Evaluation of Wind Power Forecasts from the Vermont Weather Analytics Center and Identification of Improvements

Mike Optis, George Scott, and Caroline Draxl
National Renewable Energy Laboratory

NREL is a national laboratory of the U.S. Department of Energy Office of Energy Efficiency & Renewable Energy Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Technical Report
NREL/TP-5000-70313
February 2018

Contract No. DE-AC36-08GO28308
Evaluation of Wind Power Forecasts from the Vermont Weather Analytics Center and Identification of Improvements

Mike Optis, George Scott, and Caroline Draxl
National Renewable Energy Laboratory

Prepared under Task No. GMLC.10246.09.01.08
NOTICE

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at SciTech Connect http://www.osti.gov/scitech

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
OSTI http://www.osti.gov
Phone: 865.576.8401
Fax: 865.576.5728
Email: reports@osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312
NTIS http://www.ntis.gov
Phone: 800.553.6847 or 703.605.6000
Fax: 703.605.6900
Email: orders@ntis.gov

Cover Photos by Dennis Schroeder: (left to right) NREL 26173, NREL 18302, NREL 19758, NREL 29642, NREL 19795.

NREL prints on paper that contains recycled content.
# List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASOS</td>
<td>Automated Surface Observing System</td>
</tr>
<tr>
<td>DER</td>
<td>distributed energy resources</td>
</tr>
<tr>
<td>GMW</td>
<td>Georgia Mountain Wind Farm</td>
</tr>
<tr>
<td>KCW</td>
<td>Kingdom Community Wind Farm</td>
</tr>
<tr>
<td>NREL</td>
<td>National Renewable Energy Laboratory</td>
</tr>
<tr>
<td>SCADA</td>
<td>supervisory control and data acquisition</td>
</tr>
<tr>
<td>SHF</td>
<td>Sheffield Wind Farm</td>
</tr>
<tr>
<td>VELCO</td>
<td>Vermont Electric Power Company</td>
</tr>
<tr>
<td>VTWAC</td>
<td>Vermont Weather Analytics Center</td>
</tr>
</tbody>
</table>
Executive Summary

The goal of this analysis was to assess the wind power forecast accuracy of and identify potential improvements to the Vermont Weather Analytics Center (VTWAC) forecast system. The VTWAC system is operated in partnership by the Vermont Electric Power Company and IBM and is part of a larger model that incorporates renewable power forecasting for utility-scale wind and solar power plants as well as distributed solar and electricity demand forecasting. Inputs to VTWAC include smart meter data, supervisory control and data acquisition (SCADA) data from renewable energy power plants, and weather data. The VTWAC system incorporates statistical and analytic models that are trained by the input data. Applications of VTWAC include long-term planning, short-term planning, and peak energy management at the Vermont Electric Power Company.

Data were received from four separate wind farms in Vermont. These data included measured and forecasted power at the wind farm along with turbine-level data for three of the four plants. All data were provided as 10-minute averages from June 2016 to September 2017. The forecast data provided were based on forecasts at 0000 Coordinated Universal Time (or 1900 Eastern Standard Time) and covered a 24-hour look-ahead period.

A robust assessment of forecast performance requires that the measured data to which forecasts are compared are representative of normal turbine and power plant operation (e.g., no turbine downtime or grid curtailment). Measured data should be either filtered or corrected to normal operation based on a detailed analysis of turbine-level data (e.g., fault and status codes provided from SCADA data).

Insufficient SCADA data were provided to correct periods of abnormal turbine and power plant operation. Therefore, the abnormal data were filtered from the analysis using the National Renewable Energy Laboratory’s best judgment and experience. After data filtering, only the Georgia Mountain wind power plant had sufficient data to proceed with the forecast analysis. Forecast performance at the other three plants was not analyzed.

Based on the analysis at Georgia Mountain, the National Renewable Energy Laboratory had the following recommendations for improving forecast performance:

- Determine the cause of significant negative forecast bias from February to March 2017 (50% underprediction on average)
- Improve the ability of the forecast model to capture the large amplitude diurnal cycles of wind power
- Add the ability to the forecast model to assess internal wake loss (if no such capability currently exists), particularly at sites where strong diurnal shifts in wind direction are present.

Data availability and quality limited the robustness of this forecast assessment. A more thorough analysis would be possible given a longer period of record for the data (at least one full year), multiple forecast times and look-ahead times up to 72 hours, detailed SCADA data for each wind power plant (i.e. turbine and plant-level power at 1-hour resolution or higher along with...
complete event log with documented status codes), and more detailed information on the forecast system input data and forecast methodologies.
1 Introduction

The goal of this analysis was to assess the wind power forecast accuracy of and identify potential improvements to the Vermont Weather Analytics Center (VTWAC) forecast system, which is operated in partnership by the Vermont Electric Power Company (VELCO) and IBM. The wind forecast system is part of a larger model that incorporates renewable power forecasting for utility-scale wind and solar power plants as well as distributed solar and electricity demand forecasting. Inputs to VTWAC include smart meter data, supervisory control and data acquisition (SCADA) data from renewable energy power plants, and weather data. The VTWAC system incorporates statistical and analytic models that are trained by the input data. Applications of VTWAC include long-term planning, short-term planning and peak energy management at VELCO.

Accurate wind power forecasts are important (and challenging) because of the intermittent nature of the wind resource, particularly in complex terrain (e.g., mountainous regions, land-water interfaces). Robust accounting of the diurnal cycle is particularly important because of the rapidly evolving wind regimes from day to night on a typical day. Most of Vermont’s wind farms are either located within the Champlain Valley (and close to Lake Champlain) or in other mountainous terrain. These areas are associated with complex wind flow regimes that make accurate forecasting a challenge.

Using available measured data (at the power plant level for individual turbines), forecast data, and supplemental data (e.g., weather station data) provided by VELCO, this analysis aims to benchmark the performance of wind power forecasts from VTWAC. Performance will be assessed both on seasonal and diurnal timeframes, with the goal of identifying opportunities for forecast improvement.

This report is organized as follows: Section 2 provides a description of the wind farms considered in the analysis. The data provided by VELCO is summarized in Section 3. Methods used to filter the power data for periods of normal turbine and plant operation are provided in Section 4. Results of the forecast analysis are presented in Section 5. Comments on the analysis and recommendations for forecast improvement are summarized in Section 6.
2 Wind Farm Descriptions

For this study, data were received from four separate wind farms in Vermont—one in the south and three in the north. Table 1 shows the details of each wind farm.

<table>
<thead>
<tr>
<th>Wind Farm</th>
<th>Abbreviation</th>
<th>Turbines</th>
<th>Turbine Type</th>
<th>Rotor Diameter (meters [m])</th>
<th>Hub Height (m)</th>
<th>Turbine Rated Power (megawatts [MW])</th>
<th>Wind Farm Capacity (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georgia Mountain</td>
<td>GMW</td>
<td>4</td>
<td>GoldWind GW100</td>
<td>100</td>
<td>80</td>
<td>2.50</td>
<td>10.00</td>
</tr>
<tr>
<td>Kingdom Community</td>
<td>KCW</td>
<td>21</td>
<td>Vestas V112-3.0</td>
<td>112</td>
<td>84</td>
<td>3.00</td>
<td>63.00</td>
</tr>
<tr>
<td>Searsburg</td>
<td>SBG</td>
<td>11</td>
<td>Zond Z-40FS</td>
<td>40</td>
<td>NA</td>
<td>0.55</td>
<td>6.05</td>
</tr>
<tr>
<td>Sheffield</td>
<td>SHF</td>
<td>16</td>
<td>Clipper C93/C96</td>
<td>93</td>
<td>80</td>
<td>2.50</td>
<td>40.00</td>
</tr>
</tbody>
</table>

Figure 1 shows the locations of the wind farms within Vermont. Turbine layouts are shown in Figure 2. The National Renewable Energy Laboratory (NREL) has a database with precise locations of most of the wind turbines in the United States. The database has only 3 of the 11 turbines in the Searsburg wind farm, but is otherwise complete. Each map in Figure 2 is approximately 7-by-7 kilometers (km) in scale.
Figure 1. Locations of wind farms in Vermont
Figure 2, upper left. Map showing the turbine locations in Georgia Mountain.

Figure 2. Turbine locations for the wind turbines in the four wind farms analyzed. The database has only 3 of the 11 turbines in the Searsburg wind farm, but others are complete.
3 Data Provided

3.1 Power Forecast Data
Forecast data were provided by VELCO from June 2016 through September 2017. Data were provided as CSV files.

Each file consisted of four fields: measured and forecasted power at the wind farm, and low and high uncertainty bounds (presumably on the forecast data). These uncertainty data fields were not considered in this analysis. The sampling interval was 10 minutes (min). It was unclear if data were instantaneous measurements and forecasts at the time recorded, or whether the data were averaged at the beginning, middle, or end of the reported time.

Both measured and forecasted data were missing for some time intervals. These intervals were necessarily excluded from the analysis.

The forecast system is run every 12 hours (0000 and 1200 Coordinated Universal Time (UTC), corresponding to 1900 and 0700 Eastern Standard Time (EST)) over a look-ahead period of 72 hours. However, forecast data available for this analysis represent only forecasts made at 0000 UTC (1900 EST) and with a look-ahead time of up to 24 hours.

3.2 Turbine Data
VELCO provided turbine-level data for three of the four wind farms (no data were received for Searsburg). Data were obtained from each wind farm’s SCADA system. Data typically cover the period of March 2016 through June 2017, although data gaps exist at different time periods for the different wind plants. The formatting of and information contained within the turbine-level data provided for each wind farm varied considerably, which is not unusual given the different wind farm operator/owners and turbines and the associated diversity in SCADA systems. However, nacelle wind speed and power were reported for all wind farms, as well as some indication of whether the turbine was performing in a normal state. A more detailed description of the turbine-level fields and how they were used in this analysis is provided in the next section.
4 Data Filtering and Correction

The VELCO forecast system assumes that all turbines for a given wind farm are operating normally (e.g., turbines are not offline or being operated in a derated state), and that power from the wind power plant is not being curtailed. Consequently, a robust assessment of forecast system performance requires that the measured power to which the forecast is compared is representative of normal plant operation. There are two general approaches to ensure that measured power is representative of normal power plant operation:

1. Data representing abnormal plant operation (e.g., turbine downtime, plant curtailment, and so on) are filtered out of the analysis
2. Data representing abnormal plant operation are corrected to represent normal plant operation (e.g., plant power is corrected to 100% plant availability and 0% curtailment).

These approaches require sufficient data either at the turbine level (e.g., detailed SCADA fields describing turbine state, fault codes, and so on) or the power plant level (e.g., summaries of plant availability and total curtailment for each time interval).

The only plant-level data provided for this analysis were measured and forecast power. Therefore, only the turbine-level data were used for filtering and correction. Because no turbine-level data were provided for the Searsburg wind farm, forecast performance for this wind farm was not considered in this analysis.

4.1 SCADA Fields Used for Filtering

SCADA data from a wind farm typically provide a detailed account of individual turbine performance. Output from a SCADA system can include hundreds of data fields describing common parameters, such as nacelle wind speed, direction, and ambient temperature, but also more detailed parameters, such as blade pitch angles and generator revolution speeds. Of particular importance are status/fault codes and event logs, which provide a highly detailed account of the different factors affecting turbine performance. These factors are extensive and include external power limitation (i.e., curtailment), low wind speed cut out, restart delays, scheduled maintenance, and various unscheduled downtimes (e.g., cable twisting, blade angle asymmetries, gearbox overheating). Proper accounting of status codes, fault codes, and event logs allows for the filtering or correction of time intervals in which turbines are not operating normally.

For the Georgia Mountain, Kingdom Community, and Sheffield wind farms, some information on status and/or fault codes were provided in the SCADA data. However, such information was limited (e.g., status and fault codes were not defined). Therefore, NREL researchers used their best judgement based on SCADA data experience to filter abnormal operation data. Table 2 summarizes the SCADA fields employed from each wind farm that were used to attempt data filtering and/or correction.
Table 2. SCADA Fields Used to Filter Out Abnormal Turbine Operation Data

<table>
<thead>
<tr>
<th>Wind Farm</th>
<th>SCADA Fields</th>
<th>Interpretation and Data Filtering Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMW</td>
<td>op_data_WTG_ok_time_s(h); op_d_grid_ctrl_stnd_t_(h)</td>
<td>The first field appeared to indicate the time length (in hours) during which the turbine was operating well (0.16-0.17 indicative of 10 min). The second field appeared to indicate curtailment time (measured same as first field). Data were filtered for cases where 'op_data_WTG_ok_time_s(h)' was between 0.16 and 0.17 and 'op_d_grid_ctrl_stnd_t_(h)' was equal to zero.</td>
</tr>
<tr>
<td>KCW</td>
<td>SYS_LOGS_FIRSTACTALARMNO; GRD_PROD_PWR_AVG; GRD_PROD_PSBLEPWR_AVG</td>
<td>The first field was either a state or fault code, with code 0 appearing to indicate good turbine operation. The second field appeared to be actual wind power, whereas the third field appeared to be possible power (e.g., if power was not curtailed, turbine was operational, and so on). Data were filtered for cases in which 'SYS_LOGS_FIRSTACTALARMNO' was 0 and actual power was within 5% of possible power.</td>
</tr>
<tr>
<td>SHF</td>
<td>DERATECAUSE; FAULTCODE</td>
<td>The first field appeared to denote the cause for turbine derating (code 0 appeared to indicate that the turbine was not derated. The second field appeared to denote turbine fault code, with code 0 appearing to indicate normal turbine operation. Data were filtered for cases in which both of these fields were equal to zero.</td>
</tr>
</tbody>
</table>

4.2 Data Filtering Results

Figure 3 shows a turbine power curve for Turbine 1 at the Georgia Mountain wind farm. Power curves for the other three turbines at the wind farm show similar characteristics. Data that have been filtered out based on the criteria in Table 2 are shown in red. As shown in Figure 3, the data generally follow the power curve with typical spread. The data filtering is successful in removing most of the data showing zero power but nonzero wind speeds (likely representing turbine downtime). Some of the outlier data, and in particular an apparent derating around 300 kilowatts, are not filtered out.
Figure 3. Power curve for Turbine 1 at Georgia Mountain wind farm (power in kilowatts [kW], wind speed in meters per second [m/s])

Figure 4 and Figure 5 show power curves for Turbines 1 and 18 at Kingdom Community wind farm. Power curves for other turbines show similar characteristics but are slightly less extreme. As shown in the figures, a considerable amount of data fall outside the expected power curve. The filtering criteria described in Table 2 remove most of these data. However, the apparent derating at 2,500 kW for Turbine 1 is not captured. Furthermore, a skewed power curve is apparent for Turbine 18 (and was observed for most turbines), wherein the top “knee” of the power curve is derated before reaching a peak rating of 3,000 kW above 15 meters per second (m/s).
Figure 4. Power curve for Turbine 1 at Kingdom Community wind farm

Figure 5. Power curve for Turbine 18 at Kingdom Community wind farm
Figure 6 and Figure 7 show power curves for Turbines 1 and 10 at Sheffield wind farm. Power curves for other turbines show similar characteristics but are slightly less extreme. As shown in the figures, a considerable amount of data fall outside the expected power curve. The filtering criteria described in Table 2 remove the data with zero power and derating for Turbine 10. However, a considerable amount of data well outside the power curve are not filtered.

Figure 6. Power curve for Turbine 11 at Sheffield wind farm
Given the lack of more comprehensive SCADA data with detailed state and fault code accounting, it was not possible to correct the abnormal data to those representing normal turbine operation (i.e., 100% turbine availability and zero curtailment). The abnormal data for each turbine were simply excluded from the analysis. Therefore, a robust assessment of forecast system performance was limited to time intervals in which all turbines were operating normally. For wind power plants with many turbines (i.e., Kingdom Community and Sheffield), this requirement that all turbines are operating normally resulted in considerable data exclusion (summarized in Table 3). As noted above, because no turbine-level data were provided for the Searsburg wind farm, forecast performance for this wind farm was not considered in this analysis.

### Table 3. Data Excluded Because of Turbine-Level Filtering

<table>
<thead>
<tr>
<th>Wind Power Plant</th>
<th>Data Excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georgia Mountain</td>
<td>21.5%</td>
</tr>
<tr>
<td>Kingdom Community</td>
<td>78.8%</td>
</tr>
<tr>
<td>Sheffield</td>
<td>92.6%</td>
</tr>
</tbody>
</table>

As shown in Table 3, most of the data were excluded for the Kingdom Community and Sheffield wind power plants. Such exclusions resulted in insufficient data to robustly assess seasonal and
diurnal forecast accuracy trends. Consequently, Kingdom Community and Sheffield wind farms were excluded from this forecast assessment analysis.

### 4.3 Additional Data Quality Checks

Additional checks were performed on the provided data to assess the representativeness of the data. Specifically, reported measured power in the forecast data file was compared to the sum of turbine power reported in the SCADA data over the concurrent time intervals between the two data sets. Distributions of reported and sum of turbine power are shown in Figure 8. In all cases, there are discrepancies between the two data sets. In general, the sum of turbine power tends to be less than power reported in the forecast data. There is a clear data issue with Georgia Mountain, wherein the power reported in the forecast data shows an unrealistic peak at maximum plant capacity (i.e., 10 megawatts [MW]).

These differences between the two reported power data sets raise some concerns regarding data quality.

![Figure 8. Distributions of 10-min averaged power as reported at the power plant level and as a sum of the turbines](image)
5 Forecast Analysis

Given the data quality issues described in Sections 4.2 and 4.3, only the Georgia Mountain wind farm was considered in this forecast assessment. Furthermore, the time intervals considered were limited to those in which the SCADA data indicated normal operation across all turbines.

Forecast error is defined as forecast power minus measured power (i.e., negative values denote forecast underprediction, and vice versa). Only the reported power data in the forecast file are used when calculating forecast error. Power reported in the SCADA data is not used. Only the reported 10-min data are considered (i.e., forecast error is not assessed for hourly or daily-averaged power data).

5.1 Seasonal Trends

Figure 9 shows boxplots of forecast error for Georgia Mountain wind farm. The sample size for each month is shown above each boxplot. The forecast error is, on average, accurate in the spring months (April through June) as well as September. The forecast model tends to underestimate power from July to August 2016, although this result should be interpreted cautiously because of the low data count for these months. There are clear forecast anomalies from February to March 2017, in which the forecast model underestimates power by 5 MW on average (or 50% of plant capacity). It is likely that the measured data were misreported in these months. Specifically, measured data appear to be erroneously high and would explain the abnormal distribution of measured power in Figure 8.

Figure 9. Boxplots of forecast error (forecast power minus measured power) separated by month; numbers above the boxplots denote sample size
5.2 Diurnal Trends

Diurnal forecast performance is important for utilities in which hourly look-ahead predictions for load balancing are required. Diurnal variations in the wind resource are typically strongest in the spring and summer months when the largest temperature changes from day to night are observed. Under such conditions, strong surface heating during the day—resulting in unstable stratification—can produce considerably different wind profiles than under conditions of surface cooling at night (i.e., stable stratification).

The time period from April to May 2017 provided high data availability across the entire diurnal cycle and is used in this assessment of diurnal forecast performance. All other time periods are excluded from diurnal analysis because of inadequate sample size across the diurnal cycle.

Figure 10 shows the median measured and forecast power over each hour for the April–May 2017 period. A strong diurnal cycle in power at Georgia Mountain wind farm is clearly evident. From night until early morning, median plant output is between 8.0 and 9.5 MW, or 80%–95% of total plant capacity. Power drops considerably during the day, reaching a minimum below 4 MW at 1600 EST.

The forecast model considerably underestimates the magnitude of this diurnal cycle with median forecasted power ranging from about 6 MW during the day to 8 MW at night. The drop in power throughout the day is particularly unrepresented in the forecast model. The boxplot shows that the forecast error has its largest spread during the late afternoon (when power output is lowest) and smallest spread through the night when power output peaks.

Figure 10. Diurnal cycle of median measured and forecasted power (a) and forecast error (b) at Georgia Mountain wind farm for the time period April–May 2017.
The strong diurnal trend in wind power is likely associated with the role of surface heating and cooling and the associated changes in stratification. However, a detailed analysis of the role of stratification was not possible because of a lack of information regarding the design of the forecast system and insufficient data needed to characterize stratification (e.g., measured temperature at two different heights or surface flux measurements from a meteorological tower).

The strong diurnal trend in wind power may also be associated with thermally-driven flow at Lake Champlain. Figure 11 shows the position of Georgia Mountain wind farm relative to Lake Champlain. In general, valley flow regimes are complex and have the following characteristics:

- During the day, warm air flows up along the valley walls, resulting in a pressure decrease within the valley. This decrease draws in air from the plains into the valley and a strong wind along the valley axis is common.
- At night, cold air descends along the valley walls and collects in the valley, resulting in a pressure increase. This increase forces air in the valley back into the plains.
- A lake within a valley results in a more complex flow regime. Depending on water temperature relative to air temperature, the lake can either mitigate or enhance the existing valley flow described here. The lake can also enhance or mitigate the development of stable stratification (e.g., when warm air flows over a cool lake, temperature drops at low heights, inducing stable stratification).
Figure 11. Map of Georgia Mountain wind farm (GMW) and Burlington Automated Surface Observing System location relative to Lake Champlain

Source: ©2017 Google, Image Lands at Copernicus

The flow regime at Lake Champlain can be explored in more detail by analyzing data from nearby meteorological stations. The Automated Surface Observing System (ASOS) station at Burlington International Airport (shown in Figure 11) is located within the valley and is used here. Figure 12 shows present wind speed and direction statistics at the ASOS station based on measurements at 10 m above ground. The figure demonstrates a bimodal distribution of wind direction centered around 180 and 340 degrees, indicative of flow through the valley but in opposite directions. Wind speed is strongly associated with wind direction, peaking when flowing along the valley axis (i.e., ~180 and ~360 degrees) and lowest when flowing cross-axis. Wind speeds increase in the early morning hours and peak midafternoon before decreasing again in the evening. The diurnal shift in wind direction is also notable. Winds tend to be from the south during the night and early morning, with an increasing westerly component throughout the day. The median wind direction in late afternoon is from the west.

The diurnal wind speed profile at the ASOS station is almost exactly opposite to that found at Georgia Mountain wind farm (i.e., ASOS wind speeds peak in the afternoon, Georgia Mountain wind farm power peaks during the night) despite the locations only being about 20 km apart. However, this observation is consistent with strong surface heating and cooling and the different heights for the ASOS station (10-m measurements) and the wind power plant (80-m hub height). During the day, strong convection driven by surface heating transports momentum from aloft to
the surface, resulting in a relative increase in surface wind speeds and a decrease in wind speeds aloft. The opposite occurs at night because of surface cooling and associated stable stratification and surface decoupling, resulting in lower relative wind speeds at the surface and higher winds speeds aloft.

![Wind speed and direction diurnal statistics observed at the Burlington International Airport ASOS station from April to May 2017](image)

Wind direction measurements at Georgia Mountain wind farm were not available, therefore, a direct analysis of measured and forecasted power as it relates to wind direction at the site was not possible. However, power measurements can be related to wind direction at the ASOS station, acknowledging that actual wind direction at Georgia Mountain wind farm may differ from that at the ASOS station given the different locations within the Lake Champlain valley and the different topographies.

Figure 13 shows the median measured and forecasted power at Georgia Mountain wind farm, as well as forecast error, as they depend on wind direction at the ASOS station. Power output at Georgia Mountain wind farm peaks when flow is from the south (i.e., along the valley axis) and
is lowest when flow is from the west. Both the median and spread in forecast error are lowest when flow is from the south, likely because of the fact that wind speeds are high and plant output is near maximum capacity during this time. Forecast error spread is considerably higher in the other wind directions, particularly when winds are from the west. The forecast model shows a large positive bias from this direction but is relatively accurate on average when flow is from the east and north.

Figure 13. Median measured and forecasted power (a) and forecast error (b) at Georgia Mountain wind farm by wind direction at the Burlington Airport ASOS station from April to May 2017

The diurnal shift in wind direction may have consequences for internal turbine waking at Georgia Mountain wind farm. As shown in Figure 2, the four turbines at Georgia Mountain wind farm lie approximately on a line oriented in a north-northwest to south-southeast direction. A typical 90-degree shift in wind direction over the course of a day (see Figure 12) could result in a consistent pattern of unwaked turbines during part of the day and considerable waking a few hours later.

The lack of on-site wind direction data at Georgia Mountain wind farm prevents a detailed analysis of potential waking. Furthermore, the proper identification of Turbines 1 through 4 (as labeled in the SCADA data) on the map in Figure 2 was not possible given limited data. Despite these limitations, some analysis of diurnal waking is possible by analyzing nacelle wind speeds at the four turbines. Figure 14 shows median wind speeds for each turbine as a function of time of day and wind direction. Wind speeds tend to diverge at night with Turbine 3 wind speeds considerably lower than the other turbines, suggesting that Turbine 3 is being waked. Conversely, turbine wind speeds are broadly equivalent in the afternoon, suggesting minimal to no waking. Further evidence of waking is shown in Figure 14(b). In particular, Turbine 4 typically experiences higher wind speeds than the other turbines in the 120–250 degree sector but has the lowest wind speeds in the north sector.
Figure 14. Median turbine wind speeds at Georgia Mountain wind farm as a function of time of day (a) and ASOS station wind direction (b) from April to May 2017.
6 Recommendations

6.1 Data Availability and Quality

As discussed in Section 3 and Section 4, a robust wind farm forecast analysis requires sufficient turbine-level data (i.e., SCADA data) to identify and then filter or correct periods of abnormal turbine operation. Data availability and quality were considerable issues in this analysis. The SCADA data provided were incomplete and data fields were not well described (in particular the status and fault codes). Furthermore, discrepancies were identified between the sum of turbine power from the SCADA data and measured plant power reported in the forecast data.

Kingdom Community wind farm and Sheffield wind farm turbines showed significant periods of abnormal operation or downtime. Lacking the level of SCADA data required to correct these periods resulted in the exclusion of Kingdom Community wind farm and Sheffield wind farm plants from this forecast assessment.

Georgia Mountain wind farm showed consistent normal turbine operation for most of the period of record and, despite limited SCADA data, a forecast assessment was possible. However, the diurnal analysis was restricted to only 2 months. Furthermore, results from the forecast assessment should be interpreted cautiously given data integrity issues that were identified (e.g., average 50% forecast underestimates from February to March 2017, discrepancies in turbine and plant reported power).

Only forecasts made at 0000 UTC (1900 EST) were provided and only for a look-ahead period of 24 hours. As a result, an analysis of forecast accuracy as it relates to forecast time and look-ahead period was not possible.

Finally, no information was provided on the forecast system, in particular the input data sources, forecast methodologies (e.g., extent of statistical training, use of numerical weather prediction models), or postprocessing of forecast results. Consequently, recommendations could not be made on specifically how to improve the forecast model.

A more robust assessment of forecast system performance would be possible given the following:

- A longer period of record (at least 1 full calendar year) that includes forecasts made at different times of day and with lead times up to 72 hours.
- The addition of wind speed forecasts along with power forecasts. Wind speed forecasts would eliminate the need to filter/correct the power data but would require accurate on-site measurement of wind speeds (e.g., from a meteorological tower). Nacelle wind speeds are generally not reliable because of sensitivities related to the position on the wind turbine, operating conditions (e.g., rotating blades and associated turbulence), and turbine controls.
- Comprehensive SCADA data for each turbine. A full data dump containing all output parameters from the SCADA system (including full event logs) would allow for a more in-depth analysis.
- Detailed description of the forecast system (to the extent that this is possible).
6.2 Improvements to Forecast System

Based on the forecast assessment for Georgia Mountain wind farm, the following potential areas of improvement for the forecast model were identified:

- Strong bias in winter months. Forecast error from February to March 2017 was on average around 5 MW (or 50% of plant capacity). It is likely this bias is the result of bad measured data, but nonetheless should be investigated.

- Better representation of the diurnal cycle. The complex, thermally-driven flow regime at Lake Champlain resulted in a strong diurnal cycle in wind speeds and wind power at Georgia Mountain wind farm that was not well-represented in the forecast model. Better accounting of microscale wind regimes (e.g., through local weather forecasts or historical weather station data) may improve forecast accuracy in this regime and other similarly complex regimes in Vermont.

- Accounting for internal wake loss. Large diurnal shifts in wind direction in the Lake Champlain area may be resulting in diurnal patterns in internal wake loss at Georgia Mountain wind farm and may be contributing to high forecast error. It is unclear if a wake model is built into the forecast system. If not, the incorporation of a wake model is recommended.