



## Validation of Power Output for the WIND Toolkit

J. King  
*RePPAE*

A. Clifton and B.-M. Hodge  
*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](http://www.nrel.gov/publications).

**Technical Report**  
NREL/TP-5D00-61714  
September 2014

Contract No. DE-AC36-08GO28308

# Validation of Power Output for the WIND Toolkit

J. King  
*RePPAE*

A. Clifton and B.-M. Hodge  
*National Renewable Energy Laboratory*

Prepared under Task No. WE11.0705

**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](http://www.nrel.gov/publications).

## 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](http://www.nrel.gov/publications).

Available electronically at <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  
phone: 865.576.8401  
fax: 865.576.5728  
email: <mailto:reports@adonis.osti.gov>

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

U.S. Department of Commerce  
National Technical Information Service  
5285 Port Royal Road  
Springfield, VA 22161  
phone: 800.553.6847  
fax: 703.605.6900  
email: [orders@ntis.fedworld.gov](mailto:orders@ntis.fedworld.gov)  
online ordering: <http://www.ntis.gov/help/ordermethods.aspx>

*Cover Photos: (left to right) photo by Pat Corkery, NREL 16416, photo from SunEdison, NREL 17423, photo by Pat Corkery, NREL 16560, photo by Dennis Schroeder, NREL 17613, photo by Dean Armstrong, NREL 17436, photo by Pat Corkery, NREL 17721.*

## List of Acronyms

ERCOT	Electric Reliability Council of Texas
ERGIS	Eastern Renewable Generation Integration Study
EWITS	Eastern Wind Integration Study
IEC	International Electrotechnical Commission
MISO	Midcontinent Independent System Operator
NREL	National Renewable Energy Laboratory
UCT	Coordinated Universal Time
WIND	Wind Integration National Data set
WRF	Weather Research and Forecasting
WWSIS	Western Wind and Solar Integration Study

# Table of Contents

<b>1</b>	<b>Introduction</b> .....	<b>1</b>
1.1	The WIND Toolkit.....	1
1.2	The Need for Validation .....	2
1.2.1	Validation of the WIND Toolkit Meteorological Data Set.....	2
1.2.2	Validation of the Power Data Set.....	2
1.3	About This Document.....	2
<b>2</b>	<b>Wind Turbine Performance and Locations</b> .....	<b>4</b>
2.1	Wind Turbine Power Curves .....	4
2.2	Wind Turbine Locations .....	6
<b>3</b>	<b>The WIND Toolkit Power Data Set</b> .....	<b>8</b>
<b>4</b>	<b>Power Validation Data Set</b> .....	<b>9</b>
<b>5</b>	<b>Results of the Power Validation</b> .....	<b>11</b>
5.1	Mapping Wind Speed to Power .....	11
5.2	Daily Production Profiles.....	13
5.3	Seasonal and Regional Capacity Factors .....	14
5.4	Variability Analysis .....	19
<b>6</b>	<b>Conclusions</b> .....	<b>26</b>
	<b>References</b> .....	<b>28</b>

## List of Figures

Figure 1. Power curves used to convert modeled wind speeds to power for onshore Class I through Class III locations and offshore locations.....	5
Figure 2. Locations of the 126,692 wind power plant sites .....	6
Figure 3. Map of wind power plants in the validation data set. There are also three plants in Hawaii (not shown).....	10
Figure 4. Aggregate power curve for a sample 6-MW wind power plant with theoretical power curve (red markers) .....	12
Figure 5. Aggregate power curve for a sample 300-MW wind power plant with theoretical power curve (red markers) .....	12
Figure 6. Aggregate power curve for a sample 1,008-MW offshore wind power plant with theoretical power curve (red markers) .....	13
Figure 7. Comparison of average daily production profiles of the WIND Toolkit to the measured data in ERCOT .....	14
Figure 8. Comparison of average daily production profiles of the WIND Toolkit to the measured data in MISO.....	14
Figure 9. Seasonal capacity factors by region .....	15
Figure 10. Seasonal capacity factors for onshore and offshore wind power plants in the Northeast.....	16
Figure 11. Comparison of WIND Toolkit capacity factors to those in previous study data sets.....	16
Figure 12. Comparison of WIND Toolkit capacity factors to available measurement data in ERCOT .....	18
Figure 13. Comparison of WIND Toolkit capacity factors to available measurement data in MISO .....	18
Figure 14. Frequency spectrum for power from a single site .....	20
Figure 15. Normalized variability Fourier transform for a single site .....	20
Figure 16. Frequency spectrum of power for an aggregated plant .....	21
Figure 17. Fourier transform for aggregate variability for 284 plants showing a 15-min peak.....	22
Figure 18. Daily variability for a single wind power plant.....	22
Figure 19. Daily variability for the aggregation of the entire test data set.....	23
Figure 20. Average 5-min variability profile for the Plains states region.....	24
Figure 21. Average 5-min variability and production for ERCOT.....	25

## List of Tables

Table 1. Wind Turbines Used to Create Composite IEC Power Curves .....	4
Table 2. Normalized Power Curves Used in Each Grid Cell .....	5
Table 3. Summary of Plants in the Test Data Set .....	10
Table 4. Seasonal Capacity Factors by Region .....	15
Table 5. Comparison of WIND Toolkit Regional Capacity Factors to Older Study Data Sets.....	17
Table 6. Detailed Data for WIND Toolkit and Measurement Capacity Factors.....	18

# 1 Introduction

The integration of renewable energy into the power system is an ongoing project in the United States. Important factors for the integration of variable generators, such as wind and solar power, include the geographic distribution of the generators, how the power produced at each site changes over time, and how well their production can be forecast. In the case of wind energy, this generation is driven by the variation of the wind over many different timescales; therefore, it is essential to have a clear understanding of both the wind and the turbine performance at spatial and temporal resolutions that are important for balancing authorities and utilities. This need in turn implies spatial resolutions of data from 2 km to multiple states and from a few minutes to several years. This data is not commonly available for integration studies from either public or private sources; instead, current integration studies typically simulate winds or weather over a wide area and then model the response of artificially located turbines to these simulated conditions. This modeling approach provides the resolution, scale, and realistic responses, which are essential to a useful integration study.

Previous integration studies include the Western Wind and Solar Integration Study (WWSIS) (see GE Energy 2010 and 3Tier 2010) and the Eastern Wind Integration Study (EWITS) (see EnerNex Corporation 2011). These studies used high-quality wind data sets to create time series of the power produced by a large number of wind power plants spread across the United States. The WWSIS and EWITS data sets provide realistic representations of the variability, ramping characteristics, and forecast performance of wind power plants that operate using today's technology. WWSIS covered the western United States with a seamless, 1-arc-min resolution meteorological grid at 10-min resolution for the years from 2004 to 2006 (Potter et al. 2008). The EWITS study data set had a resolution of 10 km by 10 km at 5 min for the same time period and covered the eastern United States. Although this spatial resolution was sufficient at the time, experience has shown that realistic simulation of wind integration into the U.S. grid requires data at finer resolution and for a longer period. In addition, these two data sets did not cover the entire continental United States with a consistent modeling approach; therefore, a new meteorological and power data set was required.

The Wind Integration National Data set (WIND) Toolkit, described herein, is an update for and expansion of the WWSIS and EWITS data sets, and it is intended to support the next generation of integration studies.

## 1.1 The WIND Toolkit

The WIND Toolkit includes meteorological conditions and turbine power data for more than 126,000 sites in the continental United States for the years from 2007 to 2013. Reanalysis of weather observations and mesoscale climate data using the Weather Research and Forecasting (WRF) model version 3.4.1 was used to create an underlying meteorological data set (Skamarock et al. 2008). The meteorological data set has a finer spatial (2 km by 2 km) and temporal (5 min) resolution than the existing data sets and covers a longer period of time (2006 to 2013) while modeling the entire continental United States.

As part of the WIND Toolkit project, turbine locations were selected using a 30-y mean annual wind speed data set. Wind turbine power curves were also developed for those locations. The

process of site selection and the power curve development is summarized in this document. For full details, please see *Creating the WIND Toolkit Data Sets* (3Tier forthcoming).

Three main data sets are included in the WIND Toolkit:

- The meteorological data set includes basic information on the weather conditions in each 2-km x 2-km grid cell. The meteorological data set also includes parameters such as wind profiles, atmospheric stability, and solar radiation data in those cells.
- A power data set was created using the wind data and site-appropriate turbine power curves to estimate the power produced at each of the turbine sites. This document focuses on the creation and validation of the power data set.
- A forecast data set includes forecasts for 1-h, 4-h, 6-h, and 24-h forecast horizons.

The WIND Toolkit has been funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Wind and Water Power Technologies Office and was created through the collaborative efforts of the National Renewable Energy Laboratory and 3Tier. The data set is to be made publicly available at no cost to foster the development of the next generation of wind integration studies.

## 1.2 The Need for Validation

Experience with the EWITS and WWSIS data sets showed that careful validation of the data is essential. This validation must be carried out on both the meteorological data set and the power data set. This is needed to avoid modeling artifacts that can create unreasonable data time series because of temporal or geographical seams in the modeling framework.

### 1.2.1 Validation of the WIND Toolkit Meteorological Data Set

An analysis of the meteorological data set that validated the model wind speed characteristics with actual measurements has been carried out for a small number of sites across the continental United States (Draxl et al. forthcoming). Diurnal and seasonal statistical analyses were performed and demonstrated that the model produced reasonable results when compared to actual measurements from validation sites.

### 1.2.2 Validation of the Power Data Set

The EWITS and WWSIS data sets showed that artificially produced spatially and temporally resolved power data sets could be impacted by the underlying meteorological data. For example, problems with merging short-duration meteorological simulations in those studies led to occasional artificial ramps or dips in power as simulations were overlain. Diurnal and seasonal variations in wind speed were sometimes impacted by model effects as well, leading to non-physical variations in power output through the course of the day or year. These effects could be both local and regional. Because these effects are not physical but have important impacts on the power system, it is essential to confirm that the power data set does not contain these artificial artifacts.

## 1.3 About This Document

This document focuses on the creation and validation of the WIND Toolkit power data set.



Section 2 describes the process of defining wind turbine power curves and deciding how many and what type of turbines will be installed in each of the WIND Toolkit's 2-km by 2-km grid cells. Section 3 describes how these sites are combined with the meteorological data to create the power data set. Section 4 describes the creation of a data set that is used in Section 5 to validate the WIND Toolkit power data set. Section 6 describes how the power data set can be used.

## 2 Wind Turbine Performance and Locations

The WIND Toolkit provides power time series data for sites across the United States. The power data are based upon the meteorological data set at wind power plant locations. Each wind power plant has different power curves that were chosen according to the estimated long-term wind conditions at each site. This section describes the process by which different power curves were developed and how wind power plant locations were chosen.

### 2.1 Wind Turbine Power Curves

This study used generic power curves to estimate the power that might be available at different locations across the United States. Generic turbine power curves were developed for turbines on sites with different average annual wind speeds. The turbines are described as Class I, Class II, or Class III, depending on the annual average wind speed following the classes described in the International Electrotechnical Commission (IEC) 61400-1 (IEC 2005). An offshore class has also been developed. The generic power curves were developed for an update to the EWITS data set (Pennock 2012) in preparation for the Eastern Renewable Generation Integration Study (ERGIS).

The generic power curves have cut-in wind speeds, rated power, and rated speeds that are similar to commercially available turbines (Table 1). Each turbine is assumed to have a rated power of 2.0 MW and a hub height of 100 m.

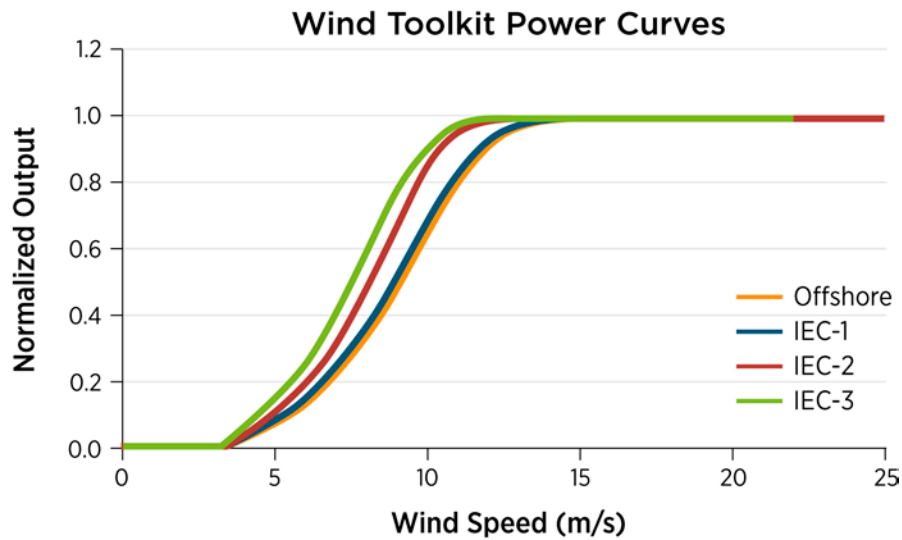
**Table 1. Wind Turbines Used to Create Composite IEC Power Curves**

Class	Turbine	Rated Power (MW)	Cut-In Wind Speed (m/s)	Wind Speed at Rated Power (m/s)	Cut-Out Wind Speed (m/s)
1	Siemens 3.0 MW	3	3	14	25
	Gamesa G80	2	4	17	25
	Nordex N90HS	2.5	4	14	25
	Vestas V90	3	4	14	25
	WIND Toolkit Class 1				25
2	Vestas V112	3	3	13	25
	Siemens 2.3 MW	2.3	3	13	25
	GE1.6-82.5	1.6	4	12	25
	GE2.5XL	2.5	3	14	25
	WIND Toolkit Class 2				25
3	Vestas V100	1.8	3	12	20
	GE1.6-100	1.6	3	12	25
	Repower 3.2 M	3.2	3	12	22
	WIND Toolkit Class 3				22
Offshore	Siemens 3.6 MW	3.6	4	14	25
	GE 4.1 M	4.1	4	14	25
	Repower 6 M	6.15	3.5	14	30
	WIND Toolkit Offshore				

The normalized power curves for the different classes are listed in Table 2 and plotted in Figure 1. A normalized power curve could apply to a single turbine, in which case the normalized power curve is the turbine output power divided by the rated power (2.0 MW for all classes), or to multiple co-located turbines of the same type, in which case the normalized power is the total output divided by the sum of the turbines' rated power.

**Table 2. Normalized Power Curves Used in Each Grid Cell**

Speed Bin	IEC Class 1	IEC Class 2	IEC Class 3	Offshore
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0.0043	0.0052	0.0054	0
4	0.0323	0.0423	0.053	0.0281
5	0.0771	0.1031	0.1351	0.074
6	0.1426	0.1909	0.2508	0.1373
7	0.2329	0.3127	0.4033	0.2266
8	0.3528	0.4731	0.5952	0.3443
9	0.5024	0.6693	0.7849	0.4908
10	0.6732	0.8554	0.9178	0.6623
11	0.8287	0.9641	0.9796	0.815
12	0.9264	0.9942	1	0.9179
13	0.9774	0.9994	1	0.9798
14	0.9946	1	1	1
15	0.999	1	1	1
16	0.9999	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	0	1
24	1	1	0	1
25	1	1	0	1



**Figure 1. Power curves used to convert modeled wind speeds to power for onshore Class I through Class III locations and offshore locations**

## 2.2 Wind Turbine Locations

Grid cells that would be populated with wind turbines were selected on the basis of geography and wind resource. The site selection process is detailed in *Creating the WIND Toolkit Data Sets* (3Tier forthcoming) and summarized here for completeness.

First, a 30-y data set from 3Tier was used to estimate the mean annual wind speed in 2-km by 2-km grid cells across the United States. Then the power curves described in the previous section were assigned to cells based on the long-term average wind speed in that cell according to the IEC class definitions. It was assumed that at most, eight 2-MW turbines, all of the same class, could be present in each grid cell. Next, the area of buildable land in each cell (accounting for buildable land and exclusions) was used to estimate how much of the resource in each cell could be harvested. Combining the wind speed distribution, power curve information, and available area gives an estimate of the energy that can be produced in each grid cell.

Having identified the energy potential in each cell, the next step was to identify likely wind power plant locations. To do this, the cells were ranked by energy production, and the 110,000 best-performing cells were chosen. Grid cells that corresponded to built sites and sites that were included in WWISS and EWITS were selected. Then the best-performing cell in every county in the lower 48 states was selected. Offshore cells at least 8 km from shore in water less than 30 m deep were also chosen. This process identified a total of 112,471 onshore cells that would be populated with one or more wind turbines and 14,221 offshore grid cells (Figure 2). The locations of these cells and the number and type of turbines in each cell were used to create the power data set.



**Figure 2. Locations of the 126,692 wind power plant sites**

This siting process ensured that cells with wind turbines are dispersed across the United States and that the turbines are sited in areas with a potentially viable wind resource. Although not all of the wind turbines might be in locations that would be considered commercially viable, the locations represent a state of high penetration of wind generation in the continental United States.

### 3 The WIND Toolkit Power Data Set

After the locations of the wind power plants were identified using the low-temporal resolution data, it was possible to create the high-temporal resolution power production data set.

The power production data set was created using wind data from the meteorological data set, and the power curves for each of the more than 126,000 wind power plant sites were identified. The process for estimating the power production was as follows:

1. Remove the bias from the wind speeds in the high-resolution WIND Toolkit meteorological data set. The adjustments to wind speed for bias and wake losses are discussed in detail in the final report for this project (3Tier forthcoming).
2. Reduce the wind speeds in each grid cell from the meteorological data set value ( $U_{Met}$ ) to account for the effects of wakes. This was done by 3Tier using an empirical function:

$$C_{wake} = 1 - \frac{1}{20} \left( \frac{n_{turbines} - 1}{7} \right)$$

where  $U_{cell} = C_{wake} \times U_{Met}$ , and  $n_{turbines}$  is the number of turbines in each grid cell. Because the maximum number of turbines per grid cell is eight, the wind speed reduction in each cell due to wakes to is limited to 5%.

3. Apply the wind power plant site-specific power curves for each site to estimate the power for each 5-min data point.

This process allows the 5-min meteorological data set to be converted into a power production time series of 5-min temporal resolution for each of the 126,000-plus sites. Similarly, the forecast data was converted into simulated operational forecasts for 1-h, 4-h, 6-h, and day-ahead forecast horizons at the 126,000-plus sites selected. No statistical adjustments were made to the power data at the time of this report.

The power data set also includes wind speed, air density, and wind direction at 100 m, 2 m temperature, and surface pressure.

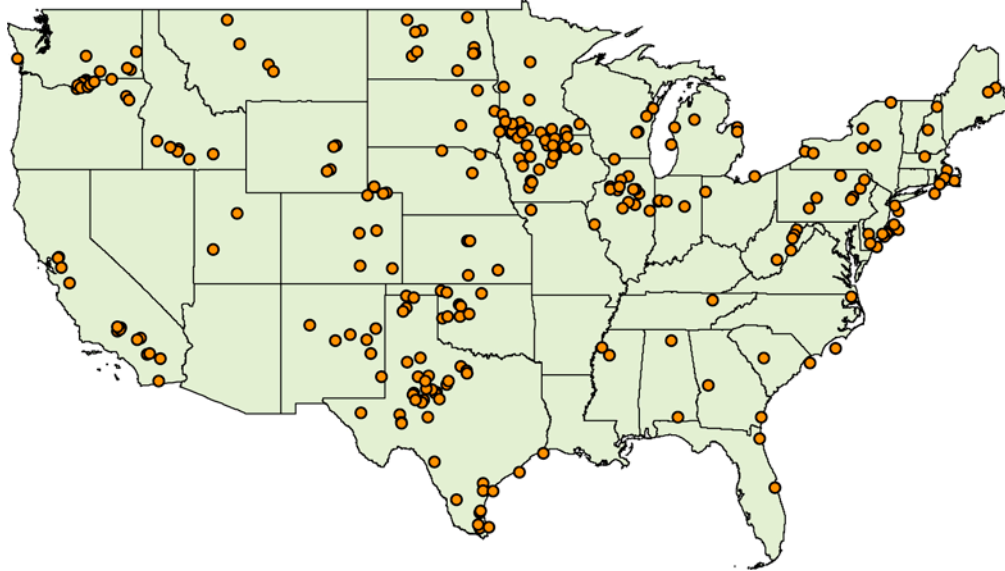
## 4 Power Validation Data Set

Very little production data from actual wind power plants is available in the public domain, which makes comparing modeled sites to actual sites very difficult. These actual data are rarely made public because of the commercial nature of the data. Broad aggregate data for several market areas are available for comparison only because the Midcontinent Independent System Operator (MISO) and the Electric Reliability Council of Texas (ERCOT) both publish wind production data. Therefore, it was necessary to create a validation data set that can be used to compare the WIND Toolkit to the limited data that is available.

A validation data set was developed by 3Tier with input from NREL to provide a representative sample of sites across the country. This validation data set was developed with a preference for locations with existing or planned wind power plants. Additionally, the validation data set was selected to have representation from most of the lower 48 states, including sites in the Southeast. The data set consists of 284 *plants* across the model footprint. These wind power plants are aggregates of individual sites of varying sizes. A plant could be as small as a single grid site from the base data set or tens of sites with the power curves of each combined into a single production profile. Some of these plants are meant to represent existing wind power facilities; others are hypothetical plants useful for the purpose of statistical analysis. As part of this project, a tool is provided that automates the process of building these plants; this tool was used to build the test data set.

When a wind power plant included different grid cells and different turbines, a composite wind speed and a composite power time series were created. The composite speed is the weighted average of the wind speeds at each site, with the weights proportional to the nameplate capacity of each site. The power is a simple summation of the expected power output for each site in the plant.

The validation data set consists of 5-min wind and power data for the 284 sites from 2007 through 2012. Figure 3 shows the location of the test plants. Of the 284 sites, 24 are offshore. Validation sites are frequently found in the upper Midwest and Texas, because these are areas with highly developed wind resources. In addition to the validation plants shown in Figure 3, three plants are located in Hawaii.



**Figure 3. Map of wind power plants in the validation data set. There are also three plants in Hawaii (not shown).**

The Northeast is dominated by offshore plants near New Jersey, Connecticut, Massachusetts, and in Lake Erie. The Plains states have a concentration of plants in Iowa and southern Minnesota. There is a relatively small representation in the Southeast states, where the wind resources are the poorest. More information about the test data set plant information is provided in Table 3.

**Table 3. Summary of Plants in the Test Data Set**

Region	Onshore				Offshore			
	Plants	Min Size (MW)	Max Size (MW)	Total Capacity (MW)	Plants	Min Size (MW)	Max Size (MW)	Total Capacity (MW)
<b>Northeast</b>	35	6	304	3,164	17	30	1,104	8,626
<b>Northwest</b>	38	16	512	5,002	1	16	16	16
<b>Plains</b>	89	12	512	10,392				
<b>Southeast</b>	11	16	320	838	2	80	608	688
<b>Southwest</b>	37	8	270	3,954				
<b>Texas</b>	47	16	418	7,640	4	16	1,008	1,840
<b>Hawaii</b>	3	32	410	522	17	30	1,104	8,626



## 5 Results of the Power Validation

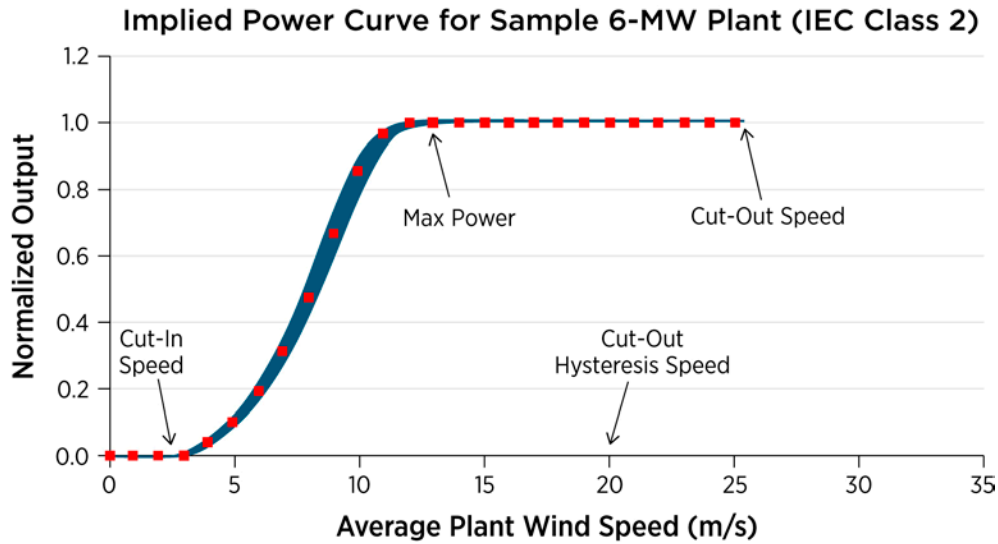
The analysis reported here deals with the broad parameters of the power data. We examined the conversion of wind speed to power and observed the plant power curves for some of the plants in the test data set. We then computed and compared capacity factor information across broad regions and compared the regional capacity factors to previous data sets. Finally, we examined the variability of the power data in close detail to look for modeling anomalies and artifacts that were observed in the EWITS and WWSID wind data sets (Pennock 2012).

### 5.1 Mapping Wind Speed to Power

Analysis during the creation of the power data set showed that there were some problems in the algorithms dealing with the conversion of wind speed to power when certain anomalous conditions were encountered in the raw database. The problems were identified and corrected, and it is assumed that the conversion of wind speed to power is correct across all of the sites in the 126,000-plus domain.

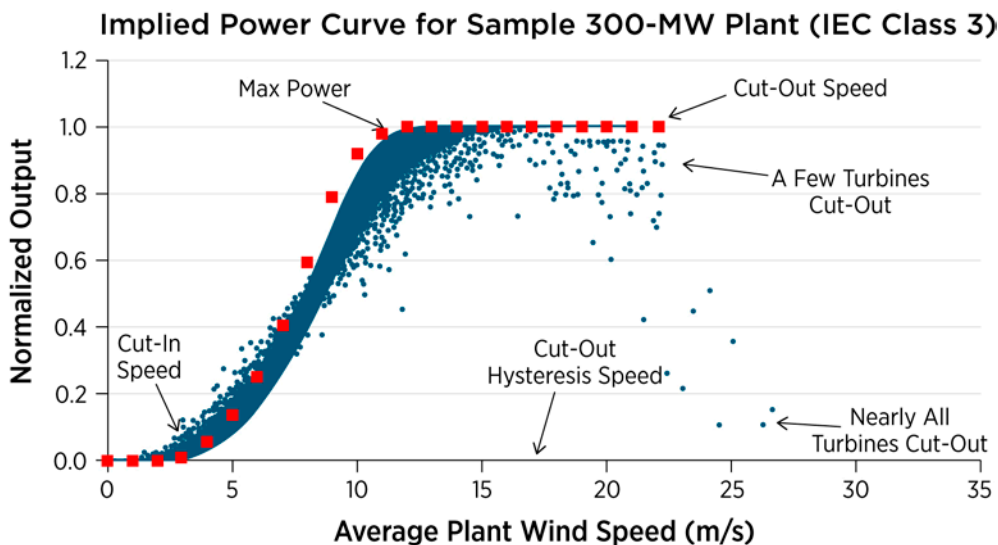
As mentioned above, the data set available for this validation does not include individual grid cell sites but rather aggregations of those sites. This made it difficult to verify the implementation of the power curves presented in the introduction. However, it was interesting to examine the power curve that was implied for the aggregated sites. This was done by plotting each average wind speed and power pair for a plant and examining the cloud of points that resulted. The wind speed is the weighted average speed experienced at each time point by each of the turbines that comprise the plant. The power is the aggregate of all of the turbines.

Figure 4 shows the implied power curve for a small 6-MW plant in the Southeast. This plant was classified as IEC Class 2. All 672,768 data points for the plant in the simulated years from 2007 through 2012 are represented on the graph. Cut-in speed is 3 m/s with a maximum power output of approximately 13 m/s. The cut-out speed is 25 m/s, and a hysteresis was applied so that after a turbine has shut down as a result of a cut-out, the wind speed must drop to 20 m/s before that turbine is permitted to restart. Each of these points can be clearly identified in the plot. There is a dispersion of points in the nonlinear portion of the power curve. This is because each of the turbines experienced slightly different wind speeds. The theoretical power curve is superimposed on the data in square red markers.



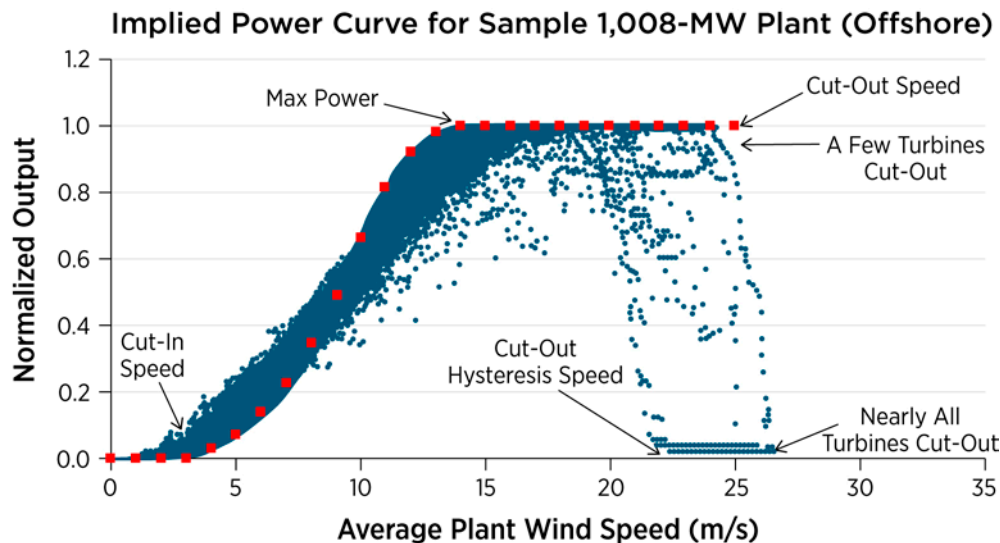
**Figure 4. Aggregate power curve for a sample 6-MW wind power plant with theoretical power curve (red markers)**

As the plant size and the number of turbines comprising the plant increases, the diversity among the turbines increases and the aggregate power curve becomes more diffuse. Figure 5 shows the power curve for a 300-MW plant. Again, the speed axis is the weighted average speed throughout the plant. The cut-in speed is no longer a clear point but an area where some turbines had started and others had not. The high wind cut-out shows that some plants continued to operate as the average wind speed exceeded the cut-out speed, because some turbines in the plant were experiencing lower wind speeds. Note that there is some discrepancy between the theoretical power curve (red dots) and the WIND Toolkit data near the top of the power curve. The differences in the results shown between the theoretical power curve and the WIND Toolkit data are because not all of the sites in this sample wind power plant that were modeled with the Class 3 turbine power curve are shown in the figure.



**Figure 5. Aggregate power curve for a sample 300-MW wind power plant with theoretical power curve (red markers)**

Figure 6 shows the implied plant power curve for a 1,008-MW offshore wind power plant. Many individual turbines comprise this plant, which covers a significant geographic area, approximately 250 km<sup>2</sup>. There was significant diversity in the wind across the plant, which led to a broad power curve. Again, each turbine in the plant experienced different wind speeds.



**Figure 6. Aggregate power curve for a sample 1,008-MW offshore wind power plant with theoretical power curve (red markers)**

## 5.2 Daily Production Profiles

Two limited data sets are available to compare production on a regional basis: those from ERCOT and MISO, both of which publish the wind production information about their systems. ERCOT provides 5-min data, and MISO provides 1-h data.

Figure 7 and Figure 8 show the comparison of the average daily production profiles of the WIND Toolkit data to the measurements for ERCOT and MISO, respectively.

The shapes of the profiles for ERCOT as shown in Figure 7 correlate well but suggest that the WIND Toolkit calculated significantly higher power output for the fleet of plants. The difference in production is explored in the next section, on capacity factor, but it is influenced by more modern turbine power curves and the assumption that all turbines have a hub height of 100 m. Also, some curtailment is included in the ERCOT measurements that would not be reflected in the Toolkit data.

There is less agreement in the shapes of the profiles for MISO, in which the measured data is much flatter than the WIND Toolkit data. Again, the measured data shows lower production than the WIND Toolkit data.

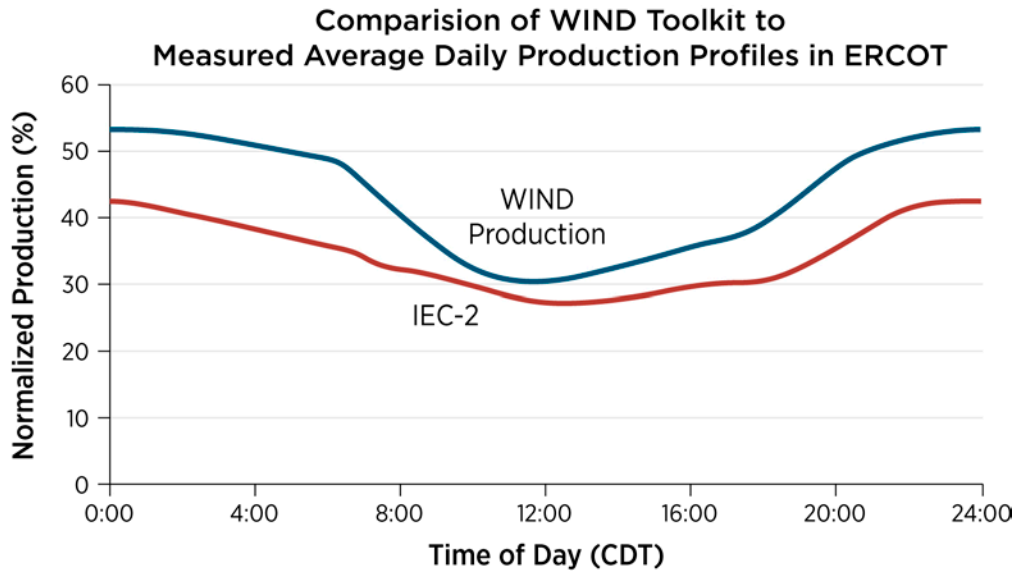


Figure 7. Comparison of average daily production profiles of the WIND Toolkit to the measured data in ERCOT

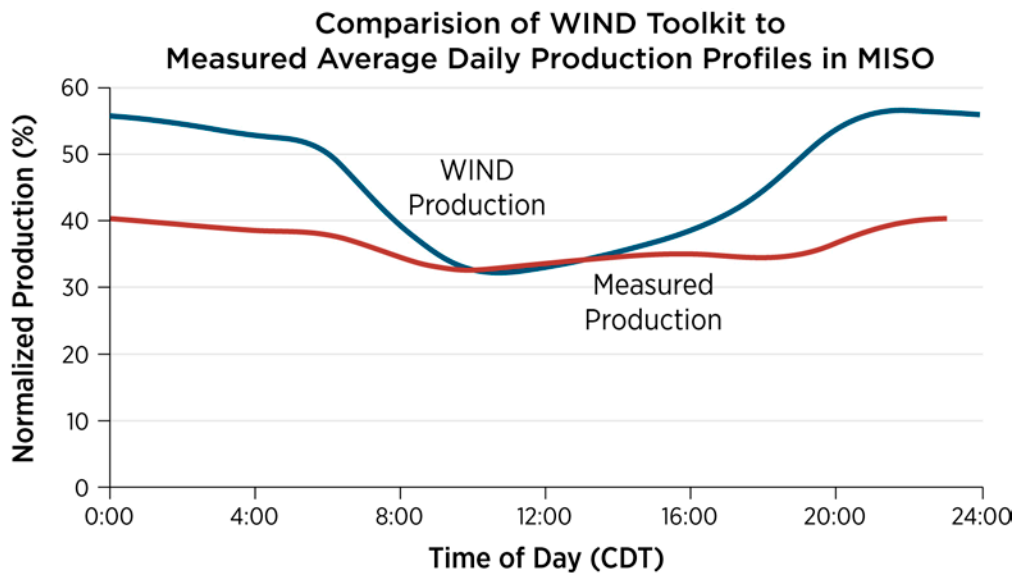


Figure 8. Comparison of average daily production profiles of the WIND Toolkit to the measured data in MISO

### 5.3 Seasonal and Regional Capacity Factors

Capacity factor is a significant characteristic of synthesized wind power data. It measures the amount of energy that a site produces over time compared to the capacity of that facility. Some earlier data sets have been criticized for low capacity factors in areas where plants exist, but only anecdotal evidence has been offered to support this. Little rigorous analysis has been possible because of the lack of measured data from existing facilities. However, meteorological validation suggests good agreement of this data set to the measured data for a number of validation sites across the country. (See meteorological validation documentation.)

The average seasonal and annual capacity factors for various regions in the United States are presented in Figure 9 and Table 4. Winter and spring present a general pattern as being the seasons of highest production in the central and eastern parts of the country. In the West, production is lowest in the winter and highest in the spring and summer.

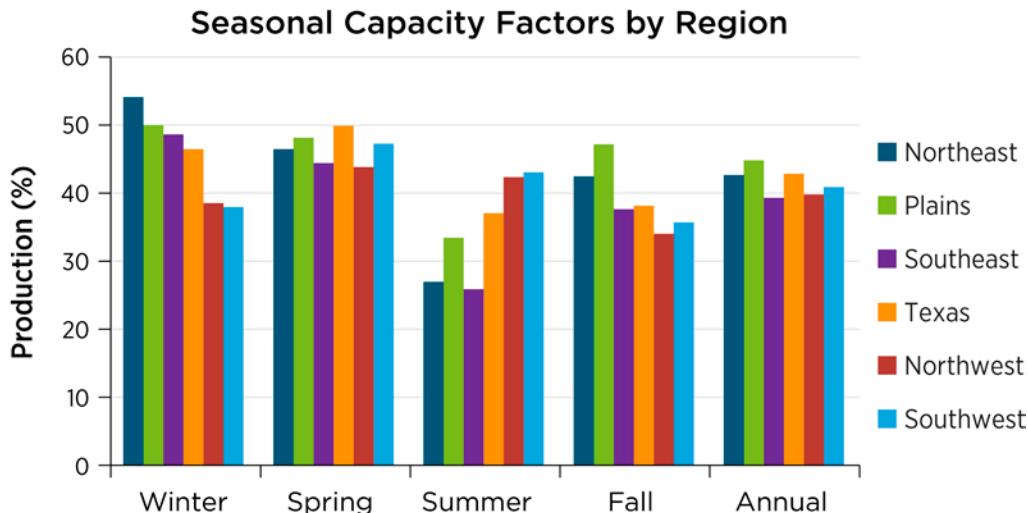


Figure 9. Seasonal capacity factors by region

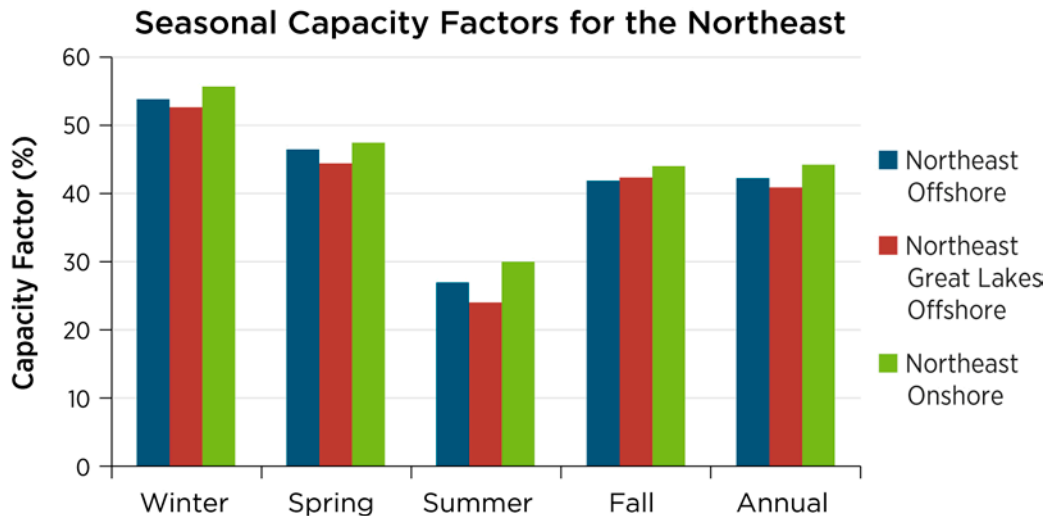
Table 4. Seasonal Capacity Factors by Region

Season	Northeast	Northwest	Plains	Southeast	Southwest	Texas
Winter	54%	39%	50%	49%	38%	47%
Spring	47%	44%	48%	45%	47%	50%
Summer	27%	42%	33%	26%	43%	37%
Fall	43%	34%	47%	38%	36%	38%
Annual	43%	40%	45%	39%	41%	43%

The high capacity factors in the Southeast are somewhat surprising, but they can be explained by the small number of sites and by the fact that those are the best sites found in the Southeast. Only 13 plants in this region, with a total capacity of slightly more than 1,500 MW. This includes 700 MW of offshore capacity with a capacity factor of more than 40%.

The Northeast region is dominated by offshore sites both off the Atlantic Coast and in Lake Erie. A comparison was done within the northeastern region to further understand the onshore and offshore capacity factors. Figure 10 shows this comparison.

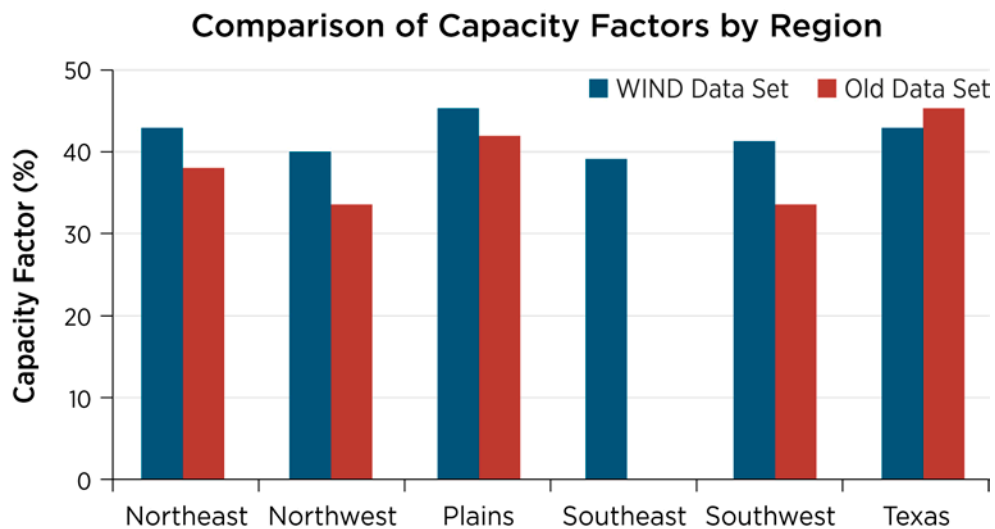
During all seasons, the capacity factor was higher for onshore wind than it was for offshore wind. This was a surprising result, because earlier studies have found that offshore wind has a higher capacity factor than onshore wind. For instance, in the NREL eastern data set (AWS Truepower 2012), the average onshore capacity factor for the Northeast was approximately 37%; whereas the average offshore capacity factor was 42%.



**Figure 10. Seasonal capacity factors for onshore and offshore wind power plants in the Northeast**

To further explore the comparison of the WIND Toolkit to other data sets, regional information was extracted from the EWITS and WWSIS wind data sets. An attempt was made to line up the 284 sites from the test data set to the sites from the older data sets to make the comparison as even as possible. The aggregate capacity factors were then computed for each of the regions shown in Figure 9. This was not possible in the Southeast, because the EWITS data set did not develop any data in this region. Also, note that the EWITS sites in Texas are concentrated in the non-ERCOT area of the panhandle of the state; whereas the Wind Toolkit covers the entire state with a large number of sites in both West Texas and the Gulf Coast.

Figure 11 shows the results of this comparison. In all regions except Texas, the WIND Toolkit sites showed significantly higher capacity factors than those in the EWITS or WWSIS sites. The difference was greatest in the Southwest, but it was also large in the Northwest. Detailed data for this chart are given in Table 5. It was assumed that the differences in Texas were because of the differing wind regimes in the WIND Toolkit and EWITS data sets, as described above.



**Figure 11. Comparison of WIND Toolkit capacity factors to those in previous study data sets**

**Table 5. Comparison of WIND Toolkit Regional Capacity Factors to Older Study Data Sets**

	<b>WIND Data Set</b>	<b>Old Data Set</b>
<b>Northeast</b>	43%	38%
<b>Northwest</b>	40%	34%
<b>Plains</b>	45%	42%
<b>Southeast</b>	39%	
<b>Southwest</b>	41%	33%
<b>Texas</b>	43%	45%

Two market regions in the United States publish historic wind data measurements: MISO and ERCOT. The data about these two system operators was obtained from their respective websites<sup>1</sup> and analyzed with information about the installed capacities of each system so that capacity factors could be calculated. Because of apparent data quality problems, we used ERCOT data from April 2012 through January 2014, which has eight months of overlap with the WIND Toolkit data. The MISO data were available for the calendar year 2012. These data sets were compiled from the websites by 3Tier.

Seasonal and annual capacity factors were calculated and averaged over the appropriate intervals. Figure 12 and Figure 13 show the comparisons of the regional values to the WIND Toolkit values. The WIND Toolkit values were approximately 8 to 10 percentage points higher in most seasons for both regions. This is a slightly larger difference than that shown in the comparisons of the WIND Toolkit data to older data sets. The additional differences can be easily explained by availability issues (curtailment, scheduled and unscheduled maintenance) and thus suggests fairly good agreement between the WIND Toolkit data set, older data sets, and measured values for MISO and ERCOT.

A number of factors contribute to the difference in capacity factors. First, the turbine power curves that were used to model the production are the most modern and efficient on the market compared to the variety of older turbines in the field. Second, the WIND Toolkit models all turbines at a hub height of 100 m, which is an aggressive assumption compared to the existing fleet. Analysis of production data has shown that the difference in production between turbines with a hub height of 80 m compared to those with a hub height of 100 m can be as much as six percentage points of capacity factor for some turbine classes. Also, because only total capacity and not available capacity are included in the ERCOT and MISO totals, it is expected that the theoretical 100% availability assumption in the WIND Toolkit data would give a higher capacity factor. Yet another factor is that electrical losses are not included in the WIND Toolkit production data. Taken together, these factors seem to explain the capacity factor differences in the WIND Toolkit data compared to the measured data.

---

<sup>1</sup> <https://www.misoenergy.org/> and <http://www.ercot.com/>

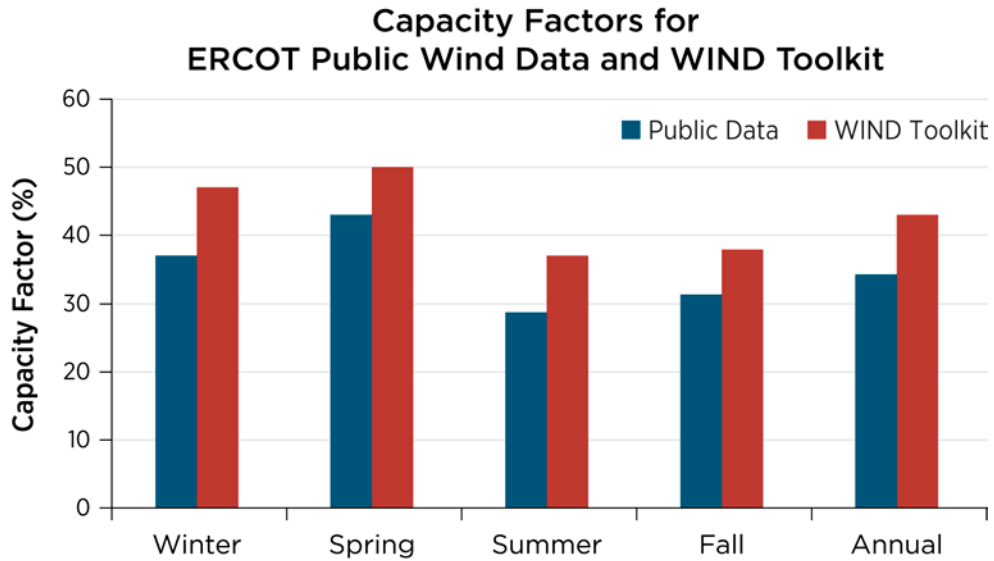


Figure 12. Comparison of WIND Toolkit capacity factors to available measurement data in ERCOT

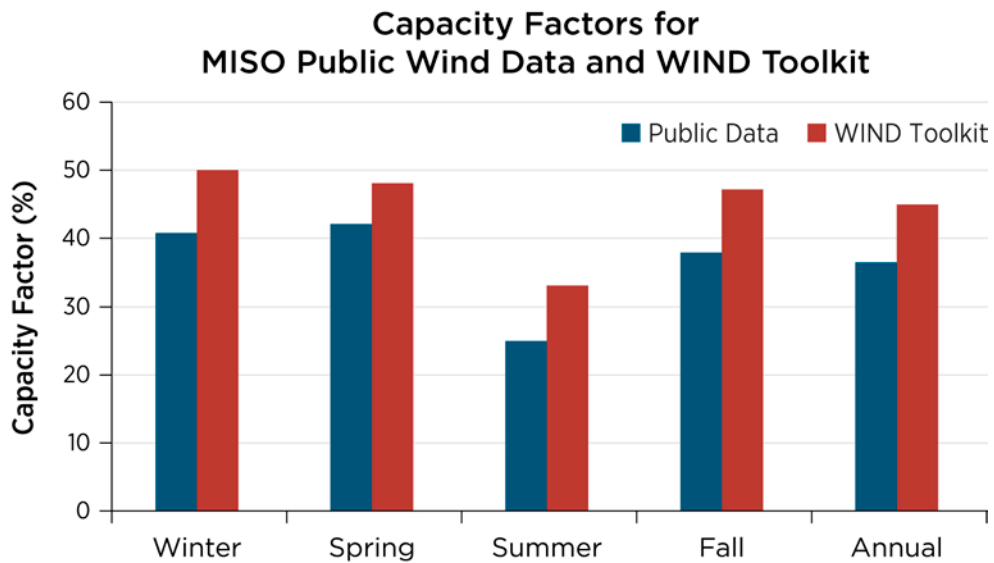


Figure 13. Comparison of WIND Toolkit capacity factors to available measurement data in MISO

Table 6. Detailed Data for WIND Toolkit and Measurement Capacity Factors

Season	ERCOT		MISO	
	Public Data	WIND Toolkit	Public Data	WIND Toolkit
Winter	37%	47%	41%	50%
Spring	43%	50%	42%	48%
Summer	29%	37%	25%	33%
Fall	31%	38%	38%	47%
Annual	34%	43%	36%	45%



## 5.4 Variability Analysis

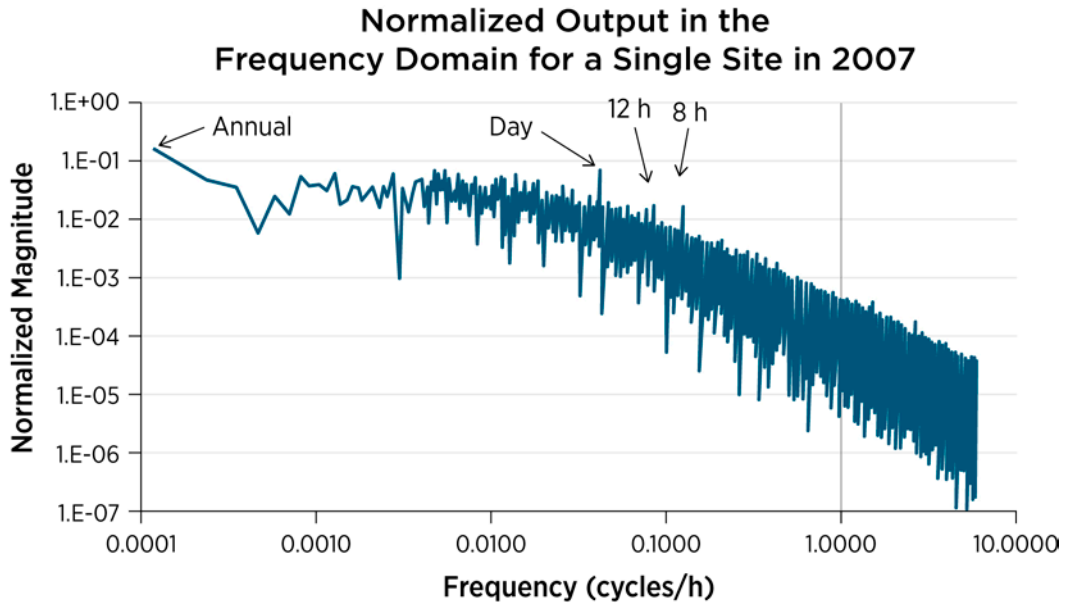
Previous efforts at producing synthesized wind production data (WWSIS, EWITS) succeeded in producing what seemed to be reasonable data; however, when the variability of that data was scrutinized, serious deficiencies were apparent. Most frequently, there were periodic spikes in variability at model geographic or temporal seams or at times when new measurement data were fed into the model. This manifested as unrealistic ramps at and around those times.

Special care was taken when designing the model for this study. First, there are no geographic seams; the model is one grid across the continental United States. Next, the model was run for one month at a time with a full day of overlap at the beginning and end of each month. This minimized the number of temporal seams while allowing ample overlap for smooth blending at those seams.

In previous efforts, variability spikes were not found when analyzing a single site or an aggregation of a small number of sites. As the number of sites included in the aggregation increased, the patterns became more pronounced. This is because the model artifacts were correlated; whereas the more natural variability was not correlated across broad areas.

Spectrum analysis was done with both the power output and on the variability to understand what periods of time would be of interest in looking for repeating phenomena in the data set. Figure 14 shows the Fourier transform for a single site from the test data set. This particular site is a 500-MW plant, but the plot was normalized so the size is not material. Note that constant components (zero frequency) were not included in the plots. The axes were plotted as log-log coordinates to show detail at shorter time frames and lower magnitudes.

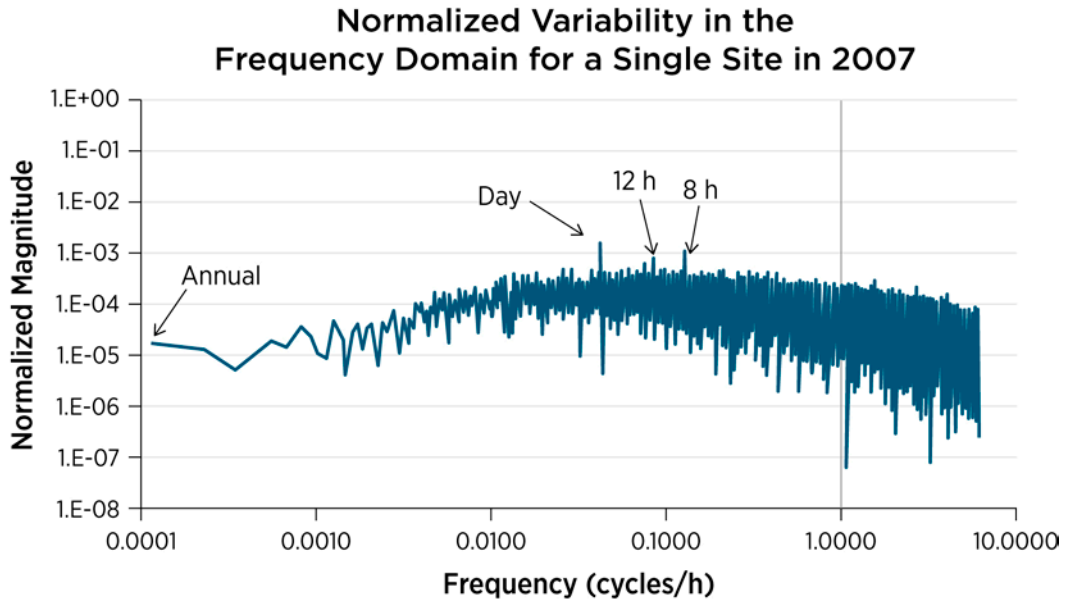
The labels on the plot show the significant peaks. These peaks suggest repeating behavior at those frequencies. When looking at the single site, there are very few peaks. The daily peak suggests that there is a diurnal pattern in the data. That is a natural outcome, because most areas do have diurnal patterns. The next peaks, at 12 h and 8 h, are not as easily explained. It is likely that these times correspond to times when the numerical model had significant events, such as feed-in of measurement data. However, these peaks are relatively small, and as long as they are not correlated across sites they should not become any more significant as more sites are aggregated.



**Figure 14. Frequency spectrum for power from a single site**

We also examined the variability associated with the power shown in Figure 14 by obtaining the interval-to-interval changes and performing a similar Fourier analysis. Figure 15 shows these results for the same plant as that shown in the previous plot.

Again, peaks at 1 d, 12 h, and 8 h were present. There was also a sharp but significant decline in variability at 1 h, which seemed to be an artifact of this plant. This decline did not appear in the larger aggregation, which suggests that this was not correlated across multiple plants.



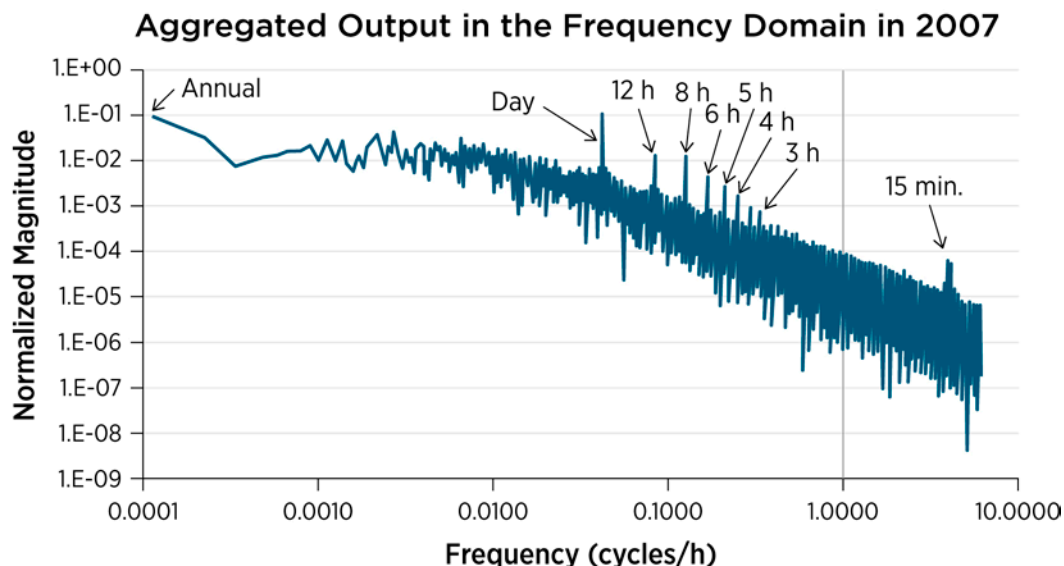
**Figure 15. Normalized variability Fourier transform for a single site**

The same analysis was repeated with the aggregated output of all 284 plants in the sample set to look for correlated phenomena. Figure 16 shows the normalized Fourier output for the aggregated power. Note that the constant component was left off of the plot.

Note that a number of new peaks appeared. There are two reasons that these peaks appeared in the aggregated output. First, the uncorrelated components were reduced in relative magnitude compared to the single plant case from above. This allowed the correlated components to be more prominent. Second, and more importantly, because they are correlated, the components increased with aggregation, which made them bigger in absolute magnitude.

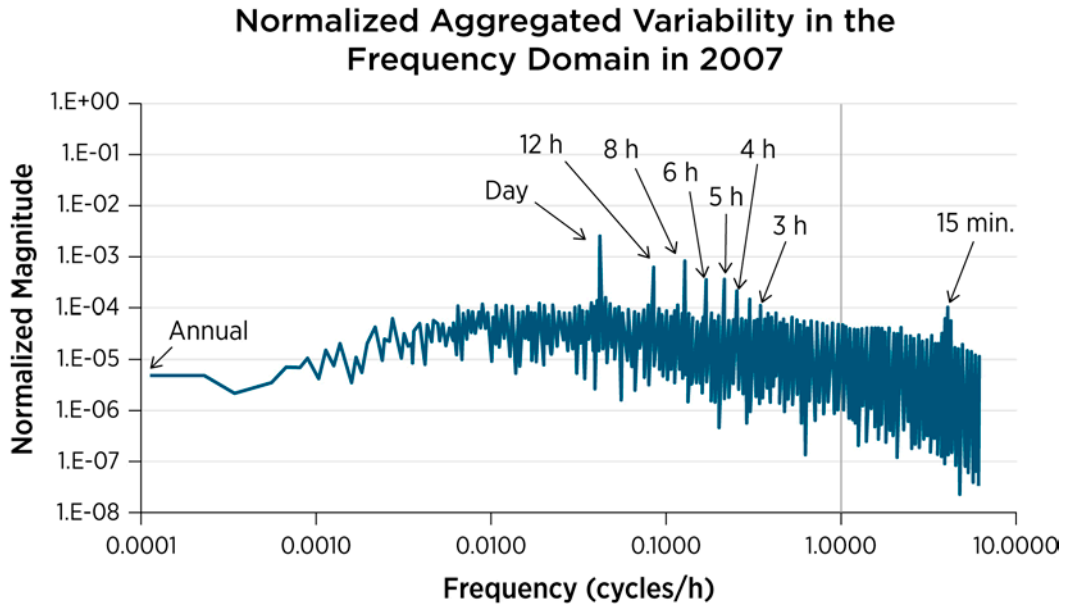
For the most part, the peaks that occurred at 3 h, 4 h, 5 h, 6 h, 8 h, and 12 h appeared to be related to the modeling, because natural processes would not normally repeat on the hour. It is odd that there was a peak at 5 h, because there were reportedly no processes within the model that were repeated at this interval. It is assumed that this peak was an artifact of resampling for the Fourier analysis.

A particularly odd peak appeared at approximately 15 min. Consultation with the scientists at 3Tier determined this peak to be an artifact of the model operation. It was explained that irradiance was forced into the model every 15 min, which caused a step change in the output at those times. This phenomenon will be explored in more detail in the time domain in upcoming discussions.



**Figure 16. Frequency spectrum of power for an aggregated plant**

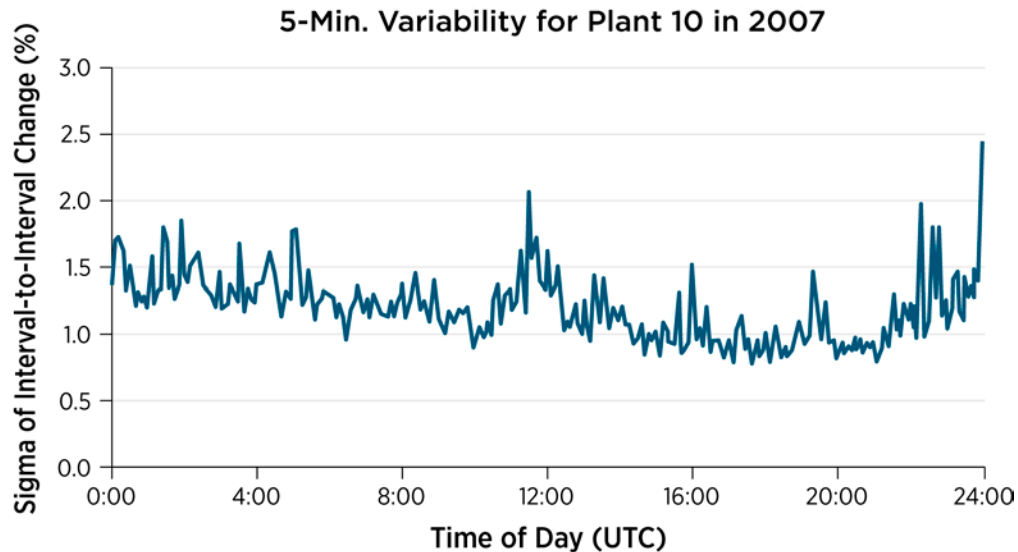
Figure 17 shows the Fourier analysis of the variability for the aggregated data. The peaks are similar to those in power. Note that the dip at 1 h in the single-site analysis is no longer apparent in the data.



**Figure 17. Fourier transform for aggregate variability for 284 plants showing a 15-min peak**

In the analysis above, we found no significant peaks above the period of 1 d. To further explore the peaks shown, we analyzed the data in the time domain. To do this, the standard deviations of the 5-min changes in production were calculated for each 5-min interval of the day throughout all of the years of data in the WIND Toolkit.

First, we observed the variability for the single 500-MW plant used in the frequency domain analysis above. Figure 18 shows these results. Note that the time axis is labeled in Coordinated Universal Time (UTC). This was done because the model was run using UTC, with various processes coordinated to that time standard. Variability peaked at noon and again at midnight, showing the 12-h cycle. This is an important feature, as will be demonstrated. The 8-h cycle was much harder to identify in the plot. There appeared to be a pattern in the afternoon that was close to a 4-h cycle.

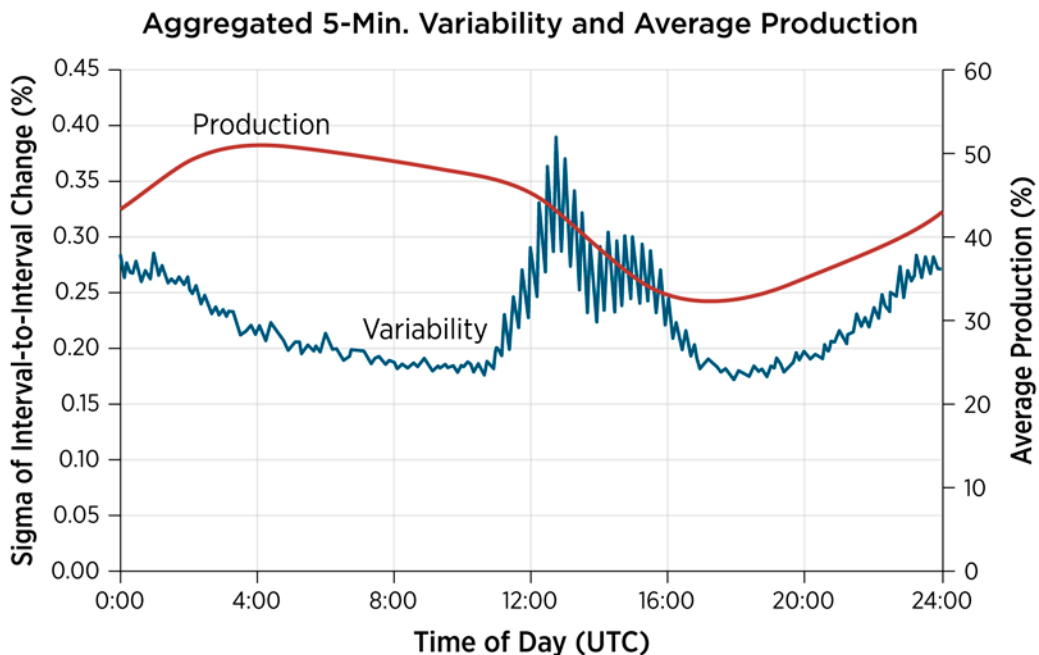


**Figure 18. Daily variability for a single wind power plant**

The average 5-min variability was calculated for the aggregate output of the 284 sample sites. Figure 19 shows the resulting data. The 15-min signal in the frequency domain when the plants were aggregated was very prominent. It was not present during the nighttime hours, but it was highly prominent during the morning hours when the solar radiation increased at its most rapid rate. Although this is clearly an artifact of the modeling, the magnitude was relatively small (.05% of rating) and should not affect the performance of the analysis using the power series. However, caution should be used in calculating 5-min ramp statistics, because this will have a small but definite effect on the results. If time averaging is applied and the data are used at a 30-min resolution or greater, this artifact does not affect the results at all.

The 12-h cycle was again clear, with peaks near noon and midnight. Note that noon in UTC is 7 a.m. Eastern Daylight Time, or approximately average sunrise in the East. The noon peak extended for 4 h or so, possibly reflecting sunrise across the footprint of the wind power plants from east to west. This would suggest a relationship of this peak to sunrise, and the behavior was correlated throughout the plants. It has been suggested that this is because of the meteorological process related to boundary layer physics surrounding dawn.

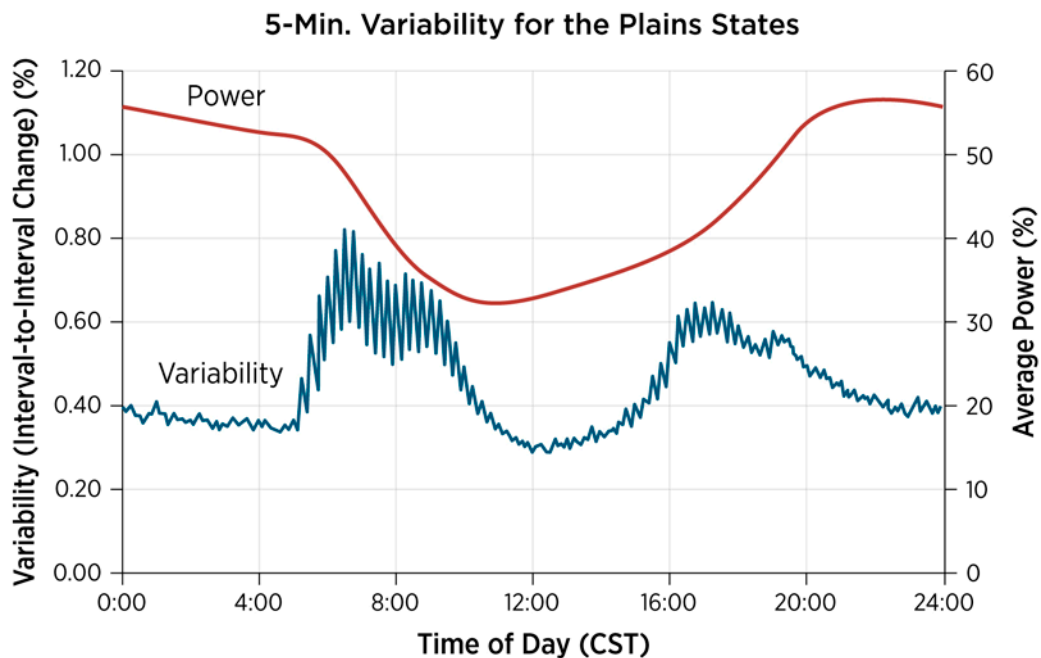
This also corresponded to the steepest decline in wind power production, as shown in the plot. This suggests that the effect is simply because of the average ramping occurring at the same time, seemingly a side effect of the meteorological processes. However, a similar but opposite ramp occurred slightly after midnight UTC that had a corresponding steep decrease in variability. This decrease in variability continued as production ramped down from 4:00 UTC and 11:00 UTC. This suggests that the production ramp rate is not necessarily the only predictor of variability.



**Figure 19. Daily variability for the aggregation of the entire test data set**

To further investigate the prominent peak near noon UTC, a subset of plants was selected to narrow the effect of sunrise time on the results. The Plains region was selected because of the significant capacity installed there, and because it spans only one time zone and is relatively compact from north to south. This should have led to a minimal range of times for sunrise and sunset throughout the region and during yearly intervals.

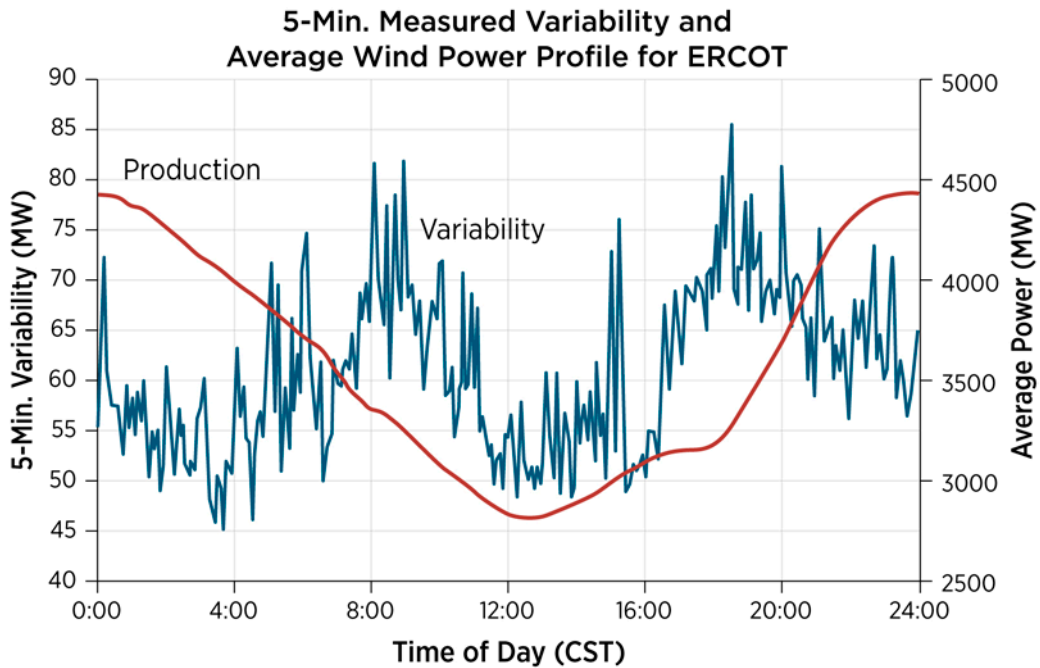
Figure 20 shows the aggregate variability and average production for the Plains states region. For comparison to Figure 19, note that the time axis is given in Central Standard Time (local time to the region) and that the time difference from UTC is 6 h, making noon UTC equal to 06:00 CST. Again, the 15-min artifact was clearly shown in the variability. In this case, however, the variability tended to follow the magnitude of the slope of the average production fairly closely.



**Figure 20. Average 5-min variability profile for the Plains states region**

The publicly available measured data from ERCOT provided a very useful comparison of the resolution and duration to the WIND Toolkit in this analysis. The data are provided at a 5-min resolution for 21 months (May 2012 through January 2014) for the aggregate wind in the ERCOT region.

Figure 21 shows the results of the variability analysis for the measured ERCOT data. In this case, we saw a similar double hump in the variability: one hump near 8:00 CST and the other at approximately 19:00 CST.



**Figure 21. Average 5-min variability and production for ERCOT**

Although the variability from the measured ERCOT data was considerably more volatile, the general shape did support the theory that the highest variability was at approximately sunrise and sunset. This helped validate the notion that the broad atmospheric processes are driving the double hump in variability and that that phenomenon is highly correlated across broad geographic areas.

## 6 Conclusions

Meteorological validation of the WIND Toolkit data sets showed that the meteorological parameters including wind speed at hub height are reasonable (Draxl et al. forthcoming). In previous data sets (EWITS, WWSIS), modeling artifacts and other data flaws have been identified by examining the data after it has been converted to power through the turbine power curves. In this report, we presented analysis of the power data.

The power data were shown to follow the theoretical turbine power curves. The implied plant power curves matched very well to the turbine curves for the Class 2 and offshore turbines. Some discrepancies were found for Class 3 turbines. Those discrepancies could be because of a mixture of turbine types in the plant model.

Capacity factors were found to be universally higher than both earlier simulated data sets and the limited measurement data available to this study. It was suggested that the differences were because of the use of advanced turbine power curves and the modeling assumption that all turbines in the WIND Toolkit are at a hub height of 100 m.

The WIND Toolkit data set exhibits some numerical artifacts, but those are deemed small relative to broader effects for the subset of data examined. Those artifacts are a direct result of the modeling techniques utilized and have been investigated by the researchers at 3Tier. Some of the artifacts are a result of the periodic introduction of reanalysis data during the simulations (using grid nudging with ERA Interim data). It was found that another was a result of the actual processes in the modeling, which caused a 15-min oscillating signal that was particularly apparent during the morning and evening sun ramps.

Variability in the data set was shown to have a double-peak pattern when a significant number of sites were aggregated together. These peaks tended to occur approximately in sync with the morning down- and evening up-ramps that were found in the average daily production profile. The ERCOT measurement data set was analyzed, and it also demonstrated this behavior.

The higher variability during predawn hours was found to be partially the result of average production down-ramps across all sites in the morning and slower up-ramps in the afternoon. The breadth of the variability hump seemed to be related to the geographic spread of the wind sites and the averaging period used in the analysis. This was deemed a real effect, because variability peaks during ramping were observed in measured data. It was correlated throughout all sites and related to broad atmospheric processes.

A few overarching considerations should guide the use of this data:

- This is a research data set. It is not meant to be a site investigation tool.
- The modeling resolution is 2 km by 2 km, which can exclude the local effects of ridges and other landscape features. The WIND Toolkit data cannot be expected to exactly model a real plant that is located in the same place.
- Caution must be used when analyzing maximum ramping statistics, because modeling artifacts and inaccuracies may be present. The maximum and minimum values should not



be utilized; instead, a confidence interval of some value less than 100% would be appropriate.

- It is important to respect the temporal limitations of the data set. For example, the smallest time step for this data set is 5 min; therefore, it would be inappropriate to utilize the data for a study of wind power plants serving on an automatic generation control signal (at 4-s resolution) by simply interpolating between the points.
- Spatial smoothing of wind power data has been extensively noted in the literature (e.g., Li et al 2007; McNerney and Richardson 1992; Nanahara et al. 2004); therefore, it would be inappropriate to scale up the power output from a single 10-MW site to a 100-MW plant by simply multiplying the power output by a factor of 10.
- The same weather phenomena that drive wind power output also drive solar and load data; therefore, load, wind, and solar data utilized should be time synchronized.
- As noted in the section on capacity factors, there are strong differences in seasonal wind power output; therefore, we recommend examining at least one full year of data.
- There is annual variability in a single site's capacity factor; this is one of the main reasons that the data set provided is 6 y long. Further examination is required before utilizing a single year as representative.

Given these caveats and the assumptions underlying the production of the data, the data set was found to acceptably model the power output of wind resources. This data can be used with reasonable confidence that it is generally representative of production by modern turbines at a hub height of 100 m.

## References

- 3Tier. (2010). *Development of Regional Wind Resource and Wind Plant Output Data Sets*. NREL/SR-550-47676. Golden, CO: National Renewable Energy Laboratory. Accessed March 2014: [www.nrel.gov/docs/fy10osti/47676.pdf](http://www.nrel.gov/docs/fy10osti/47676.pdf).
- 3Tier. (forthcoming). *Creating the WIND Toolkit Data Sets*. Boulder, CO. Golden, CO: National Renewable Energy Laboratory.
- AWS Truepower. (2012). *Updated Eastern Interconnect Wind Power Output and Forecasts for ERGIS*. NREL/SR-5500-56616. Golden, CO: National Renewable Energy Laboratory. Accessed March 2014: [http://www.nrel.gov/electricity/transmission/pdfs/ergis\\_update.pdf](http://www.nrel.gov/electricity/transmission/pdfs/ergis_update.pdf)
- Draxl, C.; Hodge, B.-M.; Clifton, A.; McCaa, J.; Jones, W. (forthcoming). *Overview and Meteorological Validation of the Wind Integration National Data Set (WIND) Toolkit*. NREL/TP-5000-61740. Golden, CO: National Renewable Energy Laboratory.
- EnerNex Corporation. (2011). *Eastern Wind Integration and Transmission Study*. NREL/SR-5500-47078. Golden, CO: National Renewable Energy Laboratory. Accessed March 2014: [www.nrel.gov/docs/fy10osti/47078.pdf](http://www.nrel.gov/docs/fy10osti/47078.pdf).
- GE Energy. (2010). *Western Wind and Solar Integration Study*. NREL/SR-5500-47434. Schenectady, NY. Golden, CO: National Renewable Energy Laboratory. Accessed March 2014: <http://www.nrel.gov/docs/fy10osti/47434.pdf>.
- International Electrotechnical Commission (IEC). (2005). *International Standard IEC 61400-1: Wind Turbines—Part I: Design Requirements*. Geneva, Switzerland.
- Li, P.; Banakar, H.; Keung, P.-K.; Far, H.G.; Boon-Teck, O. (2004). “Macromodel of Spatial Smoothing in Wind Farms.” *IEEE Transactions on Energy Conversion* (22:1).
- McNerney, G.; Richardson, R. (1992). “The Statistical Smoothing of Power Delivered to Utilities by Multiple Wind Turbines.” *IEEE Transactions on Energy Conversion* (7:4).
- Nanahara, T.; Asari, M.; Sato, T.; Yamaguchi, K.; Shibata, M.; Maejima, T. (2004). “Smoothing Effects of Distributed Wind Turbines: Part 1—Coherence and Smoothing Effects at a Wind Farm.” *Wind Energy* (7:2).
- Pennock, K. (2012). *Updated Eastern Interconnect Wind Power Output and Forecasts for ERGIS*. NREL/SR-5500-56616. Golden, CO: National Renewable Energy Laboratory. Accessed April 2014: [http://www.nrel.gov/electricity/transmission/pdfs/ergis\\_update.pdf](http://www.nrel.gov/electricity/transmission/pdfs/ergis_update.pdf).
- Potter, C.W.; Lew, D., McCaa, J.; Cheng, S.; Eichelberger S.; Gritmit, E. (2008). “Creating the Dataset for the Western Wind and Solar Integration Study (U.S.A.)” *Wind Engineering* (32:4). <http://dx.doi.org/10.1260/0309-524X.32.4.325>
- Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Barker, D.M.; Duda, M.G.; Huang, X.Y.; Wang, W.; Powers, J.G. (2008). *A Description of the Advanced Research WRF Version 3*. NCAR Tech Notes-475+STR. Boulder, CO: National Center for Atmospheric Research.