Updated Eastern Interconnect Wind Power Output and Forecasts for ERGIS

July 2012

K. Pennock
AWS Truepower
Albany, New York

NREL Technical Monitor: Kara Clark

Updated Eastern Interconnect Wind Power Output and Forecasts for ERGIS

July 2012

K. Pennock
AWS Truepower
Albany, New York

NREL Technical Monitor: Kara Clark
Prepared under Subcontract No. AFA-2-22465-01

Table of Contents

1 Introduction

AWS Truepower, LLC (AWST) was retained by the National Renewable Energy Laboratory (NREL) to update wind resource, plant output, and wind power forecasts originally produced by the Eastern Wind Integration and Transmission Study (EWITS). The new data set was to incorporate AWST's updated 200-m wind speed map, additional tall towers that were not included in the original study, and new turbine power curves. Additionally, a primary objective of this new study was to employ new data synthesis techniques developed for the PJM Renewable Integration Study (PRIS) to eliminate diurnal discontinuities resulting from the assimilation of observations into mesoscale model runs. The updated data set covers the same geographic area, 10-minute time resolution, and 2004–2006 study period for the same onshore and offshore (Great Lakes and Atlantic coast) sites as the original EWITS data set.

2 Wind Power Output

The final EWITS data set was comprised of both onshore and offshore sites within the Eastern Interconnect: 1326 onshore sites in 34 states (~580 GW) and 4948 offshore sites for 17 states (~99 GW). In review of the EWITS data for the Eastern Renewable Generation Integration Study (ERGIS), the study team discovered discontinuities in simulated diurnal variability aggregated from several wind farms. This pattern is likely due to the ingestion of observed data at known points in the grid to realign the mesoscale model with actual conditions every 12 hours (0000 and 1200 UTC).

Prior to ERGIS, wind power output for sites within the PJM Interconnect was updated with a new wind speed map, additional tall towers, and new power curves for the PRIS. The PRIS included all sites within 30 km of the PJM Interconnect (hereafter referred to as the PJM region). All EWITS sites and their relationship to the PJM region are shown in Figure 1. Due to PRIS time constraints, data within the PJM region were generated prior to data in the rest of the Eastern Interconnect. To facilitate a consistent data set over the whole of the Eastern Interconnect, the same onshore data within the PJM region delivered for the PRIS study were also delivered for the ERGIS study.
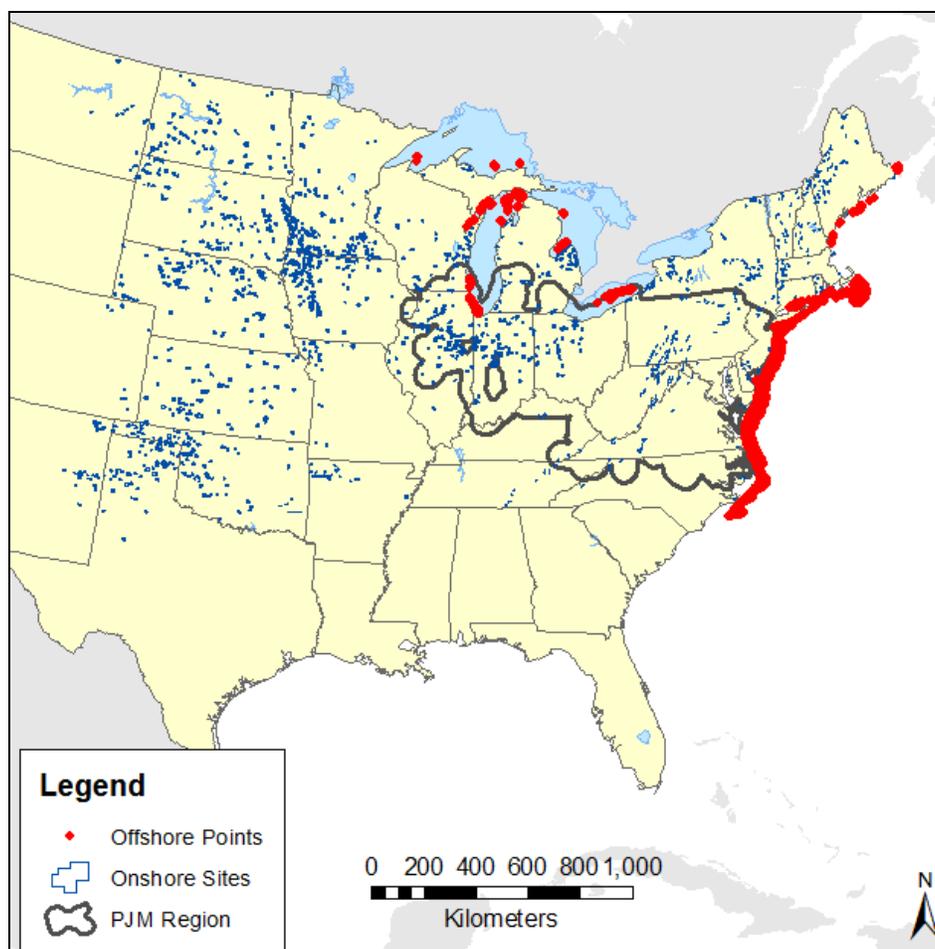
Figure 1. EWITS onshore (blue) and offshore (red) sites. The PJM region is outlined in gray.

2.1 Power Conversion

The EWITS historical mesoscale model runs and power conversion techniques were used to synthesize wind speeds and power output at each site.[1]

Minor updates to the AWST 200-m wind maps have occurred in the few years since EWITS as a result of additional validation efforts at AWST and NREL. The mean wind speeds at each site were adjusted to match these updated wind speed maps.

Wind speeds were also adjusted to diurnal mean profiles from tall towers within the study area. Ten tall towers were used for EWITS, whereas this study incorporated an additional 5 tall towers. As for EWITS, wind speeds from the tall towers were extrapolated to 80 m and 100 m hub height, and mean wind speeds for each hour of the day and month of the year were computed. The power conversion software calculates a correlation coefficient (r^2) between the simulated daily mean speeds for the site in question and the simulated daily mean speeds for each validation location. It then computes a blended adjustment matrix in which the weight given to each tower location is proportional to its correlation coefficient. The inclusion of additional

---

[1] M. Brower, "Development of Eastern Regional Wind Resource and Wind Plant Output Datasets", 2009, NREL/SR-550-46764, www.nrel.gov/wind/integrationdatasets/pdfs/eastern/2010/aws_truewind_final_report.pdf.

2

tall towers promotes better diurnal mean wind speeds in areas not previously covered by validation towers. It was found that only a diurnal adjustment need be applied, as the mesoscale model accurately predicted the variation in monthly wind speeds.

In order to account for advances in turbine technology since the EWITS data sets were created, composite power curves were created for each International Electrotechnical Commission (IEC) class using larger, more efficient turbines likely to be used for future wind farms. The composite power curves were created by averaging commercial megawatt-class wind turbine power curves normalized to their rated capacity. A separate composite was created for offshore sites. Table 1 details the power curves included in the composites, and Table 2 gives the resulting composite curves valid for the standard sea-level air density of 1.225 kg/m^3. The cut-out speed for the class 3 and offshore composites were adjusted to 22 and 30 m s^{-1}, respectively. The new IEC Class 2 power curve increased by 8-9% in the mid-range of speeds, while the Class 3 power curve increased by over 20% compared to the previous composites used for EWITS. Conversely, incorporating larger machines in the offshore composite decreased the energy in a normalized sense compared to the turbine used for the EWITS offshore sites. The old and new power curves are compared for each IEC class in Figure 2.

Table 1. Wind turbines used to create composite power curves, categorized by IEC class.

Turbine	Rated Power (MW)	Cut-in (m/s)	Max Power (m/s)	Cut-out (m/s)
Class I				
Siemens 3.0MW	3	3	14	25
Gamesa G80	2	4	17	25
Nordex N90HS	2.5	4	14	25
Vestas V90	3	4	14	25
Class II				
Vestas V112	3	3	13	25
Siemens 2.3MW	2.3	3	13	25
GE1.6 82.5	1.6	4	12	25
GE2.5xl	2.5	3	14	25
Class III				
Vestas V100	1.8	3	12	20
GE1.6-100	1.6	3	12	25
Repower 3.2M	3.2	3	12	22
Offshore				
Siemens 3.6MW	3.6	4	14	25
GE4.1MW	4.1	4	14	25
Repower 6M	6.15	3.5	14	30

Table 2. Updated composite power curves

Speed	IEC - 1	IEC - 2	IEC - 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

The power conversion software selects the most appropriate IEC class for each site based on the estimated maximum long-term annual average mean speed within the site, adjusted for air density. The IEC Class 1 and 2 turbines are assumed to have a hub height of 80 m and the IEC Class 3 turbine 100 m. It is assumed that the lower hub height will be used unless the wind resource dictates moving to a higher hub height to capture more wind.
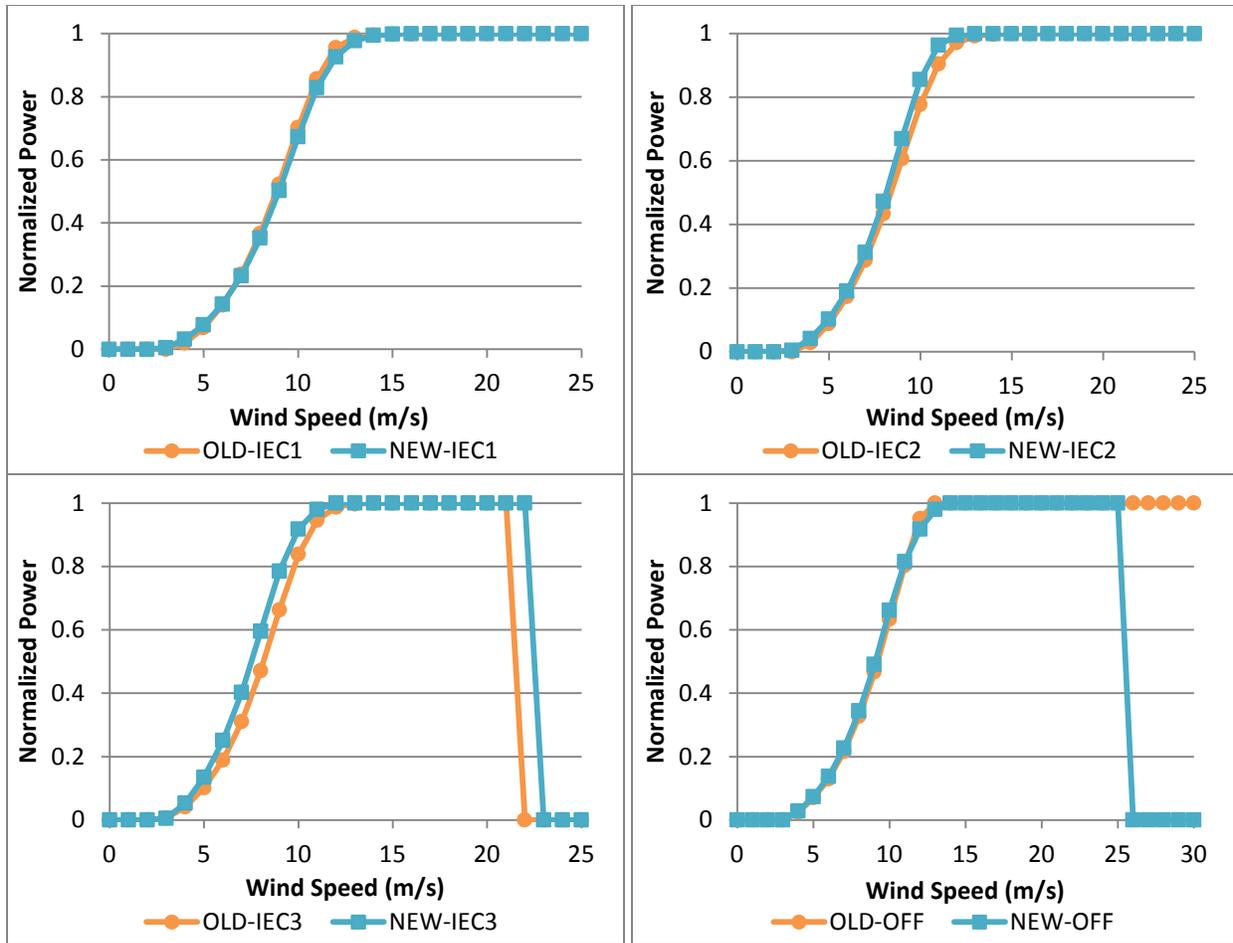
Figure 2. Comparison of EWITS (orange) and ERGIS (blue) composite power curves for IEC Class 1 (top left), Class 2 (top right), Class 3 (bottom left), and offshore (bottom right).

2.2 Diurnal Variability Adjustment

Several techniques were employed in the PRIS to minimize the impact of the data ingestion at individual wind farms, but the problem was still apparent in the aggregate of both onshore and offshore wind sites. A new technique was developed to adjust the correlated component of the wind power fluctuations by adding a proportion of the adjustment to each individual site. The result is a flatter distribution of variability throughout the day (Figure 3). Although the 12-hour data assimilation problem was addressed, an hourly signal remained. The study team determined that the hourly fluctuations were acceptable for this project. Because sites within the PJM region were updated before the entire Eastern Interconnect, the aggregate of those sites was computed separately from the aggregate of the remaining sites. Separate aggregates were also computed for onshore and offshore sites, resulting in a total of 4 separate aggregations.

Sites outside the PJM region were adjusted in this manner and delivered to NREL in May 2012. However, it was subsequently determined that while the adjustment corrected the diurnal discontinuity across the aggregate of the Eastern Interconnect, the issue still existed on a state and regional level. For this reason, all sites were re-run with the adjustment applied by state (onshore and Atlantic offshore sites) or by lake (Great Lakes offshore sites). This adjustment was deemed satisfactory, and the final data set was delivered in June 2012.
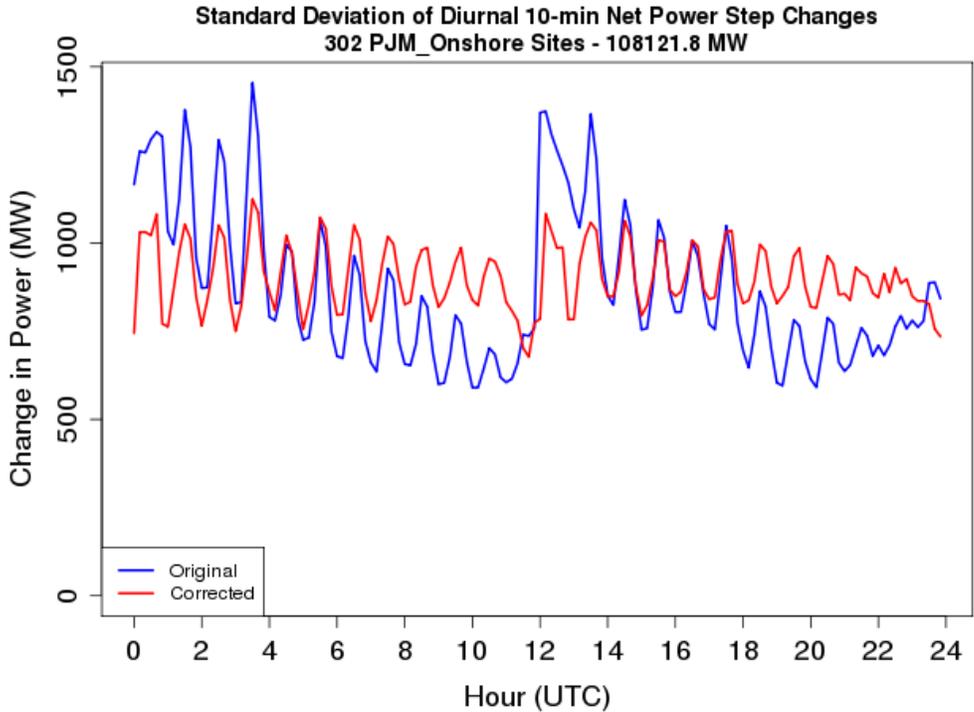
Figure 3. Diurnal mean standard deviation of 10-minute net power deltas from 302 onshore sites within the PJM region before (blue) and after (red) the diurnal correction was applied for the years 2004-2006.

Although the diurnal discontinuity in each aggregate of sites was adjusted, only small changes were made at each site. Figure 4 shows a comparison of the data at a sample onshore site before (red) and after (green) the aggregate adjustment compared to the original EWITS data (blue). There is almost no discernible change in annual or monthly mean patterns (Figure 4, left top and middle panels), and only a very slight change in diurnal mean patterns (Figure 4, bottom left). The affect of the adjustment on the diurnal variability (i.e., the standard deviation of diurnal mean 10-minute ramps), is shown in the top right panel of Figure 4. The adjustment has minimal impact on the ramp frequency distribution (Figure 4, middle right), but created a flattening of the spectrum and increased periodicity at higher frequencies (Figure 4, bottom right).

The study team expressed some concern that the adjustment negatively impacted the power spectral density, or periodic fluctuations in power output at individual sites. Upon further review, it was found that the diurnal adjustment applied to individual sites and not the aggregate was the cause of the PSD modification. For this reason, and the fact that variability is expected to vary throughout the day due to strong diurnal patterns in wind speeds in the Great Plains, the individual site adjustment was not used for sites outside of the PJM region. Because the data had already been delivered for the PRIS study, the individual site adjustment was retained for onshore sites within the PJM region to facilitate a consistent data set between studies. Figure 5 shows the result of the aggregate adjustment only at a comparable onshore site outside of the PJM region. The effect on the diurnal mean output and variability is decreased, and the flattening of the spectrum is no longer visible. The prominent peak in the power spectral density plot at 24 hours suggests that the impact of data assimilation is not entirely removed at individual sites, but is reduced. Figure 6 shows the diurnal variability for the aggregate of all onshore sites outside

6

the PJM region before and after the correction. Similar results can be found for each state or region (e.g. New England, Midwest; not shown).
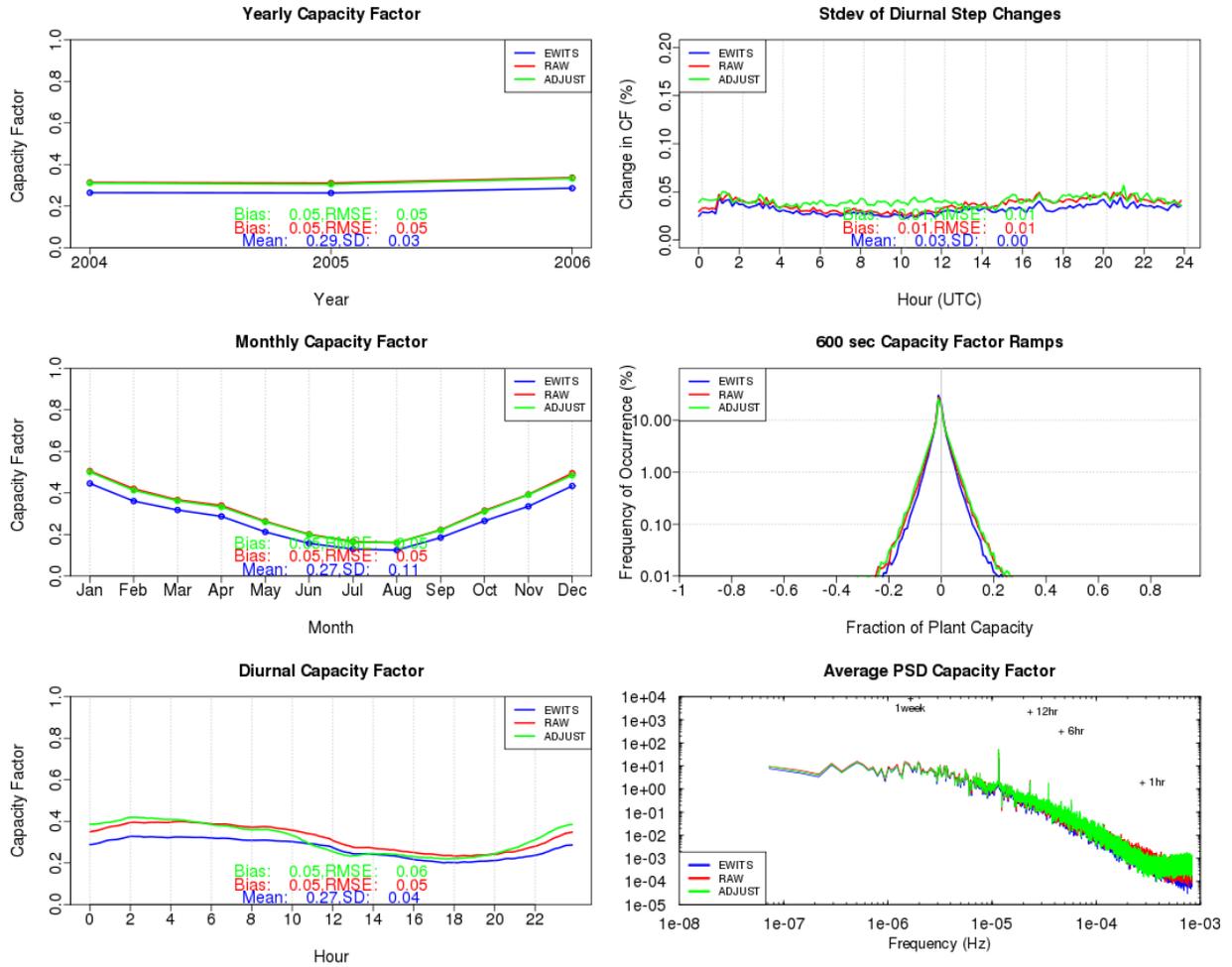
Figure 4. Comparison of net power normalized by nameplate capacity at a sample onshore site within the PJM region before (red) and after (green) the diurnal adjustment compared to the same site in the original EWITS data set (blue). The left panels depict annual, monthly, and diurnal means, while the right panels show the diurnal mean variability, frequency distribution of 10-minute ramps, and power spectral density.

Figure 5. As in Figure 4 but for a sample onshore site outside of the PJM region.
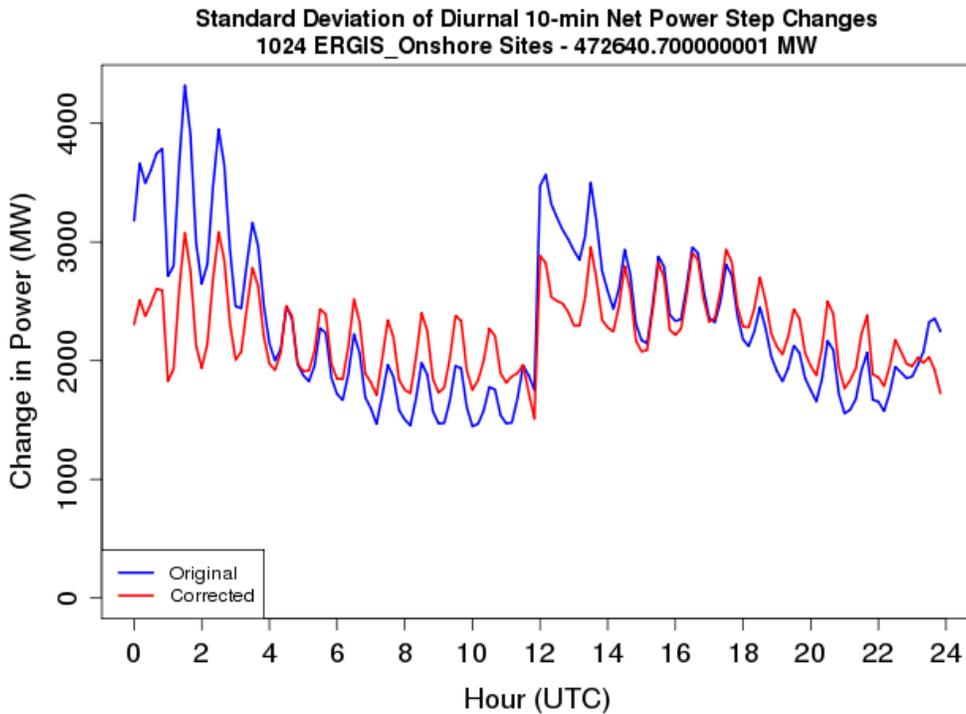
Figure 6. As in Figure 3 but for 1024 onshore sites outside the PJM region.

2.3 Discussion

Power output at individual sites can be expected to deviate from the original EWITS data mainly due to updates to the wind maps and power curves. Wind speed differences are generally small, with the greatest difference in the offshore mid-Atlantic, western Maine, and northern Great Plains (Figure 7). In some cases, the mean wind speed changed enough tha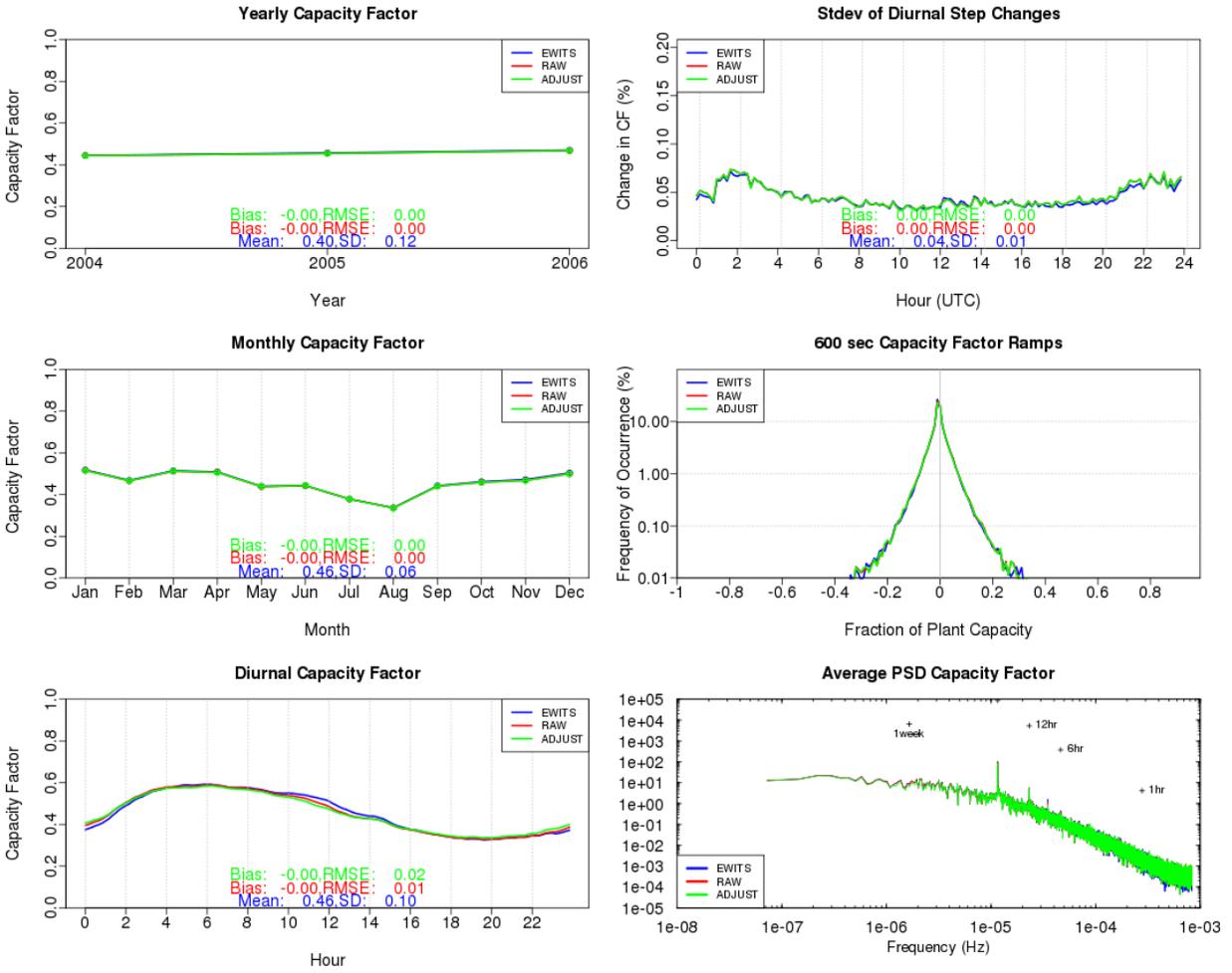t a site was assigned a different IEC class. Table 3 summarizes the onshore sites that changed classes, showing that most sites remained in their original EWITS classification. Offshore sites were run with the same power curve and hub height regardless of the IEC classification, so changes are due only to the power curve and offshore wind speed map.

Table 3. Number of onshore sites in each IEC class for EWITS and ERGIS.

ERGIS \ EWITS	IEC-1	IEC-2	IEC-3	Total
IEC-1	67	21	0	88
IEC-2	48	563	41	652
IEC-3	0	35	551	586
Total	115	619	592	1326

9

Figure 7. Site mean wind speed difference (ERGIS minus EWITS).

3 Wind Power Forecasts

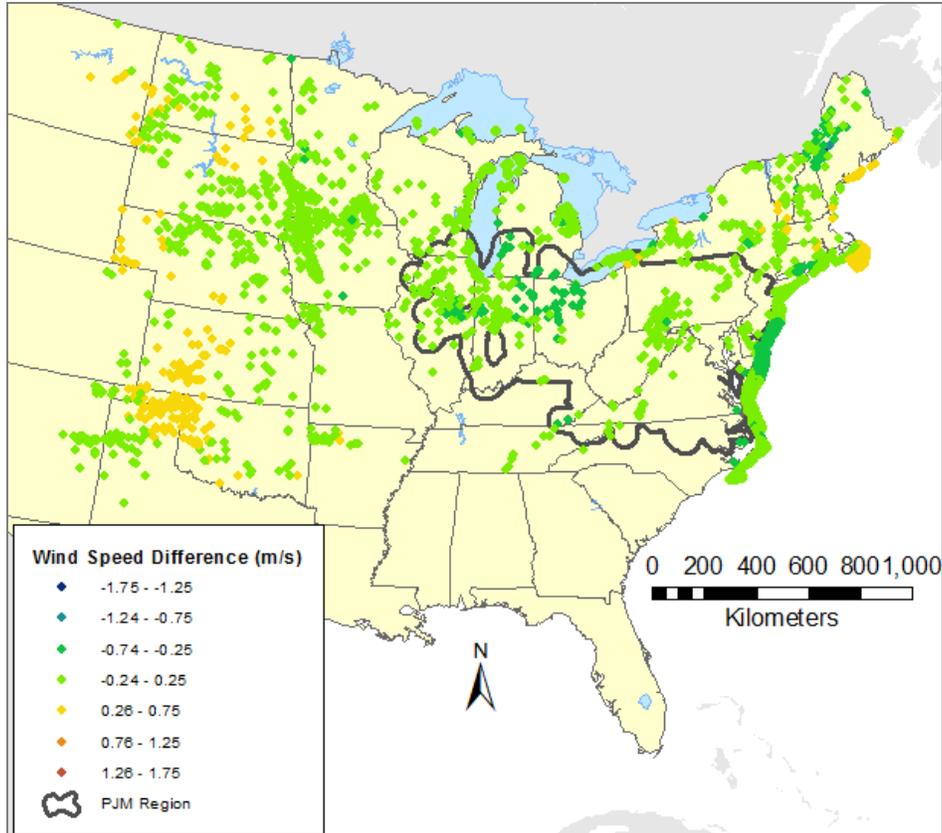Using the final updated wind power profiles, AWST generated new wind power forecasts at each onshore and offshore site for three different time horizons: next-day, four, and six hours ahead. The methodology followed EWITS with the exception that additional operational forecasts made available through PRIS were used to refine forecasts generated within the PJM region. The synthetic forecast methodology and results follow.

3.1 Forecast Synthesis Procedure

Forecasts for each time horizon were generated by running a statistical forecast synthesis tool developed by AWST: SynForecast. This tool uses actual forecasts and observed plant output to develop a set of transition probabilities. The probabilities are applied to simulated wind plant and load data, stepping forward in time from a random starting point using a mathematical tool known as a Markov chain. This process results in a synthetic forecast that imitates the statistical behavior of a real forecast.

The first step in the forecast synthesis process is to obtain a sequence of actual wind forecasts for one or more operating wind projects. AWST obtained power forecasts and observed output from 29 wind plants within the PJM region for PRIS. Due to various inconsistencies in the actual plant data provided by PJM including excessive curtailments, missing data, plant underperformance or an incomplete period of record, 11 wind plant outputs and forecasts were removed from the training dataset used in creating the transition probabilities. The period of record for the

remaining 18 PJM wind forecasts was July of 2007 through June 2011. In addition to the PJM provided forecasts, the hindcasts generated for EWITS at four Great Plains wind plants (Trent Mesa, Blue Canyon, Lake Benton, and Storm Lake) were also used. The period of record for the AWS hindcasts was February 2004 through December 2006.

From each of set of forecasts, the SynForecast program constructed a matrix of forecast probabilities of the following form:

$$P(A_t \cap F_{t-1} \cap F_t)$$

The probability P is the number of occurrences for which the actual output was A_t and the forecasted outputs were F_{t-1} and F_t, where t is a particular moment in time and $t-1$ is the previous moment (one-hour earlier). Before constructing this matrix the actual and forecasted wind plant output values were normalized to the rated capacity of the wind project and placed in 10 bins ranging in capacity factor from 0.05 to 0.95 in increments of 0.10.

For each wind project site, the SynForecast program selected, at random, one of the transition probability matrices. Starting with a random seed, the program stepped forward in time taking random draws from the transition matrix. In this manner, an hourly next-day forecast was synthesized. The same procedure was followed for 4-hour, 6 hour, and next-day forecasts.

To best capture any regional characteristics of wind power forecasts, sites located in the PJM region used transition matrices derived from the PJM forecast data while projects located outside the PJM region used transition matrices from the AWST hindcasts. The error characteristics between the AWS Truepower hindcasts and PJM forecasts are quite similar and thus the error patterns between PJM and non-PJM wind projects are similar (Table 4).

Given the lack of offshore wind plant data in North America it is not known if the forecast skill for offshore projects is similar to offshore projects. Thus, both onshore and offshore projects made use of the same transition probability matrices.

Table 4. Comparison of forecast errors between the data sets used to train the SynForecast program. MAE and RMSE are giving in % of plant capacity. Correlation represents the Pearson r statistic (forecasted vs. observed power).

Validation Period/Metric	AWST Hindcasts	PJM Forecasts
4 Hour Ahead – MAE	11.21	11.49
6 Hour Ahead – MAE	11.92	12.72
Next Day – MAE	14.33	13.70
4 Hour Ahead – RMSE	15.22	16.26
6 Hour Ahead – RMSE	16.11	18.25
Next Day – RMSE	19.12	19.39
4 Hour Ahead – Correlation	0.81	0.82
6 Hour Ahead – Correlation	0.78	0.79
Next Day – Correlation	0.76	0.76

3.2 Validation

To verify the accuracy of the synthetic wind forecast, AWST compared synthesized forecasts with the actual forecasts at the 22 validation wind projects (4 NREL sites plus 18 PJM sites). First, the time correlation of the actual and forecasted generation, root-mean-square (RMSE), and mean absolute forecast error (MAE) were considered for all proposed forecast time horizons. The results of the actual and synthesized forecasts were very similar, as shown by the next-day statistics in Table 5. It should be noted that the MAE and RMSE depend in part on the average plant output, with more productive plants experiencing greater forecast errors as a fraction of rated capacity because they spend more time in the steeply sloping parts of their power curves. Similar patterns were found for 4- and 6-hour forecasts, with lower MAE and RMSE at the shorter look-ahead periods.

Table 5. Next day wind forecast capacity factor correlation, RMSE, and MAE.

Next Day Forecasts	Correlation (R)		RMSE		MAE	
Site	Actual	SynFcst	Actual	Synfcst	Actual	Synfcst
PJM Site 1	0.768	0.678	0.230	0.237	0.167	0.180
PJM Site 2	0.695	0.636	0.234	0.238	0.160	0.172
PJM Site 3	0.826	0.796	0.175	0.185	0.121	0.133
PJM Site 4	0.790	0.735	0.186	0.198	0.129	0.146
PJM Site 5	0.825	0.813	0.158	0.155	0.106	0.111
PJM Site 6	0.759	0.688	0.203	0.210	0.141	0.155
PJM Site 7	0.726	0.705	0.193	0.187	0.128	0.131
PJM Site 8	0.787	0.760	0.205	0.206	0.146	0.156
PJM Site 9	0.778	0.797	0.165	0.158	0.114	0.115
PJM Site 10	0.803	0.778	0.203	0.205	0.142	0.151
PJM Site 11	0.761	0.746	0.148	0.149	0.099	0.106
PJM Site 12	0.681	0.659	0.223	0.218	0.150	0.163
PJM Site 13	0.813	0.771	0.198	0.207	0.140	0.153
PJM Site 14	0.720	0.747	0.198	0.184	0.137	0.137
PJM Site 15	0.792	0.719	0.184	0.207	0.126	0.152
PJM Site 16	0.755	0.724	0.191	0.196	0.131	0.145
PJM Site 17	0.777	0.779	0.182	0.180	0.125	0.133
PJM Site 18	0.813	0.773	0.192	0.201	0.133	0.147
Blue Canyon	0.770	0.736	0.206	0.218	0.158	0.168
Lake Benton	0.719	0.690	0.190	0.199	0.142	0.153
Storm Lake	0.794	0.778	0.161	0.167	0.120	0.126
Trent Mesa	0.774	0.758	0.206	0.211	0.156	0.164

Next, the autocorrelation of the plant output, forecasts, and the forecast errors was considered. The autocorrelation indicates the degree to which a particular parameter tends to persist over time. A parameter that typically changes little would have an autocorrelation factor of nearly one, whereas one that fluctuates randomly would exhibit an autocorrelation factor of nearly zero. Figure 8 compares the synthetic and actual next day forecast autocorrelation. It was found that the both the actual and synthetic forecasts exhibit strong autocorrelation over a period of one to a

few hours. The synthetic forecasts have a higher degree of autocorrelation than the actual wind project forecasts but the differences are small and the functional dependence with increasing time shift is similar. The autocorrelation of forecast errors is considerably lower than the forecasts (Figure 9), and the pattern of decreasing correlation with increasing time shift is captured quite well in the synthetic forecasts. The synthetic forecast errors have a higher degree of autocorrelation than the actual forecast errors.
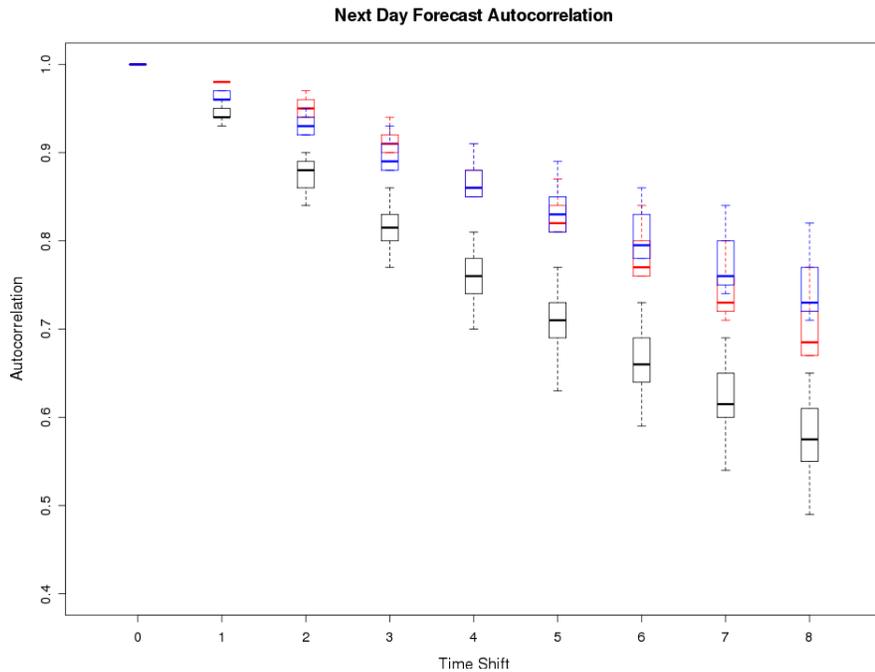

Figure 8. Autocorrelation of observed wind plant output (black), actual (red) and synthetic (blue) next-day forecasts. Time shift is represented in hours. Box and whiskers represent the minimum, quartiles, median, and maximum for the 22 plants.
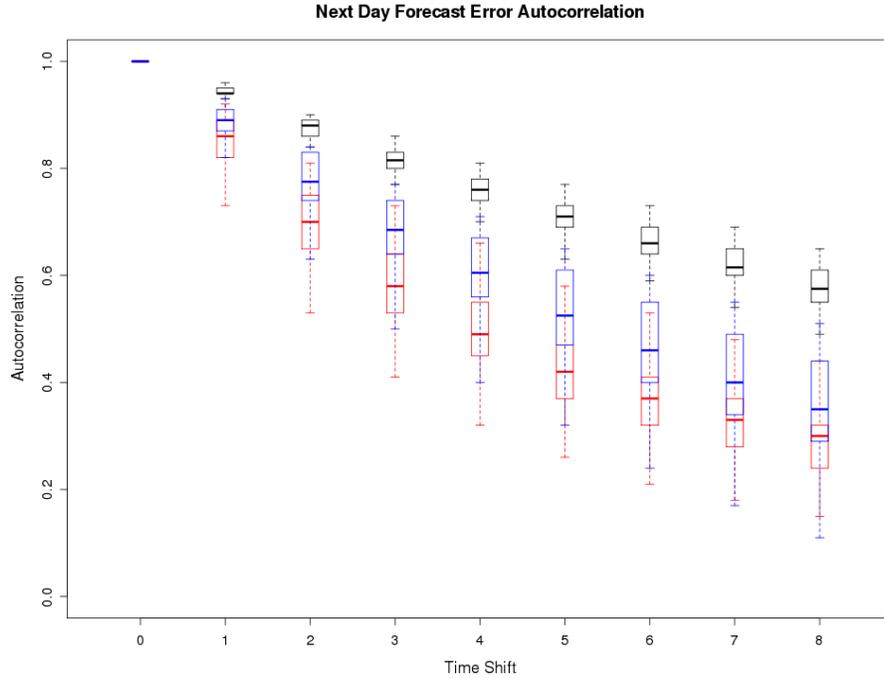
Figure 9. As in Figure 8 but for next-day forecast errors.

The correlation of forecast errors as a function of distance between projects was evaluated next. The correct modeling of the spatial correlation of forecast errors is an important consideration as it has implications on the aggregate impact of many wind projects over a large region. If the synthesized forecast errors are not correlated enough between projects, then the aggregate forecast error will be underestimated, and therefore also the impacts of those errors on system operations. Overestimating the degree of correlation between projects will have the opposite effect. With the 18 actual plant forecasts available through PRIS compared to the 4 hindcasts originally available for EWITS, a more in-depth analysis of the spatial correlations was possible. Results from the PJM forecasts only are included here. Figure 10 shows how the actual and synthetic forecast errors correlate in space. The actual forecast errors are slightly more correlated ($\bar{r}$=0.16) in space than the synthetic forecasts ($\bar{r}$=0.13), however, the Synforecast program accurately captures the observed decrease correlation with increasing distance between plants.
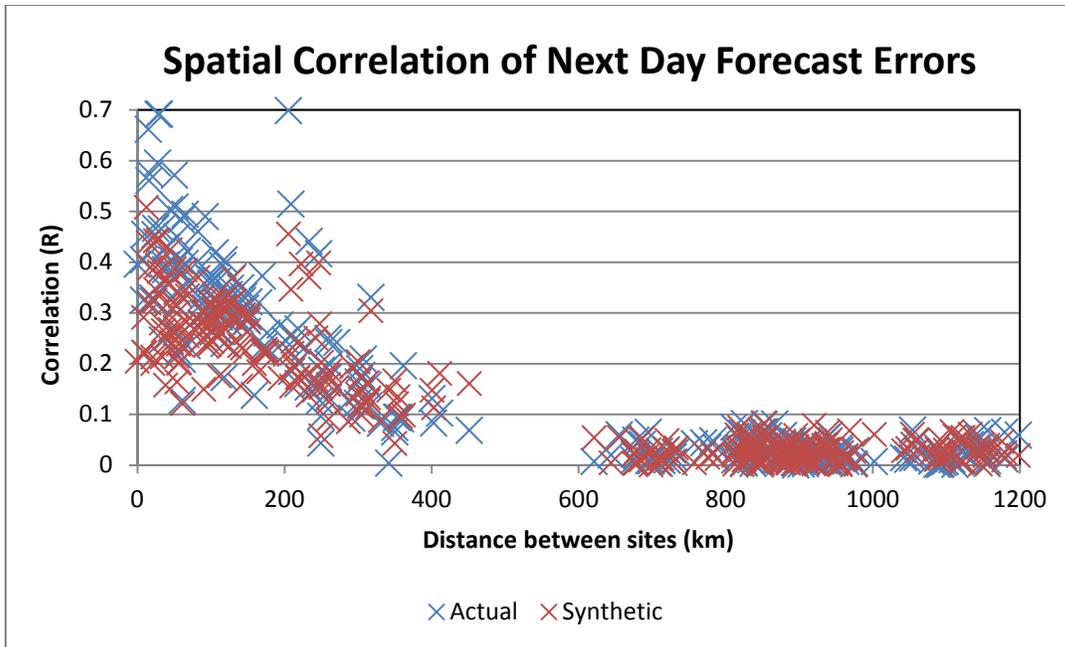
Figure 10. Next day forecast error correlation as a function of distance between projects.

Finally, the aggregation of forecast errors is considered for 18 PJM wind plants. Errors for each plant were summed, and the total error was divided by the aggregate nameplate capacity. The total error for the actual next day forecasts range from -35% to 39% of total capacity while synthetic forecast error ranges from -36% to 37% of capacity. The system-wide MAE for the actual and synthetic forecasts is approximately 7.5% and 6.9%, respectively. Figure 11 shows the total error distribution is captured quite well by the synthetic forecast process, with large forecast errors slightly less common in the synthetic forecasts compared to actual PJM forecasts.
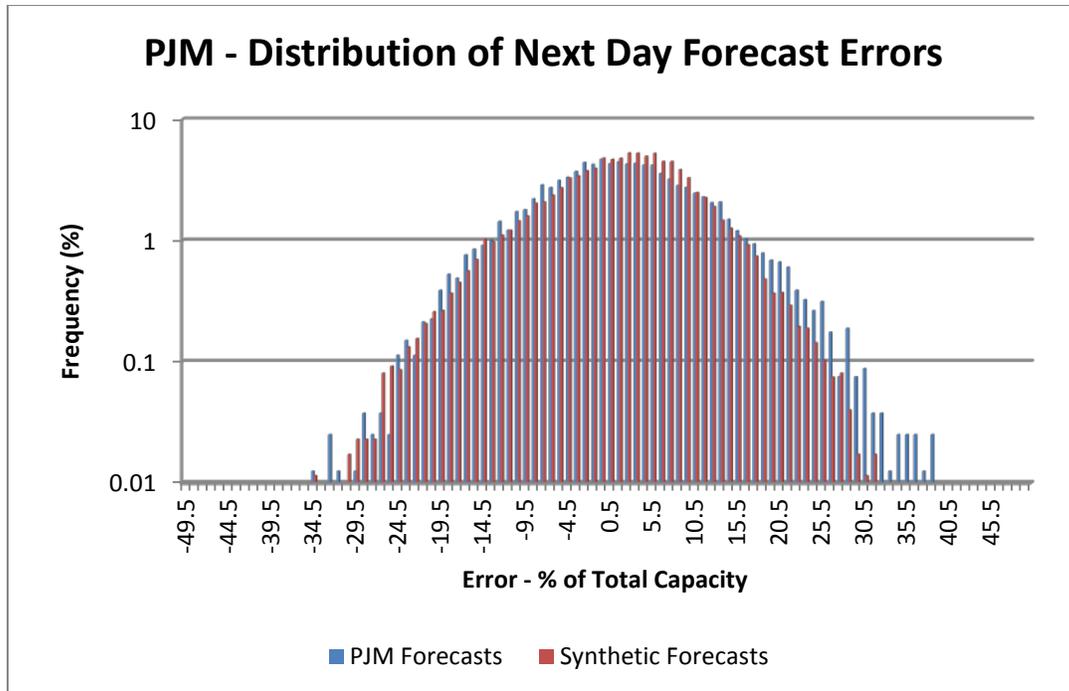
Figure 11. Distribution of actual and synthetic next-day forecast errors aggregated for 18 wind plants in the PJM region. The x-axis shows the sum of forecast errors in % of total capacity, while the y-axis shows the frequency of occurrence in %.

4 Deliverables

A separate comma-delimited text file containing 10-minute wind speed and power output for the period 2004–2006 was provided for each onshore and offshore site. Wind speeds at 80 m are given for IEC Class 1 and 2 onshore sites, whereas 100-m wind speeds are provided for Class 3 onshore and all offshore sites. A sample text file is shown in Table 6. A file containing the latitude, longitude, elevation, capacity, CF, state, and region for each site was also provided. The region is specified as either "P" (within the PJM region) or "E" (within the Eastern Interconnect outside of the PJM region). A separate comma-delimited text file was delivered for each forecast look-ahead period for each site. Each file contains the date and time, modeled wind power output, forecast power output, and forecast error (Table 7). The final version of all files was delivered to NREL via FTP in June 2012.

Table 6. Sample plant output data file.

SITE NUMBER: 05666 RATED CAP: 20.0 IEC CLASS: 2 LOSSES (%): 18.9			
SITE LATITUDE: 41.17482 LONGITUDE: -72.74722			
DATE	TIME(UTC)	SPEED100M(M/S)	NETPOWER(MW)
20040101	0010	11.817	14.79
20040101	0020	12.311	15.65
…	…	…	…
20070102	0000	12.885	16.03

Table 7. Sample forecast data file.

DATE-GMT	TIME-GMT	OBSPWR	FCSTPWR	ERR
20040101	0100	12.8	11.7	-1.1
20040101	0200	12.3	9.6	-2.6
...	...	...	...	...
20070102	0000	3.3	13.2	9.9

5 Summary

AWST provided new wind power output and forecasts for all onshore and offshore sites within the Eastern Interconnect for the years 2004–2006. Updates include scaling to adjusted wind speed maps and power conversion using new composite turbine power curves. More tall towers and actual plant forecasts were also included since EWITS. Finally, an adjustment was made to the system-wide aggregated diurnal variability to minimize the impact of assimilating observations into the mesoscale model every 12 hours. New 10-minute wind power profiles as well as hourly forecasts for 4-hour, 6-hour, and next-day look-ahead periods were provided for each onshore and offshore site. This data set is meant to replace data originally generated for EWITS.