



Final Report on the Creation of the Wind Integration National Dataset (WIND) Toolkit and API

October 1, 2013 – September 30, 2015

Vaisala, Inc.
Louisville, Colorado

NREL Technical Monitor: Bri-Mathias Hodge

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

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

Subcontract Report
NREL/SR-5D00-66189
April 2016

Contract No. DE-AC36-08GO28308



Final Report on the Creation of the Wind Integration National Dataset (WIND) Toolkit and API

October 1, 2013 – September 30, 2015

*Vaisala, Inc.
Louisville, Colorado*

NREL Technical Monitor: Bri-Mathias Hodge
Prepared under Subcontract No. AGV-2-22460-01

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

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

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

Subcontract Report
NREL/SR-5D00-66189
April 2016

Contract No. DE-AC36-08GO28308

**This publication was reproduced from the best available copy
submitted by the subcontractor and received no editorial review at NREL.**

NOTICE

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

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

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

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

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

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

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

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

NREL prints on paper that contains recycled content.

Acknowledgments

The work described herein was conducted by Vaisala, Inc. (previously 3TIER, Inc.) under NREL subcontract AGV-2-22460-01, and was supervised by two NREL technical monitors, Kirsten Orwig and Bri-Mathias Hodge, without whom it could not have had a successful completion. Additional assistance was received at NREL from Wes Jones, Caroline Draxl, Dan Getman, Julie Lundquist, John Michalakes, Kara Clark, Debbie Lew, Aaron Bloom, Patrick Moriarty, Marissa Hummon, Nate Blair, Dennis Elliott, Keith Searight, George Scott, and Avi Pukayastha.

At Vaisala, significant contributions were made by Jim McCaa, Sara Harrold, Mark Stoelinga, Eric Gruit, Paddy Fowler, Mike Dunsmuir, Rhea George, Clark Kirkman, and Pascal Storck.

In addition, we must acknowledge a generous and engaged technical review committee (TRC), comprised of the following individuals:

- Department of Energy (DOE): Joel Cline
- Environment Canada: Wei Yu
- General Electric (GE): Gary Jordan
- National Oceanic and Atmospheric Administration (NOAA): Jim Wilczak
- National Center for Atmospheric Research (NCAR): Jimmy Dudia
- RePPAE: Jack King
- Ventyx: Jason Schmidt
- Western Electricity Coordinating Council (WECC): Keegan Moyer, Dan Beckstead, Bradley Nickell
- WindLogics: Mark Ahlstrom

Table of Contents

Introduction	1
The Meteorological Dataset	3
Overview of Modeling Approach	3
Selection of Model Physics Configuration	5
Descriptions of Configurations	5
Analysis of Model Configurations	6
Validation of Model Configurations	9
Selection of Final Configuration	13
Model Simulations	14
The Wind Power Production Dataset	15
Site Selection	15
Selection of Onshore Sites	16
Selection of Offshore Sites	18
Power Production Profiles	18
Simulation of Forecast Data	21
24-Hour Lead-Time Forecasts	22
Short Lead-Time Forecasts	24
Probabilistic Forecasts	26
The Data Extraction Tool	26
Summary	27
Appendix A: WPS Namelist	28
Appendix B: WRF Namelist	29

List of Figures

Figure 1. Map showing single 2km WRF domain used for the NWP simulations	3
Figure 2. Variation of high-frequency wind speed changes over the course of individual simulations	7
Figure 3. Daily mean 80m wind speeds at Goodnoe Hills by day of year, 2008. Vertical units are meters per second.	8
Figure 4. Daily mean 2m temperature at Goodnoe Hills by day of year, 2008. Vertical units are degrees K.....	9
Figure 5. Diurnal cycle of wind speed at Crosby County, Texas.....	11
Figure 6. Monthly mean wind speed at Crosby County, Texas.....	11
Figure 7. Map of selected sites.....	15
Figure 8. ERGIS power curves	19
Figure 9. Spatial and temporal autocorrelation of the WIND Toolkit data compared to observations for a project in complex terrain	20
Figure 10. Forecast error validation locations	21
Figure 11. Example raw WRF forecasts (left panel) compared to centered 13-hr smoothing of raw WRF forecasts blended with 20% weight on the 13-hr smoothed actuals (right panel).....	24
Figure 12. Graphical example of forecast algorithm: 40min lead time	25
Figure 13. Graphical example of forecast algorithm: 80min lead time	25
Figure 14. Graphical example of forecast algorithm: 0min lead time	25

List of Tables

Table 1. Meteorological Towers Used for Configuration Selection.....	10
Table 2. Correlation of Monthly Mean Wind Speed to Observed	12
Table 3. Correlation of Diurnal Mean Wind Speed to Observed.....	13
Table 4. Forecast Error Validation Locations	22
Table 5. Iterations of Day-Ahead Forecast Smoothing.....	23
Table 6. Iterations of Short-Term Forecast Parameters	26

Introduction

One of the largest concerns with wind energy is the cost associated with integrating significant amounts of it into regional power systems. These costs largely manifest themselves through an increase in the ancillary services and/or regulation reserves required to maintain power system reliability while integrating an intermittent and non-dispatchable energy resource. To better understand these costs, wind integration studies are undertaken.

A primary requirement for completing such studies is an understanding of the dynamic behavior of the region's wind energy resource that can be synchronized with electric load and other resource data. In general, observational data are insufficient to fully support the work, and this limitation becomes more severe as the size of the study increases.

To overcome the limitations of purely observational data sets, wind integration studies have relied on weather simulation models to integrate regional observational data sets into a comprehensive representation of the region's wind resource. The result of these simulation models is a synthetic weather data set, based largely on observations, which can serve as the foundation of wind energy integration studies. As these data sets are generated by computer simulation, they are continuous in space and time, meaning that they cover the entire region of interest for the entire time period that is simulated in a consistent manner. This completeness makes these data sets ideal for integration studies.

The primary objective of this work was to create a state-of-the-art national wind resource dataset and to provide detailed wind plant output data for specific sites based on that dataset. Corresponding retrospective wind forecasts were also included at all selected locations. The combined information from these activities was used to create the Wind Integration National Dataset (WIND), and an extraction tool was developed to allow web-based data access.

The final dataset is described in the following documents:

- **Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit** -- C. Draxl, B.-M. Hodge, A. Clifton, and J. McCaa. NREL Technical Report NREL/TP-5000-61740, April 2015.
- **The Wind Integration National Dataset (WIND) Toolkit** -- C. Draxl, A. Clifton, B.-M. Hodge, and J. McCaa. Applied Energy, Volume 151, 2015, Pages 355–366.
- **Validation of Power Output for the WIND Toolkit** -- J. King, A. Clifton, and B.-M. Hodge. NREL Technical Report NREL/TP-5D00-61714, September 2014.

Vaisala's work consisted of four main tasks, described in the subsequent sections of this report:

1. Development of the meteorological dataset
2. Development of the wind power production dataset
3. Simulation of forecast data
4. Development of the data extraction tool

These activities were conducted over a two-year period spanning 2012-2014, and this report summarizes Vaisala's activities under the project.

The Meteorological Dataset

The meteorological dataset is comprised of condensed numerical weather prediction (NWP) model output covering the 48 contiguous U.S. states and offshore regions out to 50 nautical miles on a single 2km horizontal resolution grid. The time resolution of the dataset is 5 minutes, and it covers a 7-year period spanning 2007 through 2013. Some parameters were output directly from the NWP model, and others required minimal processing, e.g. computing wind speed from the meridional and zonal components. The dataset is stored as a series of NetCDF files, each containing all of the parameters listed below at a single snapshot in time.

The following parameters were output from the raw NWP meteorological simulations on the full horizontal grid (i.e. at every model grid point):

- Wind speed and direction at eight heights (10m, 40m, 60m, 80m, 100m, 120m, 140m, and 160m)
- Temperature of the ground surface and the air temperature at the same eight heights
- Specific humidity, friction velocity, and Monin-Obukhov length at 2 m above the ground
- Air pressure, direct normal irradiance, diffuse irradiance, global horizontal irradiance, and total precipitation rate at the ground surface.

Overview of Modeling Approach

The fundamental tool employed in the creation of the meteorological dataset was the Weather Research and Forecasting (WRF) modeling framework. WRF is maintained by the U.S. National Center for Atmospheric Research (NCAR), but includes contributions from institutions worldwide. Specifically, version 3.4.1 of WRF Advanced Research and Weather (ARW) version was used. Output from the model formed the foundation of the wind resource dataset.

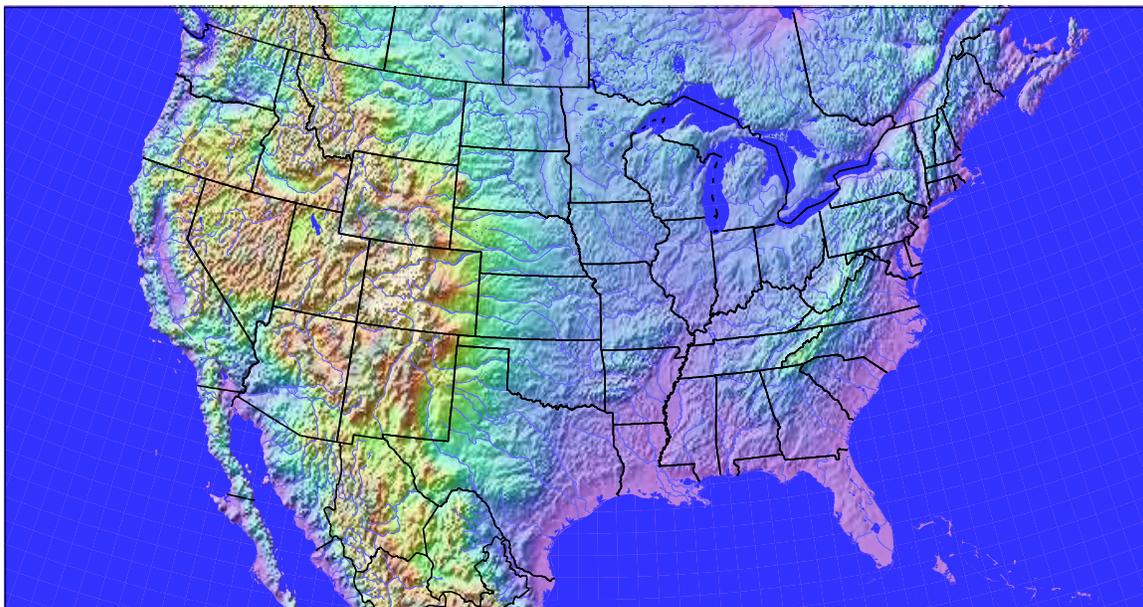


Figure 1. Map showing single 2km WRF domain used for the NWP simulations

The region was modeled at a 2km horizontal resolution, with output archived at a 5-minute interval of simulated time. To eliminate spatial ‘seams’ in the data, a single domain for the entire region was used, as shown in Figure 1. Temporal seams were minimized through month-long simulations, with only a single time discontinuity at the beginning of each month. Use of a single large spatial extent combined with the longest possible temporal simulations was motivated by a desire to create the least number of discontinuities that might affect the subsequent spatial and temporal processing. Internal model drift was minimized through a resolution-dependent data relaxation technique described below.

The WRF model was implemented as a set of nested grids, with progressively higher spatial resolutions of approximately 54, 18, 6, and 2 km from the outermost to the innermost domain. All inner grids included buffer zones in each direction to avoid grid edge effects. The WRF preprocessing namelists and associated output files describing exact locations of the grid points were delivered to NREL. For convenience, the preprocessing namelist is also included here as Appendix A.

Individual model simulations were made for a period of one month, including a 48-hour spin up period over the last days of the previous month. Additionally, simulations continued for 24 hours into the following month to allow for subsequent blending of simulations over the breaks. However, the dataset described here has a simple split between simulations occurring between 23:50 UTC on last day of the earlier month and 00:00 UTC on the first day of the following month.

The primary driver for the WRF simulations was the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis (ERA Interim) dataset. Several studies have shown this to be superior to both the first and second reanalysis datasets from the National Centers for Environmental Prediction (NCEP).

The ERA Interim dataset, like all reanalysis products, includes data from historic global weather archives, which are maintained by operational weather forecasting centers around the world. The primary reasons for its enhanced skill over earlier reanalysis efforts are improved model physics, finer spatial resolution, and the use of a sophisticated 4-dimensional variational data assimilation system. These global archives represent the overall state of the atmosphere over the entire planet and are themselves the result of a sophisticated computer analysis of available surface and upper air observations. The ERA Interim dataset was developed specifically targeting use of data from the modern satellite era.

In addition to the advances in the reanalysis modeling itself, the ERA Interim dataset was run at higher resolution (approximately 80 km) than earlier NCEP reanalysis data sets. This made it possible to more directly influence WRF simulations with higher-resolution atmospheric phenomena.

To accurately resolve the regional wind fields requires an ability to model the interaction of large scale weather systems with the varied terrain, land-use and vegetation of the region. An accurate representation of the local terrain is also important for resolving thermally driven circulations caused by differential heating and cooling of the land surface. In this case, the WRF domain was created using the U.S. Geological Survey (USGS) GTOPO30 dataset, which provides a global

30-second (roughly 900 m) digital representation of land surface topography. Land use classifications came from the 3-second National Land Cover Dataset (NCLD). A 5-minute soil texture provided by the Food and Agriculture Organization of the United Nations (FAO, www.fao.org) and a 0.15-degree monthly climatology green vegetation fraction provided by National Environmental Satellite, Data, and Information Service (NESDIS, www.nesdis.noaa.org) were employed. These datasets are available via the standard WRF Preprocessing System (WPS) distribution available from NCAR. They were used to describe the height, roughness, and soil properties of the earth's surface for the period of simulation.

Selection of Model Physics Configuration

As with any model, the WRF model has various configuration options that control the model's behavior and can affect the results of the simulation. This section describes the NWP model configuration testing for the NREL wind integration data set. Due to the large cost of the production runs for the final meteorological data set, it was important to have as much confidence as possible in the configuration of the NWP model (WRF 3.4.1) before the main simulations commenced. Keeping in mind the focus of the dataset as a tool for wind integrations studies, the configuration selection was centered on numerical stability (particularly with regard to near-surface winds) and fidelity to historical wind speed observations.

Descriptions of Configurations

Vaisala originally proposed to consider four different model configurations for the study, with the plan to test the configurations using the full domain but at a reduced model horizontal resolution of 18km. That number was expanded to eight, primarily due to suggestions from the technical review committee (TRC).

The testing of the configurations was set up as a sensitivity study, with a base configuration and five single-parameter changes to it. Two more configurations were then tested, with the final one including a combination of promising elements from the initial tests. In collaboration with the TRC, the eighth configuration was selected for subsequent tests at the full 2km resolution.

All configurations shared the following properties. The primary driver for the simulations was the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis (ERA Interim) dataset. Height, roughness, and soil properties of the earth's surface for the period of simulation were set as described above. Three full years of simulation were conducted for each model configuration.

Full namelists for each configuration were delivered to NREL. These, combined with the publicly available input files and the grid files discussed above should be sufficient to reproduce all WRF simulations used in the project. However, performance modifications to WRF implemented at NREL (such as use of parallel asynchronous I/O) were critical to the timely completion of the project

The base simulation (config 1) consisted of 31 full eta levels (at the default WRF heights) and had the following configuration:

- Planetary Boundary Layer Scheme - YSU model (MRF with entrainment)

- Surface Parameterization - Monin-Obukhov similarity model
- Land-surface option - Thermal Diffusion Scheme, 5 layer soil diffusivity model
- Nudging – traditional Newtonian relaxation.

The remaining configurations were as follows:

Configuration 2: The number of full eta levels was increased to 51, using a set of levels used in vertical level sensitivity testing conducted at the Developmental Testbed Center. See: http://verif.rap.ucar.edu/eval/wrfr vl/WRFR-R-VL_FINAL_report.pdf

Configuration 3: Scale selective nudging was enabled. While WRF simulations are underway, grids are continuously relaxed, or nudged, toward the large-scale reanalysis to prevent drift of the simulations away from the analyzed synoptic patterns. Domain-resolution-dependent nudging, or scale selective nudging, allows us to nudge very large domains based on the number of wavenumbers resolvable within the domain. Results for 3TIER using this technique have been positive, but because the scale-selective algorithm is relatively new, it was included as a configuration test.

Configuration 4: The NOAA Land-Surface Model – The NOAA LSM simulates soil temperature and moisture in four layers as well as snowpack variables, canopy water content and the energy flux and water flux terms of both the surface energy and water balance equations. By contrast the thermal diffusion scheme is much simpler and uses soil temperature only in five layers.

Configuration 5: The QNSE boundary layer scheme - This new PBL parameterization has been shown to maintain a meaningful mixed layer in the presence of strong stratification and includes a turbulent kinetic energy prediction option. The configuration test will examine how this parameterization compares to the simpler YSU model.

Configuration 6: The number of full eta levels was increased to 41, using a set of levels developed by 3TIER in an attempt to concentrate levels near the surface while maintaining numerical stability.

Configuration 7: Topographic wind enhancement – the Jiménez and Dudhia (see: <http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-11-084.1>) parameterization to improve the effects of local complex topography on near-surface winds was included. This configuration was run at the same vertical resolution as configuration 6, and departs from the one-parameterization sensitivity tests.

Configuration 8: This configuration was comprised of the base state with the changes described in configurations 3, 4, 6, and 7, based on their mildly positive impacts from the earlier testing.

Analysis of Model Configurations

Initial analysis of the log files and raw WRF output files revealed no significant numerical problems except with configuration 2, which showed significant numerical instability, resulting in CFL violations and model crashes. It was necessary to reduce the domain 1 time step from 300s to 240s in order to get all configuration 2 simulations to successfully complete. We looked for time instabilities for the other configurations by inspecting time series for individual points

within the domains, and visually inspected the grids for signs of 2-delta-x noise. As usual, some standing wave patterns were observed at outflow boundaries, but these were located within 20 grid points of the edges of the domain, and were sufficiently far from the domain of interest as to be tolerable.

We checked for drift in the simulations in two ways. The first check was on the high-frequency variability of the wind speed at various locations. Figure 2 shows the mean across all sites of the standard deviation of hourly wind speed changes as function of the day of the month (all simulations start one day prior to the beginning of the month).

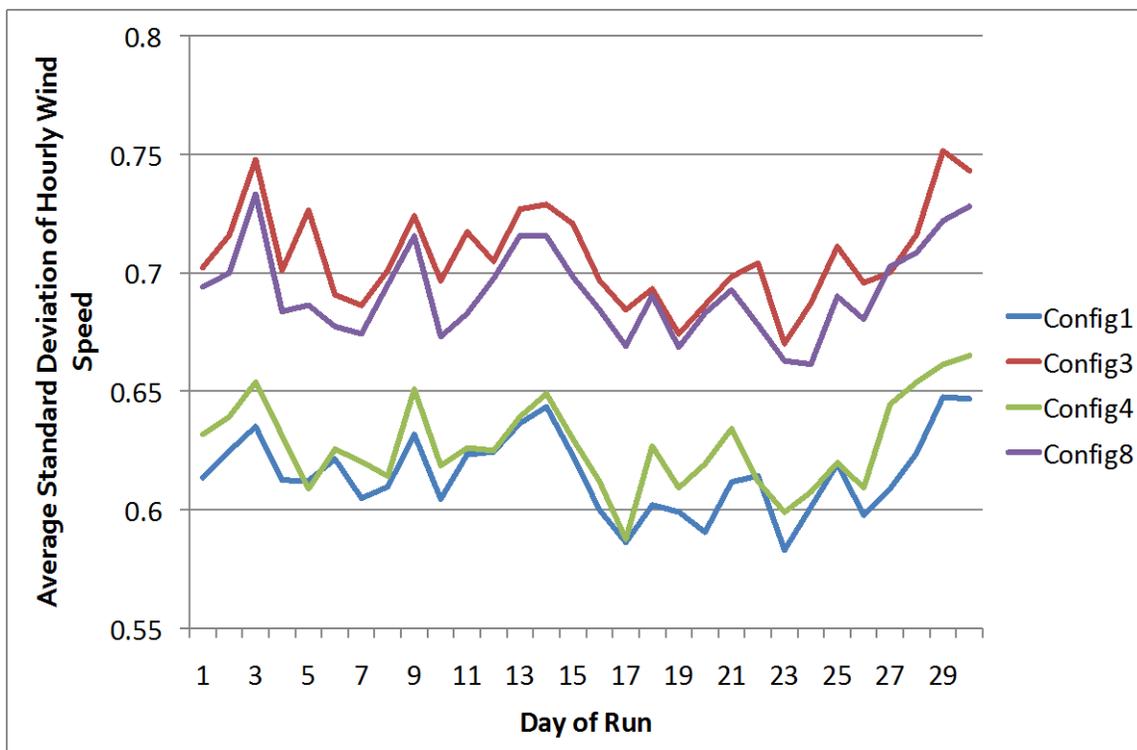


Figure 2. Variation of high-frequency wind speed changes over the course of individual simulations

Configurations 1, 3, 4, and 8 are shown, as these capture the range of results. None show a significant drift over the course of the simulations, which is an indication that the length of the simulations will be acceptable for wind integration analysis, where large change over time in this metric could lead to spurious conclusions.

The second test for model drift was inspection of daily mean values of wind speed and temperature. We were checking here for evidence of a saw tooth pattern at a one-month period that would indicate a drift away from the resolved seasonal cycle.

Figure 3 and Figure 4 show the results at one of our validation sites, Goodnoe Hills, located in the Columbia River gorge in Washington State. These are typical of the plots at all sites in that they show no distinct discontinuities at the month boundaries.

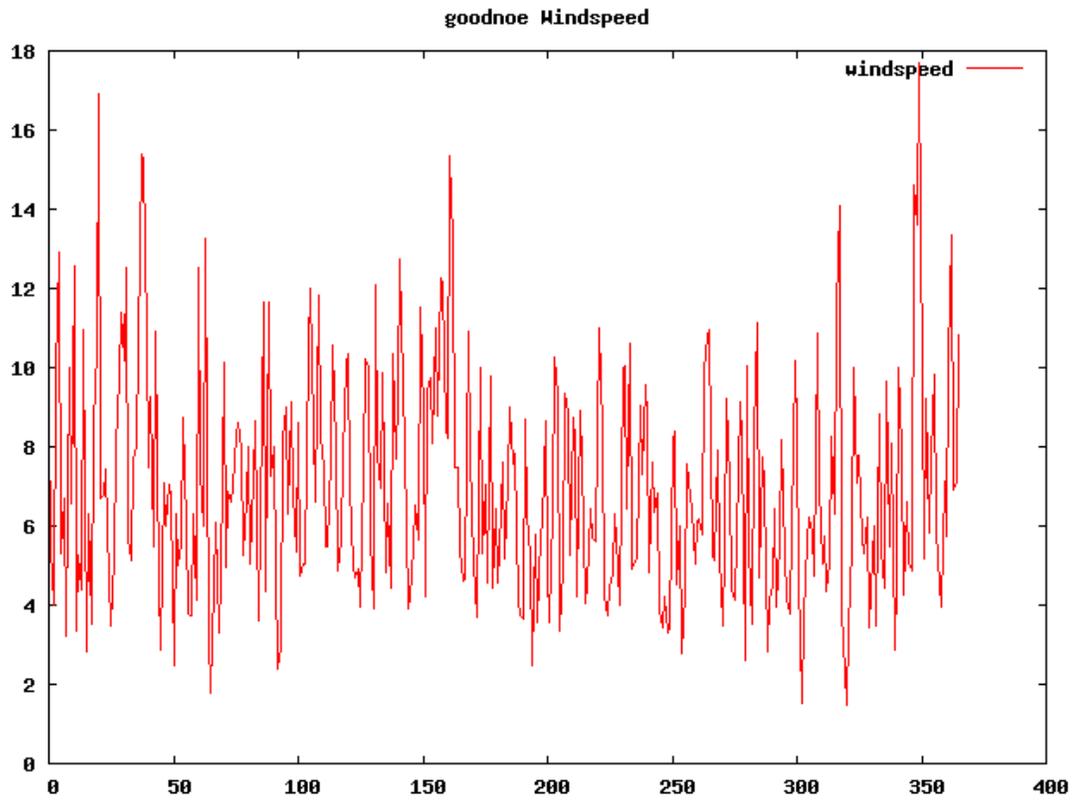


Figure 3. Daily mean 80m wind speeds at Goodnoe Hills by day of year, 2008. Vertical units are meters per second.

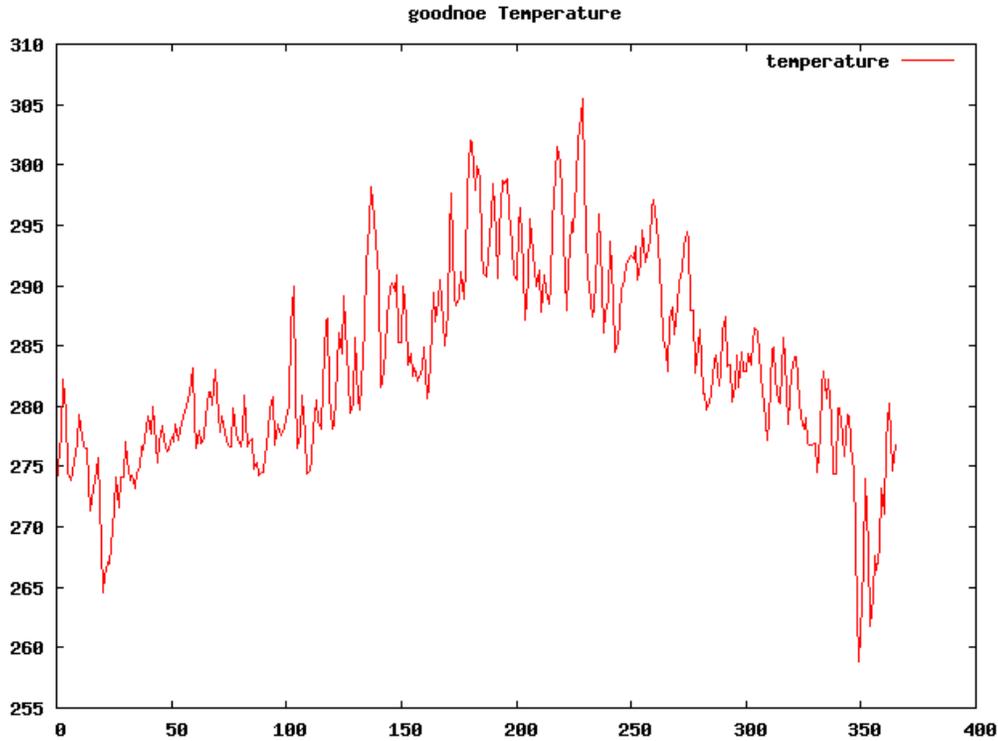


Figure 4. Daily mean 2m temperature at Goodnoe Hills by day of year, 2008. Vertical units are degrees K.

Validation of Model Configurations

In addition to the internal consistency checks described above, comparison was made to historical observations at 20 meteorological towers.

The table below lists the meteorological towers that were included for the purposes of model configuration selection. Towers were selected at sites that had meteorological observations above 20m and overlapped temporally for at least a year with other publically available stations. Towers were also selected for broad spatial coverage over the United States.

The towers below all had available data from one of three concurrent 12-month periods: 5/1991 – 4/1992 (period 1), 4/2005 - 3/2006 (period 2) and calendar year 2008 (period 3).

Table 1. Meteorological Towers Used for Configuration Selection

Name	State	Source	Max Meas. Height (m)
Goodnoe Hills	WA	Bonneville Power Adm.	59
Vansycle	OR	Bonneville Power Adm.	62
Chinook	OR	Bonneville Power Adm.	50
Cochran County	TX	West Texas A&M U.	70
Young County 3	TX	West Texas A&M U.	60
Crosby County	TX	West Texas A&M U.	58
Ft. Thompson	SD	S. Dakota State U.	70
Crow Lake	SD	S. Dakota State U.	70
Crandall	SD	S. Dakota State U.	70
Leola	SD	S. Dakota State U.	70
Summit	SD	S. Dakota State U.	70
Palmer	IL	Ill. Inst. for Rural Affairs	50
Hillcrest	IL	Ill. Inst. for Rural Affairs	50
Coffey	IL	Ill. Inst. for Rural Affairs	50
Abingdon	IL	Ill. Inst. for Rural Affairs	50
Blandford Yard	MA	U. Mass. - Amherst	60
Windland	CA	www.winddata.com	30
Gorgonio	CA	www.winddata.com	30
Flowind	CA	www.winddata.com	30
Calwind	CA	www.winddata.com	30

Figures were generated for each validation site that display the agreement between simulated and observed monthly mean wind speed, and between the simulated and observed annual diurnal cycle. An example of each is show below for the validation site at Crosby County, Texas. The full set is included in a set of two archive files that were delivered to NREL.

For most locations there was a small difference between the configurations relative to their distance to the observations. Quite frequently all configurations show values on the same side (above or below) of the observations. This is typical for physics-based simulations using WRF, so the configuration selection was largely influenced by the configurations consistently closest among the ensemble to the observations.

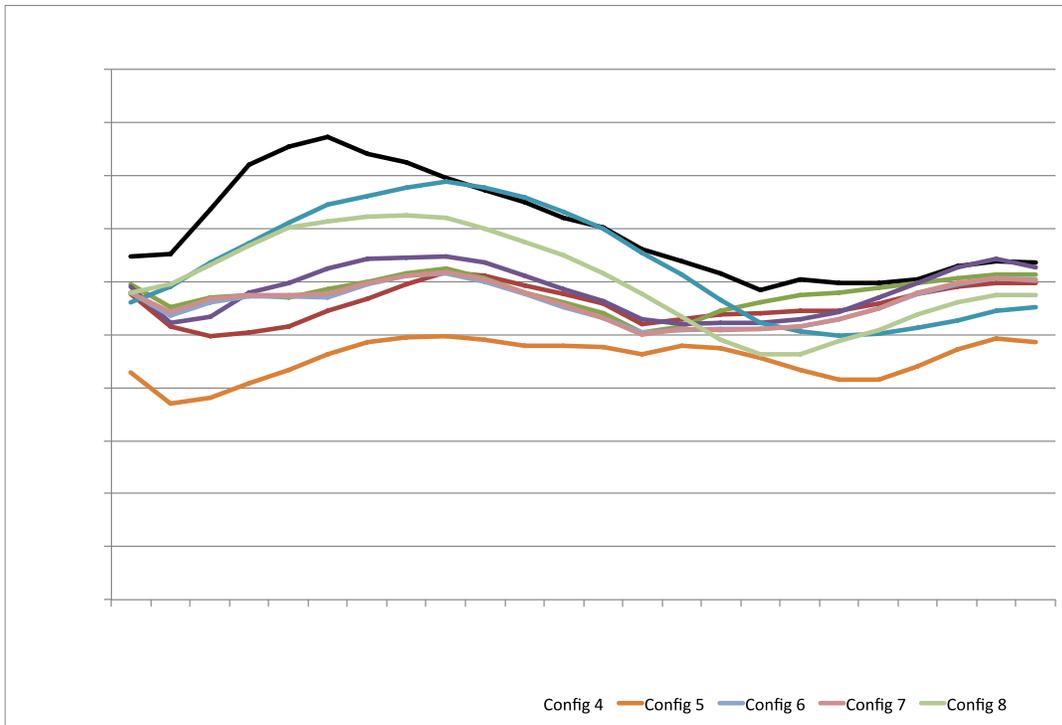


Figure 5. Diurnal cycle of wind speed at Crosby County, Texas

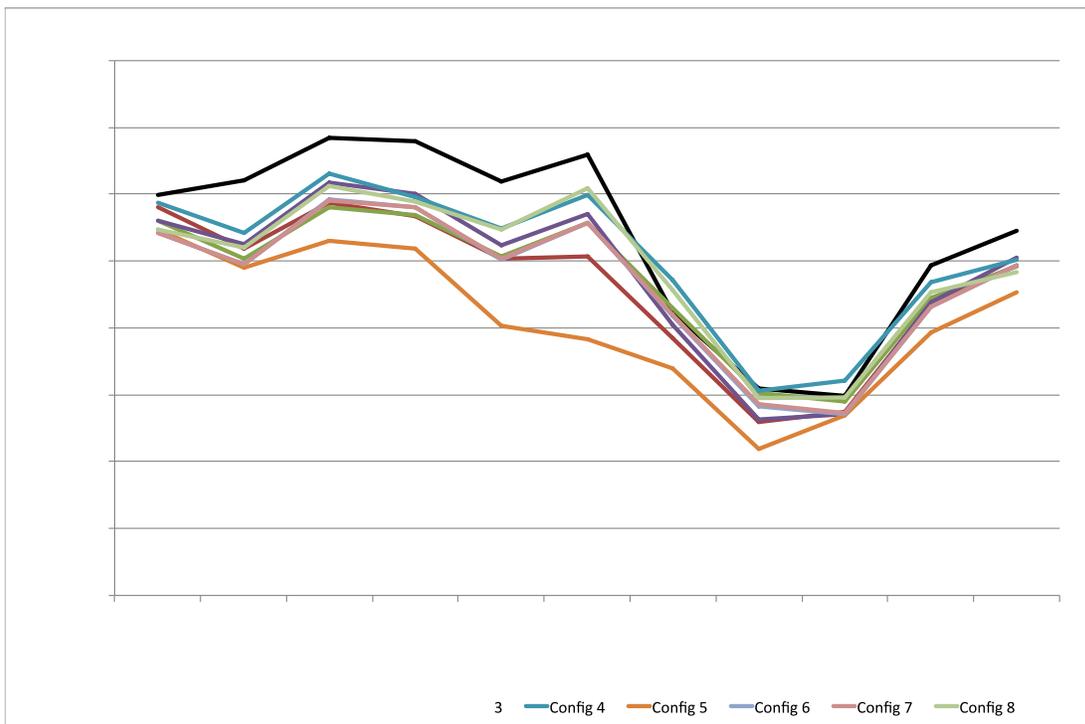


Figure 6. Monthly mean wind speed at Crosby County, Texas

Quantitative metrics were also considered. Table 2 shows correlations of monthly mean wind speeds at each site between observations and the output using the various configuration options. The average observed wind speed for each month was simply the mean of all available observations for the month, with a requirement of at least 70% data recovery for the month. Configuration 8 scores the highest on this metric.

Table 2. Correlation of Monthly Mean Wind Speed to Observed

	Configuration ID							
	1	2	3	4	5	6	7	8
Hillcrest	0.97	0.96	0.97	0.94	0.96	0.97	0.97	0.96
Vancycle	0.88	0.89	0.93	0.91	0.83	0.91	0.91	0.94
Summit	0.42	0.36	0.42	0.54	0.38	0.40	0.40	0.56
Crow Lake	0.73	0.61	0.67	0.78	0.72	0.64	0.64	0.77
Crandall	0.77	0.73	0.80	0.74	0.74	0.78	0.78	0.77
Fort Thomson	0.80	0.73	0.79	0.79	0.81	0.75	0.75	0.80
Leola	0.73	0.66	0.80	0.72	0.73	0.76	0.76	0.73
Cochran	0.96	0.90	0.96	0.95	0.89	0.95	0.95	0.92
Young County	0.99	0.97	0.99	0.95	0.97	0.99	0.99	0.96
Crosby County	0.94	0.96	0.98	0.97	0.86	0.97	0.97	0.98
Goodnoe Hills	0.59	0.70	0.62	0.63	0.52	0.66	0.66	0.71
Chinook	0.58	0.67	0.64	0.69	0.53	0.59	0.60	0.75
Abingdon	0.58	0.56	0.55	0.66	0.57	0.52	0.52	0.61
Palmer	0.71	0.71	0.74	0.67	0.71	0.73	0.73	0.71
Coffey	0.98	0.98	0.99	0.98	0.97	0.99	0.99	0.99
Blandford	0.93	0.92	0.93	0.94	0.96	0.94	0.93	0.94
Flowind	0.09	0.23	0.21	0.49	0.06	0.17	0.18	0.69
Calwind	0.58	0.71	0.68	0.90	0.56	0.60	0.60	0.96
Gorgonio	-0.34	-0.29	-0.20	0.33	-0.42	-0.19	-0.19	0.40
Mean	0.68	0.68	0.71	0.77	0.65	0.69	0.69	0.80

Similarly, Table 3 shows the correlation of diurnal mean wind speed values to observed. This is a measure of the goodness of the fit between the simulated and observed diurnal cycle. Again, configuration 8 scores the highest.

Table 3. Correlation of Diurnal Mean Wind Speed to Observed

	Configuration ID							
	1	2	3	4	5	6	7	8
Hillcrest	0.96	0.57	0.93	0.54	0.67	0.86	0.86	0.65
Vancycle	0.99	0.98	0.98	0.97	0.99	0.96	0.96	0.97
Summit	0.50	0.65	0.63	-0.24	-0.15	0.70	0.68	-0.04
Crow Lake	0.48	-0.60	0.21	0.96	0.89	-0.49	-0.46	0.82
Crandall	0.61	-0.25	0.61	0.97	0.97	-0.09	-0.08	0.92
Fort Thomson	0.67	0.52	0.57	0.62	0.30	0.49	0.49	0.87
Leola	0.25	-0.58	0.18	0.95	0.93	-0.35	-0.36	0.75
Cochran	0.88	0.83	0.86	0.96	0.96	0.76	0.76	0.99
Young County	0.92	0.75	0.91	0.85	0.80	0.73	0.70	0.78
Crosby County	0.02	0.24	0.54	0.88	0.19	0.50	0.52	0.94
Goodnoe Hills	0.83	0.76	0.85	0.97	0.85	0.77	0.77	0.90
Chinook	0.86	0.95	0.94	0.83	0.73	0.96	0.96	0.91
Abingdon	0.78	0.90	0.71	0.08	0.20	0.91	0.92	0.27
Palmer	0.86	0.55	0.82	0.48	0.67	0.83	0.84	0.64
Coffey	0.78	0.35	0.83	0.78	0.80	0.68	0.68	0.83
Blandford	0.80	0.65	0.74	0.31	0.32	0.65	0.65	0.49
Flowind	0.37	0.34	0.36	0.79	0.18	0.33	0.33	0.76
Calwind	0.43	0.39	0.45	0.82	0.27	0.38	0.39	0.81
Gorgonio	0.42	0.47	0.56	0.33	0.03	0.45	0.45	0.55
Mean	0.65	0.45	0.67	0.68	0.56	0.53	0.53	0.73

Selection of Final Configuration

It was Vaisala’s recommendation that the study proceed using configuration 8, as described above. It represented a good balance between the improved bias and diurnal cycle representation of configuration 4 (use of the Noah LSM) and the increase in hourly wind speed changes contributed by configuration 3 (the spectral nudging approach). The increase in vertical levels near the surface allowed more information to be extracted at the multiple levels of interest for integration studies, and the topographic wind adjustment reduced bias in complex terrain. Based on these arguments, the selection of configuration 8 was approved by the TRC. The namelist used by WRF for the selected configuration is included in Appendix B.

Model Simulations

The full model simulations were conducted two high performance computing (HPC) installations provided by NREL: Red Mesa, and Peregrine. As simulations completed, the WRF output was transferred to Vaisala for subsequent processing. Both the completion of the model simulations and the data transfer put considerable stress on the computing centers, and the seven years of simulations took approximately 13 months to complete.

The Wind Power Production Dataset

Because of the large size (greater than 200 TB) of the final meteorological dataset, data at a subset of locations within the domain was extracted for further analysis. Therefore, while the meteorological data set contains all the grid points, the wind power production dataset contains only the sites selected as described below. This data forms the basis of the wind power production dataset that will be used for future wind integration studies. A total of 126,693 sites were selected for this project, representing the locations of all known existing and planned wind energy plants, as well as a great number of speculative locations. At each of these, power production was estimated for each 5-minute period of the study duration. Note that in the discussion below, a ‘site’ corresponds to a single NWP model grid cell, and is collocated with a cell, so no horizontal interpolation was required.

Site Selection

The goal of the site selection process was to identify existing and probable locations for wind energy development, which may then be used to generate simulated wind plant output and forecasts. The process is motivated by the large data volumes associated with mesoscale model output; it would be prohibitively time-consuming to evaluate wind energy production at all grid locations within the model domain. Instead, for the WIND Toolkit dataset, 126,693 sites were selected as shown in Figure 7 and described below. This number represents a potential build-out of over one terawatt (TW) of wind energy capacity, dramatically more than is present in the highest wind penetration scenarios considered in integration studies to date, and thus should allow considerable flexibility in subsequent studies.



Figure 7. Map of selected sites

Sites were defined as individual 2km by 2km grid cells corresponding to the grid cells in the NWP dataset, and each site was assigned a developable capacity of up to sixteen megawatts (16 MW), based on land availability within the cell. This is generally consistent with existing and near-future wind turbine technology and plant design. For instance, approximately eight two-megawatt turbines might be placed in such an area, while larger turbines would require greater spacing, so that the areal density of nameplate capacity is relatively flat. We assumed roughly 4x12 turbine spacing, and some limitations on available land. Adjacent sites may be combined during subsequent analysis to create wind projects of any desired size. Each site was assigned a developable megawatt value (MW) ranging from 0 to 16 MW, based on the fraction of the grid cell that contains developable land.

Each potential site was assigned a turbine class. Turbine class was determined for onshore locations using the annual mean wind speed in the grid cell from a pre-existing Vaisala-created 90m dataset covering the continental U.S., in accordance with International Electrotechnical Commission standard (IEC) guidelines. For offshore locations, use of a class 1 offshore turbine was assumed.

Selection of Onshore Sites

The primary selection criteria for onshore locations were the wind resource and availability of land. The sites were determined within an Arc GIS framework, using a set of exclusion criteria developed in collaboration with the TRC. NREL provided a database of excluded land areas for use in this process. After preliminary analysis showed the exclusion of many existing wind plant locations, NREL investigated their land based exclusion criteria to minimize the exclusion of existing wind plants. The investigation revealed that some existing farms were excluded or partially excluded by the 3km buffer surrounding exclusion areas and the exclusion of slopes greater than 20%. These restrictions were removed from the exclusion layer given to 3TIER. For the purposes of the site selection, an exclusion of 100% means that no land within the area is available for wind direction, while an exclusion of 50% means that half of the land within an area is available. No analysis of which half was considered; rather the exclusion fraction simply factored into the fraction of nameplate capacity that could be installed at a site.

The exclusions that were for site selection fall into two criteria: Environmental and Land Use. The exclusions in each criterion are outlined below:

Environmental Criteria

- 100% exclusion
 - National Park service and Fish and Wildlife Service Managed lands (USGS Federal Lands shapefile, Dec 2005)
 - Federal Lands – park, wilderness, wilderness study area, national monument, national battlefield, recreation area, national conservation area, wildlife refuge, wildlife area, wild and scenic river or inventoried roadless area. (USGS Federal Lands shapefile, Dec 2005; Inventoried Roadless Areas, 2004; BLM Areas of Critical Environmental Concern (2008))

- Exclusion of state and private lands equivalent to above, where GIS data is available. (State/GAP land stewardship data management status 1, from Conservation Biology Institute Protected Lands database, 2004)
- 50% exclusion
 - Remaining USDA Forest Service (FS) lands (incl. National Grasslands) except ridgecrests (USGS Federal Lands shapefile, Dec 2005)
 - Remaining Dept. of Defense lands except ridgecrests (Military Lands boundary files, internal dataset (2007))
 - State forest land, where GIS data is available (State/GAP land stewardship data management status 2, from Conservation Biology Institute Protected Lands database, 2004)

Land-use Criteria

- 100% exclusion
 - Airfields, urban, wetland and water areas. (USGS North America Land Use Land Cover (LULC), version 2.0, 1993; ESRI airports and airfields (2006); U.S. Census Urbanized Areas (2000 and 2003))
- 50% exclusion
 - Non---Ridgecrest forest (Ridge---crest areas defined using a terrain definition script, overlaid with USGS LULC data screened for the forest categories.)

3TIER's pre-existing 90m Continental US (CONUS) wind resource data set was used during the site selection process in order to provide elevation data and annual mean wind speeds. Aggregated data from the 90m data set, along with the fractional usable area within each 2km by 2km grid cell, was used to estimate the total installable nameplate capacity for each cell.

Because existing transmission lines would place an undue constraint on the construction of such a large number of new wind plants, distance to existing transmission was not considered during site selection.

Site selection was broken down into four main steps. First, existing and proposed wind plants from DOE's OpenEI wind projects database (available at http://en.openei.org/wiki/Map_of_Wind_Farms/Data) were 'built up' by assigning the appropriate number of adjacent grid cells to match the nameplate value for the plant. Second, additional sites were added corresponding to all locations used in the Western Wind and Solar Integration Study (WWSIS) and the Eastern Wind Integration Study (EWITS) studies. Third, the highest rated cell in each U.S. county was added. This was to ensure that counties where all sites are rated below class III would have at least one site. Fourth, all remaining sites were added. This step involved randomly seeding wind plants throughout the country and then attempting to build wind plants of various sizes at these locations. The overall desire was to have widely spaced clusters of grid cells that would mimic the existing pattern of wind energy development. These sites were selected so that 60% were represented by class III turbines, 20% by class II turbines, and 10% by class I turbines.

Selection of Offshore Sites

The feasibility of offshore wind, like any generation plant, is a function of return, cost, and risk. Since forecasting future power prices and incentives are out of the scope of this work, the wind resource was the main constraint on the return side of the model. However, as with the onshore sites, a random seeder was employed to ensure the creation of separated clusters of included sites. We did not consider location-dependent permit processes. The presumption is that to make so many offshore sites feasible, substantive cooperation at the local level to reduce permitting restrictions will be necessary. Finally, at the request of the TRC, bathymetry was not considered, since DOE is investigating offshore floating platforms.

Power Production Profiles

The power production profiles were generated by running the simulated 100m hub height wind speeds through the ERGIS power curves shown in Figure 8.

For this project, the estimation of power from wind speed consisted of the following four steps, applied sequentially at each site:

1. Bias removal for wind speed. This step is listed for completeness because it was considered by the TRC. However, after comparison with RUC output, a robust systematic error was not found, and thus no bias adjustments were made.
2. Wind speed adjustment for wakes. A linear wake loss from 0% reduction in wind speed for a cell with a single turbine to a 5% reduction in wind speed for a cell with 8 turbines was assumed for each site. There was no dependence on wind direction. The equation for wind speed adjustment for wakes (C_w) is as follows:

$$C_w = \frac{N_t - 1}{8}$$

where N_t is the number of turbines in the grid cell. For selected sites, N_t varied from 1 to 8.

3. Application of power curves. ERGIS power curves for IEC classes 1, 2, and 3, plus offshore were used. A density correction (proportional to the cube root of the ratio of the modeled density to the turbine reference density) was applied to estimate an effective wind speed at the turbine reference density. A cut-out wind speed of 20 m/s was used for class 3, and 25 m/s for the other classes. Hysteresis cut-in values were set to be 5 m/s below the high wind cut-out values.
4. Statistical adjustment to power. Again this is included for completeness since it was considered by the TRC, However, after analysis of autocorrelation and spatial covariance between nearby sites, it was decided to make no adjustment.

Subsequent users of the data will need to include additional losses as required, and should consider implications for wake losses due to any aggregation effects.

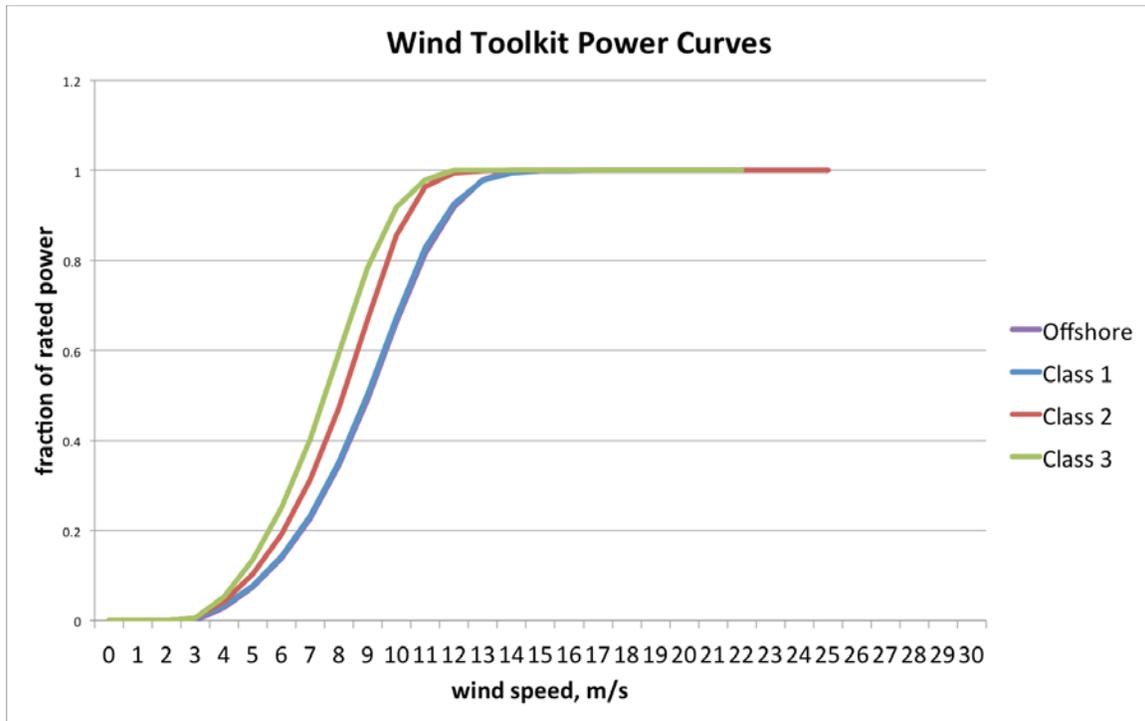


Figure 8. ERGIS power curves

Comparisons of modeled to observed spatial covariance and temporal autocorrelation were conducted using proprietary data from three projects in different wind climates. Figure 9 presents an example of the comparisons, for a project situated in an area of complex terrain in the Pacific Northwest. In order to compare gridded model output with wind speeds recorded at turbines within the plant, the turbines were aggregated into 2km x 2km boxes collocated with the model grid. The top panel shows correlations between observed and simulated power for all pairs of resulting boxes, which are separated by distances from 2 to 15km. The model has a small high bias, but mimics the trend quite well. The bottom panel shows the autocorrelation for the observed and simulated times series of power for each grid box. The bottom axis is the count of ten-minute periods. In this case, there is very little bias in the simulations. The simulations do show a small diurnal cycle (the slight hump at a lag of 144 periods) that is not present in the observations.

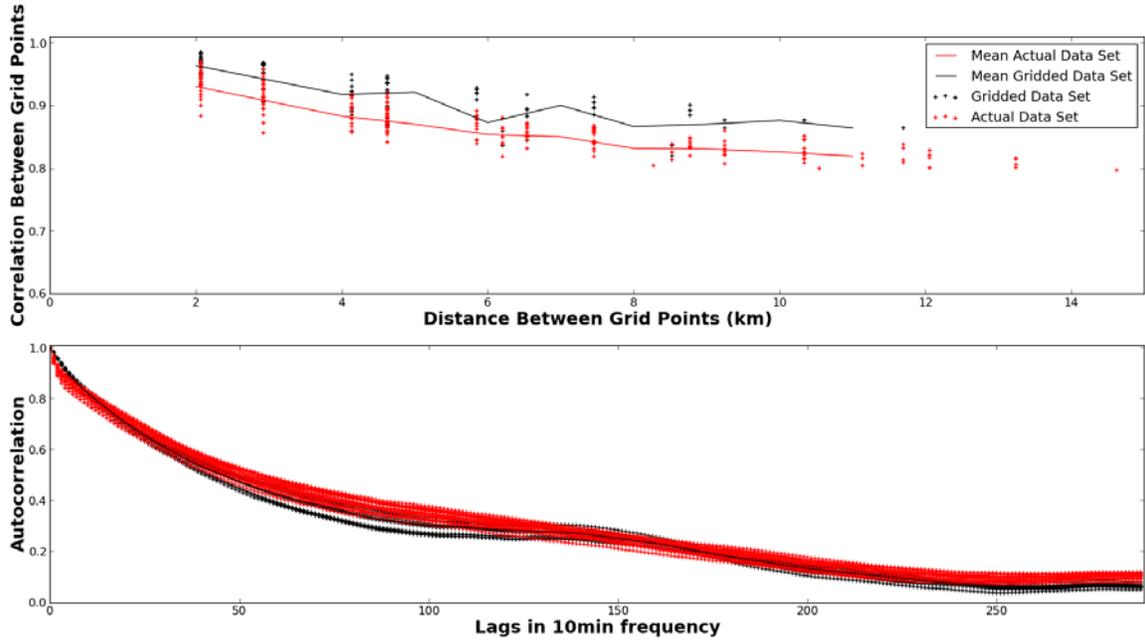


Figure 9. Spatial and temporal autocorrelation of the WIND Toolkit data compared to observations for a project in complex terrain

Simulation of Forecast Data

A key requirement for wind energy integration work is the availability of wind energy forecasts to go with the modeled estimates of production. The TRC requested that these forecasts have error characteristics that are similar to the current state of the art of operational wind forecasting, so that accurate estimates of the costs of forecast uncertainty can be computed.

Power forecasts at 1, 4, 6, and 24-hour lead times were produced to correspond to each hour of the wind power production dataset. A mesoscale NWP model-based approach was used to realistically capture the natural spatio-temporal correlations between the wind sites, i.e. to ensure that neighboring sites have forecast errors in the same direction and of similar magnitudes. Each power forecast contains a deterministic, best-estimate value, and P10 and P90 probability of exceedance values. The methods by which the forecasts were made are described below in sections 5.1 (24 hour-ahead), 5.2 (1, 4, and 6 hour-ahead), and 5.3 (probability of exceedance values).

Forecasts at all lead-times were validated for the appropriate error levels as described below by comparing to error levels from recent proprietary Vaisala operational forecasts provided at twelve representative wind plants across the United States. No publicly available data was available for comparison. The approximate locations of these plants are shown in Figure 10, and listed in Table 1.

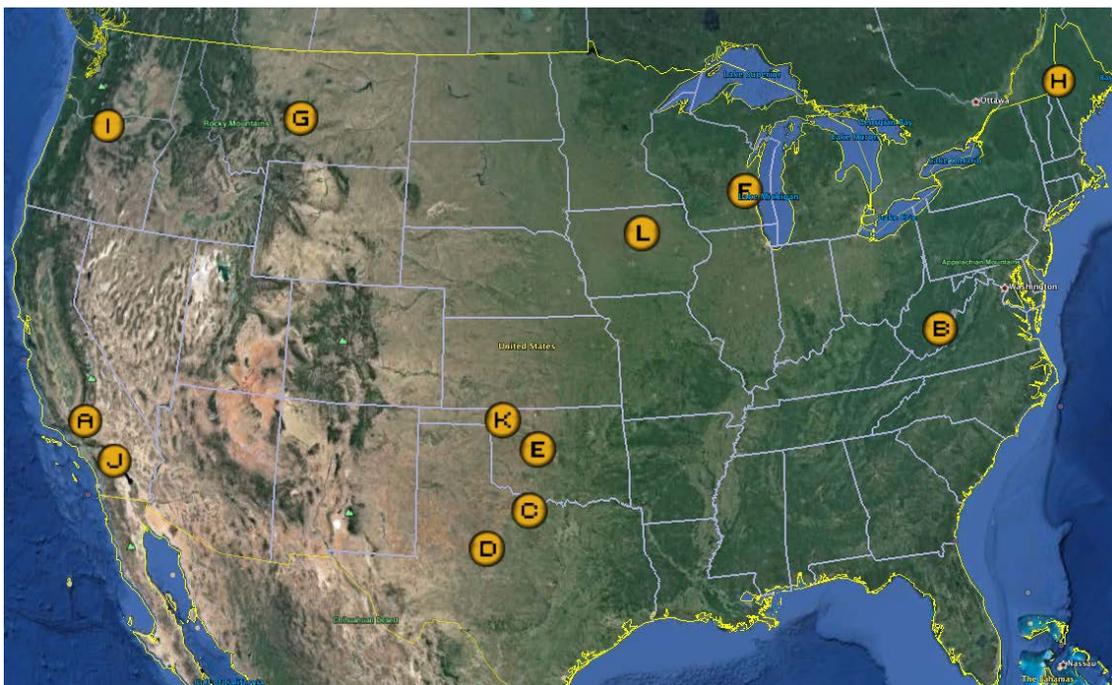


Figure 10. Forecast error validation locations

Table 4. Forecast Error Validation Locations

Plant Name	ISO Name	Simulated Nameplate Capacity (MW)
Project A	CAISO	1326
Project B	PJM	120
Project C	ERCOT	152
Project D	ERCOT	234
Project E	SPP	300
Project F	MISO	70
Project G	WECC-Montana	146
Project H	ISO-NE	132
Project I	BPA	230
Project J	CAISO	52
Project K	SPP	102
Project L	MISO	200

24-Hour Lead-Time Forecasts

To achieve reasonable 24-hour lead-time forecasts, the WRF model was again used, but at a coarser resolution and with different boundary conditions than what was used to create the meteorological data set. The motivation for this was to achieve results that were sufficiently different from the meteorological data set runs to eliminate the need to add additional noise to the forecasts at a later step to reproduce typical operational forecast errors.

At the same time, it was important to produce relatively bias-free forecasts (as expected in operations), so a very similar NWP model setup was applied. Aside from the reduced resolution, the WRF configuration was held as similar as possible to that of the meteorological simulations. All physics options were the same, and the same 54km, 18km, and 6km nests were used. The model was initialized once per day at 0000 UTC, with no spin-up period. Initial and lateral boundary conditions were provided by the 1-degree 2nd-generation NOAA Global Ensemble Forecast System Reforecast (GEFS) control simulations. Each forecast simulation was run out 48 hours to ensure that the day-ahead (i.e. next day) period was covered. The output interval was 60 minutes.

Due to computational limits, the forecast NWP simulations were initialized only once per day (at 00Z). Forecast errors are relatively flat during the 24-48 hour period, but there is a discontinuity in the forecast properties as a result.

Forecasts of simulated actuals are often overly skillful, as raw model forecasts are frequently more similar to the simulated actuals than to the real atmosphere. In this case, the use of GEFS boundary conditions and reduced resolutions resulted in forecasts that had sufficient error to proceed. To adjust this error further, as described below, we used a combination of time series smoothing and “truth” blending, which is simply a weighted average of the raw forecast and the simulated actual value at a given time. This mimicked operational model output statistics (MOS)

by reducing bias and improving skill, but not by too much. At the same time, it was computationally efficient relative to running full MOS at all of the sites.

We used an iterative approach to the statistics, and adjusted smoothing and blending amounts upward until forecast time series and error histograms appeared reasonable and bulk error metrics were similar to state-of-the-art day-ahead forecasts.

Table 5 describes the subjective characteristics of forecast error tuning throughout the iterative process of choosing the right weighting of smoothed raw NWP forecast and smoothed “truthful” simulated actuals. With the underlying goal of mimicking appropriate typical operational forecasts, Vaisala tested various configurations of raw NWP forecast smoothing windows, simulated actual smoothing windows, and blending weights. Based on industry experience with operational forecasts, Vaisala made subjective judgments based on visual inspection to attain appropriate error statistics. Some comments based on these inspections are included below, however only the final selected parameters were used in the creation of the dataset.

After the experimental iterations shown in Table 5, the 24 hour-ahead noise and error statistics were deemed appropriate when the forecast was a created from an 80%/20% mixture of a 13-hour centered moving average of the raw NWP forecasts and a 13-hour centered moving average of the simulated actuals, respectively. In this context, the simulated actuals were treated as “truth”. The mixture of simulated actuals into the raw NWP forecast is what is meant by “truth” blending.

Figure 11 shows an example forecast made using the selected configuration.

Table 5. Iterations of Day-Ahead Forecast Smoothing

Forecast Smoothing	Observation Smoothing	Observation Weight	Subjective Character	Subjective Bias	Subjective Skill
3 h	N/A	N/A	Too Noisy	Too Large	Too Low
3 h	1 h	10%	Too Noisy	A Bit Large	Too Low
3 h	3 h	10%	Noisy, But Less So	A Bit Large	Too Low
7 h	7 h	10%	A Little Noisy	A Bit Large	A Bit Low
13 h	7 h	10%	About Right	A Bit Large	About Right
7 h	7 h	30%	A Little Noisy	Much Lower	Too High
13 h	13 h	20%	About Right	Acceptable	About Right

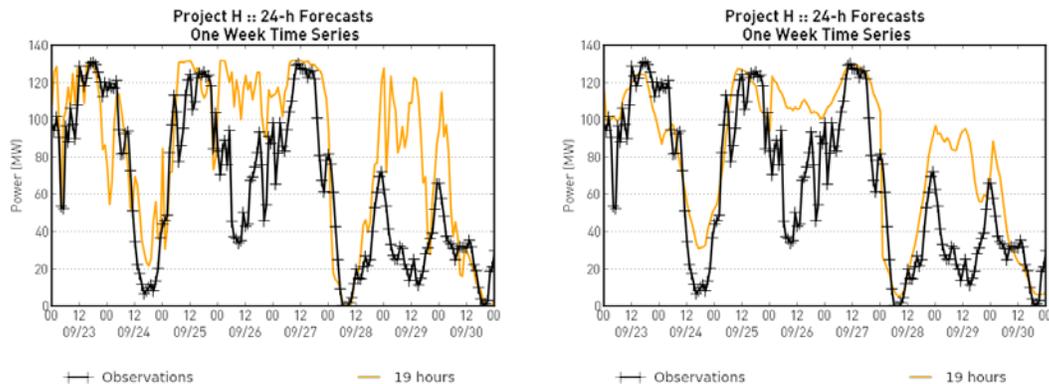


Figure 11. Example raw WRF forecasts (left panel) compared to centered 13-hr smoothing of raw WRF forecasts blended with 20% weight on the 13-hr smoothed actuals (right panel)

Short Lead-Time Forecasts

Statistical forecasts were generated for each site for the short lead-time (1, 4, and 6 hour) forecasts, based on a dispersive persistence method. Vaisala’s operational short lead-time forecasts outperform a simple persistence forecast through sophisticated statistical methods not employable for this dataset, therefore the technique described below was used with the simulated actuals to approximate the error statistics of operational short lead-time forecasts. The selected technique used shortened persistence lead times to increase the skill of the forecasts to the desired level, while maintaining temporal dispersion of errors, by using the two parameters of persistence interval and lead-time, as described below.

The most basic persistence forecast uses the current value as a prediction for some future value. The length of the time between the observation and the predicted time is referred to as the ‘lead time’, and the length of the lead time has a significant impact on the skill of the forecast. In the present convention, period-ending time was used. The enhanced skill of operational forecasts can be approximated using lead times shorter than would be available in a real time system, but using a constant lead time does not properly characterize the distribution of events in time – the errors are not properly dispersed.

For this dataset, the lead-time parameter was randomly sampled from a truncated Gaussian distributed population of 8,760 5-minute interval lead times, with a range of 0-80 minutes and a mean of 40 minutes. In other words, the population contains lead times of 0, 5, 10, 15 minutes, and so on up to 80 minutes. The most common lead time is 40 minutes. The effect of the mean 40-minute lead time is to improve average forecast skill.

The second parameter was the persistence interval. The persistence interval is the length of time over which the simulated actual (observed) data is averaged, in effect resampling the observed data from being a 5-minute average to being a 10-minute average. Three examples along with figures of the algorithm are given below:

Example #1 (shown in Figure 12). If the present time is T_0 and we are producing a forecast valid for the interval from T_{+60} to T_{+120} , and a 40-minute lead time is sampled, then the algorithm averages T_{+15} and T_{+20} and uses this value as the forecast.

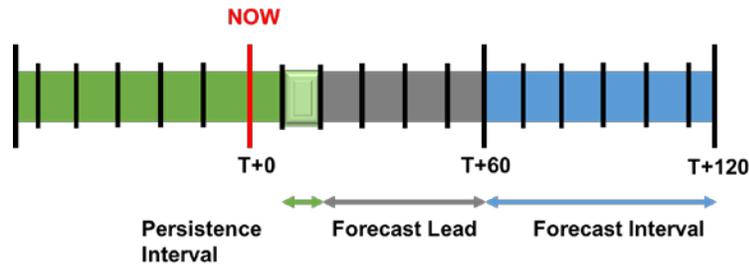


Figure 12. Graphical example of forecast algorithm: 40min lead time

Example #2 (shown in Figure 13). If the present time is T_0 and we are producing a forecast valid for the interval from T_{+60} to T_{+120} , and an 80-minute lead time is sampled, then the algorithm averages T_{-25} and T_{-20} and uses this value as the forecast.

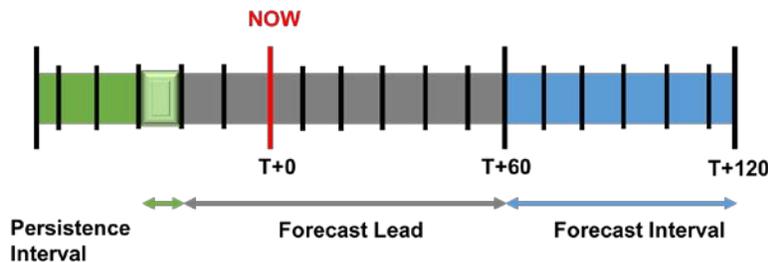


Figure 13. Graphical example of forecast algorithm: 80min lead time

Example #3 (shown in Figure 14). If the present time is T_0 and we are producing a forecast valid for the interval from T_{+60} to T_{+120} , and a 0-minute lead time is sampled, then the algorithm averages T_{+55} and T_{+60} and uses this value as the forecast.

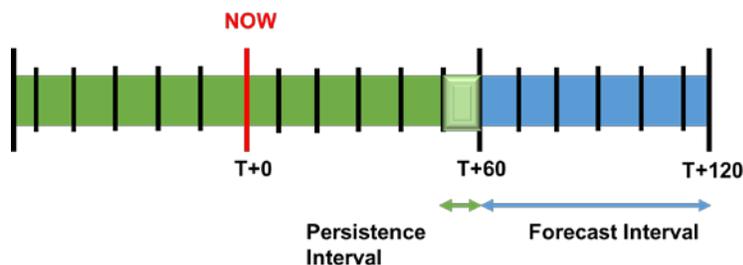


Figure 14. Graphical example of forecast algorithm: 0min lead time

To select the algorithm parameters, an iterative procedure similar to that of the 24-hour lead-time forecasts was followed, starting with a 1-hour forecast comprised of a 5-min persistence at 40-min lead. This was adjusted until the improvement over persistence was similar to current operational forecasts, as shown in Table 6.

In practice, the long-term statistics of short lead-time forecasts show a greater skill than a simple persistence forecast, however at any given time, it is possible for such a forecast to be more

skillful than normal, or less skillful than normal. A range of lead-times ensured a varying measure of success in the 1-hour forecasts.

Table 6. Iterations of Short-Term Forecast Parameters

Persistence Interval	Mean Lead Time	Lead Time Range	Auto-Correlation	Subjective Skill
5 min	40 min	N/A	N/A	A Bit High
5 min	45 min	0-90 min	0.9	A Bit Low
5 min	40 min	0-80 min	0.9	Barely High
10 min	40 min	0-80 min	0.9	Just Right

After parameters for the 1-hour forecasts were determined, the 4-hour lead-time forecasts were computed using an 80% to 20% blending of the 24-hour and 1-hour forecasts, and the 6-hour lead-time forecasts were computed using a 90% to 10% blending of the 24-hour and 1-hour forecasts. The equations used to calculate 4 (F_4) and 6 (F_6) hour-ahead forecasts are as follows:

$$F_4 = 0.8(F_{24}) + 0.2(F_1)$$

$$F_6 = 0.9(F_{24}) + 0.1(F_1)$$

Probabilistic Forecasts

A conditional, non-parametric dressing approach was used to create the probabilistic forecast intervals, using a bootstrap method to correct the quantile estimates in which we were interested. Given a power forecast, we had an empirical distribution of errors from recent and similar power forecasts. These were sampled over the previous 12-day period, using weights that were computed as a distance from the current forecast, in turn using bi-square weighting. Next, the forecasts were resampled with replacement to find the quantiles, and the probabilistic forecasts were averaged over 100 repetitions. This yielded approximate calibration, and allowed for dynamic adjustment to weather regime changes and seasonal forecast skill.

The Data Extraction Tool

Access to the forecasting and power production datasets is provided through a data extraction application programmable interface (API), that exposes programmatic access to the data for the Data Extraction GUI software and other clients. The software developed by Vaisala and delivered to NREL consists of a representational state transfer (REST) based API capable of generating time series for multiple attributes for a set of sites when provided with:

- A list of site IDs.
- A time range.
- A list of attributes.
- A file path for output.

- A parameter optionally specifying that the time series for each attribute will be aggregated across the specified sites. Aggregation techniques are explained below.

The software was developed using Jruby and targeted toward a Rails environment.

When aggregating, the aggregation technique varies by attribute:

- Power is summed.
- Wind direction is vector averaged.
- Capacity factor is weighted by assumed installation (i.e., number of turbines) per site.
- Simple averaging is performed for wind speed, density, and temperature

The attributes that are available via the API include the following (with units in parentheses):

- t - Time (UTC)
- Pw – wind power (MW)
- WS – wind speed at hub height (100m) (m/s)
- WD – wind direction at hub height (100m) (degrees)
- F1, F4, F6, F24 – wind power forecasts (MW) valid at time t for one, four, six, and twenty-four hours ahead. P10, best estimate, and P90 values will be provided at each hour for each forecast time horizon.
- T – temperature at 2m (K)
- P – surface pressure (Pa)
- ρ – density at hub height (kg/m^3)

All attributes except the wind power forecasts are available on a 5-minute interval. The wind power forecasts are available on a 1-hour interval.

Summary

The goal of this project was to produce a high-resolution national dataset that reproduces the characteristic variance, spatial and temporal correlations, and mean capacity factors for typical existing and proposed wind plants. Along with the coincident retrospective forecast dataset, we feel it represents suitable state-of-the-art resource for future wind integration studies.

Appendix A: WPS Namelist

The following lines show the namelist used by the WRF preprocessing system in generating the domain and other input files for the creation of the meteorological data set.

```
&share
  wrf_core = 'ARW',
  max_dom = 4,
  start_date = '2009-10-30_00:00:00','2009-10-30_00:00:00','2009-10-30_00:00:00','2009-10-
30_00:00:00',
  end_date = '2009-12-02_00:00:00','2009-12-02_00:00:00','2009-12-02_00:00:00','2009-12-
02_00:00:00',
  interval_seconds = 21600
  io_form_geogrid = 2,
  opt_output_from_geogrid_path = './',
  debug_level = 0
/

&geogrid
  parent_id      = 1,1,2,3,
  parent_grid_ratio = 1,3,3,3,
  i_parent_start = 1,16,18,16,
  j_parent_start = 1,18,16,17,
  s_we          = 1,1,1,1,
  e_we         = 157,379,1033,3007,
  s_sn         = 1,1,1,1,
  e_sn         = 109,223,577,1633,
  geog_data_res = '5m','2m','30s','30s',
  dx = 54000,
  dy = 54000,
  map_proj = 'lambert',
  ref_lat = 38.47240422490422,
  ref_lon = -96.0,
  truelat1 = 30.0,
  truelat2 = 60.0,
  stand_lon = -96.0,
  geog_data_path = '/home/odin1/data/wps_data'
  opt_geogrid_tbl_path = './'
/

&ungrib
  out_format = 'WPS',
  prefix = 'FILE',
/

&metgrid
  fg_name = 'FILE'
  io_form_metgrid = 2,
/

&mod_levs
  press_pa = 201300 , 200100 , 100000 ,
            95000 , 90000 ,
            85000 , 80000 ,
            75000 , 70000 ,
            65000 , 60000 ,
            55000 , 50000 ,
            45000 , 40000 ,
            35000 , 30000 ,
            25000 , 20000 ,
            15000 , 10000 ,
            5000 1000
/
```

Appendix B: WRF Namelist

The following lines show the namelist used by the WRF model in generating the meteorological data set.

```
&time_control
run_days              = 0,
run_hours             = 792,
run_minutes           = 0,
run_seconds           = 0,
start_year            = 2009, 2009, 2009, 2009,
start_month           = 10, 10, 10, 10,
start_day             = 30, 30, 30, 30,
start_hour            = 00, 00, 00, 00,
start_minute          = 00, 00, 00, 00, 00, 00,
start_second          = 00, 00, 00, 00, 00, 00,
end_year              = 2009, 2009, 2009, 2009,
end_month             = 12, 12, 12, 12,
end_day               = 02, 02, 02, 02,
end_hour              = 00, 00, 00, 00,
end_minute            = 00, 00, 00, 00, 00, 00,
end_second            = 00, 00, 00, 00, 00, 00,
interval_seconds      = 21600,
input_from_file       = .true., .true., .true., .true., .true.,
fine_input_stream     = 0, 2, 2, 2, 2, 2,
io_form_auxinput2     = 2,
history_interval      = 51840, 51840, 51840, 51840, 51840, 51840,
frames_per_outfile    = 1, 1, 1, 1, 1, 1,
restart               = .true.,
restart_interval      = 2880,
nocolons              = .true.
io_form_history       = 11
io_form_restart       = 102
io_form_input         = 11
io_form_boundary      = 2
io_form_auxinput4     = 2
auxinput4_inname      = "wrfinputp_d<domain>"
auxinput4_interval    = 360,360,360,360
auxinput4_end_h       = 816, 816, 816, 816,
debug_level           = 0
auxhist2_interval     = 60, 60, 60, 5, 60, 60,
auxhist3_interval     = 60, 60, 60, 5, 60, 60,
io_form_auxhist2      = 11
io_form_auxhist3      = 11
frames_per_auxhist2   = 1, 1, 1, 1, 1, 1,
frames_per_auxhist3   = 1, 1, 1, 1, 1, 1,
frames_per_auxinput9  = 1, 1, 1, 1, 1, 1,
frames_per_auxinput10 = 1, 1, 1, 1, 1, 1,
iofields_filename     = "wind-reduce.txt","wind-reduce.txt","wind-
reduce.txt","wind-reduce.txt","wind-reduce.txt","wind-reduce.txt",
ignore_iofields_warning = .false.
/

&domains
time_step             = 300,
time_step_fract_num   = 0,
time_step_fract_den   = 1,
max_dom               = 4,
s_we                  = 1, 1, 1, 1, 1, 1,
e_we                  = 157,379,1033,3007,
s_sn                  = 1, 1, 1, 1, 1, 1,
e_sn                  = 109,223,577,1633,
s_vert                = 1, 1, 1, 1, 1, 1,
e_vert = 41,41,41,41,41,41,
num_metgrid_soil_levels = 4,
num_metgrid_levels   = 38,
eta_levels = 1.0000, 0.9958, 0.9916, 0.9874, 0.9832,0.9790, 0.9749, 0.9707, 0.9661,
0.9609,0.9549, 0.9480, 0.9398, 0.9303, 0.9189,0.9054, 0.8894, 0.8704, 0.8481, 0.8221,0.7922,
0.7583, 0.7205, 0.6791, 0.6346,0.5877, 0.5393, 0.4900, 0.4407, 0.3922,0.3450, 0.2996, 0.2564,
0.2156, 0.1773,0.1417, 0.1086, 0.0755, 0.0475, 0.0224,0.000
p_top_requested       = 5000,
dx                    = 54000,18000,6000,2000,
dy                    = 54000,18000,6000,2000,
grid_id               = 1, 2, 3, 4, 5, 6,
parent_id              = 1, 1, 2, 3, 4, 5,
i_parent_start         = 1,16,18,16,
```

```

j_parent_start = 1,18,16,17,
parent_grid_ratio = 1, 3, 3, 3, 3, 3,
parent_time_step_ratio = 1, 3, 3, 3, 3, 3,
feedback = 0,
smooth_option = 2
use_adaptive_time_step = .false.
step_to_output_time = .false.
target_cfl = -999
max_step_increase_pct = -999
starting_time_step = -999
max_time_step = -999
min_time_step = -999
/

&physics
topo_wind = 1,1,1,1
sst_update = 1,
sst_skin = 1,
num_land_cat = 33,
mp_physics = 5, 5, 5, 5,
ra_lw_physics = 1, 1, 1, 1,
ra_sw_physics = 1, 1, 1, 1,
radt = 15, 15, 15, 15, 15, 15,
sf_sfclay_physics = 1, 1, 1, 1,
sf_surface_physics = 2, 2, 2, 2,
bl_pbl_physics = 1, 1, 1, 1,
bldt = 0, 0, 0, 0, 0, 0,
cu_physics = 1, 1, 0, 0, 0, 0,
cudt = 5, 5, 5, 0, 0, 0,
cam_abs_freq_s = 21600,
levsiz = 59,
paerlev = 29,
cam_abs_dim1 = 4,
cam_abs_dim2 = 41,
isfflx = 1,
ifsnow = 0,
icloud = 1,
surface_input_source = 1,
num_soil_layers = 4,
sf_urban_physics = 0,0,0,0,0,0,0,0
mp_zero_out = 0,
maxiens = 1,
maxens = 3,
maxens2 = 3,
maxens3 = 16,
ensdim = 144,
slope_rad = 0,
topo_shading = 0,
/

&fdda
grid_fdda = 2, 2, 2, 2, 0,
gfdda_inname = "wrfddda_d<domain>",
gfdda_interval_m = 360, 360, 360, 360, 0,
gfdda_end_h = 816, 816, 816, 816, 0,
io_form_gfdda = 2,
fgdt = 0, 0, 0, 0,
if_no_pbl_nudging_uv = 1, 1, 1, 1, 0, 0,
if_no_pbl_nudging_t = 1, 1, 1, 1, 0,
if_no_pbl_nudging_q = 1, 1, 1, 1, 0,
if_zfac_uv = 0, 0, 0, 0,
k_zfac_uv = 10, 10, 10,
if_zfac_t = 1, 1, 1, 1, 0,
k_zfac_t = 10, 10, 10,
if_zfac_q = 1, 1, 1, 1, 0,
k_zfac_q = 10, 10, 10,
guv = 0.0003, 0.0001, 0.0001, 0.00005,
gt = 0.0003, 0.0001, 0.0001, 0.00005,
gq = 0.000003, 0.000001, 0.000001, 0.0000005,
xwavenum = 21, 17, 15, 15, 0,
ywavenum = 15, 10, 9, 8, 0,
if_ramping = 0,
dtRamp_min = 0.0,
/

&dynamics
use_baseparam_fr_nml = .true.
w_damping = 1,
diff_opt = 1,
km_opt = 4,

```

```

diff_6th_opt           = 0,
diff_6th_factor       = 0.12,
base_temp             = 290.,
damp_opt              = 0,
zdamp                 = 5000., 5000., 5000.,5000.,
dampcoef              = 0.01, 0.01, 0.01, 0.01
khdif                 = 0, 0, 0, 0,
kvdif                 = 0, 0, 0, 0,
non_hydrostatic       = .true., .true., .true.,.true.
/

&bdy_control
spec_bdy_width        = 5,
spec_zone              = 1,
relax_zone             = 4,
specified              = .true., .false.,.false.,.false.,.false., .false.,
nested                 = .false., .true., .true.,.true., .true., .true.,
/

&grib2
/

&namelist_quilt
nio_tasks_per_group = 8,
nio_groups = 5,
/

```