National Renewable Energy Laboratory
A national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Innovation for Our Energy Future

Development of Eastern Regional Wind Resource and Wind Plant Output Datasets

March 3, 2008 — March 31, 2010

Michael Brower
AWS Truewind, LLC
Albany, New York

Subcontract Report
NREL/SR-550-46764
December 2009

NREL is operated for DOE by the Alliance for Sustainable Energy, LLC Contract No. DE-AC36-08-GO28308

Development of Eastern Regional Wind Resource and Wind Plant Output Datasets

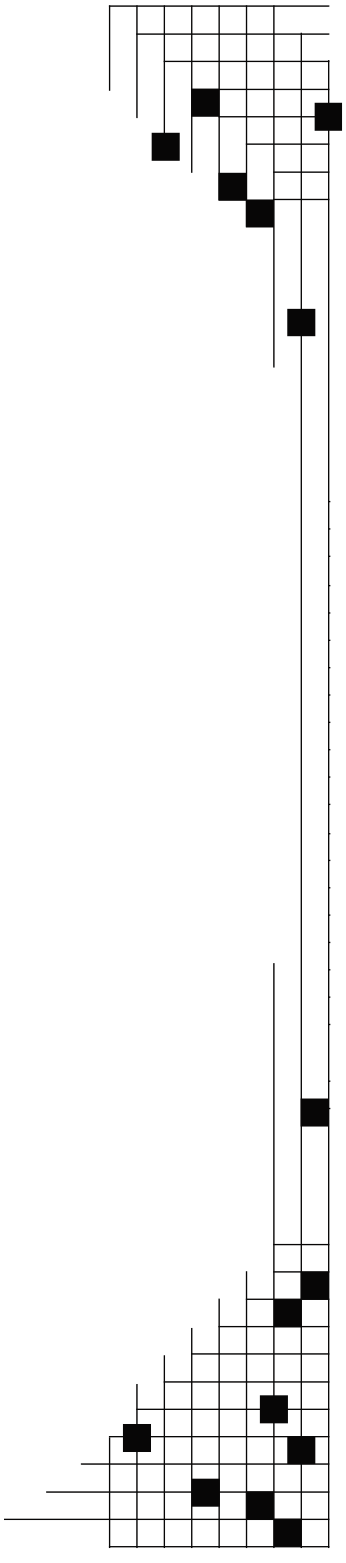
March 3, 2008 — March 31, 2010

Michael Brower
AWS Truewind, LLC
Albany, New York

NREL Technical Monitor: David Corbus

Prepared under Subcontract No. ACO-8-88500-01

Subcontract Report
NREL/SR-550-46764
December 2009



National Renewable Energy Laboratory
1617 Cole Boulevard, Golden, Colorado 80401-3393
303-275-3000 • www.nrel.gov

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

Contract No. DE-AC36-08-GO28308

Available electronically at http://www.osti.gov/bridge

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
online ordering: http://www.ntis.gov/ordering.htm

This publication received minimal editorial review at NREL

Printed on paper containing at least 50% wastepaper, including 20% postconsumer waste

Table of Contents

Development of Eastern Regional Wind Resource and Wind Plant Output Datasets

Subcontract No. ACO-8-88500-01

Final Report

1 Introduction

In March 2008, AWS Truewind was engaged by the National Renewable Energy Laboratory (NREL) to develop a set of wind resource and plant output data for the eastern United States. (The study region is shown in Figure 1.1.) The objective of this project was to provide wind resource inputs to the Eastern Wind Integration and Transmission Study (EWITS), an NREL-led effort to evaluate the potential impacts on the electric power system of the deployment of wind power plants meeting up to 30% of retail electric energy sales in a region spanning a large part of the Eastern Interconnect. To carry out this study, NREL required a set of data that would capture in a realistic fashion both the temporal and spatial variability of the wind resource and associated wind power generation of onshore and offshore projects totaling 300 GW. These data were to be based on high-resolution simulations of the historical climate performed by a mesoscale numerical weather prediction (NWP) model covering 2004 to 2006.

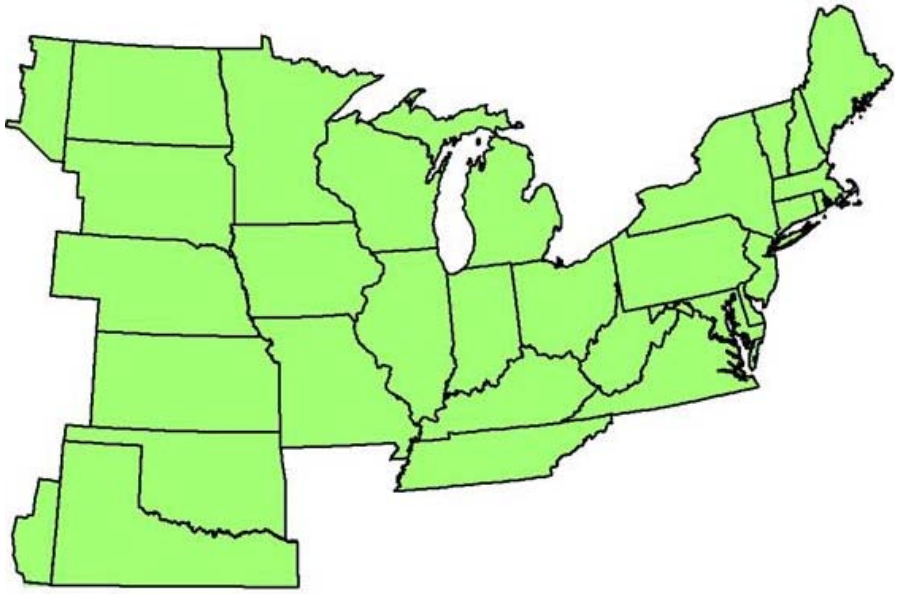


Figure 1.1. Eastern Wind Integration and Transmission Study (EWITS) region.

AWS Truewind performed this work over a period of nine months from March 2008 to November 2008. The work was divided into the following five technical tasks:

- Develop wind resource datasets based on mesoscale modeling
- Participate in the selection of wind sites

- Generate wind plant output
- Simulate forecasts
- Simulate one-minute samples of wind generation

This document presents AWS Truewind's final technical report on the methods used and results achieved for each of these tasks. A separate validation report is provided as an appendix.

2 Develop Wind Resource Datasets Based on Mesoscale Modeling

This task was divided into two subtasks: validation of the mesoscale model and selection of the "best" model configuration, and the mesoscale modeling (main runs).

2.1 Mesoscale Model Validation

The objective of the validation portion of this task was to determine the best mesoscale model and model configuration for performing the main runs. This was achieved by testing a number of different configurations covering one full year of simulations and comparing the results with wind measurements from ten tall towers in the region.

2.1.1 Validation Sites and Method

A preliminary list of seven tall towers with good data recovery in the three study years was identified by AWS Truewind. Data for the towers were extracted from AWS Truewind's archives and time-matched. Twenty-seven periods of approximately two weeks each in which data recovery was high at all seven towers were identified. Subsequently, at NREL's request, three additional towers (including one offshore) were chosen to provide better geographic coverage, bringing the total number to ten. For the offshore tower, data were not available for all three years, so a single year was used. Table 2.1 lists the periods.

The ten validation towers are located in the following general areas:

Maine (inland forested ridge top)
New York (upstate, mixed forest/cropland, complex terrain)
Kentucky (eastern mixed open/forested mountaintop)
West Virginia (mixed open/forested ridgeline)
Texas (open mesa)
Indiana (northwest, rolling hills, cropland)
Minnesota (south-central, open, flat)
North Dakota (eastern escarpment, open)
Kansas (west-central, open)
Offshore (Atlantic)

As agreed with NREL, no further information about these masts will be provided to protect the confidential client information.

Table 2.1. Dates of simulations for the ten validation sites. Each simulation was conducted from 0000 Coordinated Universal Time (UTC) of the first day to 2359 UTC of the last day.

Sites	Dates		
NY, WV, IN, KY, SD, KS, ME, TX, MN	20040114-20040128	20041028-20041111	20051125-20051209
	20040128-20040211	20041223-20041231	20051209-20051223
	20040311-20040325	20050225-20050311	20051231-20060114
	20040423-20040507	20050325-20050409	20060211-20060225
	20040618-20040630	20050507-20050521	20060409-20060423
	20040811-20040825	20050604-20050618	20060521-20060604
	20041006-20041014	20050714-20050728	20050922-20051006
	20041014-20041028	20050728-20050811	20060630-20060714
	20040825-20040908	20050908-20050922	20061211-20061225
OFFSHORE ATLANTIC	20031231-20040114	20040506-20040520	20040909-20040923
	20040114-20040128	20040520-20040603	20040923-20041007
	20040128-20040211	20040603-20040617	20041007-20041021
	20040211-20040225	20040617-20040701	20041021-20041104
	20040225-20040311	20040701-20040715	20041104-20041118
	20040311-20040325	20040715-20040729	20041118-20041202
	20040325-20040408	20040729-20040812	20041202-20041216
	20040408-20040422	20040812-20040826	20041216-20041230
	20040422-20040506	20040826-20040909	

Two models were selected for testing: the Mesoscale Atmospheric Simulation System (MASS), a proprietary model developed by AWS Truewind partner MESO, Inc., and the Weather Research and Forecasting (WRF) model, a community weather model. Each model was run in a variety of configurations, which, depending on the model, included different turbulence parameterization schemes, numbers of vertical levels, data sources for initialization of the model runs, and data sources for assimilation.

Table 2.2 lists the configurations that were tested. Unless otherwise noted, the MASS runs employed 25 vertical levels, assimilated rawinsonde data every 12 hours, and used a turbulent kinetic energy (TKE) planetary boundary layer (PBL) scheme and one-way nesting from coarser to finer grids. The WRF runs employed 28 vertical levels and used the Yonsei University (YSU) PBL scheme and two-way interactive nesting between grids. All runs employed a horizontal grid resolution of 2 km for the innermost grid. The runs were initialized either by the NCEP/NCAR Global Reanalysis (NNGR) dataset or by the North American Regional Reanalysis (NARR) dataset. “Sfc” refers to the assimilation of surface data. MYJ refers to the Mellor-Yamada-Janjic PBL scheme. Further information about these configurations can be found in the validation report in the Appendix.

Table 2.2 Validation experiments.

Experiment	Model	Gridded Data, Res	Other
1. MASS/NNGR	MASS 6.8	NNGR, 190 km	
2. MASS/NARR		NARR, 32 km	
3. MASS/NNGR/sfc		NNGR, 190 km	Surface data
4. MASS/NNGR/35 levels		NNGR, 190 km	35 vertical levels
5. WRF/NARR	WRF 2.2.1	NARR, 32 km	
6. WF/NNGR		NNGR, 190 km	
7. WRF/NARR/MYJ		NARR, 32 km	MYJ PBL scheme

2.1.2 Results

AWS Truewind compared the simulated and observed 80 m wind speeds for each tower and model configuration. (Hours for which the observed data were missing or invalid were excluded from the comparison.) For the simulations, the values were hourly samples extracted from the runs, i.e., the state of the model was extracted at the top of each hour. For the observations, the values were hourly averages. For the purposes of this study, this distinction has no practical impact.

To give the reader a sense of the diversity of results, Figures 2.1-2.3 present graphical comparisons of the diurnal and seasonal patterns for one of the configurations (MASS/NNGR/sfc) at three locations (Kansas, Maine, Minnesota). The red line on each chart is the observed average for that hour of the day or for that month. The green line is the raw model speed averaged for the hour or month. The blue line is the model speed scaled by a constant multiplicative factor so that the model average over all records equals the observed average. This last represents the scaling that will be done to match the expected average wind speed at a wind project site.

It was observed that at most sites all model configurations tended to underestimate the wind speed. The main reason for this is that the spatial resolution of 2 km was not adequate to fully resolve the elevated features on which most of the masts are located. In addition, it was noted that at several sites the models (after scaling) tended to overestimate the wind speed at night and underestimate it during the day (especially visible in Figure 2.2). This may reflect limitations of PBL schemes, which have difficulty accurately simulating the nocturnal boundary layer. Somewhat surprisingly, this problem seems insensitive to the choice of PBL scheme. (The experiments tested three: the MASS TKE scheme and two WRF schemes.)

At NREL's request, AWS Truewind investigated certain features observed in the comparison charts: (i) It was observed that the relative model speeds of the sites in Kentucky and Maine were similar, although it would normally be expected that the Maine site would be windier. However, this was explained by the fact that the Maine site is at a much lower elevation than the Kentucky site. (ii) A marked dip in the wind speeds at some sites at 1200 UTC, rather than 1300 UTC as expected based on the timing of the rawinsonde assimilation, was ascribed to an error in the AWS Truewind data assimilation script. This was later corrected in the main runs. (iii) Although the modeled and observed monthly patterns were quite similar at most towers, there was a

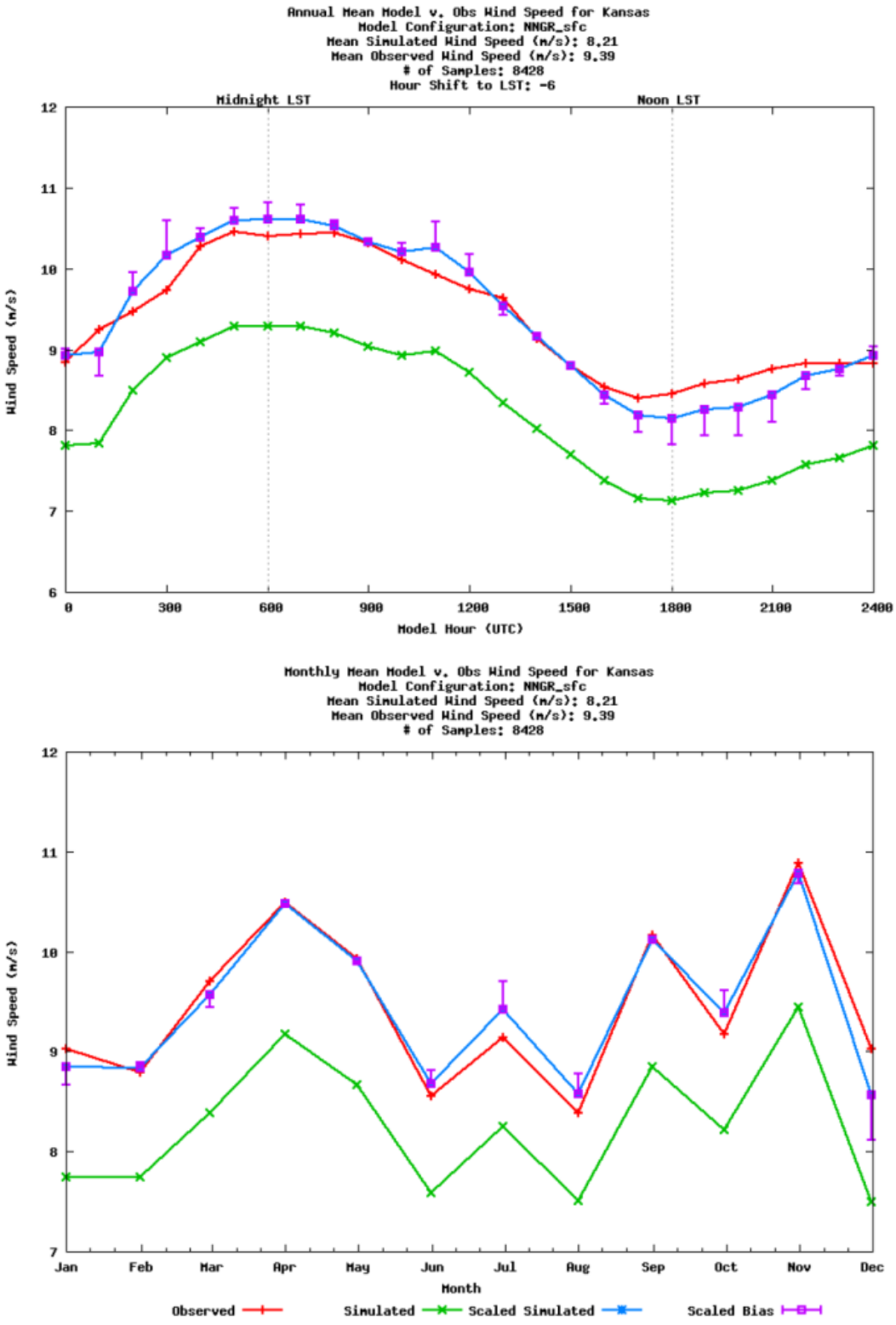


Figure 2.1. Comparison of diurnal (top) and monthly (bottom) patterns of average wind speed at the Kansas tower for the MASS/NNGR/sfc configuration.

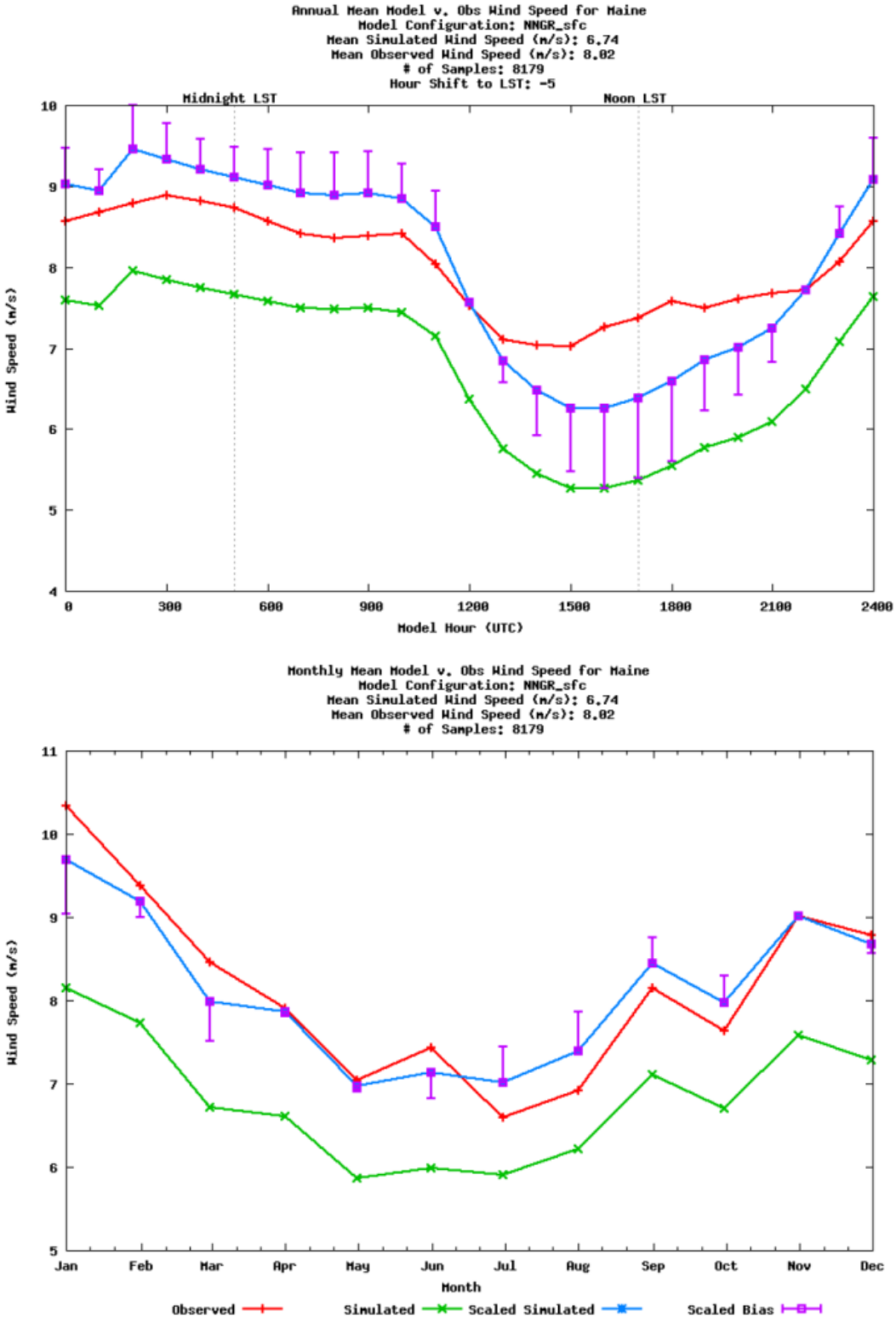


Figure 2.2. Same as Figure 2.1 for the Maine tower.

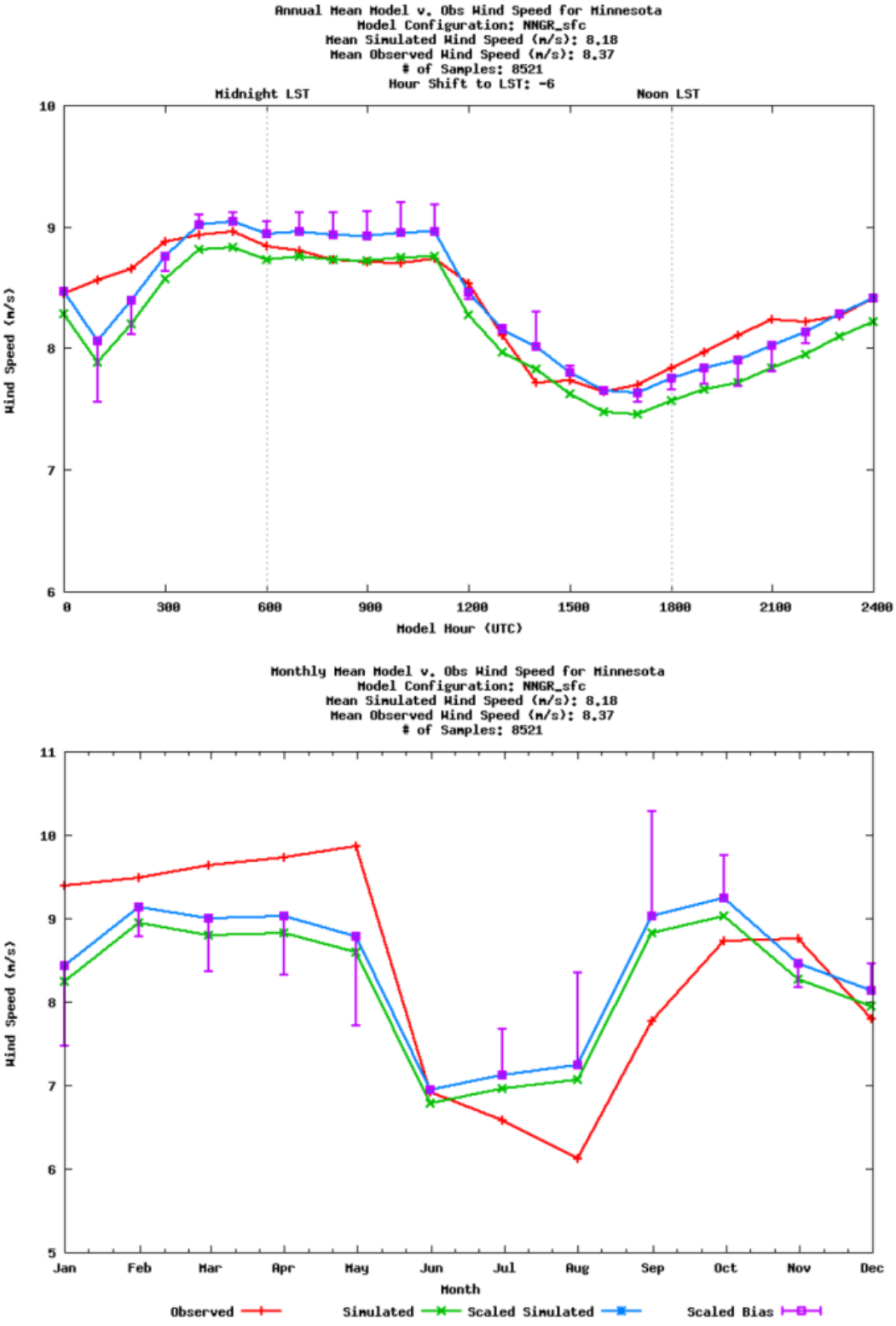


Figure 2.3. Same as Figure 2.1 for the Minnesota tower.

greater-than-normal discrepancy in some months at the Minnesota tower for all model configurations. AWS Truewind found no error in the data or analytical procedures, and also found a similar discrepancy at the nearest Automated Surface Observing Systems (ASOS) station. It was concluded that certain weeks in the validation data sample were poorly forecast at this location, and that the discrepancy would diminish with a larger sample of data.

In addition to making such qualitative comparisons, AWS Truewind calculated mean error (or bias) and mean absolute error (MAE) statistics for each tower and model configuration. It was primarily on the basis of the MAE for the scaled model speeds that the best configuration was chosen. Figure 2.4 compares the annual error statistics for all the configurations.

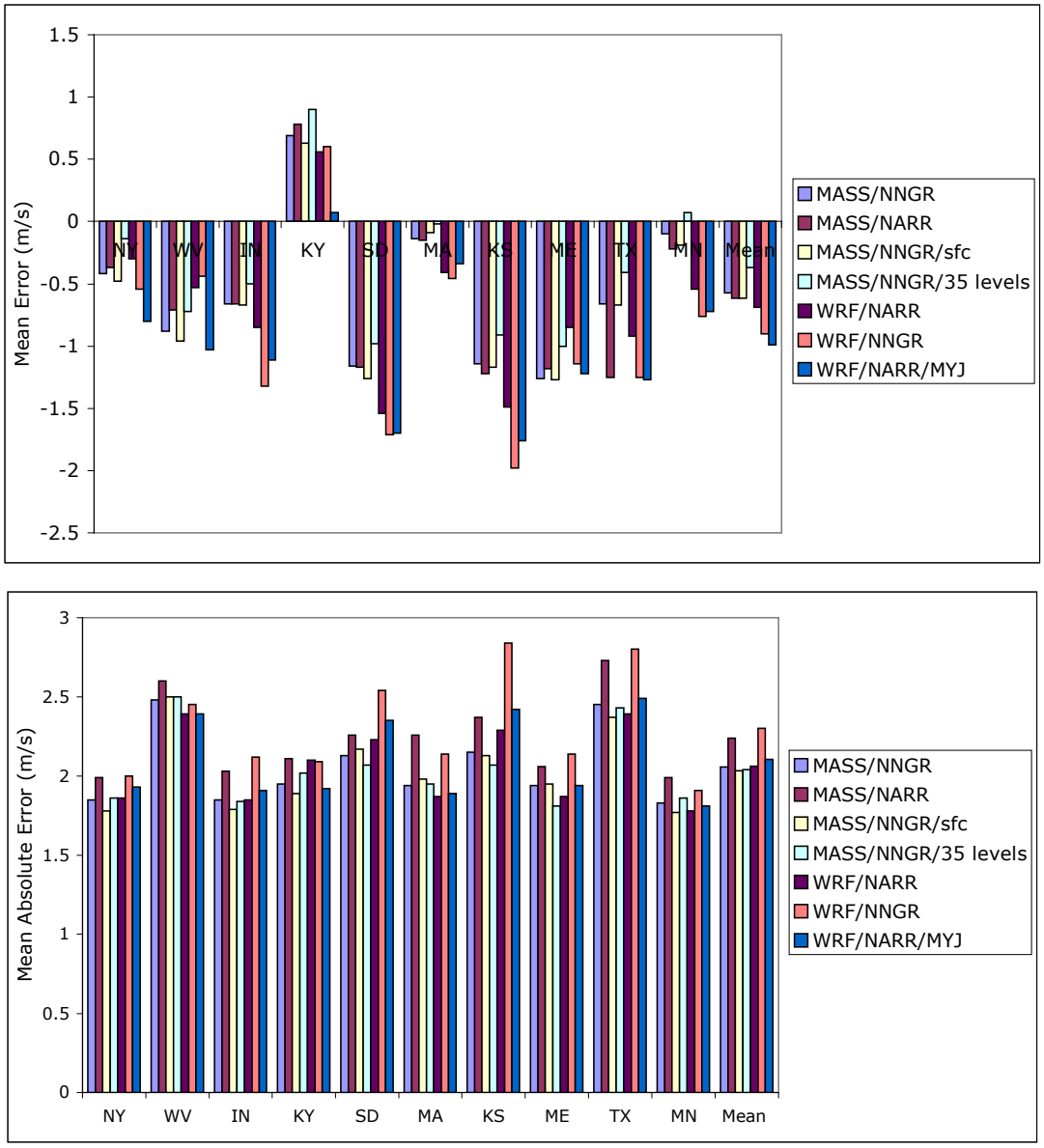


Figure 2.4. Mean error (top) and mean absolute error (bottom) over all simulations for each tower location and for all locations.

It is evident from these charts that while some model configurations performed relatively poorly at some sites (e.g., WRF/NNGR in South Dakota, Kansas, and Texas), most configurations exhibited similar errors at most sites. This reflects the underlying similarity of the models, which solve the same physical equations. It also suggests that most errors are due not to inadequacies in the model equations (e.g., turbulence parameterization) but to an insufficient amount of observational data available for initializing and updating the simulations.

2.1.3 Findings

By a small margin, it was found that the best configuration was MASS using NNGR as the initializing dataset and assimilating both surface data and rawinsonde data (i.e., MASS/NNGR/sfc).

MASS initiated with NARR data produced substantially worse results. It is suspected that this is because of an inconsistency in the definitions of land surface properties between the MASS model and the Eta model used to produce the NARR data. MASS with 35 levels did not significantly improve on MASS with the standard 25 levels, and was in some cases worse. WRF with NARR performed better than WRF with NNGR, and almost as well as MASS with NNGR. Little difference was observed in the quality of the results between the MYJ and YSU PBL schemes. Unfortunately, given the tight schedule for this project, it was not possible to configure WRF to assimilate rawinsonde and surface observations, as MASS is able to do. In addition, it was noted that the WRF model is not able to use more than four cores on a single eight-core workstation, thus making it effectively much slower than MASS. Using WRF would have therefore entailed significant additional delays in the project schedule.

2.2 Mesoscale Runs

The main runs were carried out over a five-month period from May to September 2008. The simulations covered the period 1 January 2004 to 31 December 2006. The following table summarizes the run configuration.

Table 2.3. Model configuration for main runs

Model	MASS v. 6.8
Initialization data source	NNGR
Data to be assimilated in the course of simulations (30 km and 8 km grids only)	Rawinsonde, surface observations (temperature, dew point, wind direction and speed, pressure)
Sea-surface temperatures	MODIS (Moderate Resolution Imaging Spectroradiometer)
High-resolution terrain and land cover (2 km grid only)	US Geological Survey National Elevation Dataset and Land-Use Land-Cover Dataset, both 30 m grid spacing
Cumulus scheme (30 km and 8 km grids only)	Kain-Fritsch
Spin-up	12 hours before start of valid run
Length of run	15-16 day series (e.g., Jan 1-15, Jan 16-31)
Frequency of data sampling	Every 10 min.
Data to be stored	U, V, temperature, pressure, TKE at five heights; surface temperature and pressure, specific humidity, incoming long-wave and short-wave radiation, precipitation

The runs employed a nested grid scheme with a horizontal resolution of 30 km for the parent grid and 8 km and 2 km for the child grids. The grid layout is shown in Figure 2.5. The largest black rectangle defines the 30-km parent grid (A). The three smaller black rectangles are the 8-km grids (B-1, B-2 and B-3), and the small red rectangles are the 2-km grids (C-KK).

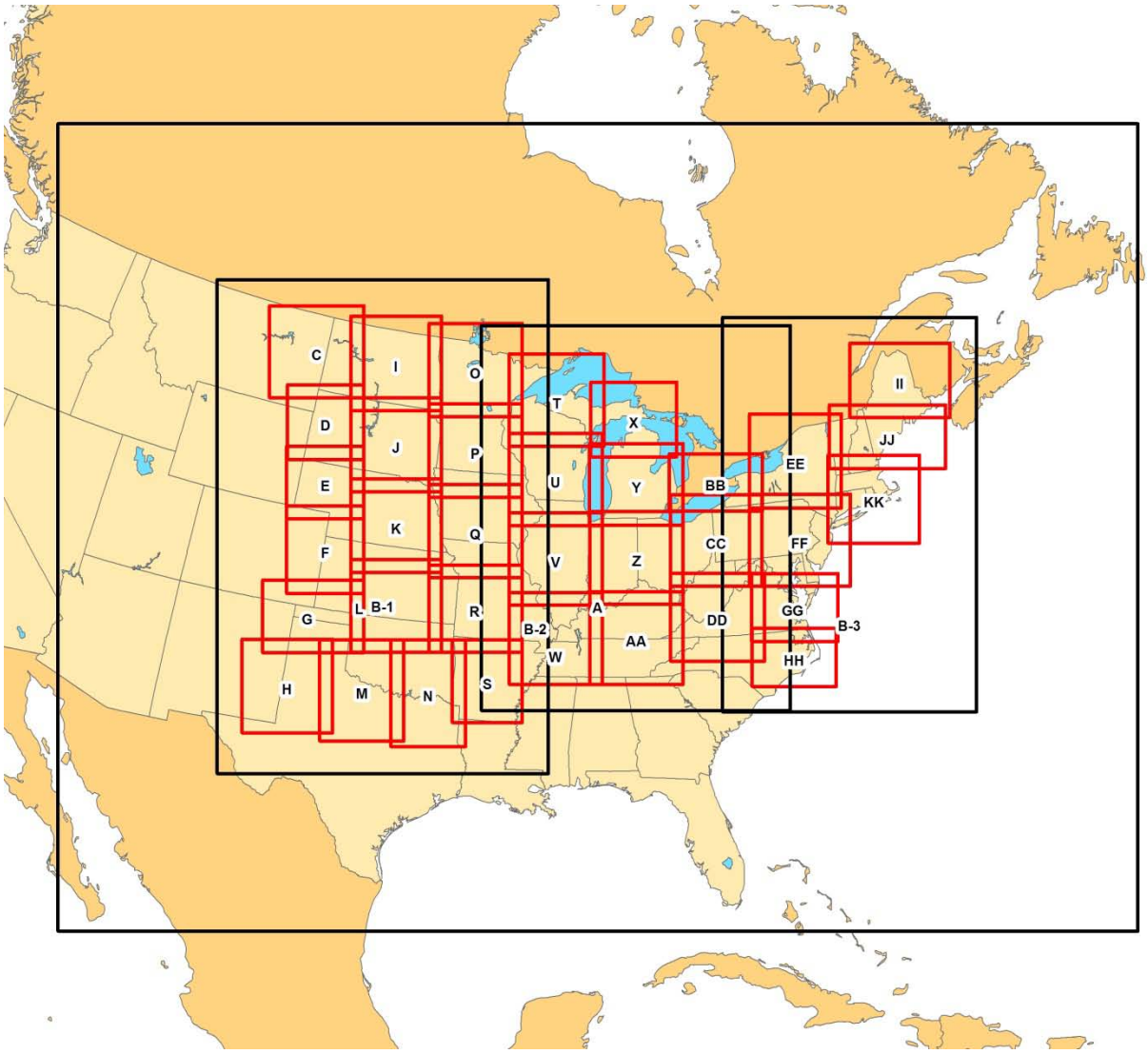


Figure 2.5. MASS grid definitions.

The simulations were carried out on a network of 640 cores on 80 dual-CPU quad-core Penguin workstations. In principle, each year should have taken about 21 days of real time on this system. However, inadequate cooling, network data congestion, and storage limits initially slowed progress to varying degrees. The cooling problem was solved with the installation of an additional 5-ton AC unit to supplement two existing 3-ton units. The network traffic issue was solved by writing more data to local scratch disks. Data storage was rectified by the installation of an additional 40 TB online RAID storage system to supplement the approximately 50 TB initially available; in addition, a high-speed tape backup system was installed to take unneeded data offline.

Once the 2004 runs were completed, three sets of mesoscale time-series files were extracted for review by AWS Truewind and NREL. The three datasets consisted of: (i) grid cells near each of the 10 tall towers used in the validation and selection of the model configuration; (ii) 22 grid cells randomly selected from 11 different grids; and (iii) grid cells associated with sites near the three existing project sites for which data were provided by NREL. The data for sets (i) and (ii) were forwarded directly to NREL. The data for (iii) were used to test the power conversion program (described below).

After NREL had reviewed and approved the sample files, the data extraction was carried out for all the grid cells associated with project sites. For each grid cell, four files were produced: (i) surface pressure, (ii) 2-m temperature, (iii) 80-m speed, direction, air density, and TKE, and (iv) 100-m speed, direction, air density, and TKE. Each file for 2004 (a leap year) contains 52704 records spanning the year in 10-minute increments. The naming convention is as follows: GRID_XX_IIIJJJ_HHHH.TXT, where XX refers to the grid number (from one to 36, except 18, one of the B grids, which was mistakenly included in the grid numbering), III and JJJ to the cell column and row numbers, and HHHH to the height (0000M = surface pressure, 0002M = 2-m temperature, 0080M = 80-m speed, and 0100M = 100-m speed). An example file is provided in Table 2.4. These files (which for each year number close to 120,000 and occupy about 300 GB of disk space) were not forwarded to NREL, but may be provided at a future date.

Table 2.4. Example time-series file for the 80-m wind at a single grid cell. The first row indicates the latitude and longitude of the grid cell. The other rows contain the data in the following columns: (i) date in YYYYMMDD format, (ii) time (GMT) in HHMM format, (iii) speed in m/s, (iv) direction in degrees from true north, (v) air density in kg/m^3, and (vi) TKE in m^2/s^2.

47.02225	-68.80990				
20040101	0010	4.89790	270.09622	1.25625	0.03305
20040101	0020	4.94108	268.34360	1.25527	0.02336
20040101	0030	4.81025	267.33597	1.25509	0.01175
20040101	0040	4.89001	267.15210	1.25468	0.00649
20040101	0050	4.55865	265.43286	1.25415	0.00473
20040101	0100	4.70651	265.82401	1.25480	0.00252
20040101	0110	4.84289	269.14575	1.25461	0.00214
20040101	0120	4.85045	266.78668	1.25462	0.00247
20040101	0130	4.76209	266.21219	1.25440	0.00268
20040101	0140	4.74387	263.26474	1.25424	0.00220
20040101	0150	4.89790	260.24161	1.25384	0.00246
20040101	0200	4.93185	256.34119	1.25321	0.00351
20040101	0210	4.87496	252.86868	1.25324	0.00413

The 2004 runs were completed by July 2008. With space and cooling issues resolved, the pace of production accelerated, and 2005 was completed in August and 2006 in September. For each year, a set of data files similar to that shown in Table 2.4 was extracted.

3 Selection of Wind Sites

3.1 Onshore Sites

The goal of this task was to select a large number of potential wind project sites with a total rated capacity of at least 300 gigawatts (GW). For this purpose, a site is defined as a near-contiguous area meeting certain exclusion criteria (described below) and having a locally maximum wind resource. Initially, roughly two-thirds of the capacity was to be onshore and one-third offshore. After conversations with NREL, the desired rated capacity was increased and specific minimum targets for each state (onshore and offshore) were established.

To provide a consistent set of resource estimates for ranking and selecting sites, a seamless map of predicted mean wind speeds at 80 m height for the EWITS region was prepared from AWS Truewind's proprietary wind maps. The seamless map has a horizontal spatial resolution of 200 m, which is sufficiently fine to reflect the influence of most terrain features and to identify specific locations for wind projects. AWS Truewind has developed a method of adjusting its wind maps using a wide array of wind resource measurements to ensure good accuracy. After considerable discussion, it was decided that AWS Truewind would use a version of its maps that had been adjusted using only publicly available wind data.

A map of the estimated net capacity factor (CF) for a composite IEC Class 2 wind turbine was then created using the seamless wind speed map and speed-frequency distributions compiled from 10 years of historical mesoscale model runs previously performed by AWS Truewind at 20 km resolution. Although IEC Class 2 turbines are not suitable for every site, the use of a single curve allowed for an objective ranking of resource potential. The composite power curve was created by taking the average of three commercial megawatt-class wind turbine power curves which had been normalized to their rated capacity. The normalized average curve was then rescaled to a rated capacity of 2 MW. The composite curve is shown in Table 3.1.

Maps of exclusion areas, including parks, wetlands, and urban areas, were created as well. The site screening took into account the following exclusions:

- From USGS National Land Cover Database (2001):
  - Open Water
  - 200-m buffer of Developed Low Intensity
  - 500-m buffer of Developed Medium Intensity
  - 500-m buffer of Developed High Intensity
  - Woody Wetlands
  - Emergent Herbaceous Wetland

- From ESRI data base:
  - Parks
  - Parks Detailed
  - Federal Lands (non-public)
  - 10,000-ft buffer of small airports (all hub sizes)
  - 20,000-ft buffer of large airports (hub sizes; medium and large)

Table 3.1. Composite power curves for IEC Class 1, 2, and 3 turbines at standard air density (1.225 kg/m^3).

Speed	Power (kW)		
(m/s)	IEC-1	IEC-2	IEC-3
0	0	0	0
1	0	0	0
2	0	0	0
3	0	0	12.6
4	39	56.6	82.4
5	136.2	176.8	204
6	280.2	347.8	378
7	474.2	574.6	621.4
8	732.6	867.8	943
9	1046.6	1213.2	1325.8
10	1404.2	1553.6	1676.6
11	1712.8	1810	1892.8
12	1911.2	1943.4	1974.2
13	1974.8	1985.2	1995.2
14	1989	1995.8	1999
15	1996.4	1999.6	1999.8
16	1998	2000	2000
17	2000	2000	2000
18	2000	2000	2000
19	2000	2000	2000
20	2000	2000	2000
21	2000	2000	2000
22	2000	2000	0
23	2000	2000	0
24	2000	2000	0
25	2000	2000	0

- Other:
  - Slopes greater than 20%
  - Areas outside the study region

At NREL's request, data from the Conservation Biology Institute (CBI) were added. The excluded layers had GAPCAT values of 1, 2, 7, and 8, which are defined as areas managed to maintain a natural state or for natural values and managed for conservation to varying degrees. Many of the areas defined in the CBI data base coincided with exclusions already defined in the other data bases. In those cases, no changes in the exclusions occurred.

The site screening occurs in two steps. In the first, the program finds all sites with a maximum output in the immediate vicinity (i.e., a local maximum) with sufficient area to support a project of the desired 100-MW rated capacity and no closer than 2 km to any neighboring site. In the second step, the program allows each of these 100 MW sites to expand so long as the output does not decrease by more than 5%. If the site encounters another site, the site that has a higher mean output is retained and the other is dropped.

In the first round of site screening, over 2000 GW of sites with a maximum capacity of 2000 MW each were selected. Upon review by NREL, it was concluded that many of the sites, particularly in the western part of the study region, were too large. At NREL's request, a second round of site screening was done, with the maximum project size set to 1000 MW. This was also deemed unsatisfactory. AWS Truewind then modified its site-screening program so that it would select sites with a range of rated capacities, even in areas where very large sites could be supported. This final approach, which produced over 7800 sites totaling over 3000 GW of rated capacity, was deemed acceptable by NREL.

An additional concern raised by NREL was the small number of sites chosen in four states: Connecticut, Rhode Island, New Jersey, and Delaware. To address this concern, AWS Truewind ran a separate site screening with a lower capacity factor threshold for those states alone, and the sites chosen were added to the others. In the last step, NREL manually selected a subset of the sites to arrive at the final site selection. This manual selection was done to ensure that there was enough wind capacity in different states and regions for the development a diverse set of wind scenarios for the EWITS analysis. The criteria used for this selection included wind capacity factor and geographic diversity. The goal was to have a "super set" of wind sites (579 GWs) from which lesser amounts of wind sites could be chosen to build the 20% and 30% wind scenarios.

The final site list contains 1326 sites totaling 580 GW. Figure 3.1 shows the locations of the sites, color-coded by predicted annual mean capacity factor, and Table 3.2 summarizes their distribution by state. In the Great Plains, the predicted capacity factor typically exceeds 40%, whereas in most eastern states it is below 30%. Figure 3.2 shows the distribution by rated capacity. The bulk of the sites fall between 100 MW and 600 MW in size. A small number of "megasites" with rated capacities exceeding 1000 MW were also chosen. All of these are in the Great Plains.

Table 3.2. State totals of onshore and offshore sites. Sizes of onshore sites range from 100 MW to over 1000 MW. Each offshore site is 20 MW.

	Onshore		Offshore	
State	Count	Total MW	Count	Total MW
Arkansas	20	4049		
Colorado	8	3760		
Connecticut	7	919	84	1680
Delaware	7	1018	221	4420
Illinois	79	42029	99	1980
Indiana	63	32591	112	2240
Iowa	92	52575		
Kansas	66	46069		
Kentucky	6	1490		
Maine	42	5863	64	1280
Maryland	9	1114	496	9920
Massachusetts	19	2166	1006	20120
Michigan	57	23944	788	15760
Minnesota	121	61480		
Missouri	19	10138		
Montana	12	5830		
Nebraska	89	48471		
New Hampshire	21	2371	1	20
New Jersey	8	1328	1343	26860
New Mexico	24	10524		
New York	66	14859	502	10040
North Carolina	10	1999	2244	44880
North Dakota	60	32138		
Ohio	34	17445	1479	29580
Oklahoma	82	40253		
Pennsylvania	56	6988	257	5140
Rhode Island	7	1039	65	1300
South Dakota	91	48547		
Tennessee	8	886		
Texas	48	31896		
Vermont	17	2019		
Virginia	16	2097	1507	30140
West Virginia	18	2376		
Wisconsin	44	20494	162	3240
Total	1326	580762.5	10430	208600

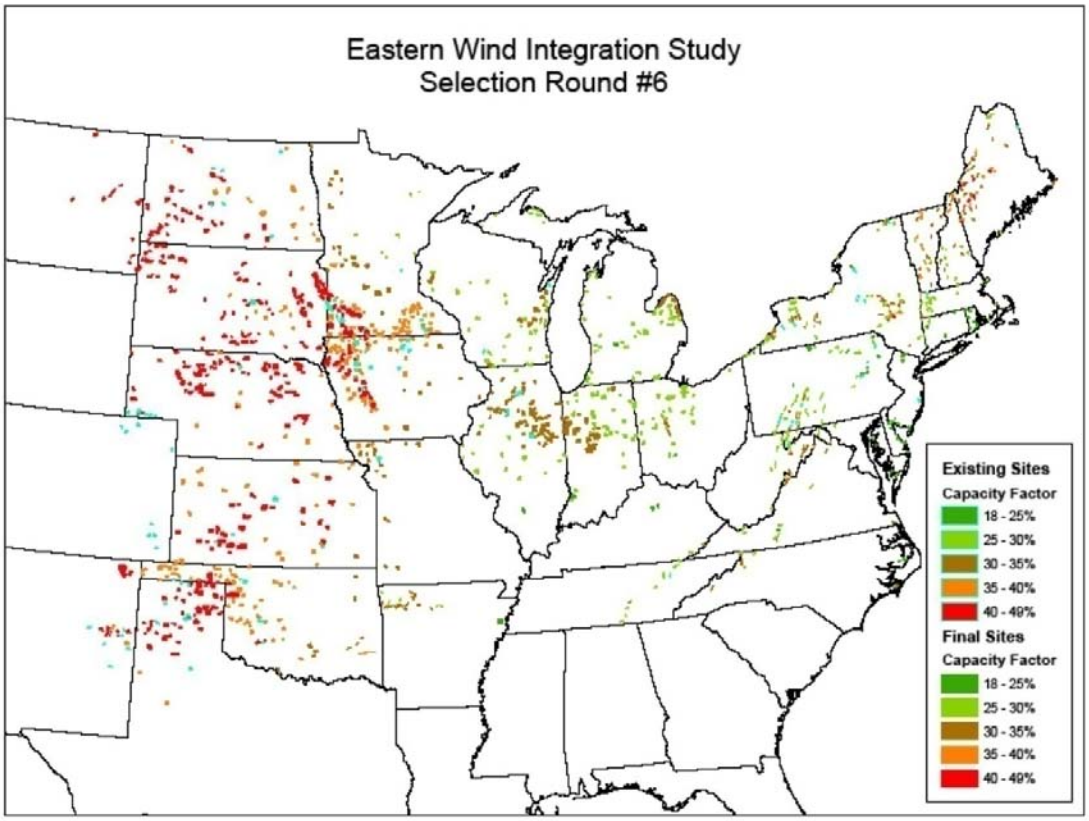


Figure 3.1. Locations of 1326 onshore sites. Light blue points show the locations of existing projects.

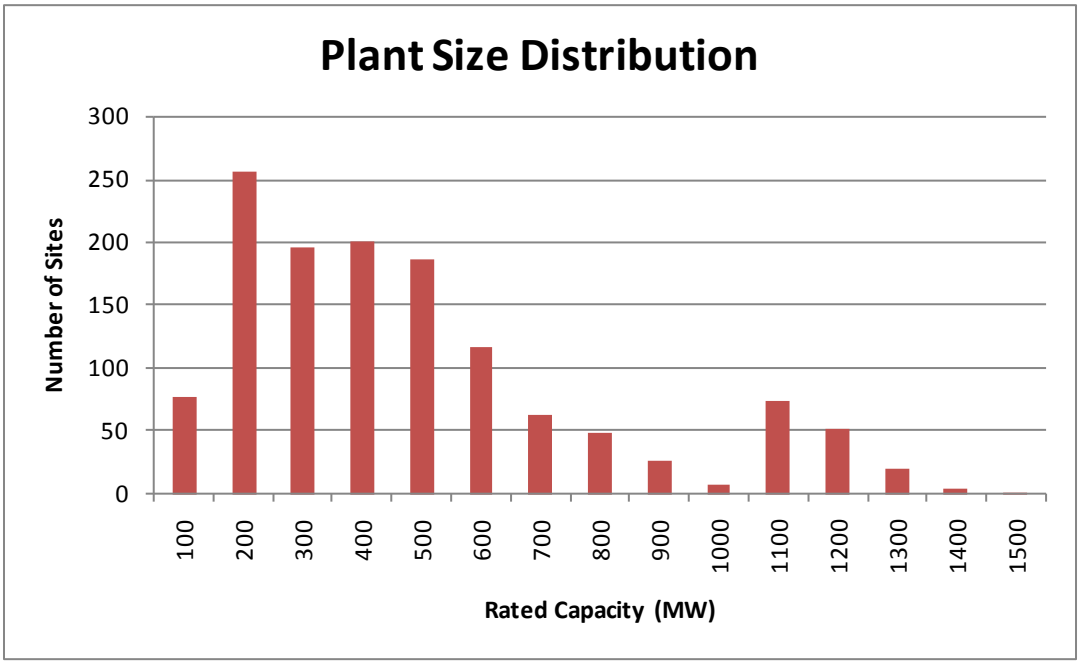


Figure 3.2. Distribution of plant sizes for 1326 onshore sites.

3.2 Offshore Sites

For the offshore site screening, NREL requested sites with a maximum water depth of 30 m and minimum distance from shore of 5 miles (8 km). In the initial round of offshore site screening, AWS Truewind selected all offshore areas meeting these criteria that were also outside federal or state protected areas and had an estimated net annual CF of at least 32%. The CF was determined from the mean speed at 80 m from AWS Truewind's wind map, speed distributions from AWS Truewind's historical 20-km resolution simulations, and the default composite IEC Class II wind turbine power curve. (The net CF is not meant to be realistic for offshore projects but was used for ranking purposes only.) The selected areas were forwarded to NREL in shape file (GIS) format.

After NREL's review, the selected regions were converted to a list of offshore "sites." The sites were presented somewhat differently from the onshore sites. Because there is so little spatial variation in the wind resource well offshore, rather than attempt to identify contiguous areas meeting a minimum capacity threshold such as 100 MW, it was decided to select points spaced approximately 2 km apart. Each point represents roughly a single mesoscale grid cell 4 km^2 in area, which is assumed to be capable of supporting 20 MW of offshore wind capacity (i.e., a mean density of 5 MW/km^2).

The list included just over 16,000 such points. In addition to latitude and longitude and nearest state, each point was characterized by the predicted mean speed at 80 m and 100 m and the predicted net CF at 80 m height. Included in the spreadsheet was a running sum of the gigawatts of wind capacity associated with each state, in order of decreasing net CF, and a table in a separate worksheet showing the grand total for each state.

In its review, NREL detected inconsistencies between the estimated site mean speed and net CF. Some sites had a lower speed than was apparently justified by the CF. AWS Truewind found that the wrong speed map had been used when extracting the speeds for the table. In addition, AWS Truewind discovered that sites closer than 5 miles (8 km) from the Atlantic coast had been mistakenly retained in the list. Both problems were corrected, and a new table containing about 10,000 offshore sites was sent to NREL. The distribution by nearest state is shown in Table 3.2. A map of the offshore sites color-coded by CF is shown in Figure 3.3.

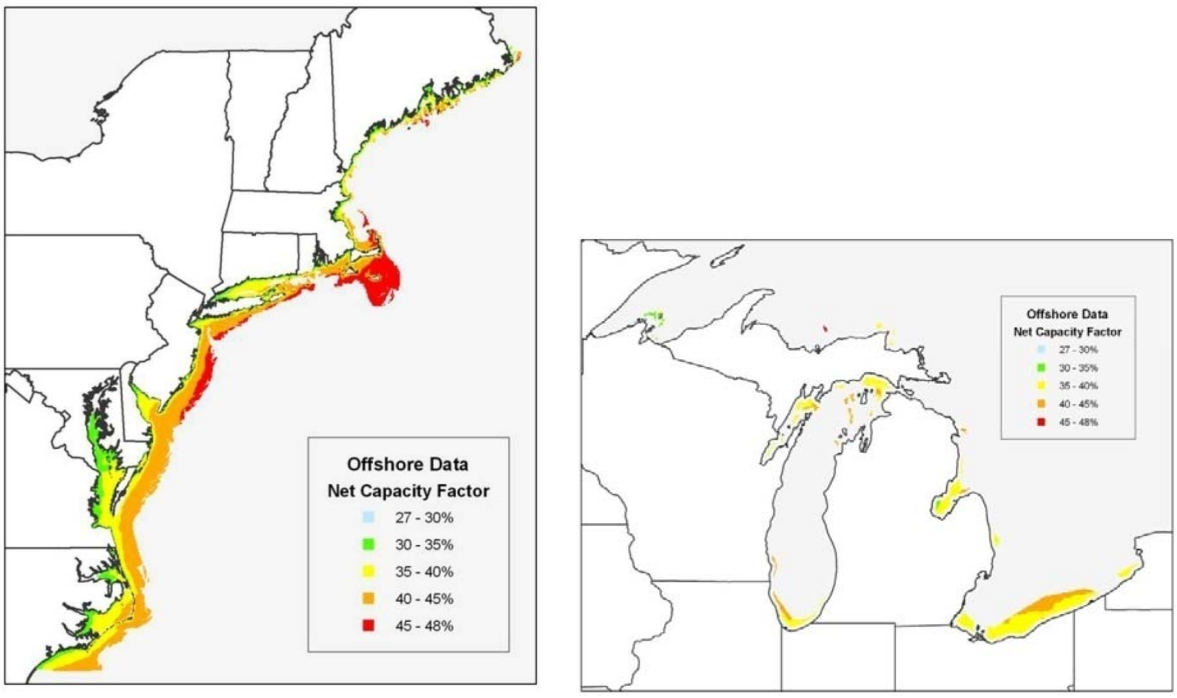


Figure 3.3. Offshore sites.

4 Generate Wind Plant Output

4.1 Conversion Procedure

Converting the meteorological data generated by the mesoscale model to wind plant output was done by a program written by AWS Truewind called SynOutput.

The program starts by reading a list of the 10 validation towers and their nearest associated grid cells (grid number and column and row position). It also reads a list of the grid cells associated with the sites in the 580-GW scenario. Up to several dozen grid cells are associated with each site, depending on its size and shape. For each cell, the list provides the latitude and longitude, expected mean speed of the part occupied by turbines, mesoscale elevation, actual mean elevation of the turbines, and relative proportion of the site's total rated capacity associated with that cell. [1] The mean speeds are based on AWS Truewind's wind maps adjusted to the year of the simulation. Since each year was produced separately, it was not possible to scale the mean speeds for all three years to the expected 2004-2006 mean at each site. AWS Truewind proposed that its 20-km historical dataset be used to establish each year's mean relative to the 1997-2007 period. It is expected that this procedure will result in only minimal error. NREL accepted this proposal.

[1] Initially there was an error in the code that truncated the proportion of the rated capacity in a grid cell to an integer. This caused a systematic underestimation of the total production for a given rated capacity. The error was corrected for the final dataset.

The program then imports the turbine power curves. There is one power curve for each IEC class. (The curves are shown in Table 2.1.) The power curves are scaled to a rated capacity of 2 MW and are valid for the standard sea-level air density of 1.225 kg/m^3. The IEC 1 and 2 curves are based on a composite of three commercial turbines (GE, Vestas, and Gamesa brands). In consultation with NREL, it was decided to base the IEC 3 curve on just two turbines (GE 1.5xle and Gamesa G90) to avoid an inconsistency in the cut-out speed of the Vestas V100. In addition, the cut-out speed of the GE turbine was changed from 20 to 21 m/s to match that of the Gamesa turbine. The IEC 1 and 2 turbines are assumed to have a hub height of 80 m and the IEC 3 turbine 100 m.

The program next reads a set of 12x24 speed matrixes, one for each of the 10 validation towers. These matrixes give the mean speed for each hour of the day and for each month of the year. For each tower there are two matrices, one for each hub height (80 m and 100 m). The program reads the mesoscale time series file for each of the grid cells nearest the validation towers. From the speed data, it creates a 12x24 mean speed matrix for each hub height. The ratio between the average observed speed and the average simulated speed is then calculated for each bin and normalized to an average of one. The result is an adjustment matrix which is used to correct model biases. Although the program calculates adjustments on a monthly basis, it was found during the validation phase that the monthly variation in speeds was accurately predicted by the model. Therefore, only an annual adjustment is actually performed.

The mesoscale time series file for each grid cell associated with a project site are then read. The speed data are scaled to match the expected mean speed and finally summed for all the grid cells associated with the site. In the sum, each cell's speeds are weighted according to the proportion of the site area associated with that cell. The result is a time series of simulated wind speeds for the site as a whole at both 80 m and 100 m.

The program 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 calculates a weighted average adjustment matrix for the site in which the weight given to the adjustment matrix for each validation location is proportional to its correlation coefficient. The program applies this blended adjustment matrix to the simulated data for the site. For example, if the time in question is 1300, the simulated speed is multiplied by the adjustment factor for 1300.

The speeds for each grid cell are then adjusted for wake losses in a manner that depends on the simulated wind direction relative to the prevailing (most frequent) direction. The loss is given by $w = w_{\min} + (w_{\max} - w_{\min})\sin^2(\theta - \theta_{\max})$, where $w_{\min}$ is the minimum loss (assumed to be 4%) when the wind is aligned with or opposite to the prevailing direction $\theta_{\max}$, and $w_{\max}$ is the maximum loss (9%) when the wind is perpendicular to the prevailing direction. The loss factors – which were arrived at by trial and error to conform with AWS Truewind's estimates for actual wind projects – account both for wake losses and implicitly for other losses such as blade soiling that affect the efficiency of power conversion for a given free-stream speed without reducing the maximum output. The method does not account for sites where there is more than one prevailing wind direction or where the prevailing energy-producing direction differs from the most frequent direction. The speed is further adjusted by adding a random factor (from -1 to +1) multiplied by the predicted TKE. This adjustment is intended to reflect the impact of gusts on the speeds experienced by the turbines in the wind project. The frequency and intensity of such simulated

gusts depends to a degree on time of day, as TKE is generally higher in the day, when the PBL is thermally unstable or neutral, than at night, when it is thermally stable.

Next, the adjusted speeds are applied to the turbine power curve for each IEC class. In the process, the power curve is corrected to the predicted air density. A time filter is then applied to mimic the effect of spatial averaging on the fluctuations of wind output over the area of a mesoscale grid cell. The time filter gives a weight of 90% to the predicted output at the current time, and divides the remaining 10% weight equally among the predicted output values of the previous 17 time records (i.e., about 3 hours of actual time). This approach has been found to reproduce the observed "variability" of wind plant output, as measured by the mean absolute deviation (MAD) as a function of time, with reasonable accuracy. Without such time filtering, the simulated plant output would tend to be more variable than the observed.

The program applies an additional power loss to account for turbine and plant availability. Based on data obtained by AWS Truewind for operating wind projects, the availability is assumed to follow a normal distribution with a mean of 94.8% and a standard deviation of 2.3%; the distribution is truncated at 100%. To avoid unrealistic rapid fluctuations in output, the availability is allowed to change at random intervals averaging only once per hour. An additional loss of 3% is subtracted from the output to represent electrical losses. For most sites, the total loss (relative to the gross production in the absence of any loss-related adjustments to speed or direct plant losses) is about 15%-17%. The range is from about 12% to about 20%.

For each site, the simulated speeds at both hub heights and the power output for all IEC classes are output to a single text file. In addition, the program selects the most appropriate IEC class based on the estimated maximum long-term annual average mean speed within the site, adjusted for air density. The power output for the selected IEC class is provided in the last column of the file. In this way, users of the data can simply import the last column rather than have to select themselves which IEC class to use. A sample text file is shown in Table 4.1.

Table 4.1. Sample plant-output data file.

SITE NUMBER: 4403 RATED CAP: 100.2 IEC CLASS: 3 LOSSES (%): 19.5 18.6 17.0							
Date (GMT)	Time (GMT)	Speed (80 m)	Speed (100 m)	IEC1 (80 m)	IEC2 (80 m)	IEC3 (100 m)	Selected
20060101	10	3.876	4.062	1.2	1.8	3.3	3.3
20060101	20	3.982	4.314	1.4	2.0	4.2	4.2
20060101	30	4.844	5.204	4.0	5.3	8.3	8.3
20060101	40	5.549	5.929	7.2	9.2	12.8	12.8
20060101	50	5.999	6.471	9.8	12.2	17.0	17.0
20060101	100	6.242	6.785	11.2	14.0	20.0	20.0
20060101	110	6.085	6.687	10.1	12.7	18.7	18.7
20060101	120	5.653	6.315	7.9	10.0	15.9	15.9
20060101	130	4.852	5.563	4.1	5.4	10.5	10.5
20060101	140	3.870	4.592	1.6	2.2	5.9	5.9
…	….	….	….	….	….	….	….

4.2 Validation

NREL provided three years of 10-minute plant output data for three projects within the EWITS area: Lake Benton (MN), Storm Lake (IA), and Blue Canyon (OK). (A fourth project, Trent Mesa, TX, is outside the EWITS area and was therefore not used.) The power conversion program was tested using mesoscale data extracted for grid cells associated with the three projects, and the results were provided to NREL both in tabular form and as diurnal and change plots. A comparison of the 10-minute output for a typical period at Blue Canyon is provided in Figure 4.1.

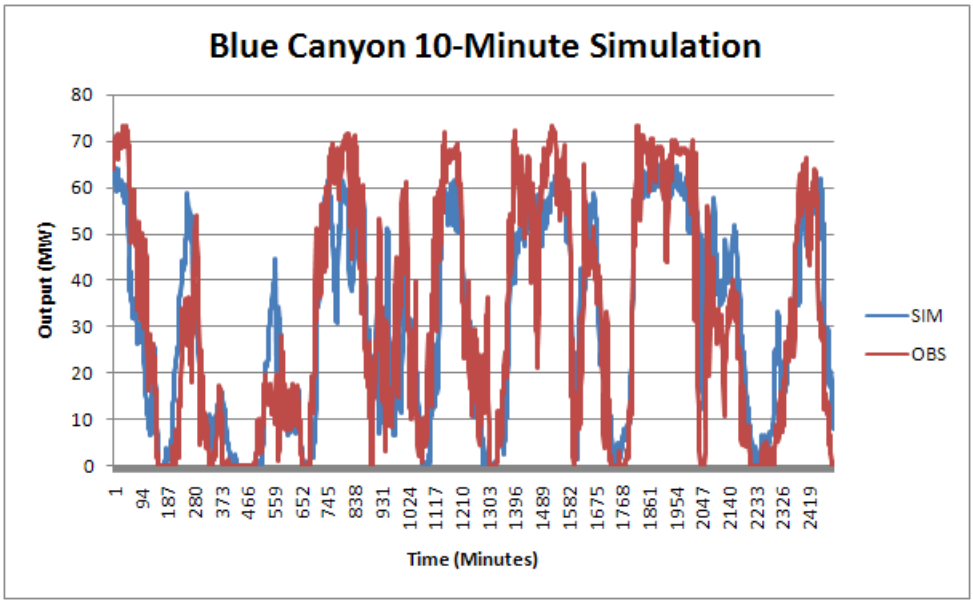


Figure 4.1. Comparison of simulated (SIM) and observed (OBS) 10-minute plant output at the Blue Canyon I project. The plot covers about 18 days.

By using an appropriate turbine power curve for each project and by adjusting the hub height of the simulated speeds and diurnal corrections to match the actual hub height, a reasonably close agreement between the predicted and observed net capacity factors and diurnal patterns was obtained (Figure 4.2). The mean absolute changes in output as a function of time lapse, an indication of the variability of output over a range of time scales, were also compared. The curves for the simulated and observed wind plant output follow similar trajectories (Figure 4.3).

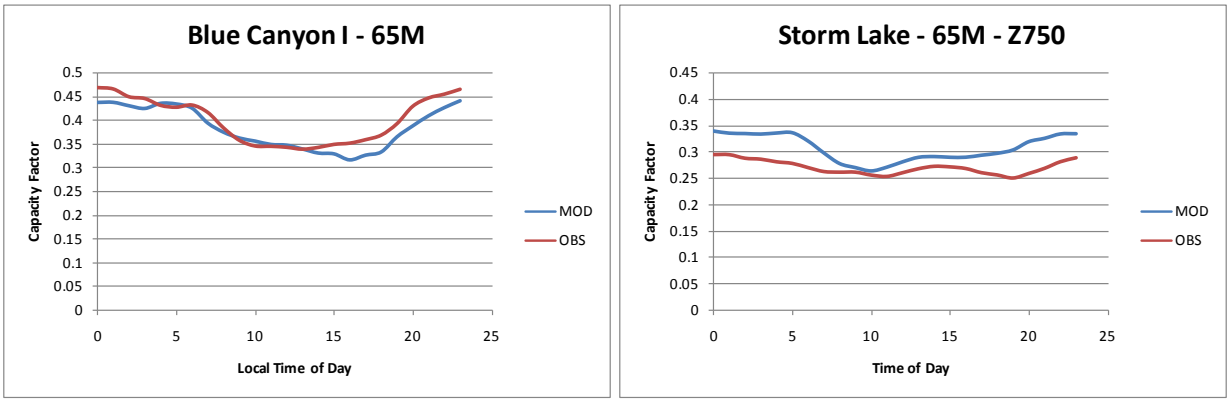


Figure 4.2. Comparison of simulated and observed diurnal mean wind speeds for two validation projects, Blue Canyon I (left) and Storm Lake (right).

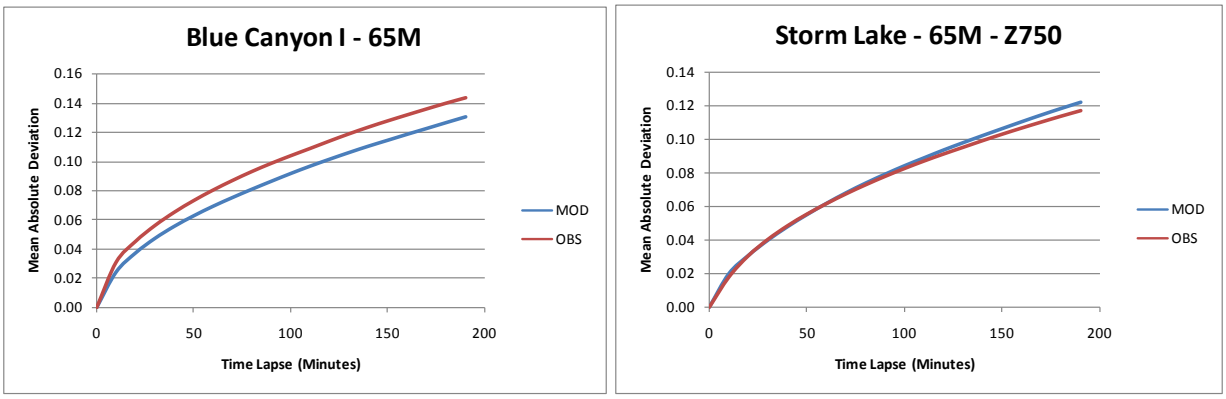


Figure 4.3. Comparison of simulated and observed mean changes in plant output as a function of time lapse for Blue Canyon I (left) and Storm Lake (right).

Figure 4.4 compares the modeled and observed ramp frequencies (step change frequency as a function of step change size) for Blue Canyon I for 10-minute and 60-minute intervals. The profiles have a very similar shape. The importance of distinguishing between averaged and sampled data is illustrated by the difference between the observed 60-minute average and sample curves. The modeled data are intended to represent sampled plant output, which is more variable than output averaged over each hour.

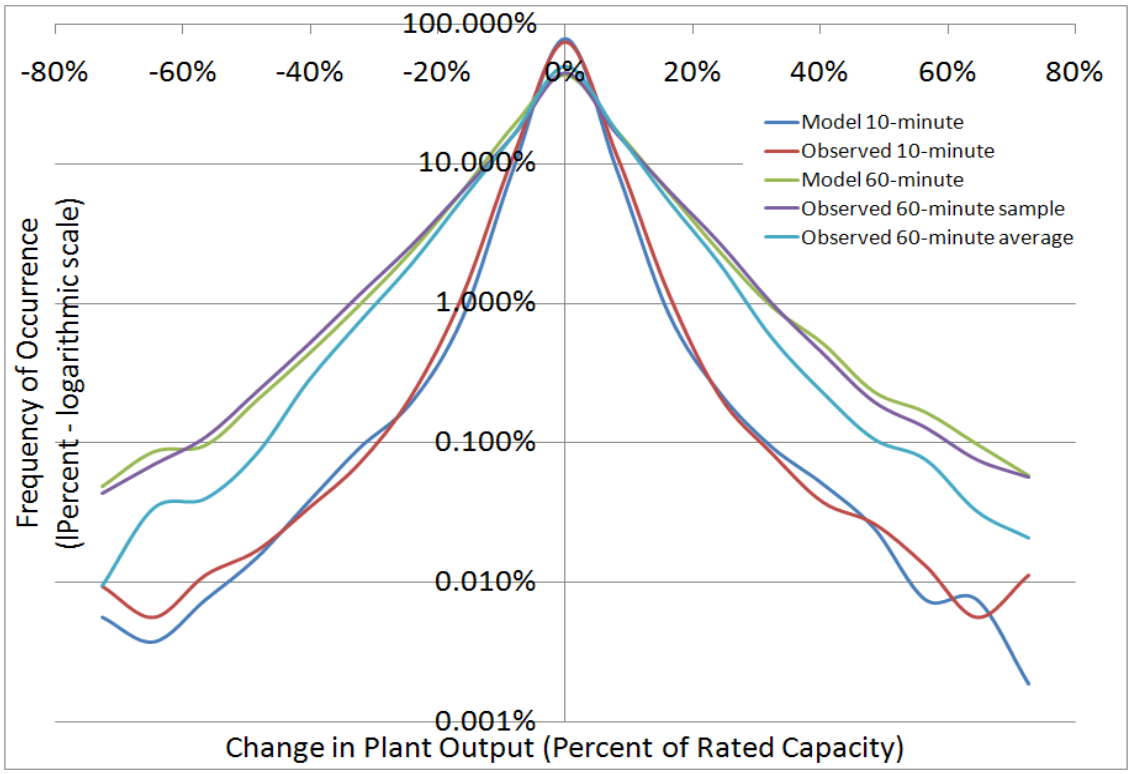


Figure 4.4. Comparison of simulated and observed step-change frequency for Blue Canyon I for 10-minute and 60-minute time intervals. Averaging over a 60-minute period significantly reduces the variability of plant output, as shown by the blue curve.

In reviewing the plant output files, MISO noted significant jumps at 0000 and 1200 UTC on many of the days. AWS Truewind concluded that the jumps were caused by the abrupt assimilation of rawinsonde and surface observations every 12 hours in the mesoscale runs. A workaround was developed, in which values spanning the affected times were replaced with synthesized data (Figure 4.5). Two other errors in the data processing were discovered at the same time, and a new set of data was produced. With these revisions, the final plant output data files were regenerated and delivered to NREL.

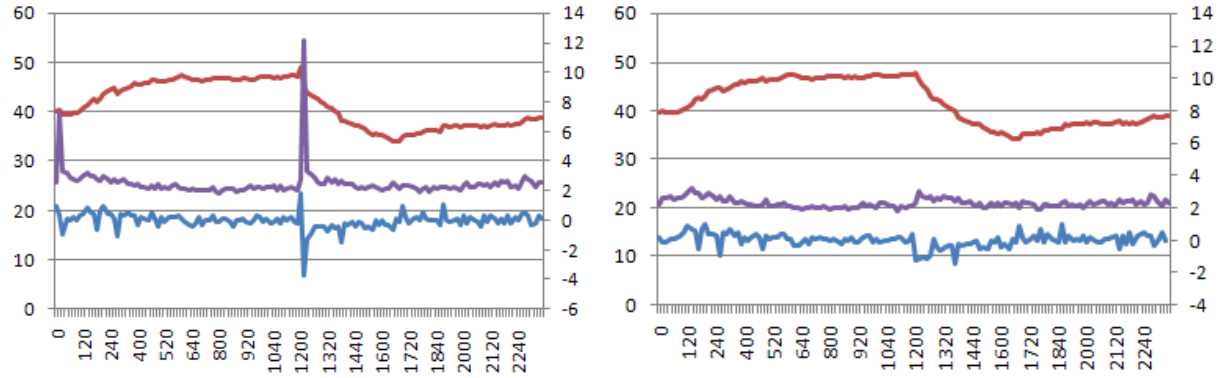


Figure 4.5. Jumps in mean speed at one site before (left) and after (right) the fix. The red curve is the mean output (left axis), the purple curve is the absolute change in output from one 10-minute record to the next (left axis), and the blue curve is the change in output (right axis).

5 Forecasts

AWS Truewind produced hourly forecasts for three different time horizons: next-day, six-hour, and four-hour. Each set of forecasts was synthesized by running a statistical forecast synthesis tool called SynForecast, which was developed by AWS Truewind. This tool uses actual forecasts and observed plant output to develop a set of transition probabilities. The probabilities are applied to simulated plant output data, stepping forward in time from a random starting point in a process known as a Markov chain. This process results in a synthetic forecast that imitates the statistical behavior of a real forecast. The procedure is described in depth in the following section.

5.1 Forecast Synthesis Procedure

The first step in the forecast synthesis process is to produce a sequence of real forecasts for one or more operating wind projects using a state-of-the-art wind forecasting system. It is assumed that these forecasts are typical or representative of what forecasts would look like at other sites in the region. For this purpose, AWS Truewind ran its eWind forecasting system[2] in "hindcast" mode for the four wind plants for which NREL had previously provided output data for 2004-2006: Trent Mesa, Blue Canyon, Lake Benton, and Storm Lake. The mesoscale model feed for

[2] eWind is a commercial wind forecasting service. It takes as its main inputs weather forecasts from a mesoscale weather model and actual plant operating data; on-site wind observations are also often used. Over a period of several days or weeks, eWind builds a statistical model relating the forecasted plant output to the actual plant output. This model is then applied to correct the wind forecasts going forward. Over time, the model "learns" from past forecast errors, and its skill gradually improves.

the forecasts was provided by 8-km resolution MASS simulations. The observed data feed was provided by the actual plant data up to the time each forecast was assumed to be generated. For the next-day forecasts, this was 5 pm local time of the day before the forecast; this effectively assumes the forecasts are generated once per day. For the four-hour and six-hour forecasts, the latest time was four hours and six hours ahead of the delivery time, which implies an hourly update schedule.

Since no on-site wind measurements were available, the eWind statistical module transformed the mesoscale model data directly into plant output by constructing a non-linear power curve based on a rolling 30-day window of simulated and observed data. The forecasts were issued in 24-hour periods from midnight to midnight of each forecast day and appended to one another to form a continuous time series of forecasts.

From each of these four sets 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_i 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, both the actual and forecasted output values are 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. Both the current and previous forecasts are included in the probability matrix to capture the autocorrelation of forecast errors, as otherwise the synthesized forecasts would fluctuate randomly about the actual output in an unrealistic fashion.

For each wind project site, the SynForecast program selected, at random, one of the four transition probability matrixes. Both onshore and offshore projects made use of the same four matrixes; given the lack of offshore plant output data, it is not known whether forecast skill will be similar for offshore projects as for onshore projects. 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.

5.2 Validation

To verify that the program was working properly, AWS Truewind compared synthesized forecasts with the actual forecasts for the four validation wind projects. First, the time correlation of the actual and forecasted generation and the root-mean-square (RMS) forecast error were considered for next-day forecasts. The results of the actual and synthesized forecasts were very similar, as shown in Table 5.1. The RMS error depends in part on the average plant output, with more productive plants experiencing higher forecast errors as a fraction of rated capacity because they spend more time in the steeply sloping parts of their power curves.

Next, the autocorrelation of the output, the 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. The following tables indicate that the observed output of the four wind projects tends to be quite strongly autocorrelated over a period of one to several hours. The eWind and synthesized

forecasts exhibit similar degrees of autocorrelation in each case. As for the forecast errors, their autocorrelation is considerably lower, and the SynForecast program seems to capture the pattern of decreasing correlation with increasing time shift quite well. However, the synthesized forecasts are slightly less correlated in time than the eWind forecasts, while the synthesized forecast errors are slightly more correlated.

Table 5.1. Comparison of correlation of forecasted and actual output and RMS forecast error for synthesized and real (eWind) forecasts.

	Correlation (Pearson *r*)		RMS Forecast Error (CF)	
Plant	eWind	SynFcst	eWind	SynFcst
Trent Mesa	0.77	0.77	0.20	0.20
Blue Canyon	0.77	0.73	0.21	0.22
Storm Lake	0.79	0.81	0.16	0.16
Lake Benton	0.72	0.71	0.19	0.19

Table 5.2a. Autocorrelation of actual output, eWind forecasts, and synthesized forecasts and forecast errors: Trent Mesa

Time Shift	Obs	eWind	SynFcst	eWind Error	SynFcst Error
0	1	1	1	1	1
1	0.93	0.94	0.92	0.78	0.81
2	0.85	0.87	0.85	0.59	0.65
3	0.77	0.80	0.79	0.46	0.51
4	0.70	0.72	0.72	0.37	0.39
5	0.63	0.65	0.66	0.30	0.30
6	0.56	0.57	0.60	0.26	0.22
7	0.50	0.50	0.55	0.22	0.15
8	0.44	0.44	0.50	0.19	0.10

Table 5.2b. Same as Table 5.2a for Blue Canyon

Time Shift	Obs	eWind	SynFcst	eWind Error	SynFcst Error
0	1	1	1	1	1
1	0.93	0.95	0.92	0.77	0.82
2	0.86	0.89	0.85	0.60	0.69
3	0.80	0.83	0.79	0.48	0.57
4	0.74	0.77	0.73	0.39	0.48
5	0.68	0.71	0.68	0.33	0.40
6	0.63	0.66	0.63	0.28	0.33
7	0.58	0.61	0.59	0.24	0.27
8	0.53	0.57	0.55	0.21	0.22

Table 5.2c. Same as table 5.2a for Storm Lake

Time Shift	Obs	eWind	SynFcst	eWind Error	SynFcst Error
0	1	1	1	1	1
1	0.92	0.94	0.91	0.73	0.76
2	0.85	0.88	0.84	0.53	0.59
3	0.78	0.83	0.78	0.41	0.46
4	0.71	0.77	0.72	0.32	0.35
5	0.65	0.72	0.66	0.26	0.27
6	0.60	0.67	0.61	0.21	0.20
7	0.54	0.62	0.56	0.18	0.15
8	0.49	0.57	0.52	0.15	0.12

Table 5.2d Same as Table 5.2a for Lake Benton

Time Shift	Obs	eWind	SynFcst	eWind Error	SynFcst Error
0	1	1	1	1	1
1	0.92	0.95	0.91	0.79	0.82
2	0.84	0.89	0.85	0.62	0.68
3	0.77	0.83	0.79	0.51	0.57
4	0.70	0.77	0.73	0.42	0.47
5	0.65	0.72	0.68	0.36	0.40
6	0.59	0.66	0.63	0.32	0.34
7	0.54	0.61	0.58	0.28	0.28
8	0.49	0.57	0.54	0.24	0.23

Last, the correlation of forecast errors between projects was evaluated, particularly as a function of distance between them. The correct modeling of the spatial correlation of forecast errors was an important consideration for this study as it looks at 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. The results are presented in Table 5.3. The synthesized forecasts exhibit a greater degree of correlation of forecast errors for a given distance, indicating a possible tendency to overestimate the errors when aggregated over large numbers of projects in a given region. However, the difference is small and the functional dependence with distance is similar.

Table 5.3. Correlation of forecast errors as a function of distance between project pairs

Distance (km)	Output	eWind Error	SynFcst Error
168	0.61	0.21	0.27
303	0.62	0.21	0.27
913	0.23	0.03	0.08
1047	0.15	0.03	0.06
1214	0.16	0.03	0.05
1342	0.11	0.04	0.03

5.3 Adjustments to Forecasts

Following the delivery of the synthesized forecasts to NREL, it was found that the difference between the next-day and six-hour-ahead forecast errors in the synthesized forecasts was not as large as that in many real forecasts. This problem was traced to the use of the same mesoscale model feed to generate forecasts for all time horizons. In reality, next-day forecasts would be based on mesoscale simulations that were run some 24-36 hours before the simulations for the four- and six-hour-ahead forecasts, and therefore the errors for the next-day time horizon would be somewhat larger. To compensate for this, AWS Truewind generated two additional sets of forecasts. One reduced the forecast error margin for the four- and six-hour-ahead forecasts by about 12% relative to the next-day forecasts; the other increased the next-day forecast error by the same margin relative to the four- and six-hour forecasts. The former may be regarded as representing the performance of an advanced forecasting system, whereas the latter represents current technology.

6 One-Minute Output

In the final task, AWS Truewind simulated one-minute plant output data for the onshore and offshore sites. The data were produced for the following time windows selected by NREL (see Table 5.4).

Table 5.4. Date Ranges (MM/DD/YY) for One-Minute Samples

Start	End
1/1/05	1/15/05
4/1/05	4/16/05
7/10/05	7/16/05
10/9/05	10/22/05
1/9/06	1/15/06
5/1/06	5/7/06
8/1/06	8/1/06
10/1/06	10/7/06

To produce the data, AWS Truewind employed a computer program to sample four-hour windows of historical one-minute data from existing wind projects. The source of the samples was two years of one-minute plant output data provided in a previous project by ERCOT for 17 substations serving seven Texas wind projects.[3] (Because of a lack of output data from offshore plants, the same dataset was used to simulate fluctuations at offshore sites.) The program removed 10-minute trends from the data using a bicubic fitting procedure, and then added the residuals to the simulated 10-minute output for each site. It did not allow the same window of residuals to be applied to two different sites in the same time period, as this would have resulted in perfect correlation of the one-minute fluctuations between those sites, whereas in reality one-minute fluctuations between wind projects are entirely uncorrelated. The program excluded from the training data one-minute changes greater than 5% of the plant rated capacity, as they correspond to plant outages, curtailments, and restarts unrelated to the wind.

Figure 6.1 shows a typical sample of one-minute simulated and actual plant output data for a single project overlaid on 10-minute data. The synthesized one-minute data display an obvious similarity to the actual data.

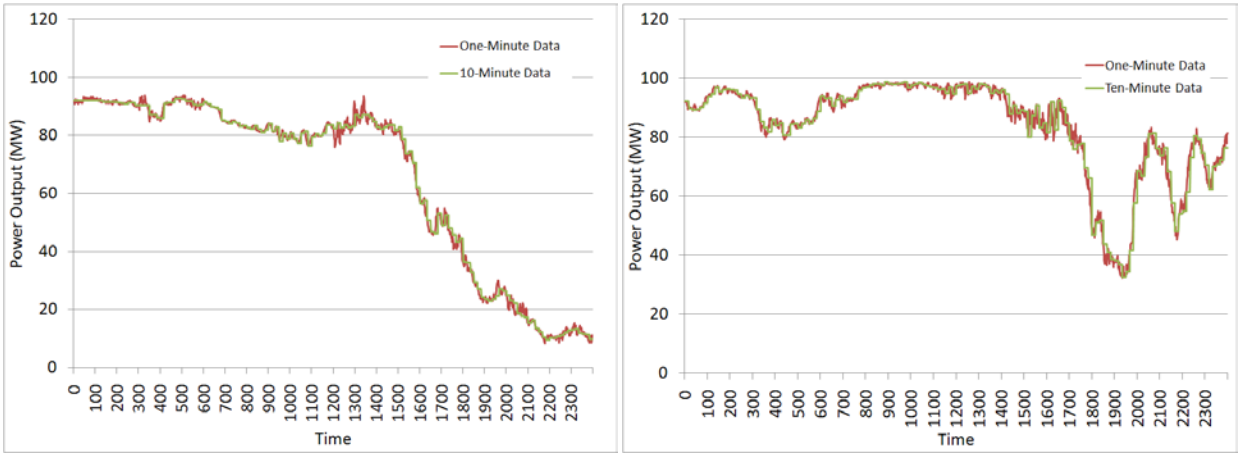


Figure 6.1. Sample of one-minute data overlaid on corresponding 10-minute data. The left-hand chart shows simulated data for a Texas site, the right-hand chart shows actual data from a Texas wind project. (Note that the sites and time periods are not the same.)

Figure 6.2 compares the modeled and observed frequency distributions for one-minute step changes for one of the plants in ERCOT dataset. It was not possible to match the site exactly to this project, since the ERCOT plants are to the south of the EWITS region. Thus, two EWITS sites are shown to illustrate the range of results to be expected in this region. The ramp distributions are similar.

[3] The files provided by ERCOT contained data for 32 substations. However, 17 of these did not have valid data. Of the remaining 15 substations, several represented different parts of the same project.

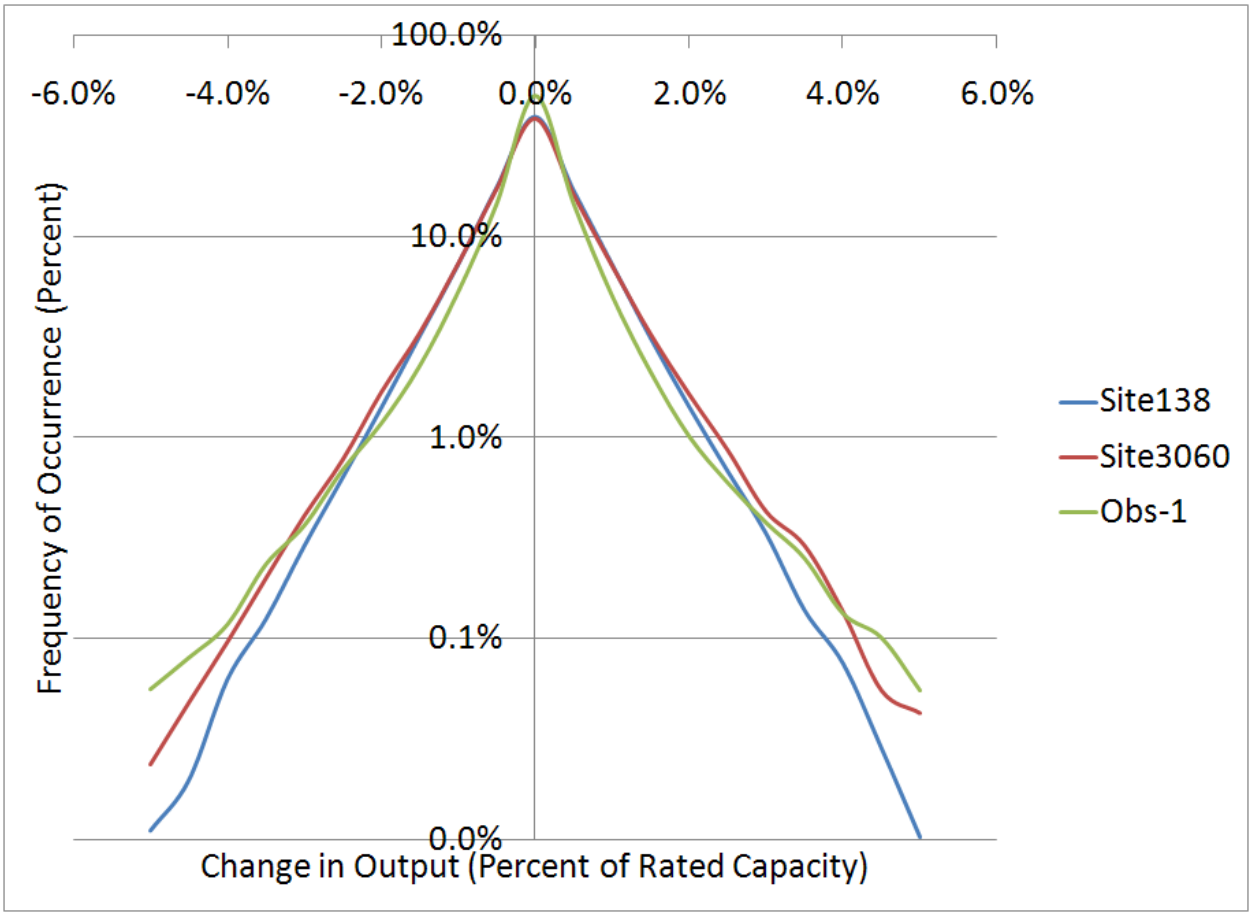


Figure 6.2 Comparison of modeled and observed step-change frequency for two EWITS sites (Site 138 and Site 3060) to the north of a wind plant in central Texas (Obs-1).

7 Conclusions

AWS Truewind has produced a wind plant output dataset spanning three years at 10-minute time resolution for over 580 GW of onshore wind projects and 208 GW of offshore wind projects in the eastern United States. Comparison of the data with observed plant output at three existing wind projects indicate that the diurnal and seasonal patterns as well as the dynamic behavior (ramp rates) are represented with acceptable accuracy. Hourly synthetic wind plant forecasts for four, six, and 24 hours ahead were also produced using a probabilistic method based on actual forecasts, and the autocorrelations and correlations between projects were found to be captured with acceptable accuracy. An adjustment was later applied to correct for an overly optimistic next-day forecast error margin relative to the other time horizons. Last, one-minute plant output data for the same onshore and offshore sites were produced for several time windows selected by NREL.

Appendix

Validation Report

Development of Eastern Regional Wind Resource and Wind Plant Output Datasets

Subcontract No. ACO-8-88500-01

Submitted by:

AWS TRUEWIND, LLC
263 NEW KARNER ROAD
ALBANY, NEW YORK

Michael Brower
Chief Technical Officer
Tel: 978-749-9591
Fax: 978-749-9713
mbrower@awstruewind.com

July 17, 2008

Development of Eastern Regional Wind Resource and Wind Plant Output Datasets

Subcontract No. ACO-8-88500-01

Validation Report

1. Introduction

To make the best possible simulations for this project, AWS Truewind tested several different configurations of two mesoscale models to choose the model and configuration for this project. AWS Truewind has successfully used the MASS (Mesoscale Atmospheric Simulation System) model to simulate wind resources for many years. The community WRF (Weather Research and Forecasting) model is widely used by many groups in the United States and around the world, and AWS Truewind has used it for some wind forecasting applications. Each of these models was run in a few different configurations and the resulting 80-m wind speeds were compared to a set of 10 tall tower locations in 10 different states. The best combination of model and configuration was chosen for the subsequent production simulations of the 2004-2006 period over the entire Eastern Region.

In the second phase of the validation process, AWS Truewind extracted 10-minute data from the mesoscale simulations in the selected configuration for points near three wind projects for which measured plant output data were provided by NREL. The SynOutput program was run to convert the mesoscale data to simulated plant output. The simulated output was then compared to the observed output. With adjustments for plant size, turbine hub height, and turbine power curve, a reasonably good fit with the observed data was obtained, and NREL approved the program for the simulation of plant data for all sites.

2. Validation of the Mesoscale Model

2.1 Experimental Configurations

High-quality tall tower data was collected from 10 sites, each in a different state within the Eastern Region area of interest. Nine of the 10 sites had data for the same 2004-2006 period that will be used for the production runs. For those nine sites, 27 two-week periods spread throughout those years were chosen, based on the best dates for overlapping availability of valid tower data. The other site (Massachusetts) had only data for 2004, so 26 two-week periods covering the entire year (except for December 31) were used instead. Table 1 shows the dates of the validation simulations for each site. For each series, the first date is a "spin-up" day – no comparisons with observations are made for those days. For each location, 30-km, 8-km, and 2-km grids were created that are centered over the tall tower site.

Table 2 shows the set of seven experiments that were tested for this project. Four experiments used MASS and three used WRF in various configurations. AWS Truewind has generally used gridded data from the NCAR-NCEP Global Reanalysis (NNGR) for initial and boundary conditions for wind mapping simulations. In recent years, a new dataset has been tested, the North American Regional Reanalysis (NARR). The NARR has much higher resolution than the NNGR (32 km instead of about 190 km), and incorporates more data sources, which should result in better surface fields (e.g., soil moisture). Each model was tested with both sources of gridded data, and the source that performed best was used for subsequent tests.

Table 1. Dates of simulations for the 10 validation sites.

Sites	Dates		
NY, WV, IN, KY, SD, KS, ME, TX, MN	20040114-20040128 20040128-20040211 20040311-20040325 20040423-20040507 20040618-20040630 20040811-20040825 20041006-20041014 20041014-20041028 20040825-20040908	20041028-20041111 20041223-20041231 20050225-20050311 20050325-20050409 20050507-20050521 20050604-20050618 20050714-20050728 20050728-20050811 20050908-20050922	20051125-20051209 20051209-20051223 20051231-20060114 20060211-20060225 20060409-20060423 20060521-20060604 20050922-20051006 20060630-20060714 20061211-20061225
MA	20031231-20040114 20040114-20040128 20040128-20040211 20040211-20040225 20040225-20040311 20040311-20040325 20040325-20040408 20040408-20040422 20040422-20040506	20040506-20040520 20040520-20040603 20040603-20040617 20040617-20040701 20040701-20040715 20040715-20040729 20040729-20040812 20040812-20040826 20040826-20040909	20040909-20040923 20040923-20041007 20041007-20041021 20041021-20041104 20041104-20041118 20041118-20041202 20041202-20041216 20041216-20041230

Table 2. The list of validation experiments. Unless otherwise noted, the MASS runs employed 25 vertical levels, assimilated rawinsonde data, and used a TKE PBL scheme and one-way nesting from coarser to finer grids. The WRF runs employed 28 vertical levels and used the YSU PBL scheme and two-way interactive nesting between grids.

Experiment	Model	Gridded Data, Res	Other
1. MASS/NNGR	MASS 6.8	NNGR, 190 km	
2. MASS/NARR		NARR, 32 km	
3. MASS/NNGR/sfc		NNGR, 190 km	Surface data
4. MASS/NNGR/35 levels		NNGR, 190 km	35 vertical levels
5. WRF/NARR	WRF 2.2.1	NARR, 32 km	
6. WF/NNGR		NNGR, 190 km	
7. WRF/NARR/MYJ		NARR, 32 km	MYJ PBL scheme

Since the NNGR is so coarse, rawinsonde data is routinely assimilated into the initial conditions for all of the MASS validation tests. Surface observations are sometimes used, but there is a risk that incorporating surface data every 12 hours during a series of simulations might disrupt higher-resolution mesoscale circulations that the model is creating, and there are also concerns about data quality (the possibility that individual erroneous observations will negatively affect one or more simulations). Experiment MASS/NNGR/sfc tested the effect of surface data in MASS. WRF was used in this project without assimilating rawinsonde or surface data by the WRF-VAR scheme, as there was not time to obtain the needed data and create suitable background error fields before the validation tests were scheduled to begin.

Compared to conventional mesoscale simulations used for weather forecasting, we generally use a modest number of vertical levels – 25 for MASS and 28 for WRF, with higher resolution near the ground (MASS has 10 vertical levels within the lowest 1 km). With these configurations, both models have enough resolution in the planetary boundary layer (PBL) to reasonably resolve low level jets and other

phenomena important for the correct prediction of low-level wind speed. Increasing the number of vertical levels should improve the overall simulation of certain types of cases; for example, added resolution at middle and upper levels of the troposphere should better resolve complex synoptic circulations such as upper level jets and cloud microphysical processes. However, additional vertical levels increase the computational cost, an important consideration when very large numbers of simulations are required, as for this project. Experiment MASS/NNGR/35 levels tests the sensitivity of the MASS simulations to an increase from 25 to 35 vertical levels.

One of the advantages of the WRF system is that, as a community model, it includes a wide range of physics options. For wind prediction, the performance of the PBL scheme is of vital importance. Both the Yonsei State University (YSU) and Mellor-Yamada-Janjic (MYJ) PBL schemes are popular with WRF users. Experiments WRF/NARR and WRF/NNGR use the YSU scheme, while Experiment 7 uses the MYJ.

2.2 Results

For each site, hourly wind speeds were extracted at 80 m above ground level from the highest-resolution (2 km) simulations and compared to tall tower wind speeds at 80 m. Three statistics were used to compare the simulated to observed wind speeds: mean error (ME), mean absolute error (MAE), and scaled mean absolute error (SMAE). The SMAE is the MAE of the simulated hourly wind speeds after each value has been scaled by the ratio of the observed to simulated mean wind speed over the entire period of comparison, to remove the overall bias. Since the wind speeds from the main mesoscale runs will be scaled in this fashion to the expected mean wind speed at each wind project site, the SMAE is arguably the most important measure for determining the best model configuration.

Figure 1 shows each statistic over the entire validation period for each site. On the far right of each plot, the values for all 10 sites are averaged together and labeled as “Mean”. In general, both models had a negative bias (i.e., the simulated wind speed was lower than the observed), by an average of about 1 m/s. The sign and magnitude of the biases at each site were generally comparable between models and configurations; in only a handful of cases did one or more experiments diverge sharply from the others at a site. The main reason for the negative bias, we believe, is that most towers are on elevated features that are not fully resolved at the model resolution.

The MAE and SMAE results show that each model performed decisively better for one gridded data source. MASS had a SMAE of 2.00 m/s using NNGR data, but 2.18 m/s with the higher resolution NARR data. This is a surprising result, but we have seen problems in earlier projects running MASS with NARR.

For the WRF model however, using NARR gridded data produced clearly superior results. Since rawinsonde and surface observations were not assimilated into the WRF runs, perhaps the 32 km resolution was the most important factor in improved performance. It’s also possible that the WRF physics schemes were able to better utilize some of the additional surface fields that come with the NARR dataset (e.g., several layers of soil moisture and temperature).

The MASS/NARR and WRF/NNGR were clearly the worst-performing configurations, but the other eight configurations were quite similar in SMAE values. The best was the MASS/NNGR/sfc; it had a SMAE value of 1.97 m/s averaged over all stations. But the next three best MASS configurations were close behind, with SMAE values between 1.99 and 2.00 m/s, and the two best WRF configurations (WRF/NARR and WRF/NARR/MYJ) were also in the same range.

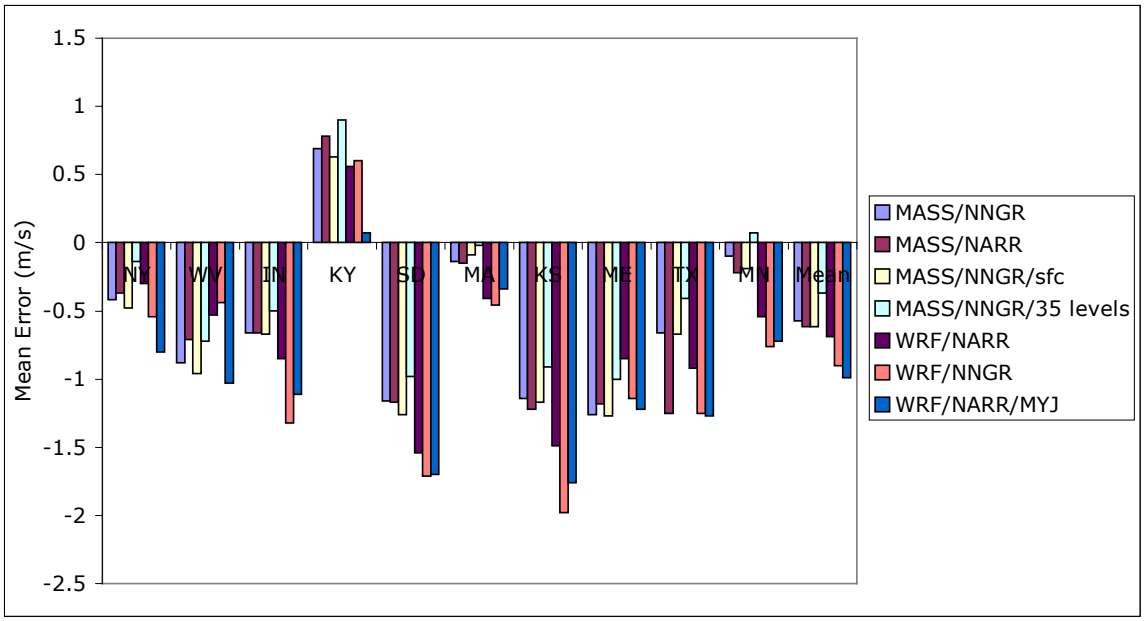


Figure 1a. Mean error over all simulations for each tower location, along with the mean of all the towers.

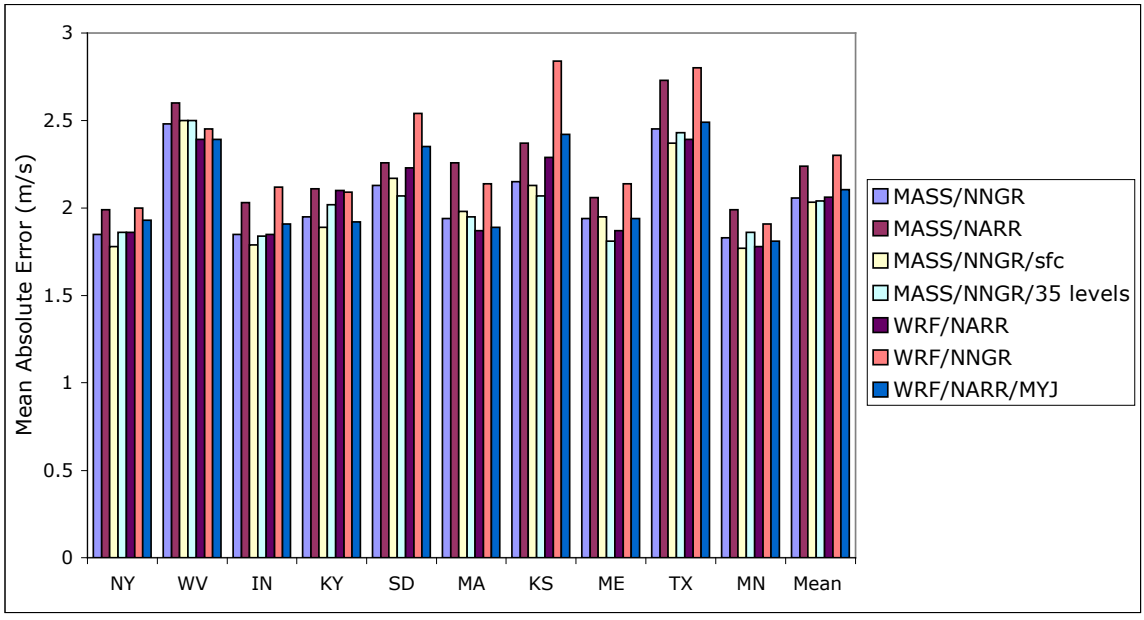


Figure 1b. Mean absolute error over all simulations for each tower location, along with the mean of all the towers.

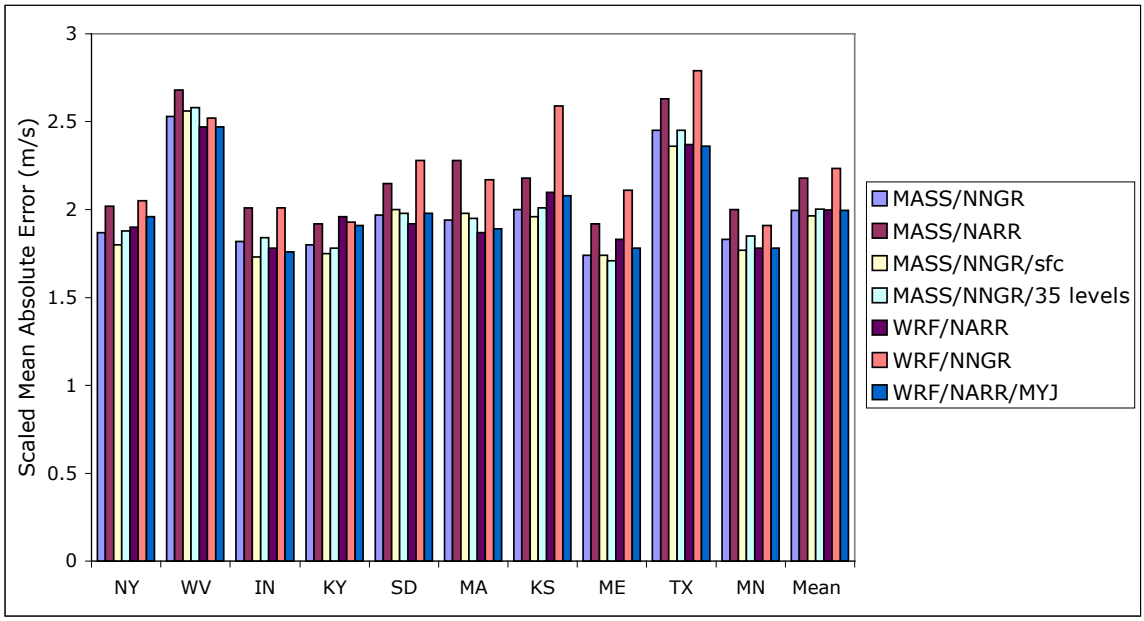


Figure 1c. Scaled mean absolute error over all simulations for each tower location, along with the mean of all the towers.

In addition to the mean biases and absolute errors of the various model configurations, the diurnal and monthly mean speed patterns were evaluated qualitatively.

Appendix A contains the validation plots for the MASS/NNGR/SFC configuration, which was judged the best overall. The most significant discrepancies in the diurnal patterns are seen at the Maine, New York, Texas, and West Virginia sites. The fact that the model cannot fully resolve the elevated features on which these towers are located may explain its tendency to overestimate the range of daily variation in the mean speed in each case. (Such patterns are caused by variations in the thermal stability and mixing of the planetary boundary layer, which are less pronounced on sharp peaks and ridge lines.)

By and large, the monthly patterns are more accurately simulated than the diurnal patterns. At only one site, in Minnesota, was a substantial discrepancy observed. After further investigation, it was concluded that poor model forecasts for the region in several of the two-week validation periods were to blame. It is expected that the discrepancy will not be as large for the full three-year period.

Among the various model configurations, no systematic differences in the error patterns were detected. Some model configurations did better at some sites than at others, but none were found to be clearly superior to the others overall. In particular, WRF produced a similar monthly error pattern at the Minnesota site as the MASS model, implying that the problem at that location was not model-dependent.

It was concluded that the scaled mean absolute error provided a reliable indication of the relative merit of the different models and their respective configurations.

2.3 Selection of Best Model Configuration

When 80-m wind speeds from the simulations were compared to tall tower data at 10 different sites, the best configuration overall used the MASS model initialized by NNGR gridded data, assimilating rawinsonde and surface data, and with 25 vertical levels (Experiment MASS/NNGR/sfc in Table 2). Three other MASS configurations and the two best WRF configurations were not far behind, but some of them have other disadvantages. The MASS runs with more vertical levels would take significantly more computational time, and the WRF runs do not take full advantage of the Intel quad-core architecture, possibly because of the use of shared cache between each set of two cores, resulting in a much lower overall throughput.

Therefore, the MASS/NNGR/sfc configuration was the final choice for the production runs.

3. Validation of the Plant Output Model

3.1 Plant Datasets

NREL provided three years of 10-minute plant production data for four wind projects: Lake Benton, Minnesota; Storm Lake, Iowa; Blue Canyon, Oklahoma; and Trew Ranch, Texas. The last was not within the bounds of the study area, however, so only the first three were employed in the validation of the plant output model.

Table 3. Wind power projects employed in the validation.

Plant Name	State	Rated Capacity (MW)	Turbine Type	Hub Height (m)
Blue Canyon I	Oklahoma	74.25	NM72 (1.65 MW)	67 m
Lake Benton	Minnesota	103.5	Zond 750	51.2 m
Storm Lake I	Iowa	112.5	Zond 750	63 m

3.2 Method

A time series of mesoscale data were extracted from all grid cells near each wind project site and near each validation tower. A computer program written in FORTRAN was then employed to convert the mesoscale data for the cells near the project sites to plant output. This program proceeds in the following steps:

1. The program reads a list of the sites and the nearest associated grid cells (grid number, I and J position). For each cell, the list includes the latitude and longitude, expected mean speed, mesoscale elevation, actual elevation, and proportion of the site's total rated capacity associated with that cell.
2. The program imports the turbine power curves. Normally, there is one power curve for each IEC class. The power curves are normalized to the rated capacity, and are valid for the standard sea-level air density of 1.225 kg/m^3. The IEC 1 and 2 curves are based on a composite of three commercial turbines (GE, Vestas, and Gamesa). In consultation with NREL, it was decided to base the IEC 3 curve on just two turbines (GE 1.5xle and Gamesa G90) to remove an inconsistency in the cut-out speed of the Vestas V100. In addition, the cut-out speed of the GE turbine was changed from 20 to 21 m/s to match that of the Gamesa turbine. (As described in the

next section, the Zond 750 power curve was also employed, which improved the fit to the data from the Lake Benton and Storm Lake projects.)

3. The program reads a set of 12x24 speed matrixes, one for each of the 10 validation towers. These matrixes give the mean speed for each hour of the day and for each month of the year. For each tower there are two matrices, one for each hub height (80 m and 100 m). (For the plant output validation, additional matrices were created for 50 m and 65 m to approximately match the actual hub heights of the turbines.)
4. The program creates a similar 12x24 speed matrix from the mesoscale speed data for the grid cells nearest the validation towers. The ratio between the average annual diurnal speed observed at each tower and that predicted by the model is then calculated. This forms an adjustment matrix, which is used to correct model biases (especially a tendency for the speed to be too high from around midnight to 5 or 6 am local time).
5. The program reads the mesoscale time series file for each grid cell associated with a project site. It then scales the data to match the expected mean speed from the wind map, and finally sums the scaled speeds for all the grid cells associated with the site. In the sum, each cell's speeds are weighted according to the proportion of the site area associated with the cell. The result is a time series of simulated wind speeds for the site as a whole, unadjusted for diurnal model biases.
6. The program 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 calculates a weighted average adjustment matrix for the site, in which the weight given to the adjustment matrix for each validation location is proportional to its correlation coefficient. The program then applies this blended adjustment matrix to the simulated data for each grid cell associated with the site. For example, if the time in question is 1300, the simulated speed is multiplied by the adjustment factor for 1300. It should be noted that although the program calculates adjustments on a monthly basis, only an annual adjustment is performed; the monthly variations in the diurnal adjustment factors are ignored. This is because it was concluded in the first validation phase that the MASS model reproduces the pattern of monthly variations with reasonable accuracy.
7. The speeds for each grid cell are then adjusted for wake and other losses in a manner that depends on the simulated wind direction relative to the prevailing (most frequent) direction. The method accounts both for wake losses and, implicitly, for other losses that affect the efficiency of power conversion for a given free-stream speed without reducing the maximum rated capacity, e.g., blade soiling and icing.
8. The speed is further adjusted by adding a random factor scaled to the predicted TKE. This adjustment is intended to reflect the impact of gusts on the speeds experienced by the turbines in the wind project.
9. Next, the adjusted speeds are applied to the turbine power curve for each IEC class. In the process, the power curve is corrected to the predicted air density.
10. A time filter is then applied to mimic the effect of spatial averaging on the fluctuations of wind project output in large plants. This approach has been found to reproduce the observed "variability" of wind plant output, as measured by the mean absolute deviation (MAD) as a function of time, with reasonable accuracy. Without such time filtering, the simulated plant output would tend to be more variable than the observed.
11. The program applies an additional power loss to account for turbine and plant availability. This power loss varies randomly from zero to 12%, and averages 6%. The variation is intended to

capture the effects of changes in plant availability over time. For most sites, the total loss is about 14-15%. The range is from about 11% to about 18%.

3.3 Results

The power conversion program was tested on the mesoscale data that was extracted for grid cells associated with the three projects for which plant data were available, and the results were provided to NREL both in tabular form and as diurnal plots. Minor problems were found with the first two versions that were submitted; version 3 appeared to be satisfactory.

For the first try, the default power curves and hub heights were used; in addition, the rated capacity of the nearest site, rather than the true rated capacity of the project, was assumed. However, NREL observed that the predicted output at two projects, Lake Benton and Storm Lake, was substantially greater, on average, than the observed. By using the correct turbine power curve at each project, and by adjusting the hub height of the simulated speeds and diurnal corrections to match the actual hub height of each project, an acceptable agreement between the predicted and observed capacity factors was obtained. Furthermore, to provide a true test of the ramp rates of project output – which depend on plant size – the rated capacities were adjusted to match the actual.

Figure 2 compares the simulated and observed diurnal mean output patterns for the Blue Canyon I, Lake Benton, and Storm Lake I projects. There are some discrepancies. The predicted output at Blue Canyon is slightly below the observed, a fact we attribute to an underestimation (by about 3-4%) of the mean wind speed in the AWS Truewind map. The predicted output at both Lake Benton and Storm Lake is, on the other hand, somewhat too high. We believe this is largely the result of poor (~90%) availability and other operational problems at both projects. For example, although the rated capacity of Lake Benton is nominally 103.5 MW, the greatest output observed in 2004 was 87.4 MW. These discrepancies were judged to be within an acceptable range for this project. In addition, the diurnal patterns appear to be reasonably well modeled.

Figure 3 compares the simulated and observed mean ramp rates of the wind projects. This is measured by the mean absolute deviation (MAD) of the capacity factor from any time T to a time T+N in the future. Once again, despite some deviations, the agreement was judged to be satisfactory for the purposes of this study.

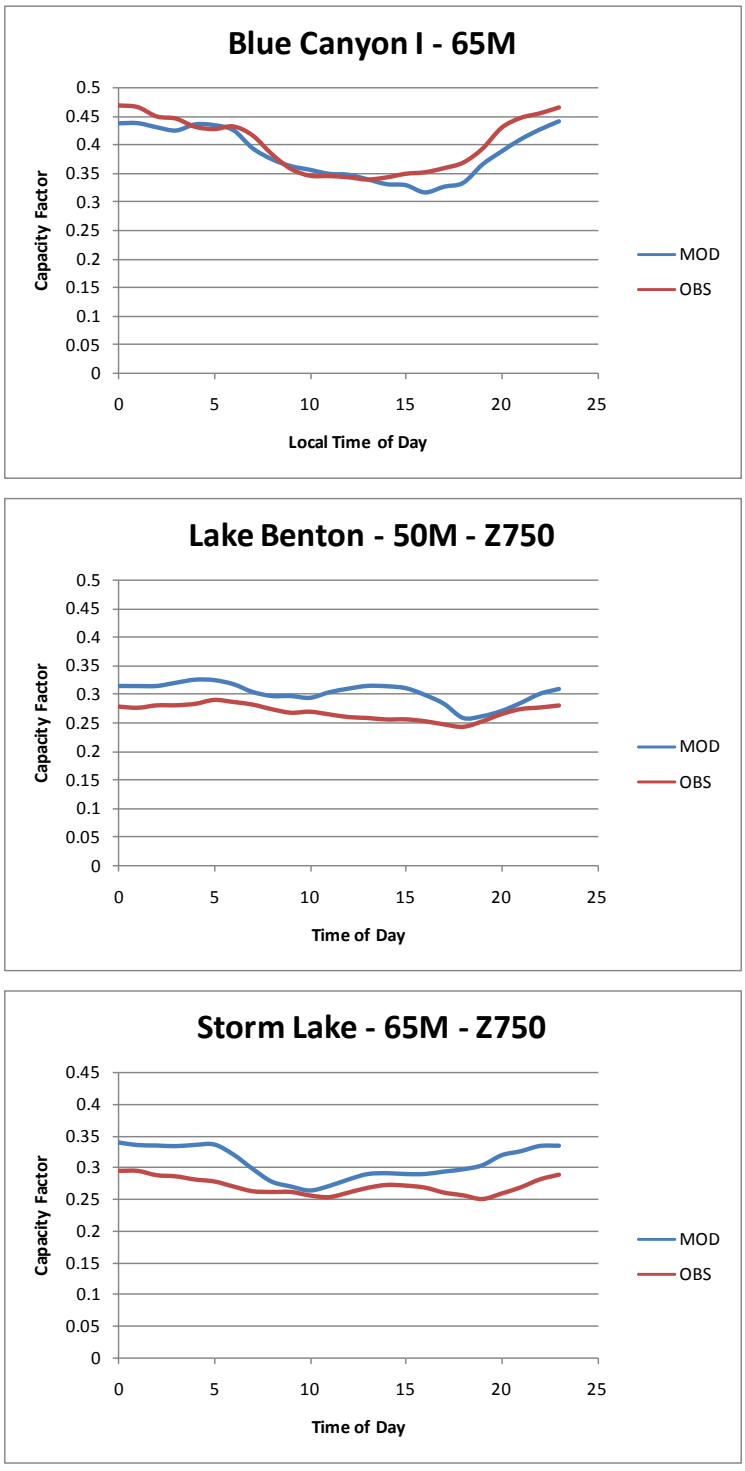


Figure 2. Comparison of simulated and observed diurnal mean output patterns for the Blue Canyon I (top), Lake Benton (middle), and Storm Lake I (bottom) projects.

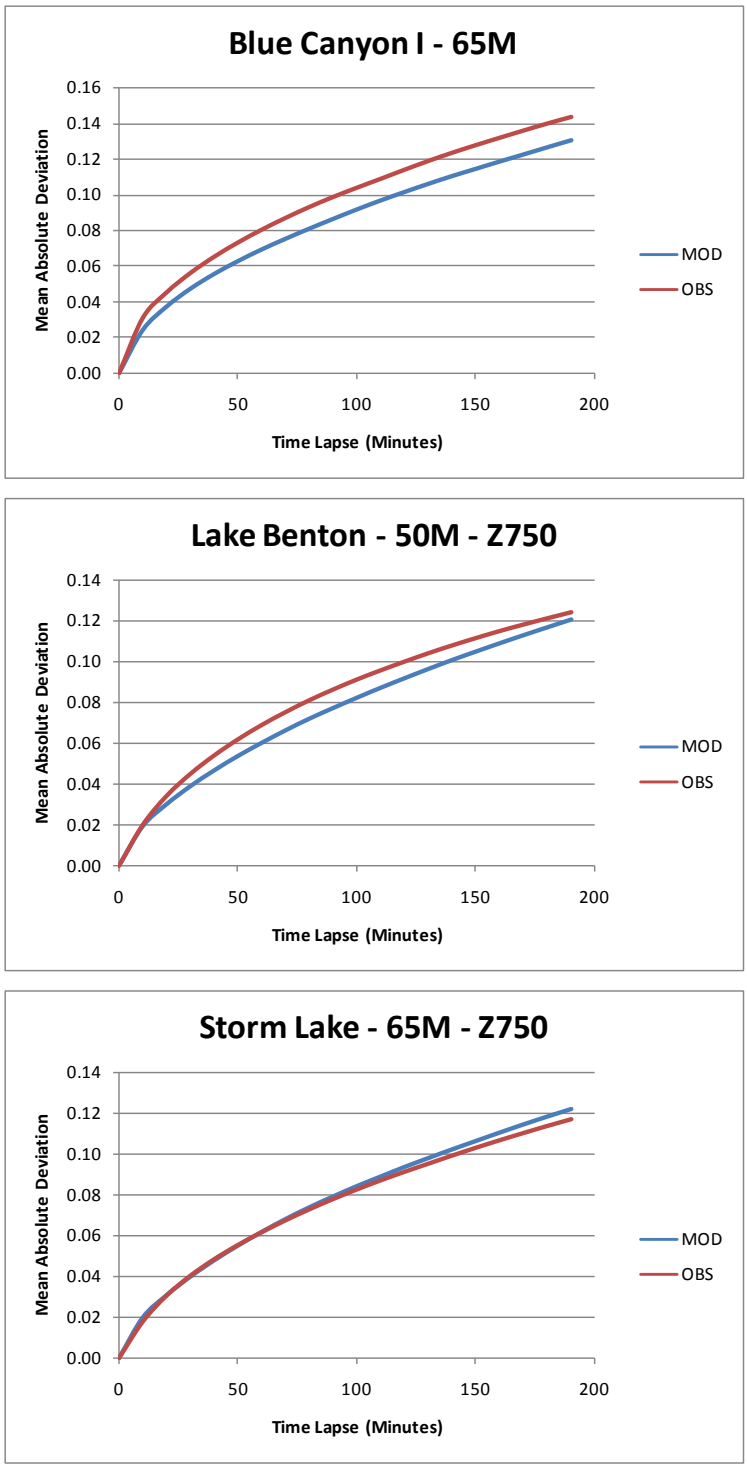
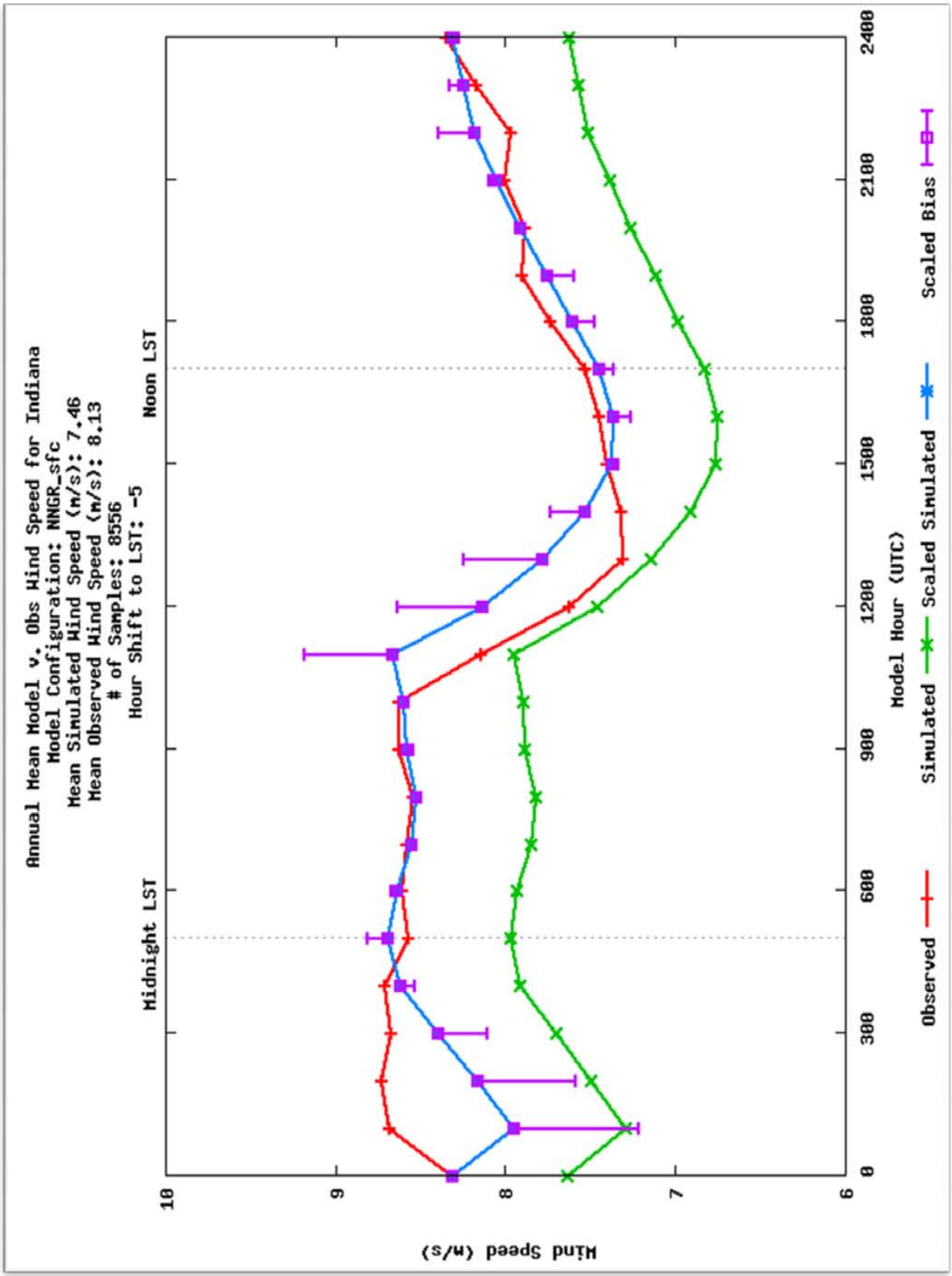


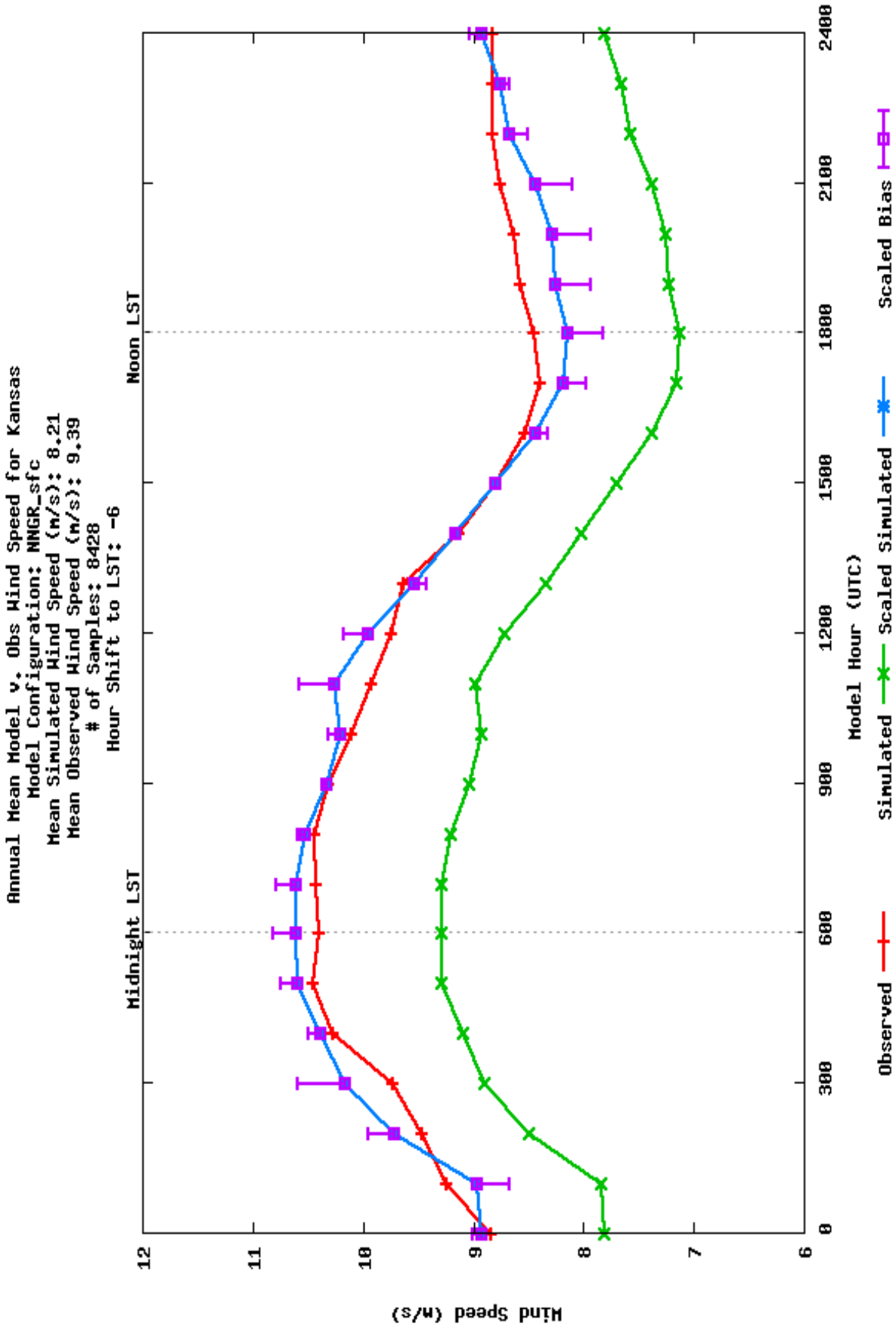
Figure 3. Comparison of simulated and observed mean ramp rates for the Blue Canyon I (top), Lake Benton (middle), and Storm Lake I (bottom) projects.

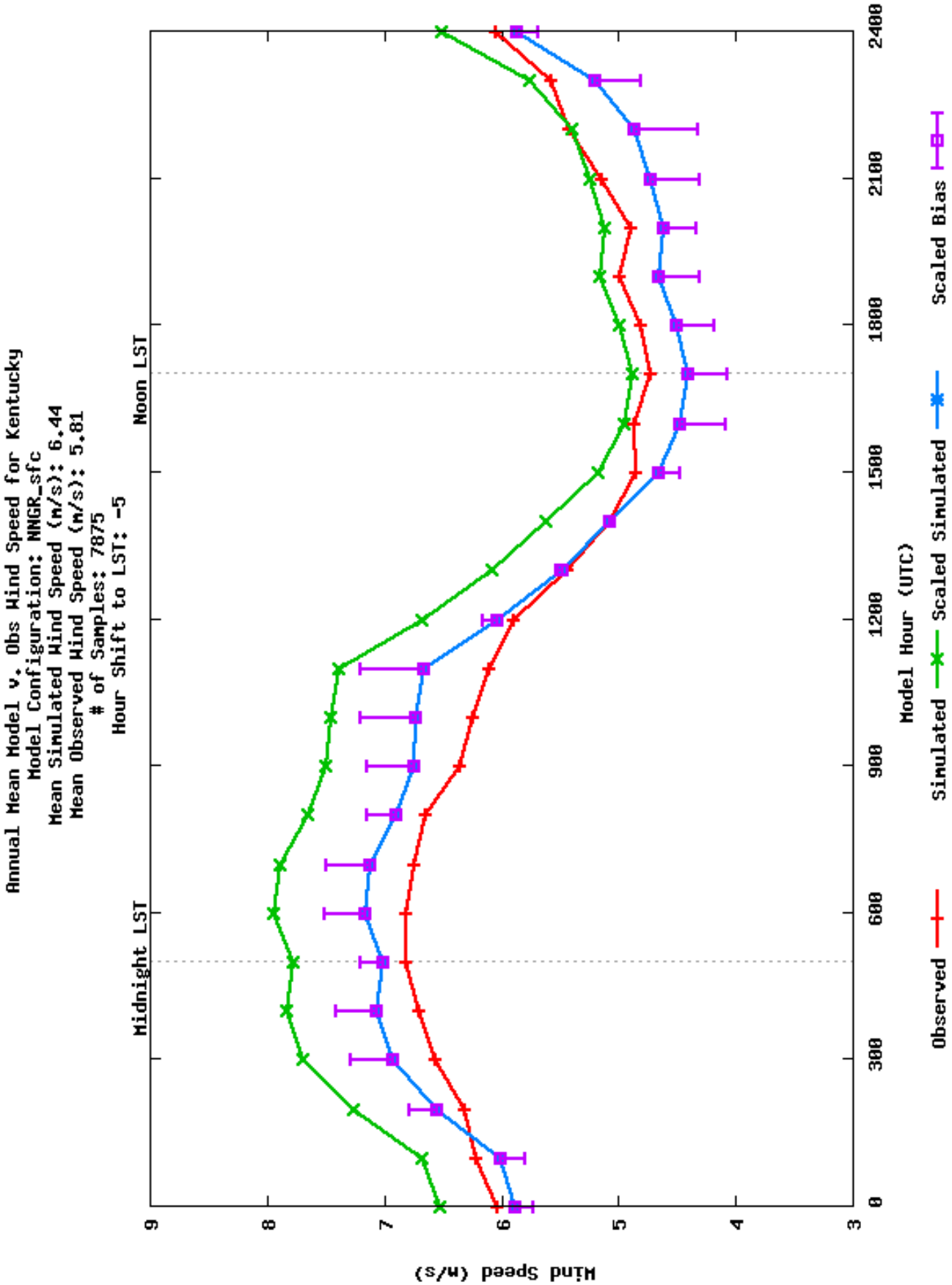
Based on the results of the validation, NREL gave its approval to proceed with version 3 of the program to convert the main mesoscale runs to plant output.

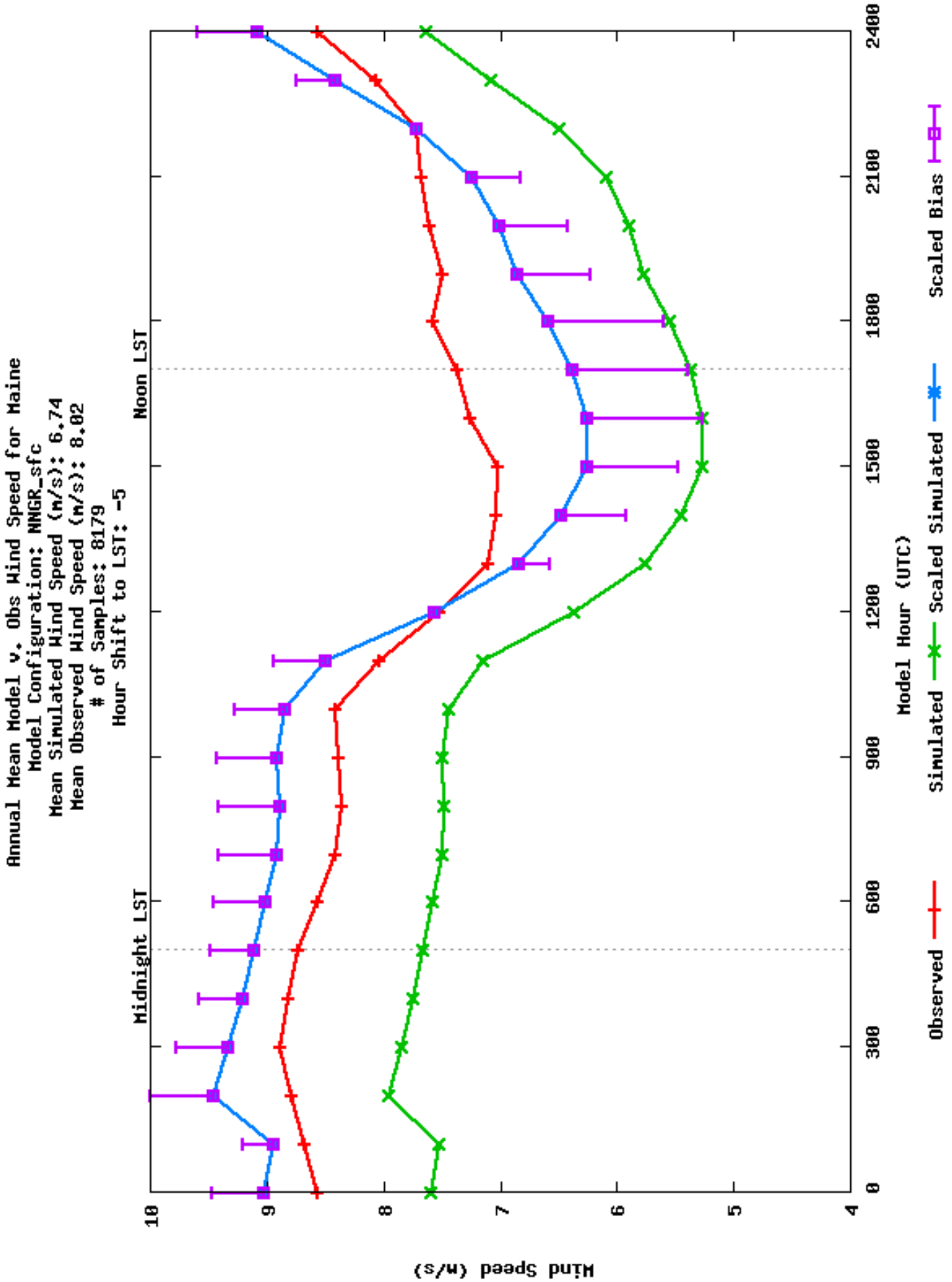
APPENDIX A

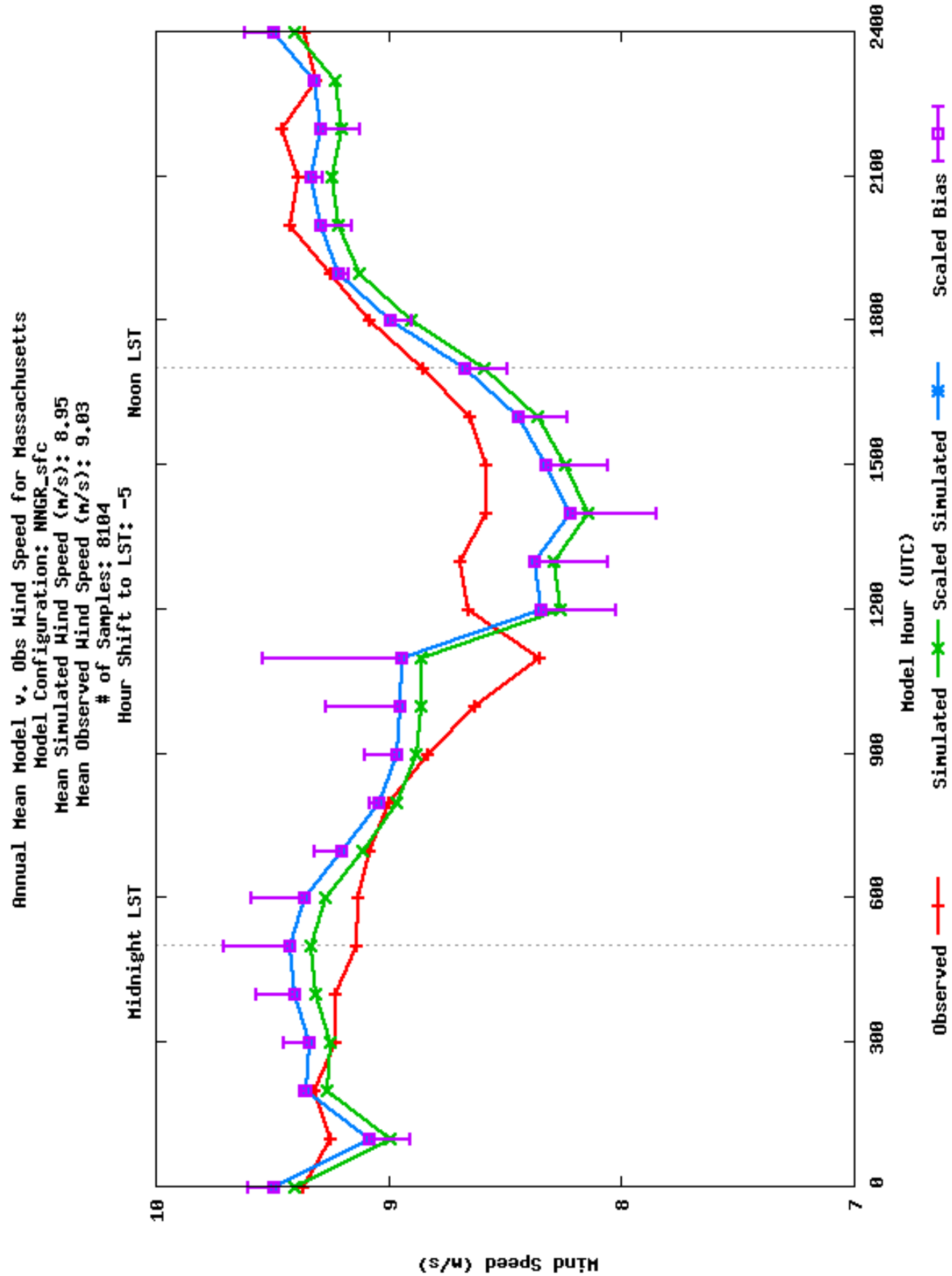
VALIDATION PLOTS FOR THE MASS/NNGR/SFC CONFIGURATION

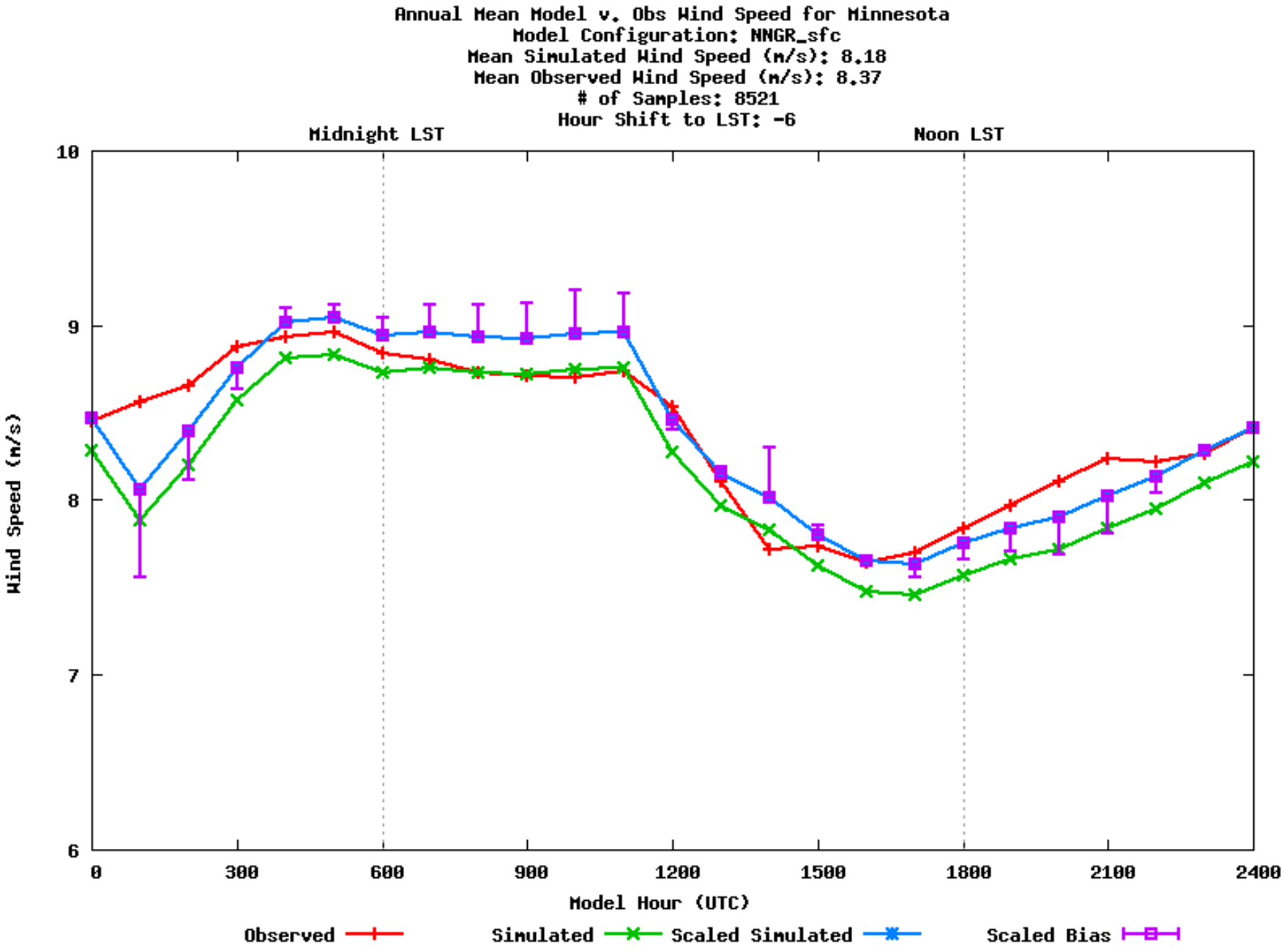

Annual Mean Model v. Obs Wind Speed for Indiana
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 7.46
Mean Observed Wind Speed (m/s): 8.13
# of Samples: 8556
Hour Shift to LST: -5
Midnight LST
Noon LST
Wind Speed (m/s)
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

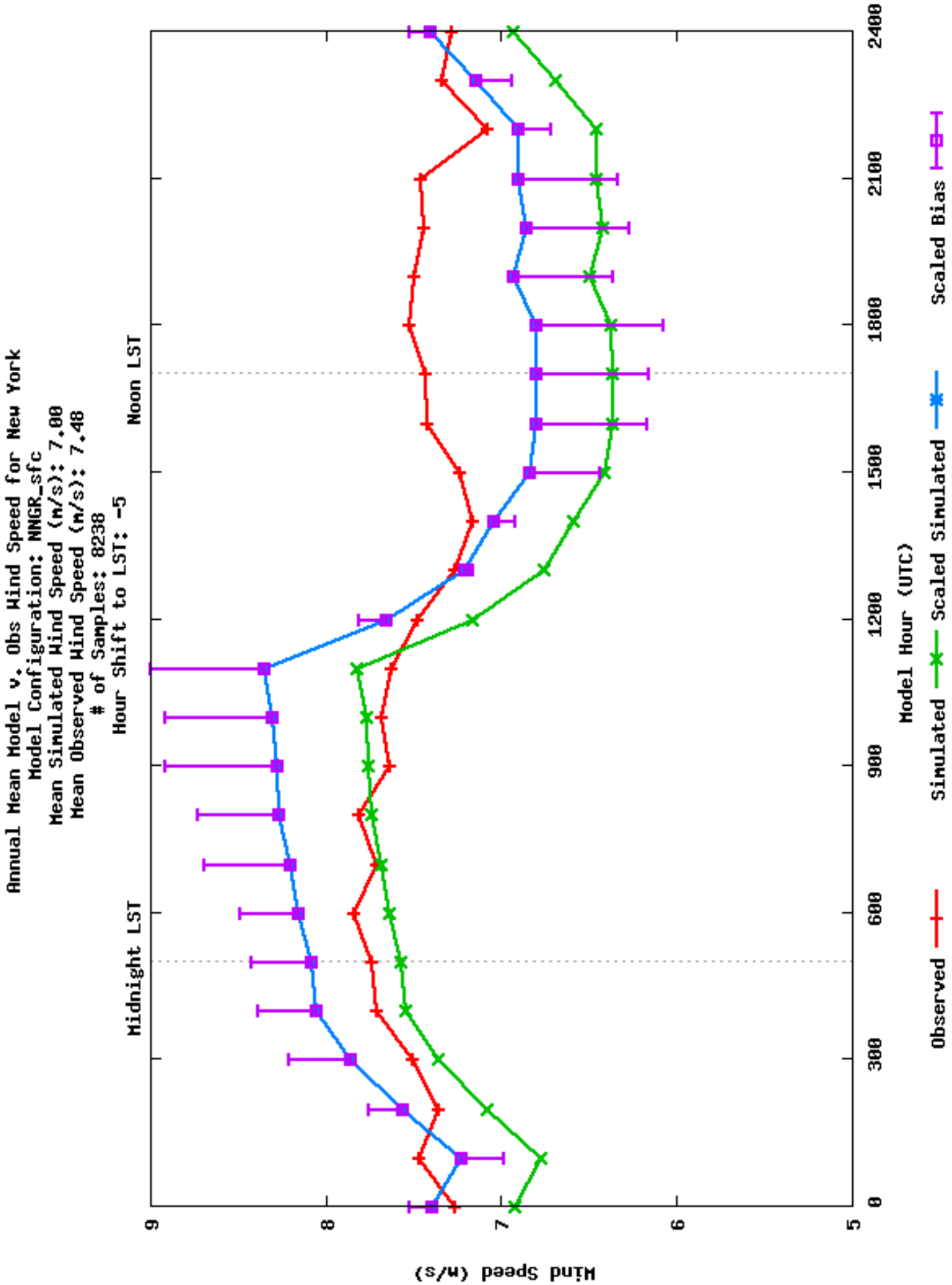

Annual Mean Model v. Obs Wind Speed for Kansas
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 8.21
Mean Observed Wind Speed (m/s): 9.39
# of Samples: 8428
Hour Shift to LST: -6
Midnight LST
Noon LST
Wind Speed (m/s)
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

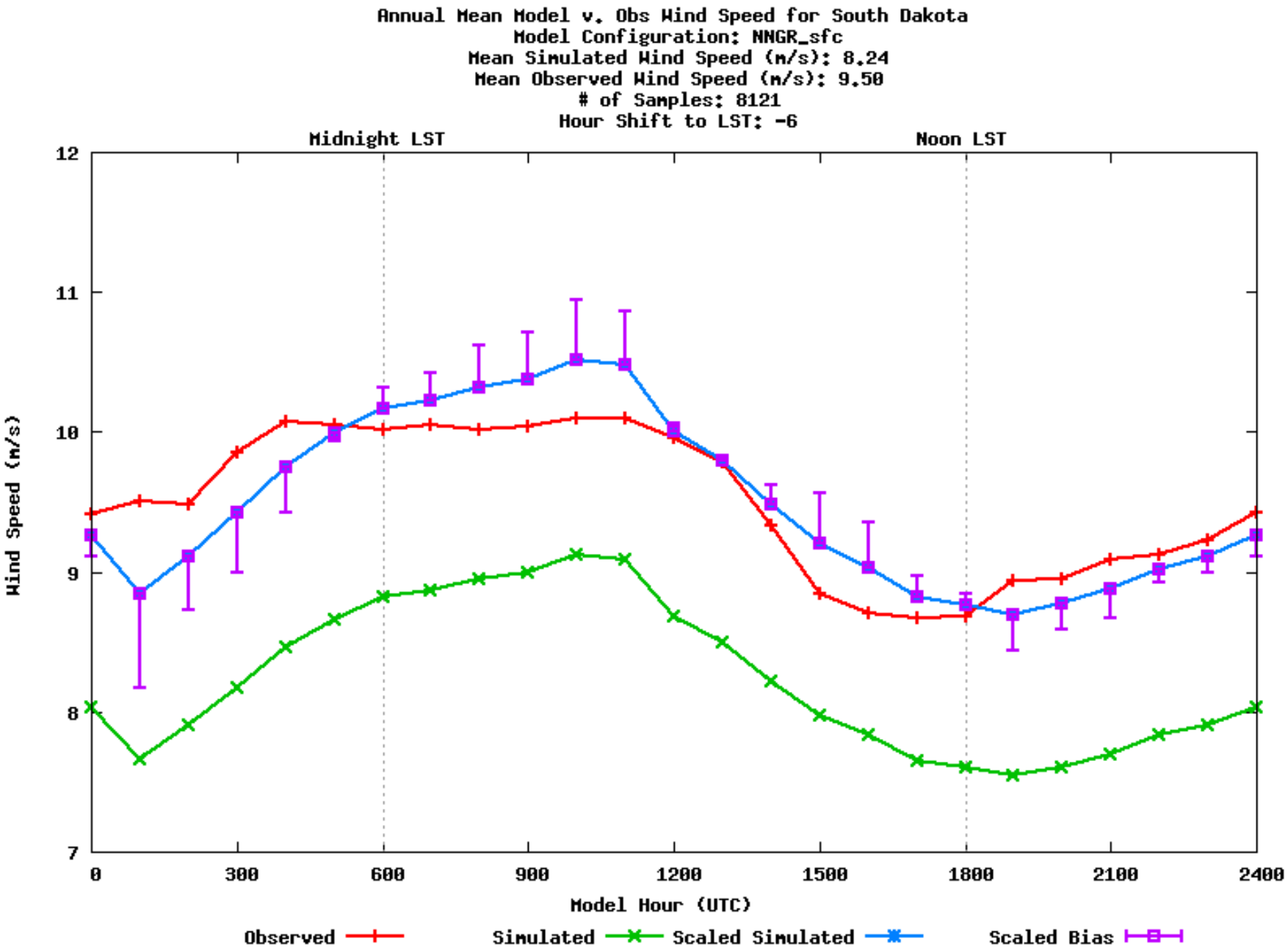

Annual Mean Model v. Obs Wind Speed for Kentucky
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 6.44
Mean Observed Wind Speed (m/s): 5.81
# of Samples: 7875
Hour Shift to LST: -5
Midnight LST
Noon LST
Wind Speed (m/s)
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

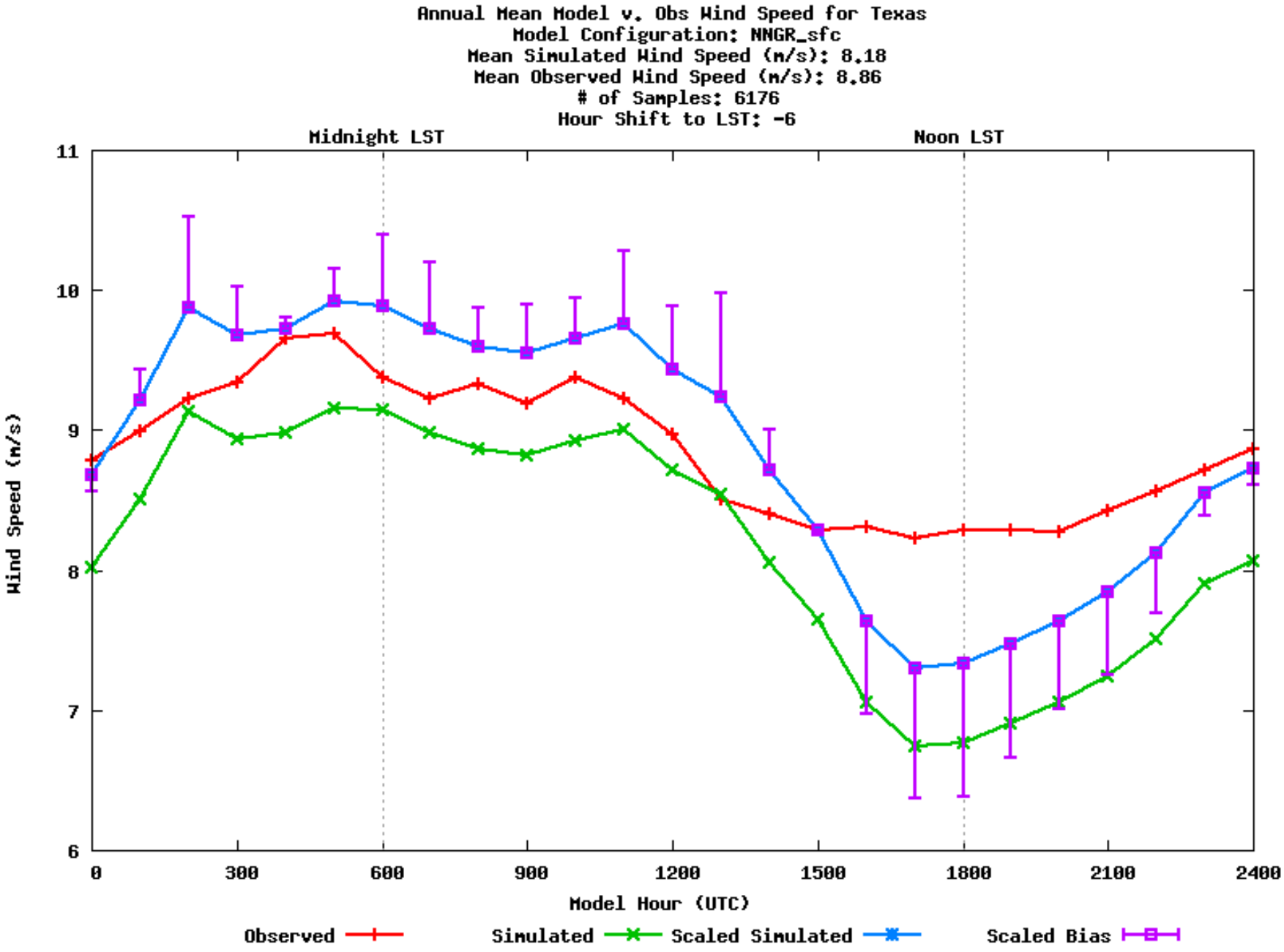

Annual Mean Model v. Obs Wind Speed for Maine
Model Configuration: MNGR_sfc
Mean Simulated Wind Speed (m/s): 6.74
Mean Observed Wind Speed (m/s): 8.02
# of Samples: 8179
Hour Shift to LST: -5
Midnight LST
Noon LST
Wind Speed (m/s)
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

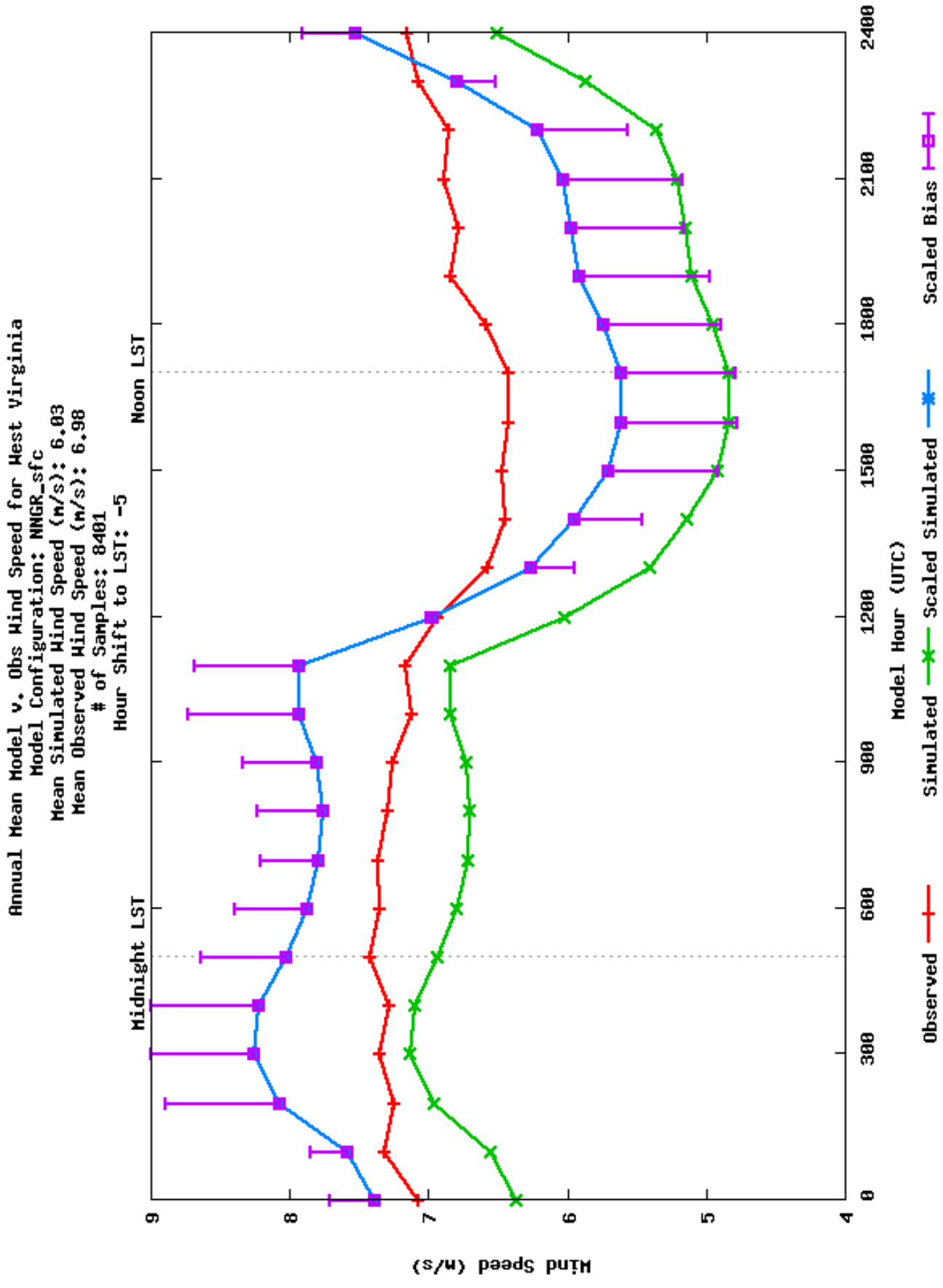

Annual Mean Model v. Obs Wind Speed for Massachusetts
Model Configuration: MNGR_sfc
Mean Simulated Wind Speed (m/s): 8.95
Mean Observed Wind Speed (m/s): 9.03
# of Samples: 8104
Hour Shift to LST: -5
Midnight LST
Noon LST
Wind Speed (m/s)
10
9
8
7
0
300
600
900
1200
1500
1800
2100
2400
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

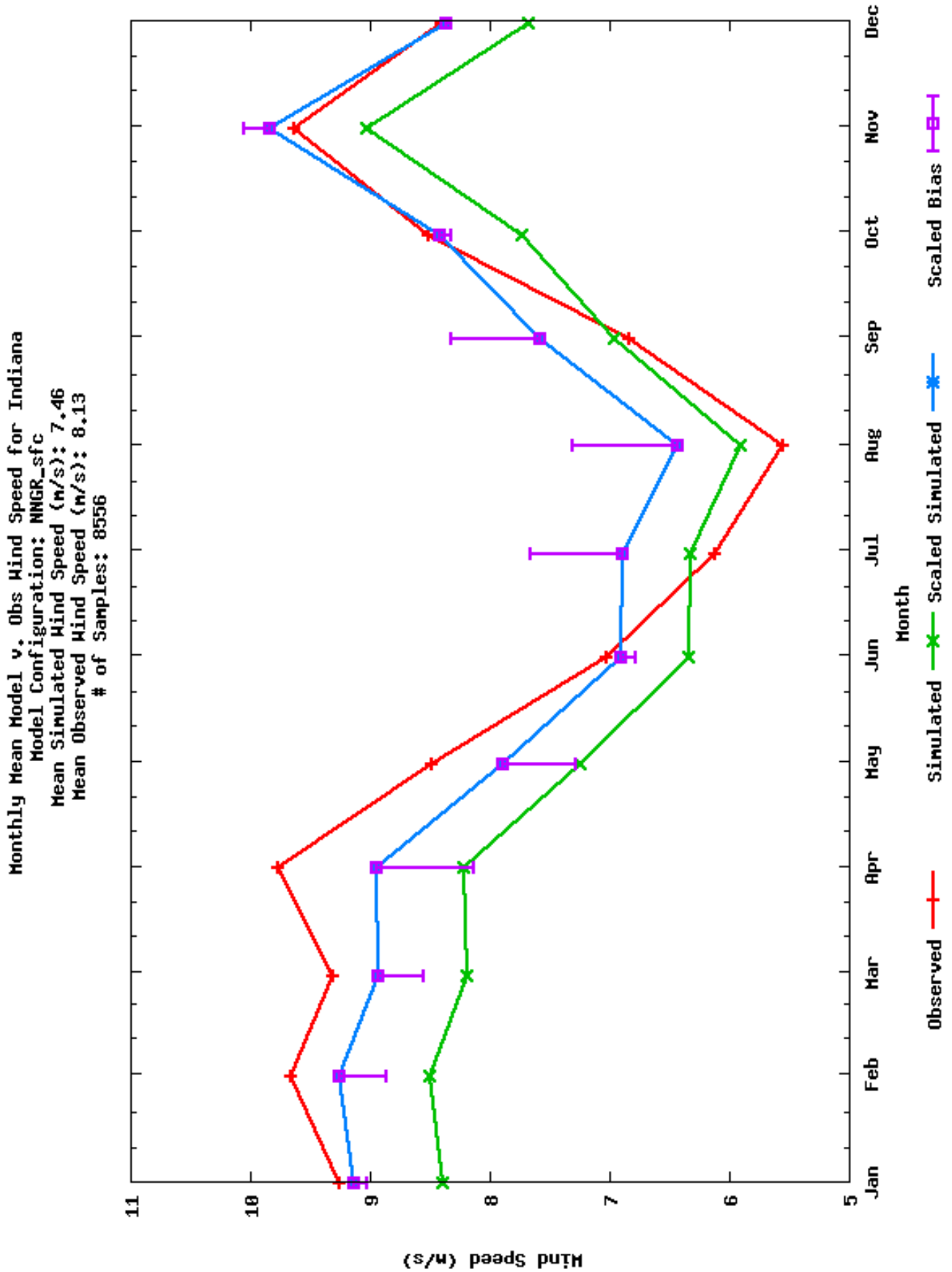

Annual Mean Model v. Obs Wind Speed for Minnesota
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 8.18
Mean Observed Wind Speed (m/s): 8.37
# of Samples: 8521
Hour Shift to LST: -6
Midnight LST
Noon LST
Wind Speed (m/s)
10
9
8
7
6
0
300
600
900
1200
1500
1800
2100
2400
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

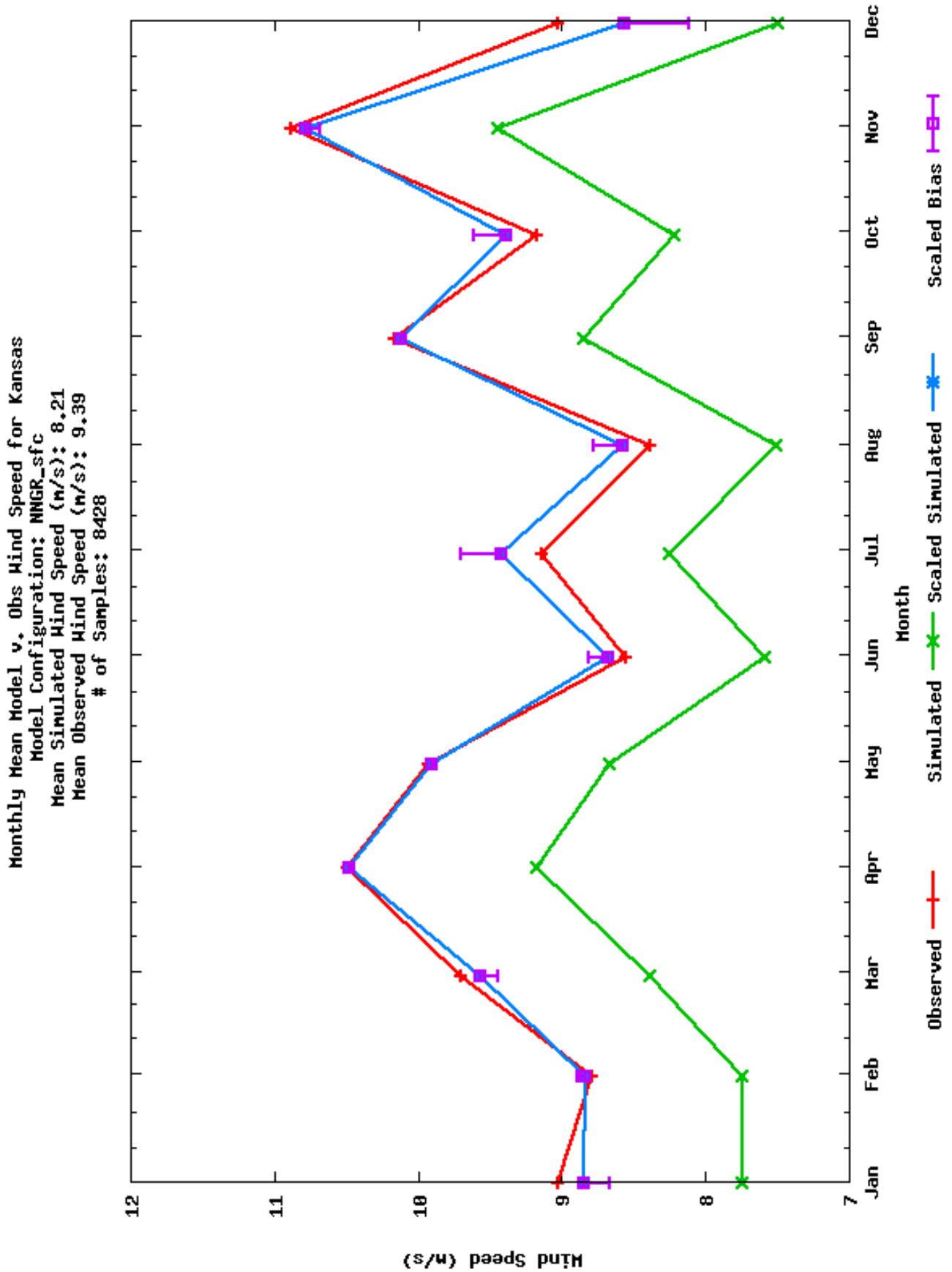

Annual Mean Model v. Obs Wind Speed for New York
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 7.00
Mean Observed Wind Speed (m/s): 7.48
# of Samples: 8238
Hour Shift to LST: -5
Midnight LST
Noon LST
Wind Speed (m/s)
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

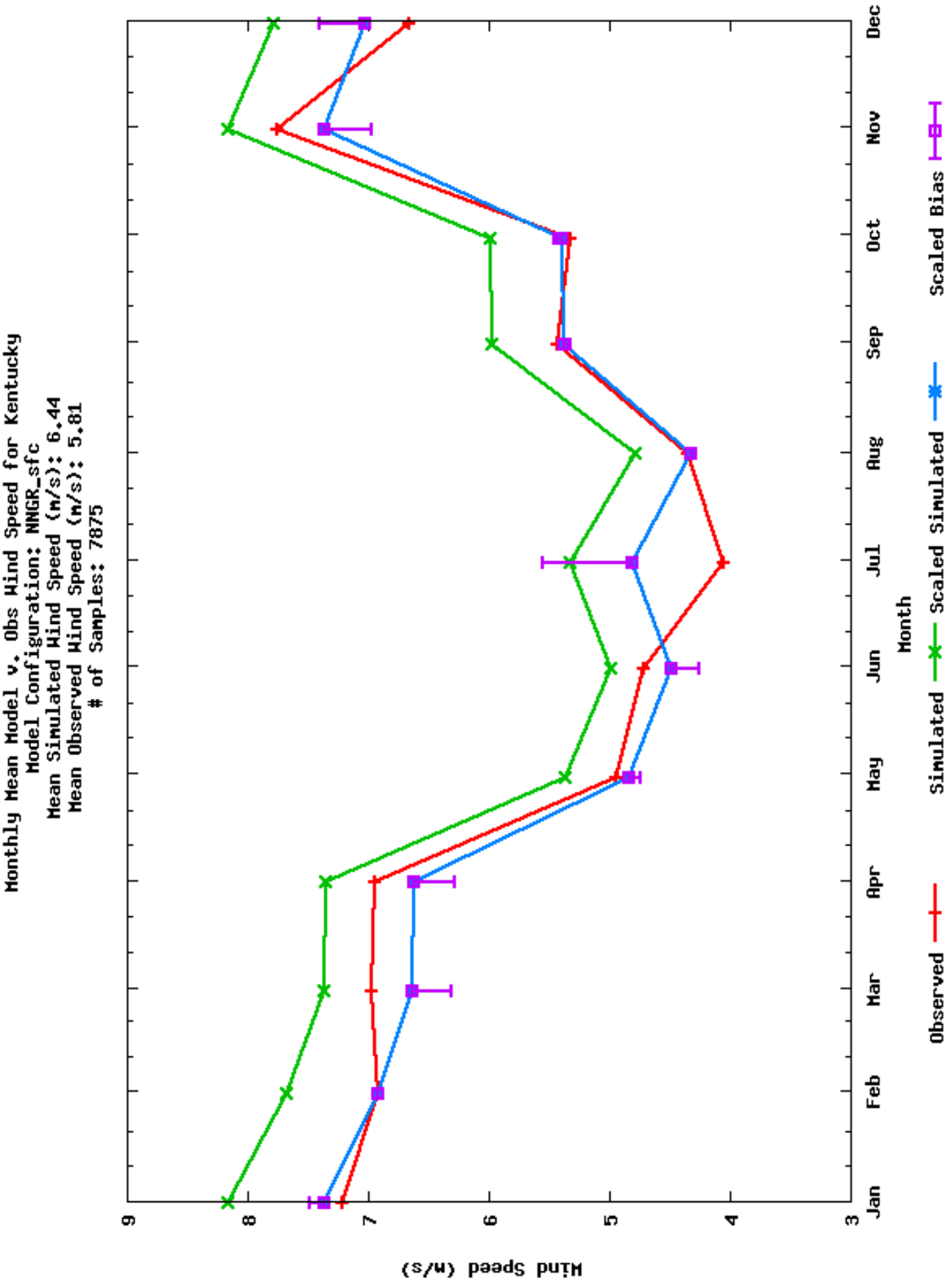

Annual Mean Model v. Obs Wind Speed for South Dakota
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 8.24
Mean Observed Wind Speed (m/s): 9.50
# of Samples: 8121
Hour Shift to LST: -6
Midnight LST
Noon LST
Wind Speed (m/s)
12
11
10
9
8
7
0
300
600
900
1200
1500
1800
2100
2400
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

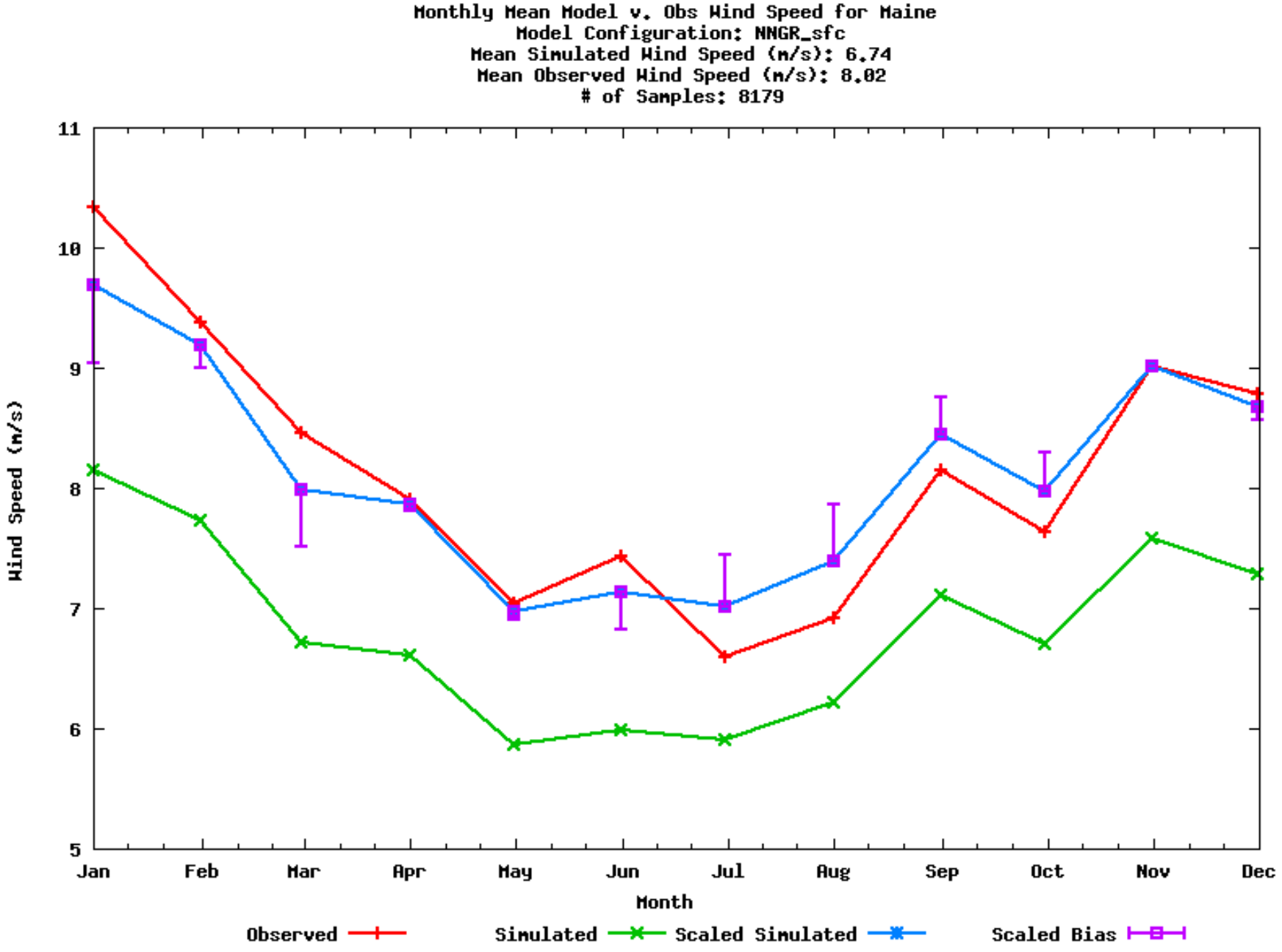

Annual Mean Model v. Obs Wind Speed for Texas
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 8.18
Mean Observed Wind Speed (m/s): 8.86
# of Samples: 6176
Hour Shift to LST: -6
Midnight LST
Noon LST
Wind Speed (m/s)
11
10
9
8
7
6
0
300
600
900
1200
1500
1800
2100
2400
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias

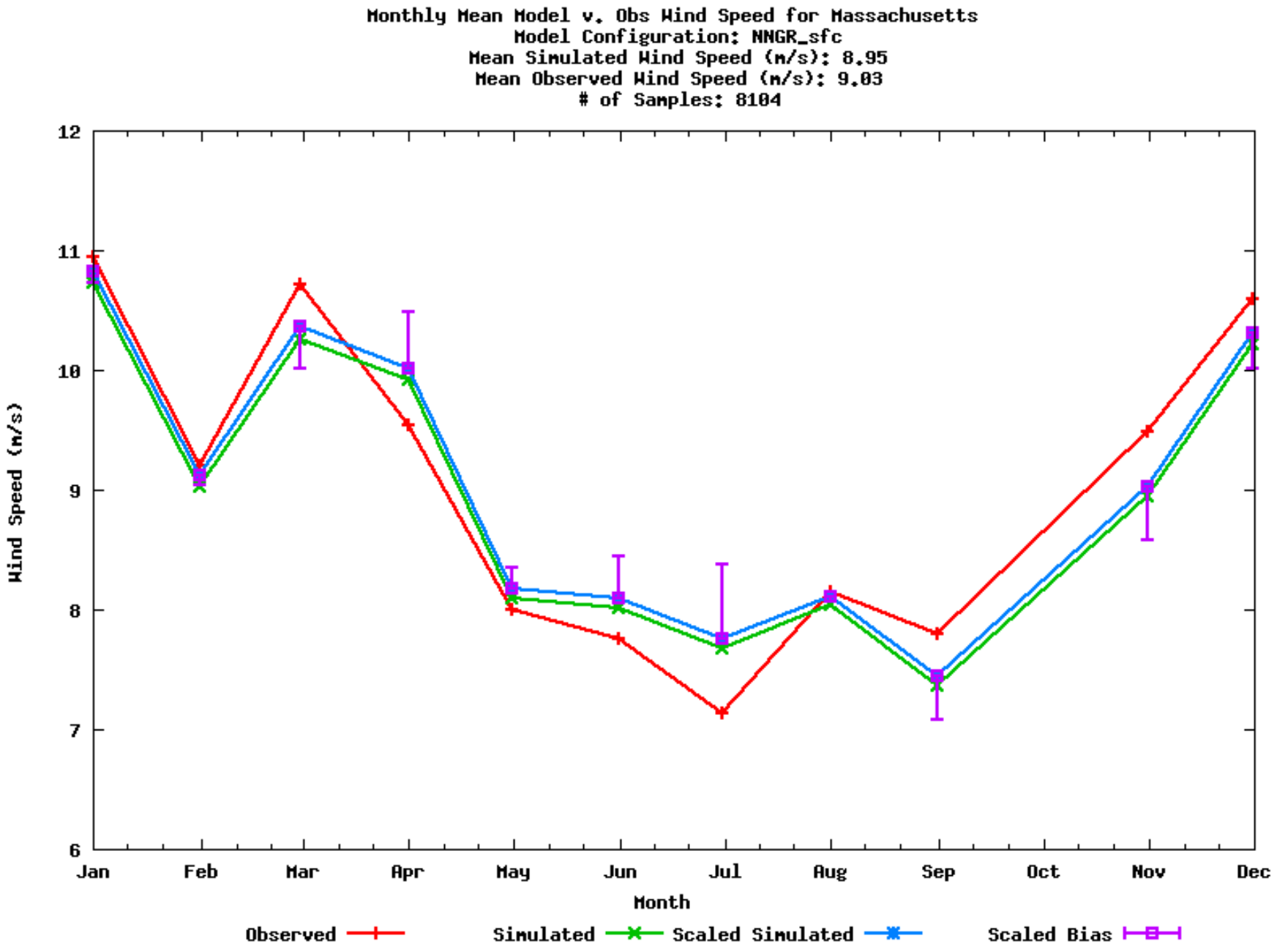

Annual Mean Model v. Obs Wind Speed for West Virginia
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 6.83
Mean Observed Wind Speed (m/s): 6.98
# of Samples: 8401
Hour Shift to LST: -5
Midnight LST
Noon LST
Wind Speed (m/s)
Model Hour (UTC)
Observed
Simulated
Scaled Simulated
Scaled Bias
9
8
7
6
5
4
0
300
600
900
1200
1500
1800
2100
2400

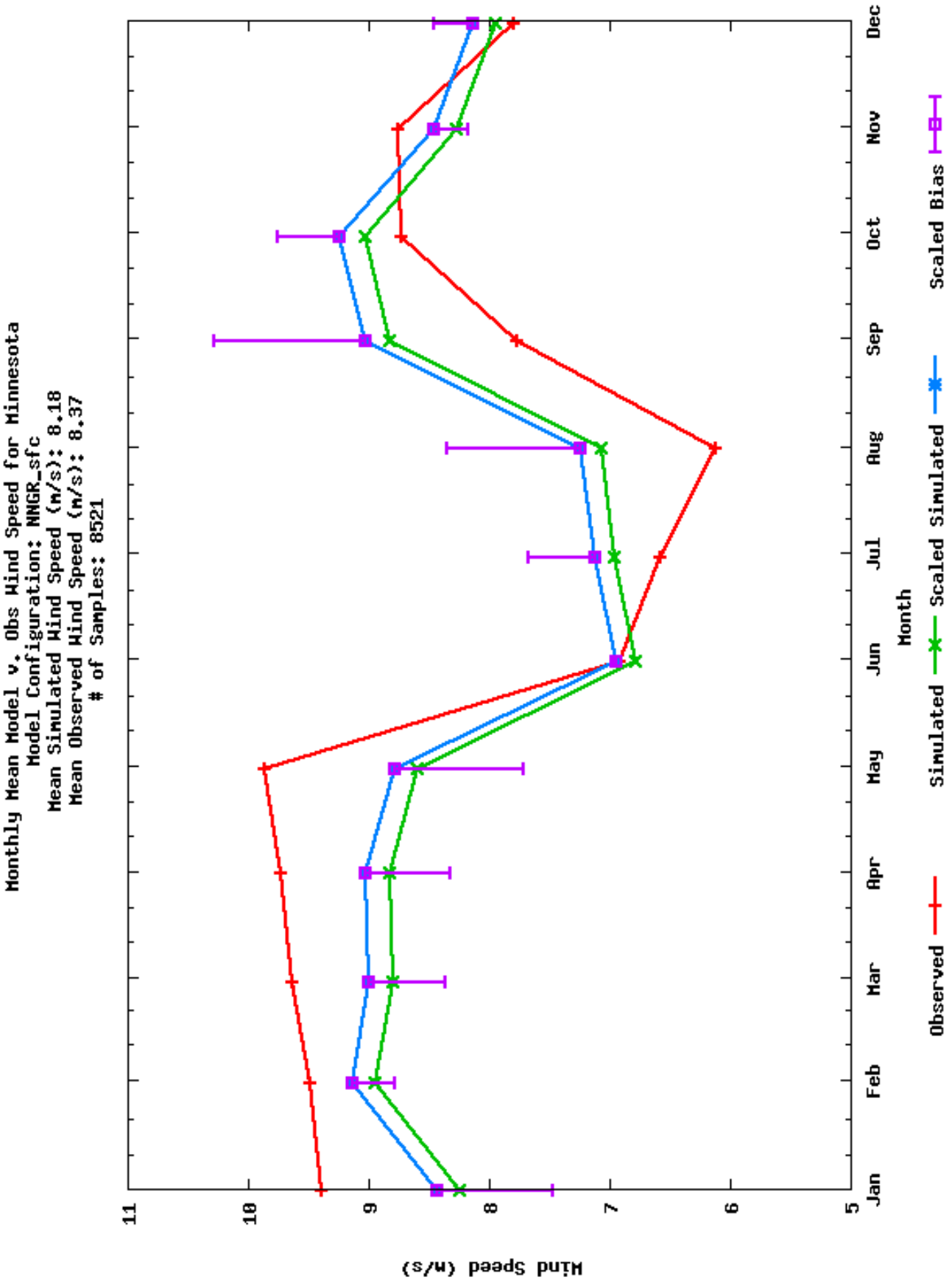

Monthly Mean Model v. Obs Wind Speed for Indiana
Model Configuration: MNGR_sfc
Mean Simulated Wind Speed (m/s): 7.46
Mean Observed Wind Speed (m/s): 8.13
# of Samples: 8556
Wind Speed (m/s)
11
10
9
8
7
6
5
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias

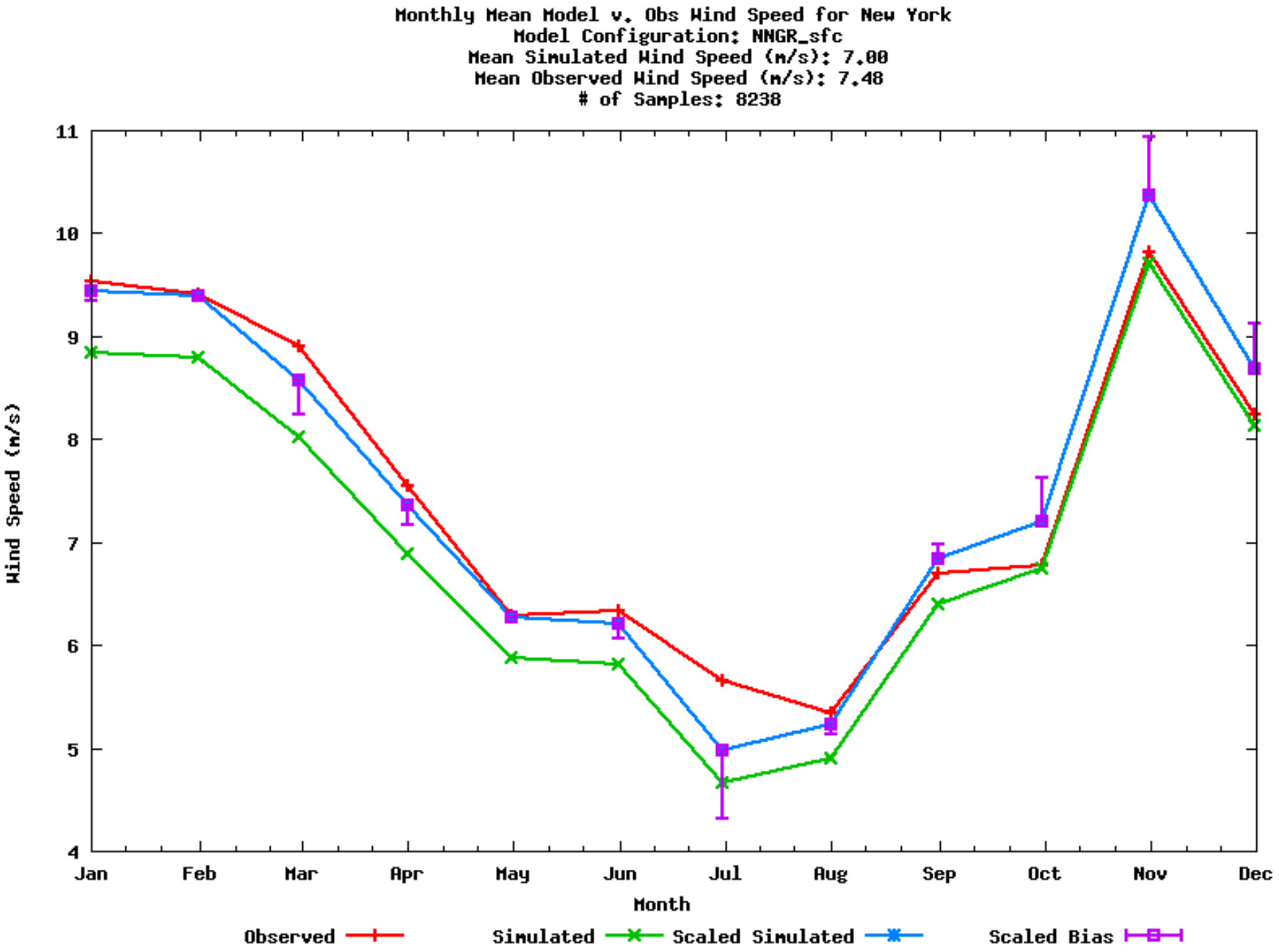

Monthly Mean Model v. Obs Wind Speed for Kansas
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 8.21
Mean Observed Wind Speed (m/s): 9.39
# of Samples: 8428
Wind Speed (m/s)
12
11
10
9
8
7
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias

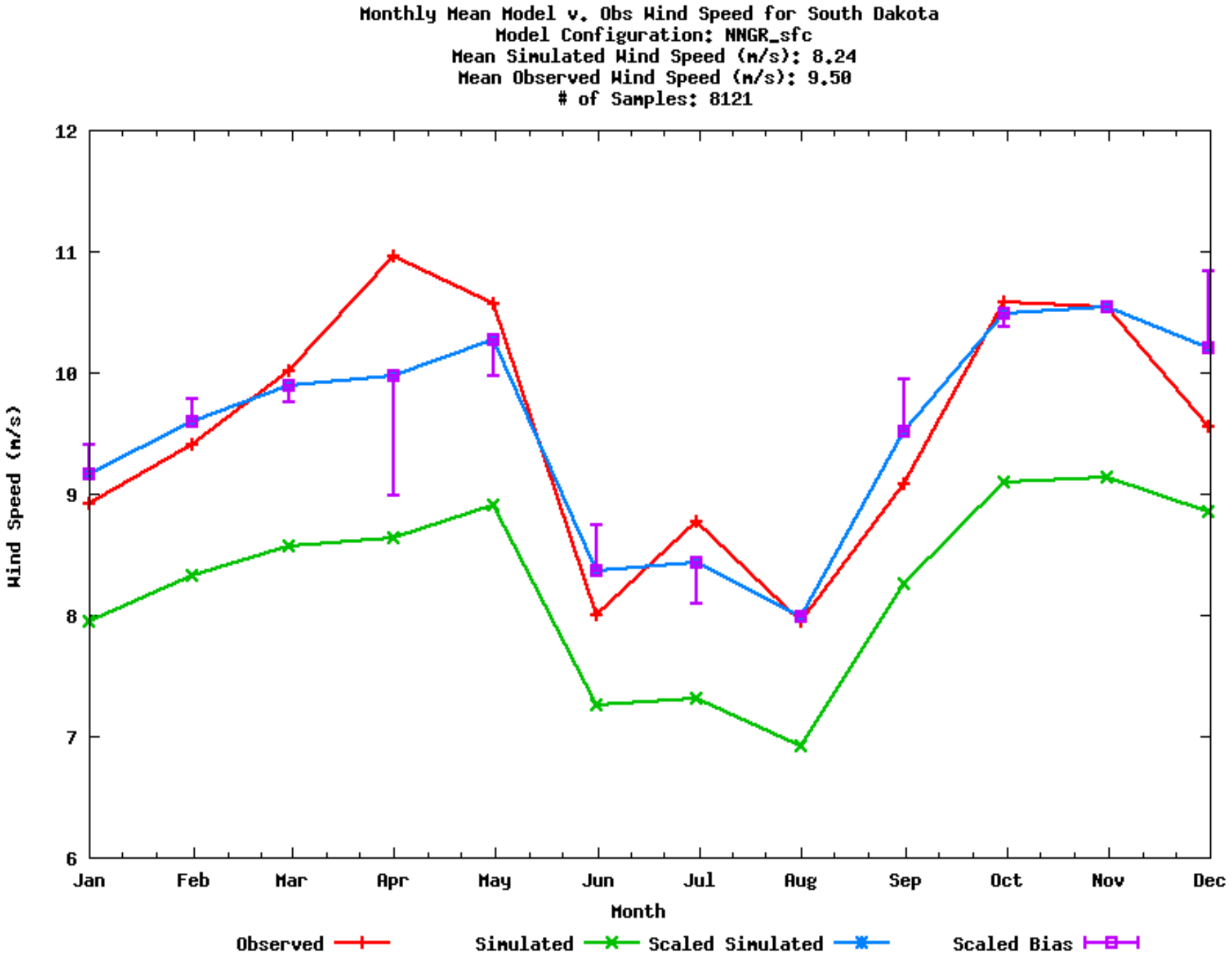

Monthly Mean Model v. Obs Wind Speed for Kentucky
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 6.44
Mean Observed Wind Speed (m/s): 5.81
# of Samples: 7875
Wind Speed (m/s)
9
8
7
6
5
4
3
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias

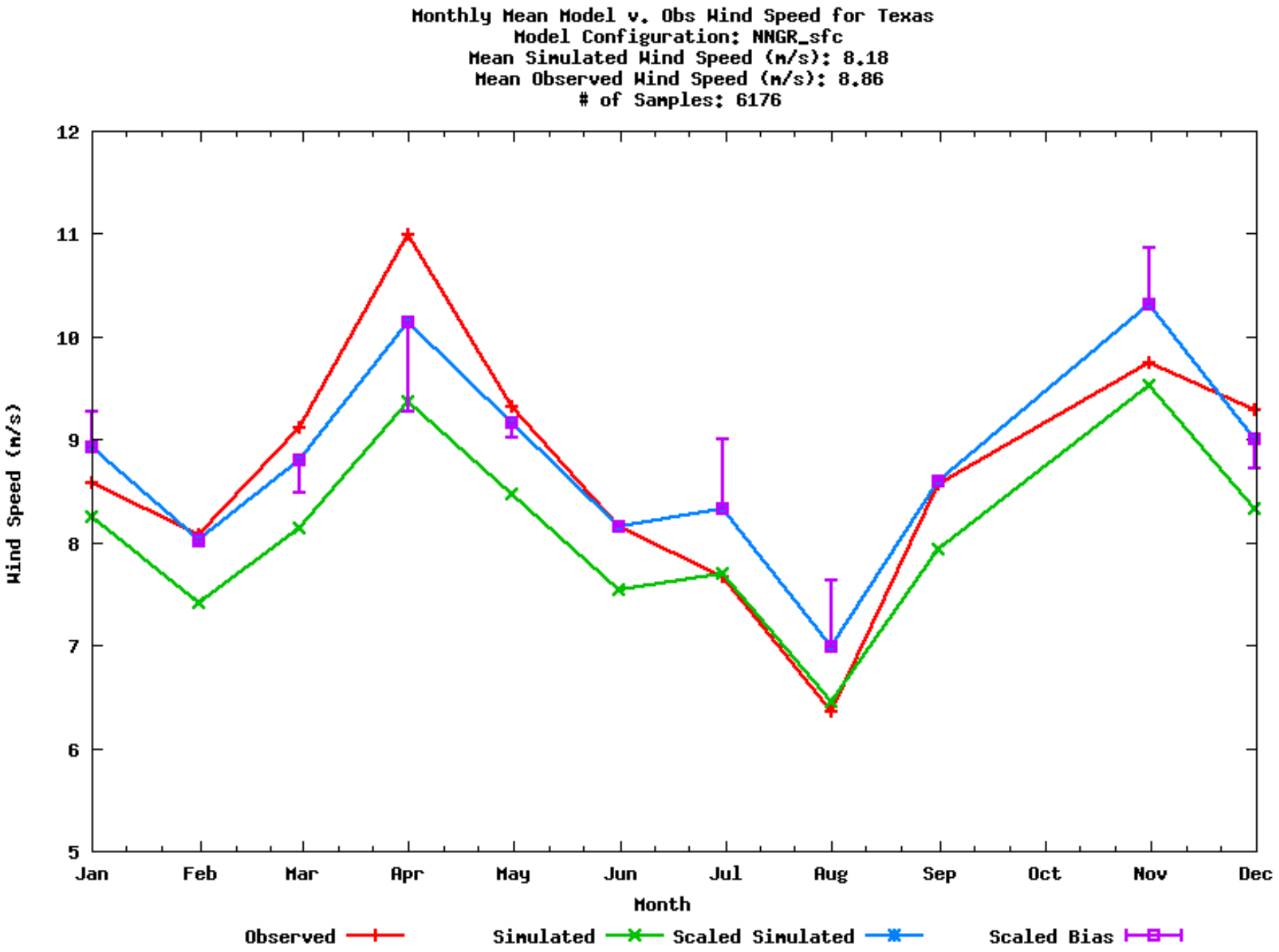

Monthly Mean Model v. Obs Wind Speed for Maine
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 6.74
Mean Observed Wind Speed (m/s): 8.02
# of Samples: 8179
Wind Speed (m/s)
11
10
9
8
7
6
5
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias

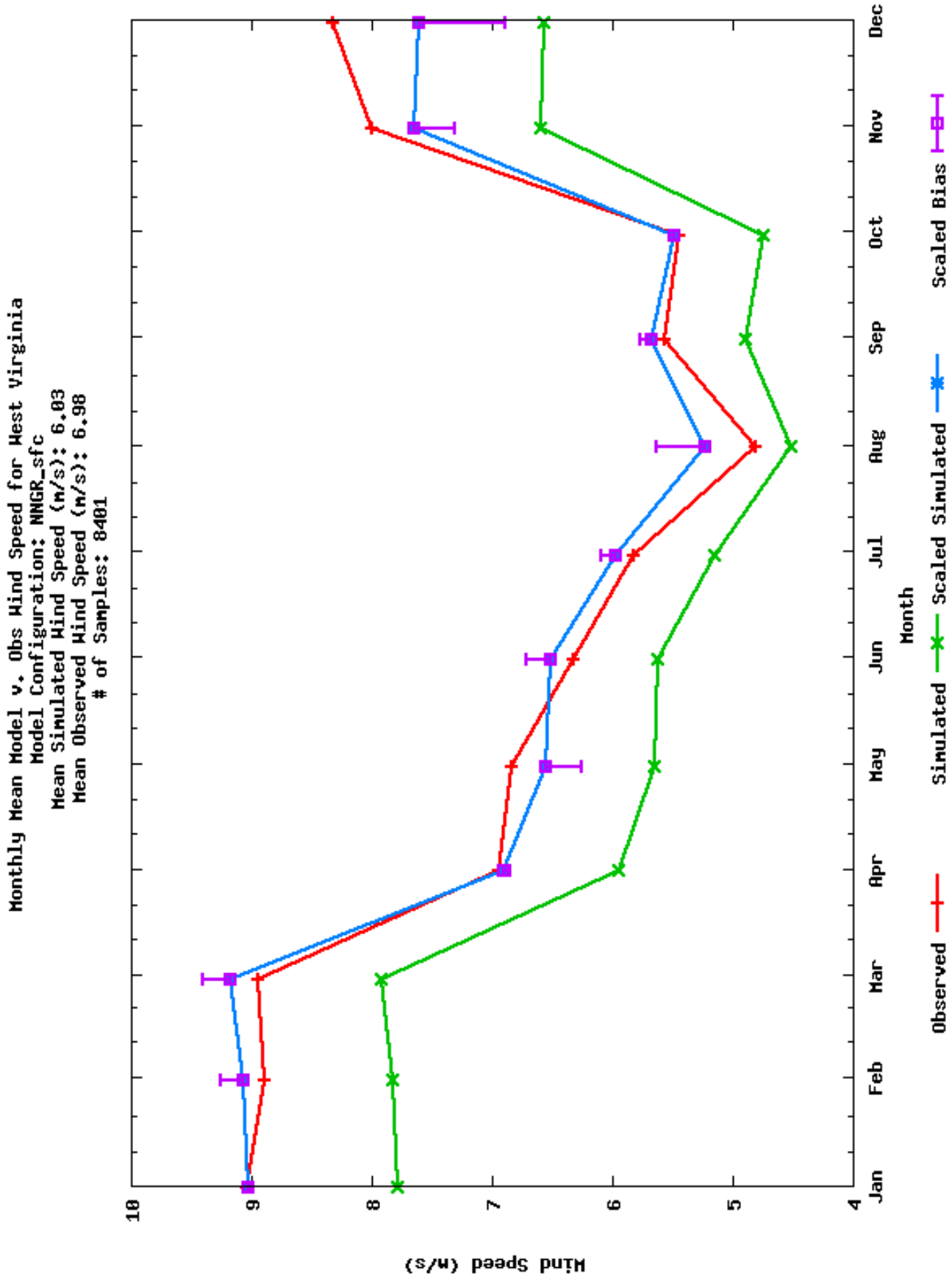

Monthly Mean Model v. Obs Wind Speed for Massachusetts
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 8.95
Mean Observed Wind Speed (m/s): 9.03
# of Samples: 8104
Wind Speed (m/s)
12
11
10
9
8
7
6
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias


Monthly Mean Model v. Obs Wind Speed for Minnesota
Model Configuration: MNGR_sfc
Mean Simulated Wind Speed (m/s): 8.18
Mean Observed Wind Speed (m/s): 8.37
# of Samples: 8521
Wind Speed (m/s)
11
10
9
8
7
6
5
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias


Monthly Mean Model v. Obs Wind Speed for New York
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 7.00
Mean Observed Wind Speed (m/s): 7.48
# of Samples: 8238
Wind Speed (m/s)
11
10
9
8
7
6
5
4
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias


Monthly Mean Model v. Obs Wind Speed for South Dakota
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 8.24
Mean Observed Wind Speed (m/s): 9.50
# of Samples: 8121
Wind Speed (m/s)
12
11
10
9
8
7
6
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias


Monthly Mean Model v. Obs Wind Speed for Texas
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 8.18
Mean Observed Wind Speed (m/s): 8.86
# of Samples: 6176
Wind Speed (m/s)
12
11
10
9
8
7
6
5
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias


Monthly Mean Model v. Obs Wind Speed for West Virginia
Model Configuration: NNGR_sfc
Mean Simulated Wind Speed (m/s): 6.03
Mean Observed Wind Speed (m/s): 6.98
# of Samples: 8401
Wind Speed (m/s)
10
9
8
7
6
5
4
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
Observed
Simulated
Scaled Simulated
Scaled Bias

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
December 2009	Subcontract Report	March 2008 to March 2010

4. TITLE AND SUBTITLE
Development of Eastern Regional Wind Resource and Wind Plant Output Datasets: March 3, 2008 – March 31, 2010

5a. CONTRACT NUMBER
DE-AC36-08-GO28308

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
M. Brower

5d. PROJECT NUMBER
NREL/SR-550-46764

5e. TASK NUMBER
WER7.7402 & WER8.5004

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AWS TRUEWIND, LLC
263 NEW KARNER ROAD
ALBANY, NEW YORK, 12205

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Renewable Energy Laboratory
1617 Cole Blvd.
Golden, CO 80401-3393

10. SPONSOR/MONITOR'S ACRONYM(S)
NREL

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
NREL/SR-550-46764

12. DISTRIBUTION AVAILABILITY STATEMENT
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

13. SUPPLEMENTARY NOTES
NREL Technical Monitor: David Corbus

14. ABSTRACT *(Maximum 200 Words)*
The objective of this project was to provide wind resource inputs to the Eastern Wind Integration and Transmission Study.

15. SUBJECT TERMS
Wind; integration; datasets; EWITS; utilities; RTO; ISO; modeling; simulation; mesoscale; AWS Truewind.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (*Include area code)*
Unclassified	Unclassified	Unclassified	UL		

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18