Oklahoma City. The exact location and name is proprietary information; for our purposes here

the farm is referred to as the Oklahoma Wind Farm. The farm consists of 140 horizontal-axis turbines that are spread across 25 km from west-to-east in rural Oklahoma. The terrain is flat and the area is used mostly for farming. The wind turbines are blade-pitch controlled, GE 1.68 MW (1.68-82.5-60 Hz). Cut-in speed is 3 m/s, rated speed is 13 m/s, and cut-out speed is 25 m/s. Hub-height is 80 m, rotor diameter is 82.5 m and blade length is 40.3 m (Table 1). This gives a minimum blade height of 40 m above ground level (a.g.l.) and maximum blade height of 120 m a.g.l.

Table 1: Specifications of the GE 1.68-82.5-60 Hz turbines at the Oklahoma wind farm

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

From November 2013 through January 2014 we deployed a continuous wave Doppler ZephIR 300 lidar (Zephir Ltd., Ledbury, UK) at the south end of the wind farm (Figure 1). Six nearby turbines were unwaked and downwind of the lidar and 80-m tall meteorological tower when winds were southerly (135-225°). Power generation on a typical winter day for these turbines is shown in Figure 2.

Figure 1: (Left) Satellite image showing the south meteorological tower and turbines, B2-B6 and C1. (Right) A ZephIR 300 lidar was co-located with an 80-m tall meteorological tower just south of the turbine row.

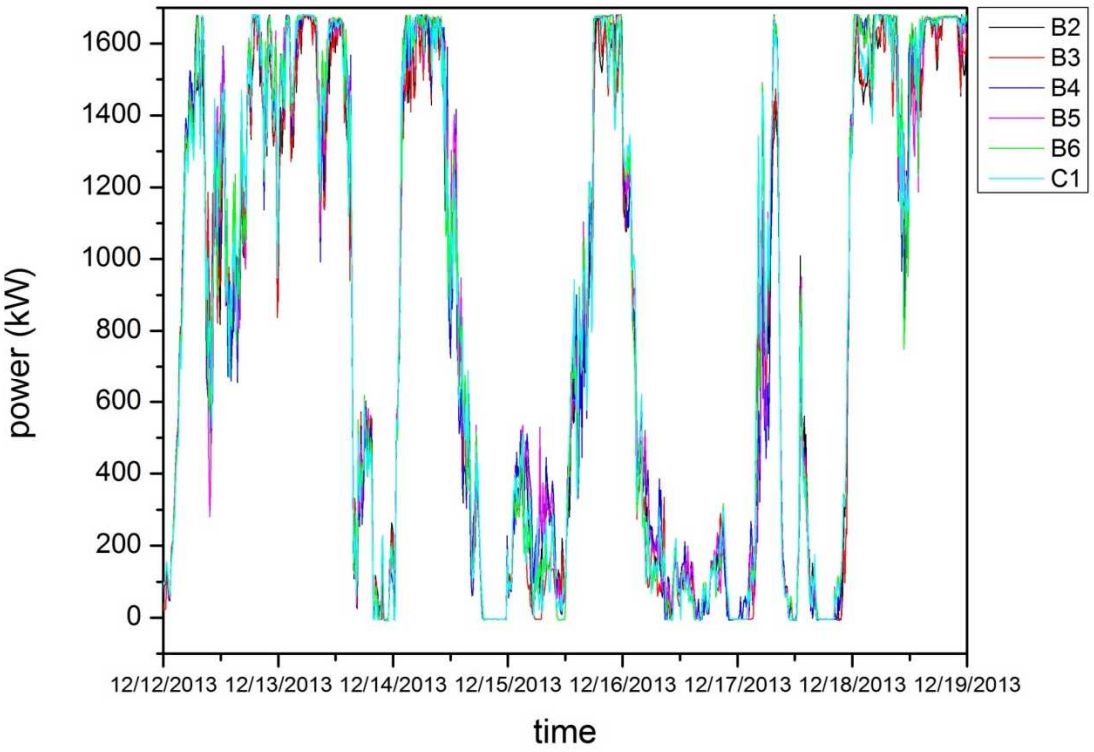

Figure 2: Power generation for a row of unwaked turbines at the south end of the Oklahoma Wind Farm for a typical week in the winter. The production of wind power appears to be dominated by synoptically-driven frontal events during this time.

During the winter the winds have a bimodal distribution with the majority of the winds coming from either the north or south. Southerly winds occurred approximately 45% of the time and put the B turbines directly downwind of the lidar. Mean wind speed during the field campaign was 8 m/s and a large portion of the wind speeds fell between cut-in and rated speed (Figure 3).

Elevation differences between the meteorological tower, turbines and lidar were negligible. Turbine B6 was selected for power modeling as this turbine was the IEC recommended distance (2-4 rotor diameters) away from the lidar for the device to measure free-stream inflow (IEC 2005). A schematic of the lidar deployment is shown in Figure 4. The lidar was programmed to collect 10-minute mean values of horizontal wind speed (m/s), wind direction (°), vertical wind speed (m/s), and turbulence intensity (%) from 40-130 m at 10 m intervals. These heights covered the entire rotor-disk at high spatial resolution. The raw high frequency u, v, and w velocity measurements were also collected and archived. Reported accuracy for the ZephIR 300 in ideal conditions is < 0.5% for wind speed and < 0.5° for wind direction (Slinger and Harris 2012).

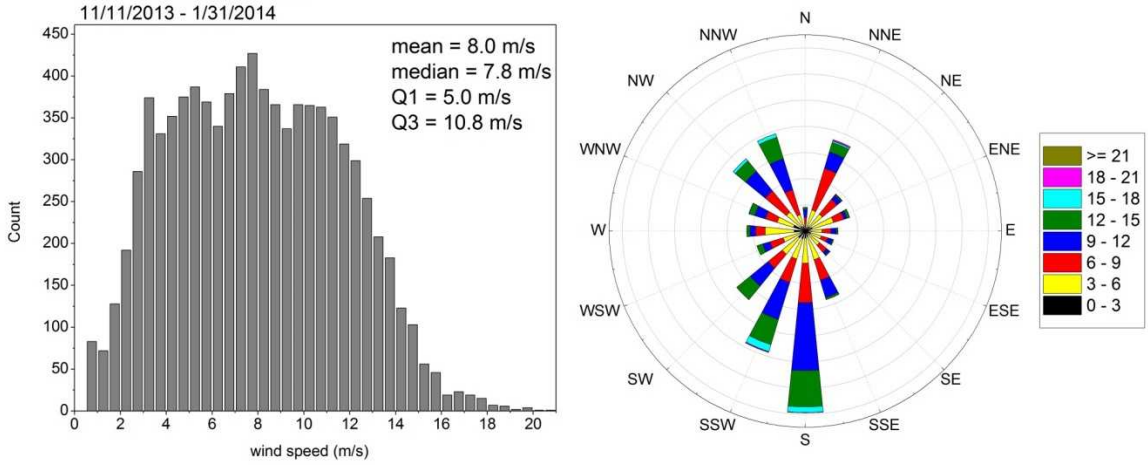

Figure 3: Wind speed histogram and wind rose based on the lidar 80-m observations from November 2013-January 2014. Westerly and easterly winds were rarely seen. Winds were predominately between cut-in speed and rated speed and from the north or south. The lidar measured free-stream, wake-free inflow measurements for turbine B6 during southerly winds.

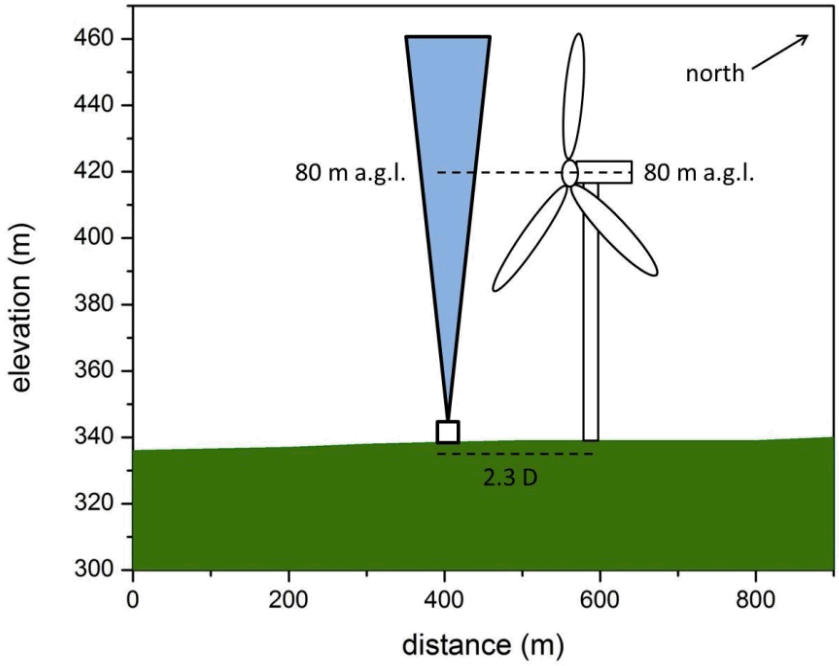

Figure 4: Schematic of the terrain profile and locations of the lidar and turbine B6 relative to one another. The lidar was located 2.3 rotor diameters (D) upwind of B6 during southerly winds. Free-steam inflow measurements were relatively easy to obtain at this site as terrain influences were negligible between the lidar and turbine.

The co-located lidar and 80-m tall meteorological tower measurements were first compared for instrument agreement in wind speed and direction at hub-height (80 m). We found excellent agreement for both wind speed (r-square = 0.99) and direction (r-square = 0.99). This gave us confidence in the lidar's accuracy. 10-minute mean air density was calculated using air temperature and barometric pressure measurements taken at the top of the meteorological tower. The tower data were available from Enel's Supervisory Control and Data Acquisition (SCADA) system. The SCADA system also provided turbine statistics for all of the turbines in the wind farm. These included 10-minute values of mean nacelle position (°), mean pitch angle (°), mean rotor speed (rpm), mean power output (kW), mean and standard deviation of hub-height wind speed (m/s), and flags for operating/fault state. Although SCADA data were available for the entire park, FY15 project scope focused on modeling power for one unwaked turbine at the Oklahoma Wind Farm.

While the local and regional terrain is relatively flat, the synoptic meteorology can be complex and is influenced by mountain systems far to the west and cold or warm frontal passages originating from the north or south. On atmospherically stable nights, a decoupled flow event, called a nocturnal low level jet (LLJ) can form over the region in the spring and summer (Klein et al. 2015) and the area is located within a climatological maximum of U.S. LLJ occurrences (Song et al. 2005). LLJs produce maximum wind speeds frequently within the lowest 500 m of the atmosphere and can reach heights down to the upper half of the rotor disk. Mechanically-generated turbulence is often observed below the jet maximum and may also penetrate the rotor-disk. Wind shear and turbulence characteristics found at turbine heights are also strongly influenced by local surface heating and cooling during the warmer months but less so during winter and autumn (Wharton et al. 2015a).

3.2 California – complex terrain wind farm

The second wind farm is located in the Altamont Pass Wind Region in Northern California, approximately 70 km east of San Francisco. The exact location and name is proprietary information; for our purposes here the farm is referred to as the California Wind Farm. The terrain is moderately complex both at regional and local scales. Nearby 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 and is used as grazing land.

The farm consists of 38 horizontal-axis turbines located along the top of parallel ridgelines. The wind turbines are blade-pitch controlled, 1 MW Mitsubishi (MWT62-1000A). Cut-in speed is 3.0 m/s, rated speed is 12.5 m/s, and cut-out speed is 25.0 m/s. Hub-height is 55 m, rotor diameter is 61.4 m and blade length is 29.5 m (Table 2). This gives a minimum blade height of 24 m above ground level (a.g.l.) and maximum blade height of 86 m a.g.l.

Table 2: Specifications of the MWT62-1000A turbines at the California wind farm

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

In June through August of 2012, we deployed a pulsed Doppler light detection and ranging (lidar) Wind Cube v2 (Leosphere, Orsay, France) at the western edge of the wind farm to obtain measurements of the free-stream, unwaked inflow as it approached a leading row of turbines (row A) (Figure 5). Power generation on a typical summer week for these turbines is shown in Figure 6.

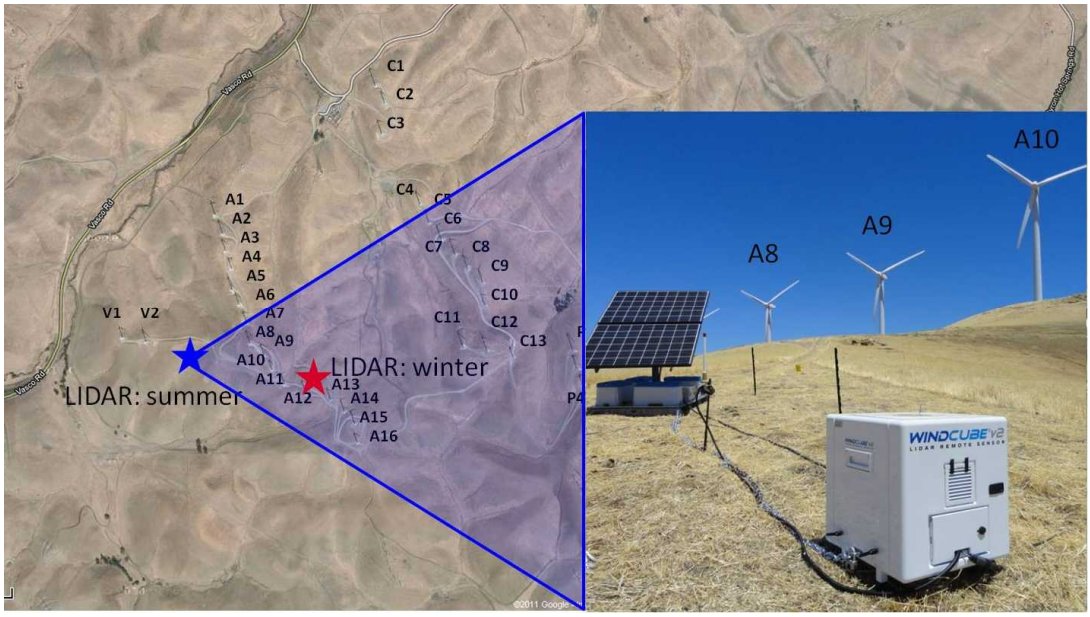

Figure 5: (left) Satellite image of the 38 MW California wind farm showing the turbines, surrounding terrain, and locations of the lidar campaigns. Here we concentrate on the summer campaign which put the lidar frequently upwind of turbines in row A (right).

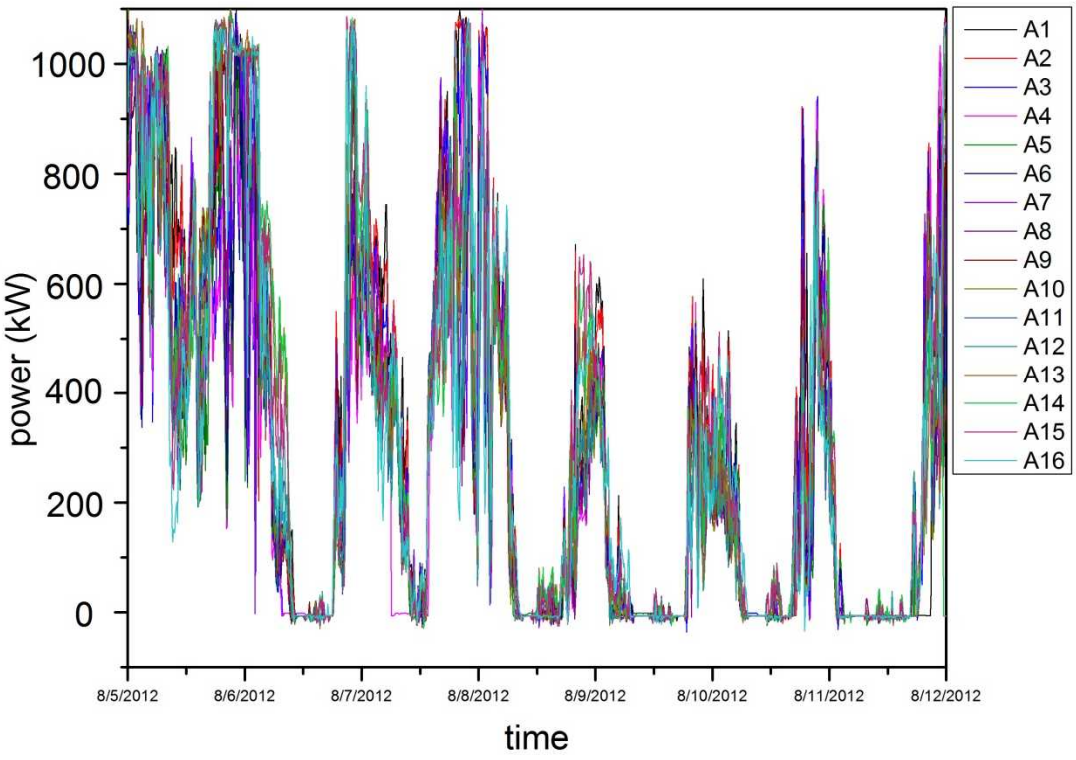

Figure 6: Power generation for a row of unwaked turbines at the western edge of the California Wind Farm for a typical week in the summer power season. The production of wind power appears to be dominated by diurnally-driven ramp events during this time.

The turbines most directly downwind of the lidar were A7, A8, A9, A10 and A11 based on wind direction (Figure 7). Prevailing strong winds from 225-250° put A8 most frequently downwind. Mean wind speed during the field campaign was 7.5 m/s and a large portion of the wind speeds fell between cut-in and rated speed.

Due to the moderately complex terrain the lidar was deployed at an elevation lower than the turbines. This was done to ensure that the instrument was placed 2 to 4 rotor diameters upwind of the turbines as recommended by the IEC Standard. A schematic of the lidar deployment is shown in Figure 8. Elevation differences as well as terrain-induced changes in flow physics made hub-height wind speed a relative term: the nacelle hub-height (55 m) wind speed was not equal to the flow velocity found at the same height above the lidar (see Figure 8). A methodology for determining a proxy for hub-height wind speed at the lidar is discussed in Section 4.1.2.

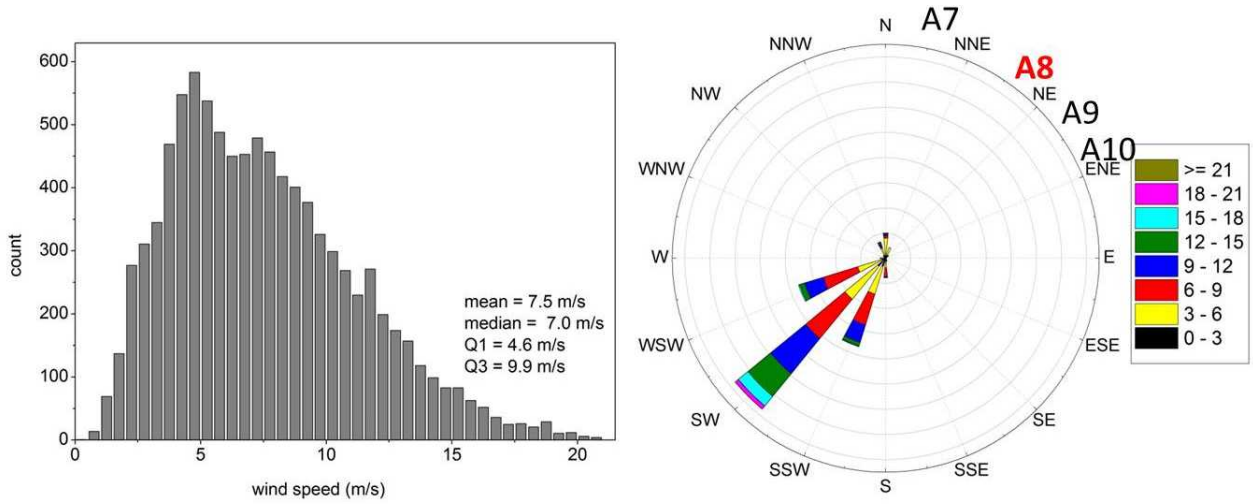

Figure 7: Wind speed histogram and wind rose based on the lidar 55-m observations from June-August 2012. The plot shows a high frequency of winds coming from the southwest (225-250°) which puts A8 (highlighted in red) most often downwind of the lidar.

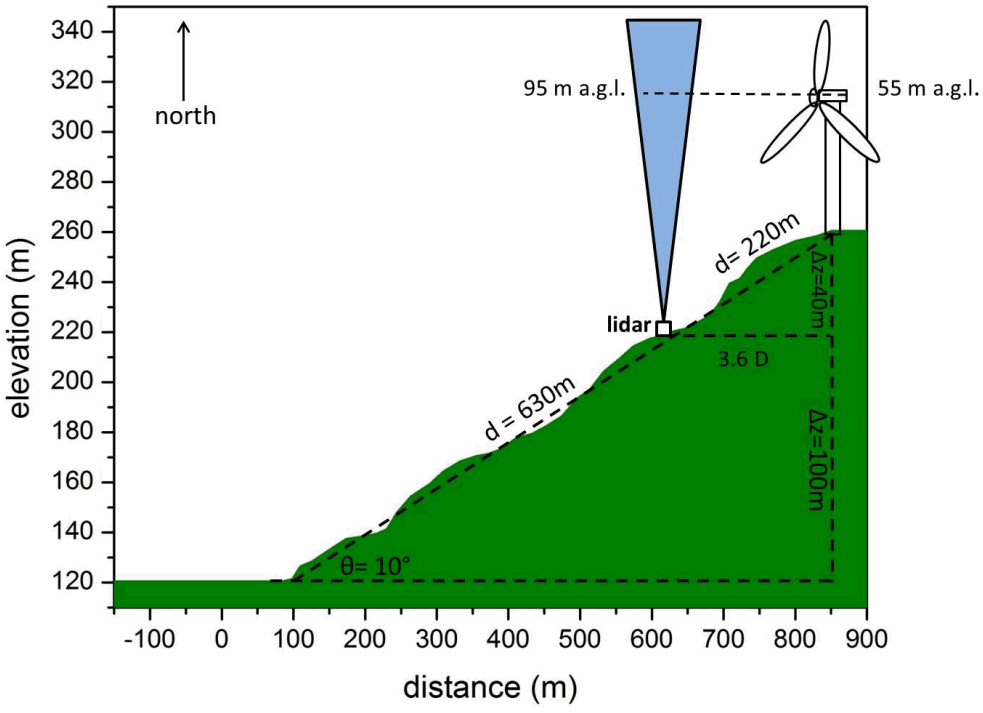

Figure 8: Schematic of the hill profile and locations of the lidar and turbine A8 relative to one another. The lidar was located 3.6 rotor diameters (D) upwind of A8. This ridge is approximately 140 m tall and has an average slope of 10%.

The lidar provided 10-minute mean horizontal wind speed (m/s), wind direction (°), vertical wind speed (m/s) and turbulence intensity (%) measurements across the entire rotor disk and at heights above. The lidar was programmed to measure from 40 m-150 m at 12 levels spaced 10 m apart. The raw 1 Hz u, v, and w velocity measurements were also collected and archived. Note that complex terrain reduces the accuracy of lidar and here uncertainty was estimated to be 3-5% for wind speed based on evaluations of test sites with similar topography (Krishnamurthy and Boquet 2014).

Infigen Energy's 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. The FY15 project scope included modeling power for all turbines at the California Wind Farm. Power modeling was done at this farm in the following order of investigation: a single unwaked turbine, a single row of unwaked turbines, and lastly, power modeling for all of the turbines at the park.

Wind speed was measured with a nacelle-mounted cup anemometer (NRG #40, NRG Systems, Inc., Vermont, USA). The site also had a single 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 because the tower was not directly upwind of turbine A8 and there were concerns about the quality of these measurements.

The summer period was chosen for study because it is the area's peak wind season. During the warm 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. See Wharton et al. (2015b) for a detailed description on how regional and local topography affects wind flow at the California Wind Farm.

4. Performance study

4.1 Data criteria

4.1.1 Oklahoma

At the Oklahoma wind farm, the lidar data were quality controlled using internal ZephIR 300 processing algorithms. The SCADA power data were quality controlled to remove any points when power generation was less than 1 kW and wind speed was above cut-in (3 m/s). Most of these times indicated periods when there was turbine maintenance or other reasons for a turbine

shutdown. The data were then restricted to a wind sector of 135-225° and time periods without any missing B6 SCADA or lidar data. Next we filtered the dataset for times when the 80 m lidar wind speed corresponded to region II of the power curve. In this region power is most sensitive to small changes in wind speed. This and all other filters resulted in 1404 data points.

While we found an excellent agreement between the lidar and met tower measurements as previously discussed, we saw a strong bias between the nacelle-based wind speed measurements and the lidar (Figure 9). Overall, the nacelle cup anemometer measured wind velocities on average 0.5-1 m/s slower than the upwind instruments. However, this was not an unexpected finding. The nacelle cup anemometer is by standard practice mounted behind the turbine blades and the instrument suffers from a reduction in wind speed when the turbine is operating and from flow distortion from the shape of the nacelle hub and rotating blades.

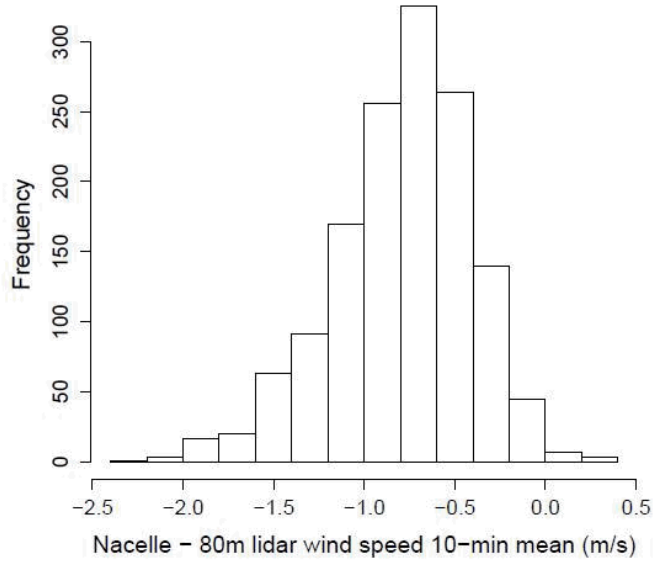

Figure 9: Distribution of the differences (m/s) between hub-height nacelle and 80 m lidar wind speed means, with a negative difference implying that the lidar mean is the larger of the two (RMSE = 0.87 m/s, correlation = 0.98).

4.1.2. California

At the California wind farm, the lidar data were quality controlled for signal-to-noise ratio and availability: data points with carrier-to-noise ratio below -23 and availability less than 25% during each 10-minute averaging period were discarded. The SCADA power data were quality controlled to remove any points when power generation was less than 1 kW and wind speed was above cut-in (3 m/s). Of all of the lidar measurement heights, the 10-minute wind speed averages at 90 m had the highest correlation and smallest root-mean-squared error (RMSE) relative to the corresponding nacelle wind speed averages (correlation = 0.94, RMSE = 0.84 m/s) (refer to Figure 8 for the elevation differences between the lidar and turbine). Figure 10 shows the distribution of the differences between the nacelle and 90 m lidar wind speeds. The nacelle wind

speed exceeds the lidar measurement by 0.4 m/s on average, indicating the frequent occurrence of accelerated hill flows at the top of the ridge. These flows, known as "hill speed-ups" are discussed in detail in Wharton et al. (2015b).

All turbine and lidar data were restricted to a wind sector of 225-250°. Wind direction was based on the lidar hub-height measurement and bias-corrected wind vane measurements from each turbine in the SCADA system. Only time periods with complete lidar and SCADA data were kept. Finally we filtered the dataset for times when the 90 m lidar wind speed corresponded to region II of the power curve. This and all other filters resulted in 1737 data points.

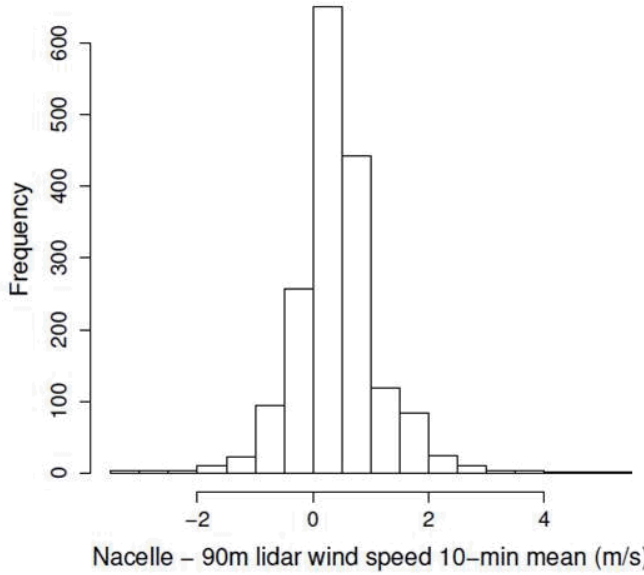

Figure 10: The distribution of the differences between the hub-height nacelle and 90 m lidar wind speed means (m/s), with a positive difference implying that the nacelle mean is the larger of the two. Here we observe a high frequency of faster wind speeds measured at the turbine than in the lidar inflow because of the site's prevalence of "hill speed-up" flows.

4.2. Atmospheric input variables

Traditional power curves model the expected 10-minute mean power for a given 10-minute mean hub-height wind speed. These curves typically assume a standard air density of 1.225 kg/m^3 and low-to-moderate turbulence intensity. However, in areas of high elevation or those with large air temperature ranges, air density can differ substantially from standard values at any given time. To account for the potential discrepancy wind speed observations were adjusted using methodology recommended by the IEC in Eq (1):

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

Where, $U_{d.adj}$ is density-adjusted lidar or nacelle wind speed (m/s), U is measured lidar or nacelle wind speed (m/s), ρ is measured air density (kg/m^3) and ρ_o is standard air density (kg/m^3). Note that this method assumes a particular relationship between wind speed, air density, and power and assumes that air density does not vary across the rotor disk or vary in time for any given 10-minute period.

In addition to using air density as described in Eq (1), we also included it as an independent variable in the suite of atmospheric inputs. These inputs included varying combinations of: mean wind speed (U) (m/s), density adjusted mean wind speed ($U_{d.adj}$) (m/s), rotor-disk equivalent wind speed (U_{equiv}, also called REWS) (m/s), wind shear (α), turbulence intensity (TI), wind veer (°), and air density (kg/m^3) (Table 3).

U_{equiv} is a variable used to estimate the average inflow seen by the entire rotor disk of the turbine and was calculated using wind speeds at all available lidar heights (see Wagner et al. 2009 for a derivation and theoretical explanation of U_{equiv}). Wind shear is defined as the value of α in the power-law profile of wind speed and indicates the shape of the profile between wind speed at the top and bottom of the rotor disk (e.g., Wharton and Lundquist 2009). On average wind shear was much higher at the Oklahoma Wind Farm. This difference is seen in the rotor disk equivalence wind speeds. At Oklahoma, the absolute difference between U and U_{equiv} exceeded 0.1 m/s for 25% of the observations, compared to 6% at the California site. Turbulence intensity is defined as the ratio of the standard deviation (σ) to the mean wind speed for a given 10-minute period. Note that most manufacturer's power curves are generated assuming a TI of 10%. In contrast, TI values at the California Wind Farm were much higher and over 94% of TI values in the dataset exceeded this value. In Oklahoma mean TI values were much closer to MPC norms and only 30% of the TI means were above 10%.

Table 3: Set of atmospheric variables used as input in the statistical power models

Variable name	Definition	Measurement source
P	Air density	Air temperature and air pressure sensors
U	Mean wind speed	Nacelle cup anemometer, lidar
$U_{d.adj}$	Mean wind speed adjusted to standard density using Eq 1	Nacelle cup anemometer, lidar
U_{equiv}	Rotor-disk equivalent wind speed	Lidar
α	Wind speed shear across the rotor disk	Lidar
TI	Turbulence intensity	Nacelle cup anemometer, lidar
Veer	Wind direction shear across the rotor disk	Lidar

The baseline method of modeling power was done using the IEC density adjusted wind speed at hub-height. Other sets of models included additional information, either in the form of the temporal variability in the wind speed (σ or TI), spatial variability across the rotor-disk (U_{equiv}, α or veer), air density, or various combinations of the variables. Because the dimensions of some of the sets were large, we used principal component analysis (PCA) to reduce dimensions and simplify fitting the statistical models. The top components which together accounted for at least 99% of the total variance were retained.

4.3 Performance metric

To assess the performance of each of the sets of inputs as well as the three statistical models, the observational data sets (SCADA power and atmospheric inputs) were randomly divided into two mutually exclusive sets. The first set was used to train the models. Each combination of the trained model and input set was then used to make wind power predictions for the remaining half of the data set and compared to the observed power values. In addition, we obtained predictions of power based on the manufacturer's power curve using training and validation sets with hub-height wind speed only to assess statistical model performance in comparison to the standard MPC method.

The data partitioning, model training and validation was repeated a total of 30 times for each input set in order 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 power data was computed as a metric of predictive skill.

5. Results and discussion

5.1. Oklahoma

We first consider the results for the input sets involving only the density-adjusted wind speed variables (hub-height wind speed, all wind speeds, standard deviation of wind speed). Figure 11 shows the RMSEs, as fractions of the rated power (1.68 MW) for each of the 30 experiments. Also shown are the RMSEs for the predictions obtained using the MPC with density-adjusted wind speed mean at 80 m as an input. The comparison of the RMSEs in Figure 11 reveals a number of important findings.

1. Lidar 80 m wind speed data provides an improvement in power prediction (RMSE = 4.3%) versus using the nacelle hub-height wind speed (RMSE = 4.8%). This suggests that the lidar provides a more accurate measurement of the inflow conditions to turbine B6 than the nacelle anemometer which is located behind the rotating blades.

2. A small improvement was found using the statistical models versus relying on the manufacturer's power curve for predicting power response.

3. The most significant improvement is found after including the additional lidar measurements. Using either a calculated rotor-disk equivalent wind speed or all wind speeds from each of the lidar measurement heights improved power prediction (RMSE = 3.3-3.5%). Interestingly, the best predictor of turbine power generation for B6 was the combination of nacelle wind speed plus all of the available wind speed measurements from the lidar (RMSE = 2.5%). With this input set the RMSE was nearly cut in half compared to relying on nacelle hub-height wind speed alone.

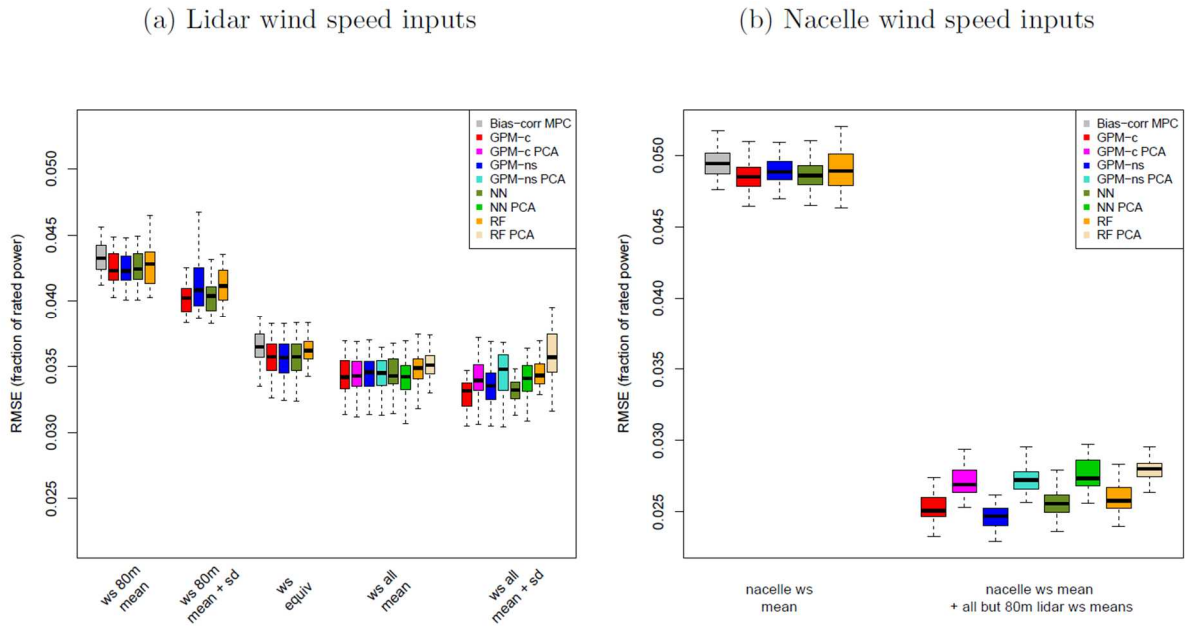

Figure 11: Boxplot distributions of RMSE (as fraction of rated power) for turbine B6 at the Oklahoma Wind Farm. Each box represents the range of RMSE for 30 experiments. Density-adjusted input variables included hub-height wind speed only (ws 80m mean), hub-height wind speed and standard deviation (ws 80m mean + sd), rotor-disk equivalent wind speed (ws equiv), all ten heights of wind speed (ws all mean), and all ten heights of wind speed and their standard deviations (ws all mean + st). RMSE in plot (a) are based on lidar measurements; RMSE in plot (b) are based either on nacelle wind speed only or on a combination of nacelle wind speed and lidar. Model type is color coded and defined in the legend.

At the Oklahoma Wind Farm it appears that shear across the rotor-disk has a greater impact on power generation than turbulence (represented in Figure 11 by standard deviation (sd)). Shear can be included in the models by either calculating the rotor-disk equivalent wind speed or by using all ten measurement heights from the lidar as input. Thirdly, shear can be included by including the shear exponent α as model input as was done in Figure 12. This figure also includes turbulence intensity as an input variable to represent the relative magnitude of turbulence in the inflow.

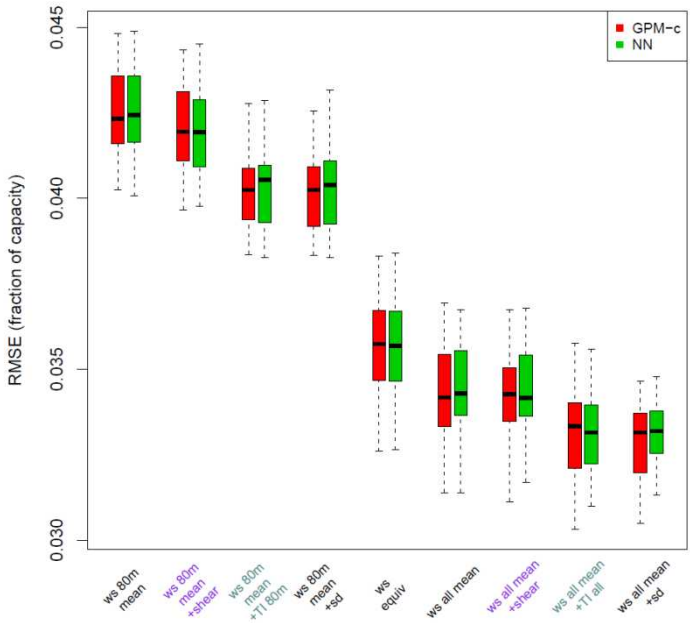

Figure 12: Boxplot distributions of RMSE (as fraction of rated power) for turbine B6 at the Oklahoma Wind Farm. Each box represents the range of RMSE for 30 experiments. This figure shows the predictive power of using the derived values of the shear exponent and turbulence intensity when added to the lidar inputs in Figure 11. The wind shear exponent α offers little benefit to power prediction in contrast to including all lidar measurement heights in the form of mean wind speed or rotor-disk equivalent wind speed. TI and standard deviation offer equal but little overall improvement. For simplicity only the best performing models are shown (GPM-c and NN).

It is interesting that the wind shear exponent is a poorer predictor of power generation than rotor-disk equivalent wind speed or all ten of the lidar wind speed measurements. While α is based on lidar wind speed, it is calculated only from two measurement heights: 40 and 120 m. The wind shear exponent is heavily relied on by the wind industry but our results at the Oklahoma farm suggest that while α reflects the shape of the wind profile, it only partially captures the information found in the entire vertical profile of the mean wind speed. The inclusion of high resolution measurements across the entire profile shows a significant advantage at the Oklahoma site for predicting power response.

We next examined the role of air density and wind veer as predictors for power. Figure 13 shows the percent reduction in RMSE due to adding wind veer to the models in Figure 11. The results show that wind veer information adds a very small improvement. Adding air density as a separate input variable provided very little improvement as well as the two ways of accounting for air density (IEC density-adjusted wind speed and wind speed + air density) performed approximately equal.

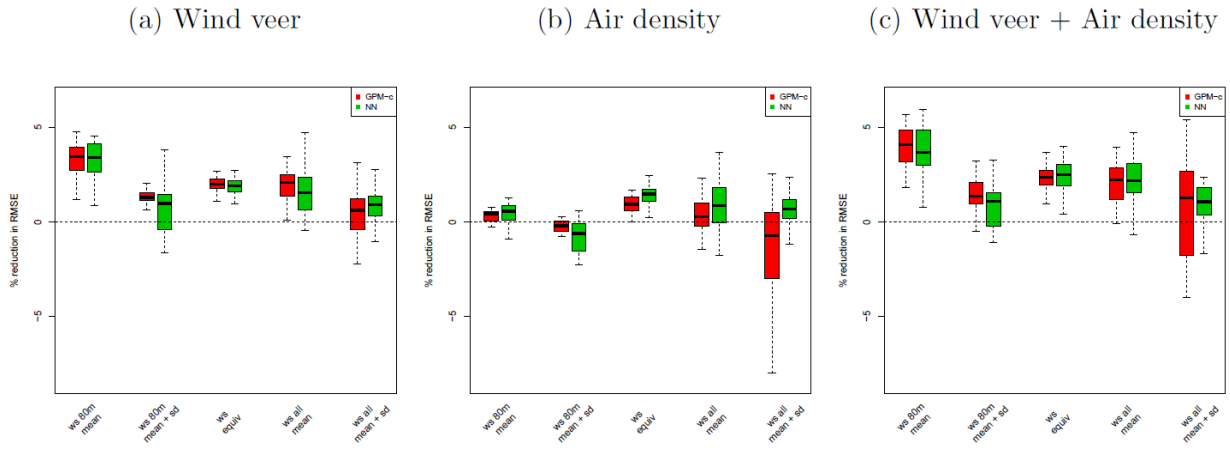

Figure 13: Boxplots of *percent reductions* in RMSE at the Oklahoma Wind Farm from adding (a) wind veer, (b) air density as a separate variable, and (c) the combination of wind veer and air density to the input sets in Figure 11.The dashed horizontal line represents a RMSE reduction of 0. A positive reduction in RMSE implies that the models with added wind veer and/or air density resulted in a lower RMSE. For simplicity only the best performing models are shown (GPM-c and NN).

Finally we evaluated the performance of the statistical models. NN, RF, and GPM performed more or less equally. The fact that GPM performs no worse than the other two models is important because it has the advantage of providing uncertainty estimates in a much more natural way than RF and NN. Given this advantage and comparable performance to that of NN and RF, GPM may be the preferred model.

5.2 California

5.2.1. Single unwaked turbine

Power prediction for the unwaked turbine A8 was earlier explored in Bulaevskaya et al. (2015). In that work we discovered that the statistical models provided greater accuracy in predictions of power generation than the reliance on the manufacturer's power curve. As such, the MPC results are not discussed below and we only present results from the suite of statistical models. We also found in the earlier study that treating air density as a separate input variable provided a significant improvement in power prediction versus using the IEC standard correction methodology (Eq 1). Hence, the results below show input sets that include air density as a free term in contrast to adjusting the wind speed based for standard density.

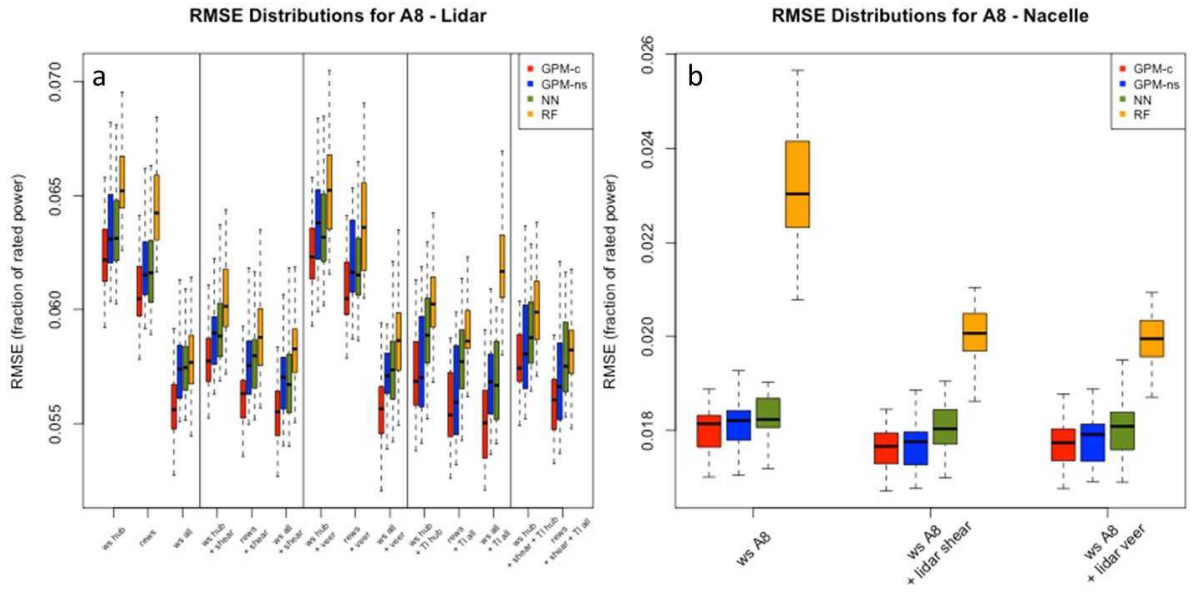

Figure 14: Boxplot distributions of RMSE (as fraction of rated power) for turbine A8 at the California Wind Farm. Each box represents the range of RMSE for 30 experiments. All models were run using air density as an additional input to the variable(s) listed on the x-axis. The variables in (a) are grouped to include: wind speeds only, wind speeds plus the shear exponent, wind speeds plus veer, wind speeds plus turbulence intensity, and lastly, wind speeds plus all three aforementioned variables. In (a) wind speed (ws) includes either hub-height (hub), rotor disk equivalent (rews) or all twelve available height measurements (all). Plot (b) shows RMSE for the following inputs: nacelle wind speed, nacelle wind speed plus the shear exponent, and nacelle wind speed plus veer. Model type is color coded and defined in the legend. Note that the scales on the y-axis are different in plots (a) and (b).

Figure 14 shows the RMSEs at A8, as fractions of the rated power (1 MW) for each of the 30 experiments using a large range of atmospheric inputs from the lidar and nacelle anemometer. Here we observe a number of significant findings:

1. There is a large difference in predictive ability between using the nacelle hub-height wind speed and the free-stream measurement taken with lidar. The lidar hub-height wind speed results in a 4% larger error compared to using wind speed measured at the turbine. This is not surprising since complex terrain is generally associated with more complex physical processes than simpler terrain, making wind power prediction at the site more difficult based on an upwind vertical profiling lidar.

2. As previously discussed in Section 4.1.2 determining the proxy "hub-height" wind speed for the lidar was not straightforward at the California site. While we found the highest

correlation between the 90 m lidar wind speed and nacelle hub-height out of all of the lidar measurements, the one-to-one agreement was not perfect. This is likely the reason why adding all of the lidar heights to the input set significantly improved power prediction in Figure 14a (compare ws hub to ws all). The entire wind profile adds significant improvement to the power prediction. In this case improvement is seen by adding either all of the wind speed heights or by using the wind shear exponent while the rotor disk equivalent wind speed added little benefit, likely because it is weighted towards our proxy for "hub-height".

3. Information about turbulence, in this case by including turbulence intensity, led to a significant reduction in RMSE in California. This is in contrast to the findings in Oklahoma where turbulence added little to no predictive advantage.

4. Veer appeared to offer no real advantage to the power prediction models either because veer is relatively small at the site or because the lidar estimate of veer was poorly correlated with veer found across the actual rotor disk.

5. It appears that for modeling a single turbine in complex terrain, having information about the nacelle wind speed leads to the highest prediction accuracy and having measurements of shear and veer from the upwind lidar provided little additional benefit (Figure 14b).

6. Prediction uncertainty was smaller when nacelle wind speed was used as model input as compared to the lidar data.

7. GPM with a constant spline (GPM-c) appears to produce the lowest RMSE out of all of the statistical models.

5.2.2. Unwaked turbine row

Next we predicted power for Row A turbines at the California Wind Farm. This row included sixteen turbines that are unwaked when winds are from the southwest (Figure 5). The statistical models were run in a similar fashion as for A8 with the exception that a much larger SCADA dataset was used and included nacelle wind speed, direction and power from each of the turbines in Row A. Figure 15 shows the distribution of modeled results for the unwaked turbine row. Here we observe a number of similar findings as we did with turbine A8. Important differences include:

1. The differences in modeled error between using lidar input and nacelle input are even greater across the unwaked row of turbines than they were for the individual turbine A8. Here we see a mean predictive error of 9.3% when using lidar input versus a 1% error for nacelle hub-height wind speed.

2. Adding lidar shear and turbulence intensity to the predictive model for the entire turbine row added less benefit than it did for the most directly downwind turbine (A8).

3. Uncertainty is 2-3 times higher for the entire turbine row than it was for the single turbine A8.

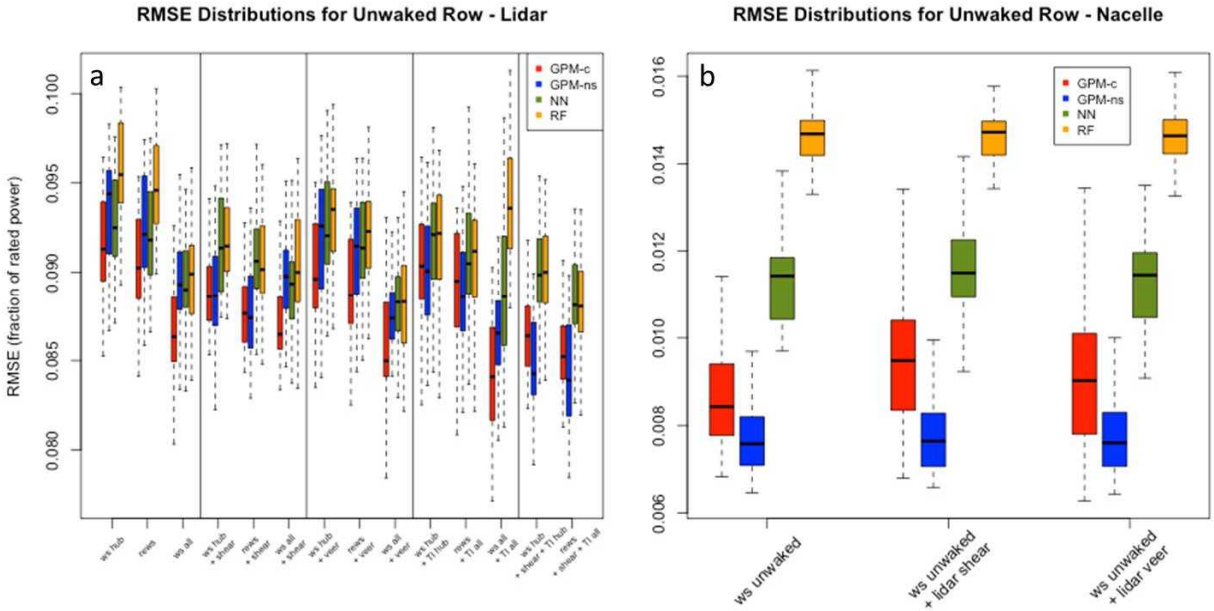

Figure 15: Same as Figure 14 except that results show RMSE distributions for all sixteen turbines (A1-A16) in an unwaked row at the California Wind Farm.

5.2.3. Wind park

Finally we predicted power for all of the 38 turbines at the California Wind Farm. The selected wind sector (225-250°) put some of the turbines directly downwind of others. As such, this model set contains both waked and unwaked turbines. Figure 16 shows the distribution of modeled results for the entire wind farm. Here we observe a number of similar findings as we did with turbine A8 and the unwaked A row. These include:

1. Nacelle wind speed is superior to lidar wind speed for predicting power generation. Lidar measured wind shear and veer add little to no additional improvement over nacelle hub-height wind speed.

2. Both mean error and uncertainty increase as the number of modeled turbines increase if atmospheric input is based only on lidar measurements. Mean error for A8 was 6%, for Row A was 9% and for the entire wind farm was 11%.

3. Information about wind speed across the entire rotor disk is superior to only using hub-height wind speed for the lidar input set.

4. The GPM produces the smallest predictive errors and is an overall better performed than NN and GF.

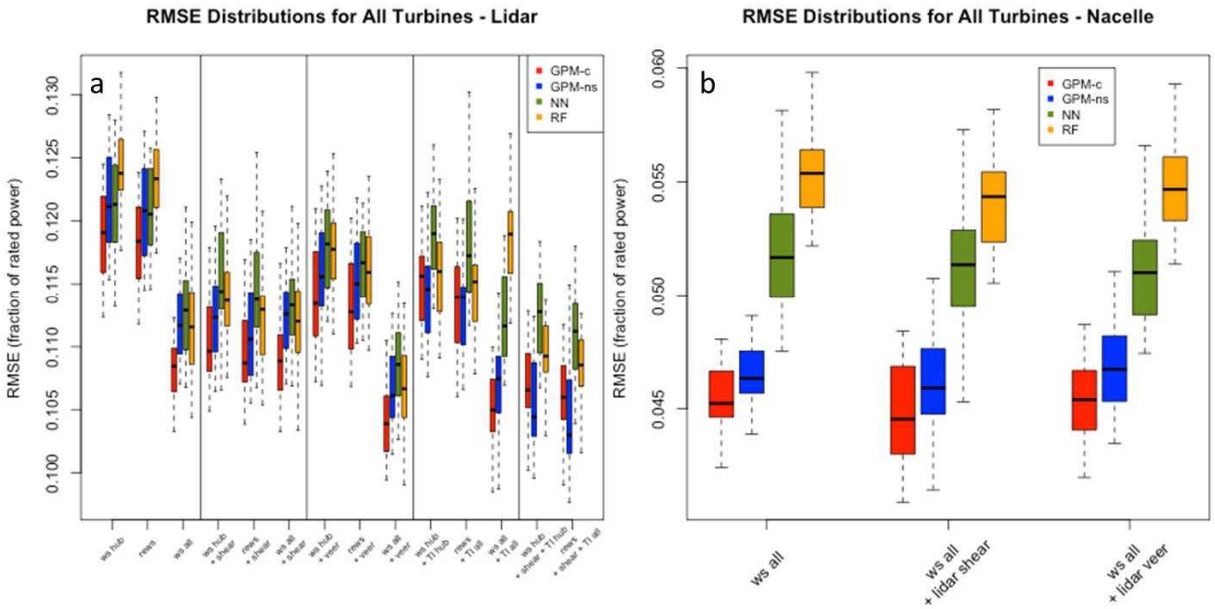

Figure 16: Same as Figures 14 and 15 except that results show RMSE distributions for all turbines at the California Wind Farm.

While the findings for turbine A8, turbine row A, and the entire wind park are similar for the most part, a few distinctions are also apparent for the best performing statistical model (GPM-c) (Figure 17):

1. Mean error did not increase monotonically from modeling a single turbine to the entire wind farm when nacelle wind speed was used as input. Instead mean RMSE was lowest to highest for Row A turbines (1%), turbine A8 (2%), and the entire wind farm (5%), respectively.

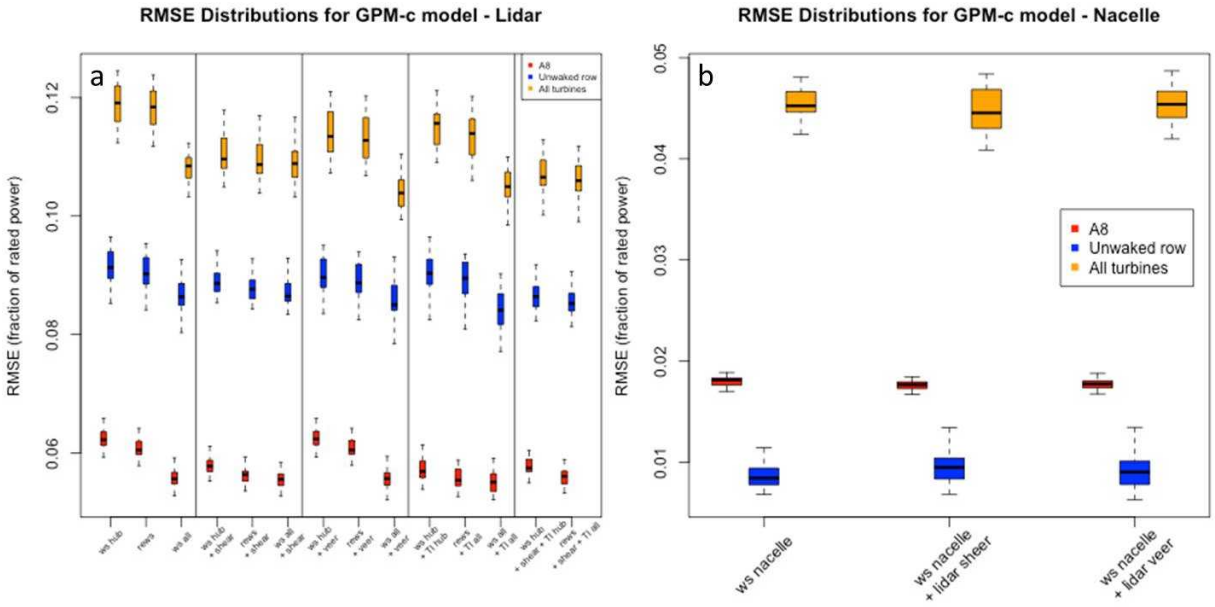

Figure 17: Distribution of RMSE for the lidar (a) and nacelle (b) atmospheric inputs from the best performing GPM-c model. The results are taken from Figures 14-16 and plotted together to compare model prediction errors for the three model sets: single turbine A8 (red), unwaked Row A turbines (blue), and all turbines (yellow). As in all the model runs, each box represents the distribution of RMSE from 30 runs for each atmospheric input set.

6. Conclusions

1. An improvement was found using the statistical models versus relying on the manufacturer's power curve (MPC) for predicting power response. Given that statistical models provide estimates of uncertainty and allow for a larger number of input variables they have a significant advantage over the MPC for predicting power generation.
2. Predictive uncertainty increases with increasing number of modeled turbines. For example, uncertainty was 2-3 times higher for the entire turbine row than it was for the single turbine at the California farm.
3. Predictive error was greater for the complex terrain wind farm (mean error = 6.2%) than the flat terrain site (mean error = 4.3%) when lidar wind speed was used as inflow.
4. Vertical profiling lidar proved most useful for predicting power response at the flat terrain site. For example, lidar measured wind shear and equivalent wind speed reduced the predictive error at the flat terrain site when used to augment nacelle wind speed but had no measurable advantage at the complex terrain farm over nacelle wind speed alone.

7. Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy (DOE) by Lawrence Livermore National Laboratory under the Contract No. DE-AC52-07NA27344. Funding came from the Wind and Water Power Technologies Office under the Office of Energy

Efficiency and Renewable Energy (EERE) at DOE (Grant No. 03.0.301). Prior funding for the lidar field campaigns came from a Laboratory Directed Research and Development Grant No. 12-ERD-069. The authors are extremely grateful for our partnerships with Enel Green Power and Infigen Energy, including data sharing, intellectual input, and field campaign assistance at the wind farms. The authors would also like to thank Cary Gellner, Maureen Alai, and John van Fossen (LLNL) for building the solar panel/battery arrays which were used in California. Lastly, sincere appreciation goes to the wind farm site operators at Enel and Infigen. Without their help and support this project would not have been possible.

8. References

Bishop, C. (2007) *Pattern Recognition and Machine Learning*, Springer, 225-281 p

Breiman, L. (2001) Random forests. *Mach. Learn*. 45, 5-32

Bulaevskaya, V., Wharton, S., Clifton, A., Qualley, G., Miller, W.O. (2015) Wind power curve modeling in complex terrain using statistical models. *J. Renew. Sustain. Energy* 7, 013103

Clifton, A., Kilcher, L., Lundquist, J.K., Fleming, P. (2013) Using machine learning to predict wind power output. *Environ. Res. Lett*. 8, 024009

Fonte, P.M., Quadrado, J.C. (2005) *ANN approach to WECS power forecast*. Proceedings of the 10th IEEE Conference on Emerging Technologies and Factory Automation, Catania, Italy

Gramacy, R.B. and H.K.H. Lee (2008) Bayesian treed Gaussian process models with an application to computer modeling. *J. Am. Stat. Assoc*. 103, 1119-1130

Hastie, T., Tibshirani, R., Friedman, J. (2009) *The Elements of Statistical Learning*, Springer, 587-602 p

International Electotechnical Commission (IEC) International Standard 61400-12-1: Wind turbine generator systems – wind turbine power performance testing (1998)

Klein, P., Bonin, T., Newman, J., Turner, D., Chilson, P., Wainwright, C., Blumberg, W., Mishra, S., Carney, M., Jacobsen, E., Wharton, S., Newsom, R. (2015) LABLE: a multi-institutional, student-led, atmospheric boundary-layer experiment. *Bull. Amer. Meteor. Soc*. doi:10.1175/BAMS-D-13-00267.1, in press

Krishnamurthy, R. and M. Boquet (2014) *Case studies of WINDCUBE measurement uncertainty for complex terrain using FLOW complexity recognition (FCR).* Conference Proceedings at the European Wind Energy Association, Barcelona, Spain

Motta, M., Barthelmie, R., Volund, P. (2005) The influence of non-logarithmic wind speed profiles on potential power output at Danish offshore sites. *Wind Energy* 8, 219-236

Power Curve Working Group (2015*) Guidelines for preparation of a turbine performance information pack* (draft), 9 p

Rasmussen, C.E. and C.K.I. Williams (2005) *Gaussian Processes for Machine Learning*, The MIT Press

Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P. (1989) Design and analysis of computer experiments. *Stat. Sci.* 4, 409-423

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

Slinger, C. and M. Harris (2012) Introduction to continuous-wave Doppler lidar. In: Proceedings of the Summer School in Remote Sensing for Wind Energy, 11-15 June 2012, Boulder, Colorado, 32 p

Song, J., Liao, K., Coulter, R.L., Lesht, B.M. (2005) Climatology of the low-level jet at the Southern Great Plains Atmospheric Boundary Layer Experiments site. *J. Appl. Meteor*. 44, 1593-1606

Wagner, R., Antoniou, I., Pedersen, S.M., Courtney, M.S., Jorgensen, H.E. (2009) The influence of the wind speed profile on wind turbine performance measurements. *Wind Energy* 12, 348-362

Wharton, S. and J.K. Lundquist (2012) Atmospheric stability affects wind power collection. *Environ. Res. Lett.* 7, 014005

Wharton, S., Simpson, M., Osuna, J.L., Newman, J.F., Biraud, S.C. (2015a) Role of surface energy exchange for simulating wind turbine inflow: a case study in the Southern Great Plains, USA. *Atmosphere* 6, 21-49

Wharton, S., Newman, J.F., Qualley, G., Miller, W.O. (2015b) Measuring turbine inflow with vertically-profiling lidar in complex terrain. *J. Wind Eng. Ind. Aerodyn*. 142, 217-231