



Arnold Schwarzenegger
Governor

CALIFORNIA REGIONAL WIND ENERGY FORECASTING SYSTEM DEVELOPMENT VOLUME 2: WIND ENERGY FORECASTING SYSTEM DEVELOPMENT AND TESTING

Prepared For:

California Energy Commission
Public Interest Energy Research Program

Prepared By:

Electric Power Research Institute, Inc.



PIER FINAL PROJECT REPORT

September 2006
CEC-500-2006-089



Prepared By:

Electric Power Research Institute, Inc.
C. McGowin and P. Hirsch
Palo Alto, CA, 94303
Contract No. 500-02-014
Work Authorization WA-111

Prepared For:

Public Interest Energy Research (PIER) Program
California Energy Commission

Michael Kane
Dora Yen-Nakafuji
Contract Manager

Elaine Sison-Lebrilla
Program Area Team Lead

Elaine Sison-Lebrilla
Manager
Energy Generation Research Office

Martha Krebs, Ph. D.
Deputy Director
Energy Research and Development Division

B.B. Blevins
Executive Director

DISCLAIMER

This report was prepared as the result of work sponsored by the California Energy Commission. It does not necessarily represent the views of the Energy Commission, its employees, or the State of California. The Energy Commission, the State of California, its employees, contractors and subcontractors make no warrant, express or implied, and assume no legal liability for the information in this report; nor does any party represent that the uses of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the California Energy Commission nor has the California Energy Commission passed upon the accuracy or adequacy of the information in this report.

CITATIONS

This report was prepared by

AWS Truwind, LLC
Albany Avenue
Albany, NY 01000

Principal Investigator
J. Zack

Lawrence Livermore National Laboratory
P.O. Box 808
Livermore, CA 94455

Principal Investigator
S. Chin

This report describes research sponsored by the Electric Power Research Institute (EPRI) and the California Energy Commission.

The report is a corporate document that should be cited in the literature in the following manner:

California Regional Wind Energy Forecasting System Development, Volume 2: Wind Energy Forecasting System Development and Testing, Electric Power Research Institute. Palo Alto, CA, and California Energy Commission, Sacramento, CA: 2006. 1013263

PRODUCT DESCRIPTION

The rated capacity of wind generation in California is expected to grow rapidly in the future beyond the approximately 2100 MW in place at the end of 2005. The main drivers are the state's 20 percent Renewables Portfolio Standard requirement in 2010 and the low cost of wind energy relative to other renewable energy sources.

As wind is an intermittent generation resource and weather changes can cause large and rapid changes in output, system operators will need accurate and robust wind energy forecasting systems in the future. In response to this need, the California Energy Commission (Energy Commission) and EPRI initiated a new project in 2003 to develop and test short- and intermediate-term (or next-hour and next-day) forecast algorithms that builds on the results of previous project completed in 2002.

Volume 2 of the final report describes the detailed results of the short- and intermediate-term forecast algorithm research and high-resolution numerical modeling of wind flow over complex terrain. Volumes 1, 3, and 4 present the executive summary of the research results and the detailed results of the wind tunnel modeling of wind flow over complex terrain and development of the California Wind Generation Research Dataset (CARD).

Results and Findings

The short-term forecast research developed and tested an artificial neural network (ANN) algorithm to generate five-minute/three-hour regional wind energy generation forecasts for the four largest California wind resource areas, updated every five minutes. The mean absolute errors of the ANN forecasts were lower than those of a simple persistence forecast algorithm after the first 15 to 20 minutes of the three-hour forecast time horizon.

The intermediate-term forecast research initially screened several forecast algorithm improvements to identify those that would yield the greatest reduction of mean absolute forecast error vs. the previous forecasting evaluation at the Altamont and San Geronio projects, completed in 2002. Relative to the previous California project, the mean absolute error of the wind energy forecasts decreased from 14.1 percent to 11.9 percent of rated capacity at Altamont and from 16.6 percent to 13.0 percent at San Geronio. The improved forecast algorithm was then tested at five wind projects in California.

Numerical modeling of wind flow over the complex terrain at Altamont Pass showed significant variations of wind speeds and directions both between individual wind turbines and between the

wind turbines and nearby met towers. In addition, atmospheric stability appears to significantly affect the wind plant power curve, for example the relationship between the power output of a group of wind turbines and the wind speed measured at the met tower.

Challenges and Objectives

Electricity systems with significant intermittent wind capacity create a challenge to the system operator. Rapid changes of wind generation relative to load require rapid dispatching of generation and transmission resources to balance generation vs. load, regulate voltage and frequency, and maintain system performance within limits established by Control Performance Standards 1 and 2 (CPS1 and CPS2). This is especially true during periods when wind generation is fluctuating rapidly relative to system load, for example during passage of thunderstorms and weather fronts. Wind energy forecasts can help the system operator anticipate rapid changes of wind energy generation vs. load and make informed decisions. The objectives are to develop and demonstrate the capabilities of wind energy forecasting technology for same-day and longer-term forecasts.

Applications, Values, and Use

The improved wind energy forecasting algorithms developed in this project have already been incorporated into regional wind energy forecasting systems in California that generate daily forecasts for several wind projects in the state and for a large utility company. The results also provide the basis for collaborating with electricity grid operators to customize wind energy forecast system content, format, and method of delivery to meet the needs of the system operator, while continuing to develop and test further algorithm improvements.

EPRI Perspective

Volume 2 of the final report from the California Regional Wind Energy Forecasting Development project describes the research and results on short- and intermediate-term wind energy forecasting system development and testing and numerical modeling of wind flow over complex terrain. The results presented in this report and the companion volumes represent significant advances in both short-term and intermediate-term wind energy forecasting technology. It is anticipated that further improvements in forecast accuracy are possible by such measures as applying the full two-stage short-term forecast algorithm developed in the project, optimizing the use of real-time wind speed and direction data from upwind meteorological towers, and ensemble forecasting. These improvements can be developed and tested in parallel with a field demo project to develop a wind forecast display for utility and system operators.

Approach

Researchers developed and tested improved wind energy forecasting technology for both the next-hour and next-day time frames and for both wind resource areas and individual wind plants; conducted wind tunnel and numerical modeling of wind flow over complex terrain; and developed a California Wind Generation Research Dataset (CARD) to provide a resource for future study of wind energy generation and forecasting in California. In each case, real-time and historical wind resource and energy generation data were collected and used to train and then test the forecast system performance. The forecast performance metric was the mean and mean absolute error of the forecasts vs. the observed data for each case.

ABSTRACT

The rated capacity of wind generation in California is expected to grow rapidly in the future beyond the approximately 2100 MW in place at the end of 2005. The main drivers are the state's 20 percent renewable portfolio standard requirement in 2010 and the low cost of wind energy relative to other renewable energy sources.

As wind is an intermittent generation resource and weather changes can cause large and rapid changes in output, system operators will need accurate and robust wind energy forecasting systems in the future. In response to this need, the California Energy Commission (Energy Commission) and EPRI initiated a new project in 2003 to conduct the research on short- and intermediate-term (or next-hour and next-day) forecast algorithms and builds on the results of previous project completed in 2002.

Volume 2 of the final report describes the detailed results of the short- and intermediate-term forecast algorithm research and high-resolution numerical modeling of wind flow over complex terrain. Volumes 1, 3, and 4 present the executive summary of the research results and the detailed results of the wind tunnel modeling of wind flow over complex terrain and development of the California Wind Generation Research Dataset (CARD).

The short-term forecast algorithm used an artificial neural network (ANN) algorithm trained using five-minute time-series data for wind energy deliveries to the grid in each of the five wind resource regions provided by the CA ISO. Testing showed that the ANN forecast algorithm reduces forecast error vs. persistence. Further testing of the algorithm is needed to assess the effect of adding real-time wind speed and direction data and rapid-update weather forecast data to the inputs to the ANN forecast algorithm on forecast performance.

Development of the intermediate-term forecast algorithm assessed the effects of several algorithm enhancements on forecast performance relative to the results of the first California forecasting project completed in 2002. Of the six enhancements tested, using improved water surface temperature data, segmented wind plant power curves, more sophisticated model operating statistics (MOS), and ensemble forecasting gave the greatest improvement.

High-resolution modeling of wind flow over the complex terrain at Altamont Pass evaluated the variation of wind speed and power generation by individual wind turbines under various conditions. The results confirm that atmospheric stability affects the variability of wind speed and therefore the plant-scale power curve used in the next-day forecast algorithms. Hence, in addition to wind speed and direction, the plant scale power curve should also be a function of stability.

Keywords

Wind, Power generation, Wind energy forecasting, Meso-scale weather models, Artificial neural networks, Screening multiple linear regression

ACKNOWLEDGEMENTS

Several organizations and individuals contributed to and provided data and information for the California Regional Wind Energy Forecasting System Development Project addressed in this and other volumes of the final report. They include Dr. Dora Yen-Nakafuji and Michael Kane of the California Energy Commission; David Hawkins of the California Independent System Operator (CA ISO); Dr. Marc Schwartz of the National Renewable Energy Laboratory (NREL); Dr. Robert Farber of Southern California Edison (SCE); Cliff Murley of the Sacramento Municipal Utility District (SMUD); Mark Smith of FPL Energy; Julie Morris of PPM Energy; Rick Koebbe and Bob Szymanski of Pacific Winds, Inc./PowerWorks; Inc.; Ed Duggan and Jeremiah Sota of Oak Creek Energy Systems; Brian Ulberg MDU Resources; John Zack, Kenneth T. Waight III, Pamela Barlow, Joe Nocera, and Dan Mead of AWS Truewind, LLC; Professor Bruce White and Dr. David Lubitz of the University of California at Davis; and Dr. Steve Chin of Lawrence Livermore National Laboratory (LLNL). The Public Interest Energy Research (PIER) Program of the California Energy Commission supported this project.

CONTENTS

- 1 INTRODUCTION 1-1**
 - Objectives and Scope 1-2
 - Project Participants 1-4
 - Report Organization 1-4

- 2 WIND ENERGY FORECASTING INTEGRATION INTO ELECTRICITY GRID OPERATIONS 2-1**
 - Characteristics of Wind Generation..... 2-1
 - How Wind Generation Affects Electricity Grid Operations 2-1
 - System Operators and Balancing Authorities 2-3
 - Area Control Error (ACE)..... 2-3
 - Control Area Objectives..... 2-4
 - System Impacts and Challenges 2-6
 - Importance of Wind Energy Forecasting 2-6
 - Forecasting Wind Generation Ramp Rates 2-7
 - Customizing Wind Energy Forecasts for System Operators 2-9
 - Conclusions..... 2-11

- 3 NEXT-HOUR REGIONAL WIND ENERGY FORECASTING SYSTEM DEVELOPMENT AND TESTING 3-1**
 - Introduction 3-1
 - Forecast System Formulation 3-1
 - Alternative Forecast Methods 3-2
 - Autoregressive Models 3-2
 - Moving Average Models 3-2
 - ARIMA (Box-Jenkins) Models 3-3
 - Artificial Neural Networks (ANN)..... 3-3
 - Support Vector Regression (SVR) 3-4
 - Short-Term Forecast System Design 3-6

Regional Power Production Data	3-10
Temporal Variability of Power Production	3-12
Initial Forecast Performance Evaluation.....	3-13
Specification of Predictors	3-13
Sensitivity Tests.....	3-18
Number and Type of Inputs	3-19
Two Input Variables	3-19
Four Input Variables - Recent Power Level and Trend	3-21
Four Input Variables - Recent Power Level and Trend on Previous Day.....	3-24
Six Input Variables	3-24
Seven Input Variables – Forecast Error	3-27
Training Sample Size.....	3-31
Number of Hidden Nodes	3-34
Multi-Region Forecast Performance Results.....	3-38
Solano Region.....	3-43
Altamont Region.....	3-43
Tehachapi Region.....	3-43
San Geronio Region.....	3-44
Conclusions.....	3-46
Recommendations	3-46

4 NEXT-DAY WIND PLANT ENERGY FORECASTING SYSTEM DEVELOPMENT AND TESTING.....4-1

Introduction	4-1
Mountain View 1 and 2 Wind Project.....	4-3
PowerWorks Wind Project.....	4-6
Screening of Improved Data and Forecast Methodologies	4-6
Forecasting Experiment Plan	4-7
Forecasting Experiment Results.....	4-9
Focus Area 1: Water Surface Temperature (WST)	4-10
Focus Area 2: Model Resolution Experiments	4-18
Focus Area 3: Next-Generation Physics-based Models.....	4-29
Coupled Ocean Atmosphere Modeling System (COAMPS)	4-29
Operational Multiscale Model with Grid Adaptivity Model (OMEGA)	4-30
Weather Research and Forecasting Model (WRF).....	4-30

MASS and WRF Model Experiments.....	4-31
Focus Area 4: Model Output Statistics Formulation	4-40
Screening Multiple Linear Regression (SMLR).....	4-40
Two-Stage SMLR (SMLR2)	4-40
SMLR2 with Direct Power Prediction (SMLR2-DP)	4-41
Stratified Two-Stage SMLR (SMLR2-ST)	4-41
Artificial Neural Network (ANN).....	4-41
Conclusions	4-49
Focus Area 5: Plant Output Model Formulation	4-50
Physical Plant Output Models	4-50
Statistical Plant Output Models	4-50
Plant Scale Power Curves	4-51
Modeling Deviations from the Power Curve.....	4-53
Testing of Improved Forecast Algorithm at Five California Wind Projects	4-59
Experimental Design	4-60
Forecast Delivery Time and Look-ahead Period.....	4-60
Forecast Parameters	4-61
Wind Plant Forecast Sites.....	4-62
Performance Metrics.....	4-62
Physics-Based Models	4-64
Mesoscale Atmospheric Simulation System (MASS)	4-64
Weather Research and Forecasting Model (WRF).....	4-67
Model Output Statistics Methods.....	4-68
Screening Multiple Linear Regression (SMLR).....	4-69
Two-stage Screening Multiple Linear Regression (SMLR2)	4-69
SMLR2 with Direct Power Production Prediction (SMLR2-DP)	4-69
Stratified SMLR2 (SMLR2-ST).....	4-69
Plant Output Models	4-70
Reference Forecasts	4-70
Persistence Forecast	4-71
Climatology Forecast	4-71
Ensemble Forecasts.....	4-71
Forecast Performance Evaluation Results by Wind Project.....	4-72
Mountain View 1 and 2 Wind Project, San Geronio Pass	4-73
Overall Forecast Performance	4-75

Forecast Performance vs. Time Horizon	4-80
Physics-Based Model Comparison	4-82
Oak Creek Energy Systems, Tehachapi	4-85
Overall Forecast Performance	4-88
Performance vs. Forecast Time Horizon	4-92
Physics-Based Model Comparison	4-95
PowerWorks Wind Project, Altamont Pass	4-98
Overall Forecast Performance	4-100
Forecast Performance vs. Time Horizon	4-105
Physics-Based Model Comparison	4-107
Forecast Ensembles	4-109
SMUD Wind Project, Solano	4-110
Overall Forecast Performance	4-113
Forecast Performance vs. Time Horizon	4-117
Physics-Based Model Comparison	4-120
Forecast Ensembles	4-122
High Winds Wind Project, Solano	4-123
Overall Forecast Performance	4-126
Forecast Performance vs. Time Horizon	4-130
Physics-Based Model Comparison	4-132
Forecast Ensembles	4-135
Summary	4-137
Screening of Improved Data and Forecast Methodologies	4-137
Forecast Evaluation at Five California Wind Projects	4-139
Forecasting Protocol	4-139
Overall Forecast Performance Results	4-139
Variability of Power Generation Forecast Error	4-141
Ensemble Forecasts	4-142
Conclusions	4-144
Recommendations	4-146
5 NUMERICAL MODELING OF WIND FLOW AND WIND PLANT POWER CURVES OVER COMPLEX TERRAIN	5-1
Introduction	5-1
The MASS Model	5-4

Ten Case Studies	5-6
Predicting Plant Power Production from Inferred Turbine Wind Speeds.....	5-8
Turbine Wind Speed Ratios From Simulations.....	5-8
The Importance of Stability.....	5-13
Calculating Plant Power Production	5-17
Evaluating Plant Power Predictions	5-17
Power Predictions Using Observed Wind Speeds	5-18
Power Predictions Using Forecasted Wind Speeds	5-23
Power Predictions with Stability Modification	5-26
Summary and Conclusions	5-28
6 HIGH-RESOLUTION WEATHER AND WIND FLOW FORECASTING	6-1
Introduction	6-1
COAMPS Model and Experiment Design.....	6-1
Measurements for Model Validation.....	6-2
Results	6-3
Summary	6-4
7 CONCLUSIONS AND RECOMMENDATIONS	7-1
Conclusions.....	7-1
Short-Term Forecast System Design and Evaluation	7-1
Next-Day Forecast Model Improvements.....	7-2
Improved Plant-Scale Power Curve Models.....	7-2
California Wind Generation Research Dataset (CARD)	7-3
Recommendations	7-3
8 REFERENCES	8-1

LIST OF FIGURES

Figure 1-1 California Mean Wind Power Map at 50-m Elevation and 2005 Rated Capacity of Wind Generation at Principal Wind Resource Areas (California Energy Commission, 2006)	1-1
Figure 2-1 Typical Variation of Total and Regional One-Minute Wind Generation in California on a Summer Day (CA ISO, 2005).....	2-2
Figure 2-2 Buffalo Ridge 115-kV Bus Voltage in kV (Top Trace and Left Scale) and Transformer Output in MW (Bottom Trace and Right Scale) (NREL, 2002)	2-2
Figure 2-3 Relationship between an Individual Control Area and Rest of Interconnection and Definition of Area Control Error (CA ISO, 2005).....	2-4
Figure 2-4 Control Area Objectives Focus on Balancing Load and Interchange vs. Generation (CA ISO, 2005)	2-5
Figure 2-5 Balancing Authority Balances Generation Resources vs. Real-Time Load (CA ISO, 2005).....	2-5
Figure 2-6 Forecast Daily Variation of Hourly Wind Generation at Tehachapi after Expansion to 4000 MW Rated Capacity, based on April 2005 Data (CA ISO, 2005)	2-8
Figure 2-7 Estimated Hourly Wind Generation (MW) and Ramp Rates (MW/hr) after Tehachapi Expansion to 4600 MW, based on Data for April 8, 2005 (CA ISO, 2005).....	2-8
Figure 2-8 Example Display of Real-Time Forecasts for Oak Creek Energy Systems Wind Project in Tehachapi, California (AWS Truewind, 2005).....	2-10
Figure 3-1 A schematic depiction of the flow of data through an artificial neural network (ANN) scheme. Each circle represents a neuron (or a computational element).	3-5
Figure 3-2 A schematic depiction of the proposed short-term (0 to 3 hour) forecast system.....	3-8
Figure 3-3 Autocorrelation of the five-minute power production for four wind power production regions in California based on data for the calendar year 2004.....	3-14
Figure 3-4 Autocorrelation of the 5-minute power production for the Altamont region for the cold (October to March) and warm (April to September) seasons of 2004	3-14
Figure 3-5 Autocorrelation of the 5-minute power production for the San Geronio region for the cold (October to March) and warm (April to September) seasons of 2004.....	3-15
Figure 3-6 Wavelet amplitude spectrum of the power production data for the Tehachapi region for December 2004 based on the Derivative of Gaussian (DOG) wavelet function.....	3-16
Figure 3-7 Wavelet amplitude spectrum of the power production data for the Tehachapi region for July 2004 based on the Derivative of Gaussian (DOG) wavelet function	3-16

Figure 3-8 Mean absolute error (% of capacity) vs. look-ahead time for the 2-input ANN and persistence forecasts for the Tehachapi region. The two inputs to the ANN algorithm were PP1 and PPTR01 from Table 3-2.3-20

Figure 3-9 Skill score relative to persistence vs. look-ahead time for a 2-input ANN forecast for the Tehachapi region. The two inputs to the ANN algorithm were PP1 and PPTR01 from Table 3-2.3-20

Figure 3-10 Mean absolute errors (% of capacity) vs. look-ahead time for the 2-input ANN and the 2-input MLR forecast methods. The inputs to both forecast algorithms were the PP1 and PPTR01 variables described in Table 3-2.3-22

Figure 3-11 Skill score of the 2-input ANN forecast relative to the 2-input MLR method vs. look-ahead time. The two inputs for both methods were the PP1 and PPTR01 variables described in Table 3-2.3-22

Figure 3-12 Skill score of the 4-input ANN forecasts (Exp. 2) relative to persistence vs. forecast look-ahead time. The four inputs to the ANN algorithm were PP1, PPO2, PPTR01, and PPTR02 as described in Table 3-2.3-23

Figure 3-13 Skill score of the 4-input ANN method (Exp. 2) relative to a 2-input ANN method (Exp. 2) vs. forecast look-ahead time. The inputs to the four-input ANN forecast were PP1, PPO2, PPTR01, and PPTR02 as described in Table 3-2.3-23

Figure 3-14 Skill score of the 4-input ANN forecast (Exp. 3) relative to. persistence vs. forecast look-ahead time. The four inputs to the ANN algorithm were PP1, PPDA, PPTR01, and PPTRDA as described in Table 3-2.3-25

Figure 3-15 Skill score of the 4-input ANN method (Exp. 3) using variables PP1, PPDA, PPTR01, and PPTRDA relative to a 2-input ANN method (Exp. 1) using variables PP1 and PPTR01 vs. forecast look-ahead time.3-25

Figure 3-16 Skill score of the 6-input ANN forecast method relative to persistence vs. forecast look-ahead. The six inputs to the ANN algorithm were P1, P2, PDA, PTR1, PTR2, and PTDA variables described in Table 3-2.....3-26

Figure 3-17 Skill score of the 6-input ANN forecast method relative to a 2-input ANN method vs. forecast look-ahead time. The inputs to the 6-input ANN algorithm were the P1, P2, PDA, PTR1, PTR2, and PTDA, and the inputs to the 2-input algorithm were the P1 and PTR1 variables described in Table 3-2.3-26

Figure 3-18 Skill score relative to persistence vs. forecast look-ahead time. The 7 inputs to the ANN algorithm were the P1, P2, PDA, PTR1, PTR2, and PTDA and ERR variables as described in Table 3-2.3-28

Figure 3-19 Skill score of the 7-input ANN forecast method relative to a 6-input ANN method vs. forecast look-ahead time. The 7 inputs to the ANN algorithm were the P1, P2, PDA, PTR1, PTR2, PTDA, and ERR variables described in Table 3-2.....3-28

Figure 3-20 Skill scores of the training sample generated using the ANN algorithm with 2, 4, 6, and 7 inputs relative to persistence vs. forecast look-ahead time for the training sample (June 2004).....3-29

Figure 3-21 Skill scores of the evaluation sample generated using the ANN algorithm with 2, 4, 6, and 7 inputs relative to persistence vs. forecast look-ahead time (July 2004).3-29

Figure 3-22 Difference in the skill scores relative to persistence between the training and evaluation samples by forecast look-ahead time for the ANN method with 2, 4, 6, and 7 inputs for the Tehachapi region.....3-30

Figure 3-23 Mean absolute error (% of capacity) vs. look-ahead time for a 6-input ANN forecast procedure with training sample sizes of 3.5 days, 1 week, 2 weeks, 1 month, and 2 months.	3-32
Figure 3-24 Skill scores relative to persistence vs. look-ahead time for a 6-input ANN forecast procedure with training sample sizes of 3.5 days, 1 week, 2 weeks, 1 month, and 2 months.	3-32
Figure 3-25 Difference in persistence-based skill scores between the training and evaluation samples vs. look-ahead time for ANN-based forecasts with training sample sizes of 3.5 days, 1 week, 2 weeks, 1 month, and 2 months for the Tehachapi region	3-33
Figure 3-26 The average CPU time (seconds) on a single Intel P4 Xeon 2.2 GHz processor required to train the 6-input ANN for each training sample size.	3-33
Figure 3-27 Mean absolute error (% of capacity) vs. forecast look-ahead time for an ANN forecast procedure with 0, 1, 2, 4, and 6 hidden nodes. All experiments used a one-month training sample (June 2004) and a one-month evaluation sample (July 2004).	3-35
Figure 3-28 Skill scores relative to persistence vs. forecast look-ahead time for an ANN forecast procedure with 0, 1, 2, 4, and 6 hidden nodes. All experiments used a one-month training sample (June 2004) and a one-month evaluation sample (July 2004).	3-35
Figure 3-29 Difference in the persistence-based skill scores between the training and evaluation samples vs. forecast look-ahead time for an ANN procedure with 0, 1, 2, 4, and 6 hidden nodes for the Tehachapi region.....	3-36
Figure 3-30 The average CPU time (seconds) on an Intel P4 Xeon 2.2 GHz processor to train an ANN with a one-month training sample and the specified number of hidden nodes.	3-36
Figure 3-31 Mean error (bias) as a percentage of region capacity vs. look-ahead period for the year of 2004 for ANN-based forecasts for each of four California wind resource areas and the aggregated areas.....	3-39
Figure 3-32 Average magnitude (% of capacity) of the monthly mean error (bias) vs. look-ahead period for 2004 for ANN-based forecasts in each of four California wind power production regions and the aggregated production for all four regions.	3-39
Figure 3-33 Mean absolute errors of ANN-power production forecasts (% of capacity) vs. forecast look-ahead period during the six cold season months (November through April), for the four California wind resource areas and the aggregated areas.....	3-40
Figure 3-34 Skill scores of ANN-power production forecasts relative to persistence vs. forecast look-ahead period for the six cold season months (November through April) for the four California wind resource areas and the aggregated areas.....	3-40
Figure 3-35 Mean absolute errors of ANN power production forecasts (% of capacity) vs. look-ahead period for the six warm season months (May through October) for the four California wind resource areas and the aggregated areas.	3-41
Figure 3-36 Skill scores of ANN power production forecasts relative to persistence vs. forecast look-ahead period for the six warm season months (May through October) for each of four California wind resource areas and the aggregated areas.	3-41

Figure 3-37 Average skill scores of the regional and aggregate ANN forecasts relative to persistence vs. forecast look-ahead period during the six warm season months for the four California wind resource areas	3-45
Figure 3-38 Average skill scores the regional and aggregate ANN forecasts relative to persistence vs. forecast look-ahead period during the six cold season months for the four California wind power resource areas.....	3-45
Figure 4-1 Power curve for a GE 1.5 MW turbine (from the GE Power website at http://www.gepower.com/prod_serv/products/wind_turbines/en/15mw/tech_data.htm)	4-5
Figure 4-2 The MAE of generation forecasts (% of installed capacity) and the percentage of hours for 20 forecasted power production bins for the Mountain View wind plant. Each bin is labeled by the forecasted production at the upper end of the bin.....	4-5
Figure 4-3 A schematic depiction of the main components of the eWind forecast system.....	4-8
Figure 4-4 A graphical depiction of the NOAA optimum interpolation WST analysis on a 1-degree global grid for March 2005	4-11
Figure 4-5 Sample MODIS water surface temperature image.....	4-12
Figure 4-6 Initial skin temperature (degrees F) specification for the August 23, 2002, forecast simulation based on the NCEP OI Water Surface Temperature (WST) analysis.	4-14
Figure 4-7 Initial skin temperature (degrees F) specification for the August 23, 2002, forecast simulation based on the MODIS Water Surface Temperature (WST) dataset.	4-14
Figure 4-8 Difference in the initial skin temperature (degrees F) specification between the NCEP OI and MODIS WST for the August 23, 2002 forecast simulation.	4-15
Figure 4-9 Forecasted and observed power production for the Mountain View wind plant and forecast period beginning at 9 a.m. PDT August 23, 2002.....	4-16
Figure 4-10 Figure 7. Change in hourly absolute forecast error (% of installed capacity) for MODIS vs. OI WSTs for the Mountain View wind plant for August 2002.....	4-16
Figure 4-11 Percentage reduction in the mean absolute error of the power production forecasts for the PowerWorks and the Mountain View wind plants.....	4-18
Figure 4-12 A depiction of the geographical area covered by the coarsest resolution grid (40 km grid cell size) used to study the impact of physics-based model grid resolution on the performance of day-ahead power production forecasts for the PowerWorks and Mountain View wind projects	4-19
Figure 4-13 A depiction of the geographical areas covered by the three nested grids used to study the impact of physics-based model grid resolution (or the horizontal size of the grid cells) on the performance of day-ahead power production forecasts for the PowerWorks wind plant in Altamont Pass.....	4-20
Figure 4-14 A depiction of the geographical areas covered by the three nested grids used to study the impact of physics-based model grid resolution (or the horizontal size of the grid cells) on the performance of day-ahead power production forecasts for the Mountain View wind plant in San Gorgonio Pass.	4-21
Figure 4-15 Mean absolute error by physics-based model grid cell size for the 1-to 48-hour raw physics-based forecasts of the hourly power production for the PowerWorks wind plant for the three forecast test months.....	4-22

Figure 4-16 Mean absolute error by physics-based model grid cell size for the 1- to 48-hour MOS-adjusted physics-based forecasts of the hourly power production for the PowerWorks wind plant for the three forecast test months.....	4-22
Figure 4-17 The incremental reduction of the mean absolute error (MAE) of the 48-hour wind energy forecasts vs. grid cell size for the PowerWorks wind plant obtained by applying a MOS procedure to the raw physics-based model data.....	4-24
Figure 4-18 Mean absolute error vs. physics-based model grid resolution for 1- to 48-hour raw physics-based forecasts of the hourly power production for the Mountain View wind plant for the three forecast test months.	4-25
Figure 4-19 Mean absolute error vs. physics-based model grid resolution for 1- to 48-hour Model Output Statistics (MOS)-adjusted, physics-based forecasts of the hourly power production for the Mountain View wind plant for the three forecast test months	4-26
Figure 4-20 The increment of reduction in the mean absolute error of the 1 to 48 hour ahead hourly power production forecasts for the Mountain View wind plant obtained by applying a MOS procedure to the raw physics-based model data.	4-27
Figure 4-21 Comparison of the mean absolute error of 1- to 48-hour ahead forecasts of the hourly power production for the three screening phase test months for the PowerWorks and Mountain View wind plants from the raw (no MOS) WRF (black bars) and MASS (red bars) physics-based model simulations.....	4-33
Figure 4-22 Comparison of the mean absolute error of 1- to 48-hour forecasts of the hourly power production for the three screening phase test months for the PowerWorks and Mountain View wind plants from the MOS-adjusted WRF (black bars) and MASS (red bars) physics-based model simulations.....	4-34
Figure 4-23 Comparison of the mean absolute error of 1- to 48-hour forecasts of the meteorological tower wind speed for the three screening phase test months for the PowerWorks and Mountain View wind plants from the raw (no MOS) WRF (black bars) and MASS (red bars) physics-based model simulation data.	4-34
Figure 4-24 Comparison of the mean absolute error of the 1- to 48-hour forecasts of the meteorological tower wind speed for the three test months for the PowerWorks and Mountain View wind plants from the MOS adjusted WRF (black bars) and MASS (red physics-based model simulation data bars).....	4-35
Figure 4-25 MAE of the power production forecasts by forecast look-ahead hour from the raw WRF (black line with rectangle markers) and MASS (red line with diamond markers) model simulation data for the PowerWorks wind plant for May 2002.	4-35
Figure 4-26 MAE of the power production forecasts by forecast look-ahead hour from the MOS-adjusted WRF (black line with rectangle markers) and MASS (red line with diamond marker) model simulation data for the PowerWorks wind plant for May 2002.	4-36
Figure 4-27 MAE of the power production forecasts by forecast look-ahead hour from the raw WRF (black line with rectangle markers) and MASS (red line with diamond marker) model simulation data for the Mountain View wind plant for April 2002.....	4-38
Figure 4-28 MAE of the power production forecasts by forecast look-ahead hour from the MOS-adjusted WRF (black line with rectangle markers) and MASS (red line with diamond markers) model output data for the Mountain View wind plant for April 2002	4-38

Figure 4-29 Difference between the hourly absolute errors of the WRF and MASS MOS-adjusted power production forecasts (red line for the PowerWorks wind plant and May 2002. 4-39

Figure 4-30 Difference between the hourly absolute errors of the WRF and MASS MOS-adjusted 1- to 24-hour power production forecasts (red line) for the Mountain View wind plant and April 2002. 4-39

Figure 4-31 A schematic depiction of the Artificial Neural Network (ANN) configuration used for the forecast experiments in Focus Area 4. 4-42

Figure 4-32 Comparison of the monthly MAE of 1- to 48-hour forecasts of the hourly power production for the three test months at the PowerWorks wind plant for the five MOS procedures listed in Table 4-5. 4-44

Figure 4-33 Comparison of the monthly MAE of 1- to 48-hour forecasts of the hourly power production for the three test months for the Mountain View wind plant for the five MOS procedures listed in Table 4-5. 4-44

Figure 4-34 Comparison of the monthly MAE of 1- to 48-hour forecasts of the hourly average wind speed for the three test months for 11 anemometer locations at the PowerWorks wind plant, for the five MOS procedures listed in Table 4-5. 4-45

Figure 4-35 Comparison of the monthly mean absolute error of 1- to 48-hour forecasts of the hourly average wind speed for the three test months for the Catellus anemometer site in the Mountain View wind plant, for the five MOS procedures listed in Table 4-5. 4-45

Figure 4-36 Mean absolute error of the power production forecasts vs. forecast “look-ahead” hour from the raw SMLR2 (blue line with rectangle markers) and stratified SMLR2 (yellow line with triangle markers) MOS procedures for the PowerWorks wind plant for May 2002. 4-47

Figure 4-37 Mean absolute error of the power production forecasts vs. forecast “look-ahead” hour from the raw SMLR2 (blue line with rectangle markers) and stratified SMLR2 (yellow line with triangle markers) MOS procedures for the Mountain View wind plant for April 2002. 4-47

Figure 4-38 Differences between the hourly absolute errors of the SMLR2 and SMLR2-ST MOS 1- to 24-hour forecasts (red line) of the hourly power production of the PowerWorks wind plant for May 2002. 4-48

Figure 4-39 Differences between the hourly absolute errors of the SMLR2 and SMLR2-ST MOS 1- to 24-hour forecasts (red line) of the hourly power production of the Mountain View wind plant for April 2002. 4-48

Figure 4-40 The percentage reduction in the MAE of the SMLR2-ST power production and wind speed forecasts relative to forecasts produced by the SMLR2 method for the sample of six test months at the PowerWorks and the Mountain View wind plants. 4-49

Figure 4-41 Hourly power output vs. measured hourly average wind speed (black triangles) for the PowerWorks M127 cluster for July 2002 and hourly power output estimates (red circles) using a plant-scale power curve derived from June 2002 data. 4-52

Figure 4-42 Hourly power output vs. measured hourly average wind speed (black triangles) for the PowerWorks cluster M225 for July 2002 and hourly power output estimates (red circles) using a plant-scale power curve derived from June 2002 data. 4-52

Figure 4-43 Hourly deviation (kW) from the median based power curve as a function of wind direction at Tower M127.	4-54
Figure 4-44 Hourly deviation (kW) from the median based power curve as a function of the wind direction at Tower M225.	4-55
Figure 4-45 Hourly deviation (kW) from the median based power curve as a function of the hour of the day (UTC) for the PowerWorks M127 cluster for July 2002.....	4-55
Figure 4-46 Hourly deviation (kW) from the median based power curve as a function of the hour of the day (UTC) for the PowerWorks M225 cluster for July 2002.....	4-56
Figure 4-47 Reduction in MAE of the power production forecasts due to the modeling of the deviations from the median-based plant-scale power curve for the three test months for the PowerWorks (POW) and Mountain View (MVW) wind plants.	4-57
Figure 4-48 A conceptual schematic of the ensemble forecast method.	4-58
Figure 4-49 The geographical domain covered by the 100 X 80 matrix of 20-km grid cells used to produce the parent coarse grid MASS-6 simulations for the project.	4-65
Figure 4-50 The geographical domain covered by the 80 X 80 matrix of 5-km grid cells used to produce the high resolution MASS-6 simulations over Northern California.....	4-66
Figure 4-51 The geographical domain covered by the 80 X 80 matrix of 5-km grid cells used to produce the high resolution MASS-6 simulations over Southern California.	4-66
Figure 4-52 The geographical domain covered by COAMPS 12-km (left) and 4-km (right) grids used to produce forecast simulation for this project.	4-67
Figure 4-53 The geographical domain covered by the 100 X 100 matrix of 10-km grid cells used to produce the WRF simulations for the project.	4-68
Figure 4-54 Number of hours for which sufficient data were available (green) and unavailable (red) to compute forecast performance statistics for the Mountain View power production forecasts	4-74
Figure 4-55 Number of hours for which sufficient data were available (green) and unavailable (red) to compute forecast performance statistics for the wind speed forecasts for the Catellus Meteorological Tower at the Mountain View wind plant.	4-74
Figure 4-56 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the forecasts for the Mountain View wind plant.....	4-78
Figure 4-57 Cumulative absolute error frequency distribution (% of hours with an absolute error less than or equal to the value on the horizontal axis) for hours 1 through 48 of the forecasts for the Mountain View wind plant.	4-78
Figure 4-58 Mean Absolute Error (MAE) of the Mountain View power production forecasts vs. the forecasted power production.....	4-79
Figure 4-59 Mean Absolute Error (MAE) of the Catellus wind speed forecasts vs. the forecasted wind speed.	4-79
Figure 4-60 MAEs of power production forecasts vs. forecast time horizon for 12 months of <i>eWind</i> (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers) forecasts for the Mountain View wind plant.	4-81
Figure 4-61 MAEs of power production forecasts vs. by forecast time horizon for 12 months of <i>eWind</i> (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) wind speed forecasts for the Mountain View wind plant.....	4-81

Figure 4-62 Number of hours for which sufficient data were available (green) and unavailable (red) to compute verification statistics for the PowerWorks power production forecasts.....4-86

Figure 4-63 Number of hours for which sufficient data were available (green) and unavailable (red) to compute verification statistics for the PowerWorks power production forecasts.....4-87

Figure 4-64 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the forecasts for the Oak Creek wind plant.4-90

Figure 4-65 Cumulative absolute error frequency distribution (% of hours with an absolute error less than are equal to the value on the horizontal axis) for hours 1 to 24 (red line) and 25 to 48 (blue line) of the seven months of power production forecasts for the Oak Creek wind plant.4-91

Figure 4-66 MAE of the Oak Creek power production forecasts vs. the forecasted power production.4-91

Figure 4-67 MAEs of the Oak Creek wind speed forecasts vs. the forecasted wind speed..4-93

Figure 4-68 MAE of power production forecasts at the Oak Creek wind plant vs. forecast time horizon the look-ahead interval) for seven months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) forecasts.....4-93

Figure 4-69 MAE of wind-speed forecasts at Oak Creek wind plant vs. forecast time horizon the look-ahead interval) for seven months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers).....4-94

Figure 4-70 Number of hours for which data were available (green) and unavailable (red) for the computation of verification statistics for the PowerWorks power production forecasts produced by the best performing forecast method (MASS-6 SMLR2-ST).4-99

Figure 4-71 Number of hours for which data were available (green) and unavailable (red) for the computation of verification statistics for the wind speed forecasts for PowerWorks Tower M427 by the best performing forecast method (MASS-6 SMLR2-ST)4-100

Figure 4-72 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the power production forecasts for the PowerWorks wind plant. The data sample includes 17,472 forecast hours.4-102

Figure 4-73 Cumulative absolute error frequency distribution (% of hours with an absolute error less than are equal to the value on the horizontal axis) for hours 1 to 48 of the power production forecasts for the PowerWorks wind plant.4-103

Figure 4-74 MAE of the PowerWorks power production forecasts vs. the forecasted power production.....4-104

Figure 4-75 MAE of the 1 to 48 hour PowerWorks Tower M427 wind speed forecasts vs. the forecasted wind speed.....4-104

Figure 4-76 MAE of power production forecasts at the PowerWorks wind plant vs. forecast time horizon for 12 months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) forecasts.4-106

Figure 4-77 MAE of wind-speed forecasts at Tower M427 at the PowerWorks wind plant vs. forecast time horizon for 12 months of <i>eWind</i> (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers) forecasts.....	4-106
Figure 4-78 Number of hours for which data were available (green) and unavailable (red) for the computation of verification statistics for the SMUD power production forecasts.....	4-112
Figure 4-79 Number of hours for which data were available (green) and unavailable (red) for the computation of verification statistics for the wind speed forecasts for the SMUD meteorological tower.	4-112
Figure 4-80 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the forecasts for the SMUD wind plant. The data sample includes 12,746 forecast hours.	4-115
Figure 4-81 Cumulative absolute error frequency distribution (% of hours with an absolute error less than are equal to the value on the horizontal axis) for hours 1 to 48 of the forecasts for the PowerWorks wind plant. The data sample includes 12,746 forecast hours.	4-116
Figure 4-82 MAE of the SMUD 1- to 48-hour power production forecasts vs. the forecasted power production. The MAEs are shown for 21 bins of forecasted production.	4-116
Figure 4-83 MAE of the 1- to 48-hour SMUD Met Tower wind speed forecasts vs. the forecasted wind speed..	4-117
Figure 4-84 MAE by forecast time horizon for 11 months of <i>eWind</i> (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers) power production forecasts for the SMUD wind plant.	4-119
Figure 4-85 MAE by forecast time horizon for 11 months of <i>eWind</i> (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers) wind speed forecasts for the 50 m level at SMUD Meteorological Tower.....	4-119
Figure 4-86 Number of hours with available and unavailable data for the computation of power production forecast verification statistics for the High Winds wind plant for the July 2004 to June 2005 period.	4-125
Figure 4-87 Number of hours with available and unavailable data for the computation of wind speed forecast verification statistics for the High Winds wind farm for the July 2004 to June 2005 period.	4-125
Figure 4-88 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the forecasts for the High Winds wind plant.	4-128
Figure 4-89 Cumulative absolute error frequency distribution (% of hours with an absolute error less than are equal to the value on the horizontal axis) for hours 1 to 48 of the forecasts for the High Winds wind plant.....	4-128
Figure 4-90 MAE of the High Winds 1- to 48-hour power production forecasts vs. the forecasted power production.....	4-129
Figure 4-91 MAEs of the 1- to 48-hour High Winds wind speed forecasts vs. the forecasted wind speed.	4-129

Figure 4-92 MAEs of power production forecasts at High Winds wind plant vs. forecast time horizon for 12 months of *eWind* (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers) forecasts.....4-131

Figure 4-93 MAEs of wind speed forecasts at High Winds wind plant based on *eWind* (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers) forecasts.....4-131

Figure 4-94 MAEs of power production forecasts for the High Winds wind plant based on the MASS-6, COAMPS, and WRF physics-based models vs. forecast time horizon the look-ahead interval) for four months (March to June 2005).....4-135

Figure 4-95 Empirical median plant-scale power curves derived from measured wind energy generation and wind speed data for three of the participating wind plants4-142

Figure 5-1 Hourly power output vs. measured hourly average wind speed (black triangles) for the PowerWorks cluster 127 for July 2002 and hourly power output estimates (red circles) using a plant-scale power curve derived from hourly data from June 2002.5-3

Figure 5-2 The area encompassed by the domain and the associated terrain representation for the 100-m MASS model grid.....5-5

Figure 5-3 A broader view of the terrain of the Altamont Pass.5-6

Figure 5-4 The 18-m wind field at 0430 UTC 16 July 2002 (8:30 PM PST), several hours into the 100 m simulation..5-8

Figure 5-5 Evolution of turbine wind speed ratios over the 6-hr 100 m simulation beginning at 0000 UTC July 16, 2002.....5-9

Figure 5-6 Evolution of turbine wind speed ratios over the 6-hr 100 m simulation beginning at 0000 UTC May 19, 2002.5-10

Figure 5-7 Evolution of turbine wind speed ratios over the 6-hr 100 m simulation beginning at 0000 UTC December 2, 2001.....5-10

Figure 5-8 Comparison of mean turbine wind speed ratios calculated from the 100-m simulations to those calculated from the UC Davis wind tunnel data for the dominant wind direction of 240 degrees.5-12

Figure 5-9 The mean wind speed ratio over all turbine locations for each of the ten 100-m numerical simulations and for simulations of two wind directions in the UC Davis wind tunnel.5-12

Figure 5-10 Hourly vertical profiles of potential temperature (in degrees Kelvin) from the 100 m simulation at Tower M127 from 0100 to 0600 UTC December 2, 2001.5-15

Figure 5-11 Hourly vertical profiles of potential temperature (in degrees Kelvin) from the 100 m simulation at Tower M127 from 0700 to 1200 UTC December 15, 2001.5-15

Figure 5-12 Hourly vertical profiles of potential temperature (in degrees Kelvin) from the 100 m simulation at Met Tower 127 from 0100 to 0600 UTC July 16, 2002.5-16

Figure 5-13 The median bulk turbulence scale (defined by Eqn 5-2) vs. the mean turbine wind speed ratio for each of the 10 high-resolution numerical simulations.5-16

Figure 5-14 Hourly power output vs. measured hourly average wind speed (black triangles) for the PowerWorks M127 turbine cluster for July 2002 and hourly power output predictions (red circles) using the numerically-simulated mean turbine wind speed ratios.....5-18

Figure 5-15 Comparison of two reference methods of predicting the plant power output using measured M127 wind speeds from the Dec 2001 to Sep 2002 dataset.	5-20
Figure 5-16 Comparison of mean error (ME) and mean absolute error (MAE) for three methods of predicting the plant power output using observed M127 wind speeds from the Dec 2001 to Sep 2002 dataset.	5-21
Figure 5-17 Comparison of mean error (ME) and mean absolute error (MAE) for three methods of predicting the plant power output using observed M127 wind speeds from the June 2001 to July 2004 dataset.	5-21
Figure 5-18 Monthly comparison of the mean absolute error (MAE) for the three methods of predicting the plant power output using observed M127 wind speeds from the December 2001 to September 2002 dataset.	5-22
Figure 5-19 Monthly comparison of the mean error (ME) for the three methods of predicting the plant power output using observed M127 wind speeds from the Dec 2001 to Sep 2002 dataset.	5-22
Figure 5-20 Comparison of the mean error (ME) and mean absolute error (MAE) produced by the empirical plant power curve for all months and those produced by employing the plant power curve for its best-performing months (April-August) and the mean turbine ratios from the numerical simulation method for the other five months.	5-23
Figure 5-21 Comparison of three methods of predicting the plant power output using forecasted M127 wind speeds from the December 2001 to September 2002 dataset.	5-24
Figure 5-22 Monthly comparison of the mean absolute error (MAE) for the three methods of predicting the plant power output using forecasted M127 wind speeds from the Dec 2001-Sep 2002 dataset.	5-25
Figure 5-23 Monthly comparison of the mean error (ME) for the three methods of predicting the plant power output using forecasted Met Tower 127 wind speeds from the Dec 2001-Sep 2002 dataset.	5-25
Figure 5-24 Comparison of four methods of predicting the plant power output using observed M127 wind speeds from the ten 6-hr periods simulated by the high-resolution numerical model.	5-27
Figure 5-25 Comparison of four methods of predicting the plant power output using observed M127 wind speeds only from the ten 24-hr periods simulated by the model.	5-28
Figure 6-1 Resolution of terrain elevation (meters) vs. grid size (km) of the nested domain: (a) $\Delta x = 12$ km (nest_2), (b) $\Delta x = 4$ km (nest_3), (c) $\Delta x = 1.333$ km (nest_4), and (4) $\Delta x = 0.444$ km (nest_5).....	6-6
Figure 6-2 Weekly mean absolute errors of NARAC forecasts averaged over the selected stations for June 1-7, 2005.	6-7
Figure 6-3 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts averaged over the selected stations for December 2004.....	6-8
Figure 6-4 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts for July 2004 (a and b), August 2004 (c and d), and September 2004 (e and f), respectively.	6-9

Figure 6-5 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts for October 2004 (a and b), November 2004 (c and d), and January 2005 (e and f), respectively.6-10

Figure 6-6 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts for February 2005 (a and b), March 2005 (c and d), and April 2005 (e and f), respectively.6-11

Figure 6-7 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts for May 2005.6-12

Figure 6-8 Vertical profiles of forecast horizontal winds at station 127 from the nest_3 domain ($\Delta x = 4$ km), (a) from 00Z of July 01 2004, and (b) from 00Z of December 01, 2004.6-12

Figure 6-9 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts with calibration of modeled terrain height for June 2005 (a and b) and December 2004 (c and d).6-13

LIST OF TABLES

Table 3-1 An example of the raw one-minute regional power production data provided by the CA ISO for 2004.....	3-11
Table 3-2 Description of predictors used in the 0 to 3 hour forecast experiments-.....	3-17
Table 4-1 <i>eWind</i> monthly and annual forecast performance statistics for the Mountain View wind plant from the previous Energy Commission-EPRI project (Energy Commission PIER EPRI, 2003)	4-2
Table 4-2 <i>eWind</i> monthly and annual forecast performance statistics for the PowerWorks wind plant from the previous Energy Commission-EPRI project (Energy Commission PIER-EPRI, 2003)	4-2
Table 4-3 Test wind resources and months for evaluation of candidate data and methodologies.....	4-7
Table 4-4 Focus areas and associated experiments	4-9
Table 4-5 MOS Techniques used in the Focus Area #4 Experiments	4-41
Table 4-6 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-ST) for the Mountain View wind plant	4-76
Table 4-7 Monthly MAE as a percentage of rated capacity for the 1- to 48-hour Mountain View power production and wind speed forecasts produced from the output of the three physics-based models used in this project.....	4-82
Table 4-8 Monthly MAE for the March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the Mountain View wind plant.	4-85
Table 4-9 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-DP) for the Oak Creek wind plant	4-89
Table 4-10 Monthly MAE as a percentage of capacity for the 1- to 48-hour Oak Creek power production and wind speed forecasts by the three physics-based models.	4-96
Table 4-11 Monthly MAE for the March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the Oak Creek wind plant.	4-98
Table 4-12 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-ST) for the PowerWorks wind plant.....	4-101
Table 4-13 Monthly mean absolute (MAE) of power production (% of capacity) and wind speed (m/s) for forecasts generated from three different physics-based models for the PowerWorks wind plant.....	4-108

Table 4-14 Monthly MAE for the March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the PowerWorks wind plant.	4-110
Table 4-15 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-ST) for the SMUD wind plant	4-113
Table 4-16 Monthly MAE of power production (% of capacity) and wind speed (m/s) for forecasts generated by the three physics-based models for the SMUD wind plant.	4-122
Table 4-17 March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the SMUD wind plant	4-123
Table 4-18 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-ST) for the High Winds wind plant	4-126
Table 4-19 Monthly MAEs of power production (% of capacity) and wind speed (m/s) for forecasts generated by the three physics-based models for the High Winds wind plant.	4-133
Table 4-20 Monthly MAE for the March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the High Winds wind plant.	4-136
Table 4-21 Summary of the net reduction in the power production forecast MAE associated with the forecast system modifications in each focus area.	4-138
Table 4-22 Mean absolute error of the best performing individual power production and wind speed forecasts for the entire 48-hour forecast period for each of the five participating wind plants.	4-140
Table 4-23 Mean absolute errors of ensemble power production (% of rated capacity) and wind speed (m/s) forecasts for the five participating wind plants	4-143
Table 5-1 Ten Cases Chosen for Numerical Simulations Using a 100-m Grid	5-7
Table 6-1 Station information for meteorological towers at Altamont Pass. Station and modeled terrain heights are also shown for different grid resolutions.	6-5

1 INTRODUCTION

California has good potential for developing new wind generation capacity beyond the approximately 2100 megawatts of rated capacity in place at the end of 2005 (American Wind Energy Association, 2006). California's Renewables Portfolio Standard, which calls for 20% renewables in the generation mix by the end of 2010, is expected to result in a large increase of the installed wind capacity in the state. Most of the current capacity is located in the five principal wind resource areas of the state (Solano, Altamont, Pacheco, Tehachapi, and San Gorgonio), shown in Figure 1-1. The new capacity is expected to be installed in these and other promising California wind resource areas in Northern and Southern California.

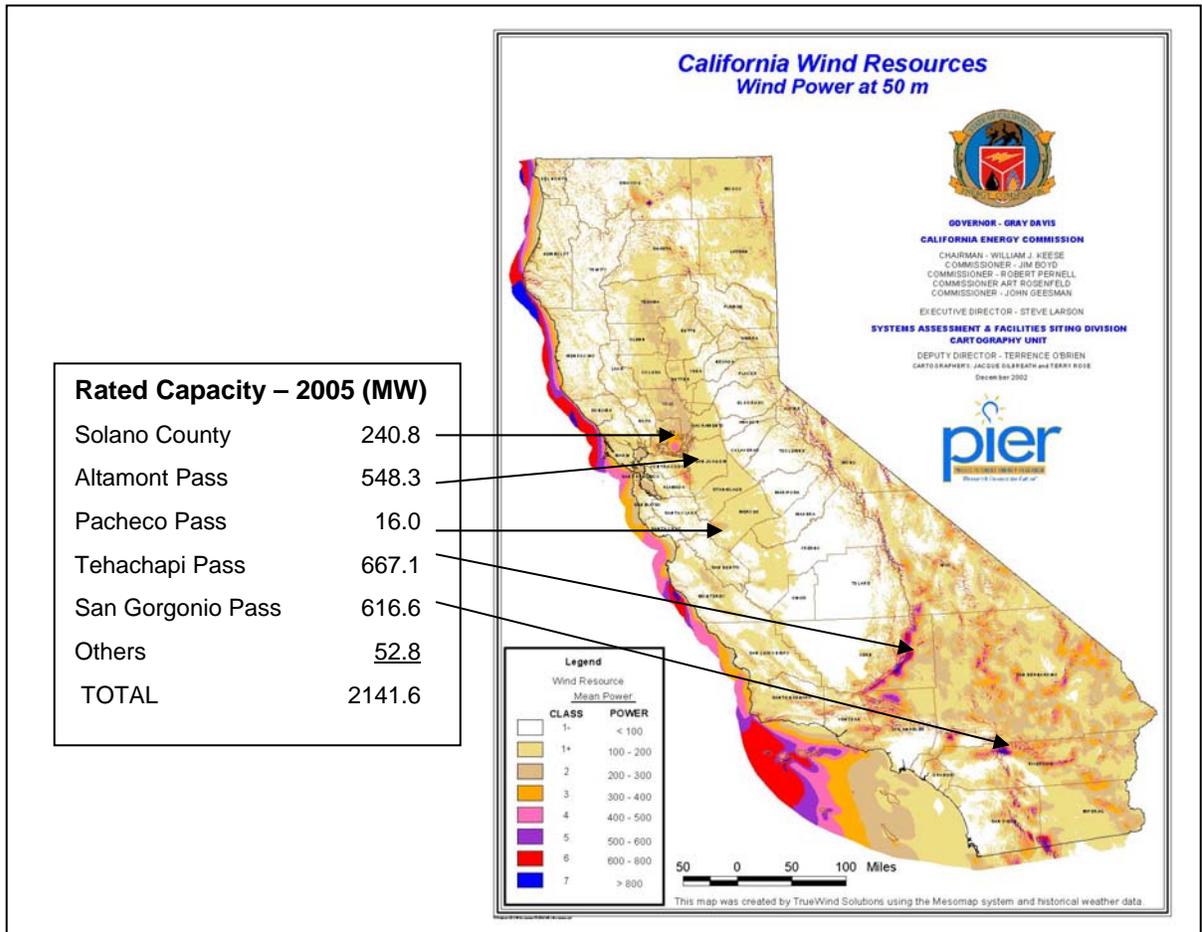


Figure 1-1 California Mean Wind Power Map at 50-m Elevation and 2005 Rated Capacity of Wind Generation at Principal Wind Resource Areas (California Energy Commission, 2006)

Because wind generation is an intermittent resource and large concentrations of wind generation can affect electricity grid operations and reserve requirements, development of accurate wind energy forecasting tools will become an increasingly critical need for managing wind and other intermittent generation resources connected to the California grid. Accurate next-hour and next-day forecasts will make it possible to optimize the response to rapid changes in wind generation to balance load and supply reserve and regulation resources to the grid.

In 2002, the California Energy Commission (Energy Commission) and EPRI completed testing of two forecasting systems at Altamont and at San Geronio (Energy Commission-EPRI, 2003a and 2003b; EPRI 2003a). Two wind energy forecasting system developers, Risoe National Laboratory and TrueWind Solutions, applied their meteorology-based, meso-scale modeling algorithms to generate twice-daily, 48-hour forecasts of hourly wind speed and energy generation during a 12-month period. The host wind projects were the 90 MW Wind Power Partners/Windworks project, operated by Powerworks at Altamont Pass, and the 66.6 MW Mountain View 1 and 2 wind project, operated by Seawest at San Geronio Pass. Based on the monthly and annual mean absolute errors (MAE) of the forecast vs. observed data, the Risoe and TrueWind forecasts performed better than simple persistence and climatology forecasts. However, the forecast errors were still significant, indicating that additional research is needed to incorporate improved forecast technology and forecast performance.

In 2004, Energy Commission, EPRI, and California Independent System Operator (CA ISO) initiated a new 18-month project to build on the first project and develop and test improved wind energy forecast algorithms for both short-term forecasts (regional five-minute forecasts over three hours) and intermediate-term forecasts (hourly wind plant forecasts over 48 hours) in the principal wind resource areas of the state. The project was completed during 2005, and the results are presented in the four-volume report *California Regional Wind Energy Forecasting System Development and Testing* (Energy Commission and EPRI, 2006a, 2006b, 2006c, and 2006d).

This report, *California Regional Wind Energy Forecasting System Development – Volume 2: Wind Energy Forecasting System Development and Testing and Numerical Modeling of Wind Flow over Complex Terrain*, describes the objectives, scope, key results, and recommendations of the project. The other report volumes include *Volume 1: Executive Summary*; *Volume 3: Wind Tunnel Modeling of Wind Flow over Complex Terrain*; and *Volume 4: California Wind Generation Research Dataset (CARD)* (Energy Commission and EPRI, 2006a, 2006c, and 2006d).

Objectives and Scope

The overall project objectives include both economic and technical goals.

The overall economic goals include:

- Support the California Independent System Operator's (CA ISO) development of a viable competitive market for intermittent wind resources.

- Pave the way for increasing market penetration of renewable resources.

The overall technical goals include:

- Leverage the experiences gained under the prior forecasting efforts to improve forecast accuracy.
- Provide capability to generate accurate forecasts for both short-term and longer-term forecast timeframes.

The specific objectives include:

- Develop and test short-term forecasting algorithms with higher accuracy than persistence forecasts to provide real-time forecasting capability and support system real-time updates to meet dispatching needs.
- Determine sources of forecast error and assess methods to reduce errors for both next-hour and next-day forecasts, for example: Improved input data, finer grid sizes in meso-scale models, and improved statistical models for short-term forecasting and model operating statistics.
- Investigate wind flow and wind plant power curve variations over complex terrain via wind tunnel and numerical modeling.

The project scope includes:

- Generate real-time weather forecasts real time over a 4-km grid in both Northern and Southern California using the COAMPS meso-scale model.
- Develop and test wind energy forecast systems to provide forecasts for two “look-ahead” time horizons: (1) short-term forecasts of five-minute wind energy generation over a three-hour period to be issued every five minutes for the principal wind resource areas of the state (Solano, Altamont, Tehachapi, and San Geronio); and (2) intermediate-term forecasts of hourly wind generation over the a 48-hour period issued twice daily or every 12 hours for wind plants in each of the principal wind resource areas.
- Conduct numerical and wind tunnel modeling of wind flow and power density at each wind turbine location vs. wind speed at a reference meteorological tower to investigate the variation of wind flow and wind plant power curve with wind speed and direction, atmospheric stability, and other conditions.
- Generate the California Wind Generation Research Dataset (CARD), a database of daily forecasts of hourly wind generation at multiple elevations over 5-km grids in Northern and Southern California.

The project was conducted over the 18-month period, July 2004 through December 2005.

Project Participants

The project participants included the California Energy Commission as program manager, the Electric Power Research Institute (EPRI) as project manager, EPRI subcontractors AWS Truewind LLC, the University of California at Davis, and UC Davis subcontractor, Lawrence Livermore Laboratory; project advisors, California Independent System Operator, National Renewable Energy Laboratory, Southern California Edison, and five wind plant operators who together with CA ISO also provided wind resource and power data for their respective wind projects, Sacramento Municipal Power District, PPM/High Winds, PowerWorks, Oak Creek Energy Systems, and BMR/Mountain View 1 and 2.

The project consisted of six major tasks: Task 1: Project Review and Reporting; Task 2: Wind Resource Data Collection and Analysis; Task 3: Rapid-Update Wind Speed and Direction Forecast Algorithm; Task 4: Regional Short-Term Wind Energy Forecasting System Development and Testing; Task 5: Long-Term Wind Energy Forecasting System Development and Testing; and Task 6: Wind Tunnel Testing Coupled with Advanced Numerical Model Data.

Report Organization

The report consists of seven chapters, including Chapter 1, Introduction.

Chapter 2, Wind Energy Forecasting Integration into Electricity Grid Operations, addresses the next steps to implement the forecast algorithms and display the forecasts and related information in a format useful to the system operator.

Chapter 3, Next-Hour Regional Wind Energy Forecasting System Development and Testing, describes the conceptual design of a two-stage short-term forecasting algorithm based on artificial neural networks and application of a portion of the algorithm to generate five-minute forecasts over three hours for the principal California wind resource areas.

Chapter 4, Next-Day Wind Plant Energy Forecasting System Development and Testing, describes testing of various improvements in data, meso-scale models, and other features of the forecast system to assess the potential improvement in forecast performance vs. the earlier work and application of the resulting improved forecast system to five wind projects in the principal wind resource areas.

Chapter 5, Numerical Modeling of Wind Flow and Wind Plant Power Curves over Complex Terrain, describes the parallel efforts to evaluate wind flow and variation of wind plant power curves with wind speed, atmospheric, and other conditions over the complex terrain at Altamont Pass by AWS Truewind and University of California at Davis.

Chapter 6, High-Resolution Weather and Wind Forecasting, describes the application of the COAMPS meso-scale model to generate real-time daily weather and wind forecasts over 4-km grids in Northern and Southern California.

Chapter 7, Conclusions and Recommendations, identifies the key advances made in short- and intermediate-term forecasting, numerical modeling of wind flow over complex terrain, and the remaining issues that should be addressed in future work, and recommends specific topics for further investigation.

Chapter 8 presents the references.

2

WIND ENERGY FORECASTING INTEGRATION INTO ELECTRICITY GRID OPERATIONS

Additions of large blocks of intermittent wind generation to the electricity grid can affect the reliability, operation, and reserve requirements of the system and create challenges for the system operator. This section addresses the characteristics of wind generation; how wind generation affects electricity system operations; why wind energy forecasting is important for next-hour and next-day dispatching of system resources; and the content, format, and method of delivery of forecasts that would be most useful to system operators.

Characteristics of Wind Generation

Important characteristics of wind energy that affect system operation include:

- Commercial wind turbines employ mechanical and electrical components to convert wind energy to electricity and use induction generators and sophisticated electronic controls to produce 50 or 60 Hz power, while thermal generation uses fuel combustion to either generate steam and drive a single large steam turbine generator or drive a combustion turbine.
- A large wind-generation facility consists of a large number of relatively small generating units, which are connected together and act like a distributed generation system.
- The “prime mover” for a wind plant is, obviously, the wind. The characteristics are determined by Mother Nature, not plant operators or system dispatchers. A particular wind regime can be characterized quite well in a statistical sense but does not lend itself well to deterministic analyses. Consequently, wind plants cannot be reliably scheduled in advance, and the capacity value is typically zero to 15 percent of the rated capacity.
- The wind resource and the resulting output of a large wind plant fluctuate over nearly all time scales of potential interest for power system designers and operators: seconds, minutes, hours, days, seasons, and so on. Especially when coupled with the uncertainties mentioned above, such behavior challenges the validity of applying well-established methodologies and analytical approaches for power system engineering.

How Wind Generation Affects Electricity Grid Operations

As shown in Figures 2-1 and 2-2, wind generation exhibits both short- and longer-term fluctuations over periods from seconds to longer periods due the intermittent nature of wind.

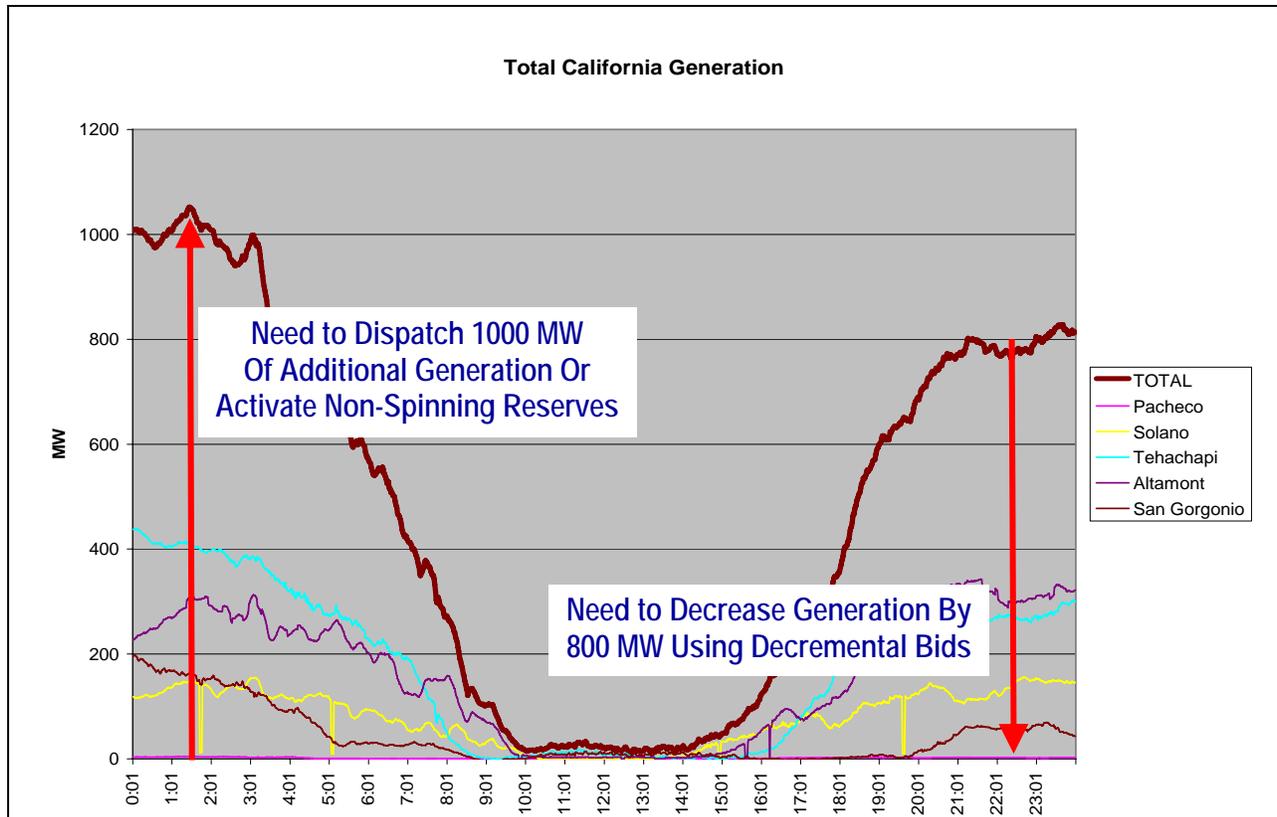


Figure 2-1 Typical Variation of Total and Regional One-Minute Wind Generation in California on a Summer Day (CA ISO, 2005)

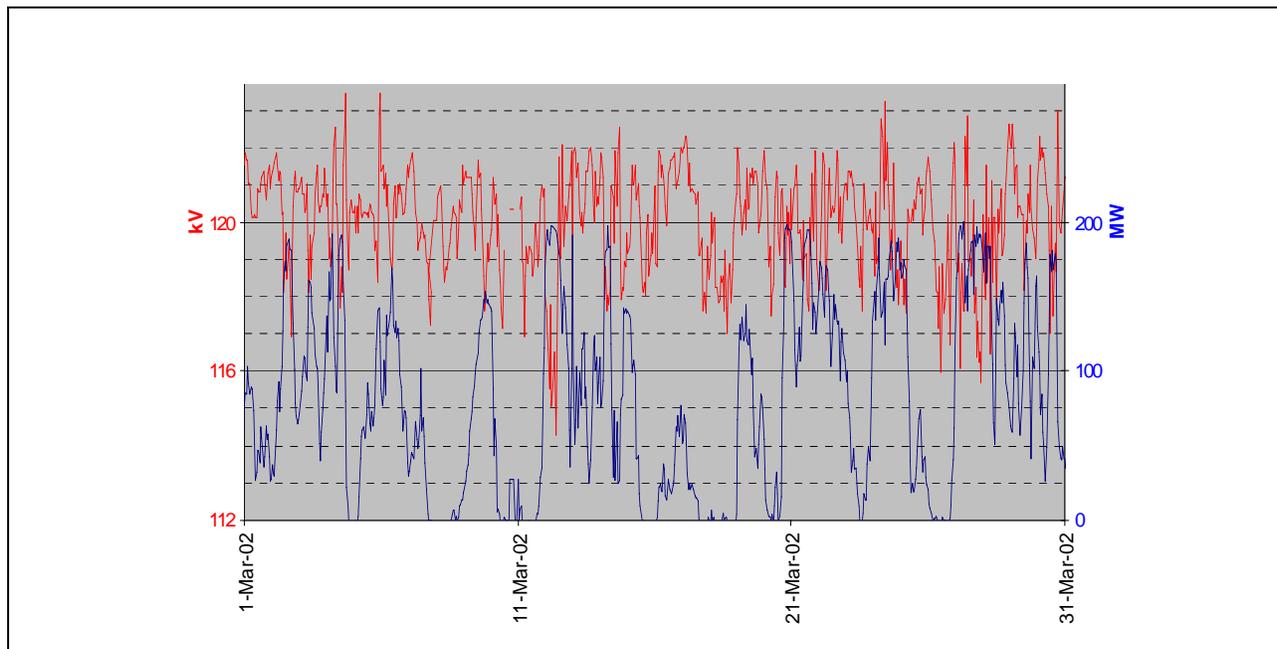


Figure 2-2 Buffalo Ridge 115-kV Bus Voltage in kV (Top Trace and Left Scale) and Transformer Output in MW (Bottom Trace and Right Scale) (NREL, 2002)

Figure 2-1 illustrates the fluctuations of one-minute wind power delivered to the California grid during a single 24-hour period on a summer day in California (CA ISO, 2005). The chart presents both aggregated wind power data for the state and data for each of the five wind resource areas (Solano, Altamont, Pacheco, Tehachapi, and San Geronio).

Similarly, Figure 2-2 illustrates the short-term fluctuations of both wind energy generation and 115-kV line voltage over a three-day period in March 2002 for 230 MW of wind generators in the Buffalo Ridge area of southwest Minnesota (EPRI, 2004). As the wind speed varied over a wide range, wind energy generation fluctuated between about 0 and 200 MW, and the line voltage varied between about 114 and 123 kV. Experience in Germany, where more than 15,000 MW of wind generation is in place, indicates that very large changes in wind output can occur in minutes as weather fronts pass through (E.ON Netz, 2004).

The California diurnal wind generation profile in Figure 2-1 is typical of a summer day in the areas affected by the coastal marine layer. Wind speed and generation build during the afternoon as the marine layer spreads from the high-pressure region over the cool Pacific Ocean to the low-pressure regions over the hot interior valleys. It then reaches a peak in early evening and begins to fall off during the early morning hours, reaching a minimum between about 10:00 a.m. and 2:00 p.m.

Unfortunately, the diurnal patterns of wind power and system load do not match well. This is especially true during the evening, when system load decreases while wind generation is reaching its peak. The system operator or balancing authority must reduce other generation to balance generation and load using decremental bids (800 MW in the example). In the morning, when load is building and wind generation diminishes to its minimum, the reverse situation exists. The system operator must dispatch additional generation or non-spinning reserve to balance the system (1000 MW in the example).

System Operators and Balancing Authorities

System operators are often single utilities responsible for operating the generation, transmission, and distribution system of single and sometimes multiple control areas. Balancing authorities like the California Independent System Operator and other regional transmission operators (RTOs) are responsible for balancing the generation and transmission resources for a group of control areas covering large regions, as illustrated in Figure 2-3.

Area Control Error (ACE)

As defined in Figure 2-3, the Area Control Error (ACE) is an algebraic function of the average deviations vs. schedule of generation, load, frequency, and net interchanges with other control areas. Each control area is responsible for maintaining ACE within two ranges, referred to as Control Performance Standards 1 and 2 (CPS1 and CPS2).

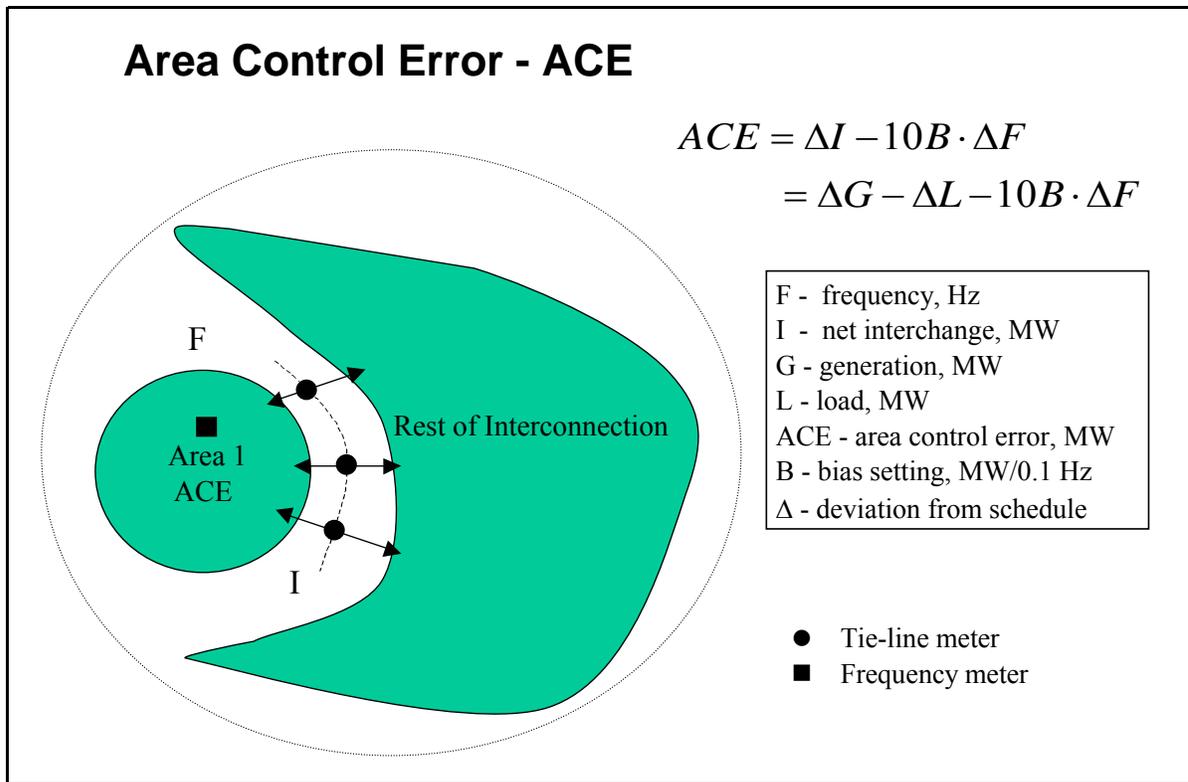


Figure 2-3 Relationship between an Individual Control Area and Rest of Interconnection and Definition of Area Control Error (CA ISO, 2005)

Control Area Objectives

As shown in Figure 2-4, the principal objectives of area and regional control operators are to operate transmission within thermal limits, maintain voltage within voltage stability limits, observe transient stability limits, and balance generation against load. In addition, the regional balancing authority balances the overall generation load, maintains scheduled interchanges, and supports interconnection frequency.

To meet these objectives, the operator is responsible for scheduling generation and transmission resources for both day-ahead and hour-ahead periods, load following, and system voltage and frequency regulation to balance generation and load and meet other objectives. Figure 2-5 illustrates the duration of system load vs. upward and downward regulation; day-ahead, hour-ahead, incremental and decremental scheduling; and five-minute dispatch of generation resources.

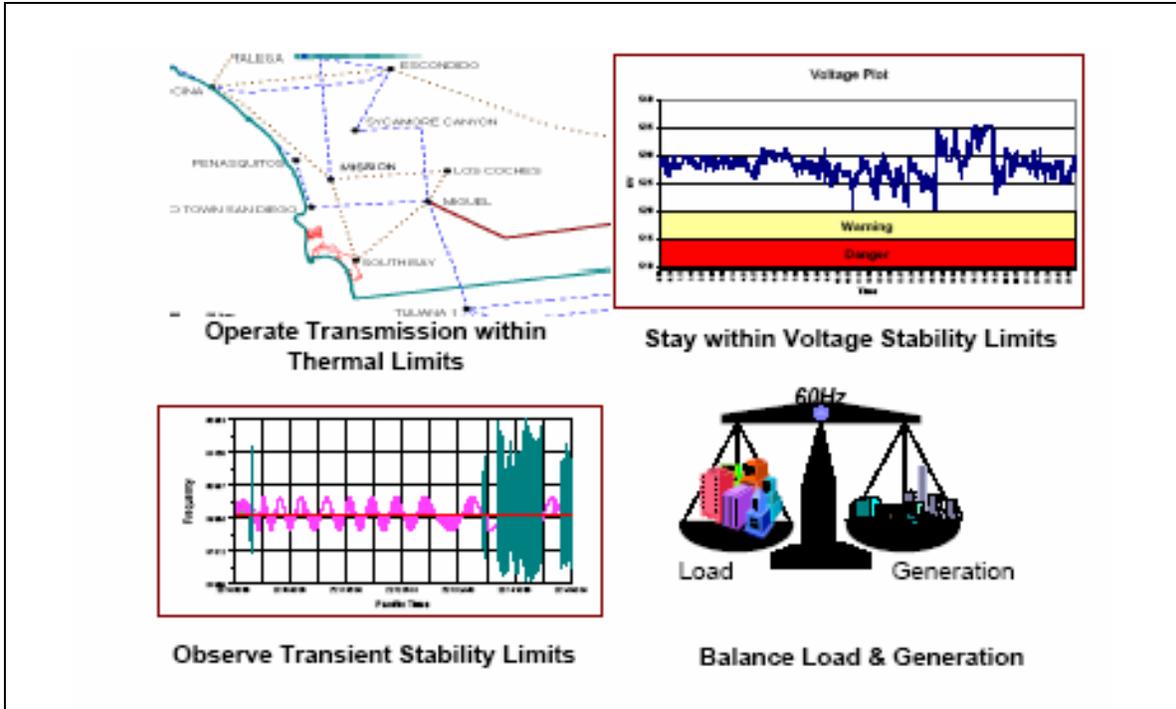


Figure 2-4 Control Area Objectives Focus on Balancing Load and Interchange vs. Generation (CA ISO, 2005)

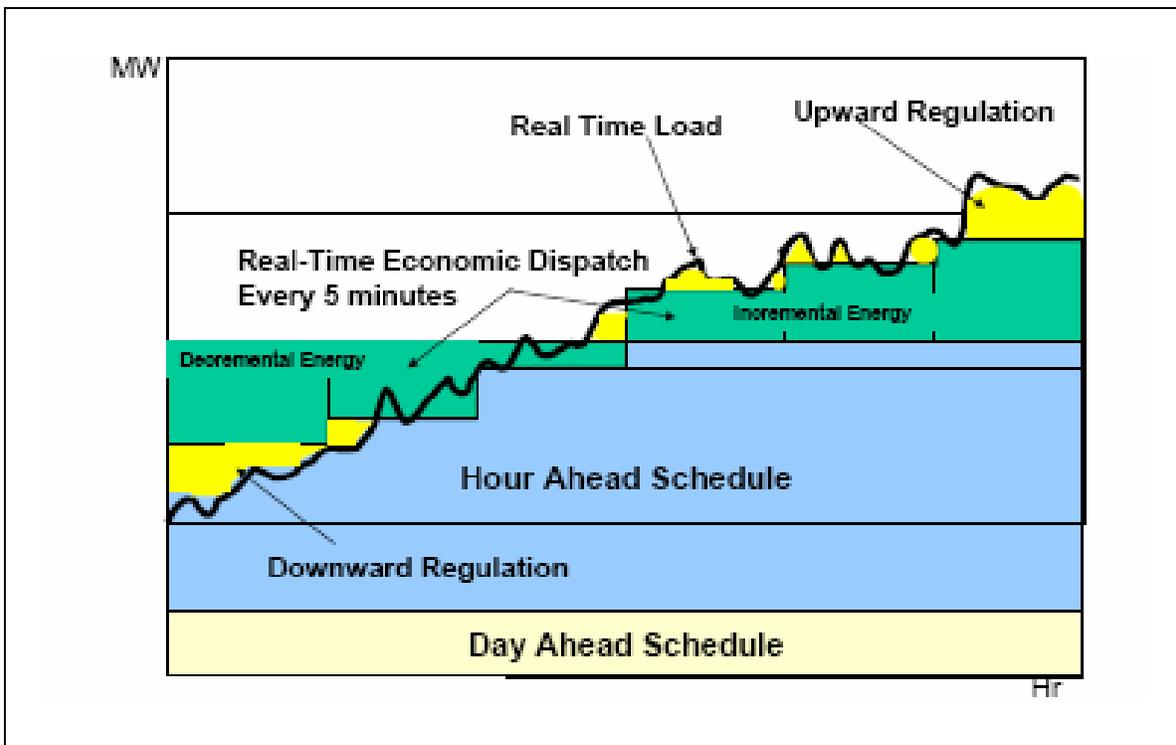


Figure 2-5 Balancing Authority Balances Generation Resources vs. Real-Time Load (CA ISO, 2005)

System Impacts and Challenges

The system impacts and operator challenges created by wind generation result from the intermittency and short- and longer-term fluctuations of wind generation, due to gusting winds, the resulting high ramp rates of wind generation, diurnal and seasonal wind speed and system load variations, and movement of weather systems.

The system impacts include:

- Unit commitments and scheduling.
- Voltage regulation/reactive power control.
- Reserve margins for security and reliability.
- Transmission bottlenecks during windy periods.
- Frequency control and regulating reserves.
- Load following and energy balance.

The magnitudes of the impacts can vary over wide ranges and depend on several important factors:

- Percentage penetration of rated wind capacity in the generation mix.
- Geographical dispersion of wind capacity.
- Diurnal and seasonal correlations between wind generation and system load.
- Penetrations and types of other generation resources in the mix.
- Presence of hydro, pumped storage hydro, and peaking capacity in the mix.
- Adequacy of transmission resources to transmit wind energy during periods of peak generation to the population centers.

Importance of Wind Energy Forecasting

Wind energy forecasting is one of several mitigation measures available to reduce the impacts of wind on power system operation and control (EPRI, 2004, 2005).

Other mitigation measures include:

- Power electronics and line compensation to absorb short-term fluctuations and control power factor.
- Integration with hydroelectric generation to absorb both short-term fluctuations and store off-peak wind energy generated at night when demand is low.
- Addition of storage and flow batteries, compressed air energy storage, pumped hydro, and other energy storage facilities.

- Transmission upgrades to relieve bottlenecks during windy periods; and wind energy forecasting.

In combination with load forecasting, wind energy forecasting can support optimal dispatching of intermediate and peaking generation, including hydro, fossil, and other dispatchable generating units; dispatching of transmission resources; scheduling next-hour and next-day wind energy deliveries to the grid; markets for green power and green certificate trading; and other uses. The ongoing CA ISO Participating Intermittent Resources Program (PIRP) provides hourly forecasts of next-hour and next-day wind generation to participating wind plant operators for use in scheduling next-hour wind energy deliveries to the California grid (CA ISO, 2005).

Balancing authorities like CA ISO need accurate next-hour and next-day forecasts of hourly wind generation to anticipate high ramp rates (CA ISO, 2005).

For the next-hour market, CA ISO uses five-hour forecasts to prepare accurate dispatch notices to send to quick-start generators and they need to receive the forecasts in time to send the notices at T-270 min. (270 minutes before the start time). In addition, the next-hour forecasts are due at T-75 min. (75 minutes before the hour) when the next-hour market closes.

For the next-day market, the forecasts are used to assess hourly energy production and anticipated hourly ramps. The wind energy forecast influences the procurement of ancillary services (regulation and operating reserves), preparation of accurate day-ahead generation dispatch notices, and effect of wind generation schedules on transmission congestion.

Forecasting Wind Generation Ramp Rates

Forecasting of wind generation ramp rates will be particularly important if the rated wind capacity in the Tehachapi Mountains increases by 4000 MW from about 670 MW to 4600MW in the future. For example, Figure 2-6 forecasts the daily variation of wind generation at Tehachapi, based on wind generation data for selected days during April 2005 (CA ISO, 2005). The daily variations cover the entire range from zero to 4600 MW during the 30-day period.

Very high positive and negative ramp rates occurred on several of the days. For example, Figure 2-7 summarizes the extreme range of hourly wind generation and ramp rates that would have occurred on April 8, 2005 with 4600 MW of wind capacity at Tehachapi. During the 24-hour period, wind generation varies between zero and 4500 MW, and the ramp rate exceeded minus 1000 MW/hr during two hours, and reached plus or minus 600 MW/hr during 11 of the 24 hours. The two minus-1000 MW/hr ramps occur at 7:00 and 8:00 AM, precisely when load is building rapidly.

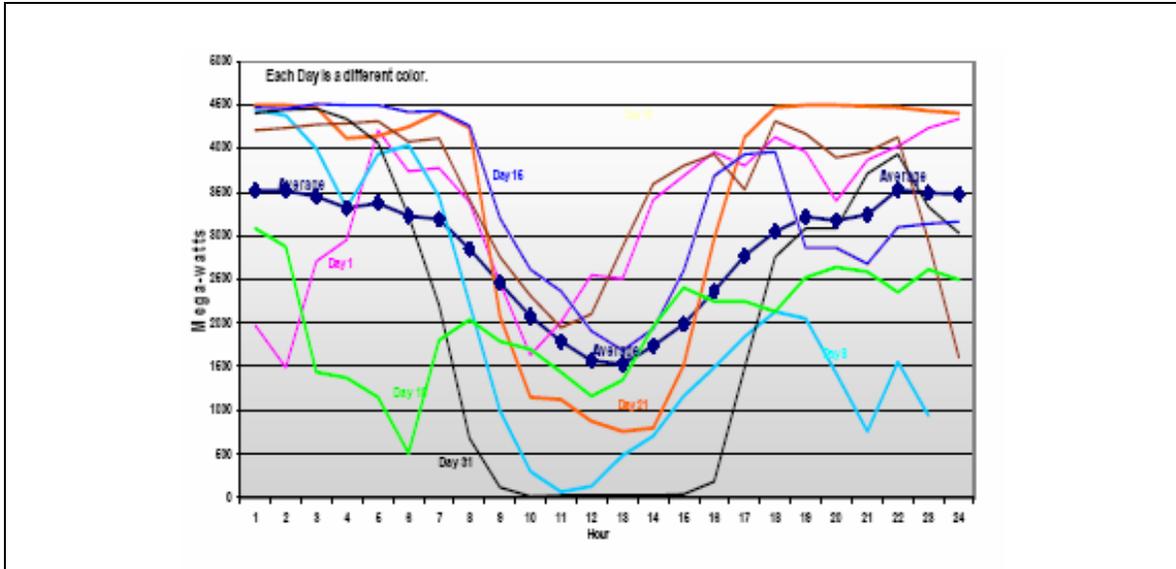


Figure 2-6 Forecast Daily Variation of Hourly Wind Generation at Tehachapi after Expansion to 4000 MW Rated Capacity. Based on April 2005 Data (CA ISO, 2005)

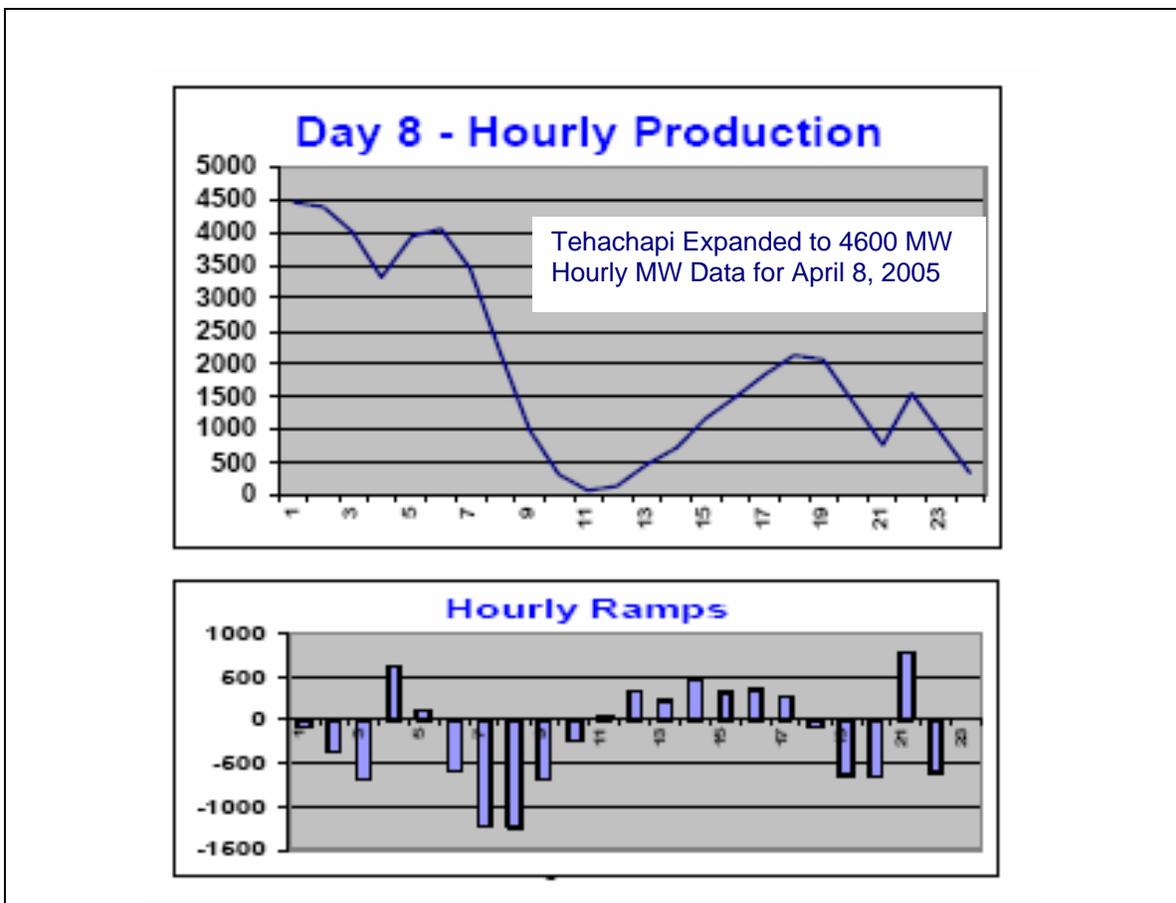


Figure 2-7 Estimated Hourly Wind Generation (MW) and Ramp Rates (MW/hr) after Tehachapi Expansion to 4600 MW. Based on Data for April 8, 2005 (CA ISO, 2005).

Because of the above and other internal studies, CA ISO has concluded that it is important to accelerate development and validation of wind energy forecast tools now so that they are ready to use when the wind generation capacity in the state grows to a level that reach two to four times the current 2200-MW level (CA ISO, 2005).

Customizing Wind Energy Forecasts for System Operators

Even the most accurate next-hour and next-day wind energy forecasts for a region will be almost useless to the system operator if they are delivered, for example, as simple tables of wind speed and wind energy generation forecasts vs. time without additional information that directly addresses the needs of the operator with regard to content, format, and method of delivery.

For example, other forecast information and data may at times be of greater interest to the system operator than the actual wind speed and energy forecast, such as the forecast ramp rate of regional wind generation, the impact of forecast errors on control area CPS1 and CPS2 compliance, and the cumulative monthly imbalance of scheduled vs. delivered wind energy.

Figure 2-8 presents an example of a web-page display developed by AWS Truewind during the development of the intermediate-term forecasting system described in Section 4, Next-Day Forecast System Development and Testing. It displays the most recent 48-hour forecast of hourly wind generation, in this case for a specific wind project, together with the observed wind generation for the site through the most recent hour, plus forecast performance metrics for various time periods through the present.

Examples of other information that could be provided in such a graphical display include:

- Relative confidence in forecast accuracy based on weather conditions.
- Range of uncertainty of forecast vs. time superimposed on forecast chart.
- Archived charts comparing forecast and observed wind speeds and energy generation.
- Archived forecast performance metrics vs. observed data, for example, mean error (bias, ME), mean absolute error (MAE), mean square error (MSE), and skill scores (Skill) over different time periods to present (day, week, month, season, and year).
- Customized forecast and other information, such as next-hour and next-day ramp rates of regional wind generation; estimated impacts of forecast errors on CPS1 and CPS2 area control error requirements, and the cumulative monthly imbalance between scheduled and delivered wind energy.

For the system operator to “buy in” and trust the forecasts, the operator should also be involved in the system customization and initial trial period. That way the forecast information provided and needs of the operator are more likely to merge.

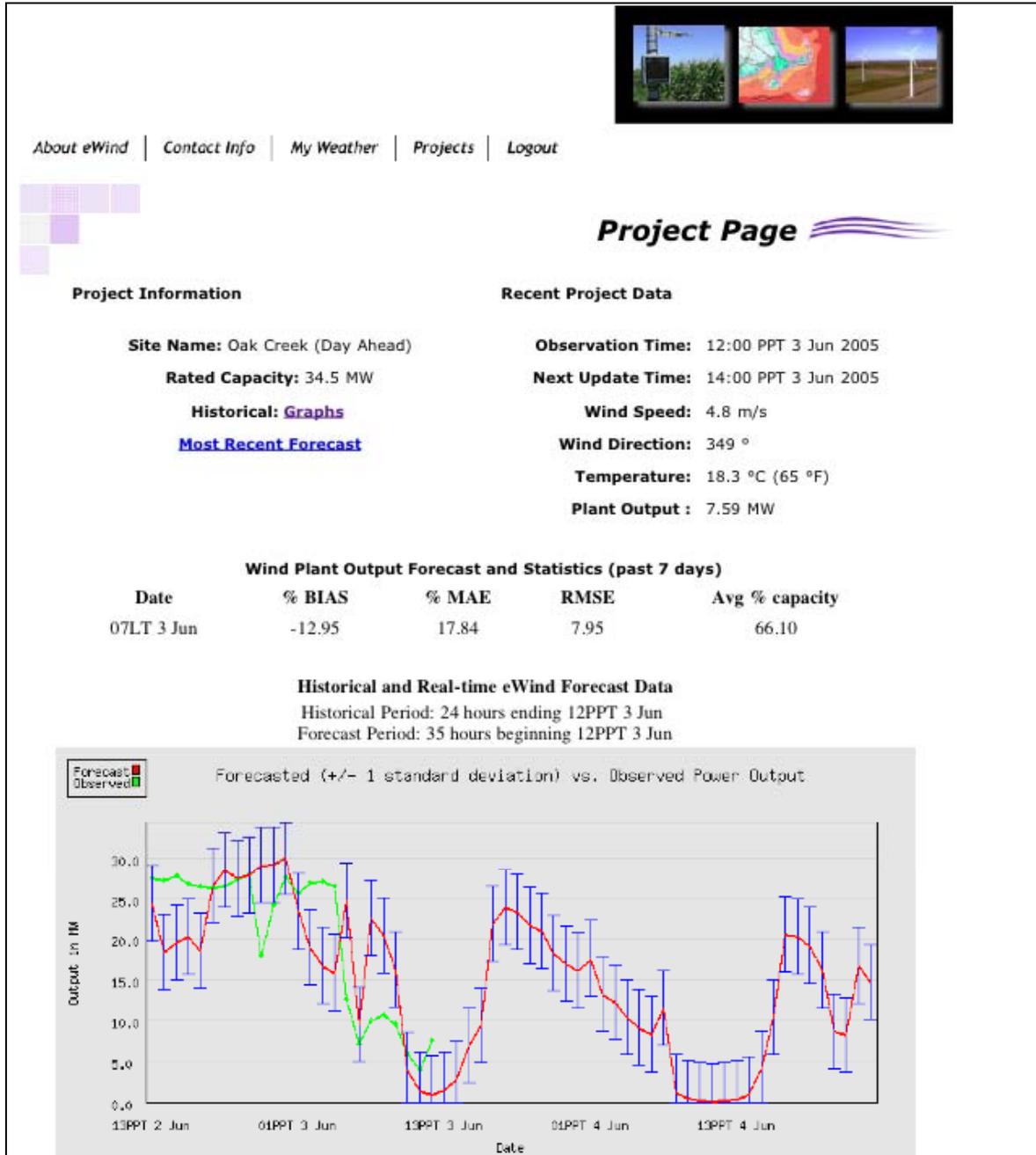


Figure 2-8 Example Display of Real-Time Forecasts for Oak Creek Energy Systems Wind Project in Tehachapi, California (AWS Truwind, 2005).

Conclusions

Wind energy forecasting will become especially important to control area operators and balancing authorities like the CA ISO in the future as wind generation concentrated in single regions reaches thousands of megawatts.

Accurate forecasting systems are needed to generate both next-hour and next-day and longer forecasts for several reasons:

- Provide early warning of high hourly ramp rates of wind generation for planning both same- and next-day dispatching of generation and transmission resources.
- Support markets for ancillary services to support intermittent wind generation.
- Support issuing accurate dispatch notices to quick-start generators
- Support scheduling of next-hour and next-day deliveries of wind energy to the system.

Development and validation of the forecasting algorithms needed to meet the needs of the control area operators and balancing authorities should begin now.

The system operators should be actively involved in the development and testing of the algorithms to ensure that their needs are met.

3

NEXT-HOUR REGIONAL WIND ENERGY FORECASTING SYSTEM DEVELOPMENT AND TESTING

Introduction

The overall goal was to produce a prototype of an integrated wind generation forecasting system that can provide short-term power production forecasts for each of the five wind resource zones in California. The specific functional objective for this short-term forecast system is that it be capable of producing forecasts of the regional power production in five-minute intervals for the next three hours after forecast delivery. Thus, each forecast should consist of predictions for the next 36 five-minute intervals. Furthermore, the system should be capable of producing an update every five minutes.

The short-term forecast system is intended to provide a low-cost regional wind generation forecasting tool to be used by control area operators and scheduling coordinators at the CA ISO for more accurate and economical supplemental energy scheduling, real-time dispatch, load following, and AGC control. The ultimate objective is to augment traditional short-term statistical forecasting techniques to include a consideration of local and regional atmospheric predictors of short-term changes in weather conditions. The intention is to develop a system that will generate regional power production forecasts even in the absence of site-specific monitored data by using all of the available local and regional forecasted and observed weather data.

The work on this task was divided into two phases or subtasks. The objective of the first phase was to formulate a design for a short-term forecast system based on a review of a variety of forecasting methods and AWS Truewind's experience with operational forecast systems and ongoing forecast system research and development for longer-term forecasts (hours and days ahead). The goal of the second phase was to implement at least a portion of the system designed in Phase 1 and produce an initial assessment of its performance using real-time wind power production data from the CA ISO system.

Forecast System Formulation

It was recognized that zero- to three-hour forecasts in five-minute intervals will necessarily rely heavily upon statistical time series prediction tools. This is because it is not possible to gather sufficient data to initialize a high-resolution physics-based atmospheric model on the scale required to make zero- to three-hour forecasts in five-minute intervals, nor is it possible to execute such models quickly enough to make their output useful in the forecast process.

However, physics-based models that run at moderately high resolution every few hours may provide some useful trend information for forecasts for the zero- to three-hour period. Therefore, the formulation of the forecast system began with a focus on statistical methods for time series prediction that could be employed in a short-term forecast system. The methods that were reviewed included a number of classical time series prediction methods as well as newer techniques based on recent advances in learning theory. A summary of the methods that were considered for use in the short-term forecasting system is presented below. The presentation begins with the classical techniques and progresses to the newer and more sophisticated techniques. After the summary of the forecast methods, a proposed design for a complete short-term forecasting system is presented.

Alternative Forecast Methods

Autoregressive Models

The most basic approach to the prediction of a future value from a time series of data is the autoregressive (AR) model. In this approach the future value of a time series (X_n) is predicted from the values at the current time ($n-1$) and previous times ($n-2$ through $n-p$) by the relationship:

$$X_n = c + \sum_{i=1}^p a_i X_{n-i} \quad (\text{Eqn 3-1})$$

where “c” and all p of the “a” variables are empirical parameters that are estimated from the training data. An autoregressive model is simply a linear regression of the current value of the series against one or more prior values of the series. The value of p is called the order of the AR model. The parameters of the AR model can be estimated in a variety of ways. The most popular method is the standard linear least squares approach that is used in linear regression. The AR model can be expanded by adding time series of other relevant variables.

Moving Average Models

Another popular approach to the prediction of future values from a time series of data is the moving average (MA) model. In this approach, the prediction is based upon a linear combination of past forecast errors rather than a linear combination of past values. In this approach the future value of a time series (X_n) is predicted from a linear combination of error values according to the relationship:

$$X_n = c + \sum_{i=1}^q b_i \varepsilon_{n-i} \quad (\text{Eqn 3-2})$$

where “c” and all q of the “b” variables are empirical parameters that are estimated from the training data and ε is the forecast error at each previous interval in the time series. The parameter q determines how many error terms are included in the model and is often called the order of the MA model.

ARIMA (Box-Jenkins) Models

Box and Jenkins (1976) developed an approach that combined the autoregressive and moving average approaches into a method called the AutoRegressive Integrated Moving-Average (ARIMA) model. Since the ARIMA approach was first popularized by Box and Jenkins, ARIMA models are often referred to as Box-Jenkins models. An ARIMA model predicts a value in a time series as a linear combination of its own past values, past errors (sometimes called shocks), and current and past values of other time series. When an ARIMA model includes other time series as input variables, the model is sometimes referred to as an ARIMAX model. The prediction equation for an ARIMA model for a time series of X values is:

$$X_n = c + \sum_{i=1}^p a_i X_{n-i} + \sum_{i=1}^q b_i \varepsilon_{n-i} \quad (\text{Eqn 3-3})$$

where a, b and c are prediction model parameters to be determined from the training data, X_{n-i} are the current (n-1) and past (< n-1) values of the time series and ε_{n-i} are the current and past values of the forecast error.

The main advantage of the ARIMA approach over other classical methods is the small number of values necessary for prediction calculation, its rapid adaptability to changing conditions, and the good performance with predicting time series that have a significant deterministic component. Its disadvantage is that it has difficulty with time series that have a significant stochastic (non-deterministic approach) component.

Artificial Neural Networks (ANN)

The most popular method currently used for predicting a time series is probably the artificial neural network (ANN). The ANN enables us to describe non-linear processes created by a set of complicated phenomena. The fundamental building block of a neural network is a computational element called a neuron. A neuron receives an input and processes it to generate an output. A neural network consists of a set of neurons that are logically arranged.

Figure 3-1 presents a schematic of a simple neural network. Each of the circles in the diagram represents a neuron. The output of a neuron is a function of the net input according to the formula:

$$\text{output} = f(\text{net input}) \quad (\text{Eqn 3-4})$$

The function that translates the net input into the output is called the activation function. Many different activation functions have been used in ANN applications. The majority of current ANN applications employ a sigmoid (S-shaped) activation function. A sigmoid function may be loosely defined as a continuous, real-valued function whose input domain is the entire set of real numbers, whose derivative is always positive and whose range is bounded. The most commonly-used sigmoid function is the logistic function. Another popular sigmoid function used in neural network applications is the hyperbolic tangent function. In practice, the

performance of a neural network prediction algorithm, but the rate of learning (training time) can be significantly affected by the choice of the activation function.

The net input to the activation function is a weighted sum of all of the input values plus an offset constant. Thus, if our n inputs to a neuron are x_i with i ranging from 1 to n , the net input is calculated from

$$netinput = \sum_{i=1}^n x_i w_i + w_0 \quad (\text{Eqn 3-5})$$

where the w_i and w_0 parameters are the weights for each input. The operational characteristics of a neuron are primarily controlled by these weights and not by the characteristics of the activation function as long as some basic criteria are met for the activation function.

In the neural network configuration shown in Figure 3-1, each hidden layer neuron receives three inputs. Therefore, the output of each hidden layer neuron (HL_n) is

$$HL_n = f(w_{in0} + INP_1 w_{in1} + INP_2 w_{in2} + INP_3 w_{in3}) \quad (\text{Eqn 3-6})$$

where INP_i are the input layer values, the w_i are the associated weights and $f()$ is the activation function. Each of the outputs from the hidden layer neurons is then combined in an analogous weighted sum to compute the final output value from

$$NetworkOutput = w_{ho} + HL_1 w_{hl1} + HL_2 w_{hl2} + HL_3 w_{hl3} \quad (\text{Eqn 3-7})$$

where the w_{hli} are another set of weights that convert the output from the hidden layer neurons to the output value for the network. It is easy to see from this example that there are a large number of weight parameters in even a simple neural network with one hidden layer and three hidden layer neurons. In this case, there are a total of 16 weights: four to calculate the net input to each of the three hidden layer neurons and four more to calculate the output value from the output of the three hidden layer neurons. The number of weight parameters will significantly increase if additional hidden layers, hidden layer neurons or inputs are added to the network.

Support Vector Regression (SVR)

One of the newest tools for creating functions from a set of labeled training data is a Support Vector Machine (SVM). The SVM concept was first developed by Vapnik and Chervonenkis (1974) to address classification problems. The unique aspect of the support vector classification function is that it only depends on a subset of the training data because the cost function for building the classification function does not care about training points that lie beyond a prescribed margin. For classification, SVMs operate by finding a hypersurface in the space of possible inputs. This hypersurface will attempt to split the positive examples from the negative examples. The split will be chosen to have the largest distance from the hypersurface to the nearest of the positive and negative data points.

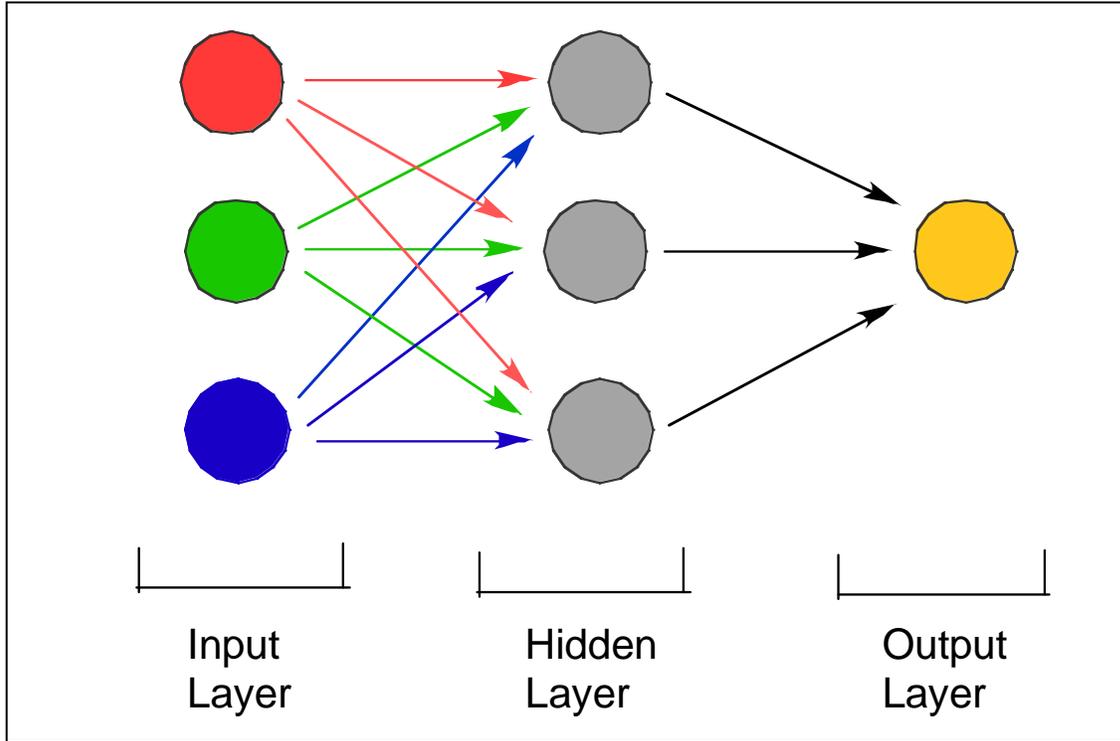


Figure 3-1 A schematic depiction of the flow of data through an artificial neural network (ANN) scheme. Each circle represents a neuron (or a computational element).

A version of SVM for regression was first proposed by Vapnik, Golowich, and Smola (1997). This method was given the name Support Vector Regression (SVR). Analogously, the regression function produced by SVR depends only on a subset of the training data, because the cost function for building the model ignores any training data that is close (within a threshold ϵ) to the regression function.

To understand the unique attributes of SVR, it is useful to compare SVR to ordinary least squares linear regression (OLSLR). However, it should be noted that SVR can be extended to the case of non-linear regression and the linear case is used here only as an illustrative example. It is useful to recall that OLSLR builds a regression function by minimizing the cost function

$$Cost\ Function = (Y_p - Y_i)^2 = (wX_i + b - Y_i)^2 \quad (Eqn\ 3-8)$$

over all variables in the training set. In this function $Y_p = wX + b$ is the prediction of the variable Y by the resulting regression equation. The regression parameters w and b are selected to minimize the cost function when summed over all members of the training set. SVR, on the other hand, specifies that the regression parameters w and b should be selected to meet the following criteria:

$$\text{minimize } \|w\|^2$$

subject to

$$y - wx - b \leq \varepsilon$$

$$wx + b - y \leq \varepsilon \quad (\text{Eqn 3-9})$$

These criteria only demand that the regression function produce values within +/- ε of the actual data values y_i . Thus, the algorithm to find the regression parameter values assigns no penalty to any points that are within +/- ε of the regression function. A penalty is assigned only to those points that lie outside of the +/- ε band around the function. The $\|w\|^2$ criterion serves to minimize the slope of the line so that resulting function maximizes its “flatness”. This criteria is related to what can be termed a “learning capacity control.” This addresses an issue that arises in many different approaches to estimating a function from a set of training data including the ANN procedure. Some classes of functions have a high capacity to learn (or fit the empirical data) due to their structure and the number of associated parameters. If the amount of independent data is small, there is a high risk that such a function will overlearn (or “overfit”) the data and result in poor forecast performance even though it fit the training data very well. The SVR approach addresses this problem by invoking a learning capacity control criteria. In the linear regression case, this is invoked by minimizing the slope of the regression line subject to the other constraints.

The SVR approach has a number of desirable attributes and has been shown to significantly outperform other time series prediction approaches in some applications. However, the approach is quite new, and its performance relative to other methods has not been thoroughly documented in a broad range of applications. A version of SVR software was tested on a set of power production data in this project. Its performance was quite erratic and it occasionally had problems converging to a solution. Thus, it was decided that it was best not to utilize SVR at this time.

Short-Term Forecast System Design

After reviewing the multitude of statistical tools that are available to serve as the core forecast method for the short-term forecast system, the Artificial Neural Network (ANN) approach was selected as the best choice to serve as the core of the short-term forecast system. The ANN method has an advantage over the traditional AR, MA, or ARIMA approaches because it can model more complex, non-linear relationships and can be easily extended to incorporate a variety of other variables with relevant predictive information. SVR and other advanced tools can also model very complex, non-linear, multi-parameter relationships. It may ultimately be recognized that some of these tools have advantages over the ANN approach for the short-term power production forecasting application. However, most of these methods, especially SVR, are still fairly early in their development process and the documentation and experience with these methods is much more limited than with the ANN approach. The ANN method has been applied to a variety of diverse applications with considerable success. As a result, the strengths and weaknesses of the ANN approach have been widely documented, which provides a solid basis for its application to predicting short-term power production.

As noted previously, the design objective for the short-term forecast system is to go beyond the basic use of power production time series information from the individual wind resource regions, subregions or wind plants. In addition to the power production time series data, it is envisioned that the following information will also be used in the forecast system: (1) time series information of meteorological parameters from meteorological towers operated by wind generators or other members of the wind energy community; (2) meteorological data from surface weather observing sites operated by the National Weather Service and other organizations; (3) meteorological data from remote sensing systems such as wind profilers, Doppler radars and satellite-based sensors; and (4) short-term forecast data from high resolution atmospheric physics-based models run in a rapid update cycle mode (assimilation of new data and execution of short-term forecasts every few hours or possibly every hour).

Figure 3-2 presents a schematic of the initial forecast system design schematic created to address these design objectives. The circles represent input or output data, while the rectangles depict algorithms (numerical models) that operate on the data. The system employs a variety of power production and meteorological data.

The four fundamental types of input data are depicted by the circles at the top of the diagram. The leftmost circle represents the time series of power production data from a wind resource region, a subregion or an individual wind plant. The next circle to the right represents the time series of meteorological data gathered by the sensors on a meteorological tower located at the wind plants or sites selected by members of the wind energy community to provide data for wind power production forecasting or wind resource assessment. For example, one participant in the California electric grid system maintains a network of 12 anemometers for its own wind power production forecasting applications. These data and similar datasets could be used in the proposed short-term forecast system if the parties that own the sensors and the data agree to provide the data for this application. The third circle to the right represents the surface weather data available from a variety of other sources. The principal source is the U.S. National Weather Service, but some additional data are potentially available from other public and private sources (for example, CalTrans, Air Quality Management Districts, and others). The rightmost circle represents the pool of meteorological data available from the diverse set of remote sensors that routinely gather information about the atmosphere. This includes ground-based remote sensing systems such as wind profilers, temperature sounders and Doppler radars, as well as space-based systems that provide vertical profiles of temperature and moisture as well as some limited wind data.

The proposed forecast system is designed to use these data sources to produce three somewhat independent short-term predictions of the power production. A separate neural network is then trained to ingest these three predictions and produce an optimal final power production prediction. The three different forecast production subsystems are represented by the three parallel vertical sequences in the schematic diagram.

The leftmost subsystem is fundamentally an autoregressive time series prediction scheme. It uses an ANN to exploit the structure contained in the power production time series to directly generate a power production forecast.

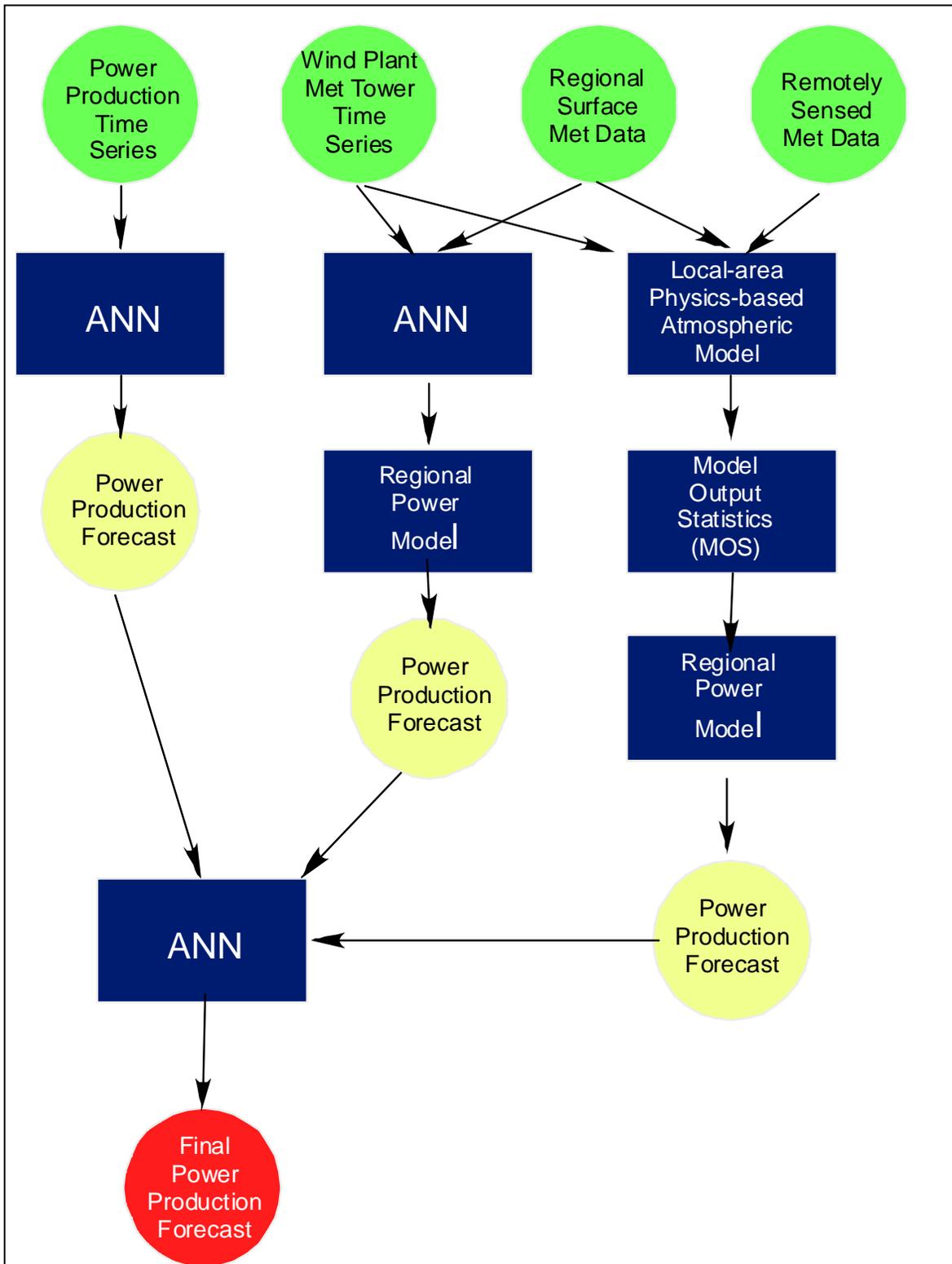


Figure 3-2 A schematic depiction of the proposed short-term (0 to 3 hour) forecast system.

The middle subsystem is also fundamentally a time series prediction algorithm, but it is much more complex. In this approach, the time series of meteorological data in the region is used to generate a wind forecast for the region. This wind forecast is then converted to a power production prediction through the use of a Regional Power Production Model (RPPM). The RPPM is a statistical relationship (a regional scale power curve in a sense) between the regional winds and possibly other atmospheric variables such as air density and the concurrent power production.

The third forecast method is based upon the use of a high-resolution physics-based model executed in a rapid update cycle mode. In this mode, the model frequently (for example, every hour or two) ingests all of the regional and local in situ and remotely sensed meteorological data to create a new initialization state for a short-term three-dimensional simulation of the atmosphere near the region to be forecasted. The grid point output from this simulation is used as input to a Model Output Statistics (MOS) procedure that adjusts the raw model predictions to account for systematic errors in the simulations due to processes that are too small to be seen on the model's grid or limitations in the model's formulation or input data. The output from the MOS procedure is a forecast of the meteorological variables that are needed for input into the RPPM. The RPPM is then used to convert the forecast of the meteorological variables to forecasts of the power production.

The final step in the forecast process is to feed the three quasi-independent predictions to an ANN scheme, which combines them into a single final power production forecast. The ANN in the final step will weight the three forecasts in a way that is dictated by their performance. This may result in selection of one of the three forecasts in certain situations and a blend of two or all three of the forecasts in others.

The use of three separate forecast subsystems has a number of advantages. First, it provides a backup capability in case one method should fail because data are not available or are corrupted due to a software failure. Second, the three systems take advantage of different data and modeling approaches and will likely have substantially different performance characteristics, which will likely enable specific methods to outperform the others in certain classes of situations.

For example, forecasts that are predominantly based on time series information will likely lack the ability to anticipate large changes in power production due to rapidly changing weather conditions, since there likely will be no signal of the upcoming large change in the recent history of the power production. However, a physics-based model initialized with regional and local meteorological data may exhibit considerable skill in forecasting major weather events. On the other hand, the physics-based models may find it difficult to forecast routine subtle changes in winds and therefore power production, since they may not provide the necessary spatial resolution of the computational grid or sufficiently detailed three-dimensional input data to accurately specify the initial conditions for the model. In those cases, which occur frequently, a time series prediction is likely to be superior.

A third advantage is that the existence of three preliminary forecasts provides an opportunity to employ an ensemble forecast approach. It has been demonstrated in other meteorological

applications that the composite of an ensemble of forecasts often statistically outperforms each of the individual members of the ensemble. In addition, the availability of an ensemble of forecasts can provide an indication of the confidence that should be placed in the forecast. If all of the members of the ensemble are in close agreement, then the user is usually justified in placing more confidence in the forecast. However, if the forecasts by the ensemble members are widely separated, the uncertainty in the forecast is typically higher.

The next step in the development is to implement and test the system. Unfortunately, the implementation of the entire system was beyond the scope of this project. In addition, a complete input dataset was not available to test the system even if it had been fully implemented. Therefore, only one of the three prediction subsystems was implemented and tested in this project. The method that was tested was the leftmost column in Figure 3-2, which applied ANN to the regional power production time series data provided by the CA ISO.

Regional Power Production Data

The CA ISO provided a regional wind power production time-series dataset consisting of one-minute aggregated wind power production and capacity data for five wind resource regions and for all of 2004. The regions were Altamont, Solano, Pacheco, Tehachapi, and San Geronio. Data were missing for the following time periods: (1) 0200 PST to 0255 PST on 4 April; (2) 0000 PDT 16 May to 0005 PDT 3 June; and (3) 2300 PST 31 October to 0000 PST 1 November. The dataset was used for all of the forecast performance experiments and evaluation that were conducted under this task.

Table 3-1 shows an example of the raw one-minute data. Several issues were noted in the initial examination of the raw one-minute data. One issue was that the actual production values for the Solano County region were always equal to or greater than the reported capacity values. Thus, according to the dataset, the Solano County wind plants always produced at 100% or more of the reported capacity for all of 2004. Another issue was that no capacity information was reported for the Pacheco region. This made the Pacheco data unusable, since the capacity data are crucial to the forecast procedure. No installed capacity information was provided with the dataset. Therefore the “installed capacity” was assumed to be the maximum availability reported for a five-minute interval during the year. These values were: San Geronio 556.71 MW, Tehachapi 659.71 MW, Altamont 827.3 MW, and Solano 171.4 MW.

The raw one-minute data were converted to five-minute data by AWS Truewind. The five-minute data intervals were synchronized with the clock hour so that a five-minute interval began and ended on each clock hour. The conversion to five-minute intervals was done by averaging all of the valid one-minute data within each five-minute interval. Any one-minute interval for which data were missing was excluded from the averaging process. If there were no valid one-minute data within a five-minute averaging interval, the five-minute interval was assigned a missing data flag. The five-minute data were used for all of the forecast experiments.

Table 3-1 An example of the raw one-minute regional power production data provided by the CA ISO for 2004.

Date/Time	Altamont		Tehachapi		San Gorgonio		Pacheco		Solano		Total EMS	
	Actual	Zone Availability	Actual	Zone Availability	Actual	Zone Availability	Actual	Zone Availability	Actual	Zone Availability	Actual	System Availability
7/6/04 0:01	270.27	608.10	153.95	316.40	238.00	554.46	9.60		166.49	152.10	919.35	921.98
7/6/04 0:02	271.24	608.10	153.81	316.40	235.07	510.06	9.60		167.87	153.48	918.63	925.15
7/6/04 0:03	268.87	608.10	149.82	316.40	229.77	554.46	9.60		168.27	153.87	907.37	917.25
7/6/04 0:04	267.83	608.10	151.62	316.40	227.67	554.46	9.60		168.39	154.00	906.15	907.58
7/6/04 0:05	269.09	608.10	159.00	316.40	222.09	554.46	9.60		168.43	154.03	909.24	912.68
7/6/04 0:06	269.32	608.10	155.71	316.40	220.48	554.46	9.60		168.46	154.06	904.59	910.26
7/6/04 0:07	270.06	608.10	150.04	316.40	222.66	554.46	9.60		168.48	154.09	901.88	903.98
7/6/04 0:08	269.34	608.10	151.63	316.40	220.63	554.46	9.60		168.51	154.12	900.75	906.87
7/6/04 0:09	270.13	608.10	150.34	316.40	219.81	554.46	9.60		168.54	154.15	899.45	904.77
7/6/04 0:10	272.29	608.10	149.29	316.40	224.21	554.46	9.60		168.57	154.18	905.00	904.55

Temporal Variability of Power Production

It is important to understand the temporal variability characteristics of any variable that is to be forecasted. Therefore, before the forecast performance experiments were conducted the five-minute regional power production data for 2004 were analyzed to understand the time scales and amplitudes of the temporal variability of the power production.

One standard approach to the analysis of the temporal variability of a parameter is the construction of a lagged autocorrelation plot. This was done for four of the five regions (excluding Pacheco) for which data were supplied. The lagged autocorrelation values represent the correlation between the current value of the power production and the value at a previous time offset from the current time by the lag period. A high autocorrelation coefficient for a particular time scale indicates that there often are trends in the data on that time scale.

Figure 3-3 shows a plot of the lagged autocorrelation values for each region. This depiction reveals that the autocorrelation is high for all four regions and drops off very slowly with time. However, the rate of decrease varies somewhat between the regions. In general, the autocorrelation decreases more rapidly with time for the two Northern California regions than for the two Southern California regions. This suggests that one- to three-hour trends in the data are more prominent in the Southern California regions than they are in the Northern California regions. The autocorrelation data also indicated that there were substantial seasonal differences in each of the regions.

Figures 3-4 and 3-5 show two examples of the seasonal variation in the autocorrelation structure. These charts show the lagged autocorrelation for the six cold season (October to March) and the six warm season (April to September) months for the Altamont (Northern California) and San Geronio (Southern California) regions. The seasonal variation is quite different in these two regions. In the Altamont region (Figure 3-3), the lagged autocorrelations decrease much more rapidly over the zero to three hour period during the cold season than they do during the warm season. However, after approximately three hours, the warm season autocorrelations decrease more rapidly than do during the cold season. As a result, the cold and warm season autocorrelations gradually approach each other during the three- to eight-hour period. After a lag period of approximately eight hours, the autocorrelation is approximately the same in both seasons. In contrast, the lagged autocorrelations decrease at approximately the same rate during the zero- to three hour period for both seasons for the San Geronio region. After three hours, the autocorrelation decreases much more rapidly in the warm season than in the cold season. As a result, the autocorrelation for the warm season during the eight- to 12-hour period is much lower than it is in the cold season. These results indicate that there are substantial differences in the temporal variability characteristics of the power production from region to region and from season to season within a region. This suggests that different forecast system configurations (for example, different predictors) may be necessary to achieve optimal performance in different regions and seasons.

Another tool that is useful in analyzing the temporal variability of a parameter is wavelet analysis. An overview of wavelet analysis is presented in Torrence and Compo (1998). Wavelet

analysis decomposes a time series into time-frequency space. This enables one to determine both the dominant modes of variability and how those modes vary in time. In essence, a wavelet analysis transforms a one-dimensional time series into a two-dimensional time-frequency image.

Figures 3-6 and 3-7 show wavelet power (the square of the wavelet amplitude) spectrum plots of the Tehachapi data for December and July 2004. The highest wavelet power during December is associated with the two- (2880 minutes) to four-day (5760 minutes) time scales and that wavelet power is highly variable over time. In fact, much of it is concentrated into two time intervals – one early in the month and one later in the month - that correspond to stormy periods. In contrast, the July analysis (Figure 3-7) indicates that there is substantial power for wave periods of about one day (1440 minutes). One possibility for future forecast systems is to use tools such as the ANN to predict the power production in frequency space and then integrate over all frequencies to obtain a prediction of the total power production. This would enable different predictors (or even different methods) to be used to make forecasts for different portions of the frequency spectrum.

Initial Forecast Performance Evaluation

This section presents the results from a series of forecast experiments designed to (1) evaluate the sensitivity of the forecast performance to a few of the most important configuration options in the ANN-based forecast system; and (2) assess the general level of performance that can be obtained from an ANN-based zero- to three-hour-ahead forecast system for the prediction of the regional wind power production within the CA ISO system.

The first objective was addressed by a series of three forecast experiments. The first series of experiments tested the impact of variation in the number and type of inputs to the ANN scheme. The second series evaluated the sensitivity of the forecast performance to the size of the training sample. The third series of experiments examined the impact of the number of hidden nodes (middle column of the schematic in Figure 3-1) used in the ANN formulation. The results of these experiments are presented below. The second objective was addressed by setting the ANN forecast system to the “best” configuration found in the sensitivity experiments and producing forecasts for each of the wind power production regions for which data were available.

Specification of Predictors

Table 3-2 presents the list and descriptions of the fixed set of predictor variables used in the forecast experiments and also assigns a label to each predictor variable. The discussion of the experimental results uses these labels to refer to the input parameters used in each experiment.

The first predictor variable is the most recent power production and is labeled *PP1* (“Power Production”). It is defined as the power production during the five-minute interval immediately preceding the forecast production time. If the forecast production time is designated as 0 minutes, then the PP1 predictor is the power production for the zero-to-five-minute interval.

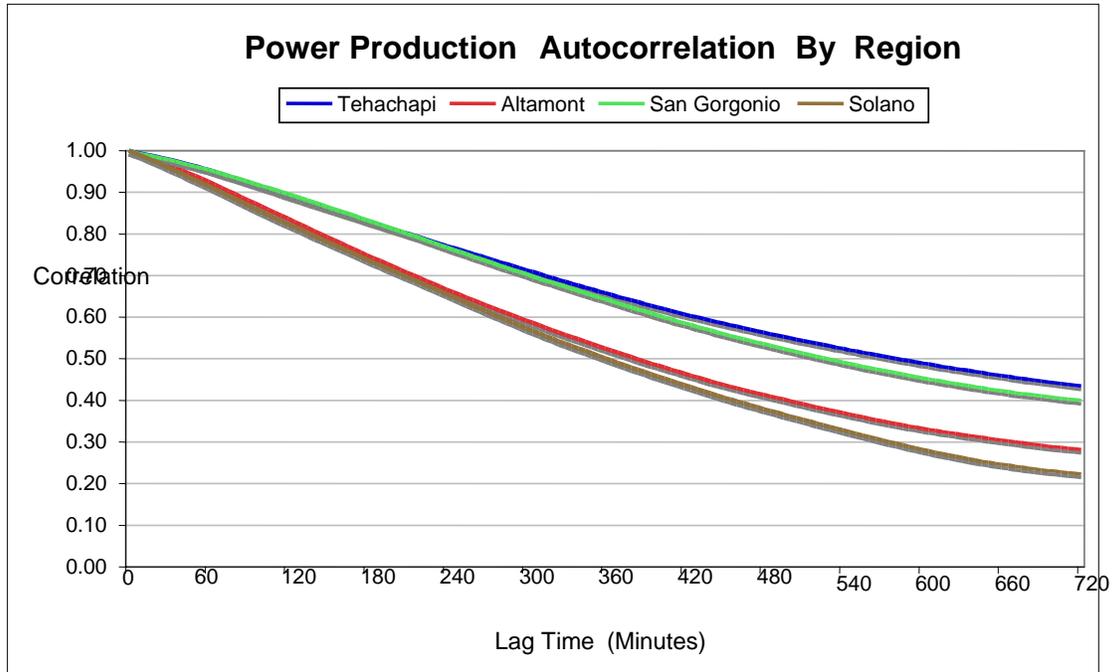


Figure 3-3 Autocorrelation of the five-minute power production for four wind power production regions in California based on data for the calendar year 2004.

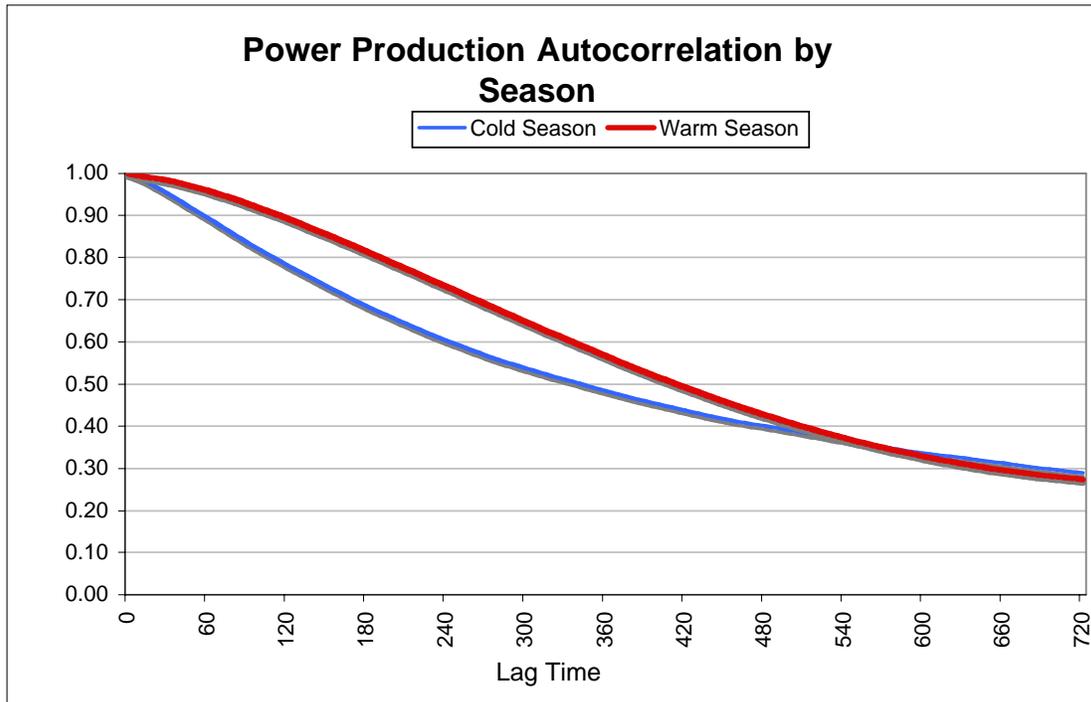


Figure 3-4 Autocorrelation of the five-minute power production for the Altamont region for the cold (October to March) and warm (April to September) seasons of 2004.

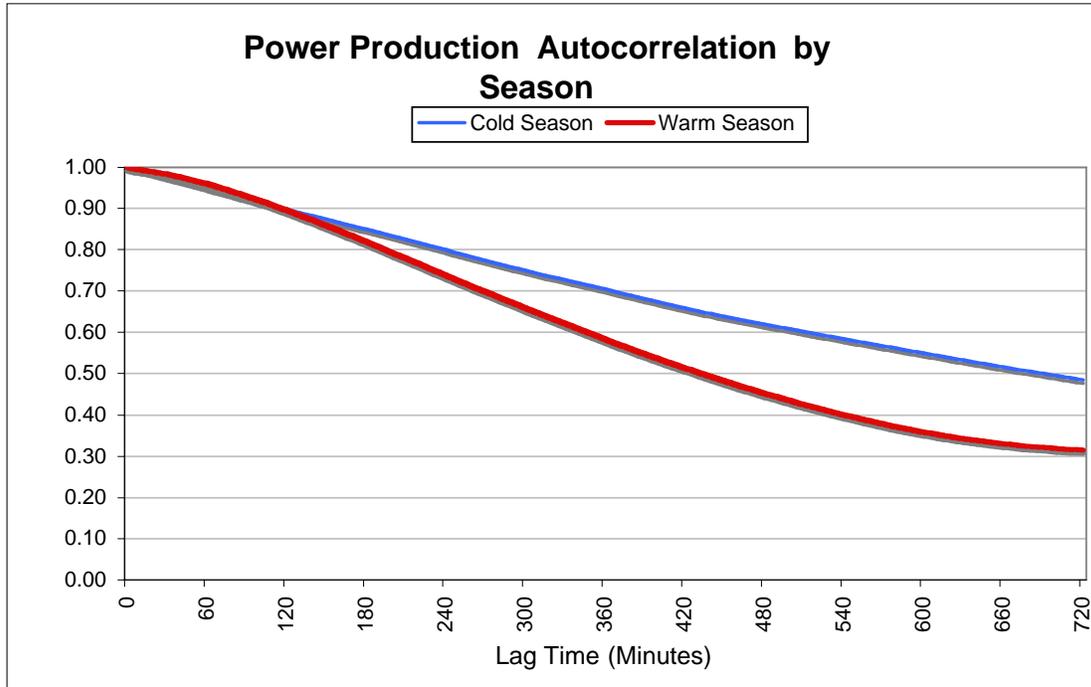


Figure 3-5 Autocorrelation of the five-minute power production for the San Geronio region for the cold (October to March) and warm (April to September) seasons of 2004.

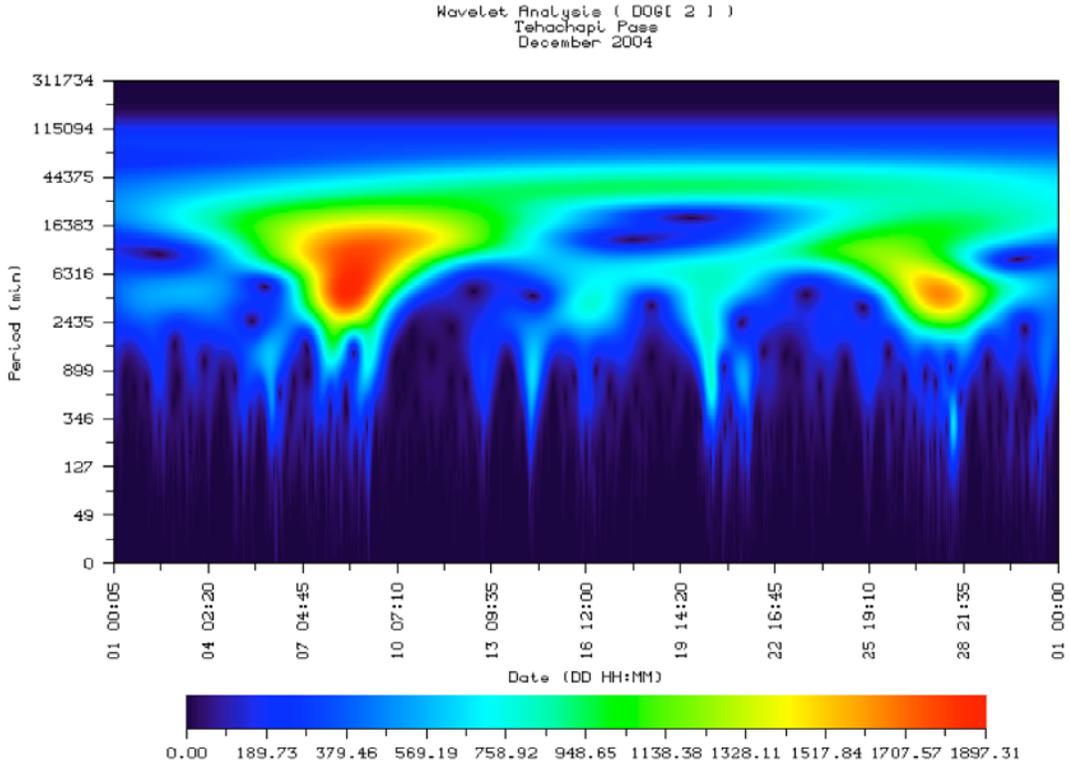


Figure 3-6 Wavelet amplitude spectrum of the power production data for the Tehachapi region for December 2004 based on the Derivative of Gaussian (DOG) wavelet function.

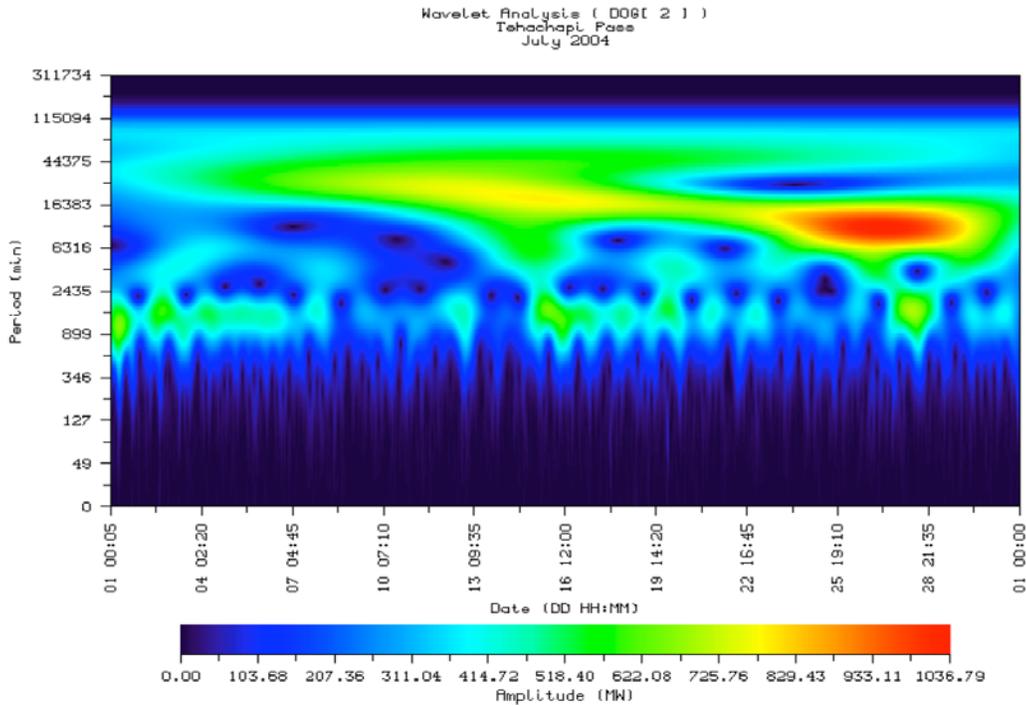


Figure 3-7 Wavelet amplitude spectrum of the power production data for the Tehachapi region for July 2004 based on the Derivative of Gaussian (DOG) wavelet function.

Table 3-2 Description of predictors used in the 0 to 3 hour forecast experiments-

Predictor Label	Description
PP1	The reported power production from the <i>most recent</i> interval prior to the forecast production time (i.e. the 0 to -5 minute interval for 5 minute data with a forecast production time of 0 minutes);
PPTR1	The change in power production between the <i>most recent</i> interval (P1) and the immediately prior interval regardless of the forecast look-ahead interval
PPTRO1	The change in power production change between the <i>most recent</i> interval and the preceding interval corresponding to the same offset time as the forecast look-ahead period (for example, the change from the -35 to -30 minute interval to the -5 to 0 minute interval for a 30-minute ahead forecast)
PPO2	The power production from the preceding interval that corresponds to the same offset time period from the forecast product time as the forecast look-ahead period (i.e. the -35 to -30 minute interval for a 30-minute forecast)
PPTRO2	The change in power production over a time period preceding the PPO2 interval by a period that is equal to the forecast look-ahead period (for example, the change from the -40 to -35 minute interval to the -10 to -5 minute interval for a 30 minute ahead forecast);
PP2	The power production from the <i>second most recent</i> interval prior to the forecast production time (i.e. the -5 to -10 minute interval)
PPDA	The power production from the <i>day-ago</i> interval corresponding to the forecasted interval (i.e. if the forecast is for the 2:30 p.m. to 2:35 p.m. interval, the production for that time period from the previous day is used as the input);
PPTRDA	The change in power production between the forecast production time and the forecasted interval from the previous day (i.e. if a forecast is made at 2 p.m. for the 2:30 to 2:35 p.m. interval, the input is the previous day's change in production between the 1:55 p.m. to 2:00 p.m. interval and the 2:30 p.m. to 2:35 p.m. interval);
ERR	The forecast <i>error</i> for the forecast with the same look-ahead period for the most recent interval preceding the forecast production time (i.e. for a forecast for the 2:30 to 2:35 p.m. interval made at 2 p.m. (a 30-minute forecast), the ERR predictor is the forecast error for the 30-minute ahead forecast (the one made at 1:30 p.m.) for the 1:55 p.m. to 2:00 p.m. interval).

The fifth predictor labeled “Power Production Trend Offset 2” or *PPTRO2* is the change in power production over a time period preceding the PPO2 interval that is equal to the forecast look-ahead period. Thus for a 30-minute forecast, *PPTRO2* is the change in power production from the -65 to -60 minute interval to the -35 to -30 minute interval. Thus, the use of *PPTRO1* and *PPTRO2* together provide information about the second derivative of the power production (the rate of change of the trend in the power production).

The sixth predictor is labeled *PP2* (Power Production 2” and is the power production for the period immediately preceding the most recent interval regardless of the forecast look-ahead period. Thus, *PP2* is always the power production for the -10 to -5 minute interval.

The seventh predictor is the power production from the previous day for the clock time of the forecasted interval. Its label is *PPDA* (“Power Production Day Ago”). Thus, if the forecast is for

the 2:30 p.m. to 2:35 p.m. interval, the PPDA input value is the power production for that time period on the previous day.

The eighth variable is the change in power production between the forecast production time and the forecasted interval from the *previous day*. That is, if a forecast is made at 2:00 p.m. for the 2:30 to 2:35 p.m. interval, the input is the previous day's change in production between the 1:55 p.m. to 2:00 p.m. interval and the 2:30 p.m. to 2:35 p.m. interval. The label for this predictor is *PPTRDA* ("Power Production Trend Day Ago").

The ninth variable is the *error* of the forecast with the same look-ahead period for the most recent interval preceding the forecast production time. Thus, for a forecast for the 2:30 to 2:35 p.m. interval made at 2:00 p.m. (a 30-minute forecast), the *ERR* predictor is the forecast error for the 30-minute ahead forecast (the one made at 1:30 p.m.) for the 1:55 p.m. to 2:00 p.m. interval).

Sensitivity Tests

The period extending from June 1 to July 31, 2004, was selected as the test period. This period was selected primarily because it is during the time of the year when the wind power production is near its peak in most parts of California, which means that the accuracy of forecasts for grid management is most critical. A secondary factor is that there are very few missing data during this period. The Tehachapi region was used as the venue for these experiments.

Three sets of experiments were executed. The first set tested the impact of the changes in the number and type of inputs to the ANN scheme. A comparison was also made between the performance of a multiple linear regression scheme and the ANN scheme. The second set of experiments evaluated the impact of changes in the training sample size. Training sample sizes ranging from 3.5 days to two months were tested. The third and final set of experiments examined the impact of the number of hidden nodes used in the ANN configuration.

All of the experiments employed a three-layer feedforward neural network. This configuration consists of an input layer, one hidden layer, and an output layer and is depicted in Figure 3-1. The training process was the same for all of the experiments and was based on the use of simulated annealing and the conjugate gradient algorithm. First, simulated annealing was used to initialize the hidden layer weights. A regression was then used to initialize the output weights. After initialization, a conjugate gradient algorithm was employed to minimize the mean-squared output error. When a minimum is found, simulated annealing is used to attempt to break out of what may be a local minimum. If annealing reduces the error, the conjugate gradient method is used again. The cycle then repeats itself indefinitely until either a termination criteria is reached or several iterations in a row produce only trivial improvement. The termination criteria consisted of an error threshold or a maximum number of iterations.

The evaluation of the forecasts was primarily based on two parameters: the mean absolute error as a percentage of the region's installed capacity and the skill score. The mean absolute error (MAE) is the average absolute error of the forecasts and is defined as,

$$MAE = \frac{\sum_{i=1}^N |F_i - O_i|}{N} \quad (\text{Eqn 3-10})$$

where F is the forecasted value, O is the corresponding reported value, and N is the number of forecast intervals in the evaluation sample. The general definition of a skill score is the percentage reduction of an evaluation parameter relative to a reference forecast. In this context, the evaluation parameter is the MAE and the reference forecast is generally a persistence forecast. However, in some cases the reference forecast will be defined as another type of forecast. Thus, the skill score is calculated from

$$Skill = \frac{(MAE_{ref} - MAE_{eval})}{MAE_{ref}} \quad (\text{Eqn 3-11})$$

where MAE_{eval} is the MAE of the forecast being evaluated and MAE_{ref} is the MAE of the reference forecast.

Number and Type of Inputs

The first series of experiments was designed to evaluate the impact of the number and type of input variables on the forecast performance. In this series, the variables listed in Table 3-2 were used in different combinations as inputs to the ANN procedure. All the experiments described in this section used a training sample size of one month, which in this case was June 2004, and were evaluated using the data for July 2004.

Two Input Variables

The first and most basic input experiment provided two inputs to the ANN algorithm. These were the most recent value of the power production (PP1) and the most recent trend in the power production (PPTR01) as described in Table 3-2. Figure 3-8 presents the mean absolute errors of the ANN and persistence forecasts as percentages of the installed rated capacity in the Tehachapi region, vs. the forecast look-ahead time.

The MAEs of both the persistence and the two-input ANN forecasts are just under 1% of the installed capacity for the first five-minute interval and are quite similar during the first 30 minutes of the forecast period. The MAEs gradually diverge after about 30 minutes with the ANN method producing the lower MAE values. At the end of the forecast period (180 minutes), the MAE of the ANN forecast is almost one full percentage point lower than that of the persistence forecast.

Another way to view the relative performance of the ANN and persistence forecasts is through the skill score parameter. Figure 3-9 presents the skill score of the ANN forecast

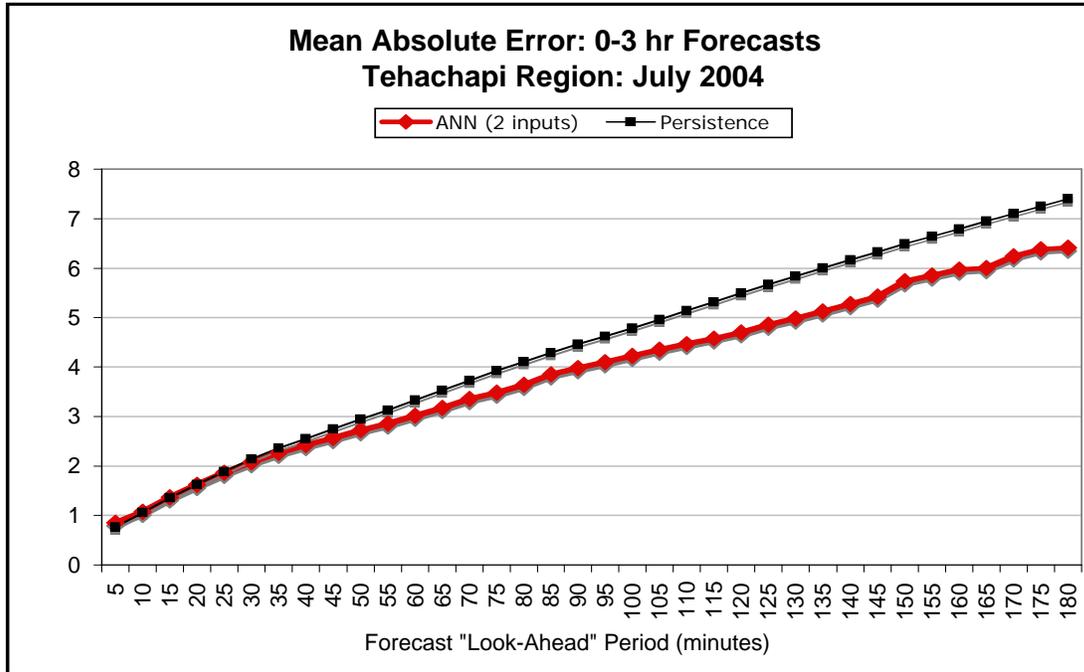


Figure 3-8 Mean absolute error (% of capacity) vs. look-ahead time for the 2-input ANN and persistence forecasts for the Tehachapi region. The two inputs to the ANN algorithm were PP1 and PPTRO1 from Table 3-2.

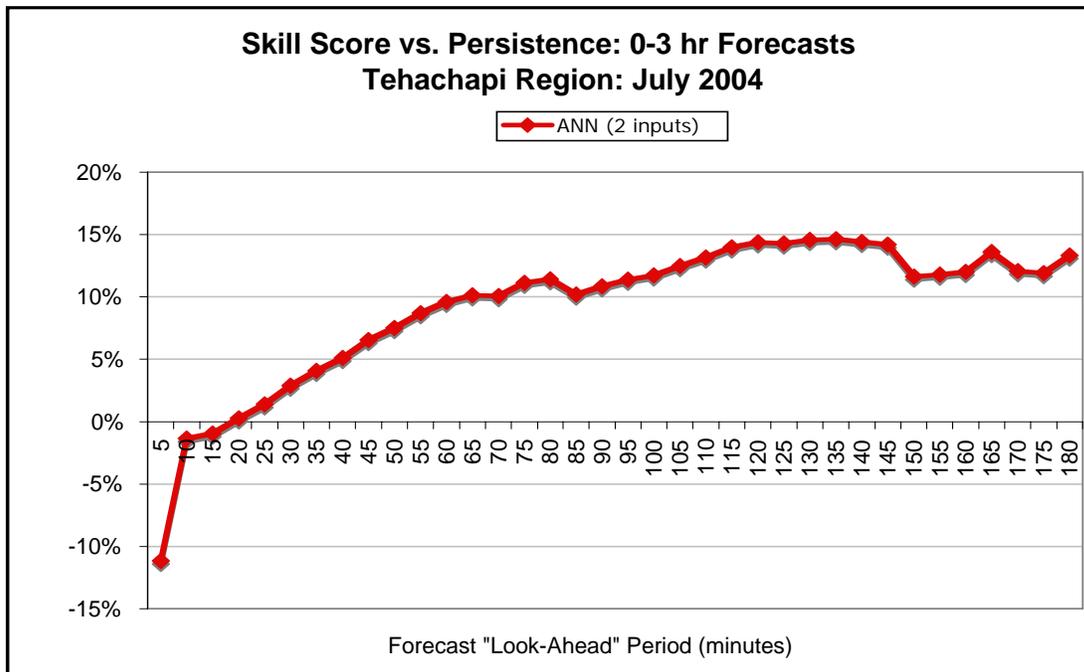


Figure 3-9 Skill score relative to persistence vs. look-ahead time for a 2-input ANN forecast for the Tehachapi region. The two inputs to the ANN algorithm were PP1 and PPTRO1 from Table 3-2.

relative to persistence vs. the forecast interval. The skill score is negative for the first three forecast intervals (0-15 minutes), which indicates that the persistence forecast has a lower MAE than the ANN forecast during this period. However, the skill scores rise rapidly from near 0% to +10% during the 20- to 70-minute period. After 70 minutes, the skill score declines slightly but remains in the +10% to +15% range.

The skill score pattern indicates that the recent trend information appears to provide little value for the first three forecast intervals (0-15 minutes). The trends on this time scale typically change rapidly and in what appears to be a fairly random fashion. As a result, this short-term trend information often misleads the forecast process and results in a MAE that is actually slightly higher than that of a persistence forecast. As the look-ahead time scale increases, the recent trend information becomes more valuable, and the skill scores exhibit a substantial increase. The peak in the skill score just after the two-hour look-ahead time frame is probably an indication that the most recent linear trend cannot be projected too far into the future.

An obvious question is: how does the forecast performance differ between a more sophisticated method, such as ANN, and a simple linear regression method, when using a set of two basic variables? To address this question, the PP1 and PPTR01 variables were used to train a two-variable multiple linear regression (MLR) model. Figures 3-10 and 3-11 present the MAEs for the two methods and the skill score of the ANN forecast relative to the MLR-based forecast.

The MAE values and the skill scores indicate that the ANN scheme generally outperforms the MLR method by an average of about 5% to 10% over the three-hour forecast period. The skill score chart (Figure 3-11) indicates that the ANN method outperforms the MLR method by the widest margin during the first two forecast intervals (0-10 minutes) and during the 100- to 140-minute interval.

Four Input Variables - Recent Power Level and Trend

The second experiment added two additional variables to the ANN input. The two additional variables were PP2 (the power production from a preceding interval offset from the forecast production time by the length of the forecast look-ahead period) and PPTR02 (the change in power production over a time interval ending at beginning of the PP2 interval [PPTR02]).

Thus, the four input variables (PP1, PPO2, PPTR01, and PPTR02) consisted of two recent power production variables and two recent trend variables. The presence of PPTR01 and PPTR02 together provides information about the second derivative of the power production (the rate of change of the linear trend in the power production).

Figures 3-12 and 3-13 respectively present the skill scores for the four-input ANN configuration relative to persistence and the two-input ANN scheme from Experiment 1. These charts indicate only modest improvements in the MAE. However, the improvement is not equally spread across the forecast period. The four-input scheme yields a 3% to 4% lower MAE during the 25- to 70-minute-ahead period and a 4% to 6% lower MAE during the 150- to 180-minute-ahead period.

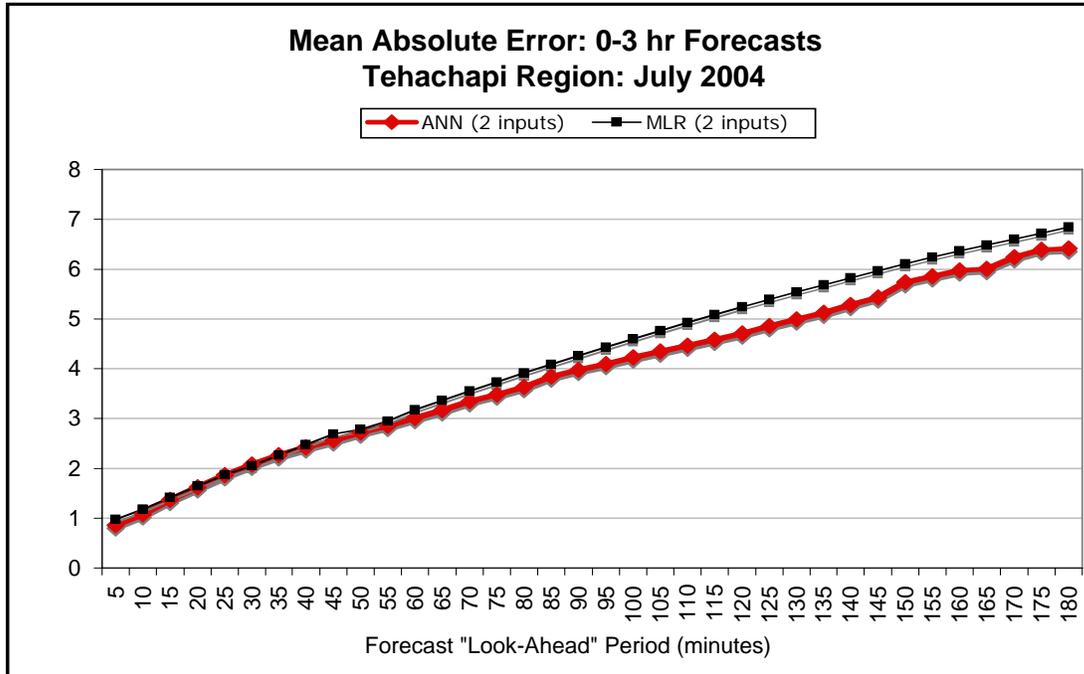


Figure 3-10 Mean absolute errors (% of capacity) vs. look-ahead time for the 2-input ANN and the 2-input MLR forecast methods. The inputs to both forecast algorithms were the PP1 and PPTR01 variables described in Table 3-2.

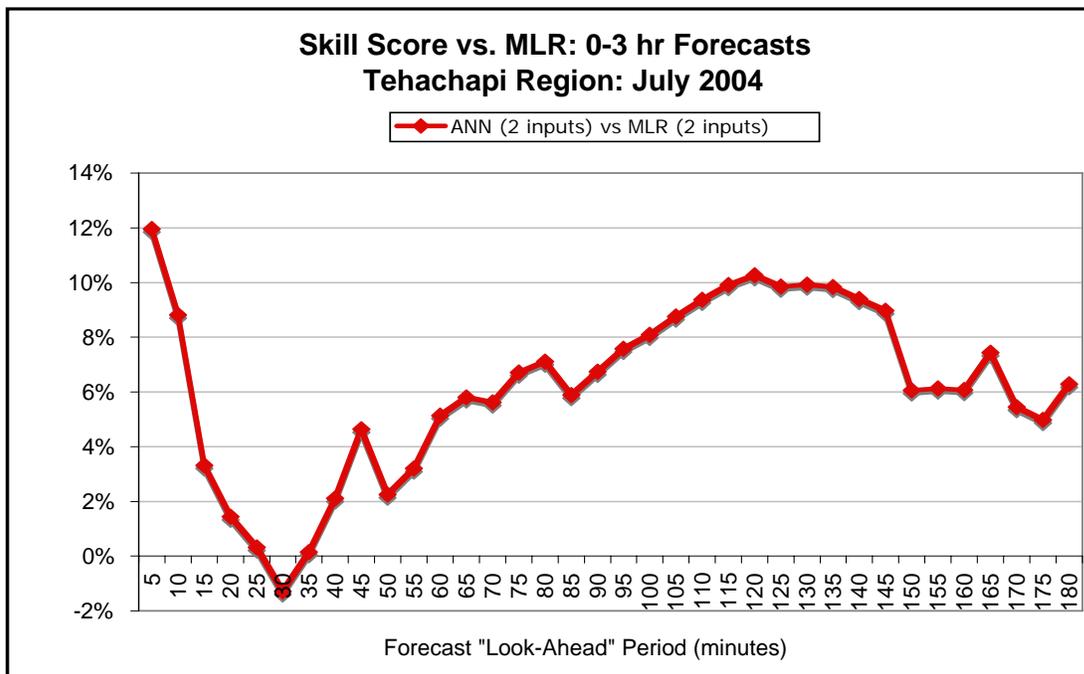


Figure 3-11 Skill score of the 2-input ANN forecast relative to the 2-input MLR method vs. look-ahead time. The two inputs for both methods were the PP1 and PPTR01 variables described in Table 3-2.

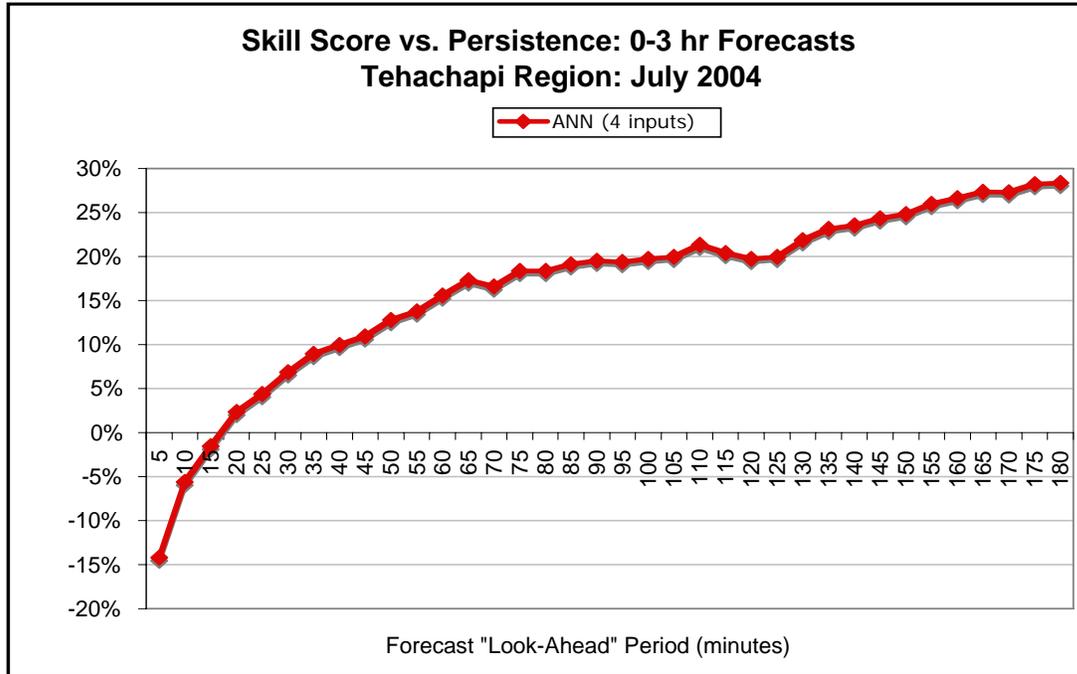


Figure 3-12 Skill score of the 4-input ANN forecasts (Exp. 2) relative to persistence vs. forecast look-ahead time. The four inputs to the ANN algorithm were PP1, PPO2, PPTR01, and PPTR02 as described in Table 3-2.

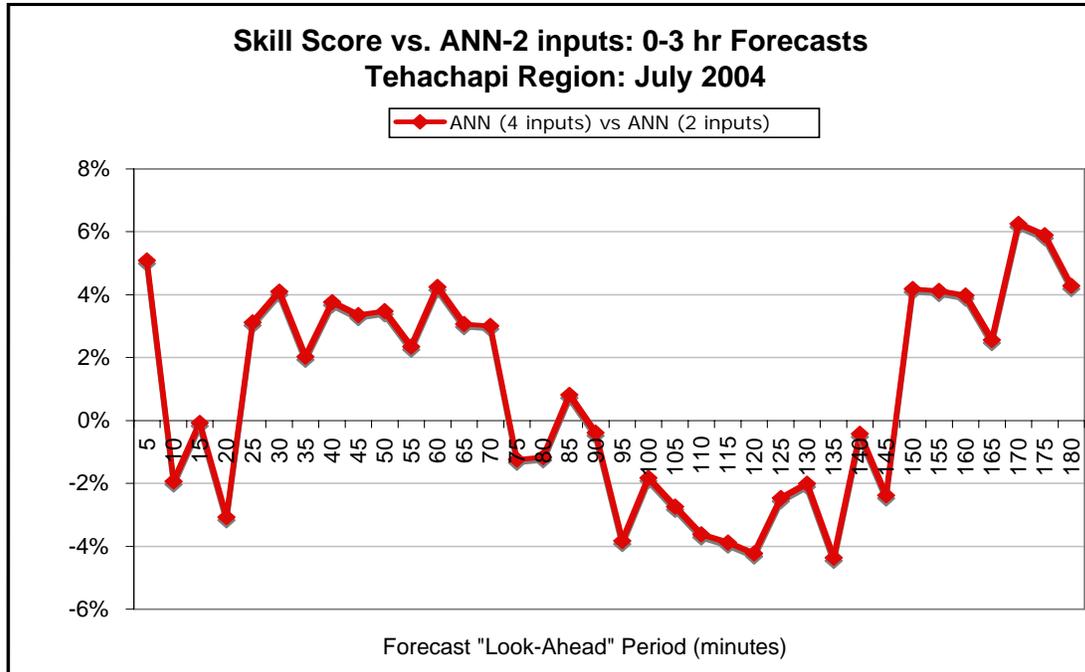


Figure 3-13 Skill score of the 4-input ANN method (Exp. 2) relative to a 2-input ANN method (Exp. 2) vs. forecast look-ahead time. The inputs to the four-input ANN forecast were PP1, PPO2, PPTR01, and PPTR02 as described in Table 3-2.

The two-input scheme actually yields a slightly lower MAE for most of the intervening 70- to 150-minute period. The end result is that the skill score of the four-input scheme relative to persistence (Figure 3-13) rises more rapidly than that of the two-input scheme during the 20- to 60-minute look-ahead period and reaches a peak just over +14% at the 60-minute look-ahead time. The skill vs. persistence then declines slightly to near the +10% level and remains near that level through the 120-minute mark. It then rises gradually to about +17% near the end of the three-hour forecast interval. The peak skill score is about 2 percentage points higher and occurs later in the forecast period than for the two-input ANN scheme. Thus, in general, the forecast performance of the four-input scheme is better during the first hour and the last half hour of the forecast period, but about the same or worse during the intervening hour and a half.

Four Input Variables - Recent Power Level and Trend on Previous Day

The third experiment also made four inputs available to the ANN procedure. Two of the predictors were the same as those used in Experiments 1 and 2 (PP1 and PPTR01). However, instead of providing additional information about the recent power production level and trend as in Experiment 2, information about the power production level and trend from the previous day was provided via the variables PPDA and PPTRDA.

Figures 3-14 and 3-15 present the skill scores for this four-input ANN scheme relative to persistence and the two-input scheme from Experiment 1. These skill scores indicate that the information from the previous day yields much more of a reduction in the MAE than the additional information about the recent trends (Experiment 2). The skill score relative to the two-input scheme from Experiment 1 (Figure 3-15) is slightly negative for the first two five-minute intervals but then rises fairly continuously to a high around +18% near the end of the three-hour period. Clearly, the previous day's information is more valuable than the recent trend information, especially for the longer look-ahead periods. The only advantage for the additional recent trend information employed in Experiment 2 is in the very short-term look-ahead periods.

The resulting skill-score vs. persistence for the four-input ANN with the previous day's information exhibits a rapid rise to +10% in the first 40 minutes and then continues to rise at a slower rate to a peak around +28% at the end of the three-hour period. This peak is much higher than the +15% and +17% persistence-based skill score peaks seen in Experiments 1 and 2. The skill score is above the Experiment 1 and 2 peaks for all of the last two hours of the forecast period. This underscores the value of the information from the previous day, at least during July in the Tehachapi region.

Six Input Variables

The fourth input experiment combined the four inputs from Experiment 2 and two additional variables added in Experiment 3. This resulted in a total of six ANN input variables—PP1, PP20, PPDA, PPTR01, PPTR02, and PPTRDA.

Figures 3-16 and 3-17 present the skill scores of the six-input configuration against persistence and the two-input scheme from Experiment 1. The six-input configuration yields a substantial positive skill score relative to the two-input scheme from Experiment 1 throughout

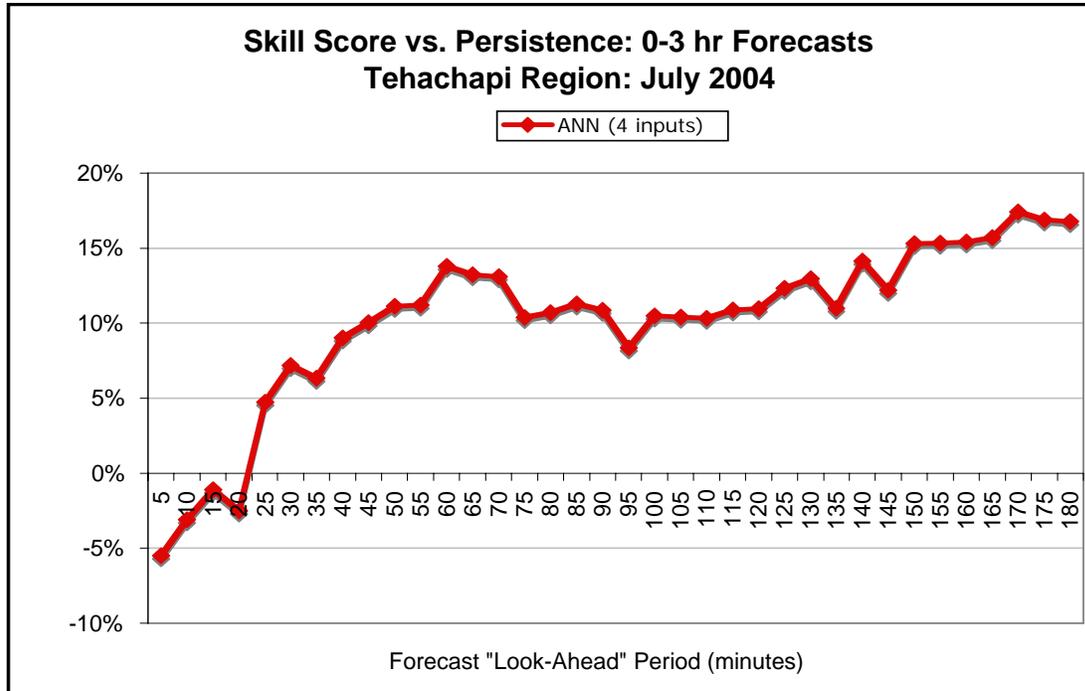


Figure 3-14 Skill score of the four-input ANN forecast (Exp. 3) for persistence vs. forecast look-ahead time. The four inputs to the ANN algorithm were PP1, PPDA, PPTRO1, and PPTRDA as described in Table 3-2.

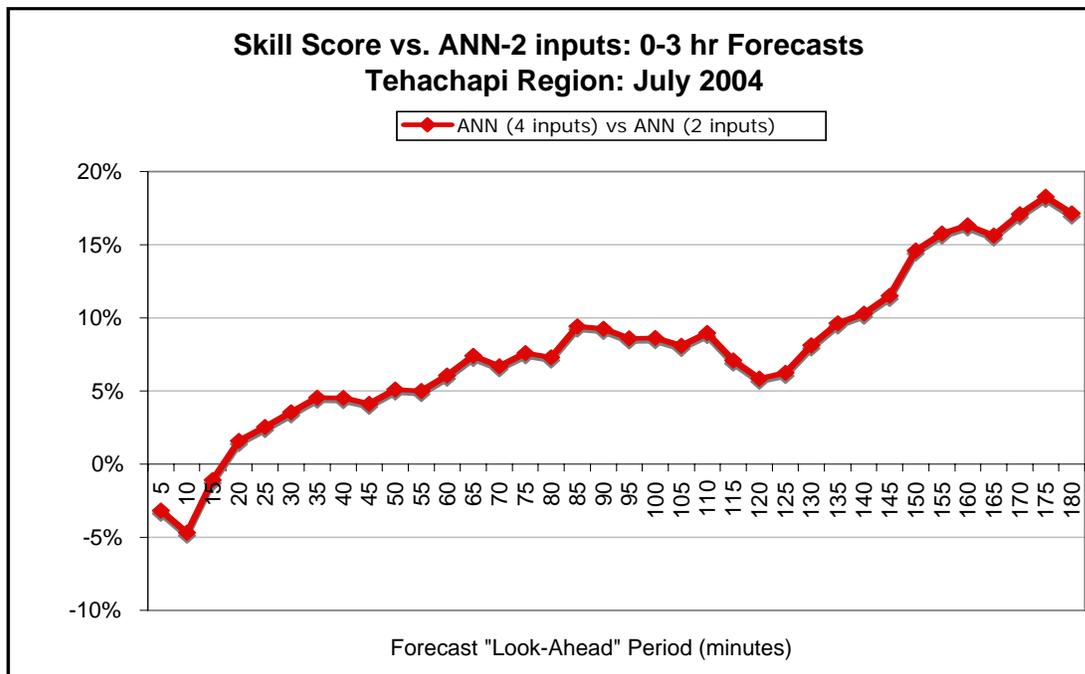


Figure 3-15 Skill score of the four-input ANN method (Exp. 3) using variables PP1, PPDA, PPTRO1 and PPTRDA against a 2-input ANN method (Exp. 1) using variables PP1 and PPTRO1 vs. forecast look-ahead time.

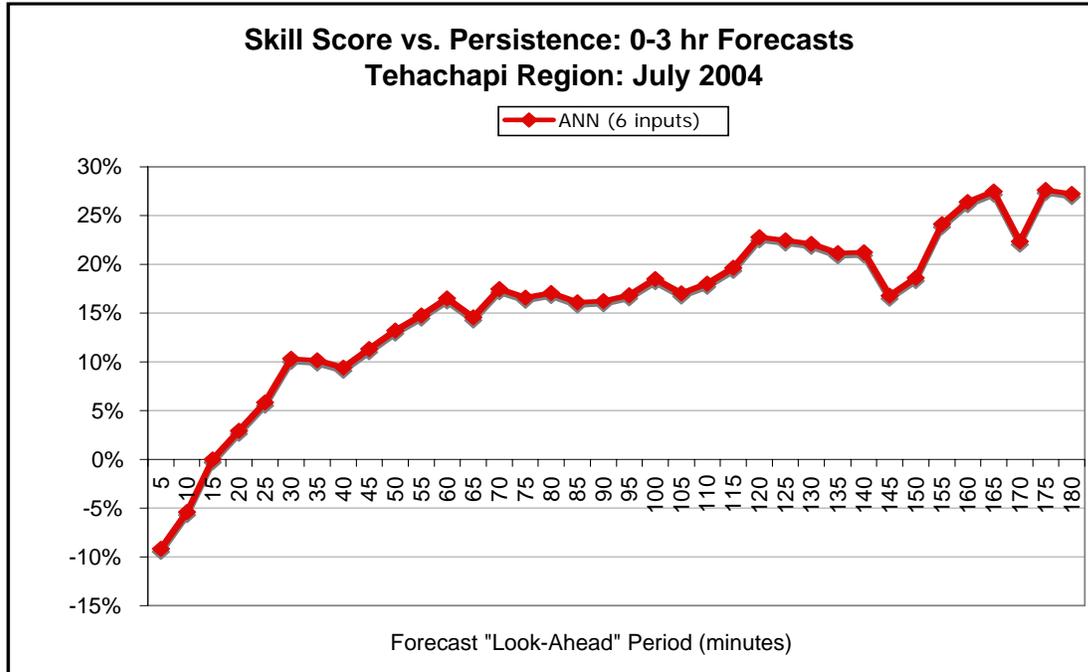


Figure 3-16 Skill score of the six-input ANN forecast method relative to persistence vs. forecast look-ahead. The six inputs to the ANN algorithm were P1, P2, PDA, PTR1, PTR2, and PTDA variables described in Table 3-2.

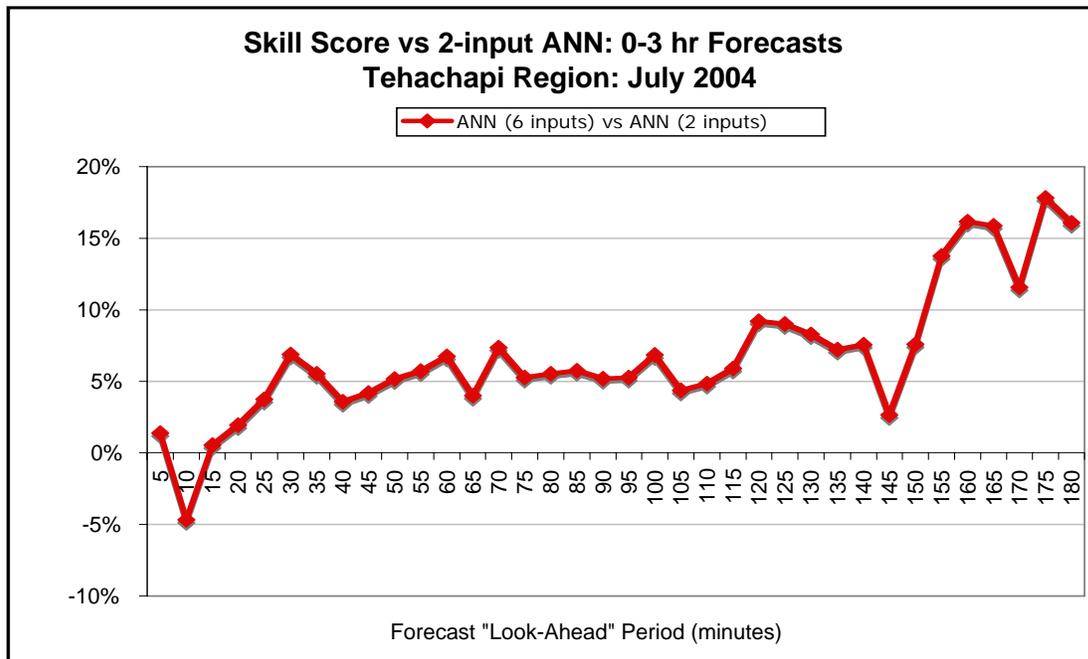


Figure 3-17 Skill score of the six-input ANN forecast method relative to a two-input ANN method vs. forecast look-ahead time. The inputs to the six-input ANN algorithm were the P1, P2, PDA, PTR1, PTR2, and PTDA, and the inputs to the two-input algorithm were the P1 and PTR1 variables described in Table 3-2.

most of the forecast period, except for the first two intervals. However, the forecast performance relative to the Experiment 1 scheme is generally no better than and, for some intervals, is worse than that of the four-input scheme using the previous day's information in Experiment 3. The only forecast improvements of the six-input scheme compared to the four-input method from Experiment 3 occurred during first interval (0 to 5 minutes) and in the vicinity of the 120-minute mark. Overall, the performance of this six-input scheme is not much different from that of the four-input scheme in Experiment 3.

Seven Input Variables – Forecast Error

The fifth and final input experiment was built upon the six-input scheme from Experiment 4 by adding the error variable ERR as an additional input. This variable provides information about the recent forecast errors to the prediction scheme.

Figures 3-18 and 3-19 present the skill scores for this seven-input forecast vs. persistence and the six-input configuration from Experiment 4. It is clear from these charts that the seven-input configuration yields substantially lower skill scores (the MAE is higher) relative to the six-input scheme used in Experiment 4 for almost every interval in the forecast period.

The fact that the addition of the error parameter makes the forecast performance somewhat worse for almost all look-ahead periods is rather puzzling. To gain some insight into why this occurred, researchers conducted a further analysis of the forecast performance for all input configurations of the ANN scheme. The focus was on the comparison between how well a specific configuration performed on the training sample versus how well it performed on the evaluation sample.

Figure 3-20 shows the training skill scores for four of the ANN configurations for each forecast look-ahead period. This chart indicates that the training skill score increases for every forecast look-ahead period as inputs are added. The four-input scheme always has a higher training skill score than the two-input scheme. In fact, the training skill scores for the seven-input scheme are very high, with scores in the +40% to +50% range in the latter half of the forecast period.

However, the corresponding chart for the evaluation sample presents a very different picture (Figure 3-21). The six-input configuration achieves a higher peak skill score than the seven-input scheme for almost all intervals, even though its performance in the training sample was much better than that of the six-input configuration.

The occurrence of better performance in the training sample and worse performance in the independent evaluation sample is a classic symptom of “overfitting.” That is, the seven-input ANN scheme was able to more closely fit data in the training sample, but the relationships resulting from this better fit to the training data were not representative of the relationships in the evaluation period. Thus, either the seven-input scheme was fitting noise in the data or it was modeling relationships that were not independent of time they changed from the training to the evaluation samples). The result is poor performance in the evaluation sample.

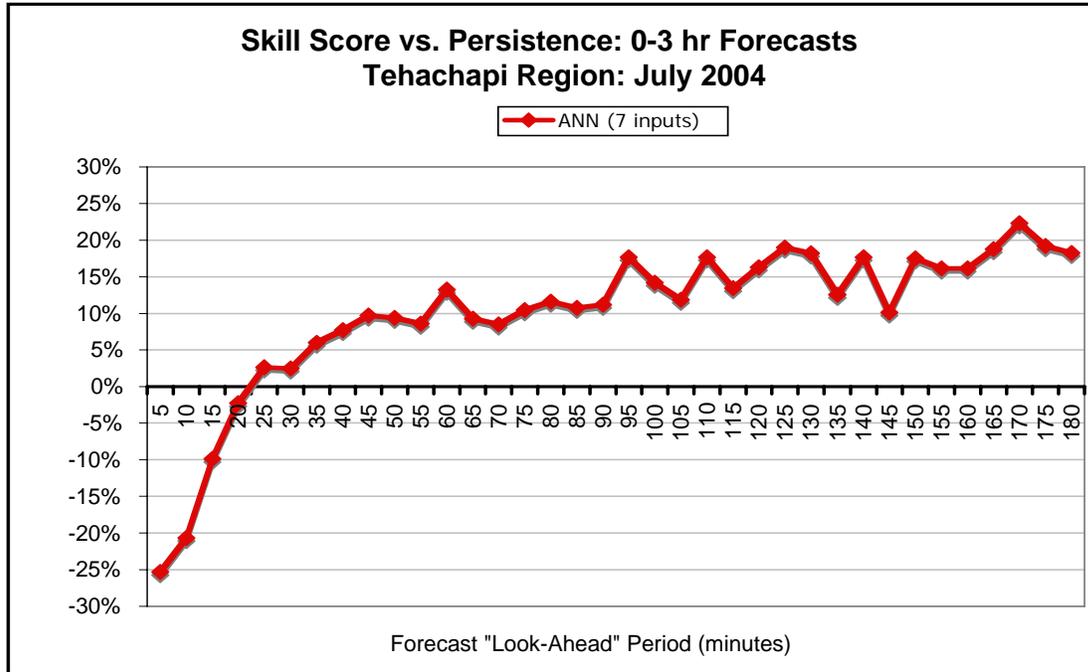


Figure 3-18 Skill score relative to persistence vs. forecast look-ahead time. The inputs to the ANN algorithm were the P1, P2, PDA, PTR1, PTR2, and PTDA and ERR variables as described in Table 3-2.

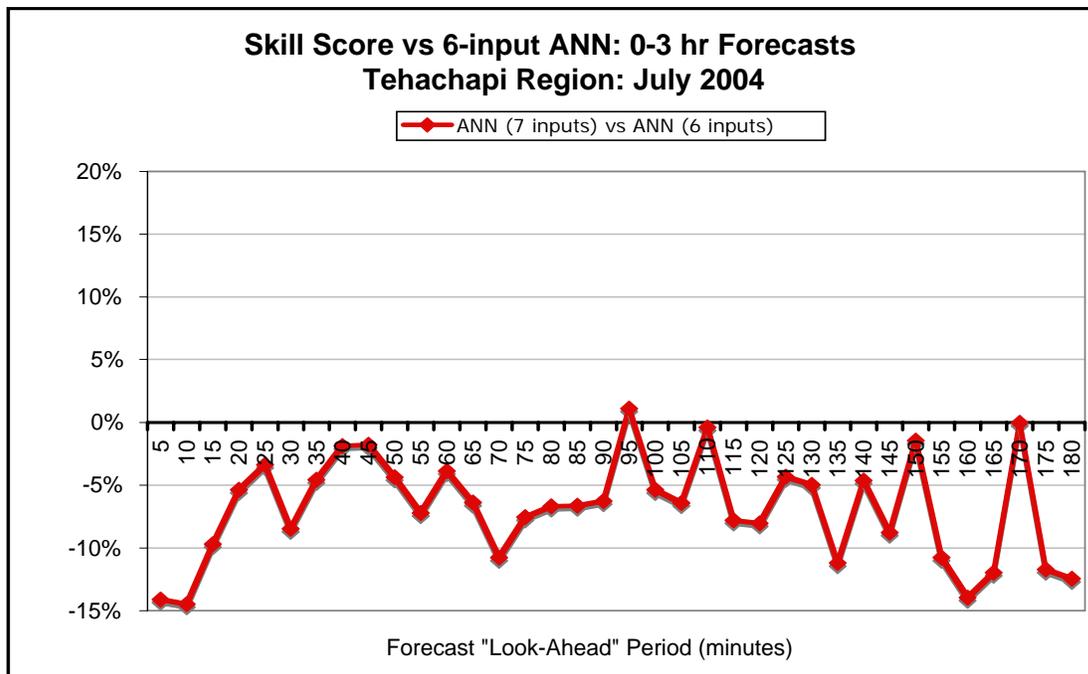


Figure 3-19 Skill score of the seven-input ANN forecast method relative to a 6-input ANN method vs. forecast look-ahead time. The inputs to the ANN algorithm were the P1, P2, PDA, PTR1, PTR2, PTDA, and ERR variables described in Table 3-2.

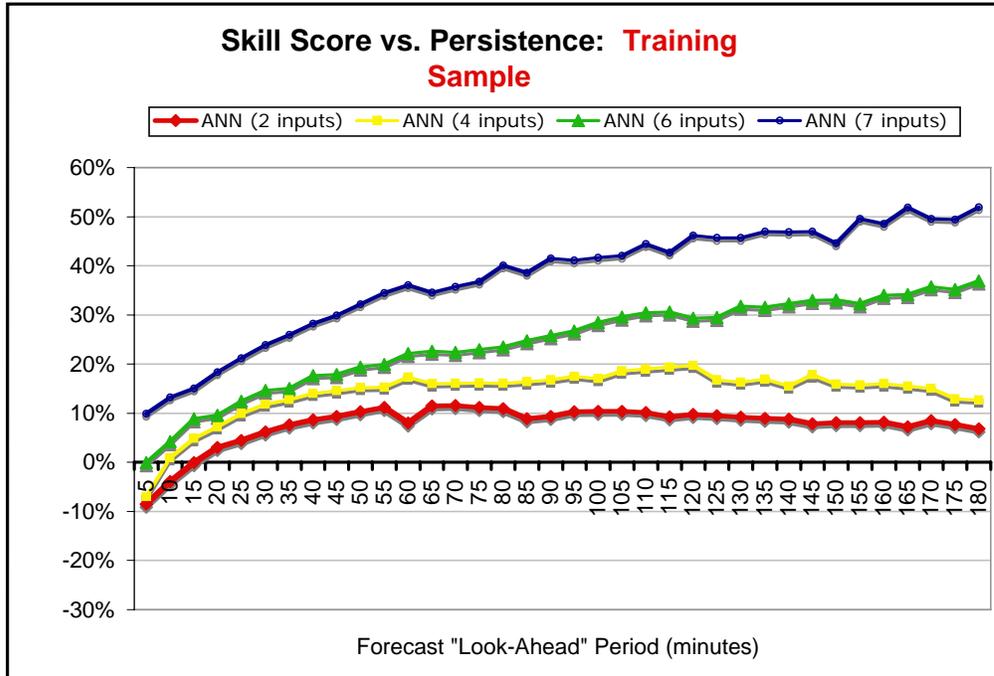


Figure 3-20 Skill scores of the training sample generated using the ANN algorithm with two, four, six, and seven inputs relative to persistence vs. forecast look-ahead time for the training sample (June 2004).

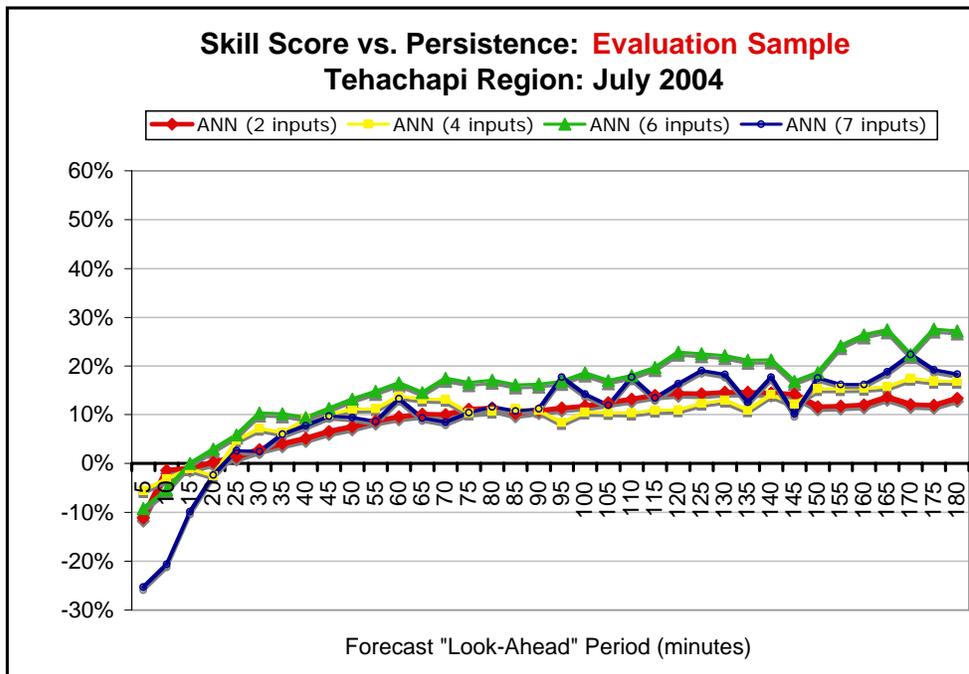


Figure 3-21 Skill scores of the evaluation sample generated using the ANN algorithm with two, four, six, and seven inputs relative to persistence vs. forecast look-ahead time (July 2004).

Figure 3-22 presents the skill score differences between training and evaluation samples for different input configurations and vs. the forecast look-ahead time. It illustrates how the skill score from the training sample to the evaluation sample for each of four ANN configurations. Interestingly, the change for the two-input configuration is near zero and actually slightly positive (it performed better in the evaluation sample) for the latter half of the period. Ideally, this is the relationship one would want to see between the training and evaluation samples. The four-input and six-input configurations generally exhibit decreases in skill scores between 0% and 10% for most intervals, and the six-input scheme exhibits a generally more negative, especially during the third hour of the forecast period. However, the skill score difference of seven-input configuration is much more negative than those of the three other configurations for all of the forecast intervals. It is clear that the relationships that were identified in the training sample do not hold in the evaluation sample for this configuration. This is most likely an acute case of overfitting.

However, it is curious that the addition of one additional input variable was able to trigger such a significant increase in overfitting. This may be attributable to the nature of the seventh parameter, the forecast error (ERR), which in essence provides information about how well the forecast scheme fits the data. This feedback on forecast performance may implicitly reduce the number of degrees of freedom in the data and result in greater overfitting than another type of parameter. Typically, overfitting issues can be addressed by increasing the training sample size. This is addressed in the next set of sensitivity experiments.

The conclusions of the input sensitivity experiment are that information from the previous day is very important for forecast performance, and that “overfitting” may be a problem with as few as

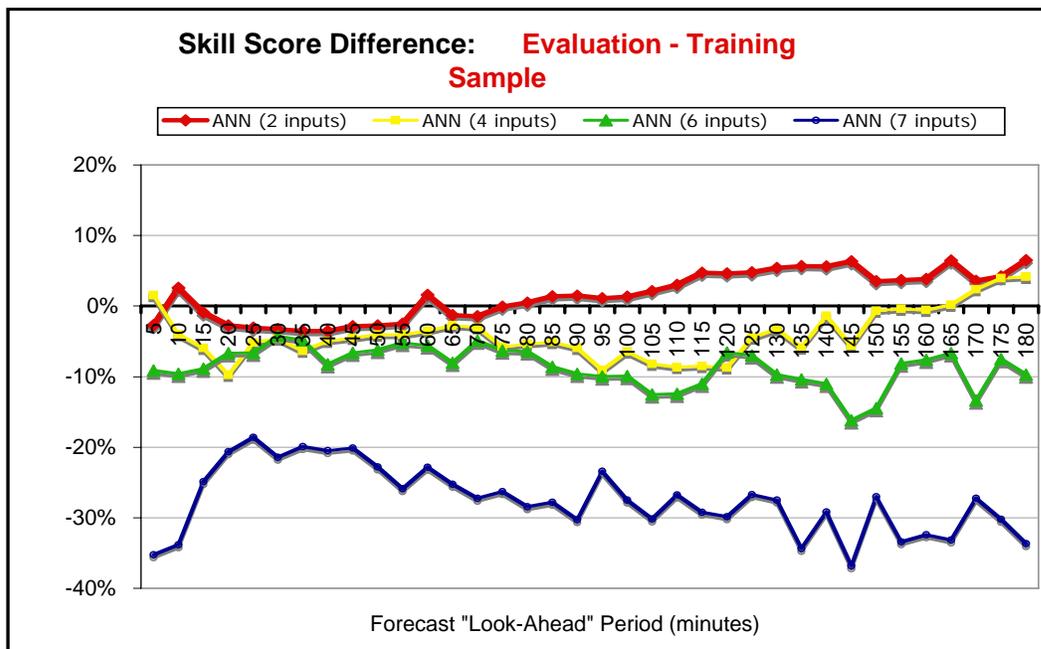


Figure 3-22 Difference in the skill scores relative to persistence between the training and evaluation samples by forecast look-ahead time for the ANN method with two, four, six, and seven inputs for the Tehachapi region.

seven inputs with a one-month training sample. The conclusion of the input sensitivity experiment is that information from the previous day is very important for forecast performance and that “overfitting” may be a problem with as few as seven inputs with a one-month training sample.

Training Sample Size

The second series of forecast experiments assessed the impact of changing the size of the training sample. The version of the ANN scheme used for this set of experiments was the six-input configuration used in Experiment 4 of the “input number and type” experiment series. There were six hidden nodes. This configuration used the PP1, PPO2, PPDA, PPTR01, PPTR02 and PPDA variables as input. The ERR variable was not used. The evaluation sample was identically the same for all of the training sample size experiments. It consisted of the entire month of July 2004. Five different training sample sizes were used: 3.5 days (1/2 week), one week, two weeks, one month and two months. All of the training samples ended on the last day of June 2004. Thus, the 3.5-day sample employed the last 3.5 days of June as the training sample, while the two-month sample consisted of the entire months of May and June 2004. While the 3.5 day training sample may appear to be small, it must be remembered that five-minute data are being used. A 3.5 day training sample consists of 1008 intervals. As in the “input number and type” experiments, the Tehachapi region data were used.

Figures 3-23 and 3-24 present the MAE and skill score relative to persistence vs. look-ahead period for each of the training sample size experiments are depicted in Figures 3-23 and 3-24.

The most striking feature of these two charts is that the performance of the forecasts based on the 3.5-day training sample is very poor. Clearly, that is an inadequate training size for this ANN configuration. As the sample size increases, the MAE decreases and the skill scores vs. persistence increase for all look-ahead periods. In general the largest improvement for most look-ahead periods is between the 3.5-day and one-week sample sizes. However, the improvement from one week to two weeks is also substantial for most intervals and in some cases is greater than for the 3.5-day to one-week change. The best performance occurred using the two-month training sample size for all look-ahead periods, although the improvement over the one-month sample size is typically modest.

An obvious question is the degree to which the overfitting symptoms change as the training sample size changes. This issue is addressed by Figure 3-25. As noted previously, a large decrease in the skill score from the training sample to the evaluation sample is a symptom of overfitting. Ideally, the performance should be similar in the training sample and the evaluation sample. The forecasts based on the 3.5-day training sample show a huge decrease in skill between the training sample and the evaluation sample. It is typically in the -70% to -100% range. This a strong indication of overfitting. The forecasts based on the one-week and two-week sample sizes show a more modest (-20% to -50%) but still substantial decrease in skill from training to evaluation sample. The decrease for the one-month training sample is in the 5% to -15% range and one might consider this to be an acceptable decrease in skill considering in the evaluation sample. Thus, one could justifiably argue that a two-month training sample is needed for combination of this ANN configuration and the characteristics of this dataset.

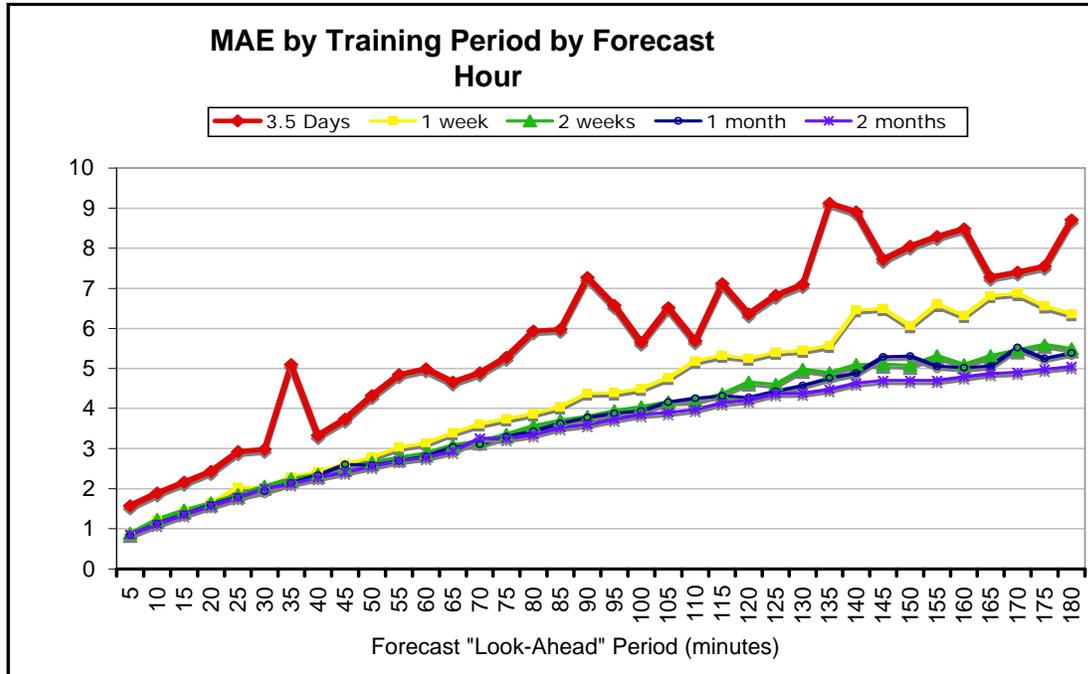


Figure 3-23 Mean absolute error (% of capacity) vs. look-ahead time for a six-input ANN forecast procedure with training sample sizes of 3.5 days, 1 week, 2 weeks, 1 month, and 2 months.

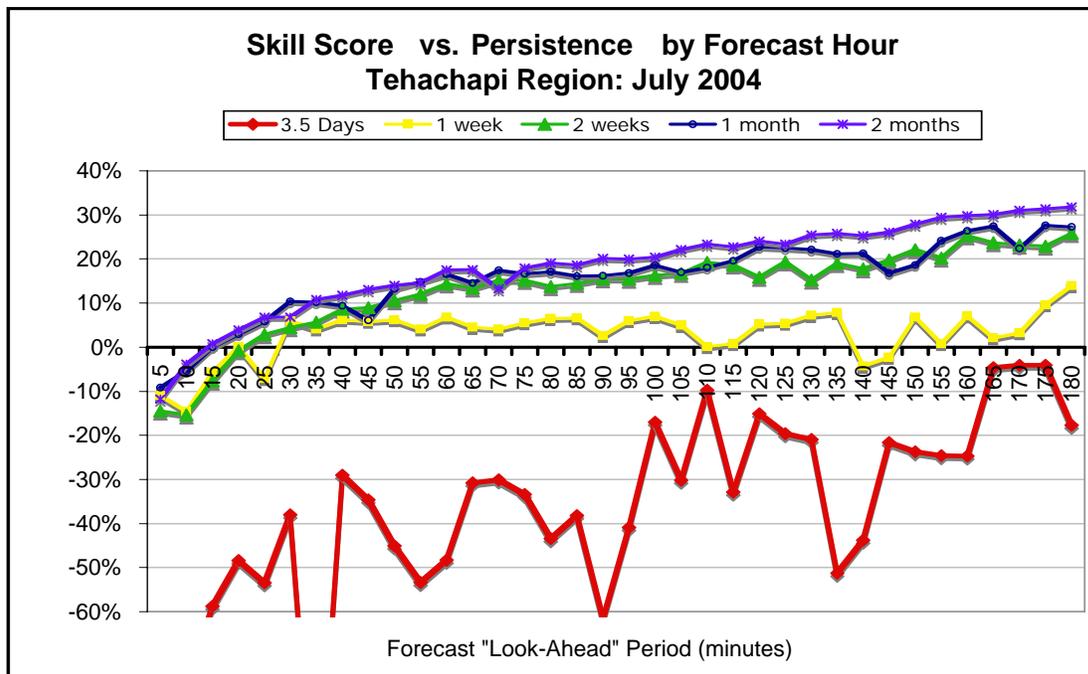


Figure 3-24 Skill scores relative to persistence vs. look-ahead time for a six-input ANN forecast procedure with training sample sizes of 3.5 days, 1 week, 2 weeks, 1 month, and 2 months.

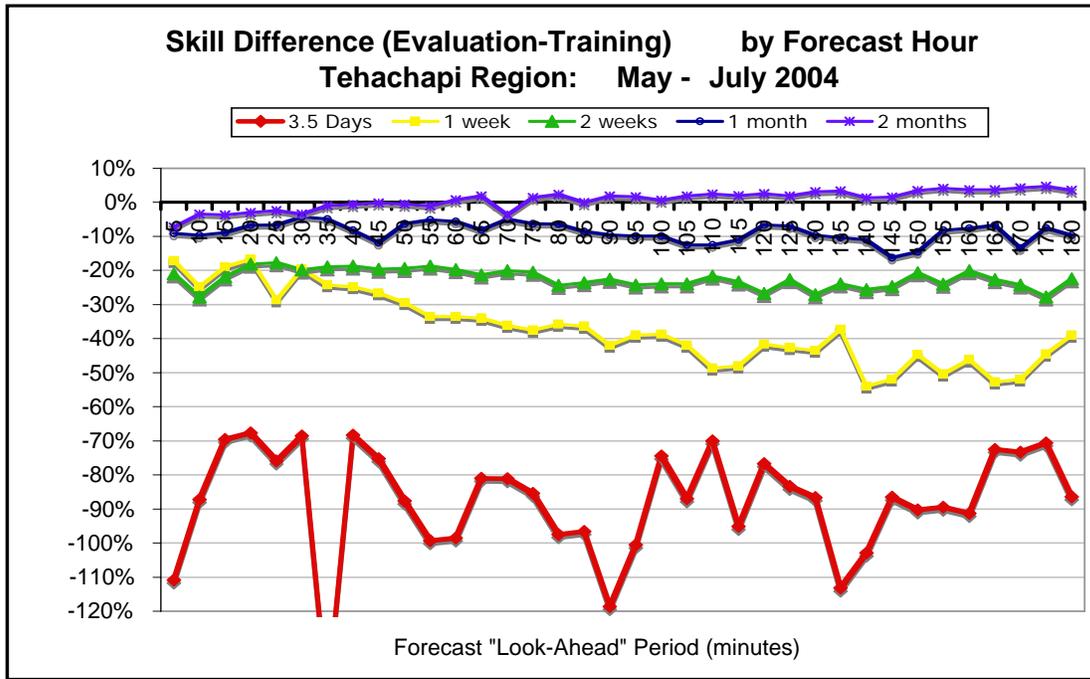


Figure 3-25 Difference in persistence-based skill scores between the training and evaluation samples vs. look-ahead time for ANN-based forecasts with training sample sizes of 3.5 days, 1 week, 2 weeks, 1 months and 2 months for the Tehachapi region.

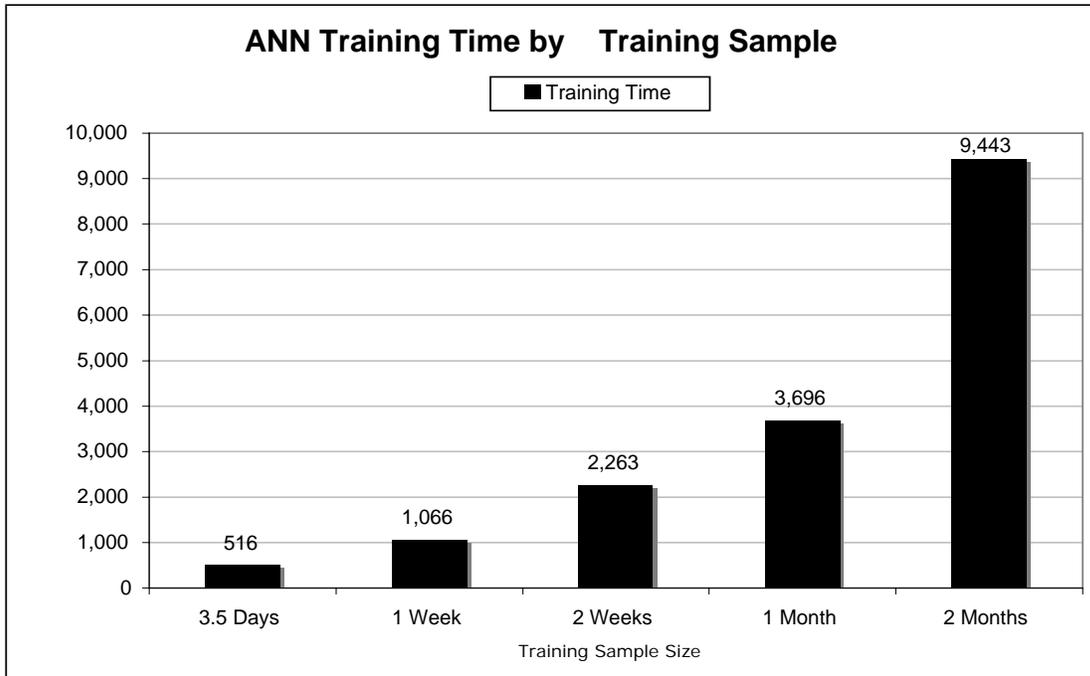


Figure 3-26 The average CPU time (seconds) on a single Intel P4 Xeon 2.2 GHz processor required to train the six-input ANN for each training sample size.

Although the performance data suggest that the two-month training sample is the best choice in this case, it is necessary to consider the computational time necessary to train the network if an application to a real-time forecast environment is contemplated. Figure 3-26 depicts the average CPU time required to train the network for all look-ahead periods and for each sample size. The CPU used is an Intel P4 Xeon with a 2.2 GHz clock speed. This data indicate that the training of this ANN configuration with a two-month dataset is a very substantial computational task. Of course, the clock time required for this task could be greatly reduced by performing the calculations in a parallel processing mode on multiple CPUs.

Number of Hidden Nodes

The third series of experiments evaluated the impact of changing the number of hidden nodes in the ANN configuration on the forecast performance. The number of hidden nodes in an ANN configuration determines the amount and complexity of the interaction among the input variables during the generation of the output from the network. When a low number of hidden nodes is employed, the interaction among the input variables is implicitly restricted to be simpler and more linear. The use of a zero hidden node configuration results in essentially a linear model in which the output is a weighted linear combination of all of the inputs. The use of a higher number of hidden nodes means that the interaction of the input variables is potentially more complex and non-linear. It should also be noted that the use of a larger number of hidden nodes means that there are more weighting coefficients in the network. Since these are parameters whose values are chosen to enable the network fit the training data, the use of a higher number of hidden nodes means that there is a higher risk of overfitting since the larger number of training parameters enables a closer fit to the sometimes noise-related details of the training sample.

The version of the ANN scheme used for this set of experiments was the same six-input configuration used in the training sample size experiments. The six inputs were the PP1, PPO2, PPDA, PPTR01, PPTR02 and PPDA variables. The ERR variable was not used. The evaluation and training samples used in the hidden-nodes experiments were identical to those used in the previous experiments. June 2004 was used as the training sample and July 2004 served as the evaluation sample. Seven experiments were executed to evaluate the impact of using zero to six hidden nodes.

Overall, the forecast performance differences among the number of hidden node configurations tested in this series of experiments were modest, as illustrated by the MAE and persistence-based skill scores vs. forecast hour curves illustrated in Figures 3-27 and 3-28 for five of the hidden node experiments. The experiments with three and five hidden nodes are not shown on the charts to enhance the readability of the charts. However, those configurations produced results that were quite similar to the four-node forecasts.

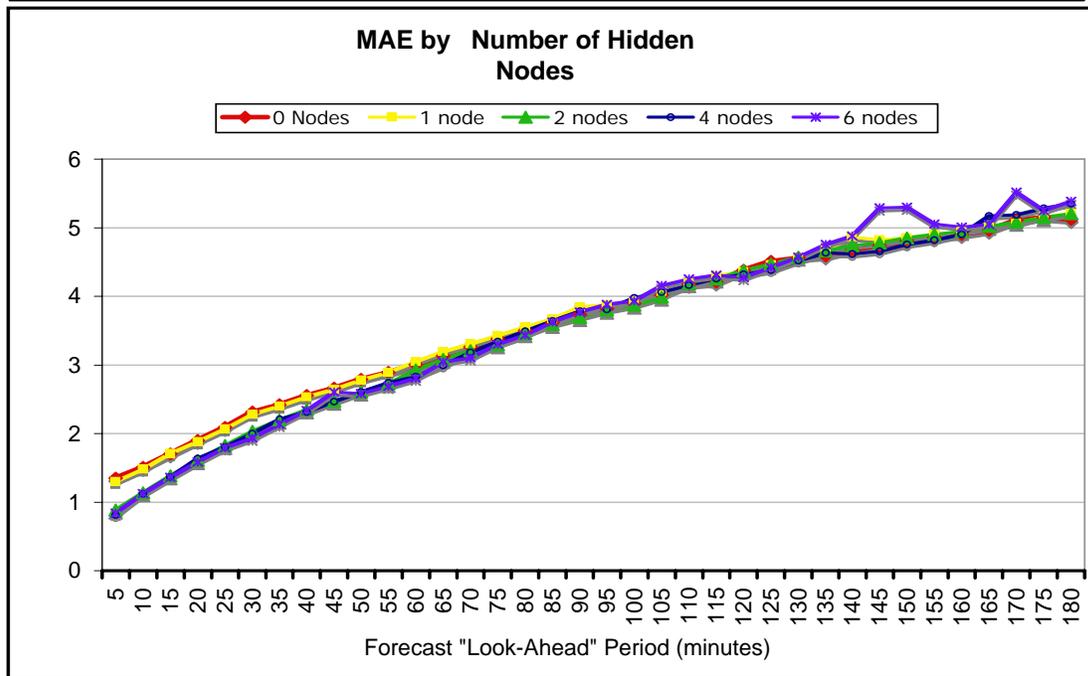


Figure 3-27 Mean absolute error (% of capacity) vs. forecast look-ahead time for an ANN forecast procedure with zero, one, two, four and six hidden nodes. All experiments used a one-month training sample (June 2004) and a one-month evaluation sample (July 2004).

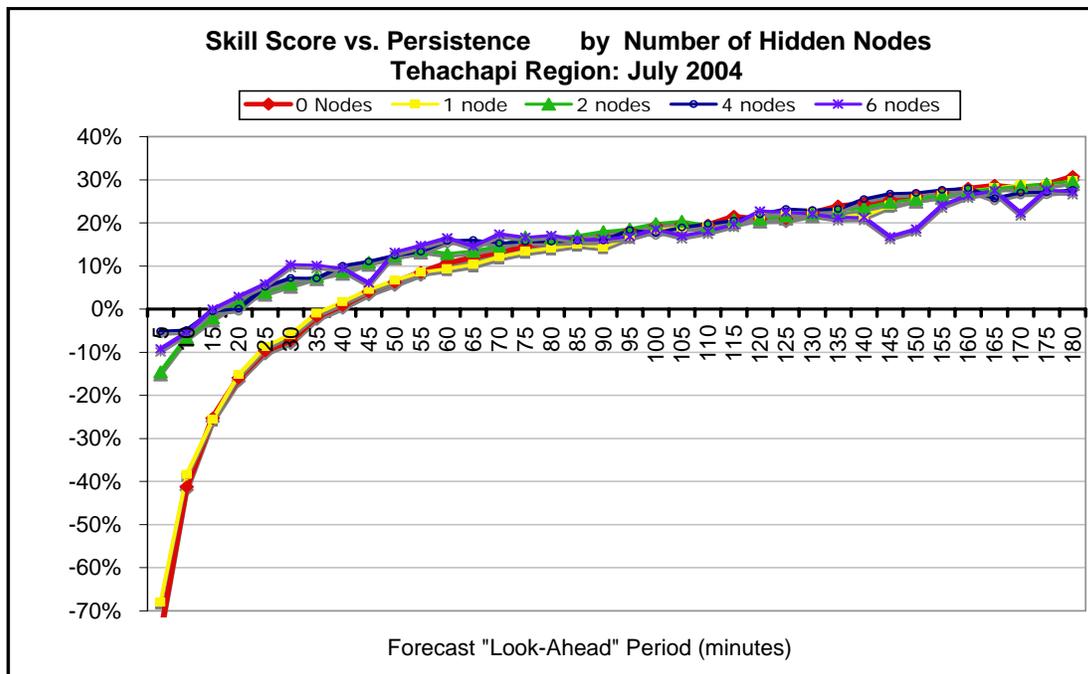


Figure 3-28 Skill scores relative to persistence vs. forecast look-ahead time for an ANN forecast procedure with zero, one, two, four and six hidden nodes. All experiments used a one-month training sample (June 2004) and a one-month evaluation sample (July 2004).

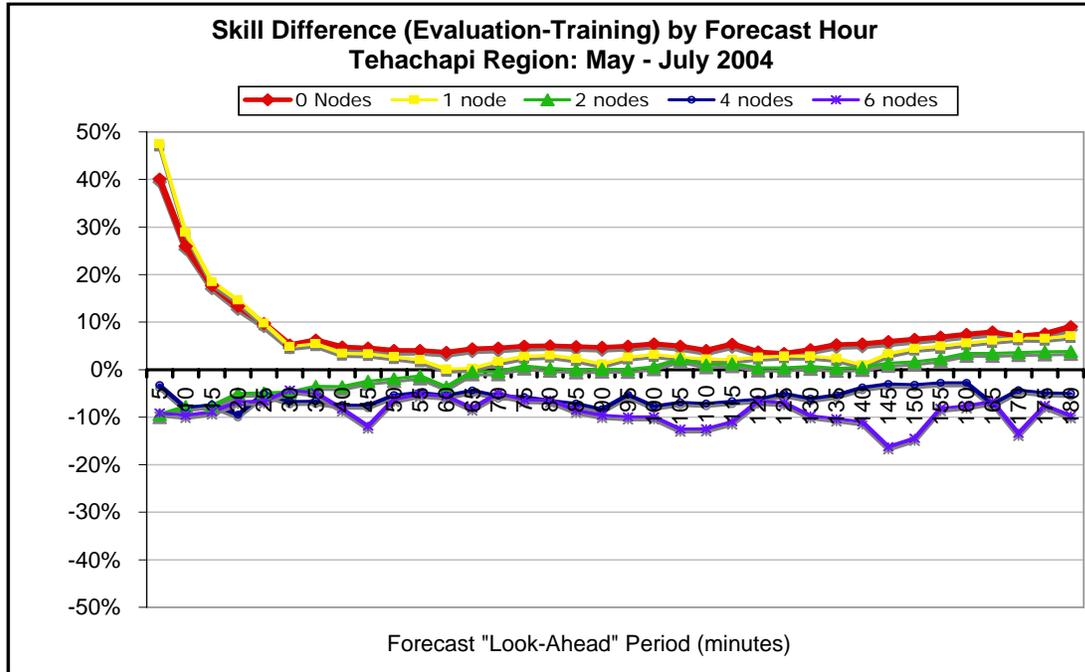


Figure 3-29 Difference in the persistence-based skill scores between the training and evaluation samples vs. forecast look-ahead time for an ANN procedure with zero, one, two, four and six hidden nodes for the Tehachapi region.

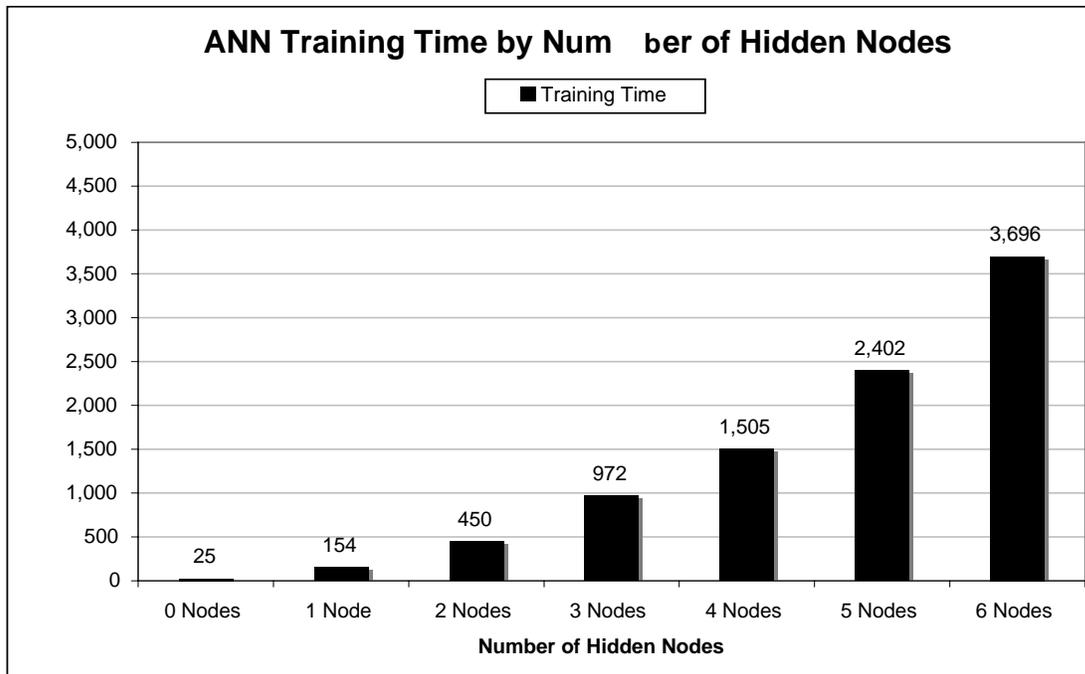


Figure 3-30 The average CPU time (seconds) on an Intel P4 Xeon 2.2 GHz processor to train an ANN with a one-month training sample and the specified number of hidden nodes.

Although MAE and skill scores for the all of the hidden node configurations exhibited a high degree of similarity, there were a couple of notable exceptions.

One is the significantly higher MAEs and lower skill scores for the zero- and one-node forecasts during the first hour. The departure from the performance of the higher hidden node number configurations is greatest for the first forecast interval (0 to 5 minutes ahead) and gradually decreases as the forecast look-ahead period increases until the zero- and one-node performance is almost indistinguishable from the two-, four- and six-hidden node configurations for look-ahead periods of about 80 minutes and longer. The poor performance of the zero- and one-node configurations during the early portion of the three-hour forecast period suggests that the performance of the very short-term forecasts is more dependent upon complex non-linear relationships that can't be modeled by a simple linear-type structure than the forecasts for the longer look-ahead periods in the three-hour forecast window. This is also supported by the fact that the two-input ANN scheme outperformed a multiple linear regression model by a greater margin during the first hour than the second hour of the forecast window (recall Figure 3-11).

The other departure from the overall similar performance is the higher MAE and lower skill score of the six-node configuration for a couple of intervals during the third hour of the forecast period. These appear to be anomalies, possibly attributable to a modest degree of overfitting in the six-node training process.

Some insight into the overfitting issue can be gained by the examining the change in skill score from the training sample to the evaluation sample for the hidden node experiments. This change in skill score is shown for five of the hidden node configurations in Figure 3-29. The zero-node and one-node forecasts have a positive change for all forecast intervals. This means the forecasts performed better in the evaluation sample than in the training sample. The positive changes are generally around 5% except for the first 25 minutes of the forecast period. There are very large positive differences during the initial 25 minutes of the forecast period for both the zero- and one-node configurations. The cause of this is not clear. However, it is not present in any of the configurations with two or more hidden nodes. The two-node configuration has the greatest similarity between the performance in the training and evaluation samples. The skill difference between the two samples is very close to zero for most of the intervals during the three-hour forecast period with a tendency for negative differences during the early portion of the forecast period and positive differences towards the end of the period. The skill score differences are negative for all forecast intervals for the four- and six-node configurations. The negative values are large for the six-node configuration, especially during the latter half of the three-hour forecast window. This may be an indication that the six-node configuration is on the verge of experiencing overfitting issues.

Of course, in an operational forecast environment, the time it takes to train the network can be an issue especially if the training is to be done frequently. Thus, in planning a forecast system, the computational requirements for different forecast system configurations should be considered. The average CPU time required to train each hidden-node configuration of the ANN using a one-month training sample is depicted in Figure 3-30. These CPU times are for a single Intel P4 Xeon processor with a 2.2 GHz clock speed and 1 GB of RAM. The ANN executable used in these tests was compiled from C++ source code. A separate network was trained for each

forecast interval. There are 36 forecast intervals in the three-hour forecast period, so each bar in the chart represents the average training time for 36 different networks with the specified number of hidden nodes. Thus, the total CPU time required to train networks for all 36 forecast intervals in the three-hour forecast window is 36 times the values shown in Figure 3-30. The zero-node configuration required an average of only 25 seconds to train. The two-node configuration required more than 10 times that amount, while the six-node configuration required more than 100 times more CPU time for training. In an operational environment, it would be possible to use multiple CPUs in a parallel processing environment to accelerate the training of the networks.

In summary, the “number-of-hidden-node” experiments indicated that for a one-month training sample from the Tehachapi region, a two- or three-node configuration produced a good combination of performance and reasonable CPU times for training. The use of a zero- or one-node configuration produced much worse performance during the first hour of the forecast period. The six-node configuration did not perform much better than the configurations with two to four hidden nodes and required substantially more CPU time for training. There was also some indication that the five- and six-hidden node configurations experienced some overfitting issues. This issue might be addressed by employing a larger training sample, but of course that would further increase the training time.

Multi-Region Forecast Performance Results

To assess the forecast performance over a broader spectrum of regions and establish an estimate of the performance level of a forecast for the aggregated power production of all wind generation on the CA ISO system, forecasts were generated for 2004 for four of the five regions for which data were supplied by CA ISO. The Pacheco region was not addressed because availability and capacity information were not available in the data files for this region.

The same ANN configuration was used to generate forecasts for each of the four regions. This configuration included a total of six inputs. They were the PP1, PPO2, PPDA, PPTRO1, PPTRO2, and PPDA variables listed in Table 3-2. The number of hidden nodes was set to two, and a training sample size was one month. As was the case in the sensitivity experiments, a separate network was trained for each five-minute look-ahead interval. Thus, 36 different networks were used to generate forecasts for the 36 five-minute intervals in the three-hour forecast period.

Figures 3-31 through 3-36 present selected performance statistics from this year of simulated short-term forecasting. The ME (bias) and MAE are expressed as a percentage of regional capacity. The capacity value used in the calculation of these statistics is based on the maximum available capacity reported for each region for any reporting interval during the calendar year of 2004. This may not be the same as the installed capacity for a region, since all of a region’s installed capacity may not have been available for any interval during the year and the installed capacity may have changed during the year. The maximum annual availability values used in the ME and MAE as a percentage of capacity calculations were: San Geronio 556.71 MW, Tehachapi 659.71 MW, Altamont 827.3 MW and Solano 171.4 MW, and the aggregated mean rated capacity of the four regions was 2215.2 MW.

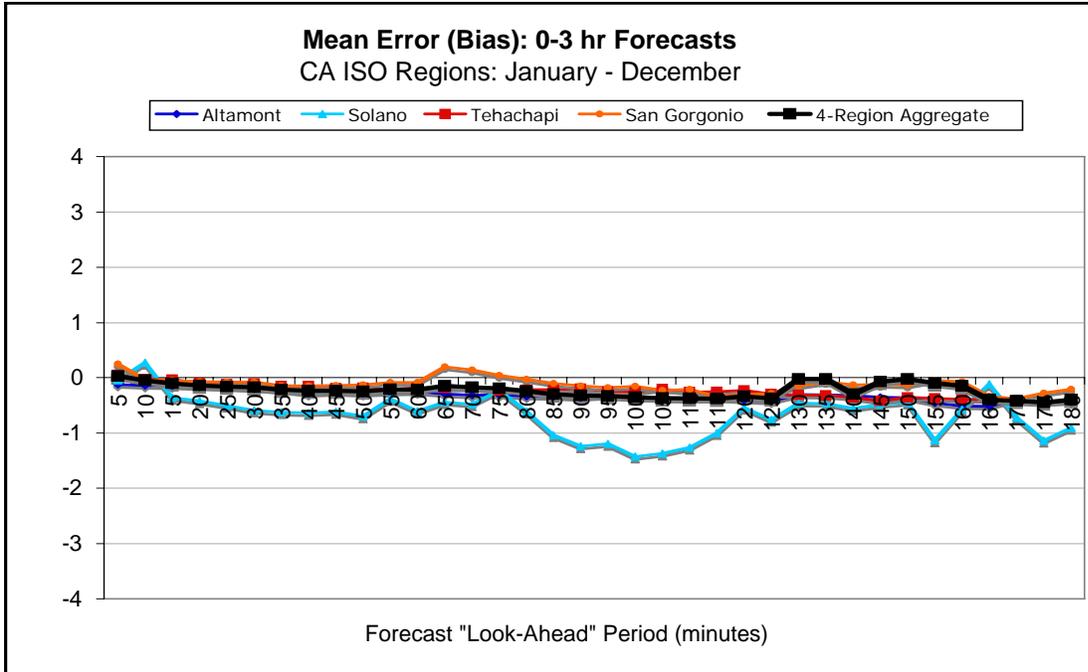


Figure 3-31 Mean error (bias) as a percentage of region capacity vs. look-ahead period for 2004 for ANN-based forecasts for each of four California wind resource areas and the aggregated areas.

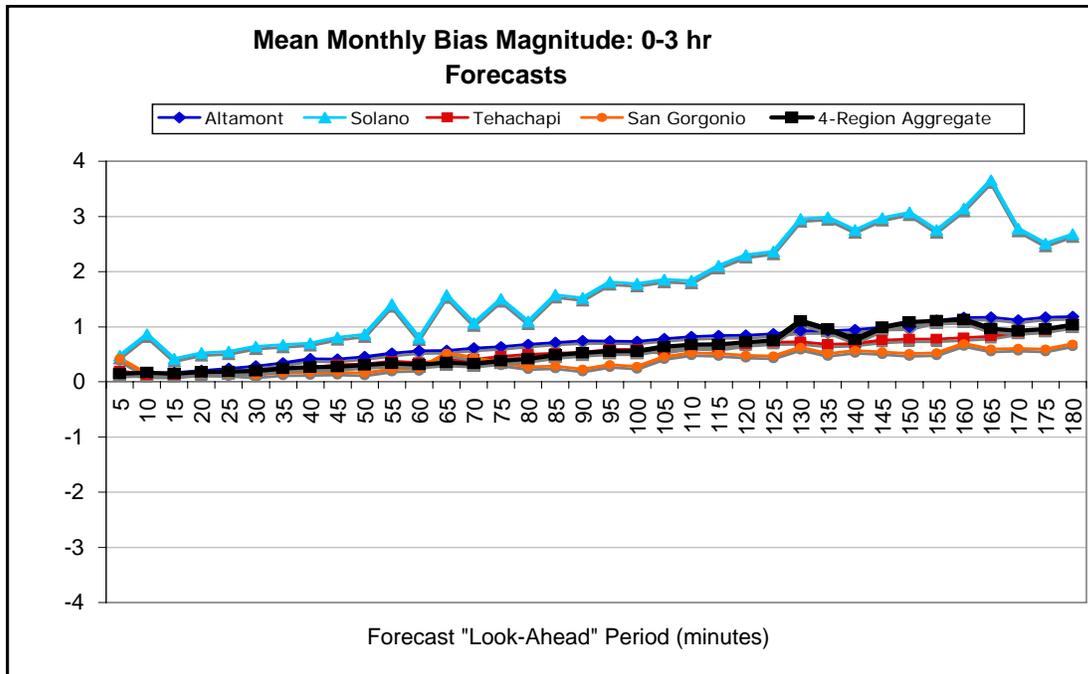


Figure 3-32 Average magnitude (% of capacity) of the monthly mean error (bias) vs. look-ahead period for 2004 for ANN-based forecasts for each of four California wind power production regions and the aggregated production for all four regions.

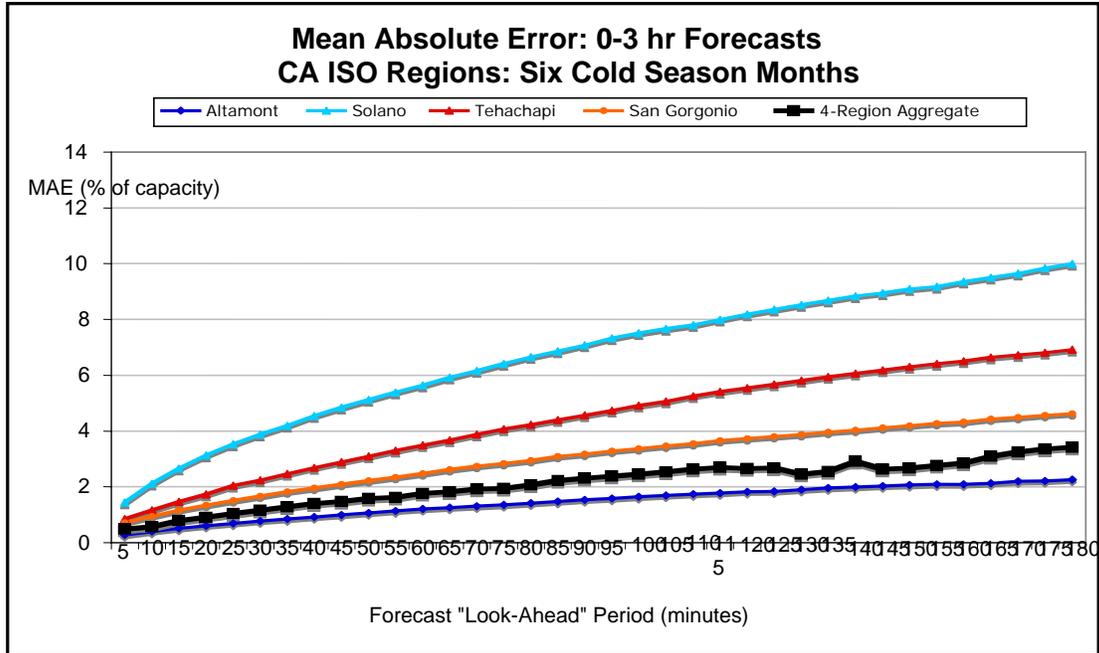


Figure 3-33 Mean absolute errors of ANN-power production forecasts (% of capacity) vs. forecast look-ahead period during the six cold season months (November through April), for the four California wind resource areas and the aggregated areas.

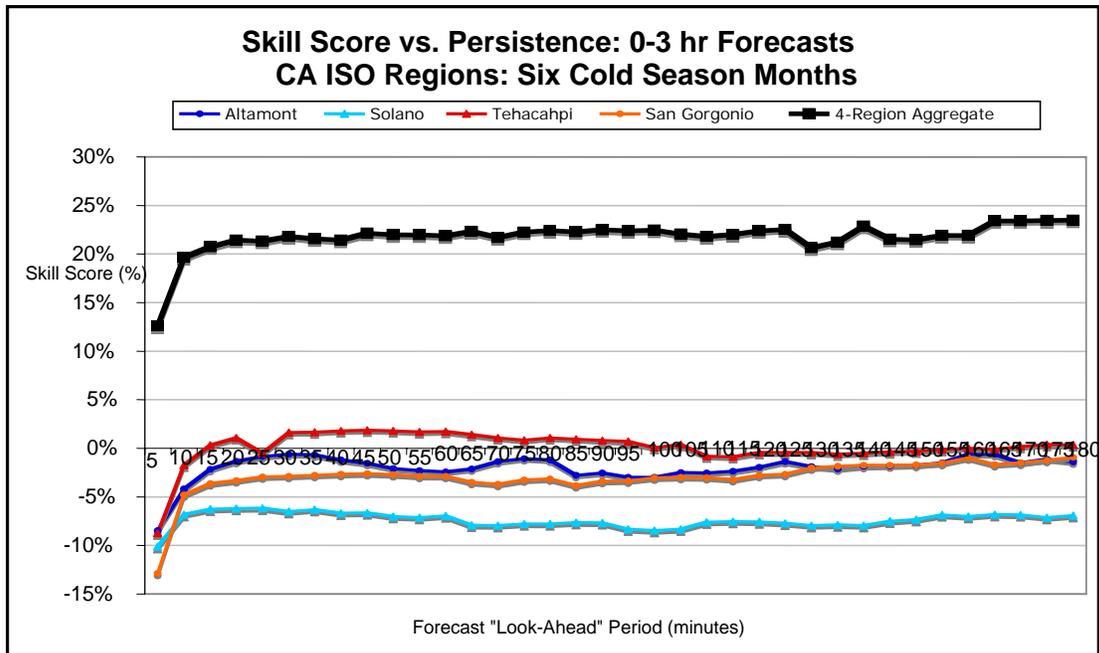


Figure 3-34 Skill scores of ANN-power production forecasts relative to persistence vs. forecast look-ahead period for the six cold season months (November through April) for the four California wind resource areas and the aggregated areas.

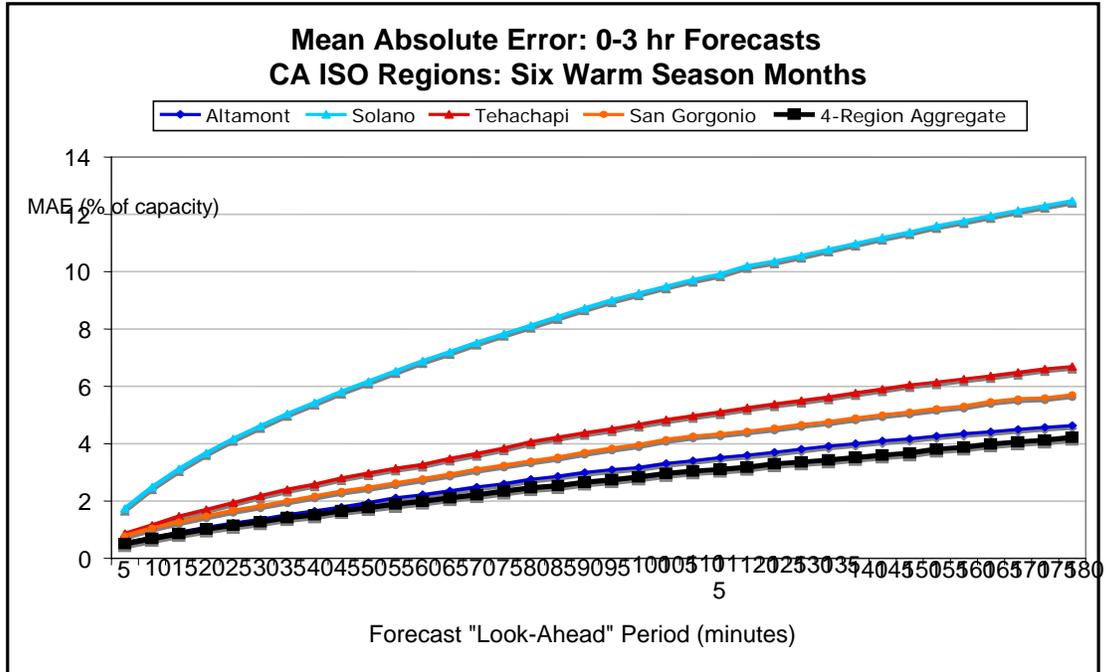


Figure 3-35 Mean absolute errors of ANN power production forecasts (% of capacity) vs. look-ahead period for the six warm season months (May through October) for the four California wind resource areas and the aggregated areas.

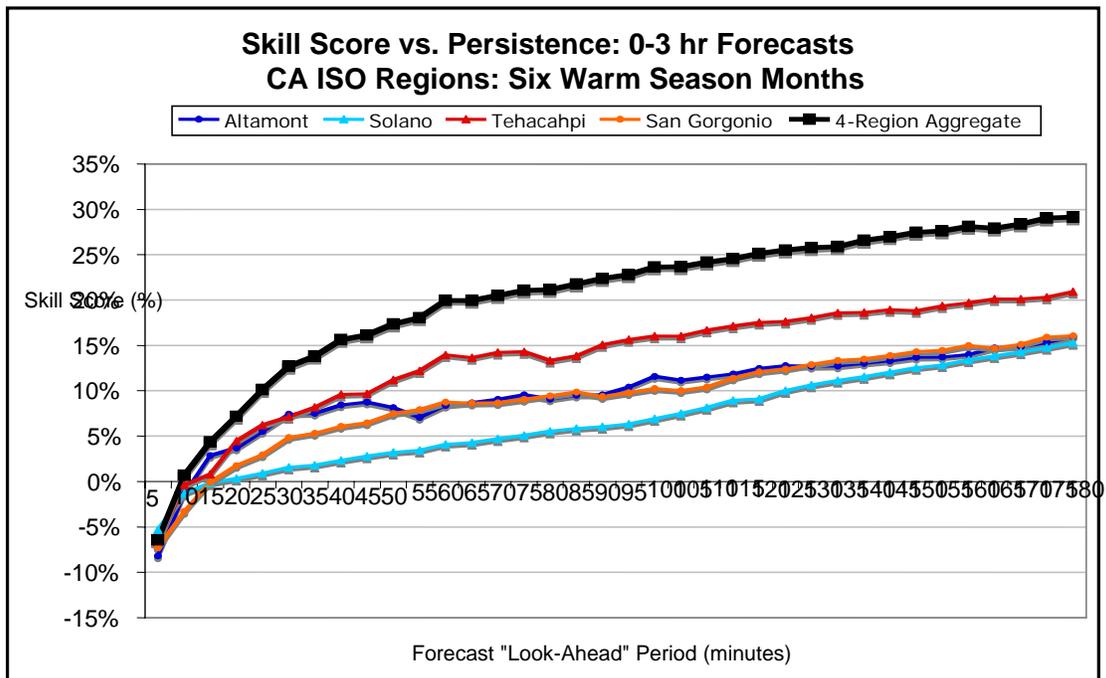


Figure 3-36 Skill scores of ANN power production forecasts relative to persistence vs. forecast look-ahead period for the six warm season months (May through October) for each of four California wind resource areas and the aggregated areas.

Figure 3-31 shows the forecast bias (mean error) for the year vs. look-ahead period for each region. This chart indicates that even though no bias compensation scheme was employed, the forecast bias for the year is quite low for most look-ahead periods and regions. The bias is generally near or below 0.5% of capacity for almost all forecast look-ahead periods for the San Geronio, Tehachapi and Altamont regions as well as the four-region aggregate. The exception is the Solano region. The magnitude of the annual mean error exceeded 1.0% for several forecast look-ahead periods for Solano.

For some market and grid management applications, the magnitude of the forecast bias over intervals shorter than one year is important. Therefore, the average magnitude of the monthly bias for each region and look-ahead period was computed. This is depicted in (Figure 3-32). The pattern is very similar for the San Geronio, Tehachapi and Altamont regions and the four-region aggregate. The average magnitude of the monthly forecast bias is near zero for all of these regions for approximately the first 30 minutes of the forecast period. After this time, the average magnitude of the monthly bias rises slowly and reaches the 1% level near the end of the three-hour forecast period. Once again the Solano region is the exception. The average magnitude of the monthly bias is larger for Solano for all look-ahead periods and rises much more rapidly after the first 30 minutes than it does for the other regions.

There was a substantial difference in the MAE and skill score values between the cold season and warm season months. This is not surprising since there is a marked difference in the characteristics of the warm season and cold season wind regimes over most of California. Therefore, the MAEs and skill scores for the forecasts are presented separately for the two seasons. The cold season is defined as November through April, and the warm season is the period from May to October. Figures 3-33 and 3-34 present the MAE and skill scores for the cold season months, and Figures 3-35 and 3-36 present those for the warm season months.

To understand the variation in forecast performance between the cold and warm seasons, it is necessary to understand the differences in wind regimes between the two seasons. During the cold season the wind over most of California is predominantly controlled by large-scale weather systems. The temporal patterns in the wind speed in most parts of California during the cold season are characterized by relatively long periods (several days) with low wind speeds and little variability that are interrupted by short periods with high wind speeds and considerable variability in the wind. The short periods of high winds are typically associated with large-scale storm systems moving through the state. The longer periods of relatively light winds are associated with the intervals between storms. In contrast, the variability of the wind during the warm season is mostly controlled by small-scale circulation resulting from the differences in the heating and cooling of the earth's surface (for example, land vs. water) on the regional scale. These circulations have a strong diurnal pattern associated with the daily cycle of daytime heating and nighttime cooling. The wind patterns tend to be similar each day, but the amplitude and phase of the diurnal cycles are modulated by often subtle changes in the larger scale atmospheric environment

The MAE and skill scores in Figures 3-31 through 3-36 indicate that there is indeed a fairly significant difference in the forecast performance between the two seasons. However, the

characteristics of the MAE differences between the two seasons vary substantially among the regions.

Solano Region

The Solano region is clearly an outlier in both seasons, although it is less of an outlier in the cold season. The MAEs for the Solano region are substantially higher than those for the other three regions in both seasons. In the cold season, the MAE for the Solano region rises from about 1.5% for a five-minute ahead forecast to approximately 10% for a three-hour forecast (Figure 3-34). The Solano MAEs are substantially higher for all look-ahead periods for the warm season (Figure 3-36). For the warm season the MAEs rise from just under 2% for a five-minute ahead forecast to about 12% for a three-hour ahead forecast. For both seasons the MAEs for the Solano region are more than twice as large as the average MAE for the other three regions. The higher MAEs for Solano appear to be attributable to two factors.

First, the capacity of the Solano region is substantially lower than that of the other regions. In addition, the Solano region hosts only two wind plants, which are next to each other. This combination means that there is a smaller number of turbines in the region, and their deployment has less geographic diversity. This results in a reduced “ensemble effect,” in which uncorrelated variations in the output of individuals or clusters of turbines offset each other to produce a less volatile aggregated output.

A second factor is that the winds have a higher degree of variability at the location of the two wind plants in this region than in many other wind production regions in California.

A third factor is that there appears to be a larger number of hours with wind speeds in the middle of the turbine power curve where the power production is most sensitive to changes in the wind speed.

Altamont Region

The behavior of the MAE is quite different for the Altamont region, which is only a short distance to the south of the Solano region. However, the rated capacity of the Altamont region is much greater than that of the Solano region (827 MW vs. 171 MW). During the cold season, the MAEs for the Altamont region rise from about 0.5% for a five-minute ahead forecast to approximately 2% for a three-hour ahead forecast. The MAEs for the warm season are somewhat higher, but well below those recorded for the Solano region.

Tehachapi Region

The MAEs for the two Southern California regions have more similarity to each other than those for the two Northern California regions. The Tehachapi region exhibits the least difference in MAEs between the two seasons. During the cold season, the MAE for the Tehachapi region rises from about 0.5% for a five-minute ahead forecast to just under 7% for a three-hour ahead forecast. The MAE values are slightly lower during the warm season, when they rise from about

0.5% for the 5-minute forecast to just over 6% for a three-hour ahead forecast. This is the only region for which the MAEs are not lower during the cold season.

San Geronio Region

The MAEs for the San Geronio region increase from about 0.5% for a five-minute forecast to just over 4% for a three-hour forecast for the cold season while they increase from 0.5% to about 6% for the warm season. The MAEs for the four-region aggregated power production forecasts are lower than the MAE for any individual region for the warm season and are lower than the MAEs for three of the four regions for the cold season. Only the Altamont region has a lower MAE for the cold season.

The MAEs for the cold season are generally lower because there are many forecast intervals during the cold season (the quiet periods between storms) for which the forecasted and the reported power production are zero because the wind speed is below the turbine start-up speed. These periods produce many intervals for which the forecast error is zero and reduce the overall MAE for those months. The short-term forecast MAE during the storm periods is relatively high, but these periods are relatively short and the overall MAE is more heavily weighted toward the errors during the longer quiet periods.

During the warm season, a substantial portion of the forecast performance variations are attributable to the diurnal cycle. The MAE is higher when significant changes in wind speed and production occur during ramp up and down periods, and it is lower the portions of the day when the wind speed and production are changing more slowly.

The skill score charts (Figures 3-34 and 3-36) depict a more dramatic difference between the seasons. The skill score is near 0% for all four regions during the cold season. The Tehachapi region has the highest skill score (generally 0 to +5%) for almost all look-ahead periods during the cold season. The Solano region has the lowest skill score for the cold season with skill scores of -5% to -10% for most look-ahead periods. Surprisingly, the skill score for the aggregated power production during the cold season is rather high with values slightly over 20% for most look-ahead times after the first 10 minutes. The skill scores for all regions and most look-ahead periods are much higher for the warm season (Figure 3-36). Once again, the Tehachapi region has the highest skill scores, and the Solano region produced the lowest scores.

The bottom line from the one-year of testing of the autoregressive component of the short-term forecast system is that the MAEs are higher during the warm season than during the cold season, but the skill scores were much higher during the warm season. The forecasts of the aggregated power production for all four regions had considerable skill in both the warm and cold seasons, although the skill was somewhat higher during the warm season.

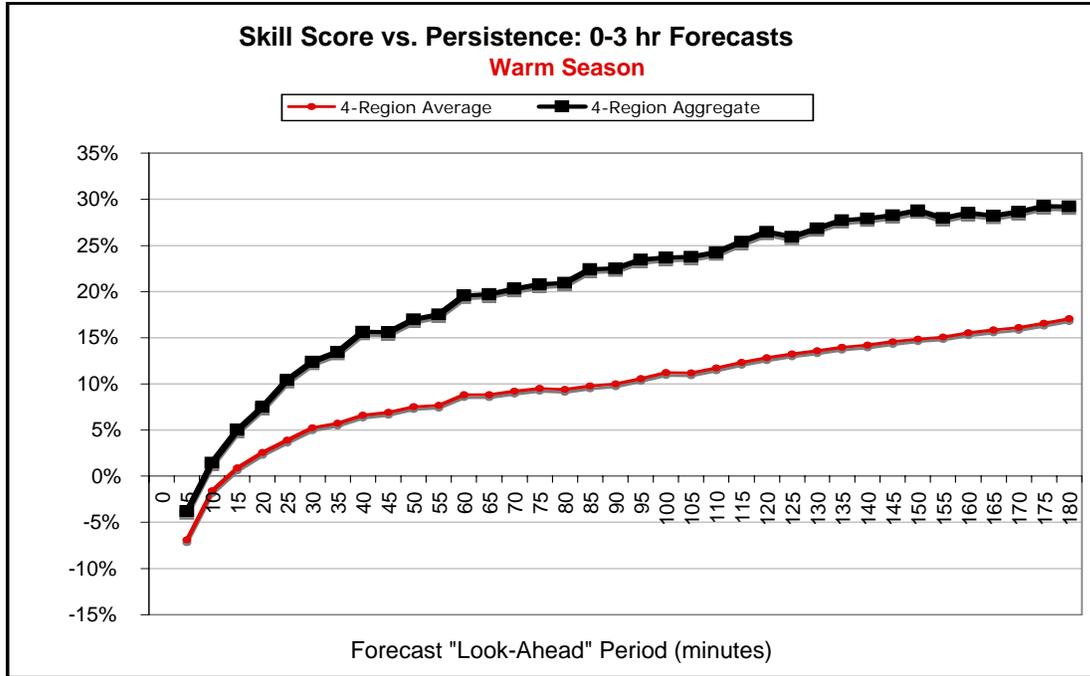


Figure 3-37 Average skill scores of the regional and aggregate ANN forecasts relative to persistence vs. forecast look-ahead period during the six warm season months for the four California wind resource areas.

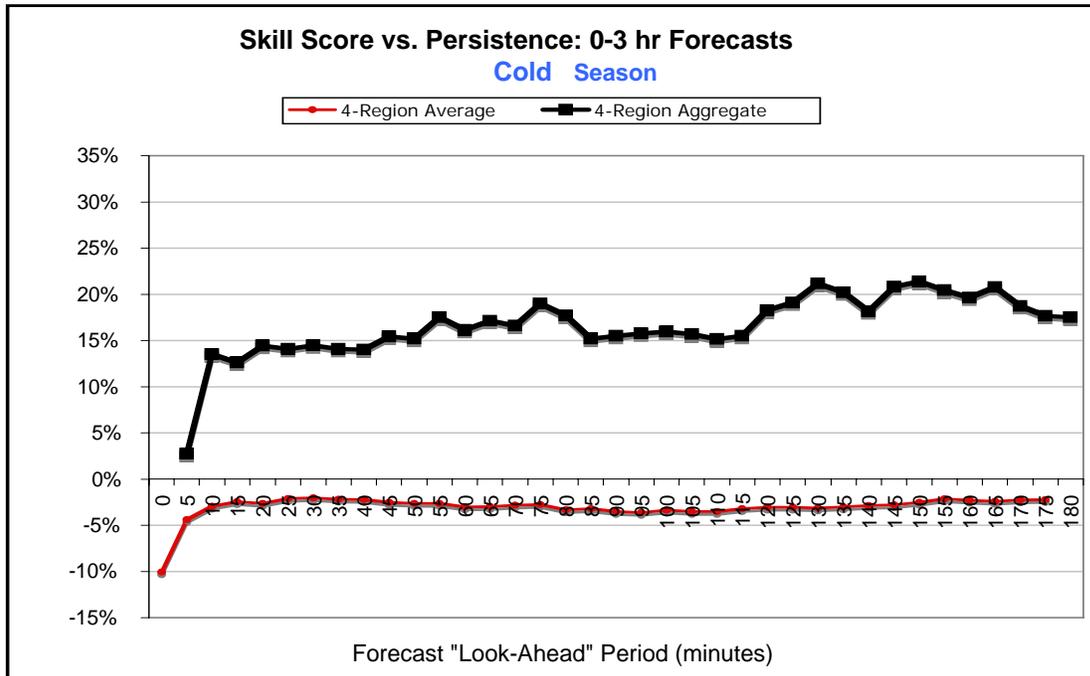


Figure 3-38 Average skill scores the regional and aggregate ANN forecasts relative to persistence vs. forecast look-ahead period during the six cold season months for the four California wind power resource areas.

Conclusions

The key results of the short-term forecast system development and testing are:

1. Review of statistical and physics-based forecast methods and analysis of the spatial and temporal characteristics of the wind speed and direction variability to provide the foundation for the design of a robust short-term forecasting system capable of producing forecasts of the 5-minute regional energy production for the next 3-hour period updated every 5 minutes;
2. Design of a two-stage forecast system with the first stage consisting of a mini-ensemble of three forecast methods that each produces an independent power production forecast, and a second stage that weights each of the three forecasts from the first stage based on their recent performance to produce a single composite forecast and an estimate of forecast uncertainty; and
3. Initial testing of one of the three forecast methods (the autoregressive method) in the first-stage of the proposed forecast system with one year (2004) of regional wind power production data supplied by the CA ISO.

The main conclusions are:

1. The autoregressive component of the three-method first stage of the proposed forecast system produced a reduction in the regional power production forecast error of a persistence forecast for the warm season (May –October) that ranged from near 5 to 10% for the first 30 to 60 minutes of the 3-hr forecast period to the 15% to 20% range during the latter stages of the period; and
2. The autoregressive method showed no improvement over persistence for the cold season, but that is to be expected due to the character of the wind during the California cold season and is one of the reasons for incorporating two other methods into the forecast system.

Recommendations

The recommended next steps are:

1. Implement and evaluate the two other component forecast subsystems and the ensemble compositing subsystem that were not implemented in this project;
2. Implement and evaluate the performance of the entire forecast system for several months in different seasons for each of the wind resource areas in California; and
3. Test the system in an operational environment and obtain feedback from CA ISO personnel.

Another approach would be to test and implement the same subsystem tested in this project and then develop the other components while the single subsystem forecast system is in production.

The results produced in this project suggest that this approach could yield a 5% to 20% improvement over a persistence forecast, especially during the warm season.

4

NEXT-DAY WIND PLANT ENERGY FORECASTING SYSTEM DEVELOPMENT AND TESTING

Introduction

The objective was to formulate and test new forecast methods and datasets that have the potential to improve the forecast performance of one- to 48-hour wind power production forecasts relative the performance in the previous Energy Commission-EPRI wind energy forecasting project (Energy Commission and EPRI, 2003a and 2003b).

Previous California Wind Energy Forecasting Project Results

In the previous project, 48-hour forecasts were generated for two California wind plants twice each day and for one year (October 2001 to September 2002) by two forecast providers: AWS Truewind, LLC (AWST) and Risoe National Laboratory (Risoe) of Denmark. Each forecast provider generated and delivered two 48-hour forecasts per day of the hourly average power production (kW) and the average hourly wind speed and direction for one meteorological tower for each wind plant. A morning forecast was delivered at 8:00 a.m. PST. The first forecast hour from this forecast was the hour ending at 9:00 a.m. PST and the forecast period extended until the hour ending at 8:00 a.m. PST two days later. The evening forecast was delivered 12 hours later at 8:00 p.m. PST and had a forecast period that was offset by 12 hours from the morning forecast.

It is important to note that these forecasts were generated in a “next-day” mode. This meant that the forecasts were produced without real-time data from the wind plants. The real-time data from the wind plants is very important for the performance of forecasts during approximately the first six to nine hours of the forecast period. However, the real-time data from the plant has little impact the quality of the forecasts after this period.

The two wind plants that participated in this project were the Mountain View 1 and 2 wind plant in the San Gorgonio Pass of Southern California and the PowerWorks plant located in the Altamont Pass. The rated capacities of the Mountain View and PowerWorks wind plants are 66.6 and 90 MW, respectively. Because the AWST forecast performance was as good as or better than that of Risoe, the AWST forecasts from the previous project are used as the baseline to measure improvements in forecast performance in the current project.

Tables 4-1 and 4-2 present the forecast performance statistics for AWST’s *eWind* forecast system from one year of forecasting in the previous project are listed in. The performance statistics include the monthly and annual mean absolute errors (MAE) of the wind energy

and speed forecasts and the skill scores of the forecasts relative to persistence and climatology (Skill- P and Skill-C). The skill score is the percentage reduction of the MAE relative to the MAE for the persistence or climatology forecast; Skill Score = $(1 - \text{MAE}/\text{MAE-P}) \times 100\%$.

Table 4-1 eWind monthly and annual forecast performance statistics for the Mountain View wind plant from the previous Energy Commission-EPRI project (Energy Commission PIER EPRI, 2003)

Month	Power Forecast				Wind Speed - M V I - 1		
	MAE (kW)	%MAE (1)	Skill-P	Skill-C	MAE (m/s)	Skill-P	Skill-C
Oct-01	10,295	15.5%	44.2%	44.9%	2.68	42.6%	31.6%
Nov-01	9,293	14.0%	34.0%	47.4%	2.68	42.2%	43.1%
Dec-01	10,366	15.6%	31.5%	29.9%	3.50	28.9%	9.5%
Jan-02	6,918	10.4%	48.1%	50.4%	2.47	43.3%	28.7%
Feb-02	9,446	14.2%	27.4%	56.1%	3.27	24.1%	28.4%
Mar-02	10,995	16.5%	45.2%	49.2%	4.17	39.0%	29.6%
Apr-02	15,095	22.7%	22.3%	18.8%	3.68	36.3%	32.2%
May-02	9,523	14.3%	49.7%	40.8%	2.73	57.9%	41.6%
Jun-02	10,786	16.2%	42.3%	27.9%	2.58	54.2%	42.9%
Jul-02	12,433	18.7%	41.0%	22.1%	2.93	44.0%	35.0%
Aug-02	13,346	20.0%	35.1%	14.9%	2.83	43.1%	36.4%
Sep-02	13,679	20.5%	18.8%	33.1%	3.15	26.0%	16.3%
Annual	11,037	16.6%	37.5%	36.4%	3.05	41.0%	31.7%

Table 4-2 eWind monthly and annual forecast performance statistics for the PowerWorks wind plant from the previous Energy Commission-EPRI project (Energy Commission PIER-EPRI, 2003)

Month	Energy - All Clusters				Wind Speed - Tower #438		
	MAE (kWh)	% MAE (1)	Skill-P	Skill-C	MAE (m/s)	Skill-P	Skill-C
Oct-01	13,725	15.3%	34.6%	37.6%	1.92	32.5%	25.8%
Nov-01	5,664	6.3%	-9.8%	43.2%	1.91	34.4%	22.8%
Dec-01	6,319	7.0%	3.9%	38.1%	2.44	23.5%	0.3%
Jan-02	2,167	2.4%	28.6%	76.4%	2.08	8.3%	30.2%
Feb-02	2,234	2.5%	35.2%	77.4%	1.91	23.4%	31.6%
Mar-02	10,372	11.5%	28.6%	36.5%	2.17	31.0%	22.7%
Apr-02	18,350	20.4%	30.7%	18.9%	2.11	30.8%	15.8%
May-02	20,567	22.9%	36.2%	21.5%	1.89	38.3%	18.9%
Jun-02	16,445	18.3%	39.8%	30.0%	1.88	38.2%	20.9%
Jul-02	17,676	19.6%	31.0%	10.1%	1.55	28.0%	8.0%
Aug-02	17,248	19.2%	31.0%	32.9%	1.63	28.2%	21.9%
Sep-02	18,734	20.8%	20.2%	17.2%	1.69	31.6%	25.5%
Annual	12,702	14.1%	30.9%	30.9%	1.93	29.9%	20.9%

These performance statistics represent all forecast hours (one to 48 hours) and both forecast cycles (morning and afternoon). Thus, the statistics for a 30-day month are based upon a forecast data sample consisting of 48 forecast hours per forecast cycle, two forecast cycles per day, 30 days per month. This yields a total sample size of 2,880 hours for a 30-day month. However, plant power production and meteorological data were not available for all of these hours due to various communication and measurement instrumentation problems.

Therefore, the monthly and annual statistics are based on (in some cases substantially) less than the maximum number of hours.

Mountain View 1 and 2 Wind Project

In the previous project, the AWST forecasts achieved an annual mean absolute error (MAE) of 16.6% of the 66.6 MW rated capacity for the Mountain View plant. The monthly MAE values ranged from 10.4% in January 2002 to 22.7% in April 2002. A large amount of the variability in the monthly MAE values is related to the relationship between the error in the power production forecasts and the forecasted power production due to the nature of a typical turbine power curve.

This relationship can be understood by considering a typical power curve for a wind turbine. Figure 4-1 shows an example power curve for a GE 1.5 MW turbine, which has the same general shape as that of almost all wind turbine power curves. The power curve shape indicates that the sensitivity of the predicted power production to errors in the prediction of the wind speed varies over the wind speed range. At wind speeds less than 6 m/s near the lower end, and greater than 11 m/s at the upper end of the wind speed range, a 1-m/s error in the wind speed prediction results in a much smaller error in the power production than when the wind speed is in the steeply sloped portion (within the 6- to 11-m/s range) of the power curve. Thus, forecasts that have the same error in the wind speed prediction will generate larger power production errors if the wind speeds are in the steeply sloped portion of the power curve.

Figure 4-2 illustrates the impact of wind speed prediction error on the forecast errors in the previous Energy Commission-EPRI project. It shows the MAE for each of 20 forecasted power production bins for all the verified forecast hours for the Mountain View plant during Oct 2001 through Sept 2002. Each bin has a width of 5% of installed capacity. The MAE is below 15% for forecasted production levels that are below 10% or more than 80% of installed capacity (near the bottom or top of the power curve). However, the MAE is above 20% when the forecasted production is between 15% and 65% of installed capacity the middle of the power curve). This chart indicates that there is a factor of two or more difference between the typical MAE of a forecast near the lower or upper end of the power curves and those in the middle of the curve.

This relationship is responsible for a substantial portion of the monthly variability of the MAE. For example, during January the number of hours with a forecasted production at the low end of the power curve is much higher than in July. Therefore, the MAE will tend to be much lower in January than in July. A good example of the impact of this effect can be seen by comparing the wind speed and power production MAE for February 2002 and August 2002. The wind speed MAE for August (2.83 m/s) is 0.44 m/s lower than the MAE for February (3.27 m/s). However, the MAE of the power production forecasts is 20.0% of rated capacity for August and only 14.2% for February. Thus, based on MAE, the wind speed forecasts are much better in August, but the power production forecasts are much better in February. As one might expect from the above discussion, the explanation is that there are many more hours in February with wind speeds near the bottom end of the power

curve (and hence creating less sensitivity to error) than during August. Thus, when analyzing the performance of a wind power production forecast system, one must consider the distribution of the forecasted hours on the plant-scale power curve for the particular months under consideration and look at the relative errors of the wind speed and the power production forecasts.

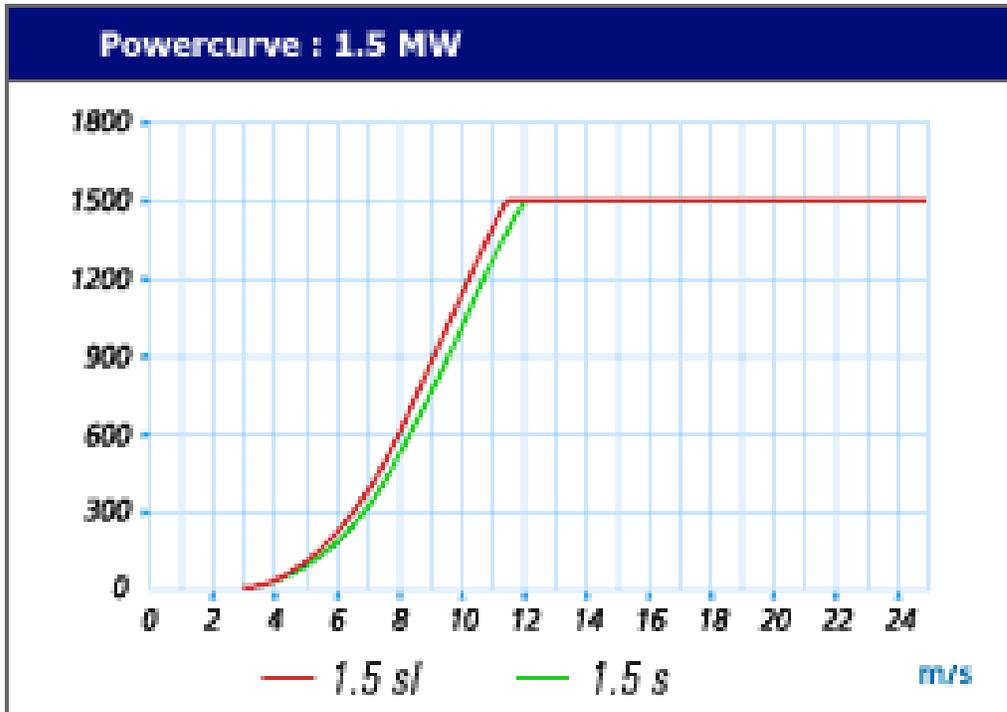


Figure 4-1 Power curve for a GE 1.5 MW turbine (from the GE Power website at http://www.gepower.com/prod_serv/products/wind_turbines/en/15mw/tech_data.htm)

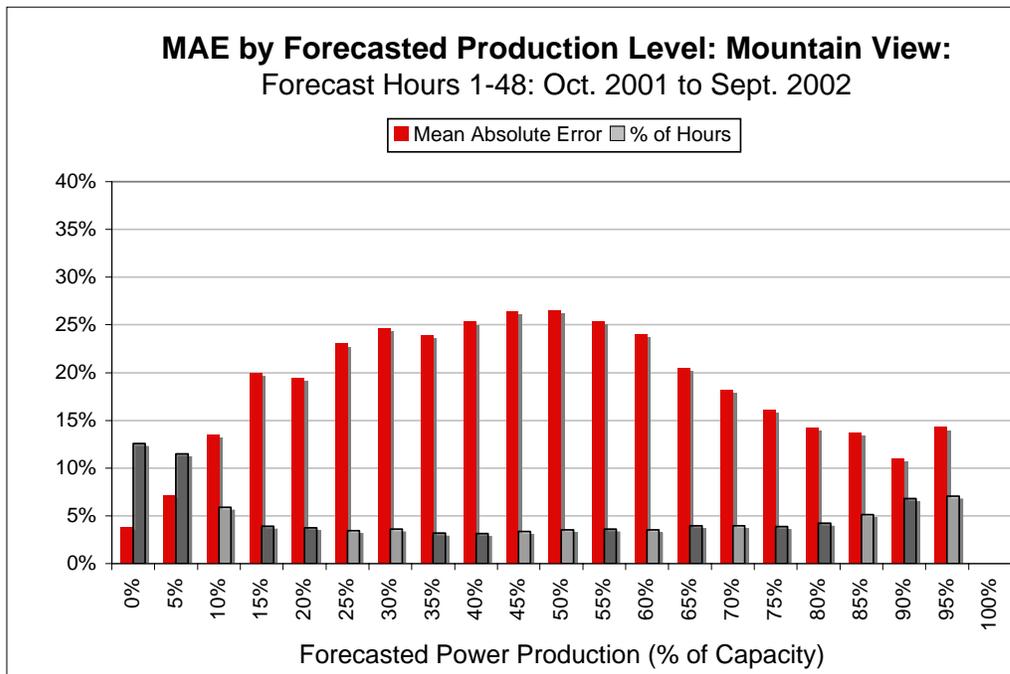


Figure 4-2 The MAE of generation forecasts (% of installed capacity) and the percentage of hours for 20 forecasted power production bins for the Mountain View wind plant. Each bin is labeled by the forecasted production at the upper end of the bin.

PowerWorks Wind Project

The forecast performance for the PowerWorks wind plant was generally similar to that for the Mountain View plant. The annual MAE of the power production forecasts 14.1%, which is slightly lower than that achieved for the Mountain View plant. However, the range was substantially greater. The lowest monthly MAE was 2.4% during January 2002, and the highest was 22.9% for May 2002. Interestingly, the MAE of the wind speed forecasts for meteorological tower M438 was lower for May (1.89 m/s) than it was for January (2.08 m/s). In part, this is another example of the power curve effect discussed in the previous paragraphs. However, it is also a result of the fact of the greater variability of the wind speed within the area of the PowerWorks facility during the warm season than during the cold season.

Screening of Improved Data and Forecast Methodologies

The objectives were to identify and test candidate forecasting data and methodologies to determine the impact of each improvement on forecast performance and to identify those that offer the greatest potential for improved performance for a wide range of weather regimes and locations. The methodologies were tested using the data from the previous Energy Commission-EPRI project for the Mountain View 1 and 2 and PowerWorks wind projects.

It was not practical to execute a large number of forecast method experiments for the full 12 months of the previous forecast evaluation period for both wind plants. Therefore, it was decided to test each forecast system improvement using a three-month data sample for each of the two wind plants. To further reduce the computational task, it was decided that only the morning forecast cycle would be used in the test period. This was considered to be a reasonable approach because the performance of the afternoon cycle is generally highly correlated with the performance of the morning cycle. Therefore, little additional information would be gained by testing the changes on both the morning and evening forecast cycles for the same months.

The test months were selected independently for the PowerWorks and Mountain View plants. The two main selection criteria were that (1) plant data were available for a large fraction of the hours in the months (a low level of missing data), (2) the forecast performance was below the average for that plant; and (3) the months should be distributed so that there is a winter, summer, and transition-season month in the test sample for each plant.

Table 4-3 shows the test months selected for each plant. December 2001 was the winter month selected for both plants. This was the winter month with the highest power production and wind speed MAE for both plants. The transition month was May 2002 for PowerWorks and April 2002 for Mountain View. In each case, this was the month of the

highest monthly MAE for each wind plant. The summer test month was July 2002 for PowerWorks and August 2002 for Mountain View.

Table 4-3 Test wind resources and months for evaluation of candidate data and methodologies.

Mountain View						
Month	Power Production			Wind Speed @ Catellus		
	MAE (kW)	MAE (%)	Skill - P	MAE (m/s)	MAE (%)	Skill - P
Dec-01	9,657	14.6%	35.0%	3.43	54.9%	32.7%
Apr-02	15,638	23.7%	29.2%	3.69	26.1%	41.9%
Aug-02	13,199	20.0%	41.7%	2.73	23.0%	49.2%
PowerWorks						
Month	Power Production			Wind Speed @ Tower 438		
	MAE (kW)	MAE (%)	Skill - P	MAE (m/s)	MAE (%)	Skill - C
Dec-01	6,579	7.3%	1.2%	2.51	46.4%	18.5%
May-02	20,370	22.6%	38.2%	1.83	23.3%	45.6%
Jul-02	18,083	20.1%	40.0%	1.54	19.8%	32.3%

In the case of PowerWorks, the performance for the three summer months was fairly similar, but July had a slightly higher power production forecast MAE, and the skill scores for both the power production and wind speed forecasts were lower than for the other summer months. For Mountain View, August 2002 was selected because it had the highest power production forecast MAE and lowest skill score for the summer months.

Table 4-3 also shows the power production and wind speed MAEs and skill scores for the morning forecast cycle from the previous project. These provide the reference points for comparison of the forecast performance of the candidate methodologies in the current project.

Forecasting Experiment Plan

The forecast system experiments were designed to test possible improvements to each part of the forecast system. Figure 4-3 is a schematic overview of the components of the *eWind* forecast system. The system is divided into two major components: data and numerical models. There are three fundamental types of data used in the forecasting process: (1) regional weather data; (2) time series of power production and meteorological data from the wind plant; and (3) off-site local meteorological data. There also three major types of numerical models used in the forecast system: (1) physics-based atmospheric models; (2) statistical prediction models and (3) a plant output model.

There are, of course, almost an infinite number of different configuration and new techniques that can be implemented within the general framework of the *eWind* system depicted in Figure 4-3. Obviously, it was possible to test only a very small subset of the possible changes within the resource limitations of this project. The improvements tested within this task were selected based upon two criteria: (1) those that had the greatest

promise of improving the forecasts; and (2) those that could be tested with the data available for this project and within the resource limitations of the project.

Table 4-1 lists the six focus areas of the forecasting improvements that were tested.

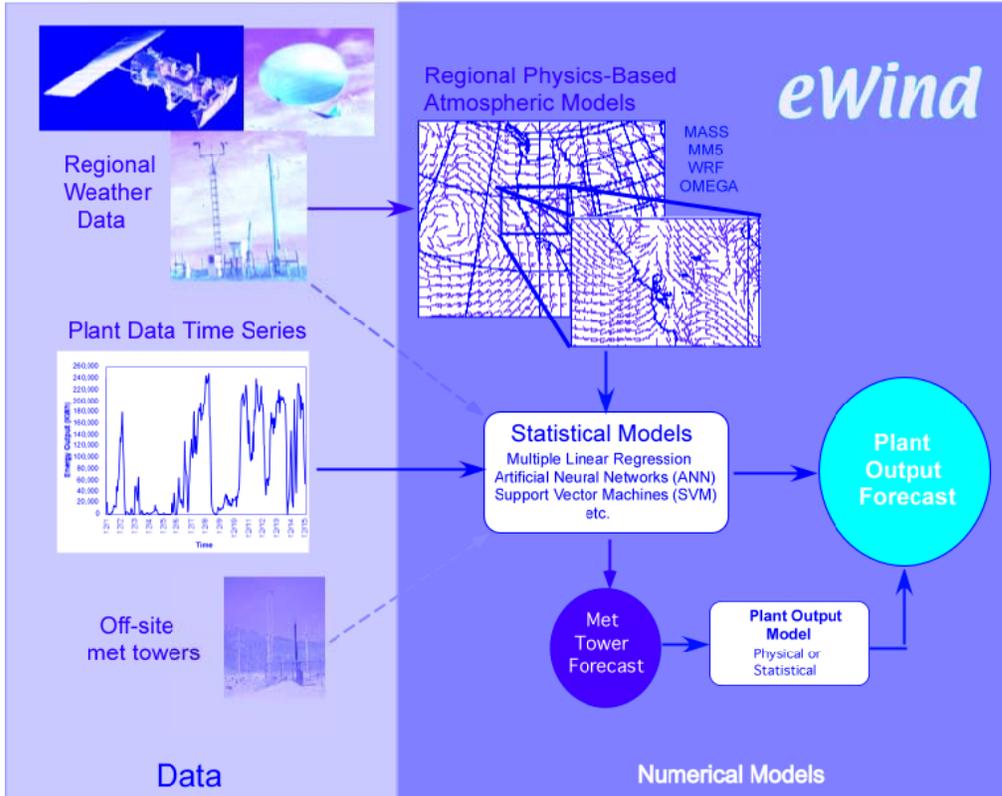


Figure 4-3 A schematic depiction of the main components of the *eWind* forecast system.

Table 4-4 Focus areas and associated experiments.

Focus Area	Experiments
(1) Additional or improved input data for the physics-based simulations	4-km MODIS and Pathfinder Water Surface Temperature (WST) was used to initialize the physics-based model
(2) Higher resolution (i.e. smaller grid cells) for physics-based model	4-km and 1-km physics-based simulations were executed
(3) “Next-generation” physics-based models	Forecast simulations were executed with the WRF model
(4) Advanced statistical models for MOS	Stratified multiple linear regression and artificial neural network methods were used for MOS in place of the screening multiple linear regression
(5) More sophisticated formulation for plant output model	Median-based power curve plus residual formulation was used
(6) Forecast Ensembles	An ensemble of MASS and WRF model forecasts was used in place of the single MASS forecast

Focus Area 1 addresses the input data into the physics-based model. There are many emerging datasets that have the potential to improve the simulation of winds in physics-based models. In this project the impact of higher-resolution water surface temperatures from the current generation of satellite sensors was evaluated.

Focus Area 2 assessed the impact of the resolution of the physics-based model grid. Smaller grid cells mean that the physics-based model can simulate more of the local-scale weather features that determine the evolution of the wind at the wind plant.

Focus Area 3 assesses the impact of employing the “next generation” of physics-based atmospheric models in place of the MASS model currently used in the *eWind* system.

Focus Area 4 examines the impact of using more sophisticated statistical techniques for the Model Output Statistics (MOS) component of the forecast system.

Focus Area 5 investigates the sensitivity of the forecast performance on the formulation of the plant scale power curve.

Finally, Focus Area 6 examines the impact of employing a forecast based on a composite of individual forecasts from different methods.

Forecasting Experiment Results

This section documents the results of the forecast performance experiments in each of the six focus areas. The results are reported below in separate sections for each focus area.

Focus Area 1: Water Surface Temperature (WST)

The strongest power-generating winds in California typically occur during the warm season because they are driven by the large-scale sea breeze circulation created by the contrast between the relatively cool Pacific Ocean and the warm, usually dry interior. The strength of the circulation is directly related to the magnitude of the temperature difference, so the day-to-day temperature variations of both water and land surfaces are important. Some of the most productive wind energy regions are close enough to the ocean that the onshore flow during the day sometimes brings a distinct marine boundary layer to the wind generation area (for example, Altamont Pass and San Geronio Pass). Variations in the temperature and depth of these marine layers are likely to have a considerable effect on the wind speeds through these mountain passes. In addition to the ocean itself, the temperatures of bodies of water such as the San Francisco Bay and the Salton Sea may be significant.

Water surface temperatures are also a factor in the formation and transport of the marine stratocumulus cloud layer, which plays an important role in modulating the sea breeze circulation. When the marine cloud layer moves onshore in the San Francisco Bay Area or the Los Angeles basin, the winds in the Altamont and San Geronio Passes tend to strengthen. For all of these reasons, accurate estimates of sea surface temperatures are necessary for the correct simulation of power-generating winds in California.

Along with many other surface and atmospheric properties, physics-based atmospheric models require the specification of an initial water surface temperature (WST) field. It has long been recognized that the accurate representation of spatial and temporal variations of WST is very important for the correct simulation of sea breezes, coastal cyclogenesis, and a range of other phenomena. The simplest way to specify WSTs for a physics-based model simulation is to access a database of climatologically averaged values. Weekly or monthly climatological WST datasets are available from a variety of sources – NCAR, NCEP, USGS, the Naval Research Laboratory, and others. The obvious weakness of climatological data is that actual WSTs can vary significantly from climatological values on global (for example, El Niño-Southern Oscillation) scales, and transient smaller-scale features can produce significant atmospheric effects.

A better option is to use WST fields based upon measurements at or just before a forecast simulation is initialized. This approach permits WST anomalies to be included in the model's initial state. A dataset often used in meso-scale models is the [NOAA Optimum Interpolation Sea Surface Temperature Analysis](#), which uses available *in situ* and satellite SST observations to perform an objective analysis on a 1-degree (~110 km) global grid (Reynolds and Smith, 1994). These analyses have been archived from 1981 to the present, and weekly and monthly analyses are available in near-real time. Figure 4-4 shows an image from this dataset.

Beginning in the mid-1980s, global AVHRR Oceans Pathfinder SST datasets have been derived from Advanced Very High Resolution Radiometer (AVHRR) instruments, which have been

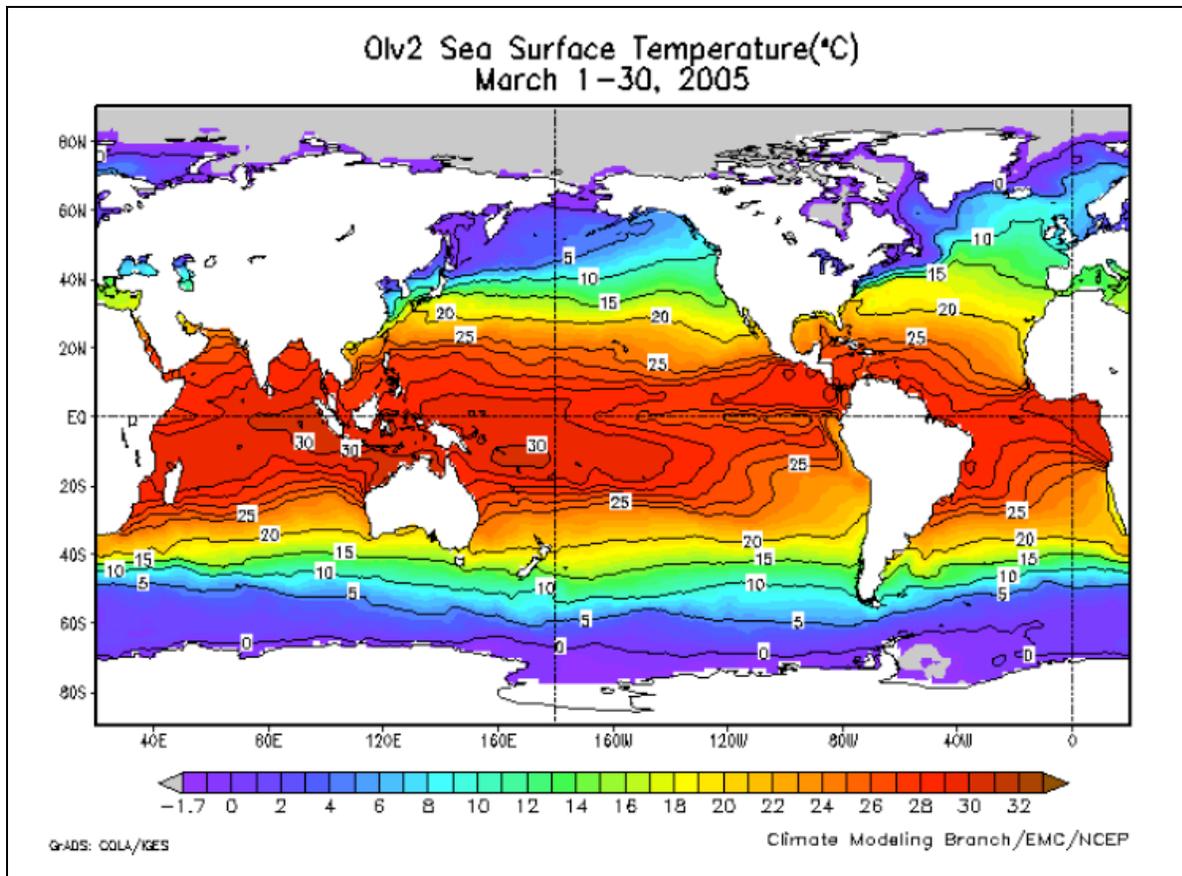


Figure 4-4 A graphical depiction of the NOAA optimum interpolation WST analysis on a 1-degree global grid for March 2005.

carried aboard a series of NOAA polar-orbiting satellites (Kilpatrick et al., 2001). A joint effort by NOAA, NASA, and the University of Miami has resulted in successive versions of Pathfinder datasets that have been quality-controlled and distributed in a variety of resolutions as high as 4 km. Temporal averaging is done at daily, eight-day and monthly intervals. A significant problem with purely satellite-derived datasets (as opposed to objectively-analyzed fields such as the NCEP OI) is that some data pixels are flagged as missing, or classified as low quality. These quality problems can be caused by issues such as persistent cloud cover, problems with atmospheric aerosols, and uncertainty regarding sea ice coverage. A model using this kind of dataset must include logic to override questionable satellite-derived values with climatological or other information. NOAA has a program called CoastWatch, which distributes real-time AVHRR SST data for U.S. regions at resolutions as high as 1 km. At least one meso-scale modeling group is preparing composites of these data and using the data to initialize WSTs in MM5 simulations.

A similar but improved satellite-derived WST source, the MODIS Sea Surface Temperature dataset, has been available since 2002. The MODerate Resolution Imaging Spectroradiometer (MODIS) instrument (Brown and Minnett, 1999) has been

aboard NASA's Terra and Aqua satellites. Building on knowledge gained from the earlier AVHRR missions, the MODIS instrument and algorithm produce more accurate values of WST. Datasets are produced on a similar global grid for a similar range of resolutions, as high as 4 km. Cloud contamination and other issues continue to result in flagging of some areas and time periods. Figure 4-5 is an example MODIS image, and it illustrates the fine-scale structure of the WST field.

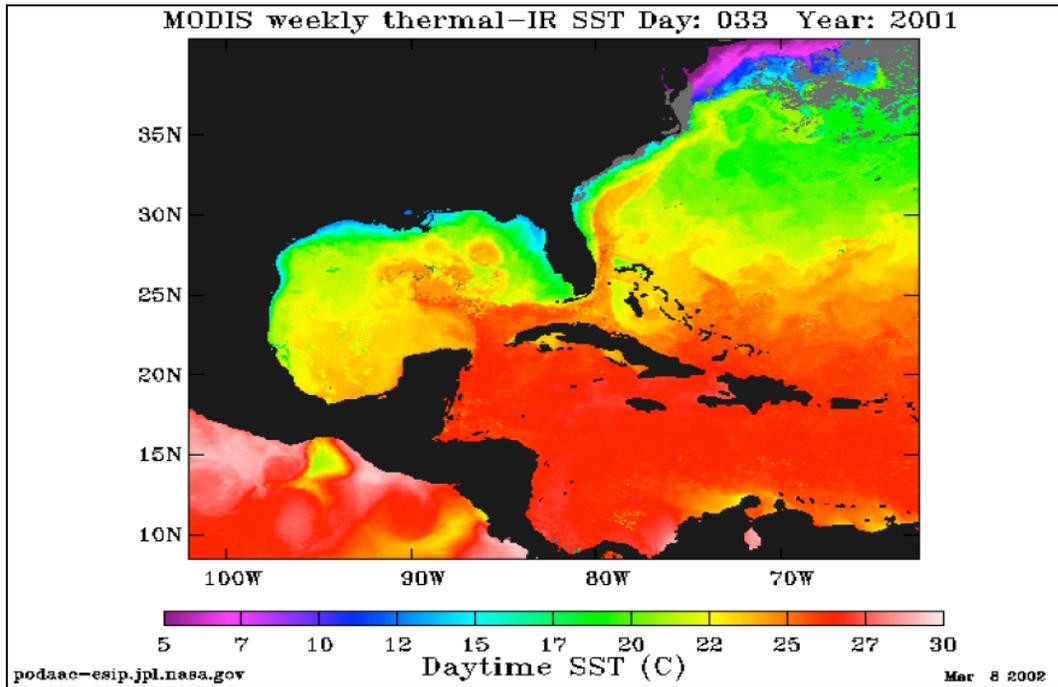


Figure 4-5 Sample MODIS water surface temperature image.

During the previous wind power production forecast project, the NCEP OI database of historical WST at one-degree (about 110 km) resolution and weekly intervals was used to specify the WST fields. The one-degree NCEP OI data is too coarse to properly resolve important gradients of WSTs in inland lakes and coastal oceans. The result is that the water surface temperature for lakes as large as the Salton Sea in Southern California is poorly known, and the initialization process may make an assumption that differs significantly from reality.

In an attempt to assess the impact of WST specification on the day-ahead power production forecasts, a new set of day-ahead power production forecasts that utilized the new higher resolution specification of the WSTs were produced for the PowerWorks and Mountain View wind plants for the three test months at each plant. The new set of forecasts was based upon the use of either the (1) AVHRR Pathfinder global data at 4-km resolution, which is available from the mid-1980's to the present; or the (2) Aqua MODIS global data at 4-km resolution, which is available from July 2002 to the present. The Pathfinder data was used for the test months before July 2002, and the MODIS data was used for July 2002

and later months. Both datasets can be downloaded from a NASA ftp site in the form of global Hierarchical Data Format (HDF) files.

The most common problem with satellite-derived data is that persistent cloud cover over some locations may make it difficult for the instrument to cleanly infer WST, and instrument problems can sometimes result in missing data for some areas. One issue is that these datasets consist of both weekly and monthly composites. After examining the MODIS weekly data for July 2002, it was discovered that the data quality was very poor for one of the weeks. This would result in a poor initialization of WST if used in real-time forecast runs. It is unclear whether the missing WST values were caused by persistent cloud cover, instrumentation problems, or something else. The WST data in the monthly file were much more complete; the monthly data should probably be preferred when both are available. Another problem was encountered – only a fraction of the MODIS WST data values were valid for the grid points over the Salton Sea of Southern California; the rest of the grid points have flagged values. To prevent a subsequent surface temperature analysis from spreading land temperature values (which might differ significantly from water temperatures) over the flagged portions of the Salton Sea, a change was made to the model's data preprocessor to spread the valid WST values to nearby water points before doing an analysis of surface temperature.

Figures 4-6 through 4-8 illustrate how the NCEP OI and MODIS datasets may differ. These figures depict the WST field used to initialize a physics-based model forecast simulation at 0000 UTC August 23, 2002 (4:00 p.m. PST 22 August, 22). The forecast simulation from this initial state is used to produce the 48-hour power production forecast delivered at 8:00 a.m. PST (1600 UTC) on August 23.

Figures 4–6 and 4-7 respectively show the WST specification from the NCEP OI (110 km) and the MODIS datasets. Figure 4-8 presents the difference between the MODIS and NCEP OI WST values for the coastal area of the Pacific Ocean immediately to the west of the Los Angeles Basin.

A comparison of Figures 4-6 and 4-7 reveals the considerable amount of additional detail that is present in the MODIS WST dataset. This is not surprising since the NCEP OI employs a grid of 110 km while the MODIS dataset has a resolution of 4 km. The added detail most likely provides a much better representation of the surface water temperatures near the coast, which is probably the most significant region for the prediction of winds in the wind power production areas of California. In addition to the added detail for the coastal waters, small inland bodies of water, such as the Salton Sea are represented in the MODIS dataset but not in the NCEP OI dataset. The impact of this on the specification of the temperature of the Salton Sea is very significant. When the NCEP OI is used to specify the WST, the temperature of the Salton Sea is specified to be similar to that of the coastal waters, because the initialization procedure uses the only available information (the ocean water temperatures) to estimate the Salton Sea temperature. However, the MODIS dataset indicates that the actual temperature of the Salton Sea is much warmer than that of the ocean water.

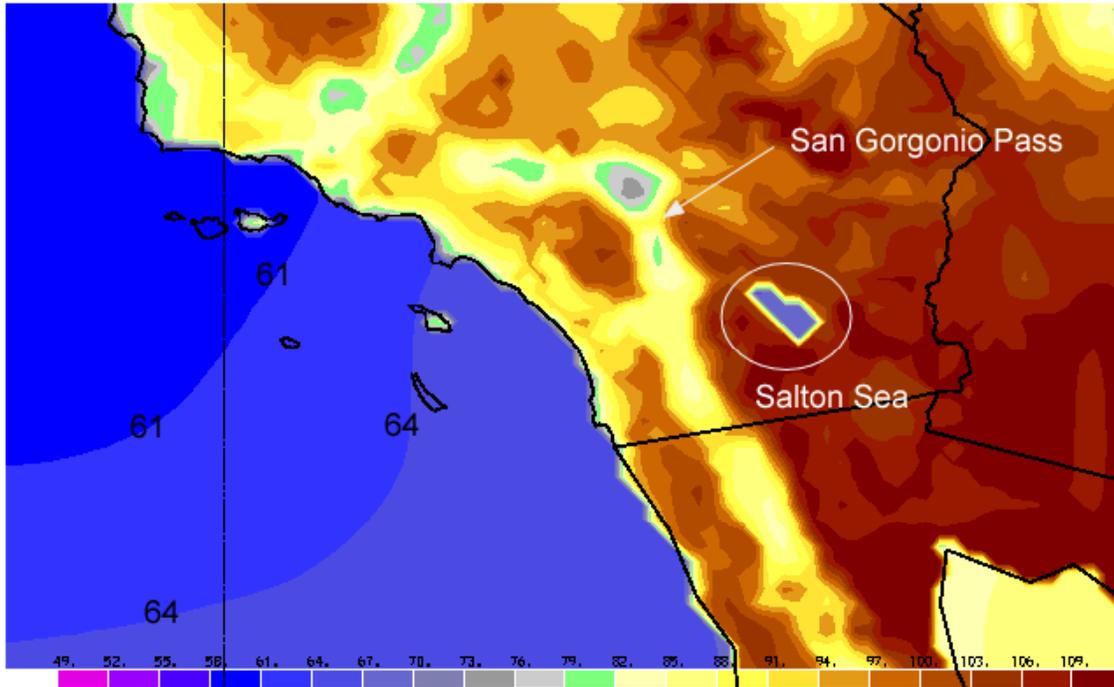


Figure 4-6 Initial skin temperature (degrees F) specification for the August 23, 2002, forecast simulation based on the NCEP OI Water Surface Temperature (WST) analysis.

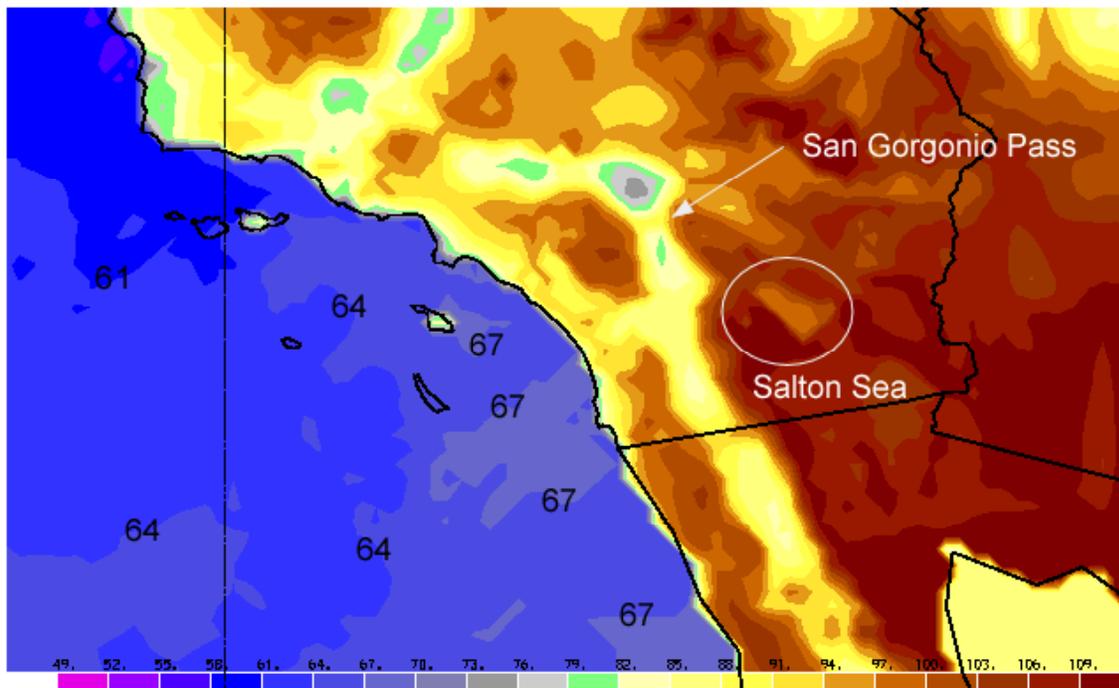


Figure 4-7 Initial skin temperature (degrees F) specification for the August 23, 2002, forecast simulation based on the MODIS Water Surface Temperature (WST) dataset.

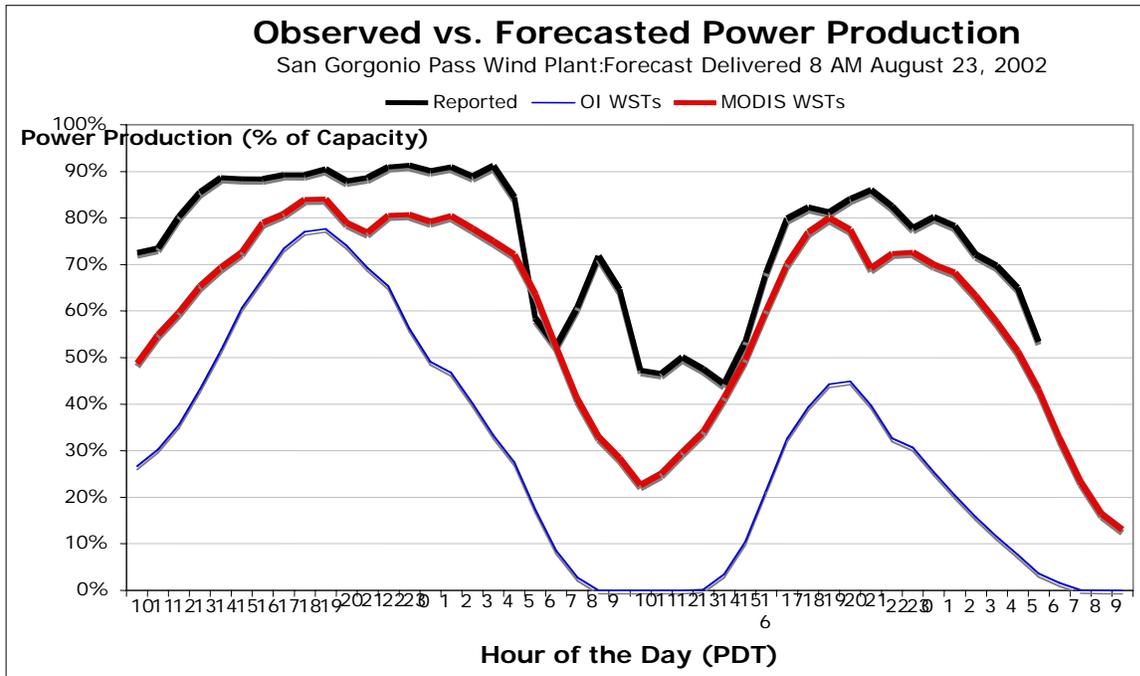


Figure 4-9 Forecasted and observed power production for the Mountain View wind plant and forecast period beginning at 9 a.m. PDT August 23, 2002.

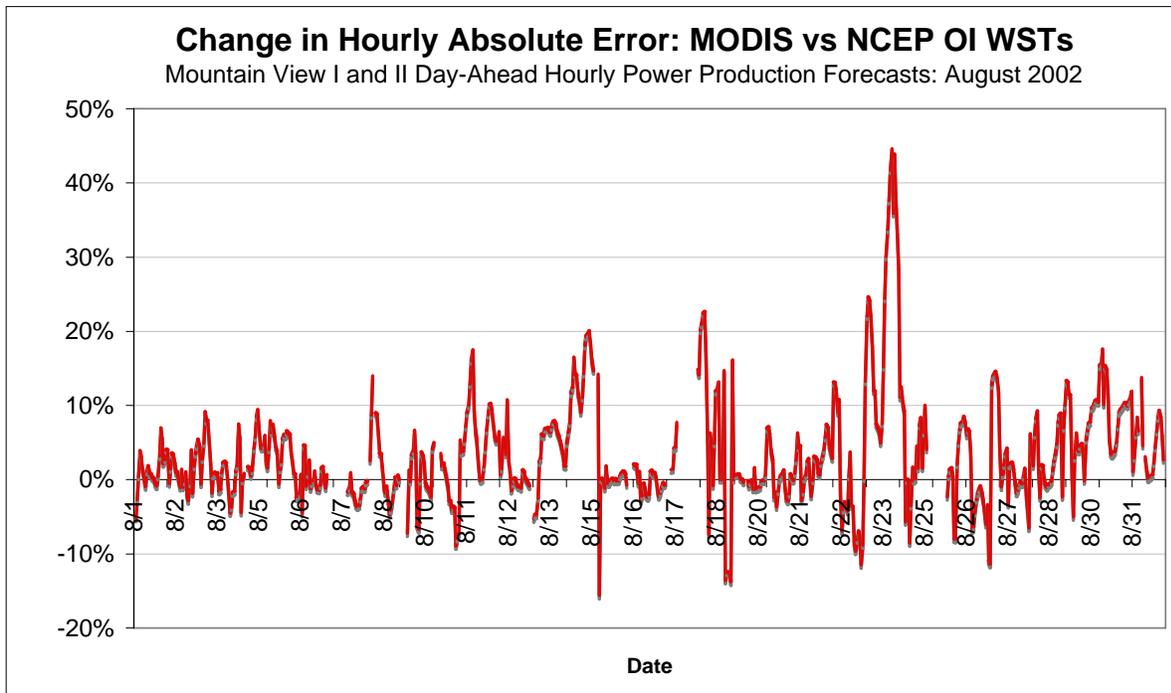


Figure 4-10 Change in hourly absolute forecast error (% of installed capacity) for MODIS vs. OI WSTs for the Mountain View wind plant for August 2002. Positive values indicate forecasts based on MODIS data have a smaller absolute error.

wind energy forecast errors for the forecasts using the NCEP OI data vs. those using the MODIS data for the Mountain View wind plant and August 2002. Positive values indicate the forecasts based on the MODIS WST data have a lower hourly absolute error than those based on the NCEP OI data. The chart indicates that the MODIS-based forecasts are better for a substantial majority of the hours and are rarely worse. The largest improvement occurs on August 23, which is, not by coincidence, the case example shown in Figures 4-6 through 4-8.

The comparison of forecast performance for August 2002 in Figure 4-10 certainly provides a more persuasive argument about the positive impact of more detailed WST data on wind power production forecasts in California than a single case. However, one might still argue it is for only one month at one wind plant. Figure 4-11 presents the change in the mean absolute error (MAE) for all six of the test months, which include three months of forecasting for the PowerWorks wind plant and three months for forecast production for the Mountain View wind plant. As in Figure 4-10, a positive change means that the MODIS-based forecasts are better, and a negative change indicates the OI-based forecasts are better. The black bars depict the change in the MAE of the raw physics-based model forecasts (without the MOS procedure). These forecasts are produced by interpolating the forecasted grid point values of wind speed and direction from the physics-based model grid to the wind plant site. The red bars show the change in the MAE for the forecasts that used an MOS procedure on the physics-based model data. The same MOS procedure was used for both sets of forecasts.

Interestingly, the forecasts derived directly from the raw physics-based model data showed little change in performance. The change for the three months of PowerWorks forecasts exhibited only noise-level differences from zero. There was a slightly more significant positive change for two of the three months for Mountain View with the largest impact being a 3.2% improvement for August 2002. However, the results for the forecasts that also employ the MOS procedure were much more impressive. All six months showed a substantial positive change (MODIS-based forecasts are better) in the MAE, ranging from a minimum improvement of 7.3% for December for PowerWorks to a maximum of 18.5% for Mountain View for the same month. The reduction in MAE is more significant for the Mountain View plant than for the PowerWorks plant. In fact, the reduction in MAE for all three of the Mountain View tests months is greater than the largest monthly reduction in MAE achieved for the PowerWorks wind plant.

The fact that the forecasts based on the raw physics-based model data showed little improvement indicates that the direct simulation of the winds by the physics-based model at the site of the plants was not significantly improved by the use of the more detailed WST data. However, the MOS procedure employs a number of physics-based model parameters (not just the near-surface wind speed and direction) to find signals that predict the behavior of the winds at the wind plant sites. The results indicate that some combination of the physics-based model parameters from the MODIS-based simulations are a better set of predictors for the wind plant winds than any set from the OI-based simulations. These results indicate that the use of more detailed WST data can have a significant positive

impact on wind power production forecasts for plants in the Altamont and San Geronio Passes and likely in other wind resources areas in California.

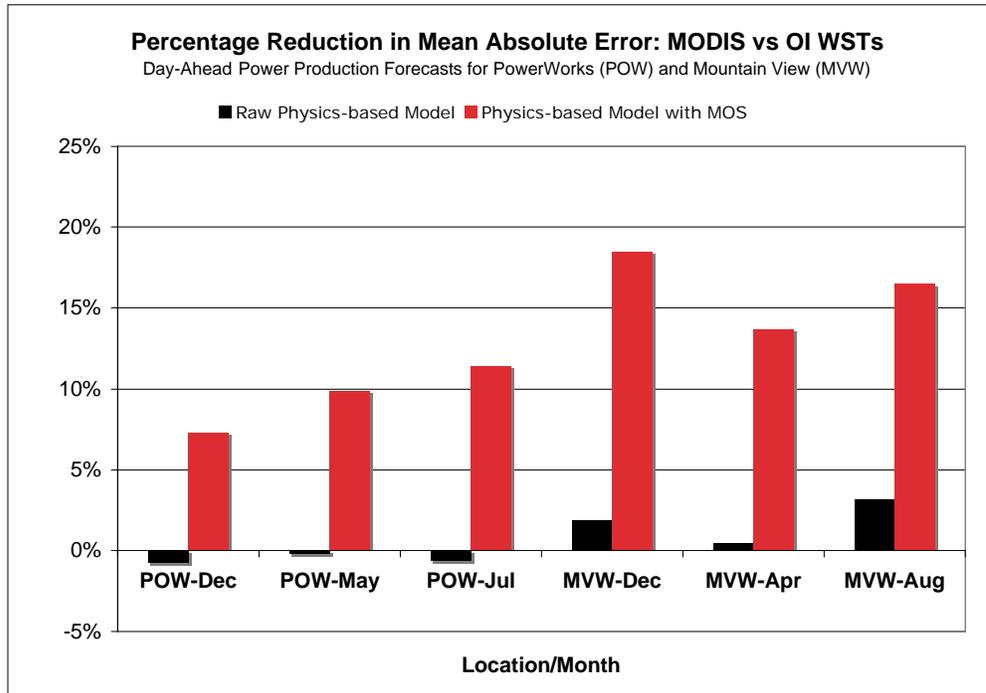


Figure 4-11 Percentage reduction in the mean absolute error of the power production forecasts for the PowerWorks and the Mountain View wind plants. Positive values indicate a reduction in MAE when switching from the use of OI WST data to MODIS data while negative values indicate an increase.

Focus Area 2: Model Resolution Experiments

Most of the wind power production regions in California are located within areas of complex terrain that exhibit large variation in elevation and slope over relatively short distances. In addition, the areas of complex terrain are typically near the coast lines of bays and the ocean that have a complex structure. The numerical simulation of the wind in this type of environment requires a computational grid with small grid cells to resolve the fine scale structure of the terrain and coastlines. Therefore, one might expect that forecasts generated from a physics-based model that employs a grid with smaller grid cells (higher resolution) will achieve a higher level of forecast performance than those generated from a coarse grid.

The impact of physics-based model grid resolution on forecast performance was addressed through a set of numerical experiments. Power production forecasts were produced for the three test months at the PowerWorks and Mountain View wind plants based from simulations produced on four different resolutions of the physics-based model grid. Grids with cell sizes of 40 km, 10 km, 4 km, and 1 km were used. Figure 4-12 shows the single 40-km grid used to produce forecasts for both wind plants. However, separate 10-km, 4-km

and 1-km grids were used to produce the higher resolution forecasts for each plant. Figures 4-13 and 4-14 show the sets of higher resolution grids used to generate forecasts for the PowerWorks plant in Altamont Pass and the Mountain View wind plant in San Geronio Pass.

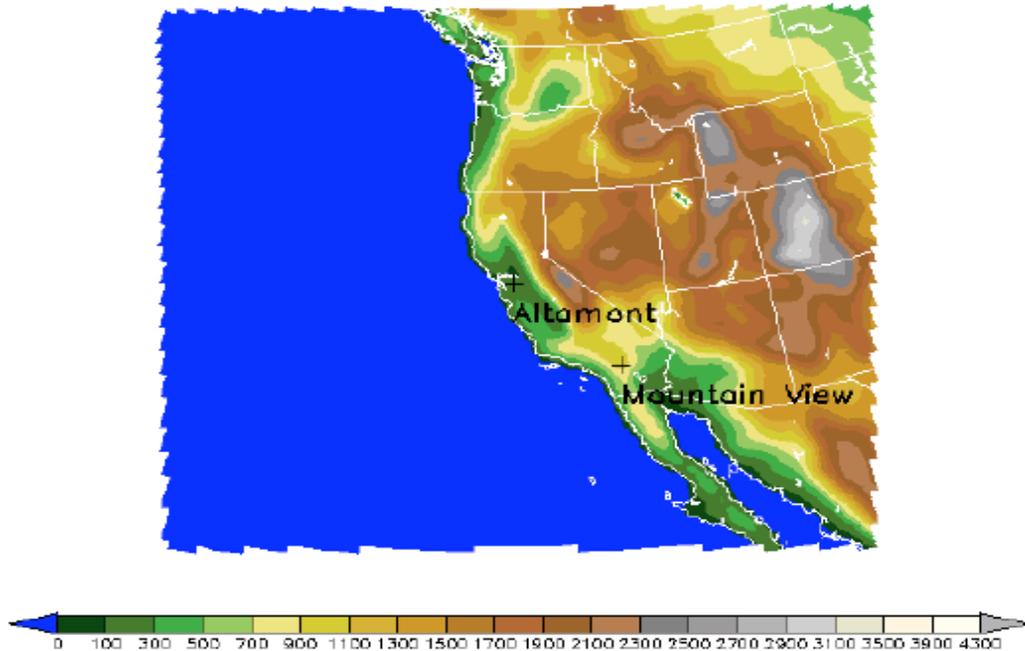


Figure 4-12 A depiction of the geographical area covered by the coarsest resolution grid (40 km grid cell size) used to study the effect of physics-based model grid resolution on the performance of day-ahead power production forecasts for the PowerWorks and Mountain View wind projects.

Two different types of forecasts were generated from each grid. The first type was based directly upon an interpolation of the wind speed and direction from each model grid to the locations of the anemometers for each plant. The power production was then directly estimated by using a plant output model. The second set of forecasts used a MOS procedure to predict the wind speed and direction at the wind plants from the physics-based model data from each grid.

Figures 4-13 and 4-14 present the MAE as a function of the physics-based model grid cell size for the raw and MOS-adjusted power production forecasts at the PowerWorks wind project. A casual glance at Figures 4-13 and 4-14 indicates that there is a substantial variation in the dependence of the MAE on the grid cell size of the physics-based model among the three months. The MAE has very little dependence on the grid cell size for December 2001.

Actually, a close look at the MAE values indicates that there is a tendency for the MAE to decrease as the grid cell size decreases. At PowerWorks, the monthly MAE for December ranges from a high of 6.4% of installed capacity for the 40-km model to a low of 5.7% of

installed capacity for the 1-km model. The MAEs of the 10-km and 4-km models are within this range of 5.9% and 6.0%, respectively.

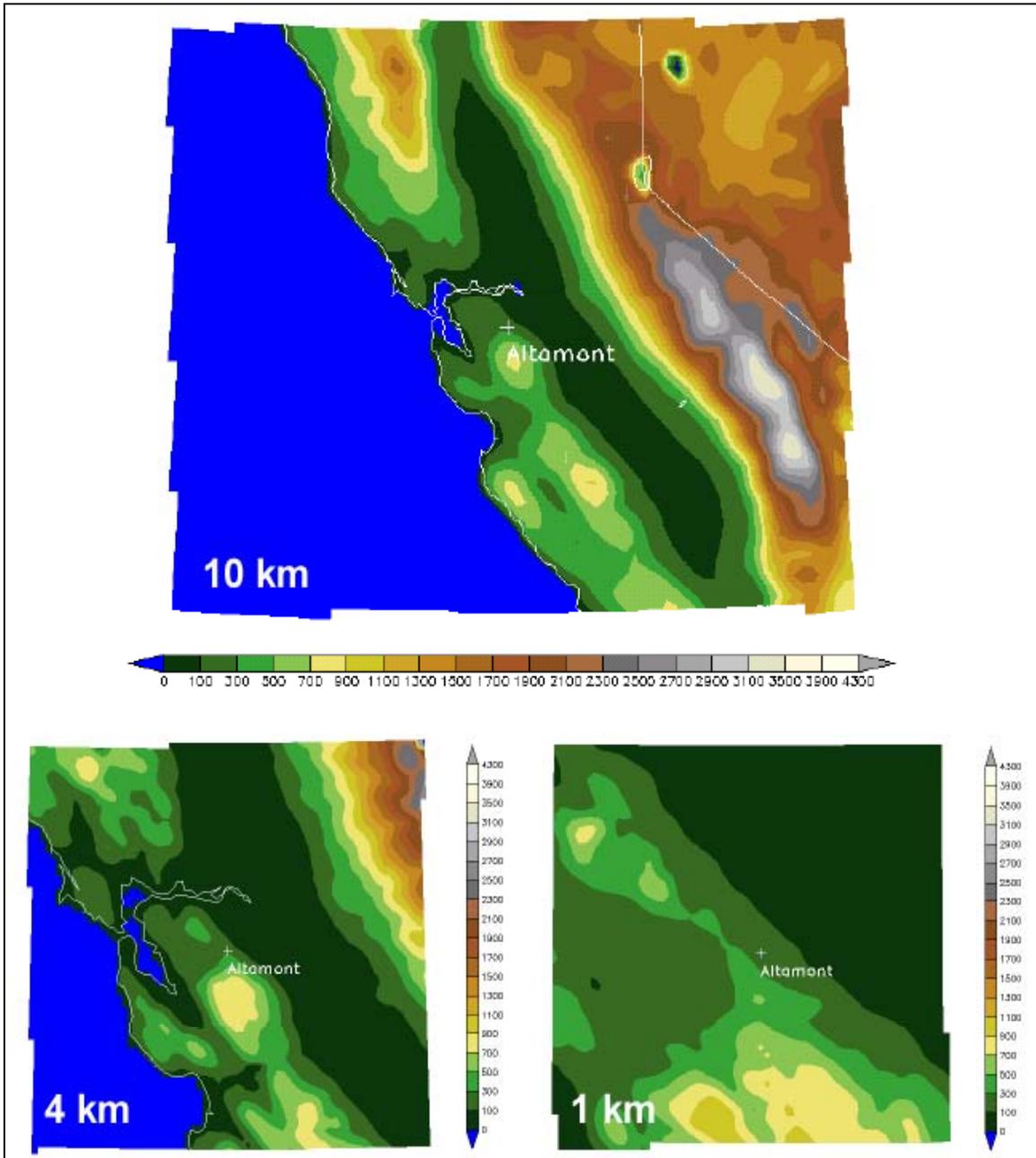


Figure 4-13 A depiction of the geographical areas covered by the three nested grids used to study the impact of physics-based model grid resolution (or the horizontal size of the grid cells) on the performance of day-ahead power production forecasts for the PowerWorks wind plant in Altamont Pass. The upper image depicts the domain of the 10-km grid (Grid B). The lower left image illustrates the area covered by the 4-km grid (Grid C), and the lower right image shows the region encompassed by the 1-km grid (Grid D). A contour plot of the terrain representation on each grid is shown in each image. The terrain elevations are in meters above sea level.

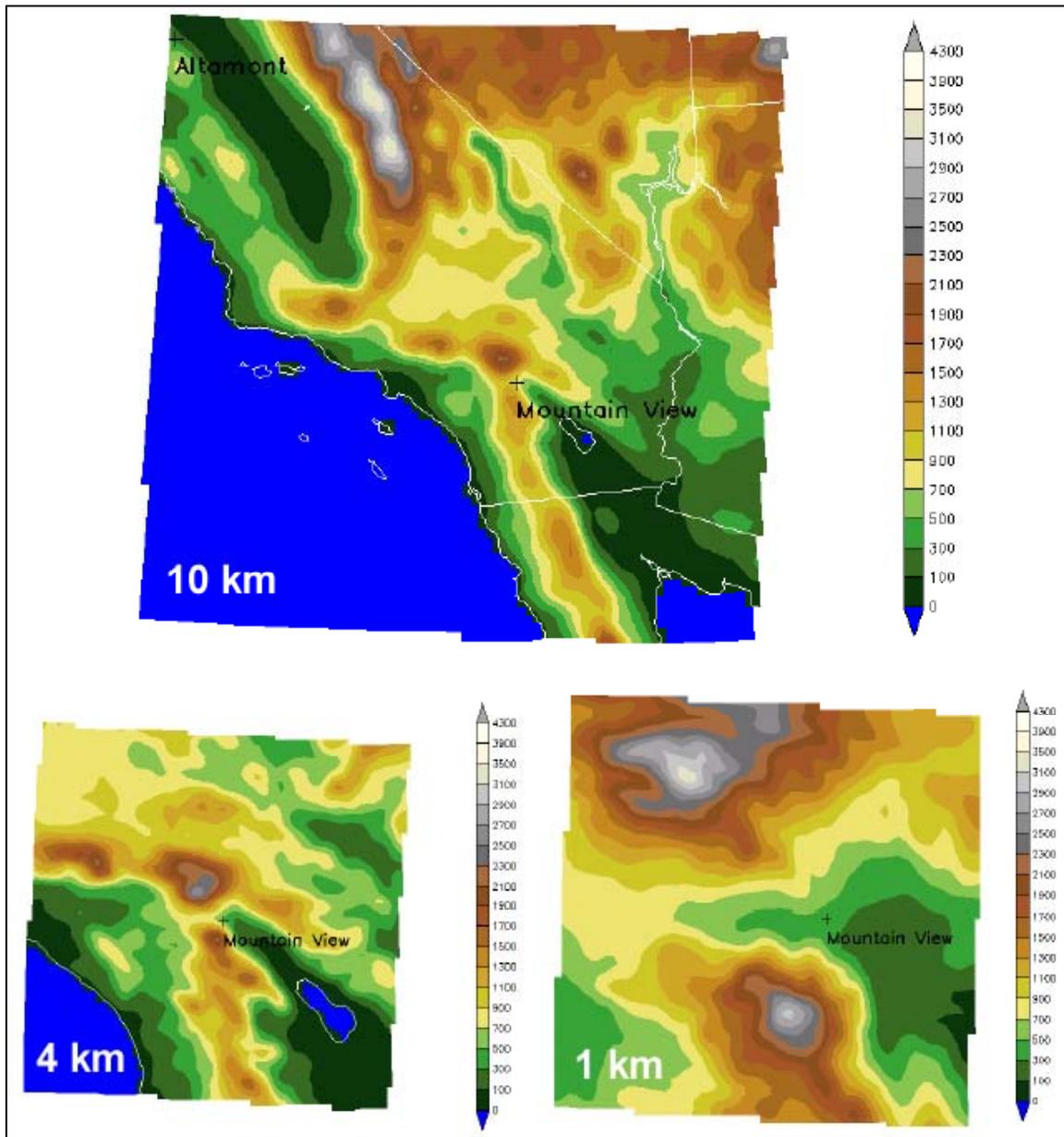


Figure 4-14 A depiction of the geographical areas covered by the three nested grids used to study the impact of physics-based model grid resolution (or the horizontal size of the grid cells) on the performance of day-ahead power production forecasts for the Mountain View wind plant in San Gorgonio Pass. The upper image depicts the domain of the 10-km grid (Grid B). The lower left illustrates the area covered by the 4-km grid (Grid C), and the lower right shows the region encompassed by the 1-km grid (Grid D). A contour plot of the terrain representation on each grid is shown in each image. The terrain elevations are in meters above sea level.

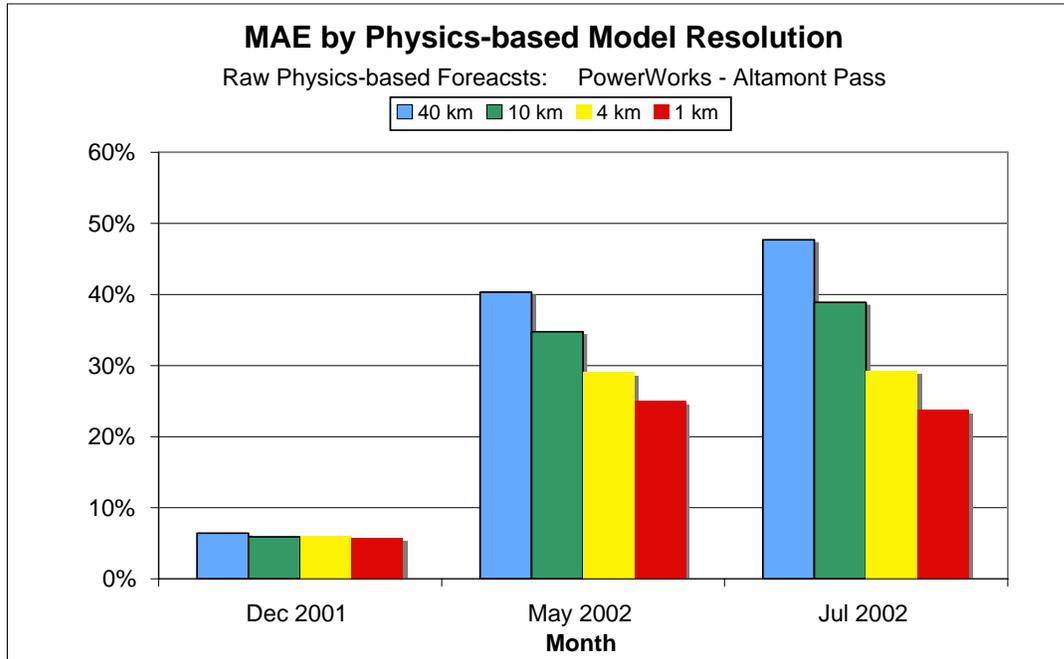


Figure 4-15 Mean absolute error by physics-based model grid cell size for the 1-to 48-hour raw physics-based forecasts of the hourly power production for the PowerWorks wind plant for the three forecast test months

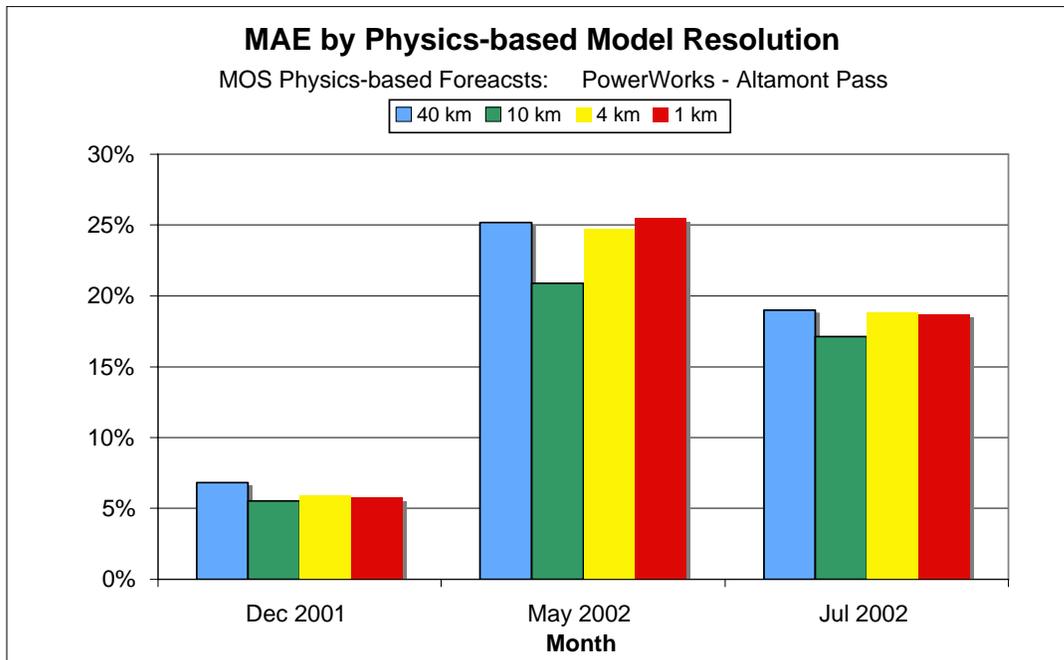


Figure 4-16 Mean absolute error by physics-based model grid cell size for the 1- to 48-hour MOS-adjusted physics-based forecasts of the hourly power production for the PowerWorks wind plant for the three forecast test months

One issue that could be raised when attempting to interpret the power production forecast MAE values is that there are a large number of hours during December when the wind speed falls near or below the turbine start-up speed. It might be argued that the lack of dependence of MAE on grid cell size is due to the fact that if the wind speeds are below the start-up speed, the power production forecast is not sensitive to errors in the wind speed forecast since the forecast will be for zero production regardless of the forecasted wind speed as long as it is forecasted to be below the start-up speed. Hence, it might not matter much if one version of the model has a higher wind speed error than the other.

However, the composite MAE of the wind speed forecasts for the 11 PowerWorks anemometers indicate that the relationship of MAE to grid cell size is similar for the underlying wind speed forecasts. The MAE of the 40-km model wind speed forecasts is 2.18 m/s, while the MAEs for the 10-km, 4-km, and 1-km models are 2.19 m/s, 2.21 m/s, and 2.28 m/s, respectively. Thus, there is actually a slight increase in the wind speed forecast MAE as the grid-cell size is reduced.

In contrast, there is a very strong relationship between the MAE and the grid cell size for of May 2002 and July 2002, with July having the strongest relationship. For July, the MAE of the 40-km model was 47.6% of installed capacity, while the 1-km model produced an MAE of 23.7%. Similarly, for May, the MAE decreased from 40.3% of installed capacity for the 40-km model to 24.9% for the 1-km model. Interestingly, for May and July, there is a roughly linear trend in the reduction in MAE as the grid cell size is decreased. The MAE of the wind speed forecasts for the 11 PowerWorks anemometers exhibit a similar pattern with a fairly linear decrease from near 5 m/s for the 40-km grid to just above 3 m/s for the 1-km grid.

The reason for this dramatic difference in the dependence of the MAE of the raw forecasts on the model grid cell size is most likely tied to the differences in the character of the weather that affects the Altamont Pass area in the cold and warm seasons.

During the winter, most of the variations in wind speed and direction in this area are driven by large-scale weather systems that typically move into the area from the Pacific Ocean. These large-scale systems can be fairly well simulated by a physics-based model with a relatively large grid cell size.

In contrast, during the warm season, the wind variability in the Altamont Pass area is predominantly controlled by the small-scale circulations that are driven by the differences in temperature and pressure between the air over the cold waters of the San Francisco Bay and the adjacent Pacific Ocean and the inland areas. This results in a strong diurnal wind speed cycle as the inland areas heat during the day and cool at night, causing large changes in the temperature differences between the coastal and inland areas. The small-scale circulations that result from these temperature gradients require a physics-based model with very small grid cells because of the complex nature of the terrain and land-water boundaries that determine the details of the temperature and wind patterns and the shallow nature of these near-surface flows.

Figure 4-17 presents the incremental change in MAE resulting from application of the MOS procedure to the output of the physics-based model as a function of grid size. In general, the

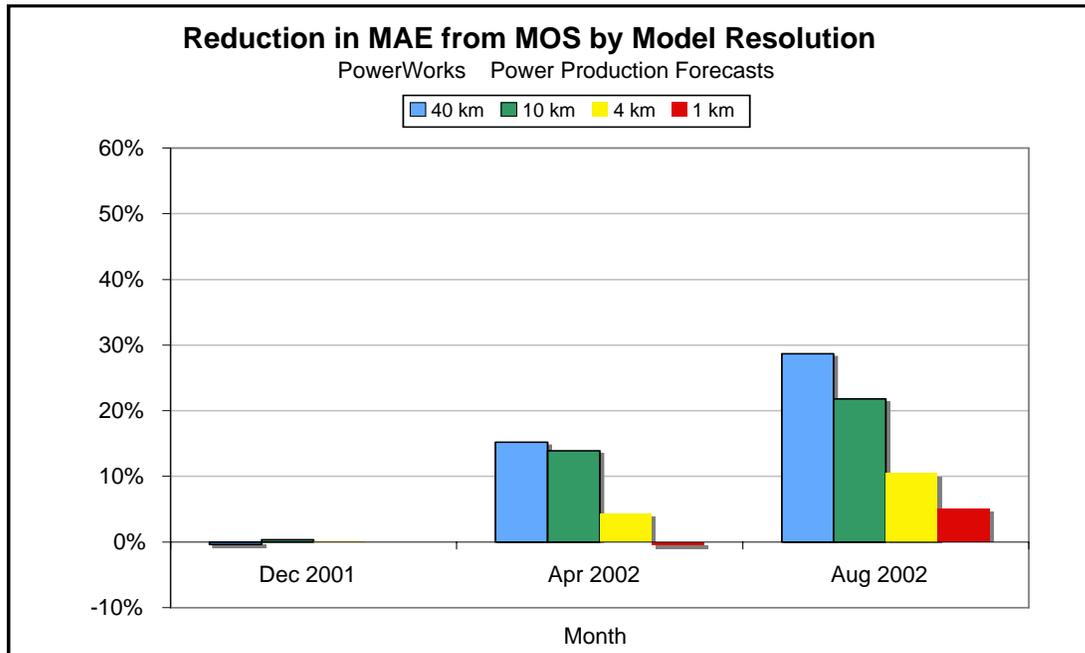


Figure 4-17 The incremental reduction of the mean absolute error (MAE) of the 48-hour wind energy forecasts vs. grid-cell size for the PowerWorks wind plant obtained by applying a MOS procedure to the raw physics-based model data. Positive values represent a decrease, negative values an increase in MAE.

MOS procedure reduces the MAE values. The MOS procedure has very little impact on the MAE for December 2001, regardless of the grid cell size. In contrast, the MOS procedure provides a much more significant improvement, for at least some grid-cell sizes, for May 2002 and July 2002, and the reduction in MAE is generally greater in July than in May. The impact of the MOS procedure is greatest for the 40-km physics-based model,

Figures 4-18 and 4-19 present the MAE vs. physics-based model grid cell size for the raw and MOS-adjusted power production forecasts at Mountain View.

Interestingly, the MAE forecast pattern is somewhat different from that seen in the MAE data for the PowerWorks plant. For the Mountain View plant, the MAE of the power production forecast for December 2001 shows somewhat more dependence on the grid cell size than it did for the PowerWorks plant. The MAE of the forecasts based on the 40-km grid was 23.9% of the installed capacity.

The MAEs produced by the forecasts from 10-km, 4-km, and 1- km grids were all substantially lower than this value. The lowest MAE for this month was produced by 4-km grid, which yielded an MAE of 15.0% of installed capacity. However, the MAEs for the 10-km and 1-km grids were only modestly higher with MAE values of 16.8% and 15.9%, respectively.

The MAE patterns for the two warm season months, April and July, were also somewhat different for the Mountain View wind plant. The pattern for April was the most similar to the pattern seen in the warm seasons for the PowerWorks plant. There is a well-defined monotonic decrease in MAE as the grid-cell size decreases. The PowerWorks plant exhibited the same pattern during both May and July. However, the decrease in MAE was not as linear for the Mountain View plant as it was for the PowerWorks plant. There is a large decrease from the 40-km to the 10-km grid, and then a more linear decrease for the 10-km through 1-km grids. The Mountain View MAE vs. grid-cell size pattern for August is quite different from that seen for the PowerWorks plant.

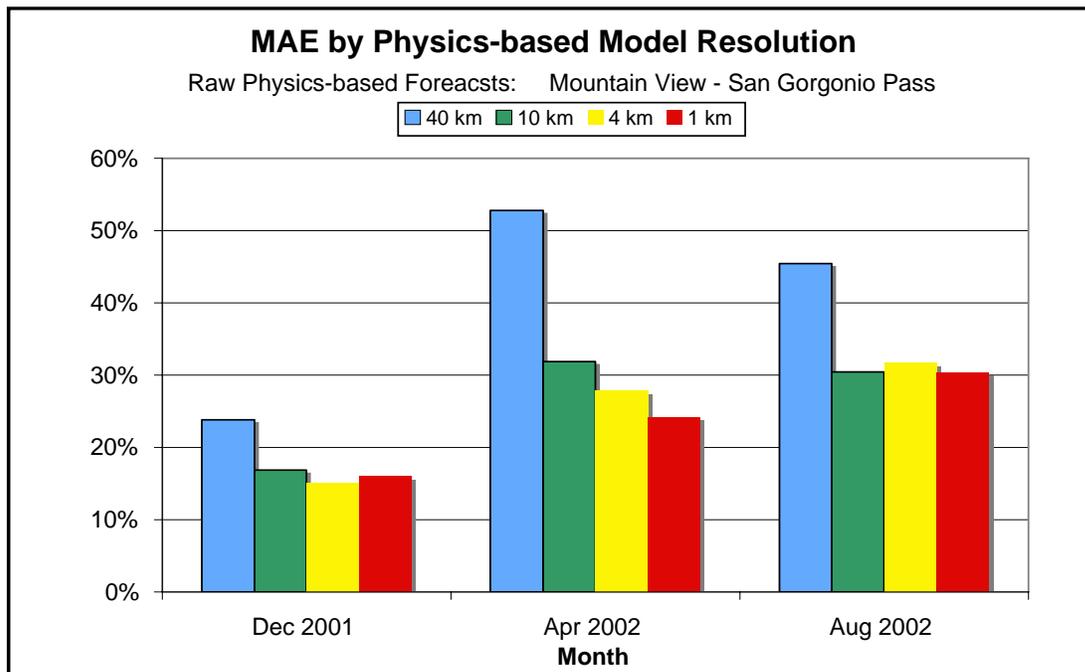


Figure 4-18 Mean absolute error vs. physics-based model grid resolution for 1- to 48-hour raw physics-based forecasts of the hourly power production for the Mountain View wind plant for the three forecast test months.

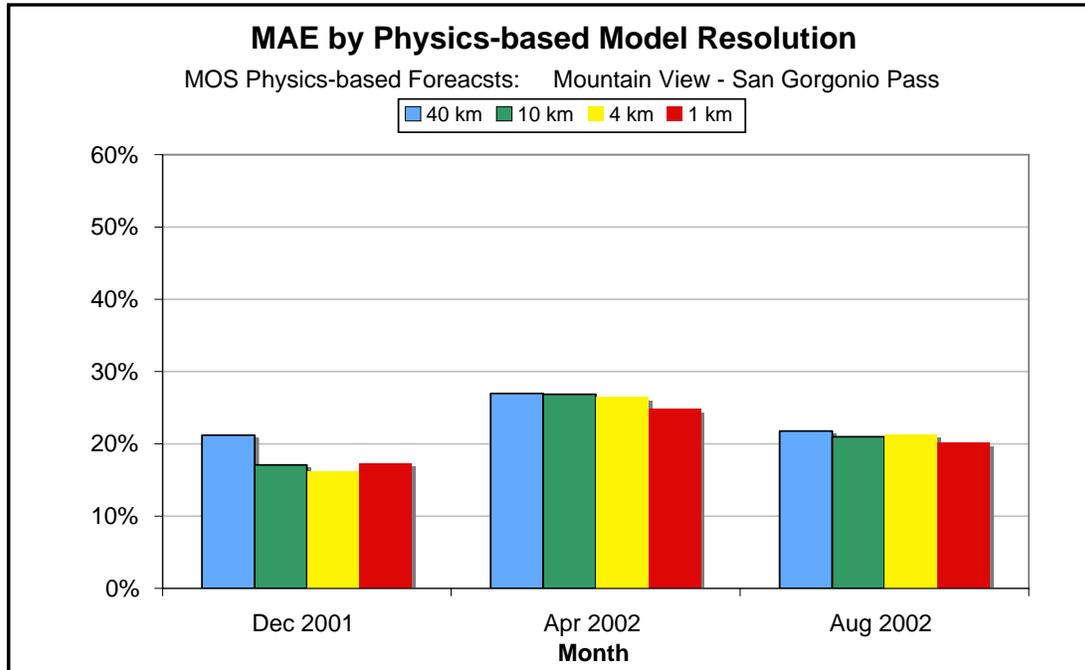


Figure 4-19 Mean absolute error vs. physics-based model grid resolution for 1- to 48-hour Model Output Statistics (MOS)-adjusted, physics-based forecasts of the hourly power production for the Mountain View wind plant for the three forecast test months

Thus, the main characteristic of the relationship of grid-cell size to power production forecast MAE for the Mountain View plant is that there is substantial difference in MAE between the 40-km grid and all of the finer scale grids, but there is less difference when the grid cell size is reduced below 10 km. The reason for this is not clear.

Even a cursory glance at Figure 4-19 indicates that the MOS-adjusted forecasts show relatively little dependence on the grid-cell size of the physics-based model. The largest change in MAE occurred for December, when the MAE for the MOS-adjusted forecasts from the 40 km grid was 21.2%, and those from the other grids were all 4 to 5 percentage points lower. The 4-km grid yielded the lowest MAE of 16.2% of installed capacity.

The dependence of MAE on grid cell size was even less prominent during April and August. During both months, the MAE decreased monotonically with grid-cell size. However, the rate of decrease was fairly small. For April, the MAE decreased from 26.9% for the 40-km grid to 24.7% for the 1-km grid. For August, the MAE declines from 21.7% to 20.1% from the 40-km grid to the 1-km grid.

Figure 4-20 presents the MAE of the Mountain View power production forecast after MOS adjustment vs. month and grid cell size. As in the case of the PowerWorks forecasts, there is little improvement in the MAE after applying the MOS procedure for December 2001. In fact, the MOS procedure actually increased the MAE slightly for all grid-cell sizes below 40 km. The MOS benefit was larger for April 2002, but most of the benefit occurred for the 40-km grid, and there was modest or no improvement using the higher resolution grids. The

forecasts for August received the largest and broadest benefit from application of the MOS procedure. The MAE was reduced by more than 20% of the installed capacity for the 40-km grid, and the three grids with smaller grid cell sizes all yielded MAE reductions of about 10%.

Two significant conclusions can be drawn from the forecast experiments that tested the effect of the physics-based model grid cell size on the performance of the power production and wind speed forecasts. The first conclusion is that the dependence of performance of the raw physics-based model forecasts on the grid-cell size has a strong seasonal dependence. The MAE of the forecasts produced directly from the raw physics-based model data (the raw forecasts) had much less dependence on the model grid-cell size during the cold season than during the warm or transition seasons for both the Mountain View and PowerWorks wind plants.

The second conclusion is that MOS-adjusted forecasts generally show much less dependence on the grid cell size than forecasts made directly from the underlying physics-based model data. In a sense, the MOS procedure serves to compensate for subgrid scale effects that are missing from the physics-based model simulations. Thus, the improvements associated with the MOS procedure are larger for the large grid cell sizes (for example, 40 km), which have a larger range of features that fall into the subgrid category, than for grids with small grid cells. Thus, the MOS procedure acts as a grid resolution equalizer.

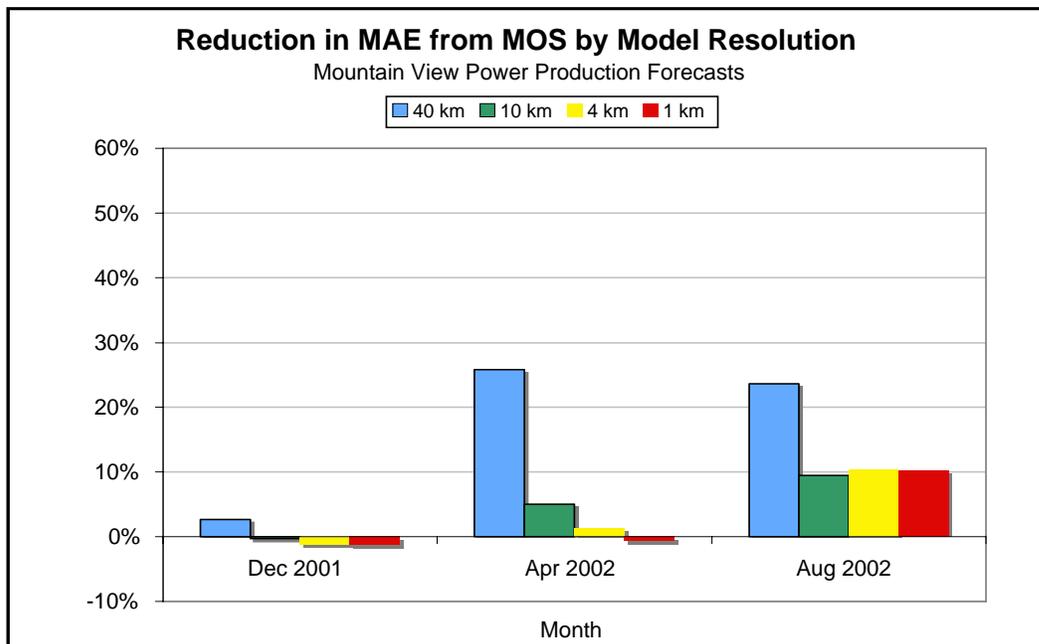


Figure 4-20 The increment of reduction in the mean absolute error of the 1-to-48 hour ahead hourly power production forecasts for the Mountain View wind plant obtained by applying a MOS procedure to the raw physics-based model data. Positive values represent a decrease in MAE from the raw physics-based model forecasts, negatives values denote an increase in MAE.

The overall implication is that there is not a strong incentive to run the physics-based models with a very high resolution grid. While forecasts derived directly from physics-based models that employed a finer grid (smaller grid cells) often yielded substantially lower errors (especially during the warm season) for the two wind plants used in this experiment, the MOS adjustment eliminated a substantial amount of that advantage by reducing the errors more for the coarse grid models. Since there is a much greater computational cost to execute a physics-based model with a fine grid than to run a coarse grid model and apply the MOS adjustment, the coarse grid model with MOS is the more economical choice.

Of course, this conclusion is based on forecast experiments for a total of six months for two wind plants. Therefore, it is difficult to frame this as a general conclusion that can be applied to a broad range of wind plants in different locations. It will certainly be necessary to conduct similar experiments for wind plants in other locations and climate regimes before a more general conclusion can be rigorously justified.

Focus Area 3: Next-Generation Physics-based Models

A key component of most wind power production forecast systems is typically a physics-based model, which is based upon the fundamental principles of physics: the conservation of mass, momentum, and energy and the equation of state for air.

In the previous Energy Commission-EPRI forecast project, the *eWind* forecasts used the Mesoscale Atmospheric Simulation System (MASS) model (Kaplan et al., 1982). The MASS model was developed during the 1980s as part of NASA's research and development of new remote sensing systems. The model has evolved over the ensuing 20 years by incorporating new representations of various physical processes as they became available. However, certain components of the model are difficult to modify because they require a fundamental change in the basic structure of the modeling system and associated software. Several new models have been developed in recent years to take advantage of recent advancements in numerical techniques, software structures, and representations of atmospheric physics. Three prominent "next generation" physics-based atmospheric models are the Coupled Ocean Atmosphere Modeling System (COAMPS) (Hodur, 1997), Weather Research and Forecasting (WRF) model (Michalakes et al., 2004), and the Operational Multiscale Model with Grid Adaptivity (OMEGA) (Bacon et al., 2000).

Coupled Ocean Atmosphere Modeling System (COAMPS)

COAMPS was developed by the Marine Meteorology Division (MMD) of the Naval Research Laboratory (NRL). The atmospheric components of COAMPS are used operationally by the U.S. Navy for short-term numerical weather prediction for various regions around the world. The COAMPS website provides additional information at <http://www.nrlmry.navy.mil/coamps-web/web/home>.

The atmospheric portion of COAMPS is a complete three-dimensional data assimilation and forecast system that includes data quality control, analysis, initialization, and forecast model components. Features include a globally relocatable grid, user-defined grid resolutions and dimensions, nested grids, an option for idealized or real-time simulations, and code that allows for portability between mainframes and workstations.

The nonhydrostatic atmospheric model includes predictive equations that address momentum, non-dimensional pressure perturbation, potential temperature, turbulent kinetic energy, and mixing ratios of water vapor, clouds, rain, ice, graupel, and snow, and contains advanced parameterizations for boundary layer processes, precipitation, and radiation.

The portions of COAMPS that are still under development and testing include the ocean quality control, analysis, initialization, and forecast model and the NRL Atmospheric Variational Data Assimilation System (NAVDAS). The latter is a three-dimensional variational data assimilation system that will replace the current analysis method in COAMPS, which is based on three-dimensional multivariate optimum interpolation method.

The COAMPS model was not available at AWS Truewind for this project. However, the Lawrence Livermore National Laboratory (LLNL) used the COAMPS model to generate simulations for the real-time forecasting phase of the work. Unfortunately, COAMPS simulations were not available for the test months used in the test phase, shown in Table 4-1. Therefore, a comparison of the performance of COAMPS-based forecasts with those based on other physics-based models was deferred until the forecasting evaluation phase of the project began.

Operational Multiscale Model with Grid Adaptivity Model (OMEGA)

OMEGA is a novel multiscale atmospheric simulation system for advanced, high-resolution weather forecasting and the prediction of dosage and hazard levels due to the atmospheric release of aerosols and gases. It has a horizontal grid resolution that ranges from 100 km to 1 km and a vertical resolution that ranges from a few meters to 1 km. OMEGA represents a significant advance in the field of weather prediction. Operational forecast models in current use are scale-specific. Their fixed rectangular grid structure limits the resolution of both the input boundary conditions and the resulting atmospheric simulation. The OMEGA model grid, which is unstructured in the horizontal, adapts to the underlying surface features and can dynamically adapt to atmospheric phenomena as they evolve.

The major advantages of OMEGA over the current state-of-the-art include the ability to resolve the surface terrain down to scales of less than 1 km and the resulting local perturbations of the larger scale wind field. This local wind field perturbation is of extreme importance in wind energy applications as well as in a number of other applications such as the dispersion of atmospheric aerosols. However, to calculate the local perturbation, it is important to include all of the physical parameters and processes, that affect the local flow. These include, not only topography, but also land use, land/water composition, vegetation, soil moisture, snow cover (if appropriate), and surface moisture and energy budgets. The inclusion of this additional physics, some of which is only appropriate because of the increased spatial resolution, represents an additional advance in the state-of-the-art. The OMEGA model website provides additional information at <http://vortex.atgteam.com/>.

The original plan was to generate forecast simulations using the OMEGA model for the six test months. However, the time required to setup the OMEGA software and the computational requirements of OMEGA simulations made it impossible to execute OMEGA simulations within the available resources. Therefore, no forecast experiments were conducted using the OMEGA model.

Weather Research and Forecasting Model (WRF)

The Weather Research and Forecasting Model (WRF) is a next-generation meso-scale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. It features multiple dynamical cores, a three-dimensional variational (3DVAR) data assimilation system, and a software architecture that allows for computational parallelism and system extensibility. WRF is suitable for a broad spectrum of

applications across scales ranging from meters to thousands of kilometers. More information is available on the WRF website, which can be found at <http://www.wrf-model.org/>. This site provides information on the WRF effort and its organization, references to projects and forecasting involving WRF, and links to the WRF users' page, real-time applications, and WRF-related events.

The WRF development effort has been a collaborative partnership, principally among the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (the National Centers for Environmental Prediction (NCEP) and the Forecast Systems Laboratory (FSL), the Air Force Weather Agency (AFWA), the Naval Research Laboratory, Oklahoma University, and the Federal Aviation Administration (FAA). WRF allows researchers the ability to conduct simulations reflecting either real data or idealized configurations. WRF provides operational forecasting a model that is flexible and efficient computationally, while offering the advances in physics, numerics, and data assimilation contributed by the research community. The WRF has a rapidly growing community of users, and workshops and tutorials are held each year at NCAR. The latest version of the model is Version 2.0, and the most recent release is WRF V2.0.3.1 (December 2004).

MASS and WRF Model Experiments

As noted previously, it was not possible to execute experiments with COAMPS and OMEGA models in the screening phase of this task. Thus, the only “next generation” model for which forecast simulations could be produced was the WRF model.

The WRF software is still in the later stages of its development, and all of the planned capabilities were not available at the time the experimental simulations were generated. One of the capabilities that was not available for this project was the ability to execute a one-way nested simulation. This capability is necessary to permit the use of embedded higher resolution grids within a large scale coarse grid. This was the capability used to execute the multiple resolution MASS simulations under Focus Area 2. Without the availability of this functionality, it was only possible to execute a WRF simulation on a single grid.

A 40-km grid was used to make the results comparable with the simulations from the MASS 40-km grid. Unfortunately, at this scale, the physics-based models do not capture most of the meso-scale and local-scale features within the vicinity of the wind plants.

Another limitation of the WRF software at the time of the screening phase work was that only a limited number of grid point datasets could be used for the specification of the initial conditions and lateral boundary conditions for the simulation. Unfortunately, the NCEP Global Forecast System (GFS) dataset used for the MASS simulations in this project could not be used by the WRF software. Hence, a different dataset had to be used for the initialization and the specification of the lateral boundary conditions for the WRF simulations. The dataset used for this purpose was based on the output from the NCEP Eta model on a 40-km grid.

Thus, the comparison between the MASS-based forecasts and WRF-based forecasts is clouded by the fact that the two models used different initialization and lateral boundary condition datasets.

A third issue was the fact that the Eta data required for the WRF initialization could not be obtained in a timely manner for all of the test months. Thus, the test months for this experiment had to be modified. The months of December 2001, April 2002 and June 2002 were used for the Mountain View wind plant and the months of December 2001, May 2002 and June 2002 were used for the PowerWorks plant.

Figures 4-21 and 4-22 compare the monthly MAEs of the wind energy forecasts based on the MASS and WRF 40-km simulations for the six evaluation months. No comparison could be made for June 2002 for the Mountain View wind plant because no corresponding 40-km MASS simulation data were available for the month.

The wind energy forecasts from the raw (for example, interpolated directly from the model grid) WRF wind speed data yielded a lower monthly mean absolute error than those based on the raw MASS wind speed data in four of the five test months (Figure 4-21). The MASS-based forecasts yielded a lower MAE only for the PowerWorks wind plant for December 2001. The WRF also yielded a lower MAE for four of the five comparison months for the MOS-based forecasts. Curiously, they were not the same months as for the raw forecasts.

Interestingly, the relative wind speed errors for the forecasts from the two models were somewhat different (Figures 4-23 and 4-24). The MAEs are lower for the MASS-based forecasts for all three of the PowerWorks evaluation months and for both the raw and MOS-adjusted forecasts. The WRF-based wind speed forecasts yielded a lower MAE for the two months for which a comparison was possible for the Mountain View wind plant.

Analysis of the MAE variation with forecast hour provides further insight into the nature of the differences in the forecast errors between the two models.

Figures 4-25 and 4-26 present the MAEs vs. forecast hour for the raw and the MOS-adjusted WRF and MASS forecasts at the PowerWorks wind plant during May 2002. The MAEs of the raw WRF and MASS forecasts are virtually identical throughout the entire 48-hour period. The only exception is that the raw WRF forecasts exhibit a slightly lower MAE during the 3 to 15 hour look-ahead period. This period accounts for virtually all of the difference in the monthly MAE for the PowerWorks plant in May shown in Figure 4-21. It is also interesting to note that there is a strong diurnal oscillation in the MAE patterns for both models. There are MAE minimums at 3 and 27 hours, which correspond to 11:00 a.m. PDT, and peaks at 15 and 39 hours, which correspond to 11:00 p.m. PDT.

The most obvious difference between the raw and MOS-adjusted forecasts at Mountain View in Figures 4-25 and 4-26 is that the diurnal MAE cycle is almost completely absent in the MOS-adjusted forecasts. The MOS-adjusted WRF and MASS forecasts are less similar than the corresponding raw forecasts. Furthermore, over the entire 48-hour period the

MOS-adjusted MASS forecasts have a lower MAE than the MOS-adjusted WRF forecasts even though the relationship for the raw forecasts was the opposite. The majority of the MAE advantage of MASS over the WRF in the MOS-adjusted forecasts occurs during the 18- to 33-hour look-ahead period.

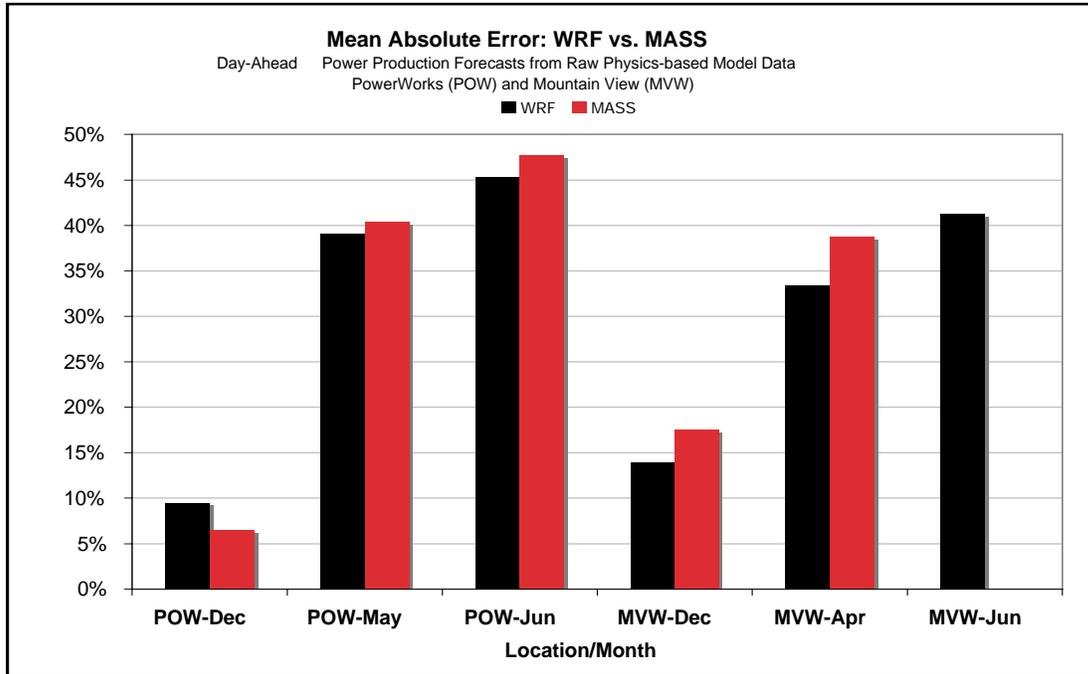


Figure 4-21 Comparison of the mean absolute error of 1- to 48-hour ahead forecasts of the hourly power production for the three screening phase test months for the PowerWorks and Mountain View wind plants from the raw (no MOS) WRF (black bars) and MASS (red bars) physics-based model simulations.

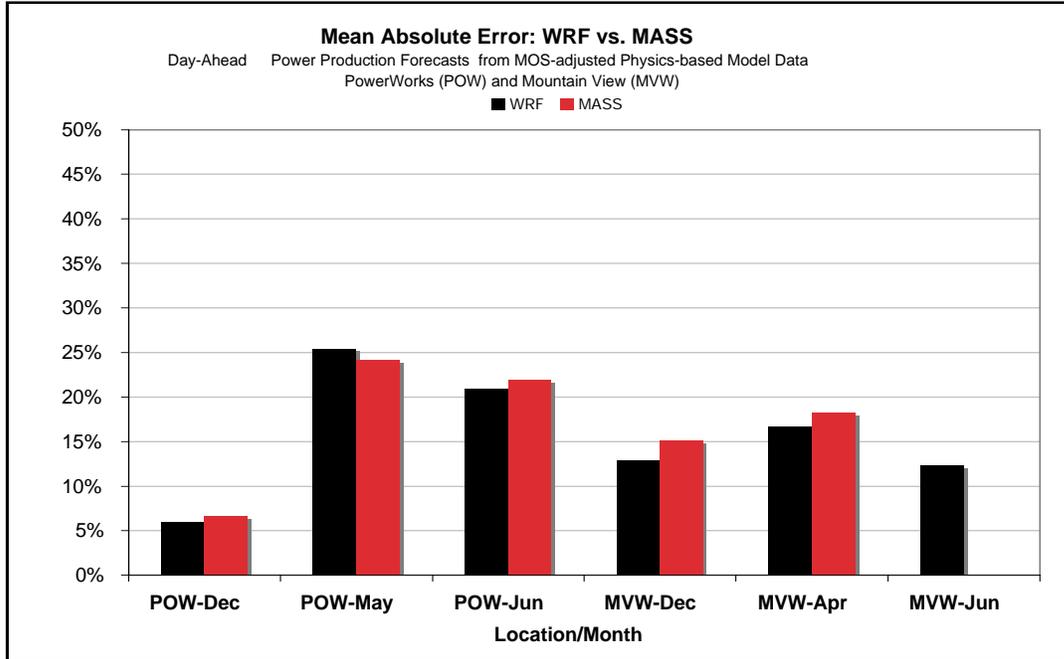


Figure 4-22 Comparison of the mean absolute error of 1- to 48-hour forecasts of the hourly power production for the three screening phase test months for the PowerWorks and Mountain View wind plants from the MOS-adjusted WRF (black bars) and MASS (red bars) physics-based model simulations.

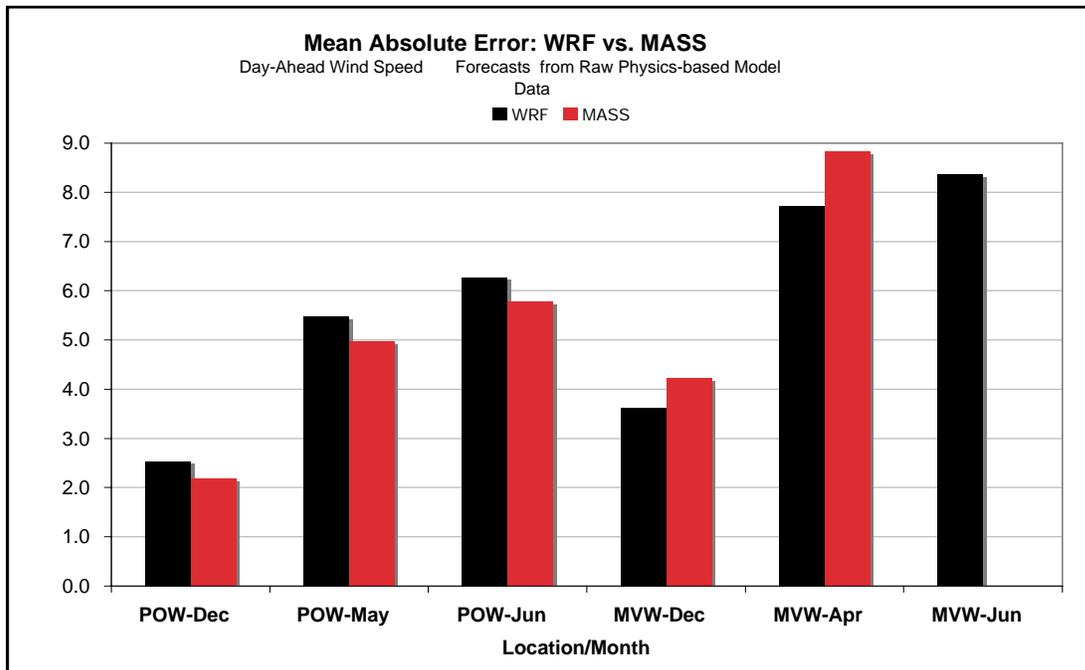


Figure 4-23 Comparison of the mean absolute error of 1- to 48-hour forecasts of the meteorological tower wind speed for the three screening phase test months for the PowerWorks and Mountain View wind plants from the raw (no MOS) WRF (black bars) and MASS (red bars) physics-based model simulation data.

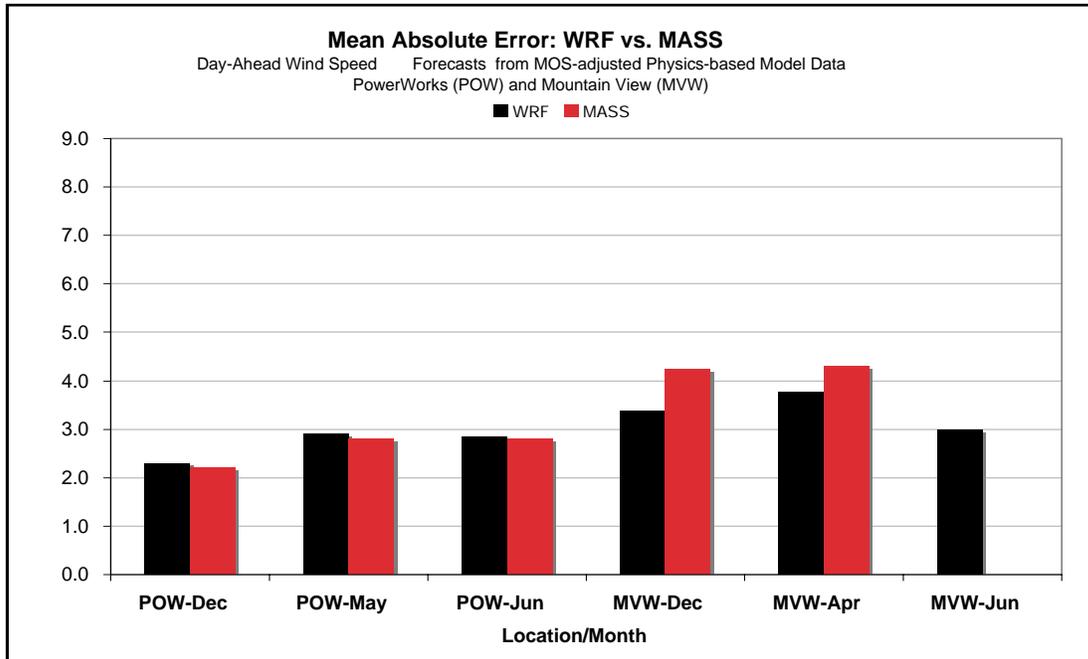


Figure 4-24 Comparison of the mean absolute error of the 1- to 48-hour forecasts of the meteorological tower wind speed for the three test months for the PowerWorks and Mountain View wind plants from the MOS adjusted WRF (black bars) and MASS (red physics-based model simulation data bars).

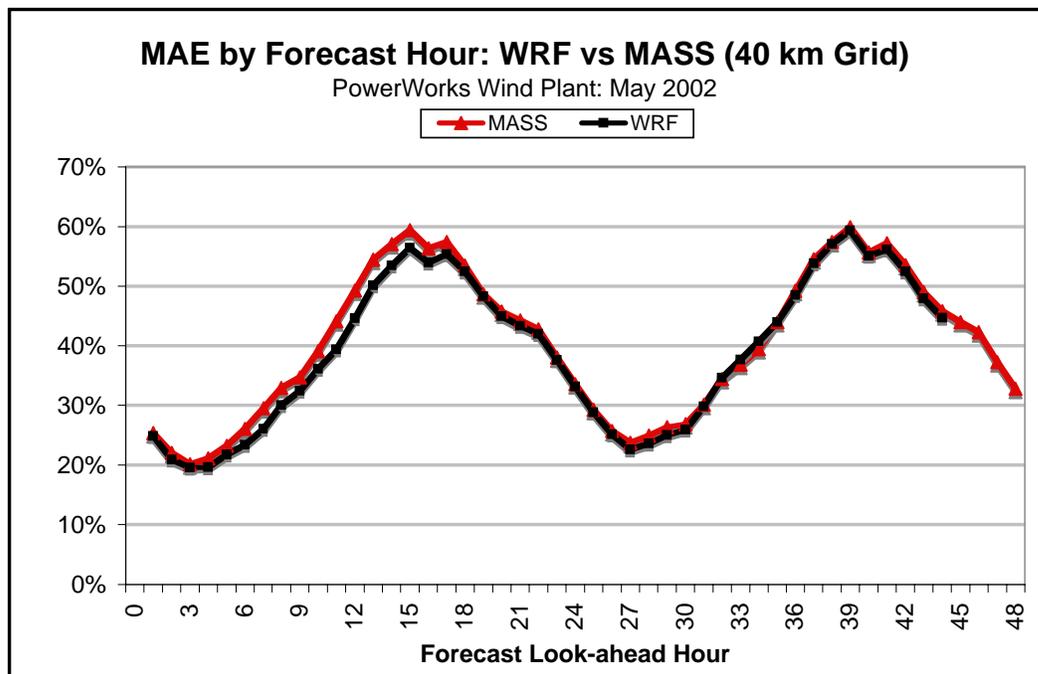


Figure 4-25 MAE of the power production forecasts by forecast look-ahead hour from the raw WRF (black line with rectangle markers) and MASS (red line with diamond markers) model simulation data for the PowerWorks wind plant for May 2002.

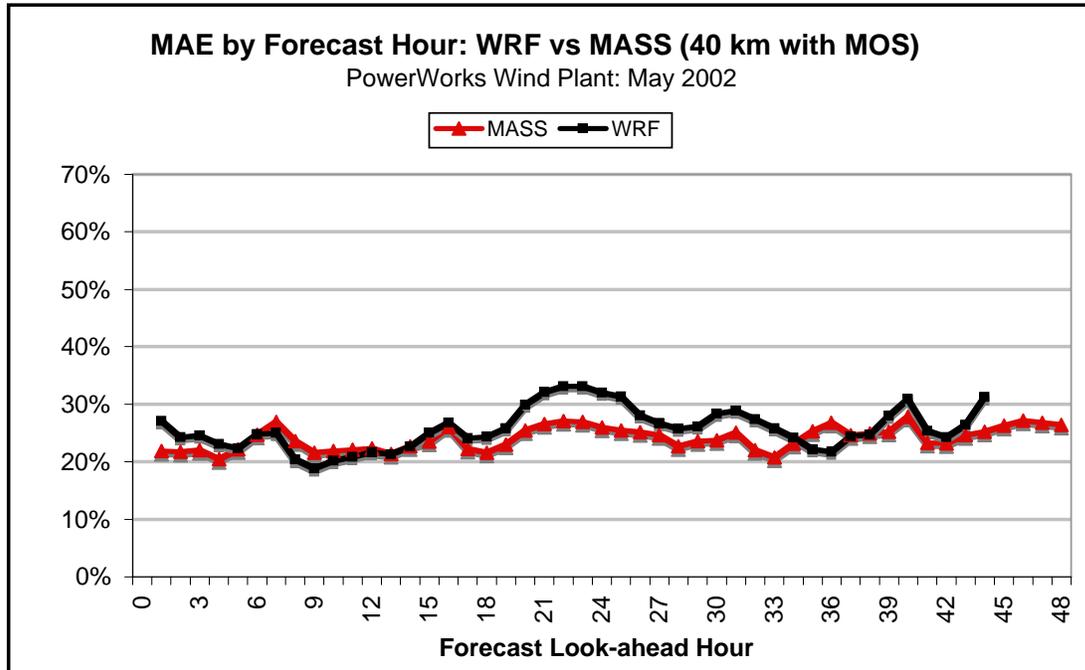


Figure 4-26 MAE of the power production forecasts by forecast look-ahead hour from the MOS-adjusted WRF (black line with rectangle markers) and MASS (red line with diamond marker) model simulation data for the PowerWorks wind plant for May 2002.

Figures 4-27 and 4-28 show another snapshot of the differences between the MASS and WRF forecasts via the MAE vs. forecast hour depictions for the Mountain View wind plant for April 2002. The MAEs of the raw WRF forecasts are significantly lower MAE for a majority of the forecast hours (Figure 4-27), although there appears to be a diurnal cycle to the MAE differences. The MAE differences between the two models are less prominent in the MOS-adjusted forecasts (Figure 4-28). The MOS-adjusted WRF forecasts exhibit a substantially lower MAE in the 24-to-33 hour look-ahead period, but the WRF MAEs are almost the same as the MOS-adjusted MASS MAEs for almost all of the other hours in the forecast period. The 24- to 33-hour period accounts for almost all of the monthly MAE advantage of the MOS-adjusted WRF for April 2002 Mountain View forecasts shown in Figure 4-22.

In addition to the differences in the overall performance between two different forecast methods, it is important to evaluate the performance variations during a period of time. This makes it possible to determine if (1) there are well-defined situations in which one method performs better than the other; or (2) there is substantial temporal continuity in performance patterns so that once it is determined that one method is performing better, that information can be used to select the best performing method for some future time period.

Figures 4-29 and 4-30 show the differences in the hourly absolute error of the MOS-adjusted WRF and MASS forecasts for one of the evaluation months for the PowerWorks and the Mountain View wind plants. Positive differences in both charts indicate that the WRF-based forecasts have a lower absolute error. Addition of the 24-hour moving average

(black line) of the hourly values is also increases the visibility of the patterns on times scales of a day or more.

The character of the performance differences for the May 2002 forecasts for the PowerWorks wind plant (Figure 4-29) appears to change in the middle of the month. For the first 18 days of the month, the WRF-based forecasts have a slight tendency to outperform the MASS-based forecasts but, except for a big advantage to MASS for a few hours on May 10 and a similar one to the WRF on May 15, the differences were consistently small. The amplitude of the differences was much greater during the second half of the month with alternating periods of one to three days in which each model significantly outperformed the other. The reason for this change in the character of the differences is not clear. It may be related to a change in the weather regime or other factors.

The temporal pattern of the error differences is somewhat different in the other illustrated month: April 2002 for the Mountain View plant (Figure 4-30). This month is characterized by short periods (one to two days) in which the WRF-based forecasts significantly outperform the MASS-based forecasts, interspersed with longer periods in which the differences have low amplitude oscillations around zero difference. Most of the monthly MAE reduction produced by the WRF-based forecasts is accumulated in seven or eight days of the month when the WRF-based forecasts had a substantial performance advantage. The performance difference is generally insignificant on the other days.

Overall, the performance of the two models was fairly similar, especially when the MOS procedure was applied. However, the limited test sample suggests that the WRF may have a slight performance advantage over MASS. This advantage was more evident for the Mountain View wind plant than for the PowerWorks plant. The results of the screening phase suggest that it will be worthwhile to examine the performance of the WRF model over a larger sample and to explore the benefits of using both models in an ensemble forecast system. This was done in the second phase of this task.

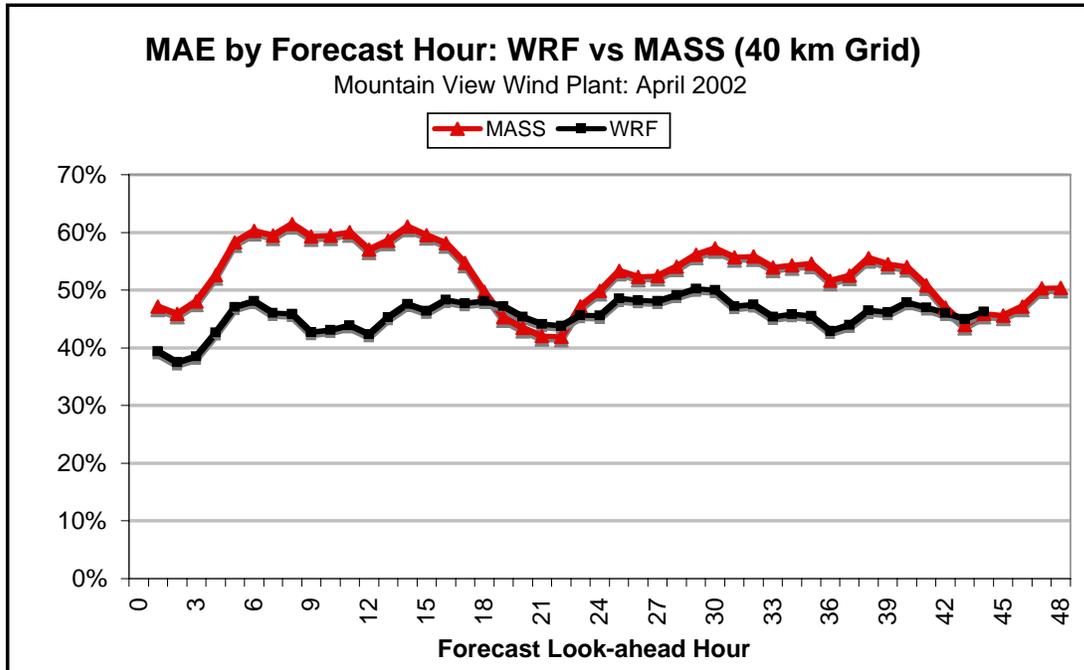


Figure 4-27 MAE of the power production forecasts by forecast look-ahead hour from the raw WRF (black line with rectangle markers) and MASS (red line with diamond marker) model simulation data for the Mountain View wind plant for April 2002.

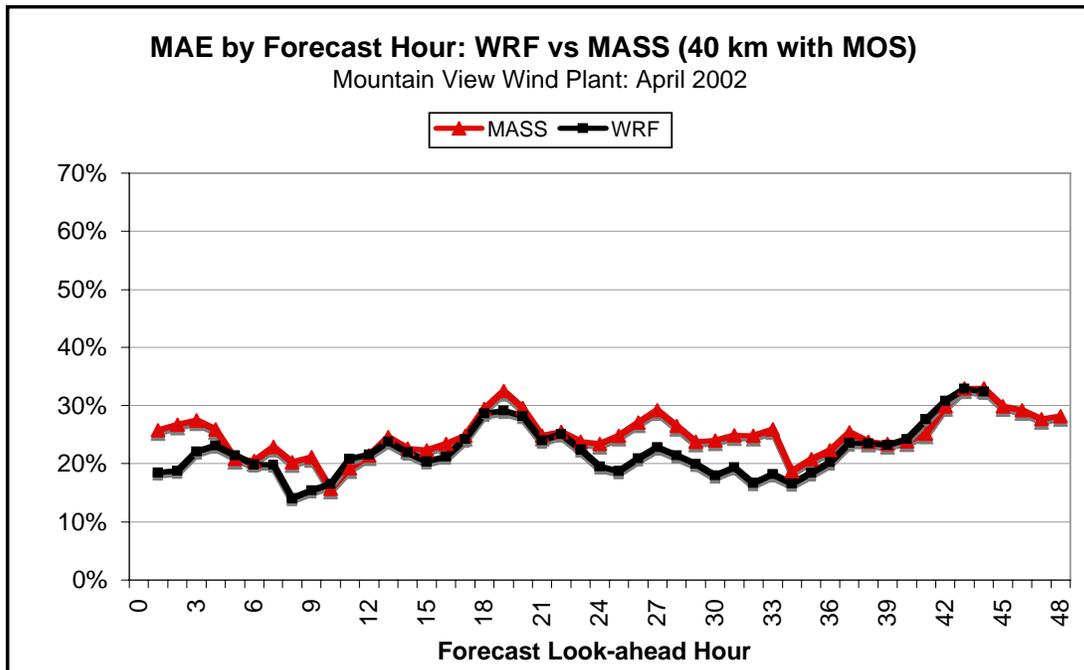


Figure 4-28 MAE of the power production forecasts by forecast look-ahead hour from the MOS-adjusted WRF (black line with rectangle markers) and MASS (red line with diamond markers) model output data for the Mountain View wind plant for April 2002

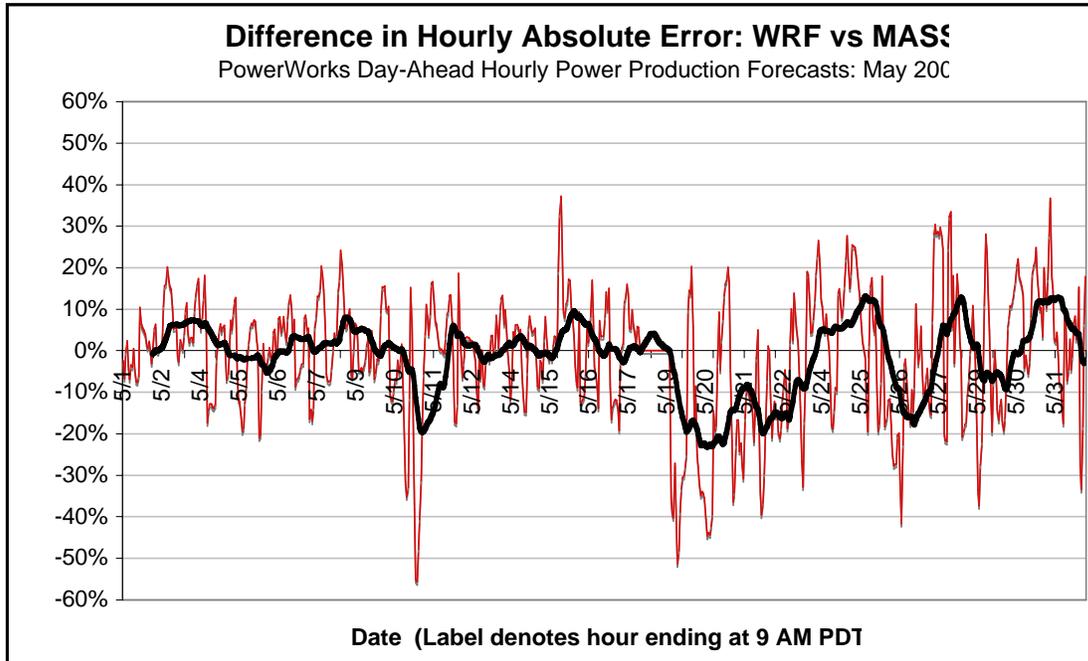


Figure 4-29 Difference between the hourly absolute errors of the WRF and MASS MOS-adjusted power production forecasts (red line for the PowerWorks wind plant and May 2002). Positive differences indicate that the WRF-based forecasts have a lower absolute error. The black line is a lagged 24-hour moving average.

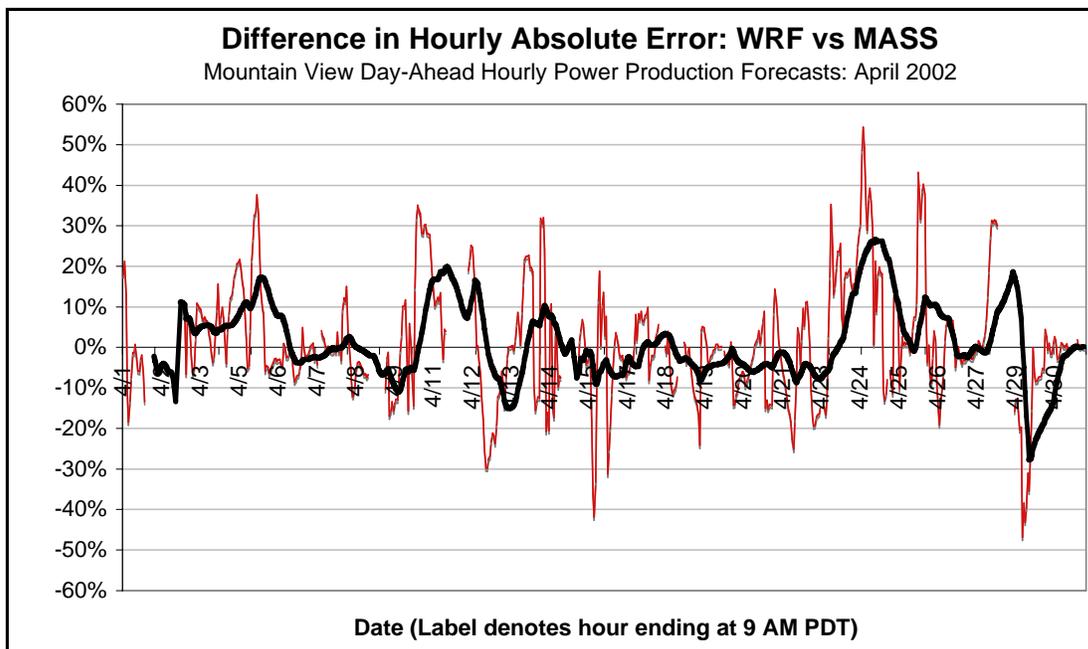


Figure 4-30 Difference between the hourly absolute errors of the WRF and MASS MOS-adjusted 1- to 24-hour power production forecasts (red line) for the Mountain View wind plant and April 2002. Positive differences indicate that the WRF-based forecasts have a lower absolute error. The black line is a lagged 24-hour moving average.

Focus Area 4: Model Output Statistics Formulation

The Model Output Statistics (MOS) component of a day-ahead wind power production forecast system converts the grid-based predictions of a physics-based atmospheric model to point-specific (for example, at a wind plant) forecasts of the variables of interest for a particular application. A well-structured MOS procedure will correct systematic physics-based model errors that result from the omission of atmospheric features that are too small to be resolved by the physics-based model grid, as well as those from deficiencies in the representation of physical processes within the physics-based model.

Screening Multiple Linear Regression (SMLR)

In the previous Energy Commission-EPRI project, AWST used a Screening Multiple Linear Regression (SMLR) procedure for the MOS component of the forecast system. In this approach a large set of physics-based model variables is selected as a pool of candidate predictors. The SMLR algorithm then uses a training sample to select the candidate predictor with the most predictive power. This is defined as the candidate predictor with the highest R^2 value when evaluated over the entire training sample. The candidate predictor that is selected through this process becomes the first predictor of the ultimate statistical prediction equation. The selected predictor is then removed from the pool of candidate predictors and the process is repeated. A second predictor is then selected based on the criteria of the maximum increase in R^2 . The process is repeated until an iteration yields an increase in R^2 that is below a specified threshold. At this point, it is assumed that no additional significant predictive information is available in the training sample and the process is terminated. The selected predictors and associated coefficients form the MOS prediction equation.

There are, of course, many more sophisticated and elegant procedures that could be used in place of this basic SMLR approach. Table 4-5 lists several more sophisticated approaches tested under Focus Area 4. These are listed in Table 4-5. The first method is the SMLR technique described above, which was also the MOS method used in the previous project. This method became the baseline method to evaluate the performance of the other methods in the current project.

Two-Stage SMLR (SMLR2)

The first alternate method tested was a two-stage SMLR, which is assigned the label SMLR2 in Table 4-5. In the first stage, only physics-based model variables are used to create one prediction equation for all forecast hours (hours 1 to 48 in this case). The first stage results in a wind speed forecast for all forecast hours. The second stage uses the wind speed prediction from stage 1 and additional predictors, such as the recent forecast errors, to generate a separate forecast equation for each forecast look-ahead hour. Ultimately, the wind speed forecast from stage 2 is converted to a power production forecast using a plant output model.

Table 4-5 MOS Techniques used in the Focus Area No. 4 Experiments

Technique Name	Identifier	Salient Characteristics
Screening Multiple Linear Regression	SMLR	Predictors selected sequentially from a pool of 56 physics-based model variables
Two-stage SMLR	SMLR2	Two stages of SMLR; recent error parameters included in second stage
SMLR2 with Direct Power Prediction	SMLR2-DP	Direct prediction of power production from physics-based model variables (i.e. no plant output model)
Stratified Two-Stage SMLR	SMLR2-ST	Training samples are stratified by forecasted wind speed
Artificial Neural Network	ANN	Artificial neural network with one hidden layer

SMLR2 with Direct Power Prediction (SMLR2-DP)

The second alternate method is a variant of the SMLR2 scheme. In this scheme (designated as SMLR2-DP) the power production is predicted directly from the physics-based model variables instead of first predicting the wind speed at the wind plant's anemometer site and then using a plant output model to forecast the power production as in the SMLR and SMLR2 schemes. The reasoning is that some model variables may implicitly provide useful information about variations in power production.

Stratified Two-Stage SMLR (SMLR2-ST)

The third approach is a stratified two-stage SMLR. The approach is conceptually similar to the SMLR2 scheme except that separate prediction equations are derived for subsets of the training sample. The content of the subsets are determined through the use of one or more stratification parameters. The selection of stratification parameters was subjective although, in principle, some type of cluster analysis could be used to define the subsets optimally. In these experiments, the physics-based model wind speed for the wind plant site was selected as the stratification parameter. This parameter was used to divide the training sample into four subsets, each having approximately the same sample size. The SMLR procedure was then independently applied to each subset.

Artificial Neural Network (ANN)

The fourth alternative is a method based upon an artificial neural network (ANN) model. The basic computational element of an ANN is called a neuron. A neuron accepts input

values and processes them into an output (a response) value. An ANN is typically composed of a set of inputs, one or more hidden layers of neurons and a set of outputs. Figure 4-31 shows an example of a neural network. Each circle represents a neuron. The net input into each neuron consists of a weighted sum of all the individual inputs into that neuron. The output of the neuron is obtained by allowing an activate function to operate upon the net input. The training of the network consists of finding an optimal set of weights from the training sample data.

In the application of ANN for the day-ahead MOS forecast in this project, a feed forward network with one hidden layer and one output was used. Within this general structure, several different network configurations were evaluated on a couple of the test months. The alternate ANN configurations were generated by varying the number of input and hidden layer neurons. A total of 56 different variables were available from the physics-based model to serve as input into the ANN. Experiments were conducted by using a subset of the set of 56 variables as the inputs as well as varying the number of hidden layer neurons. After evaluating the results, it was decided that all 56 variables would be used as input and that a total of three hidden layer neurons would be used. It was thought that the use of all 56 physics-based model variables as inputs would provide the best comparison to the SMLR schemes since it used all 56 variables in their screening process. The selection of three hidden layer neurons was based on the fact that the use of larger number of hidden layer neurons in the preliminary tests yielded no better and often worse performance. The use of large numbers of hidden neurons with a relatively small training sample size can result in overfitting problems, since each additional neuron introduces new parameters (or the weights) into the ANN prediction system.

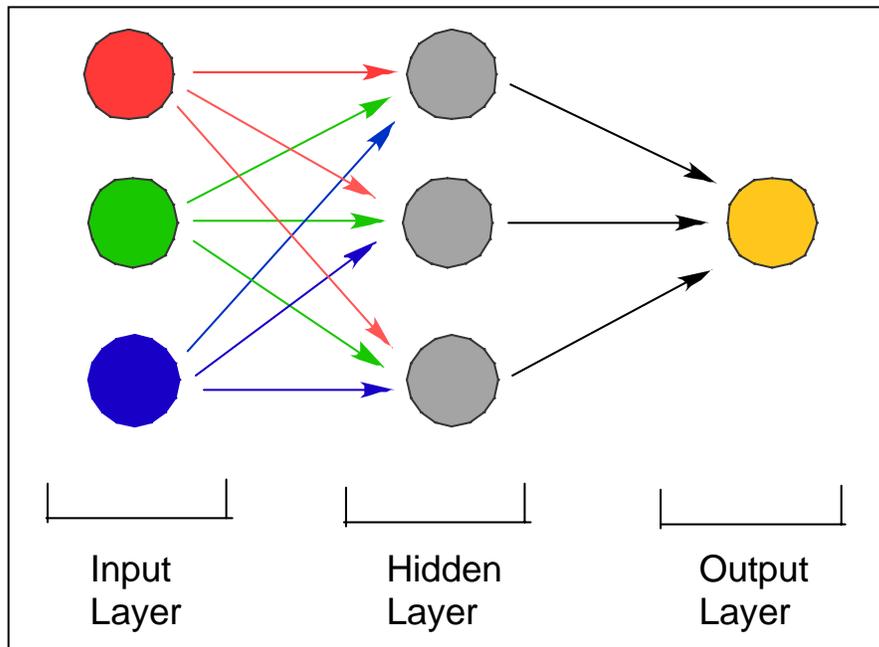


Figure 4-31 A schematic depiction of the Artificial Neural Network (ANN) configuration used for the forecast experiments in Focus Area 4.

Figures 4-32 and 4-33 present the monthly MAEs of the hourly power production forecasts for the entire 1- to 48-hour forecast period for all five MOS techniques for the PowerWorks and Mountain View wind plants, respectively. These results are based on the 4- km MASS simulations produced in the Focus Area 3 experiments.

There was generally very little difference between the performance of the SMLR2 method and that of the SMLR method with the exception of August 2002 for the Mountain View wind plant. Curiously, the SMLR method performed very poorly for this month, and all of the other methods including the SMLR2 method performed much better than it did. However, in view of the results for the other five months, this appears to be the result of an anomaly in the SMLR performance and is probably not representative of the typical performance relationship between SMLR and the SMLR2 methods.

The SMLR-DP method, which directly predicts the power production from the physics-based model output data, generally did not perform quite as well as the SMLR2 method. The monthly MAE for the SMLR-DP method was higher during four of the six test months, including all three of the months for the PowerWorks wind plant. The MAE was slightly lower during April and August 2002 for the Mountain View wind plant. In the view of the small sample and the small difference, the appropriate conclusion is most likely that there is no significant performance difference between the SMLR-DP and SMLR2 methods.

The stratified SMLR2 approach (SMLR2-ST) produced a substantial and consistently lower MAE than the unstratified SMLR2 scheme. The SMLR2-ST methods yielded a lower MAE than the SMLR2 method in all six of the test months, and the average magnitude of the improvement was 10%. Although the sample size is limited, the results clearly suggest that the SMLR2-ST method produces a significant improvement over the SMLR2 scheme.

The ANN method yielded net forecast performance results that were very similar to those produced by the SMLR2-ST method. Each method exhibited a lower MAE in three of the months. The actual net improvement statistics indicated that the ANN method had a slight performance advantage over the SMLR2-ST method with a net percentage reduction in MAE of 0.35%, and the net MAE reduction increment was 0.55% of installed capacity. However, these net error reductions are small relative to the monthly variations of MAE changes.

For example, the ANN method outperformed the SMLR2-ST method by an increment of 0.9% of installed capacity or 5.6% of the SMLR2-ST MAE of 16.1% of installed capacity for July 2002 for the PowerWorks plant. However, the ANN method performed worse than the SMLR2-ST method by an increment of 0.5% of installed capacity (4.1% of the SMLR2-ST MAE) In view of the limited sample size, the small net difference in the results, and large variability in the monthly results, the most reasonable conclusion from these results is that there is no significance difference in the performance between the two methods.

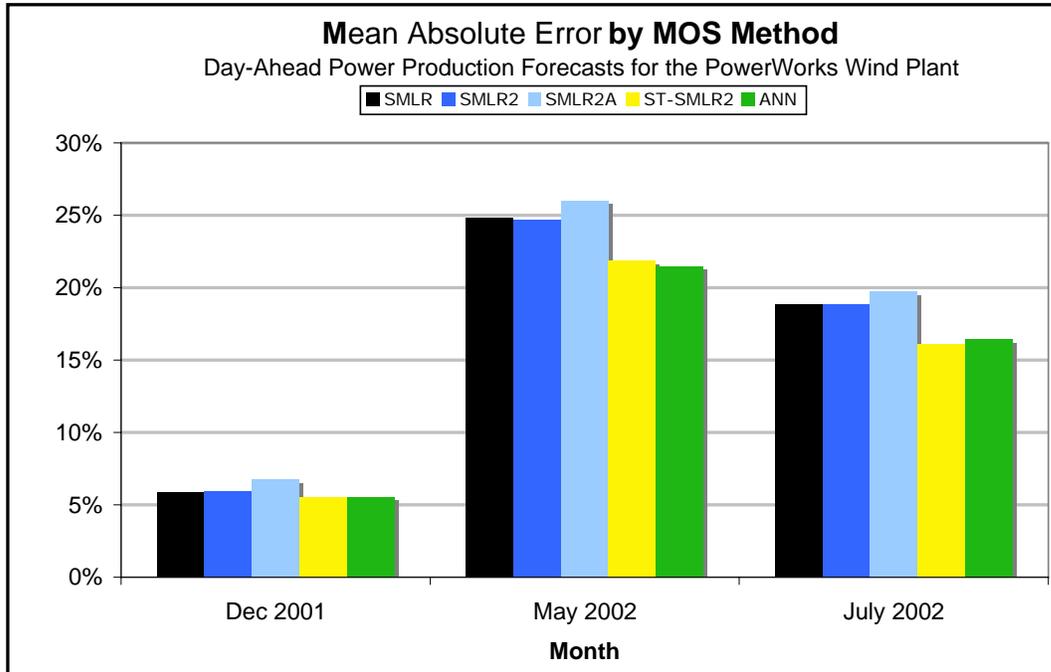


Figure 4-32 Comparison of the monthly MAE of 1- to 48-hour forecasts of the hourly power production for the three test months at the PowerWorks wind plant for the five MOS procedures listed in Table 4-5.

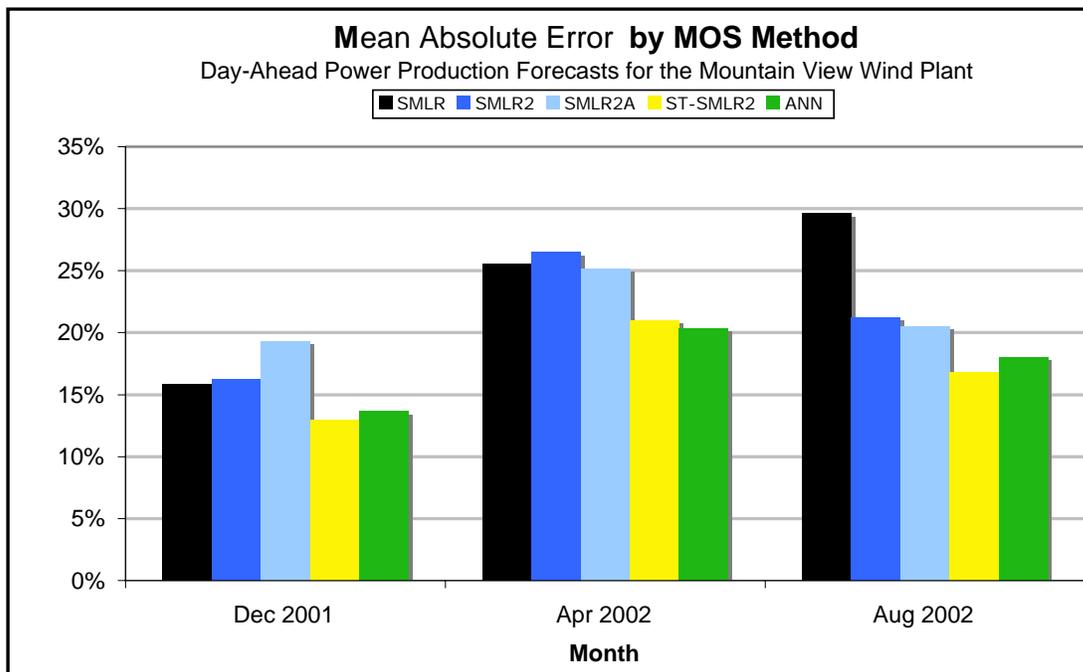


Figure 4-33 Comparison of the monthly MAE of 1- to 48-hour forecasts of the hourly power production for the three test months for the Mountain View wind plant for the five MOS procedures listed in Table 4-5.

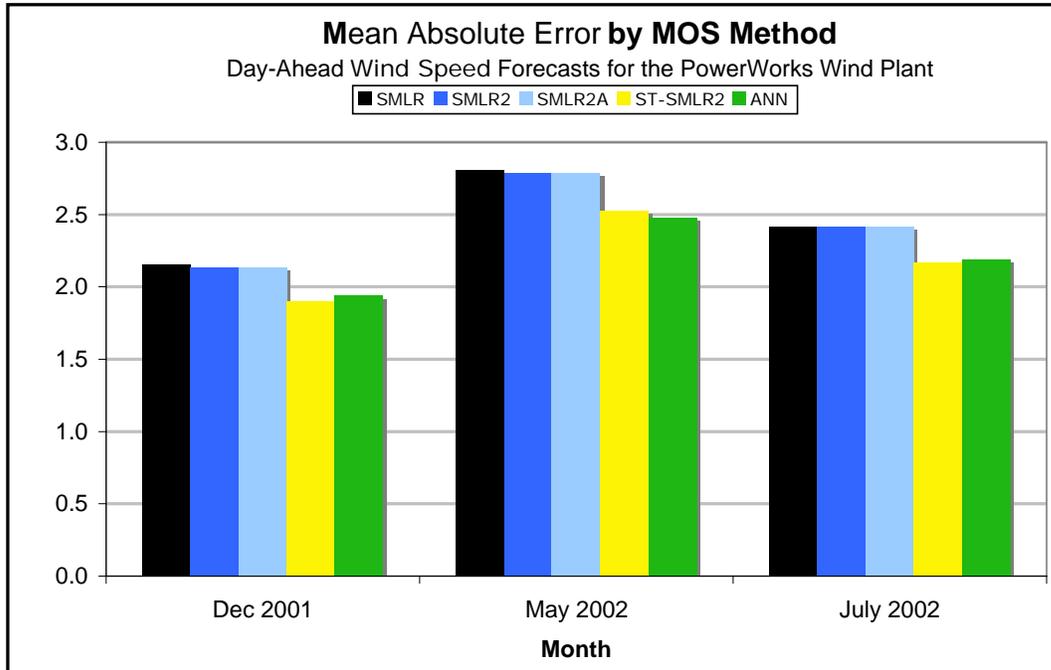


Figure 4-34 Comparison of the monthly MAE of 1- to 48-hour forecasts of the hourly average wind speed for the three test months for 11 anemometer locations at the PowerWorks wind plant for the five MOS procedures listed in Table 4-5.

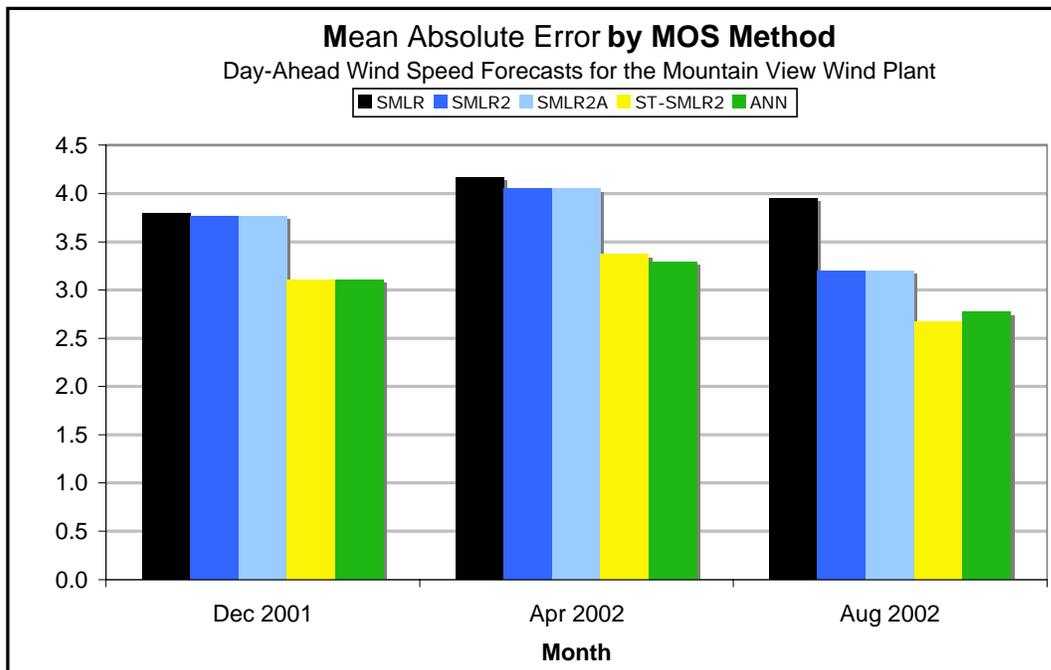


Figure 4-35 Comparison of the monthly mean absolute error of 1- to 48-hour forecasts of the hourly average wind speed for the three test months for the Catellus anemometer site in the Mountain View wind plant for the five MOS procedures listed in Table 4-5.

Figures 4-34 and 4-35 summarize the performance of the same five MOS methods for the wind speed forecasts is depicted. Not surprisingly, the performance is quite similar to that of the power production forecasts. The SMLR and SMLR2 methods yield very similar results except for August 2002 at the Mountain View wind plant. The net differences of the monthly MAEs for the other five months is only -0.19 m/s or an average of approximately -0.04 m/s per month. The SMLR2 method yielded an MAE that was 0.75 m/s lower for August 2002, but it appears this occurred because the SMLR method yielded an anomalously high MAE for that month. The SMLR and SMLR-DP methods exhibited identical MAEs for all six months. This is because they use the same procedure to produce the wind speed forecasts. The crucial difference is that the SMLR-DP method directly predicts the power production from the physics-based model variables (without the use of a plant output model,) and the SMLR2 method uses a separate plant output model to create the power production forecasts from the wind speed forecasts. As in the case of the power production forecasts, the SMLR2-ST and the ANN methods both produce a significant MAE reduction relative to the SMLR, SMLR2, and SMLR-DP methods during all six test months.

Analysis of the MAEs. vs. forecast look-ahead period in Figures 4-36 and 4-37 can provide an understanding of the characteristics of the MAE reduction produced by the SMLR2-ST scheme relative to the SMLR2 scheme. These charts indicate that a substantial portion of the MAE reduction for both wind plants is attributable to significantly lower errors during the first 24 hours of the forecast period. The performance differential is especially prominent for the Mountain View wind plant (Figure 4-37).

During the 1- to 24-hour period, the SMLR2-ST method reduces the MAE by an increment of 4.3% of installed capacity, which is a 23.7% improvement over the SMLR2 method. However, during the 25- to 48-hour look-ahead period, the MAE reduction increment is only 1.7% of installed capacity, which corresponds to an improvement of 9% over the SMLR2 method. This MAE pattern is also present in the May 2002 forecasts for the PowerWorks wind plant, but it is less prominent. The improvement percentage of SMLR2-ST over SMLR2 is 13.3% for the first 24 hours and 9.7% for the second 24 hours. Another noteworthy feature on the charts of Figures 4-36 and 4-37 is that the structures of both the SMLR2 and SMLR2-ST curves are quite similar for both wind plants.

Some insight into the variability of the performance differences between the SMLR2 and SMLR2-ST methods can be gained from the time series of the hour-by-hour differences in the absolute error between the two methods. Figures 4-38 and 4-39 show the time series for the first 24 hours of the forecast period for May 2002 for the PowerWorks and Mountain View plant, respectively. These charts indicate that the SMLR2-ST method consistently outperforms the SMLR2 scheme (for example, the 24-hour running average rarely falls below zero). In addition, there are a number of episodes in which the SMLR2-ST scheme outperforms SMLR2 by a large margin.

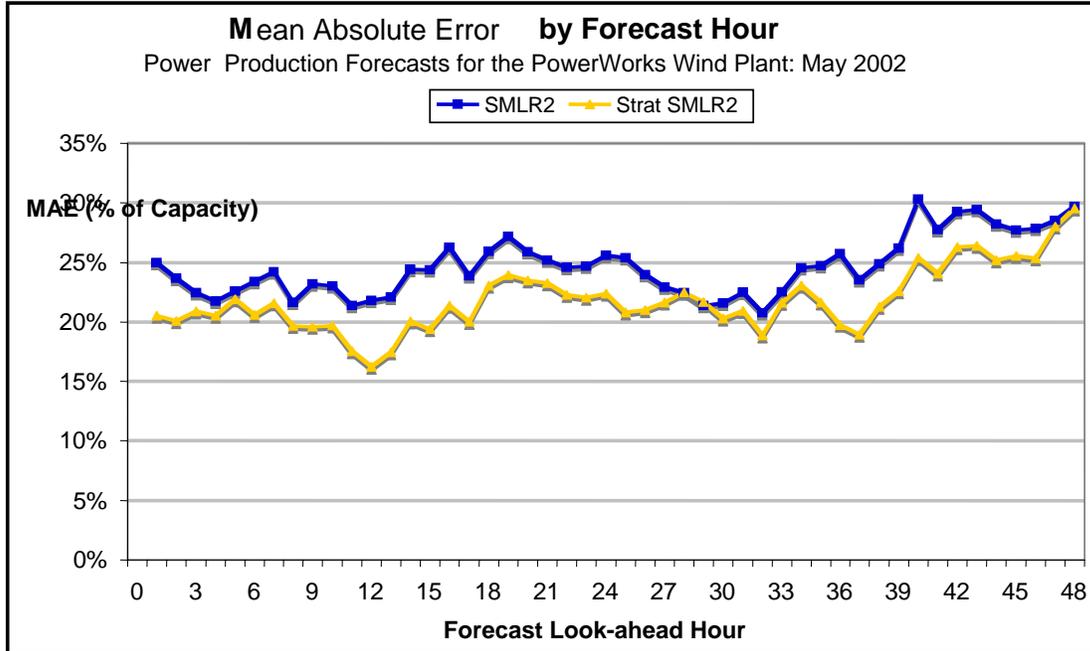


Figure 4-36 Mean absolute error of the power production forecasts vs. forecast “look-ahead” hour from the raw SMLR2 (blue line with rectangle markers) and stratified SMLR2 (yellow line with triangle markers) MOS procedures for the PowerWorks wind plant for May 2002.

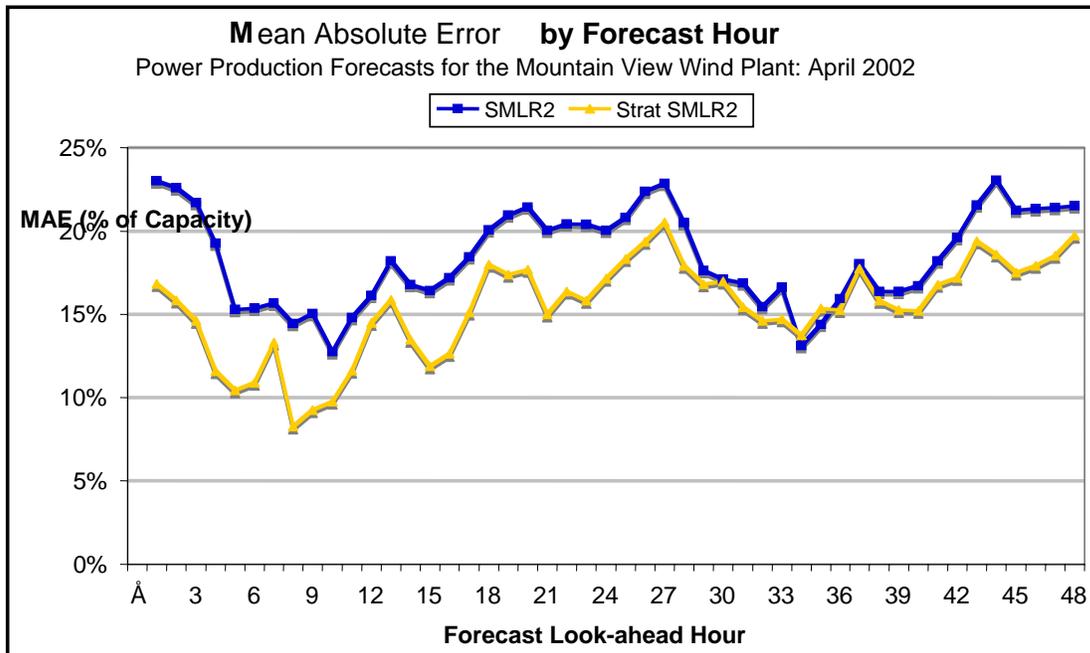


Figure 4-37 Mean absolute error of the power production forecasts vs. forecast “look-ahead” hour from the raw SMLR2 (blue line with rectangle markers) and stratified SMLR2 (yellow line with triangle markers) MOS procedures for the Mountain View wind plant for April 2002.

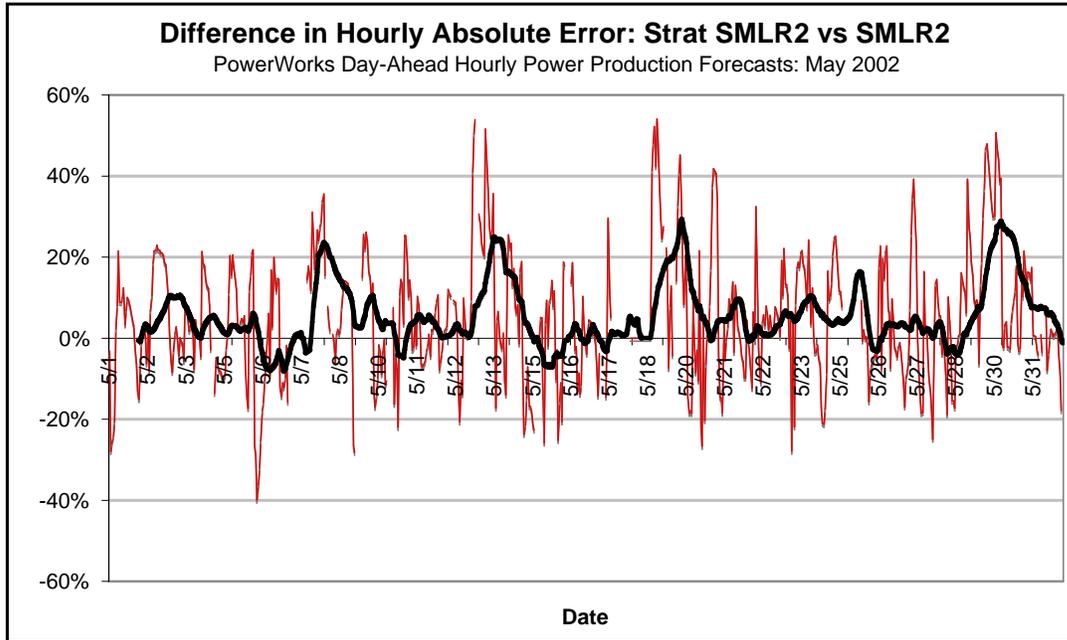


Figure 4-38 Differences between the hourly absolute errors of the SMLR2 and SMLR2-ST MOS 1- to 24-hour forecasts (red line) of the hourly power production of the PowerWorks wind plant for May 2002. Positive differences indicate that the SMLR2-ST forecasts have a lower absolute error. The black line is the lagged 24-hour moving average.

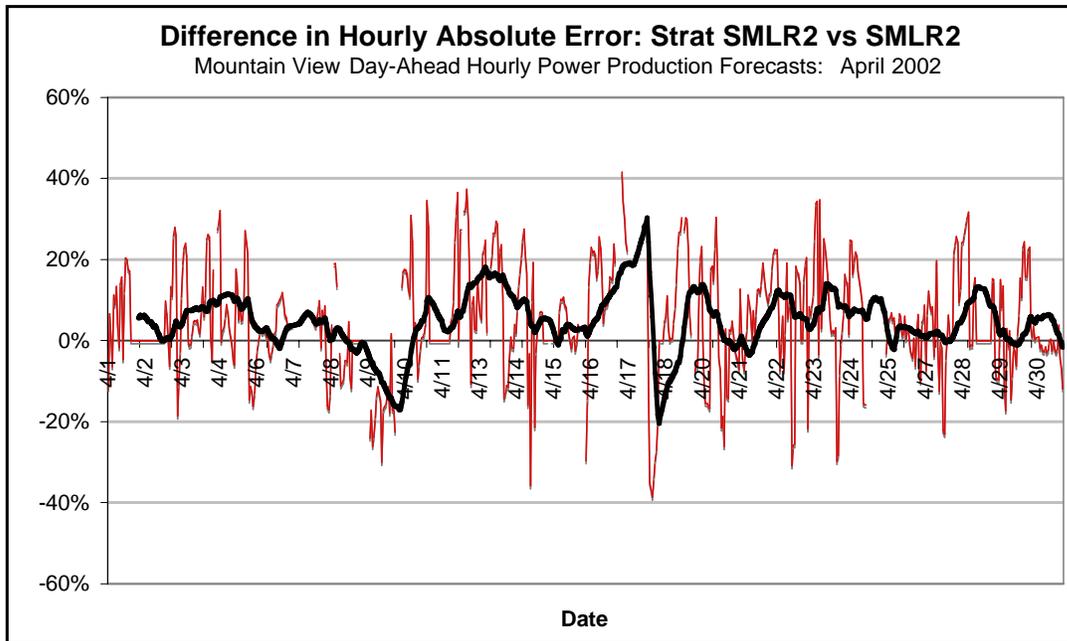


Figure 4-39 Differences between the hourly absolute errors of the SMLR2 and SMLR2-ST MOS 1- to 24-hour forecasts (red line) of the hourly power production of the Mountain View wind plant for April 2002. Positive differences indicate that the SMLR2-ST forecasts have a lower absolute error. The black line is the lagged 24-hour moving average.

Conclusions

Three conclusions were derived from the six months of experimental forecasts using the set of four alternate MOS procedures.

First, the two-stage screening multiple linear regression method (SMLR2), which included the use of parameters that measured the recent forecast error in the second stage, yielded no significant improvement over the baseline SMLR scheme.

Second, the use of a direct power prediction version (no explicit plant output model) of the SMLR2 scheme did not yield any significant improvement over the SMLR2 scheme, which used a plant-scale power curve to predict the power production. In fact, the overall performance of SMLR2-DP was generally slightly worse for the six test months.

Finally, the most significant conclusion is that both the stratified SMLR scheme (SMLR2-ST) and the ANN scheme produced a significant reduction in the MAE relative to the baseline SMLR scheme for all six of the test months.

Figure 4-40 presents a summary of the percentage MAE reduction achieved by the SMLR2-ST method relative to the SMLR2 method for each of the six months. These two schemes appear to offer substantial potential to yield significant forecast improvements. However, the ANN scheme requires much more computational power for training and, based on the results from this limited sample, does not produce any additional benefit in forecast performance over the SMLR-ST scheme. Therefore, the SMLR2-ST method was selected for use in the one-year evaluation period in the second phase of this task.

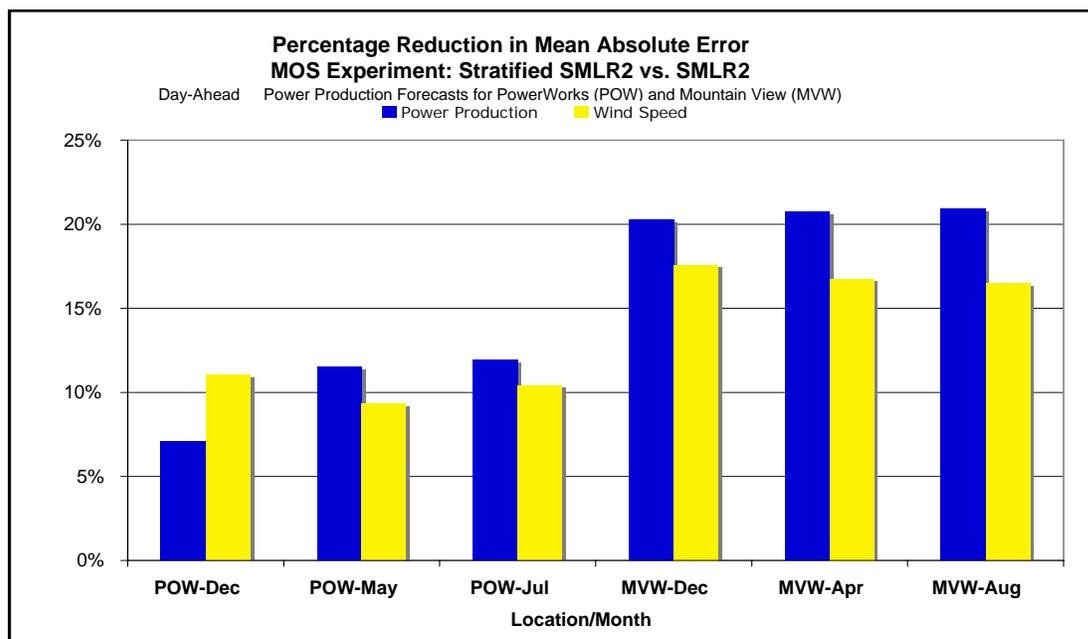


Figure 4-40 The percentage reduction in the MAE of the SMLR2-ST power production and wind speed forecasts relative to forecasts produced by the SMLR2 method for the sample of six test months at the PowerWorks and the Mountain View wind plants.

Focus Area 5: Plant Output Model Formulation

The plant output model is the link between the meteorological conditions at a wind plant and the plant's power production. There are several ways to formulate a plant output model. In fact, it is possible to totally avoid the use of an explicit plant output model by training the MOS procedure discussed in the previous section to directly forecast the power production from the output data of the physics-based model. This is the approach used by the SMLR2-DP method. However, the results in the previous section indicated that this approach generally did not perform as well as the technique of first forecasting the wind speed at one or more anemometer locations within the wind plant and then using a plant output model to forecast the power production. This is generally consistent with AWST's overall experience, which indicates that it is usually better to use the MOS procedure to make a prediction of the relevant meteorological parameters at the wind plant site and then to use those predicted quantities as input into an explicit plant output model to produce the power production predictions. The plant output model can be formulated as a statistical or a physical numerical model.

Physical Plant Output Models

In the physical model approach, a representation of the turbine layout and an explicit model of the wind flow throughout the area of the wind plant are employed. The model accounts for flow variations within the plant domain due to terrain and roughness effects and also calculates the impact of the effects of turbine wakes on nearby turbines. The wake effects may become particularly important for certain wind directions that result in air flow along the rows of turbines rather than across the turbine rows. Once the relevant meteorological conditions (for example, wind speed, direction, and air density) at each turbine are known, the power production for each turbine can be calculated from the manufacturer's power curve for the particular wind turbine. The individual power production values for each turbine can then be aggregated to obtain a prediction of the power production for the entire wind plant. Physical models of this type can provide a great deal of insight into the reasons for variations in power production but they are very complex, often use significant computational resources, and require detailed knowledge of the wind plant layout and the terrain and roughness variations in the vicinity of the plant. This makes physical plant output models difficult to configure and operate.

Statistical Plant Output Models

The other major type of plant output model is the statistical model. In this approach, power production and meteorological data from the plant are used to construct an empirical relationship between the plant-scale power production and the meteorological conditions. The simplest model of this type is commonly known as a "plant-scale power curve." This is simply an empirical equation derived from the plant's power production data and the wind speed data from one or more of the plant's anemometers. The power curve is used in the forecast process to calculate the forecasted power production from meteorological predictions of the wind speed at the plant's anemometer locations. This is the approach that

was used in the first Energy Commission-EPRI forecasting project. However, data from wind plants indicate that there are substantial deviations from the behavior modeled by this basic type of “plant-scale power curve.” Thus there is an opportunity to improve the wind power production forecasts if a significant fraction of the variability of the deviations from the wind-speed-based power curve can be forecasted.

Plant Scale Power Curves

Figures 4-41 and 4-42 show two examples of plant-scale power curves. The black triangles show the hourly power production and average hourly wind speed data from turbine clusters M127 and M225 in the PowerWorks wind plant for July 2002. The red circles denote the power production prediction from a basic plant-scale power curve applied to the wind speed measured at the anemometer associated with each cluster. The plant-scale power curve used in this calculation is based on a polynomial curve fit to the median hourly power production values for a set of 1-m/s wind speed bins based on data from June 2002. Thus, any systematic changes in the operating characteristics of the wind plant between June 2002 and July 2002 will result in a systematic difference between the hourly point values (black triangles) from July and the power curve predictions (red circles) based on the June data. The data from the PowerWorks wind plant in Altamont Pass were used to analyze the magnitude of this source of forecast error and to experiment with ideas to reduce this type of forecast error.

The relationship between the predictions from the power curve and the July 2002 data is quite good for the M127 cluster (Figure 4-41). The predictions from the power curve are roughly in the center of the cloud of points formed by the July power production and wind data. This is confirmed by the fact that the overall bias in the power curve predictions is +0.3% of the installed capacity. This indicates that the relationship between the wind speed measured by the anemometer and the power production was roughly the same in June and July for the M127 turbine cluster. In addition, the vertical spread of power production values for a particular wind speed is relatively small. This means that an accurate prediction of the anemometer wind speed will yield a good prediction of the power production for this cluster. For the M127 cluster, during July, a perfect forecast (MAE = 0.0 m/s) for hourly wind speed at the anemometer site would have yielded an MAE for the power production prediction of 4.2% of the installed capacity.

In contrast, the power curve predictions of the July 2002 power production from turbine cluster M225 are not as good (Figure 4-42). There are many more points above the power curve than there are below the curve. This indicates that the power curve tended to provide a power production value that was too low for a particular wind speed during July. In fact, the prediction bias (predicted minus reported values) over all hours of the month was -6.7%. This indicates that there was an apparent change in the relationship between the wind speed measured at the anemometer site and the power production between June and July. In addition to this change, the spread of the power production values at a particular wind speed is generally much greater for cluster M225 than it was for cluster M127.

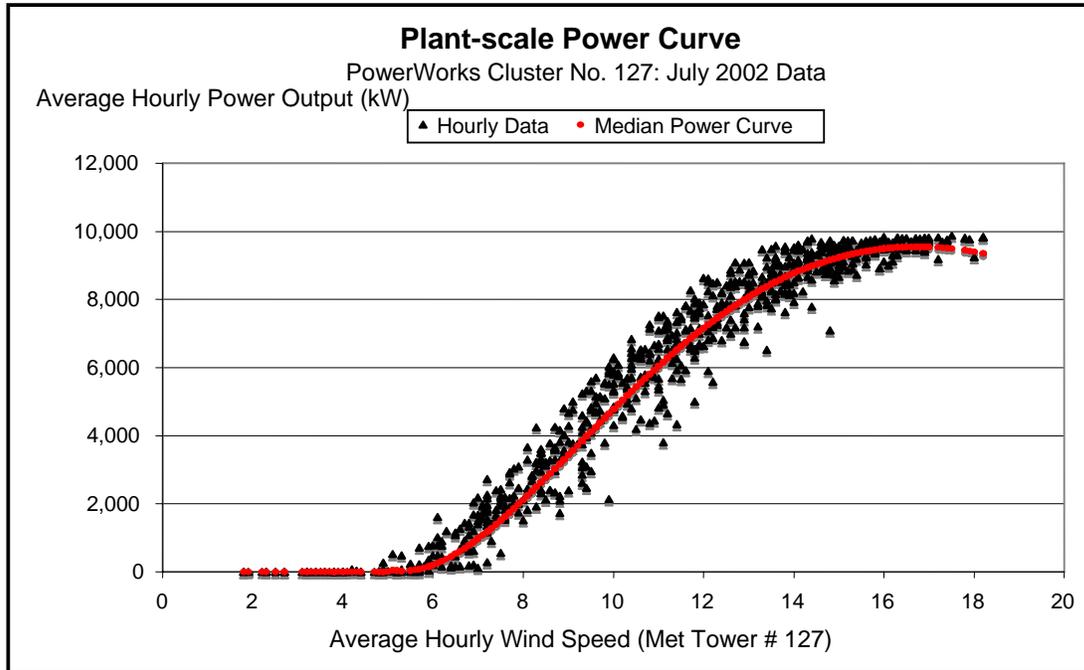


Figure 4-41 Hourly power output vs. measured hourly average wind speed (black triangles) for the PowerWorks M127 cluster for July 2002 and hourly power output estimates (red circles) using a plant-scale power curve derived from June 2002 data.

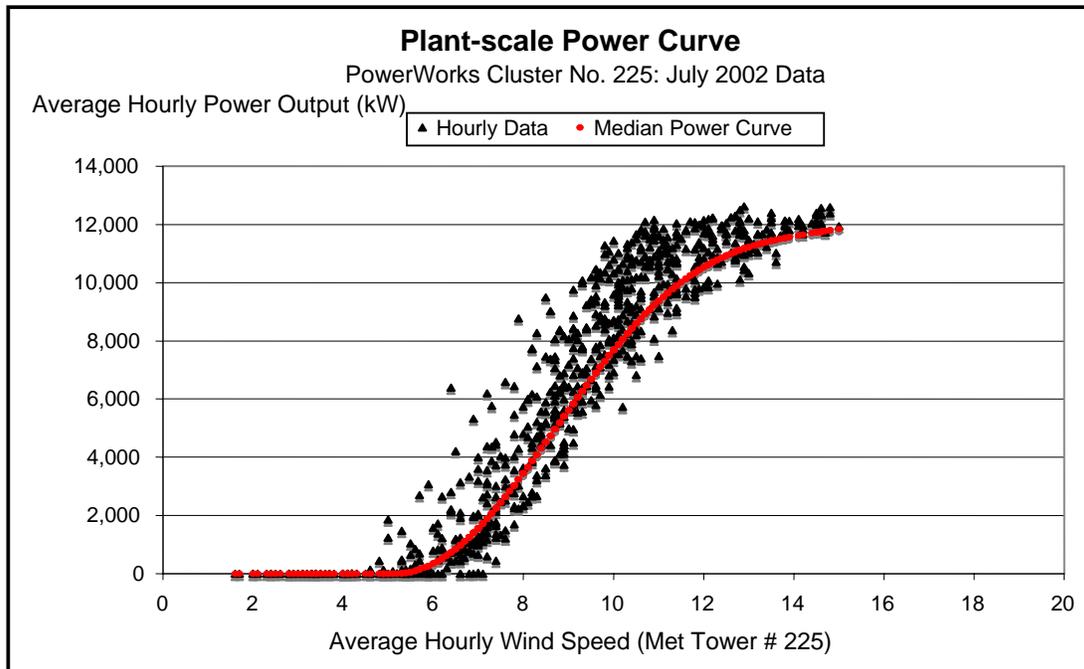


Figure 4-42 Hourly power output vs. measured hourly average wind speed (black triangles) for the PowerWorks cluster M225 for July 2002 and hourly power output estimates (red circles) using a plant-scale power curve derived from June 2002 data.

The reason for the offset relative to the June values and the greater spread relative to the cluster no. 127 data is not known. It is most likely related to variations in the wind patterns within the arrays of cluster M225 turbines. It is also possible that it could be related to measurement issues (for example, an anemometer that is out of calibration or has other maintenance issues). Regardless of the cause, both of these factors introduce additional error into the power production predictions for cluster M225. In this case a perfect wind speed forecast for the anemometer site will result in a mean absolute error of 8.5% of installed capacity for the power production predictions. This is very large in comparison with the typical day-ahead forecast error of about 15% to 20% of installed capacity. It is approximately twice the MAE associated with the M127 cluster for the same month. This illustrates how the variability in the relationship between the wind speed measured at the anemometers sites and the concurrent power production can have a significant impact on the accuracy of the power production forecasts.

Modeling Deviations from the Power Curve

One possible approach to improve the power production predictions is to model the deviations from the power curve. To do this, it is necessary to identify parameters that explain a substantial amount of the variance from the power curve predictions. One possible parameter is the wind direction.

Figures 4-43 and 4-44 are scatter plots that show the relationship of the deviation from the power curve to the wind direction for PowerWorks clusters M127 and M225. The most striking characteristic of these plots is that there is a very limited range of wind directions for both of these anemometer sites during July 2002. This is quite typical for these sites, especially during the summer months. The range is slightly greater for tower M225 than it is for M127. There appears to be very little relationship between the deviation from the empirical plant-scale power curve and the wind direction for cluster M127 (Figure 4-43).

However, the pattern is somewhat different for cluster M225 (Figure 4-42). In cluster M225 the more southerly wind directions (180 to 210 degrees) tend to have a significant positive deviation, while the more westerly directions (210 to 240 degrees) tend to yield deviations that are near zero or negative. Thus, it appears that wind direction may be a useful predictor for cluster M225 but not for M127.

Another possible parameter is the hour of the day. Figures 4-45 and 4-46 show the relationship of the deviation between the empirical plant-scale power curve and the hour of the day. These charts indicate that there is a distinct relationship between the hour of the day and the deviation from the power curve for both the M127 and M225 clusters. For the M127 cluster (Figure 4-45), negative deviations from the power curve predictions tend to occur during the mid-afternoon to late evening period (3:00 p.m. to midnight PDT). Positive deviations tend to occur during the period from slightly after sunrise to early afternoon (7:00 a.m. to 2:00 p.m. PDT). The red line in Figure 4-45 is a least squares polynomial curve fit to the deviation and hour of the day data. This indicates that the amplitude of the diurnal oscillation in the deviations is about 600 kW (about 6% of installed capacity). The diurnal pattern for the M225 cluster is somewhat different from that for the

M127 cluster. The large positive deviations for the M225 cluster tend to occur during the afternoon and evening (2:00 p.m. to midnight PDT), while the smaller positive or negative deviations tend to occur during the morning hours. The least squares polynomial curve fit (red line in Figure 4-46) indicates that the amplitude of the oscillation for the M225 cluster is about 1900 kW (14.6% of installed capacity). The analysis of the data from the M127 and M225 clusters indicate that the use of the wind direction and the hour of the day parameters to model the variation in the deviations from the power curve has some potential to reduce the forecast error associated with these deviations. The quality of the deviation relationships varies substantially from cluster to cluster within the PowerWorks plant and also between wind plants. Therefore, a thorough analysis of the wind and power production data from each plant must be done to identify the parameters that can explain the variance of the power curve variations.

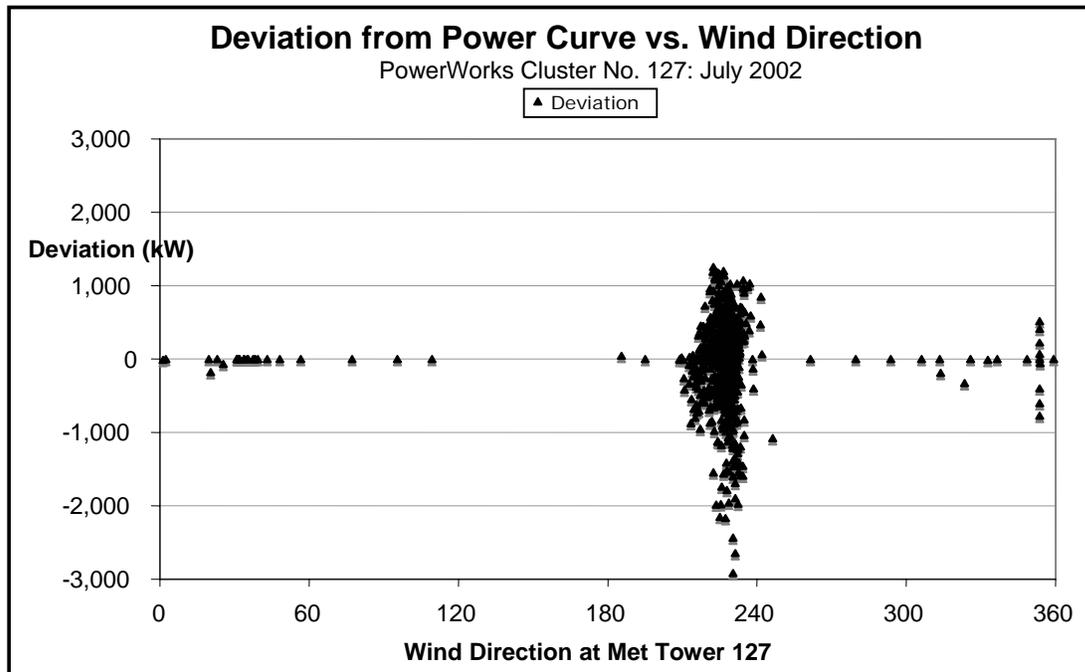


Figure 4-43 Hourly deviation (kW) from the median based power curve as a function of wind direction at Tower M127.

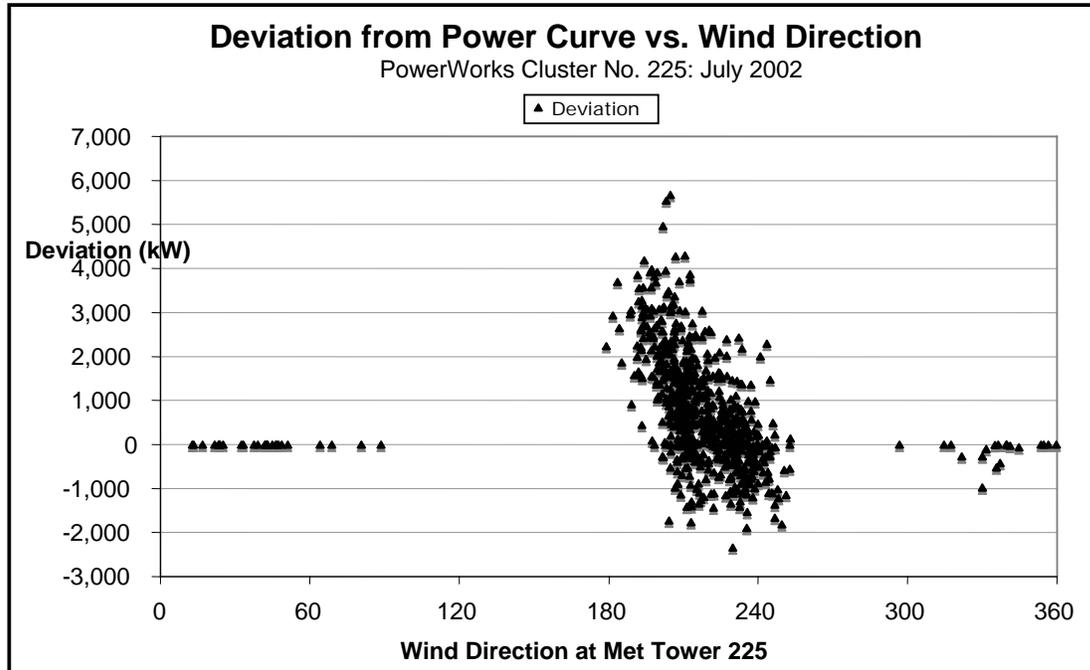


Figure 4-44 Hourly deviation (kW) from the median based power curve as a function of the wind direction at Tower M225.

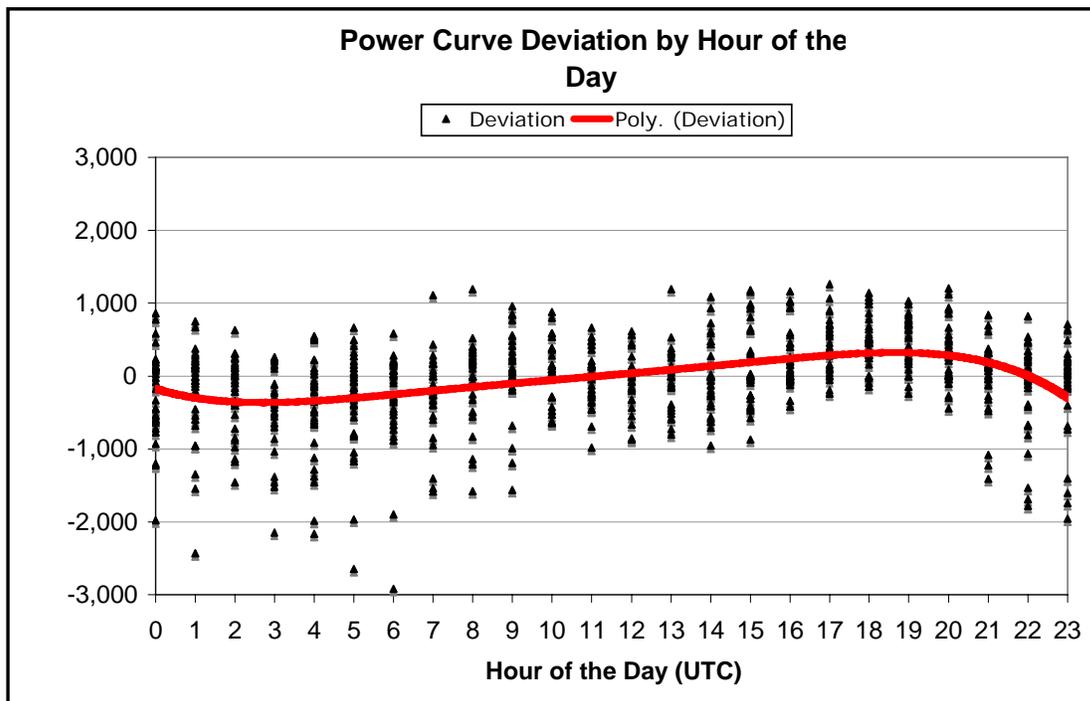


Figure 4-45 Hourly deviation (kW) from the median-based power curve as a function of the hour of the day (UTC) for the PowerWorks M127 cluster for July 2002.

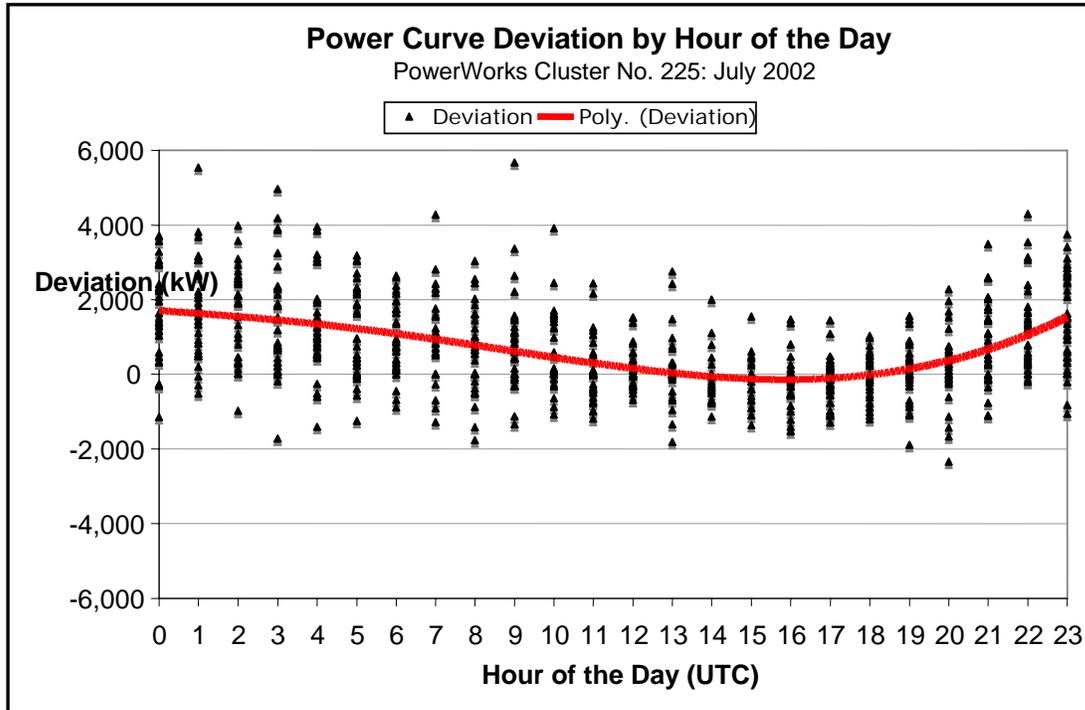


Figure 4-46 Hourly deviation (kW) from the median-based power curve as a function of the hour of the day (UTC) for the PowerWorks M225 cluster for July 2002.

To estimate the effect of a more sophisticated approach to the plant output model on the accuracy of the power production forecasts, the deviations from the median-based power curve were modeled for both the PowerWorks and Mountain View wind plants through the use of the wind direction and the time of day parameters. Forecasts with this new plant output model were then produced for all six of the evaluation months. Figure 4-47 illustrates the resulting reduction in the power production forecast MAE for each plant and evaluation month. Interestingly, the percentage MAE reduction for the PowerWorks wind plant is in the 6% to 9% range for all three of the evaluation months, whereas it is 2% or less for all three months for the Mountain View wind plant. The reductions are relative to a plant output model that did not model the deviations.

The significantly greater effect for the PowerWorks plant may be because of the more complex terrain within the turbine arrays at that site. It may also be because the wind direction and time of day are not the best parameters to model the deviations from the power curve for the Mountain View plant. Thus, there may further performance gains to be realized by further extending the sophistication of the plant output model. However, this initial attempt at improving the plant output model produced a significant improvement in forecast performance.

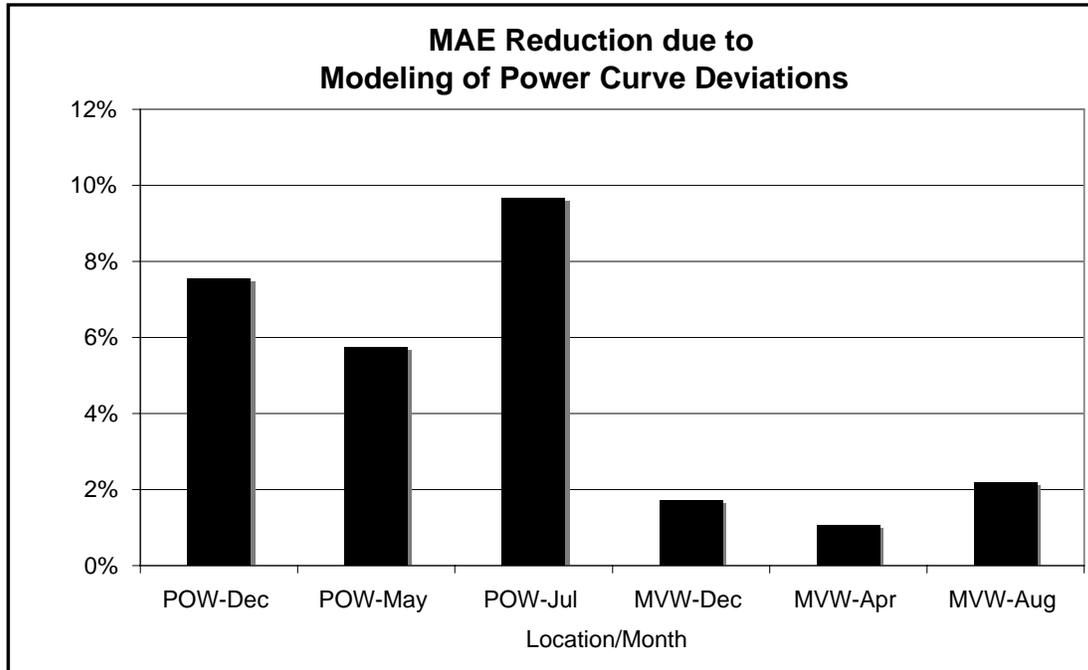


Figure 4-47 Reduction in MAE of the power production forecasts due to the modeling of the deviations from the median-based plant-scale power curve for the three test months for the PowerWorks (POW) and Mountain View (MVW) wind plants.

Focus Area 6: Forecast Ensembles

Forecast ensemble models are statistical models that produce an optimal forecast by compositing forecasts from a number of different techniques (Sivillo, et al., 1997). The use of forecast ensemble models is based on research that has demonstrated that a composite of forecasts from an appropriate ensemble of forecast-generating techniques is often superior to those produced by any one member of the ensemble. A conceptual schematic of the ensemble forecasting process is depicted in Figure 4-48.

The fundamental concept is that if the errors in the forecasts produced by the different methods are unbiased and have a low degree of correlation with one another, the random errors from the individual forecasts will tend to offset each other, with the result that a composite of the forecasts will have a lower error than any individual forecast. If all of the input forecasts are highly correlated, the impact of an ensemble forecast will be minimal. This means that the underlying forecast methods must be quite different in how they construct the relationships between the raw observational data and their forecasts or the type, amount, or quality of input data going into the methods must be significantly different. This "ensemble effect" is a well-known technique used by meteorologists in intermediate and extended-range forecasting. The spread of the forecasts produced by the ensemble may also be related to the forecast uncertainty if the differences in the ensemble members are related to the primary factors that introduce uncertainty into the forecasts.

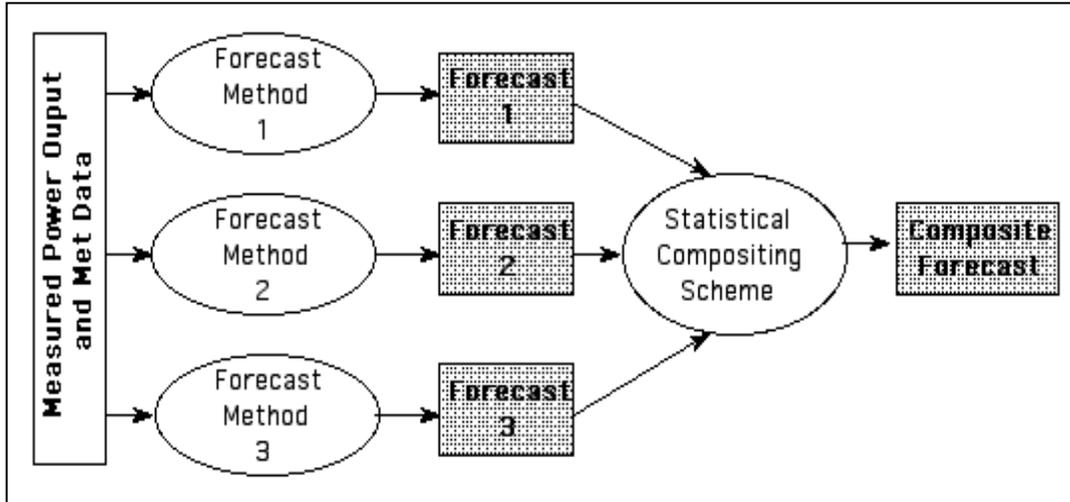


Figure 4-48 A conceptual schematic of the ensemble forecast method.

This ensemble approach can be applied to most forecast look-ahead time periods (hour-ahead, day-ahead). There are two fundamental strategies that can be used to generate an ensemble of forecasts. One strategy is to use the same forecast models and vary the input data within their range of uncertainty. Complex methods such as “breeding methods” are sometimes used to determine the initial perturbations that lead to greater forecast differences (Toth and Kalnay, 1997). The other strategy is to use the same input data and employ different forecast models or different configurations of the same model. The relative value of either strategy depends upon the sources of uncertainty in the forecast procedure. In practice, the sources of uncertainty vary with location, season, and other factors, and thus the choice of the ensemble components and the number of members must be determined from experience and experimentation.

The National Centers for Environmental Prediction (NCEP) has an ongoing program of experimentation with ensembles, mostly for global-scale forecasting. Ensembles for medium-range forecasts (0-14 days) have been used operationally in North America and Europe since the 1990s. Short-range ensembling (0-3 days) is promising but less advanced. It continues to be the subject of active research. Hamill et al. (2000) list as the three most important current issues: “(1) how or how best to use ensemble information to improve data assimilation strategies; (2) how to address model errors in ensemble forecasts, and (3) how to appropriately use and interpret the voluminous information from ensembles”.

For some types of forecasting problems, a need for extremely high resolution or physical complexity in the physics-based simulations may make ensemble forecasting impractical. But the promise of both a better consensus forecast (for example, the ensemble mean forecast) and a useful estimate of uncertainty for each forecast continues to drive development. The relationship between ensemble spread and forecast skill is complex (Whitaker and Louge, 1998) and is sensitive to the details of the ensembling methods.

A comprehensive investigation of the potential benefits the ensemble method in wind power production forecasting is certainly beyond the scope of this project. However, it is possible within the scope of this project to investigate the benefits of constructing an ensemble of the different methods explored in this project's experiments. One of the most basic tests of the ensemble concept for wind power production forecasting is to construct an ensemble mean forecast from forecasts made by two different physics-based models. In the screening phase of this project, forecasts from the MASS and WRF models were generated under Focus Area 3. An ensemble mean forecast was constructed by averaging those two forecasts after the MOS adjustment procedure was applied to each. The result was that the ensemble mean forecast reduced the overall MAE of the best overall forecast by 0.8% for the five test months for which both forecasts were available. This is hardly a substantial improvement, but the size of the ensemble (two) was very small, and evidence from other studies indicates that the benefit from constructing an ensemble mean forecast is more significant for larger ensemble sizes.

A somewhat more extensive investigation of the benefits of the ensemble approach will be possible in the evaluation phase, since a greater variety of methods will be available during the one year of experimental forecasting under the second part of the next-day forecast development and testing task.

Testing of Improved Forecast Algorithm at Five California Wind Projects

The objective of the second phase of the day-ahead wind power production forecast development and evaluation task was to apply and evaluate the performance of the most promising next-day wind energy forecast methodologies identified in the previously described first phase of the effort over a one-year period at wind plants in the principal California wind resource areas.

Based on the substantial improvements in forecast performance relative to that in the previous Energy Commission-EPRI project, the improved forecast methodologies used in the one-year evaluation period included:

- Higher-resolution water surface temperature data from the MODIS sensor.
- “Next generation” physics-based atmospheric models.
- A stratified two-stage screening multiple linear regression (SMLR2-ST) scheme for the Model Output Statistics (MOS) component of the forecast system.
- Modeling deviations the empirical plant-scale power curve.
- Ensemble forecasting using a portfolio of models.

Experimental Design

There are a number of factors that must be considered when a forecast evaluation project is designed. These include (1) the forecast delivery time and the length of the forecast look-ahead period (also called the forecast time horizon); (2) the specific parameters that will be forecasted and evaluated; (3) the data that will be available for input into the forecast process; (4) the sites for which forecasts will be made; and (5) the performance statistics that will be used to evaluate the forecasts. The following subsections provide a summary of these factors for the one-year forecast evaluation project conducted in the second phase of this task. The design of this forecast evaluation experiment was formulated to be as similar as possible to the forecast evaluation project conducted within the previous Energy Commission-EPRI project to facilitate a comparison of the results from the two projects.

Forecast Delivery Time and Look-Ahead Period

The delivery time of operational forecasts can be a critical factor in an operational environment since power production schedules often have to be submitted within a particular time frame to adhere to market or grid management rules. Although power production forecasts can be produced and delivered for any delivery time, there are certain forecast production times that are better suited to take advantage of the most recent set of comprehensive atmospheric data. This is because the most comprehensive set of atmospheric measurements is gathered twice per day at 0000 UTC (4:00 p.m. PST) and 1200 UTC (4:00 a.m. PST). For this project the forecast delivery times were set at 8:00 a.m. PST (1600 UTC) and 8:00 p.m. PST (0400 UTC). This means that the first hour of the forecast period is the hour ending at 1700 UTC (9:00 a.m. PST) and 0500 UTC (9:00 p.m. PST). This is the same forecast delivery time as that used in the earlier Energy Commission-EPRI forecasting project.

The primary time scale of the forecast, which is also known as the look-ahead period, determines, in part, the methods that will be used in the production of the forecasts. The primary focus of this task, as was the case in the previous Energy Commission-EPRI project, is the day-ahead time frame. The forecast time period was set to the 1- to 48-hour period after forecast delivery time to be consistent with the time frame of the previous project. Thus, the morning forecasts (those delivered at 8:00 a.m. PST) extended from the hour ending at 9:00 a.m. on the day of forecast delivery to the hour ending at 8:00 a.m. two days after the day of forecast delivery. Similarly, the evening forecasts extended from the hour ending at 9:00 p.m. on the day of forecast delivery to the hour ending at 8:00 p.m. two days after the day of forecast delivery.

Input Data

It was assumed that no real-time data from the wind plants was available to the forecast system. The same assumption was made in the previous project. This assumption has a large impact on the performance of the forecasts for the first few hours of the forecast period. This is because the wind power production and the trends in the production typically have at least a moderate degree of persistence over time scales of a few hours.

Thus, knowledge of the wind speed and direction and the power production at times immediately preceding the forecast delivery time provides a considerable amount of information about the power production over the next few hours. Forecasts that do not have access to this information will typically not perform even as well as a simple persistence forecast for the first few hours of the forecast period. However, the effect of recent data from the wind plant on the performance of longer period forecasts (beyond about six to nine hours) is very minimal.

The decision not to use real-time wind plant data in the forecasting process was based on three factors: (1) near real-time data was not available from all of the participating wind plants; (2) the focus of the task was on day-ahead forecasting for which real-time plant data does not play an important role; and (3) a desire to maintain consistency with the forecast evaluation done in the previous project, which did not use real-time data from the wind plant. The bottom line is that the forecasts produced for the first few hours of the forecast period in this task were not as good as they could have been had they used recent data from the wind plant.

Forecast Parameters

The forecasted parameters consisted of the hourly total power production from each wind plant and an average hourly wind speed and direction for one or more anemometer site on each wind plant for each hour in the 48-hour forecast period. In some cases, power production data were available for clusters of turbines within the wind plant. In those cases, forecasts of the power production were made for each cluster for which data was available, and then the predictions for each cluster were aggregated to produce the power production prediction for the entire wind plant. However, only the forecasts of the production of the entire wind plant were evaluated in this task.

The wind speed and direction was forecasted for a site within the wind plant at which wind speed and direction measurements were available. In some cases, the wind speed and direction data were available from several meteorological towers within the wind plant. If sufficient information was available to associate each tower with a cluster of turbines for which power production data were available, then wind speed and direction forecasts were made for each meteorological tower to support the power production predictions for each cluster. In those cases, the wind speed and direction predictions for only one of the towers were used in the forecast evaluation process. In cases in which there were more than one meteorological tower but no information about the relationship of the towers to turbine clusters was available, a composite of wind speed and direction from all of the towers was constructed, and the wind speed and direction forecasts were made for the composite wind speed and direction data. The composite was constructed from all of the available data for a particular hour. Thus, if data from one or more towers were missing for a particular hour, the composite was constructed from the data received from the other towers.

Wind Plant Forecast Sites

The number of wind plants for which forecasts were generated in this plant was determined by the number of California wind plants that were willing to permit the use of their data in the forecast production and evaluation process. Five California wind plants agreed to provide power production and meteorological data for use in the generation and evaluation of power production and wind speed forecasts for their plants. The participating plants included: (1) the Mountain View plant, a 66.6-MW facility in the San Geronio Pass of Southern California; (2) the Oak Creek plant, a 34.5-MW plant in the Tehachapi Pass, which is adjacent to the Mojave Desert; (3) the PowerWorks wind plant with a rated capacity of 90 MW in the Altamont Pass and located just to the east of the San Francisco Bay Area; (4) the 15.18-MW SMUD wind plant located in the Montezuma Hills in Solano County; and (5) the 162-MW High Winds Energy Center, which is adjacent to the SMUD plant in Solano County. Additional details about each wind plant are provided in the forecast performance section.

Performance Metrics

A variety of metrics can be used to assess the performance of forecasts of the wind power production and related meteorological variables such as wind speed, wind direction, and temperature. The basic performance metrics used in this study are similar to those used in the previous Energy Commission-EPRI project (Energy Commission and EPRI, 2003). These included the mean error (ME), the mean absolute error (MAE), and the skill scores with respect to climatology and persistence. The definition of each of these metrics and some insight as to what they actually measure is provided in the following paragraphs.

The mean error (ME) is defined as

$$ME = \left(\sum_{i=1}^{i=N} (F_i - O_i) \right) / N \quad (\text{Eqn 4-1})$$

where F_i is the forecasted quantity for a particular hour i , and O_i is the corresponding observed (measured) quantity for the same hour. The summation is over all N hours in the verification sample under consideration. Since positive errors (forecast too high) and negative errors (forecast too low) will offset each other in the summation, the ME represents the net error, which is often referred to as the bias.

A parameter that is widely used and provides more information about the typical level of error in a forecast is the mean absolute error (MAE). The definition of the MAE is

$$MAE = \left(\sum_{i=1}^{i=N} |F_i - O_i| \right) / N \quad (\text{Eqn 4-2})$$

where, as before, F_i is the forecasted quantity for a particular hour i , and O_i is the corresponding observed (measured) quantity for the same hour. The summation is over all

N hours in the verification sample under consideration. Since the absolute value of the difference between the forecasted and observed quantity is used in the summation, errors of opposite sign do not offset each other, and the MAE represents the average magnitude of the forecast error. For the power production forecasts, the MAE can be expressed in several ways. One approach is to express the MAE in standard power units such as MW or kW. However, these will vary substantially from plant to plant because of the variations in the size of the plant. Hence, to facilitate comparison in forecast performance among wind plants, the MAE is typically expressed as a percentage of the plant's installed capacity. This is sometimes referred to as the Mean Absolute Percentage Error (MAPE). Since the capacity factor (the ratio of actual production to the plant's installed capacity) varies substantially among plants and often from season to season at a particular plant, the MAE is sometimes expressed as a percentage of the actual production for the period that is being used to evaluate the forecast performance. Since capacity factors are often in the 30% to 40% range, the MAE as a percentage of production is typically about 2.5 to 3 times larger than the MAE as a percentage of capacity. However, capacity factors are sometimes significantly higher or lower than the 30% to 40% range, and the ratio of MAE-capacity to MAE-production may in some cases be substantially larger or smaller than the typical value of three.

Another metric that is widely used in forecast performance evaluations is the skill score. In general, the skill score measures the percentage improvement in a forecast performance metric relative to a reference forecast. The concept can be applied to any performance metric and reference forecast combination. In this investigation, the performance metric to be used in the skill score computation is the MAE, and the reference forecasts are persistence and climatology forecasts. This yields two difference skill scores. The skill score relative to a persistence forecast (S_p) is defined as:

$$S_p = 1 - \frac{MAE_{eWind}}{MAE_{persistence}} \quad (\text{Eqn 4-3})$$

The skill score relative to a climatology forecast (S_c) is defined as:

$$S_c = 1 - \frac{MAE_{eWind}}{MAE_{climatology}} \quad (\text{Eqn 4-4})$$

Persistence and climatology forecasts are often used as reference forecasts in skill score computations because they are simple forecasts that can be generated at lost cost, require very little knowledge, and represent basic forecast concepts – a forecast of no change (persistence) from the value at the time the forecast is made and a forecast of the long-term average of the quantity (climatology). The details of how the persistence and climatology forecasts were produced in this project are presented in the next section, which describes all of the forecast methods.

Forecast Methods

Forecasts were generated for each participating wind plant from an ensemble of forecast methods. The ensemble approach served two purposes. First, it provided a way to evaluate the performance of a number of different forecast methods over an extended period for several wind plants. Second, it provided an opportunity to evaluate the effect of constructing a composite forecast from a suite of individual forecast methods. This has proven to be a useful forecast tool in a number of other meteorological applications. The use of the ensemble forecast method was actually beyond the original scope of the project. It was originally envisioned that modifications would be made to the forecast system used by AWST in the previous Energy Commission-EPRI project (Energy Commission and EPRI, 2003) and that a single forecast made by a modified forecast system would be produced and evaluated in this task. However, it was felt that with a modest additional effort and no increase in the project budget, a number of different forecast methods could be evaluated, and an initial assessment of the ensemble approach could be made.

The ensemble of methods employed in this project included three different physics-based models, four different MOS approaches, and one explicit plant output model formulation. This produced a potential total of 12 different forecasts for each forecast cycle for each participating wind plant. However, not all methods were available for the entire forecast evaluation period. This was primarily due to limitation in obtaining or generating numerical forecast simulation data from the different physics-based models. Consequently, the comparison between some of the methods and construction of the ensemble forecasts had to be limited to a subsample of the one-year evaluation period.

Physics-Based Models

Three physics-based atmospheric models were used to produce the wind speed and energy forecasts for each of the wind plants. These models were the Mesoscale Atmospheric Simulation System (MASS), the Weather Research and Forecasting (WRF) model, and the Couple Ocean Atmosphere Mesoscale Prediction System (COAMPS) model. A brief description of each modeling system and how it was employed to generate the forecasts for this project is presented in the following subsections.

Mesoscale Atmospheric Simulation System (MASS)

The MASS model was developed in the 1980s as part of NASA's research activities in the development of new remote sensing systems (Kaplan et al, 1982). Version 5 of the MASS model was used to generate the *eWind* forecast simulations for the previous Energy Commission-EPRI forecast evaluation project (Energy Commission and EPRI, 2003). Version 6 was employed in this project.

The MASS-6 forecast simulations were generated on a nested grid system. An outer matrix of 100 by 80 grid cells with a cell size of 20 km was used to simulate the larger scale flow over the southwestern United States and the adjacent Pacific Ocean and Mexico. This is

referred to as the “A” grid. Figure 4-49 shows the geographical domain covered by the A grid. Two higher-resolution grids were nested inside of the A grid. Both high-resolution grids employed an 80 by 80 matrix of grid cells and a grid cell size of 5 km.

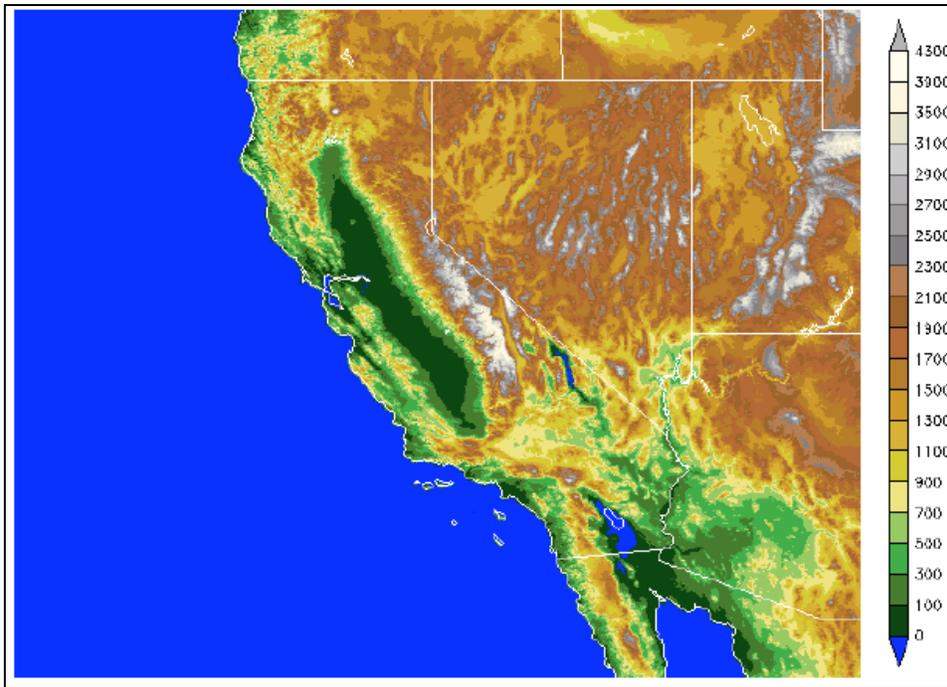


Figure 4-49 The geographical domain covered by the 100 X 80 matrix of 20-km grid cells used to produce the parent coarse grid MASS-6 simulations for the project.

One of the nested grids was centered over the Bay Area of Northern California and was used to generate forecast data for the PowerWorks, SMUD and High Winds wind plants. The other high resolution grid was centered over Southern California and was used to produce forecast data for the Oak Creek and Mountain View wind plants. Figures 4-50 and 4-51 present the two grids.

Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS)

The Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) was developed by the Marine Meteorology Division (MMD) of the Naval Research Laboratory (NRL) (Hodur, 1997). The atmospheric components of COAMPS are used operationally by the U.S. Navy for short-term numerical weather prediction for various regions around the world. Additional information about the COAMPS modeling system was presented earlier in this report and is also available at <http://www.nrlmry.navy.mil/coamps-web/web/home>.

The COAMPS simulations used to produce the power production and wind speed forecasts in this project were generated at the Lawrence Livermore National Laboratory (LLNL). These simulations were run twice per day using a 12-km and a nested 4-km grid. The areas covered by these grids are shown in Figure 4-52. The 4-km grid did not cover the area required to forecast for all five wind plants. Therefore, only the output from the 12-km grid was used to produce the forecasts for the evaluation phase of this task.

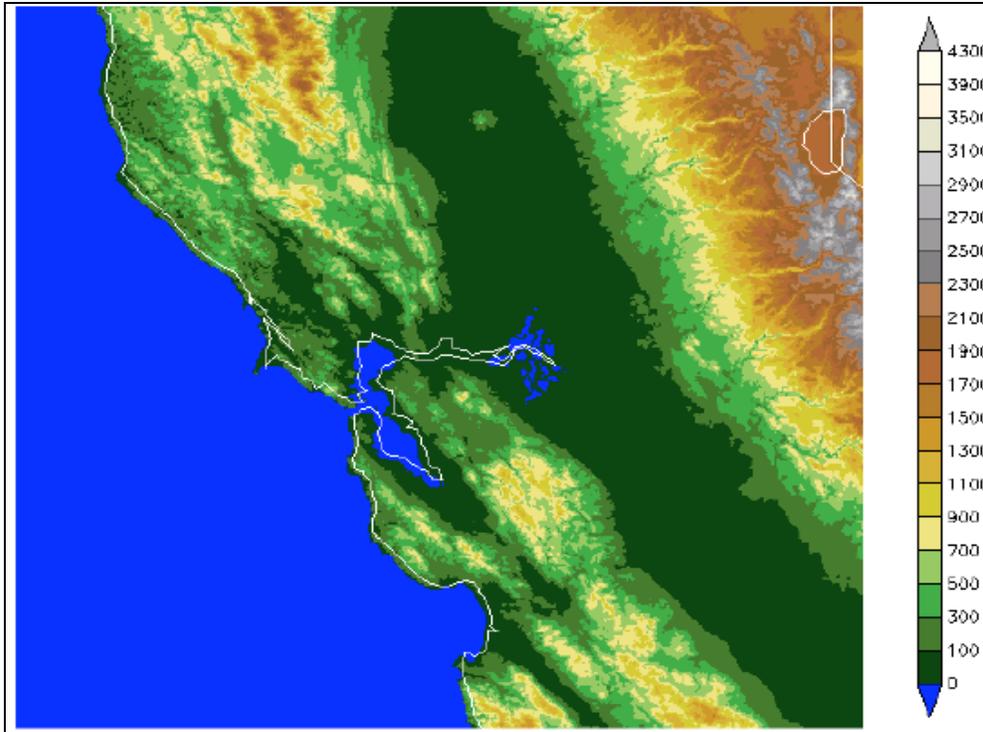


Figure 4-50 The geographical domain covered by the 80 X 80 matrix of 5-km grid cells used to produce the high resolution MASS-6 simulations over Northern California.

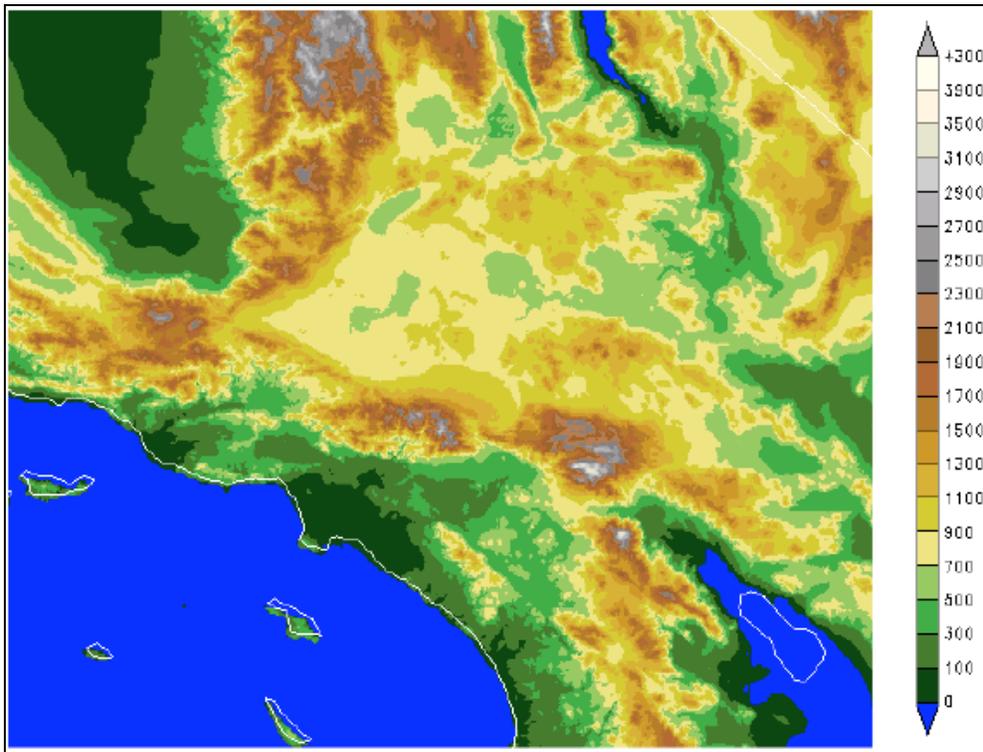


Figure 4-51 The geographical domain covered by the 80 X 80 matrix of 5-km grid cells used to produce the high resolution MASS-6 simulations over Southern California.

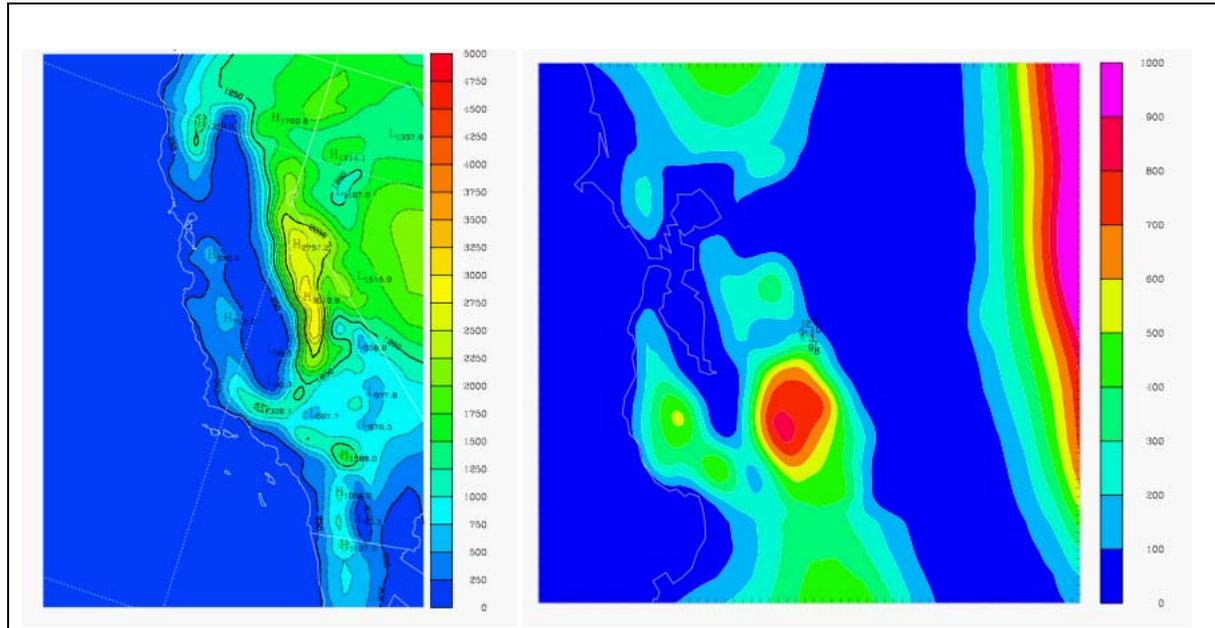


Figure 4-52 The geographical domain covered by COAMPS 12-km (left) and 4-km (right) grids used to produce forecast simulation for this project.

Weather Research and Forecasting Model (WRF)

The Weather Research and Forecasting (WRF) Model is a next-generation meso-scale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. It features multiple dynamical cores, a three-dimensional variational (3DVAR) data assimilation system, and a software architecture allowing for computational parallelism and system extensibility. WRF is suitable for a broad spectrum of applications across scales ranging from meters to thousands of kilometers. Additional information about the WRF model was presented earlier in this report and is also available on the WRF website, which can be found at <http://www.wrf-model.org/>. This site provides information on the WRF effort and its organization, references to projects and forecasting involving WRF, and links to the WRF users' page, real-time applications, and WRF-related events.

Real-time simulations for the California region were generated with Version 2.0.3.1 of the WRF model for this project. Unfortunately, the version of the WRF software available at the time the forecast systems were setup did not provide for one-way nested simulations such as those used in the MASS and COAMPS configurations. Therefore, the WRF simulations were restricted to a single-grid domain. The domain used for the WRF forecast simulations is depicted in Figure 3-5. The WRF grid was a 100 X 100 matrix of grid cells overlaid on this domain. The grid-cell size was 10 km, which yields a physical domain size of 1000 km by 1000 km. The model grid also included 25 vertical layers.

A 40-km grid point dataset from NCEP's Eta model (the Eta 212 dataset) was used for the specification of the initialization and lateral boundary conditions for the WRF simulations. This dataset is actually generated from a version of NCEP's Eta model that employs a 22 km grid for the model calculation, but NCEP interpolates the model to the 40-km grid for dissemination.

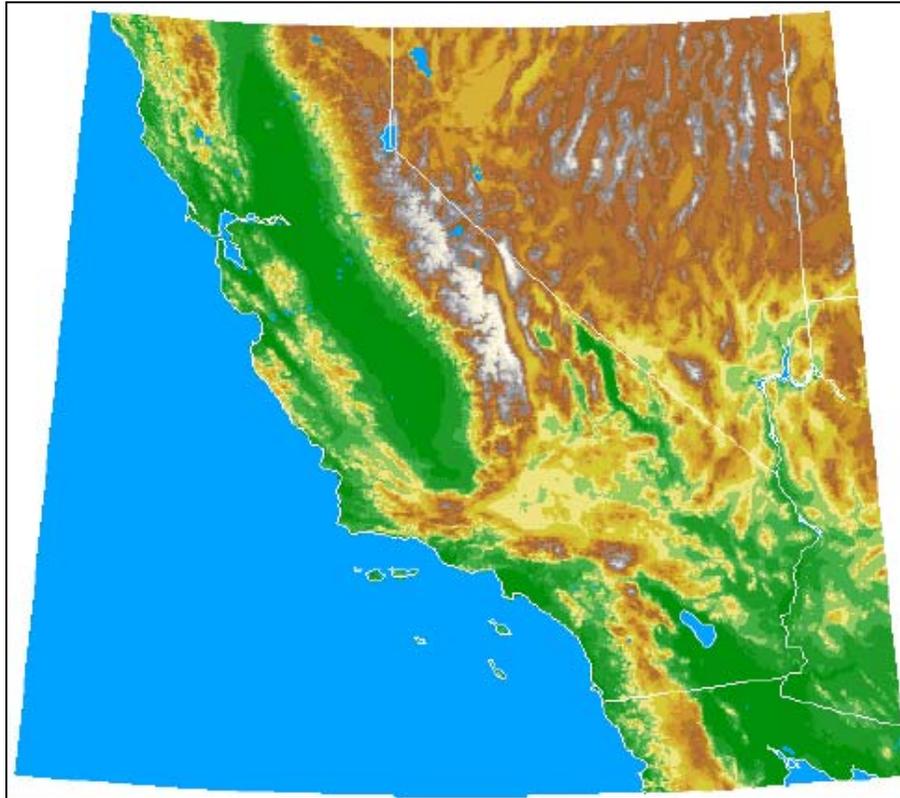


Figure 4-53 The geographical domain covered by the 100 X 100 matrix of 10-km grid cells used to produce the WRF simulations for the project.

Model Output Statistics Methods

The Model Output Statistics (MOS) component of a day-ahead wind power production forecast system serves to convert the grid-based predictions of a physics-based atmospheric model to point-specific (for example, at a wind plant) forecasts of the variables of interest for a particular application. In the process of doing this, a well-structured MOS procedure will correct systematic physics-based model errors that result from the omission of atmospheric features that are too small to be resolved on the physics-based model grid and deficiencies in the representation of physical processes within the physics-based model. It is possible to use a variety of statistical techniques and configurations of those techniques to serve as the MOS component of a wind power production forecast system. Four different approaches were used in this project. These are described in the following subsections.

Screening Multiple Linear Regression (SMLR)

In the previous project (Energy Commission and EPRI, 2003), AWST used a Screening Multiple Linear Regression (SMLR) procedure for the MOS component of the forecast system. In this approach, a large set of physics-based model variables is selected as a pool of candidate predictors. The SMLR algorithm then uses a training sample to select the candidate predictor with the most predictive power. This is defined as the candidate predictor with the highest R^2 value when evaluated over the entire training sample. The candidate predictor that is selected through this process becomes the first predictor of the ultimate statistical prediction equation. The selected predictor is then removed from the pool of candidate predictors, and the process is repeated. A second predictor is then selected based on the criteria of the maximum increase in R^2 . The process is repeated until it yields an increase in R^2 that is below a specified threshold. At that point it is assumed that no additional significant predictive information is available in the training sample and the process is terminated. The selected predictors and associated coefficients form a linear regression equation that serves as a MOS prediction equation.

Two-Stage Screening Multiple Linear Regression (SMLR2)

A second method employed in this project is a two-stage SMLR, which is assigned the label SMLR2. In the first stage, only physics-based model variables are used to create one prediction equation for all forecast hours. The first stage results in a wind speed forecast for all forecast hours. The second stage uses the wind speed prediction from Stage 1 and additional predictors such as the recent forecast errors to generate a separate forecast equation for each forecast look-ahead hour. Ultimately, the wind speed forecast from Stage 2 is converted to a power production forecast using a plant output model.

SMLR2 with Direct Power Production Prediction (SMLR2-DP)

A variant of the SMLR2 scheme was also employed as a separate forecast method. In this scheme (designated as SMLR-DP) the power production is directly predicted from the physics-based model variables instead first predicting the wind speed at the wind plant's anemometer site(s) and then using a plant output model to forecast the power production as in the SMLR and SMLR2 schemes. The reasoning is that some model variables, such as the turbulent kinetic energy, may provide useful information about variations in power production.

Stratified SMLR2 (SMLR2-ST)

A fourth method was a stratified SMLR2 scheme. The approach is conceptually similar to the SMLR2 scheme except that separate prediction equations are derived for subsets of the training sample. The content of the subsets are determined through the use of one or more stratification parameters. The selection of stratification parameters in this application of the scheme was subjective although in principle some type of cluster analysis could be used to define the subsets in an optimal manner. In this project, the physics-based model wind

speed for the wind plant site was selected as the stratification parameter. This parameter was used to divide the training sample into four subsets, each having approximately the same sample size. The SMLR2 procedure was then independently applied to each subset.

Plant Output Models

The plant output model is the link between the meteorological conditions at a wind plant and the plant's power production. There are a large number of ways in which a plant output model can be formulated. In fact, it is possible to totally avoid the use of an explicit plant output model by training the MOS procedure discussed above to directly forecast the power production from the output data of the physics-based model. This is the approach used by the SMLR2-DP method. The results obtained in the screening phase of this task indicate that this approach generally did not perform as well as the technique of first forecasting the wind speed at one or more anemometer locations within the wind plant and then using a plant output model to forecast the power production. This is generally consistent with AWT's overall experience, which indicates that it is usually better to use the MOS procedure to make a prediction of the relevant meteorological parameters at the wind plant site and then to use those predicted quantities as input into an explicit plant output model to produce the power production predictions.

The plant output model can be formulated as a statistical or a physical numerical model. A discussion of the merits of each approach was presented earlier in this report. The statistical approach was employed for the plant output model used for all forecast methods that utilized an explicit plant output model in the one-year forecast evaluation experiment. In the statistical approach power production and meteorological data from the plant are used to construct an empirical relationship between the plant-scale power production and the meteorological conditions. The simplest model of this type is commonly known as a "plant-scale power curve." This is simply an empirical equation derived from the plant's power production data and the wind speed data from one or more of the plant's anemometers. This type of curve is used in the forecast process to calculate the forecasted power production from meteorological predictions of the wind speed at the plant's anemometer locations. This is the approach that was used in the previous Energy Commission-EPRI forecasting project. The approach was extended in this project to include the statistical modeling of the deviations from the plant-scale power curve based on the wind direction and the time of day.

Reference Forecasts

Two types of reference forecasts were employed in this project: persistence and climatology. These reference forecasts represent simple forecast techniques that are based on simple forecast concepts such as "no change" or "long-term averages". These forecasts are used to calculate the skill scores and provide a reference point for the value of the more advanced and complex forecast methods.

Persistence Forecast

The persistence forecast is based on the concept that there will be no change in the forecasted quantity throughout the forecast period. In this project the persistence forecast was assumed to be value of the forecasted quantity for the hour ending at the forecast delivery time. Thus, the persistence power production forecasts for the morning forecast cycle (delivered at 8 AM) used the reported power production for the 7:00 a.m. to 8:00 a.m. hour as the forecast for all hours in the 48-hour forecast period. Similarly, the persistence forecast for the evening forecast cycle (delivered at 8:00 p.m.) employs the reported production for the 7:00 p.m. to 8:00 p.m. hour as the forecast for each of the 48 hours in the forecast period.

Climatology Forecast

The climatology forecast is based on the concept that the forecasted quantity will revert its long-term average value for the entire forecast period. The long-term average can be defined in many ways. For the purposes of this project, the long-term average consisted of a separate estimate of the long-term average for each hour of the day for each month of the year from the available data. Ideally, climatological averages should be constructed from many years of data, but this is not typically available for most wind plants. Multi-year monthly diurnal climatological averages of the power production and wind speed were available for the Mountain View and PowerWorks wind plants, and these were the basis for the climatological forecasts for these wind plants. Unfortunately, multi-year climatologies were not available for the Oak Creek, SMUD, and High Winds plants. For these plants, a pseudo-climatology was constructed by computing the monthly average value of the power production and wind speed for each hour of the day from the available measured data for each month. This pseudo-climatology forecast had a big advantage over a typical climatology forecast since the average was from the actual month for which the forecasts were made. The forecast errors associated with this climatological forecast represented only the deviations from the average diurnal values for the month. This is a substantial advantage, and one would expect that the skill score relative to this type of climatological forecast would be substantially lower than that relative to a true climatological forecast.

Ensemble Forecasts

Forecast ensemble models are statistical models that produce a forecast by compositing forecasts from a number of different forecast techniques (Sivillo, et al, 1997). The availability of forecast based on three physics-based models and four different MOS procedures provides an opportunity to at least obtain an initial estimate of the effect of employing an ensemble forecasting approach to the wind power production forecast application. To address this objective, an ensemble-mean forecast was generated by constructing an average of all of the 12 individual forecasts available for the final four months of the forecast evaluation period.

Forecast Performance Evaluation Results by Wind Project

This section presents the performance results for each of the five wind plants for the one-year forecast evaluation period extending from July 2004 to June 2005. The forecast performance results are presented by wind plant beginning with the southernmost plant, Mountain View in San Geronio Pass and proceeding northward to the High Winds and SMUD wind plants in Solano County.

As noted previously, 48-hour forecasts of the hourly power production and wind speed were generated on a twice-daily cycle with scheduled delivery times of 8:00 a.m. and 8:00 p.m. PST each day. To construct a more detailed evaluation of the forecast performance, a comprehensive analysis was done only for the forecasts generated during the morning forecast cycle (the forecasts scheduled for delivery at 8:00 a.m. each day). A review of the performance statistics indicates that the differences in forecast performance between the morning and afternoon cycles were not significant and that all of the significant conclusions from the analysis of the performance of the morning forecast cycle would apply to the combined pool of the afternoon and morning forecast cycles. Furthermore, it is the morning forecast cycle that is typically most important to forecast users since the required scheduling of output for the next day typically occurs early in the day and utilizes a forecast for the next day that is available early on the day that the schedules are to be submitted. As a result, only the results from the morning forecast cycle are presented in this report.

There were a total of 17,520 forecast hours in the forecast evaluation pool for the morning forecast cycle, which represents the product of the 365 days in the forecast period and 48 forecast hours per forecast cycle. Unfortunately, the actual number of hours in the verification pool varied significantly by wind plant and type of forecast and in many cases was substantially less than the maximum possible number of forecast hours. This reduction in the size of the verification pool was due to several factors. The most prominent factor was unavailable (or missing) observational data from the wind plants. The amount of unavailable data varied substantially among the participating wind plants and is documented in the sections that present the forecast performance results for each plant.

Another factor that limited the size of the verification pool in some cases was the availability of output data from the physics-based models. The MASS-6 forecast simulations, which were generated by AWST, were available for the full 48-hour forecast period for every day of the forecast evaluation period. The WRF model forecast simulations, which were also produced by AWST, were available only for the four-month period from March to June 2005. The COAMPS model forecast simulations, which were generated by Lawrence Livermore National Laboratory, were available for all 12 months of the forecast evaluation period, but the simulations did not extend far enough in time to cover the full 48-hour forecast period. In fact, the COAMPS data covered only the first 32 hours of the forecast period. Thus, all of the COAMPS forecast performance results presented in this report are for a 32-hour forecast period, whereas those for the MASS-6 and WRF models are for a 48-hour period.

Mountain View 1 and 2 Wind Project, San Gorgonio Pass

The Mountain View wind plant is located on the eastern side of the San Gorgonio Pass of Southern California. SeaWest, Inc. installed it in 2001. The project consists of 111 turbines manufactured by Mitsubishi Heavy Industries. Each turbine is rated at 600 kW, which yields a rated capacity of 66.6 MW for the wind plant. The project is spatially divided into two separate groups of turbines, one group with a rated capacity of 44.4 MW (Mountain View I) and a second group with a rated capacity of 22.2 MW (Mountain View II). The turbines are mounted on 50-meter and 60-meter tubular steel towers.

The power production and meteorological data for this wind plant were obtained through the CA ISO PIRP data communication system. The power production data consisted of the energy output of the Mountain View I and Mountain View II turbine clusters at 10-minute intervals. Data were available from three meteorological towers. The Catellus tower is the westernmost of the three towers and since the prevailing direction of the strongest winds at this site is from the west, this tower provides the best representation of the winds entering the domain of the wind plant. Wind speed and direction forecasts were made for all three meteorological towers, but the forecast performance analysis were done only for the Catellus tower.

Figures 4-54 and 4-55 show the number of hours available to the forecast verification pool for each month for the power production and Catellus wind speed forecasts. The available hours are the hours for which both forecast and observational data are available, which enables forecast errors to be calculated. The forecast data used to compile this chart were from the MASS-6 forecasts, which were available for every possible forecast hour for the entire year. Therefore, the unavailable hours shown in these two charts are solely attributable to missing observational data. There were a total of 17,520 possible forecast hours (365 days times 48 forecast hours for each day's forecast). A quick glance at these two charts indicates that there was a substantial amount of unavailable observational data during the one-year forecast evaluation period. In fact, power production data were not available for 26.8% of the hours, and the Catellus anemometer data were not available for 17% of the hours during the one-year period. Therefore, the MASS-6 forecast verification statistics for the Mountain View wind plant shown are based on 12,820 hours of data for the power production forecasts and 14,544 hours for the wind speed forecasts. The size of the verification pool was smaller for the forecasts based on the COAMPS and WRF models since output from those models was not available for all of the possible forecast hours.

Although power production data were not available for 26.8% of the hours during the one-year period, the distribution of missing data hours was quite uneven. August 2004 (62% unavailable) and November 2004 (46.3% unavailable) had particularly large fractions of missing power production data. On the other hand, July 2004, February 2005, and June 2005 had missing data for less than 10% of the forecast hours.

There was also considerable variability in the availability of the wind speed data. The worst month was August 2004, when data was unavailable for 73.9% of the forecast hours.

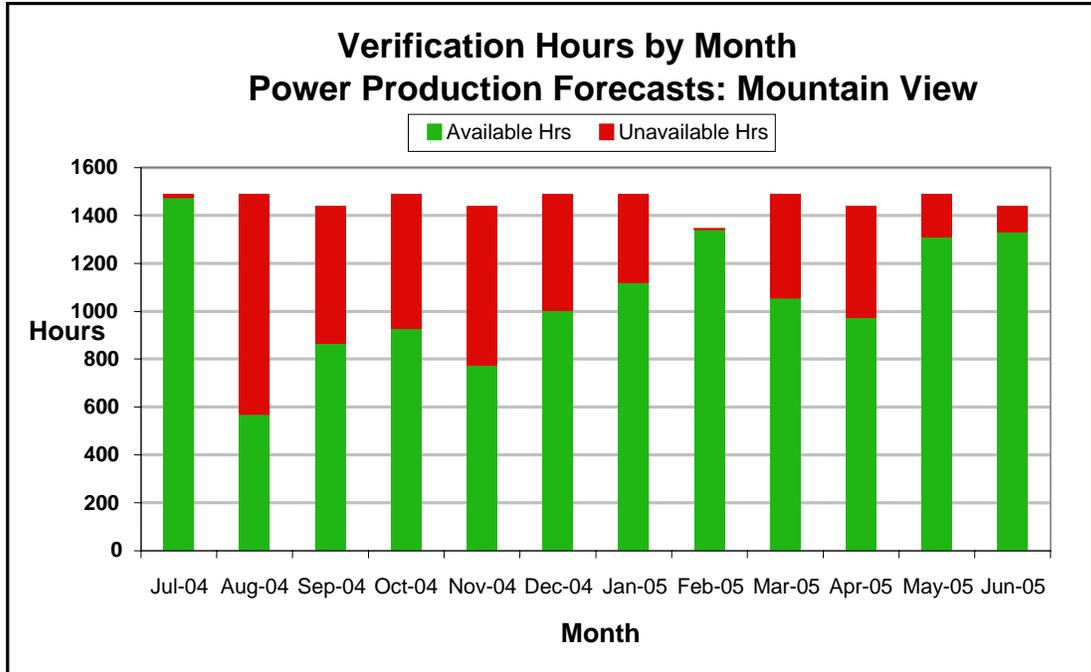


Figure 4-54 Number of hours for which sufficient data were available (green) and unavailable (red) to compute forecast performance statistics for the Mountain View power production forecasts.

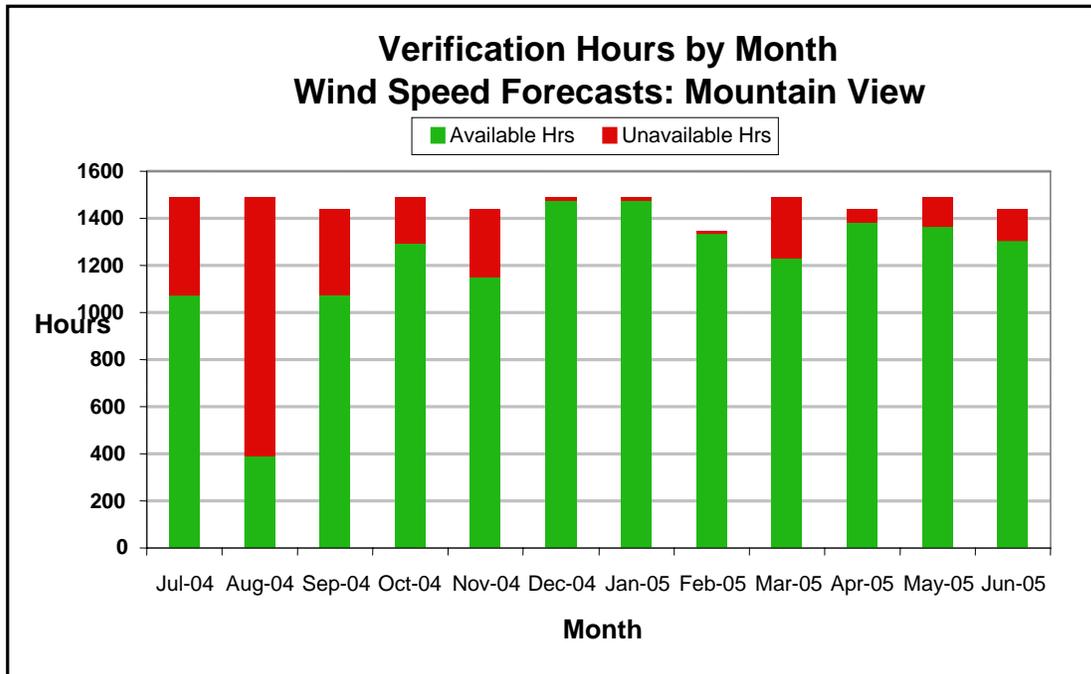


Figure 4-55 Number of hours for which sufficient data were available (green) and unavailable (red) to compute forecast performance statistics for the wind speed forecasts for the Catellus Meteorological Tower at the Mountain View wind plant.

Overall Forecast Performance

Table 4-6 presents the forecast performance statistics by month for the best performing MASS-6 forecast MOS method, which was the SMLR2-ST method for the Mountain View plant. The MAE of the power production forecasts for all verifiable forecast hours during the one-year period (12,728 hours) was 13.0% of the installed capacity. This is approximately 3.6 percentage points lower than the annual MAE achieved by the *eWind* forecast system for the one-year period from October 2001 to September 2002 in the previous Energy Commission-EPRI project (See Table 5-2 in Energy Commission and EPRI, 2003 for statistics comparable to those in Table 4-6). This represents an MAE decrease of about 21.7% from the previous project.

A substantial portion of this decrease is most likely attributable to the improvements made to the forecast system, but some of it may be attributable to differences in wind regime characteristics between the 2001-02 evaluation period and the 2004-05 period. The monthly MAE values for the 2004-05 period ranged from 3.6% for December to a high of 17.4% for March.

The annual skill score relative to persistence was 40.6%. This was somewhat higher than the persistence-based skill score of 37.5% during the 2001 to 2002 period. The monthly skill scores ranged from -1.0% for December 2004 to a high of 56.1% for May 2005. The annual skill score with respect to climatology was 47.9%, which was substantially higher than the climatology-based skill score of 36.4% achieved in the previous project.

The annual MAE of the wind speed forecasts was 2.65 m/s, which is 27.5% of the mean annual wind speed. The monthly MAE values ranged from a low of 2.21 m/s for February 2005 to a high of 3.15 m/s for September 2004. The annual skill score with respect to persistence for the wind speed forecasts was 40.1%. The lowest monthly skill score vs. persistence was the 14.1% for December 2004, while the highest was 55.3% for April 2005. The annual skill score with respect to climatology was 32.0% and ranged from 14.1% for December 2004 to 48.6% for April 2005.

Table 4-6 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-ST) for the Mountain View wind plant

Month	Power Production Forecast				Wind Speed Forecast - Catellus			
	% MAE(1)	%MAE (2)	Skill-P	Skill-C	MAE (m/s)	% MAE (3)	Skill-P	Skill-C
Jul-04	14.9%	27.4%	50.0%	28.6%	2.26	19.1%	50.2%	28.5%
Aug-04	16.0%	36.4%	36.5%	22.6%	2.52	21.0%	45.5%	36.4%
Sep-04	16.9%	54.9%	25.4%	50.8%	3.15	35.9%	29.1%	29.3%
Oct-04	14.4%	61.9%	40.1%	50.6%	3.07	38.8%	31.1%	23.0%
Nov-04	11.4%	78.1%	38.8%	52.5%	2.80	40.6%	37.0%	22.1%
Dec-04	3.6%	159.6%	-1.0%	78.5%	2.67	48.1%	14.2%	14.1%
Jan-05	6.2%	85.8%	23.6%	69.0%	2.51	44.8%	34.4%	26.9%
Feb-05	10.1%	62.5%	29.8%	64.6%	2.21	36.1%	31.6%	42.9%
Mar-05	17.4%	40.6%	30.7%	47.3%	3.09	30.6%	39.1%	41.4%
Apr-05	17.1%	29.2%	52.9%	46.7%	2.80	22.8%	55.3%	48.6%
May-05	13.6%	20.5%	56.1%	40.9%	2.49	18.4%	52.2%	24.9%
Jun-05	15.7%	23.2%	27.6%	18.1%	2.27	14.9%	48.4%	32.9%
Annual	13.0%	36.5%	40.6%	47.9%	2.65	27.5%	40.1%	32.0%

(1) MAE as % of rated capacity (2) MAE as a % of production; (3) MAE as % of average wind speed

Figure 4-56 shows the error distribution for the one year of power production forecasts. Separate error distributions for the first and second 24 hours are shown on this chart. The error distributions for the two periods are quite similar. Both are characterized by a strong central peak in the -5% to +5% range with a rapid decrease in the frequency on either side of this peak and a long tail of very low frequencies for medium and large negative and positive errors. In general, the positive and negative sides of the distribution are fairly symmetric with positive errors of a particular magnitude occurring about as frequently as negative errors of the same magnitude. There is a substantial asymmetry in the +5% to -5% range of errors. The frequency of errors in the -5% to 0% bin is much greater than that in the 0 to +5% bin. This is because errors of exactly zero are included in the -5% to 0% bin. There are a large number of hours during the winter months when the forecasted and observed wind speeds are both below the turbine start-up threshold and the forecasted and observed power production are both zero. This results in many hours of exactly zero error and accounts for this asymmetry.

Figure 4-57 shows the cumulative error distribution for the power production forecasts. The vertical axis represents the percentage of forecasts where the absolute error is less than or equal to the value on the horizontal axis. This chart can be used to estimate an overall confidence level for the forecasts. For example, 90% of the forecasts have an absolute error that is less than 40% of capacity. Once again, separate error distributions are shown for the first and second 24 hours of the forecast period, and the distributions for these two periods are very similar.

Previous studies have indicated that there is a strong relationship between the forecasted level of power production and the magnitude of the forecasted error. This is predominantly a result of the shape of the turbine power curve. The sensitivity of power production error to the error in the wind speed forecasts is greatest in the middle portion of the power curve

and least at the top and bottom of the curve. As a result, one would expect that even if the wind speed forecast error were the same for all wind speeds, the power production forecast error would peak for wind speeds and power production values that are in the middle of the power curve. The chart in Figure 4-58 indicates that this is indeed the case for the Mountain View plant. This chart presents the MAE for 21 bins of forecasted production from 0% to 100%. Each represents a 5% interval of forecasted production. The MAE is near 25% of installed capacity when the forecasted production is in the 35% to 65% range but falls to near or under 10% when the forecasted production is near zero or the plant capacity.

Figure 4-59 shows that the magnitude of the wind speed forecast error has much less dependence on the forecasted wind speed than the power production error has on forecasted power production. This chart shows the MAE for 20 bins of width one m/s for forecasted wind speeds ranging from zero to 20 m/s. The bin labeled “0” represents wind speed forecasts of exactly zero. All other bins are labeled by the wind speed value at the upper end of the bin. The magnitude of the error is generally lower for forecasts of low wind speeds. The MAE is under 2 m/s for wind speeds under 3 m/s. The MAE values rise as the forecasted wind speeds increase and are near 3 m/s for forecasted wind speeds in the 6 to 13 m/s range. The MAE is generally lower above the 13 m/s mark. There are some rather large MAE values in the 18 m/s and 19 m/s bins, but the sample size for these bins is very small. The fact that the magnitude of the wind speed errors peak in the 6 to 13 m/s range greatly exacerbates the power production forecast errors because this is the portion of the power curve that has the greatest sensitivity to wind speed forecast error.

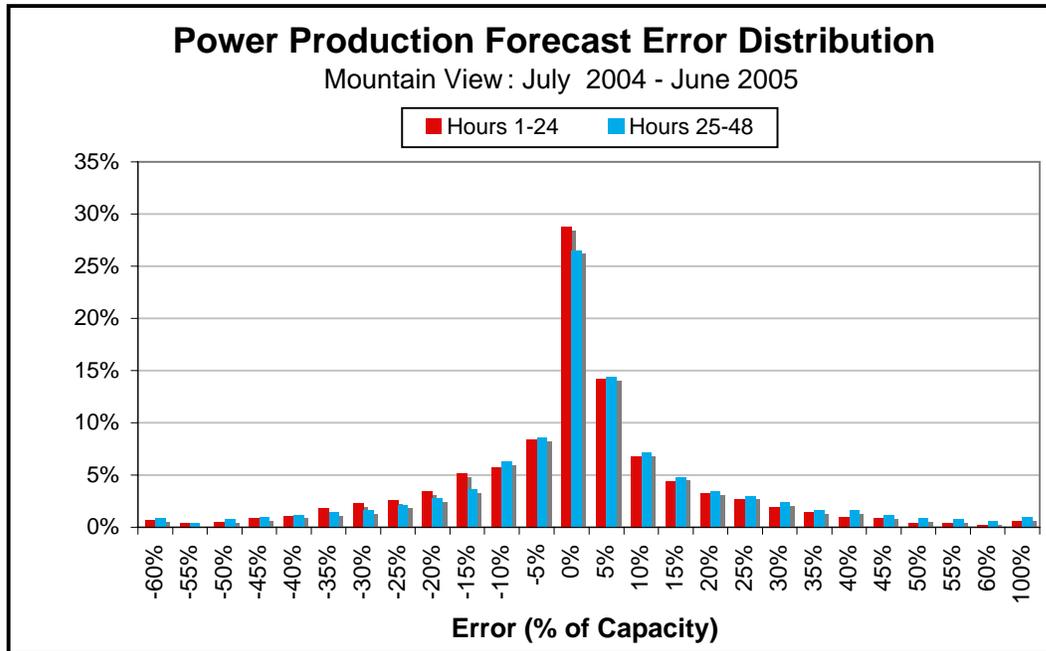


Figure 4-56 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the forecasts for the Mountain View wind plant. The data sample includes a total of 12,820 forecast hours during the July 2004 to June 2005 period and are included in the data sample.

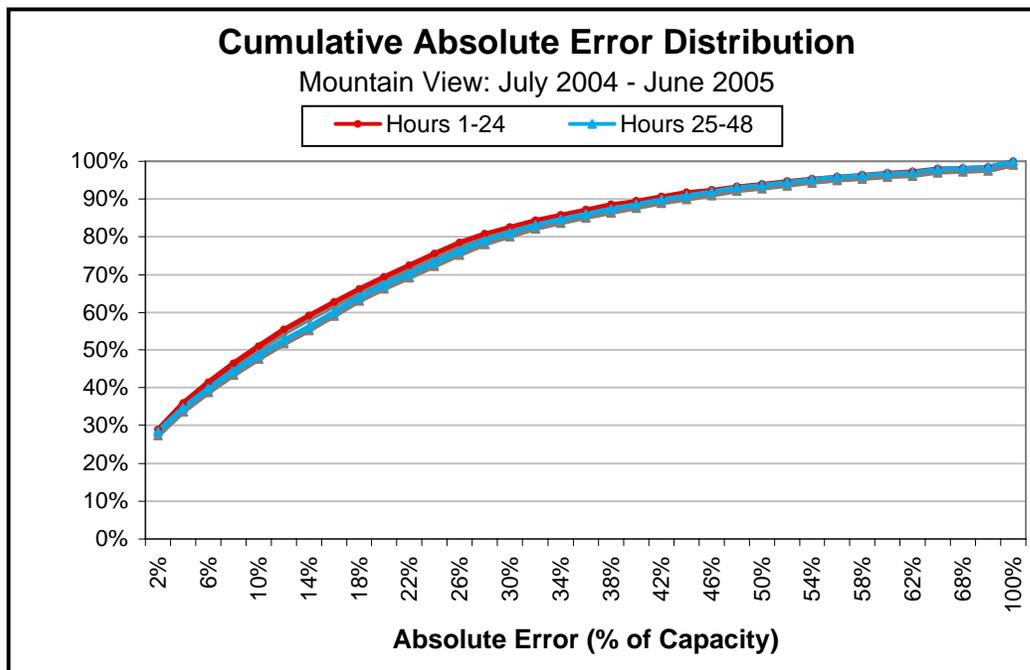


Figure 4-57 Cumulative absolute error frequency distribution (% of hours with an absolute error less than or equal to the value on the horizontal axis) for hours 1 through 48 of the forecasts for the Mountain View wind plant. The data sample includes a total of 12,820 forecast hours are included in the data sample.

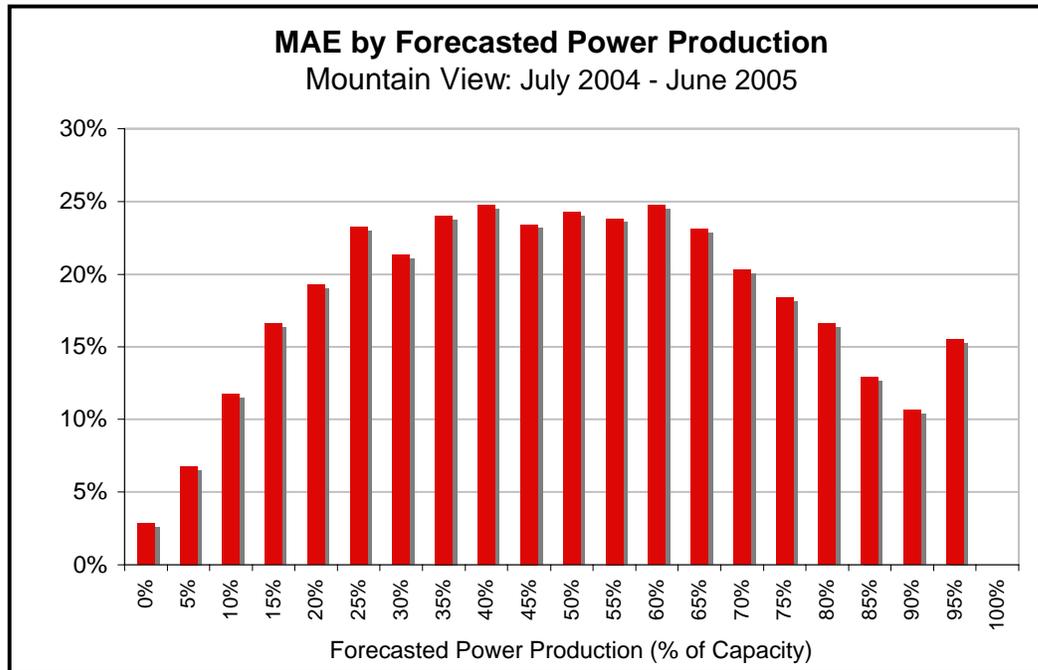


Figure 4-58 Mean Absolute Error (MAE) of the Mountain View power production forecasts vs. the forecasted power production. The MAE is shown for 21 bins of forecasted production. Each has a width of 5% of installed capacity and is labeled by the value at the upper end of the bin.

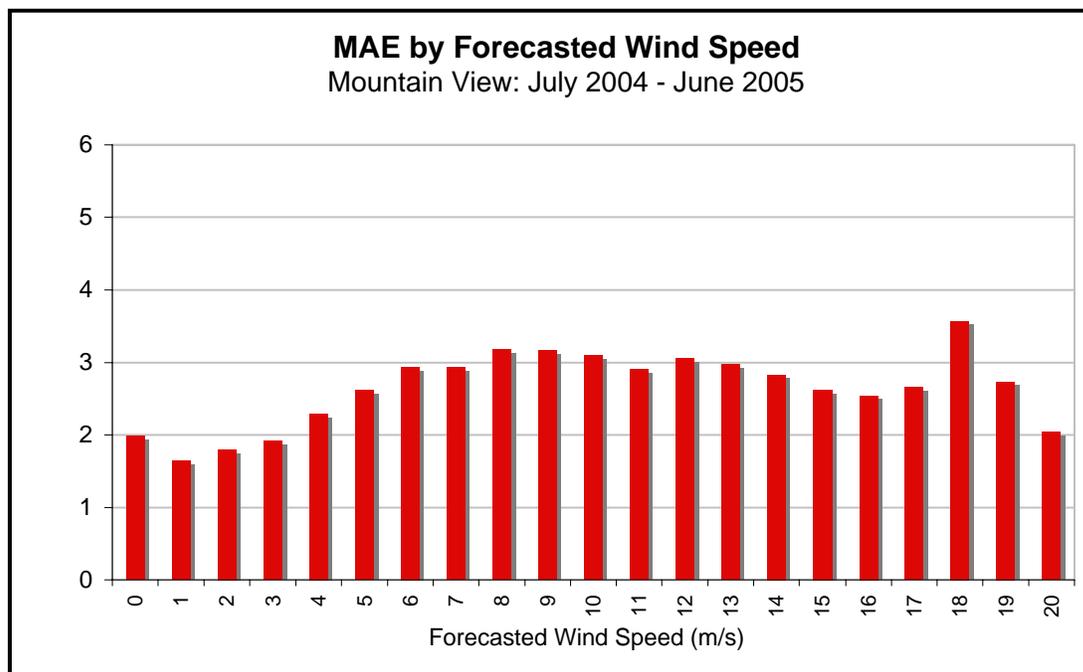


Figure 4-59 Mean Absolute Error (MAE) of the Catellus wind speed forecasts vs. the forecasted wind speed. The MAE is shown for 21 bins of forecasted wind speed. Each has a width of 1 m/s and is labeled by the value at the upper end of the bin.

Forecast Performance vs. Time Horizon

An important consideration for using wind power production forecasts is the rate of degradation of the forecast performance as the forecast time horizon (look-ahead period) increases. If the performance of the forecasts decreases rapidly with increasing look-ahead period, then it is useful and perhaps critical to have frequent forecast updates. On the other hand, if the degradation is slow, then the need for frequent updates is reduced.

This issue was addressed in the analysis of the forecast performance by constructing an MAE vs. forecast time horizon chart. Figures 4-60 and 4-61 show the charts for the Mountain View power production and the Catellus wind speed forecasts. The charts show the MAE for each look-ahead hour for the MASS-6 SMLR2-ST forecasts (labeled *eWind*) and the persistence and climatology reference forecasts.

The MAE of the power production forecasts rises in a fairly linear fashion from 11% to 12% during the first few hours of the forecast period to 14% to 15% by the end of the forecast period. The MAE of the persistence forecasts rises rapidly from about 5% for the first hour to near 25% by the 12th hour of the forecast period. The MAE of the persistence forecasts then actually decreases to a minimum of about 20% near the 24-hour mark. This minimum is a reflection of the fact that there is a pronounced fairly regular diurnal cycle in the Mountain View power production, especially during the warm season. The persistence forecasts outperform the *eWind* forecasts for the first couple of hours of the forecast period. The *eWind* forecasts begin to outperform the persistence forecasts at forecast hour three and outperform persistence by a wide margin after that time. As noted previously, the fact that persistence outperforms the *eWind* forecasts for the first couple of hours is the result of not using the most recent data from the wind plant and its vicinity in the forecast process. The MAE of the climatology forecasts exhibits a significant diurnal pattern ranging from a high near 27% at forecast hours 6 and 30 (3:00 p.m. PST) to a minimum of about 22% near forecast hours 12 and 36 (9:00 p.m. PST).

A linear fit to the *eWind* MAEs indicates that the average rate of error growth over the 48-hour period is 0.06% of installed capacity per hour or about 1.4% of installed capacity per day. This is a growth rate of 11.1% per day relative the overall MAE of 13.0% for the 48-hr period.

The MAE of the Catellus wind speed forecasts increases from near 2.6 m/s during the first 12 hours of the forecast period to about 3.0 m/s at the end of the 48-hour period. The MAE of the persistence forecasts rises rapidly from about 1.5 m/s for the first hour to more than four m/s by forecast hour 12. It decreases slightly in the few hours after the 12-hour mark but then slowly increases for the remainder of the forecast period, reaching 5 to 6 m/s during the last 12 hours of the forecast period. It is interesting to note that there is not a well-defined MAE minimum near the 24-hour mark as there is in the persistence MAE pattern for the power production forecasts. The MAE of the climatology forecasts remains near 4 m/s throughout the forecast period. A linear least squares fit of the *eWind* hourly MAE data indicates that the average rate of increase in MAE is 0.0074 m/s per hour or about 0.18 m/s per day.

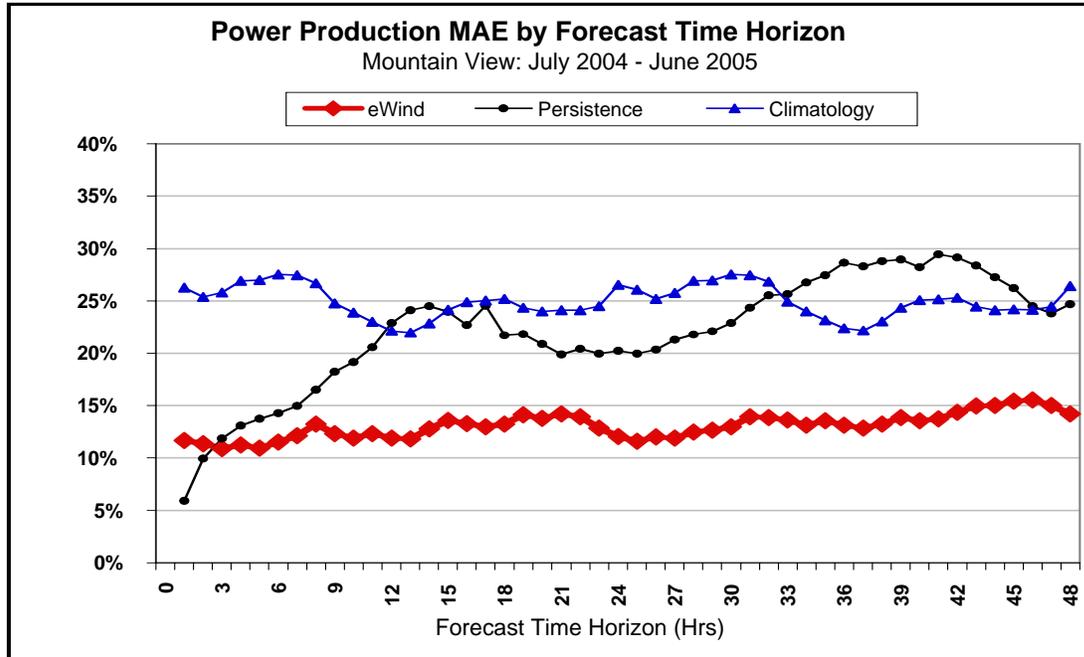


Figure 4-60 MAEs of power production forecasts vs. forecast time horizon for 12 months of *eWind* (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers) forecasts for the Mountain View wind plant.

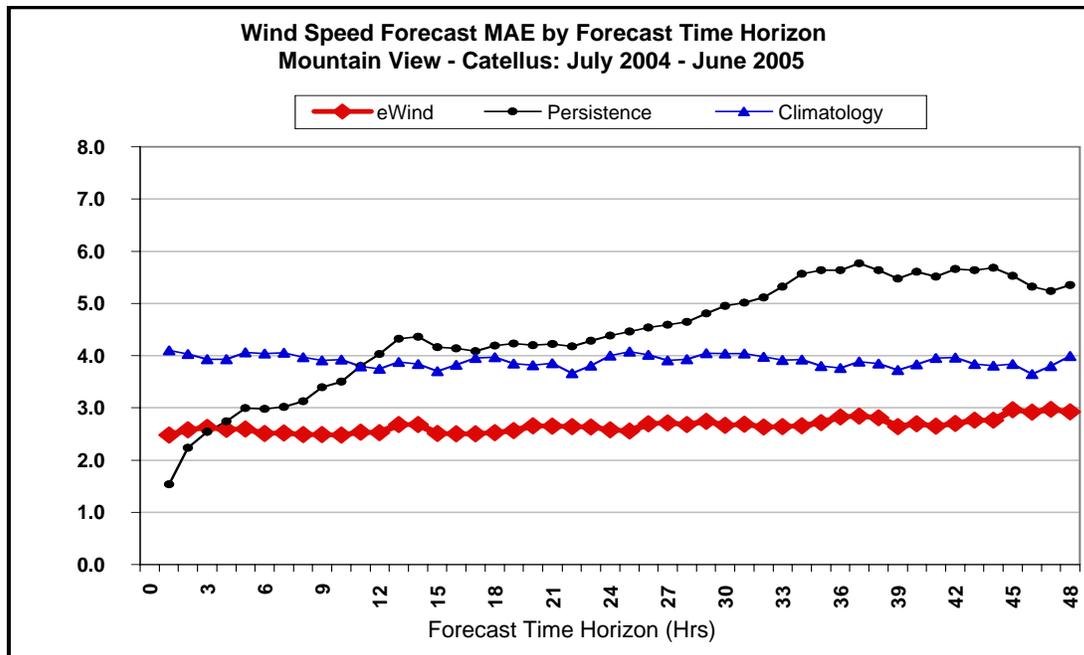


Figure 4-61 MAEs of power production forecasts vs. by forecast time horizon for 12 months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) wind speed forecasts for the Mountain View wind plant.

This is a growth rate of 6.7% per day relative to the overall MAE for the entire 48-hr period of 2.65 m/s. This is little bit more than half of the percentage growth rate in power production forecast error. The lower percentage rate for the wind speed forecast error is probably a reflection of the fact that wind speed forecast errors are magnified when the wind speeds are in the steeply sloped portion of the power curve.

Physics-Based Model Comparison

The generation of forecasts from three different physics-based models in this project provided an opportunity to determine the sensitivity of the forecast performance to the choice of physics-based model and also to assess the relative performance of the three different models. The performance comparison analyzed the forecast error statistics for forecasts generated using the same MOS method from all three models. The selected method was the SMLR2-ST scheme.

Table 4-7 presents the performance statistics for the SMLR2-ST forecasts for all three models. The comparison of the monthly forecast performance statistics is made more difficult by the fact that the forecasts from the three models were not generated for the entire one-year forecast period and did not have the same forecast length. As noted previously, the MASS-6 forecasts were produced for all 12 months of the evaluation period and for all 48 hours. The COAMPS forecasts were also produced for all 12 months but these only extended to 32 hours beyond the forecast delivery time. The WRF forecasts were only available for 4 of the 12 months but did extend throughout the entire 48-hour forecast period.

Table 4-7 Monthly MAE as a percentage of rated capacity for the 1- to 48-hour Mountain View power production and wind speed forecasts produced from the output of the three physics-based models used in this project.

Month	Power Production Forecast MAE			Wind Speed Forecast MAE		
	MASS 6	COAMPS	WRF	MASS 6	COAMPS	WRF
Jul-04	14.9%	19.7%		2.26	2.96	
Aug-04	16.0%	16.8%		2.52	2.53	
Sep-04	16.9%	18.0%		3.15	3.26	
Oct-04	14.4%	19.3%		3.07	3.94	
Nov-04	11.4%	12.2%		2.80	3.19	
Dec-04	3.6%	4.0%		2.67	2.56	
Jan-05	6.2%	4.9%		2.51	2.23	
Feb-05	10.1%	12.7%		2.21	2.59	
Mar-05	17.4%	22.4%	22.7%	3.09	3.45	3.86
Apr-05	17.1%	17.7%	16.5%	2.80	3.11	2.82
May-05	13.6%	17.0%	16.3%	2.49	2.69	2.61
Jun-05	15.7%	16.8%	18.1%	2.27	2.45	2.62
Annual	13.0%	15.0%	18.4%	2.65	2.92	2.98
Mar - Jun	16.0%	18.5%	18.4%	2.66	2.92	2.98

Since the COAMPS statistics do not include the last 16 hours of the forecast period, when errors are expected to be slightly higher, the calculated COAMPS errors would be expected to be lower than those for the other two models if all models performed similarly.

The performance statistics indicate that the forecasts based upon the MASS-6 model significantly outperformed the other two models. Over the entire 12-month period, the MAE of the MASS-6 power production forecasts was 2 percentage points (13.3%) lower than the forecasts from the COAMPS output data. In addition the MASS-6 forecasts had a lower MAE in 11 of the 12 months. Similarly, the MAE of the MASS-6 wind speed forecasts was 0.27 m/s (9.2%) lower than the COAMPS forecasts, and the MAE was lower in 10 of the 12 months.

A comparison with the WRF-based forecasts was, of course, only possible for the four months of March through June 2005. The MAE of the MASS-6 power production forecasts was 2.4 percentage points lower for those four months (13.0%). In addition, the MASS-6 forecasts yielded a lower MAE during three of the four months. The exception was April when the WRF-based forecasts yielded a lower MAE by 0.6 % of capacity.

The relative performance was similar for the wind speed forecasts. The MAE of the forecasts based on the MASS-6 output was 0.32 m/s (10.7%) lower. The MASS 6-based forecasts yielded a lower MAE during all four months, although the difference for April 2002 was only 0.02 m/s.

Interestingly, the difference in the MAE of the raw wind speed forecasts was even greater. The raw forecasts are based on the wind speeds that are directly interpolated from the physics-based model's computational grid. The MAE of the raw MASS-6 wind speed forecasts for the Catellus anemometer site for the entire four-month comparison period was 3.16 m/s while the MAE of the raw WRF forecasts for the same period was 3.60 m/s. This represents a 12.2% lower MAE for the raw forecasts from the MASS-6 simulations. Thus, the MOS procedure improved the WRF-based forecasts more than it improved the MASS-6 forecasts.

The performance of the forecasts based on the COAMPS and WRF models was much more similar. For the four months for which a comparison was possible, the MAE of the COAMPS and WRF forecasts was essentially the same. The WRF-based forecasts had an MAE of 18.4% while the COAMPS-based forecasts recorded a slightly higher MAE of 18.5%. On a monthly basis, each model yielded a lower MAE during two of the months. The MAE of the wind speed forecasts was also quite similar, with the COAMPS model having a slightly lower MAE of 2.92 m/s compared to 2.98 m/s for the WRF-based forecasts.

The performance statistics clearly indicate that the forecasts based on the MASS-6 model had a distinct advantage for the Mountain View site while the forecasts based on the WRF and COAMPS models performed similarly. The reason for this is not readily apparent. These results are somewhat consistent with the results from the earlier Energy Commission-EPRI project in which the power production and wind speed forecasts for the Mountain

View plant based on forecast simulations from the MASS model significantly outperformed those from Risoe's Prediktor forecast system. In that project, the Prediktor system employed the U.S. National Weather Service's Eta model as the physics-based model input for its forecasts. In the context of this project, the superior performance of the MASS-6 model for this wind plant may be related to the fact that the MASS-6 model was run at a higher resolution than the other two models. However, in experiments done in the screening phase of this task, it was shown that the use of higher physics-based model resolution beyond about 10 km did not yield a substantial improvement in the MAE of the wind speed or power production forecasts after an MOS scheme was applied. Another possibility is that the MASS-6 model system and associated databases have been modified to better account for the geophysical properties of the Earth's surface near San Geronio Pass, including an improved representation of the vegetation and soil moisture associated with the extensive irrigation in the Coachella Valley east of San Geronio Pass. Experiments conducted by AWST have indicated that a detailed representation of the effect of irrigation on the geophysical surface properties in a physics-based model is an important factor in the accurate simulation of the winds in San Geronio Pass. Unfortunately, it is difficult to reach a definitive conclusion on this issue without conducting a set of additional experiments.

Forecast Ensembles

An ensemble mean forecast of the power production and the wind speed was generated from the 12 forecasts produced for the Mountain View wind plant for the four months (March to June 2005) for which all 12 forecasts were available. The 12 ensemble members consisted of each combination of four MOS methods and the three physics-based models.

Table 4-8 presents the MAEs for the power production and wind speed forecasts for several combinations of the 12 ensemble members as well as for the ensemble-mean forecast. For the power production forecasts, the MAE of the ensemble-mean forecast was 0.4% of rated capacity higher than the MAE of the best performing individual forecast (MASS-6 SMLR2-ST) for the entire four-month period. It was also 0.7% of rated capacity higher than the MAE of a composite forecast constructed by using the best performing forecast for each month. However, the MAE was 1.5% of capacity lower than the average MAE for all 12 of the individual methods.

The MAE of the ensemble-mean wind speed forecasts for the four-month period was 0.04 m/s lower than the MAE of the best performing individual forecast method. This represents an MAE reduction of 1.5% relative to the best individual forecast. In addition, the MAE of the ensemble mean forecast was lower than that of the best overall forecast during three of the four months. The ensemble-mean forecast performed much worse than the best overall forecast during March, and this prevented the overall MAE reduction from being substantially higher than 1.5%. The MAE reduction by the ensemble-mean method was approximately 6% for the other three months. The MAE of the ensemble-mean forecast was also slightly lower (0.01 m/s) than the average MAE of the composite of the best monthly forecast methods.

Table 4-8 Monthly MAE for the March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the Mountain View wind plant.

Month	Power Production Forecast MAE				Wind Speed Forecast MAE			
	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs
Mar-05	19.7%	17.4%	17.3%	20.8%	3.38	3.09	3.09	3.58
Apr-05	16.8%	17.1%	16.5%	17.9%	2.63	2.80	2.77	2.93
May-05	14.2%	13.6%	13.6%	16.0%	2.25	2.49	2.39	2.58
Jun-05	15.0%	15.7%	15.3%	16.9%	2.24	2.27	2.27	2.47
4-months	16.4%	16.0%	15.7%	17.9%	2.62	2.66	2.63	2.89

Note: The ensemble forecasts are the: (1) average of all forecasts in a 12-member ensemble ("Ensemble-12 Mean"), (2) individual method with the lowest MAE for the 4-month period ("Best Overall Method"), (3) composite of the individual methods with lowest MAE in each month ("Best Monthly Method"); and (4) average MAE of all 12 members of the ensemble ("Average of MAEs").

In summary, the ensemble-mean forecast method had a skill score of -0.4% relative to the best performing individual power production forecast method and a skill score of +1.5% relative to the best wind speed forecast method for this four-month period. These overall skill scores would have been more favorable had the ensemble-mean method not performed very poorly during March. During this month the forecasts based on the MASS-6 model significantly outperformed forecasts based on the other two physics-based models, which created a great disparity in performance among the ensemble members and resulted in a degradation of the performance of the ensemble-mean forecast.

Oak Creek Energy Systems, Tehachapi

The Oak Creek wind plant is located on a series of ridges in the Tehachapi Pass of Southern California. The rated capacity is 34.5 MW and the plant consists mostly of NEG-Micron turbines.

The power production and meteorological data for this wind plant were obtained through a direct ftp from an Oak Creek computer system. The power production data consisted of the energy output of the entire plant at one-minute intervals. The meteorological data consisted of wind speed and direction data from a single meteorological tower. The one-minute power production and wind data were aggregated to hourly data by AWST. The hourly data were used to train the MOS algorithms and compute the verification statistics.

Unfortunately, the data link between Oak Creek and AWST could not be set up until November 2004. The data for November 2004 were used as the initial training data for the MOS component of the forecast system and to construct the statistical plant output model. Therefore, forecasting began on December 1, 2004, and continued through the end of the forecast evaluation period on June 30, 2005. Thus, only seven months of evaluation statistics were produced for the Oak Creek wind plant.

Figures 4-62 and 4-63 present the number of hours available to the forecast verification pool for each month for the Oak Creek power production and wind speed forecasts. The available hours are the hours for which both forecast and observational data are available,

which enables those hours to be included in the verification pool. The forecast data used to compile this chart were from the MASS-6 forecasts, which were available for every possible forecast hour for the entire year. Therefore, the unavailable hours shown in these two charts are solely attributable to missing observational data. There were a total of 10,272 possible forecast hours (214 days in the December 2004 to June 2005 period times 48 forecast hours for each day's forecast) during the seven months of forecasting for the Oak Creek wind plant.

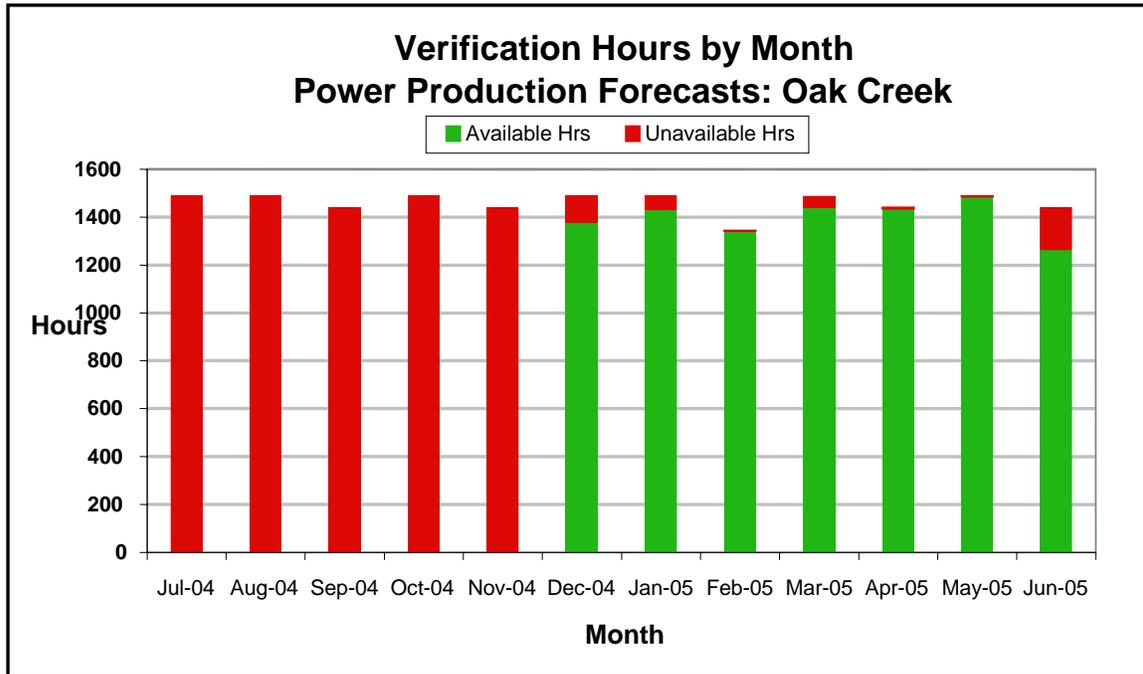


Figure 4-62 Number of hours for which sufficient data were available (green) and unavailable (red) to compute verification statistics for the PowerWorks power production forecasts.

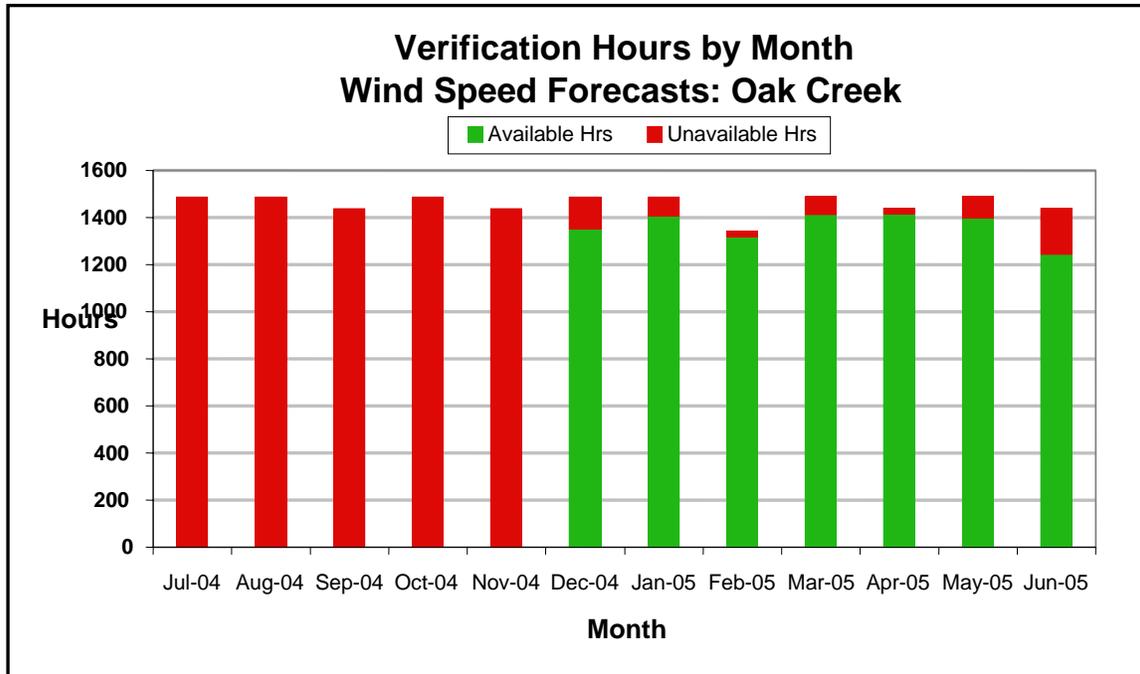


Figure 4-63 Number of hours for which sufficient data were available (green) and unavailable (red) to compute verification statistics for the PowerWorks power production forecasts.

A quick glance at these two charts indicates that there was only a very small fraction of missing data during this period. The power production data was missing for only 501 of the 10,272 forecast hours (4.9%) during the seven-month period. Therefore, the MASS-6 power production forecast verification statistics for the Oak Creek wind plant are based on 9,771 hours of data (an average of 45.7 hours per forecast cycle). A total of 435 of the 501 missing data hours were in the months of December, February, and June with June having the highest percentage of missing data hours at 12.3%. March and May were the best months for data availability with unavailable data for less than 1% of the forecast hours for each month. The size of the verification pool was smaller for the forecasts based on the COAMPS and WRF models since output from those models was not available for all of the possible forecast hours during the seven-month period

The availability of the wind speed and direction data was not quite as high as that of the power production data. The wind speed data was missing for 733 of the 10,272 forecast hours (7.1%). This meant that a total of 9,539 forecast hours (an average of 44.6 hrs per forecast cycle) was included in the wind speed verification pool for the Oak Creek wind plant. As in the case of the power production data, December, February, and June accounted for a majority of the missing data. A total of 504 of the 733 (68.8%) missing data hours were in those three months. The largest percentage of missing data hours was in June when data were not available 13.8% of the hours.

Overall Forecast Performance

Table 4-9 presents the forecast performance statistics by month for the best performing forecast method that used the output from the MASS-6 model. In contrast to the Mountain View wind plant, the best forecast performance for the Oak Creek wind plant over the entire seven-month period was achieved by the SMLR2-DP method. This is the method that uses a two-stage linear regression and directly predicts the power production without the use of an explicit plant output model. All of the MOS methods that used an explicit plant-scale output model performed significantly worse for the Oak Creek plant than the SMLR2-DP method. This was true for all three of the physics-based models. This appeared to be related to the fact that the plant-scale output model for the Oak Creek plant was constructed from data only for November 2004, which were the only Oak Creek data available at the time forecasting began in December 2004. This limited sample apparently yielded a somewhat unrepresentative plant-scale output model. As a result, the direct prediction method, which constructs an implicit plant output model each forecast cycle from a rolling lagged 60-day sample of physics-based model output and observed power production data, yielded better results.

The MAE of the power production forecasts for all verifiable forecast hours during the seven-month period (9,771 hours) was 15.0% of the installed capacity. The monthly MAE values ranged from a low of 11.7% for December 2004 and February 2005 to a high of 19.6% for April 2005.

The annual skill score vs. a persistence forecast was 33.2%. The monthly skill scores ranged from -0.9% for December 2004 to a high of 42.0% for May 2005. The annual skill score with respect to climatology was 27.1% and ranged from -34.7% for December 2004 to 54.1% for June 2005. It should be noted that the skill score relative to a climatology forecast was not based upon true climatology data, because none was available for this plant. Instead, a pseudo-climatology was constructed by computing averages for each hour of the day from the months immediately before and after the month for which the climatology was desired.

The annual MAE of the wind speed forecasts was 2.03 m/s, which is 40.6% of the mean wind speed for the seven-month period. This is actually quite a bit lower than the MAE of 2.65 m/s recorded for the Mountain View wind plant. Thus, although the power production forecasts had a higher MAE than the Mountain View wind plant, the wind speed forecasts actually had a substantially lower MAE. The monthly MAE values range from a low of 1.68 m/s for February 2005 to a high of 2.34 m/s for March 2005. The annual skill score with respect to persistence for the wind speed forecasts was 32.8%. The lowest monthly skill score was the 9.9% recorded for December 2004 while the highest monthly skill score was 44.7% for April 2005. The skill score with respect to climatology was 43.8% and ranged from -1.4% for December 2004 to 62.1% for April 2005. As in the case of the power production forecasts, the climatology-based skill score was based on a pseudo-climatology using data from the 2004-05 period and not on a true climatology.

Table 4-9 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-DP) for the Oak Creek wind plant.

Month	Power Production Forecast				Wind Speed Forecast			
	% MAE(1)	%MAE (2)	Skill-P	Skill-C	MAE (m/s)	% MAE (3)	Skill-P	Skill-C
Jul-04								
Aug-04								
Sep-04								
Oct-04								
Nov-04								
Dec-04	11.7%	141.8%	-0.9%	-34.7%	2.12	89.4%	9.9%	-1.4%
Jan-05	12.2%	87.6%	26.2%	15.8%	1.80	58.5%	30.9%	28.4%
Feb-05	11.7%	85.9%	12.9%	3.1%	1.68	53.2%	17.6%	39.6%
Mar-05	17.2%	62.6%	38.5%	31.1%	2.34	46.0%	38.9%	22.5%
Apr-05	19.6%	48.0%	40.2%	33.3%	2.26	36.1%	44.7%	62.1%
May-05	17.1%	30.1%	42.0%	21.4%	2.00	25.2%	32.0%	18.6%
Jun-05	15.4%	35.3%	35.4%	54.1%	1.95	27.7%	37.1%	70.3%
Annual	15.0%	57.1%	33.2%	27.1%	2.03	40.6%	32.8%	43.8%

(1) MAE as % of rated capacity (2) MAE as a % of production; (3) MAE as % of average wind speed

Figure 4-64 shows the error distribution for the seven months of power production forecasts. The chart shows two error distributions. One is for the first 24 hours of the forecast period, and the other is for the second 24 hours. The error distributions for the two periods are quite similar. Both are characterized by a strong central peak in the -5% to +5% range, with a rapid decrease in frequency on either side of this peak and a long tail of very low frequencies for medium and large negative and positive errors. There is a very noticeable asymmetry between the negative and positive tails of the distribution. Negative errors of a particular magnitude occur substantially more frequently than positive errors of a similar magnitude. For example, errors in the range of -30% to -25% of capacity occur about 4% of the time during forecast hours 1 to 24. However, positive errors of a similar magnitude occur during only about 1% of the hours.

This asymmetry is actually slightly greater during forecast hours 1 to 24 than during hours 25 to 48. Curiously, there is very little evidence of a similar asymmetry in the wind speed error distribution (not shown).

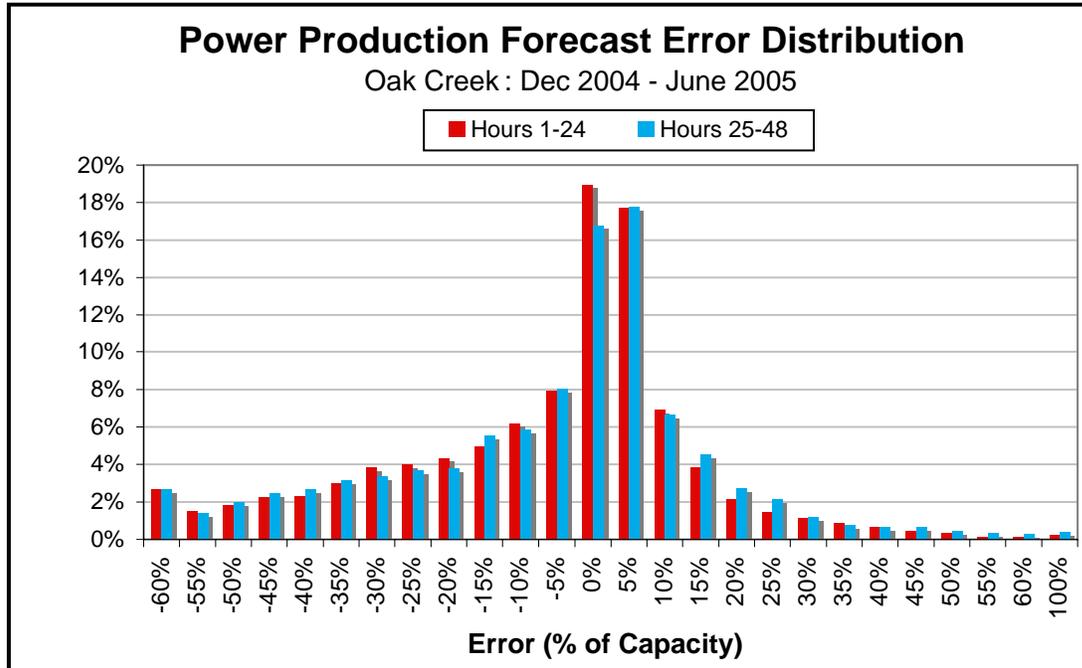


Figure 4-64 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the forecasts for the Oak Creek wind plant. The data sample includes 9,771 forecast hours during the December 2004 to June 2005 period.

Figure 4-65 shows the cumulative absolute error distribution for the 9,771 power production forecast hours that could be verified. Once again, separate error distributions are shown for the first and second 24 hours of the forecast period. As noted previously, the distributions for the first and second halves of the forecast period are very similar. The chart indicates that 50% of the errors (or the median error) are less than 10% for forecast hours 1 to 24 and 11% for hours 25 to 48.

As noted earlier, there is a strong relationship between the forecasted level of power production and the magnitude of the forecasted error because of the impact of the shape of the power curve on the sensitivity of power production forecast error to the error in the wind speed forecast.

Figure 4-66 indicates that the MAE is very low when the forecasted level of production is low and rises to a broad peak near 25% of capacity when the forecasted production is between 20% and 50% of installed capacity. Above a forecasted production level of about 50%, the MAE decreases as the forecasted production level decreases, and the MAE declines to the 15% to 20% of capacity range for forecasted production levels above 70% of capacity.

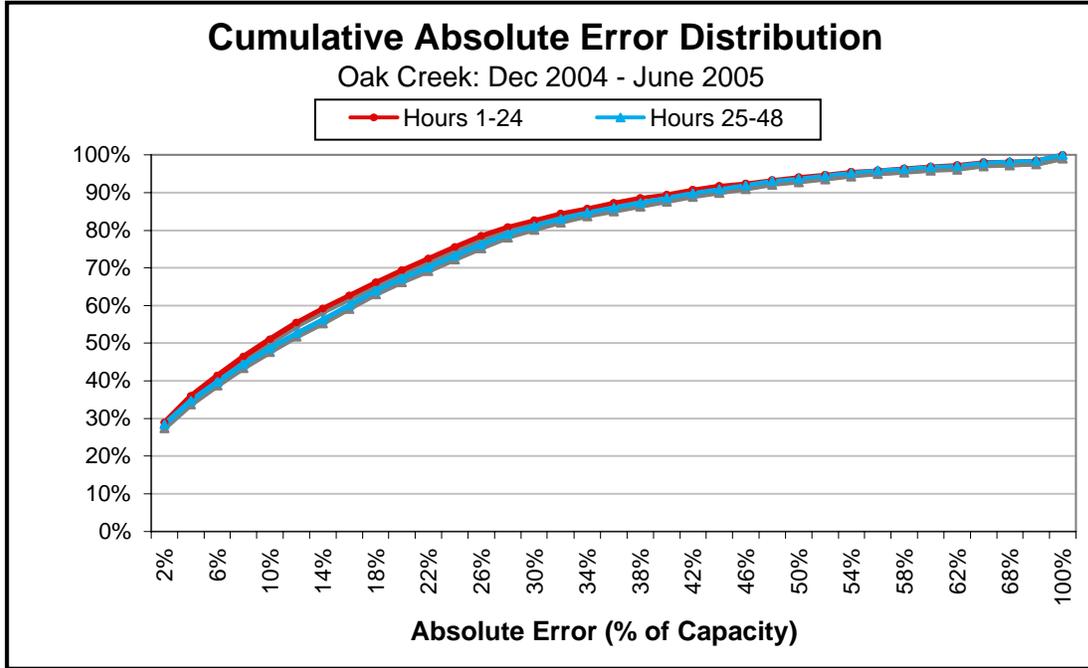


Figure 4-65 Cumulative absolute error frequency distribution (% of hours with an absolute error less than are equal to the value on the horizontal axis) for hours 1 to 24 (red line) and 25 to 48 (blue line) of the seven months of power production forecasts for the Oak Creek wind plant. The data sample includes 9,771 forecast hours.

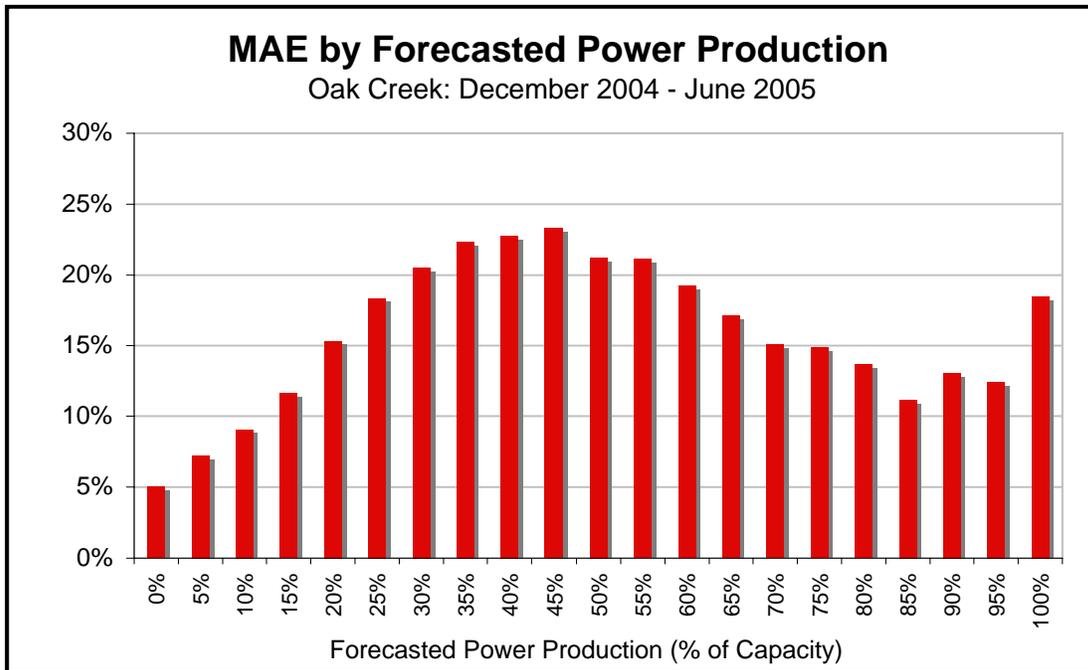


Figure 4-66 MAE of the Oak Creek power production forecasts vs. the forecasted power production. The MAE is shown for 21 bins of forecasted production. Each has a width of 5% of installed capacity and is labeled by the value at the upper end of the bin.

Figure 4-67 illustrates that the magnitude of the wind speed forecast error has much less dependence on the forecasted wind speed than the power production error has on the forecasted production. The chart depicts the MAE for 20 bins with a width of 1 m/s for forecasted wind speeds ranging from 0 to 20 m/s. The bin labeled “0” represents wind speed forecasts of exactly zero. All other bins are labeled by the wind speed value at the upper end of the bin. The magnitude of the error is generally lower for forecasts of low wind speeds. The MAE is under 2 m/s, which is below the overall MAE of 2.03 m/s for wind speeds under 4 m/s. The MAE of the wind speed forecasts gradually rises as the forecasted wind speeds increase and approaches 3 m/s for forecasted wind speeds above 10 m/s. The MAE values are somewhat erratic above 14 m/s. As indicated by the lighter shading of the bars in the chart, this is mostly attributable to the fact that the sample size in those bins is too small to yield statistically significant results. For example, the 17 m/s bin has an anomalously large MAE of 8.2 m/s, but there are only two hours that fall into that bin while the 16 m/s bin has a very low MAE of 1.93 m/s but is based on a sample size of four.

Performance vs. Forecast Time Horizon

An important consideration is the rate of degradation of the forecast performance (increase in forecast error) as the forecast time horizon increases. If the performance of the forecasts decreases rapidly with increasing look-ahead period, then it is useful and perhaps critical to have frequent forecast updates. On the other hand, if the degradation is slow then production and wind speed forecast MAEs based on the MASS-6 SMLR2-DP forecast method, the best performing method for the Oak Creek wind plant.

The MAE of the power production forecasts rises from just under 15% at the start of the forecast period to just more than 15% at the end of the forecast period. The rate of increase is somewhat obscured by oscillations of the forecast performance. The MAE of the persistence forecasts rises rapidly from about 5% for the first hour to near 15% by the 6th hour of the forecast period. The MAE then rises more slowly and levels out near 30% towards the end of the forecast period. The persistence forecasts outperform the *eWind* forecasts for the first couple of hours of the forecast period. The MAEs of the persistence and the *eWind* forecasts are very similar by hour three and remain similar through hour nine. After that time the MAE of the persistence forecasts rises much more rapidly than the *eWind* forecasts.

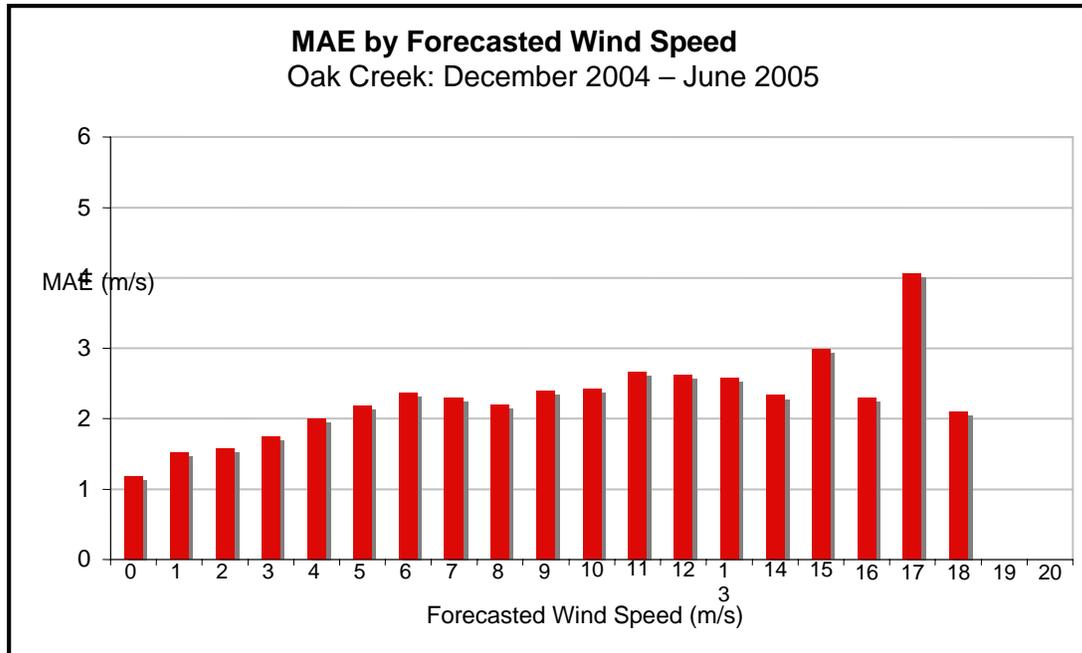


Figure 4-67 MAEs of the Oak Creek wind speed forecasts vs. the forecasted wind speed. The MAE is shown for 21 bins of forecasted wind speed. Each has a width of 1 m/s and is labeled by the value at the upper end of the bin. Bins represented by lightly shaded bars have a sample size of less than 30.

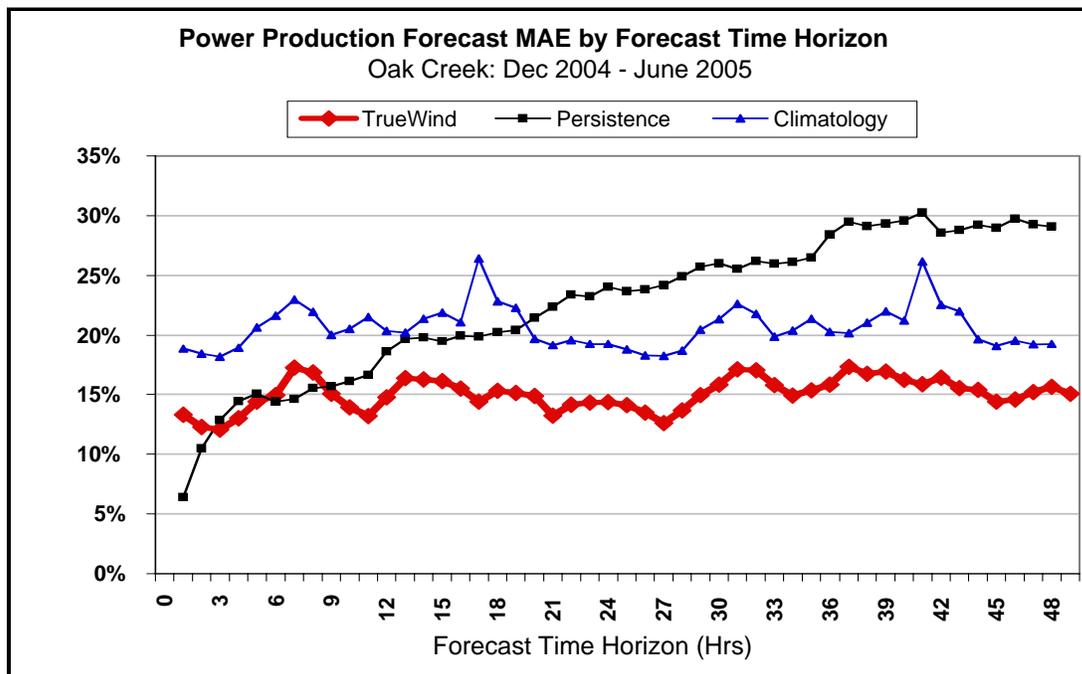


Figure 4-68 MAE of power production forecasts at the Oak Creek wind plant vs. forecast time horizon (the look-ahead interval) for seven months of *eWind* (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers) forecasts.

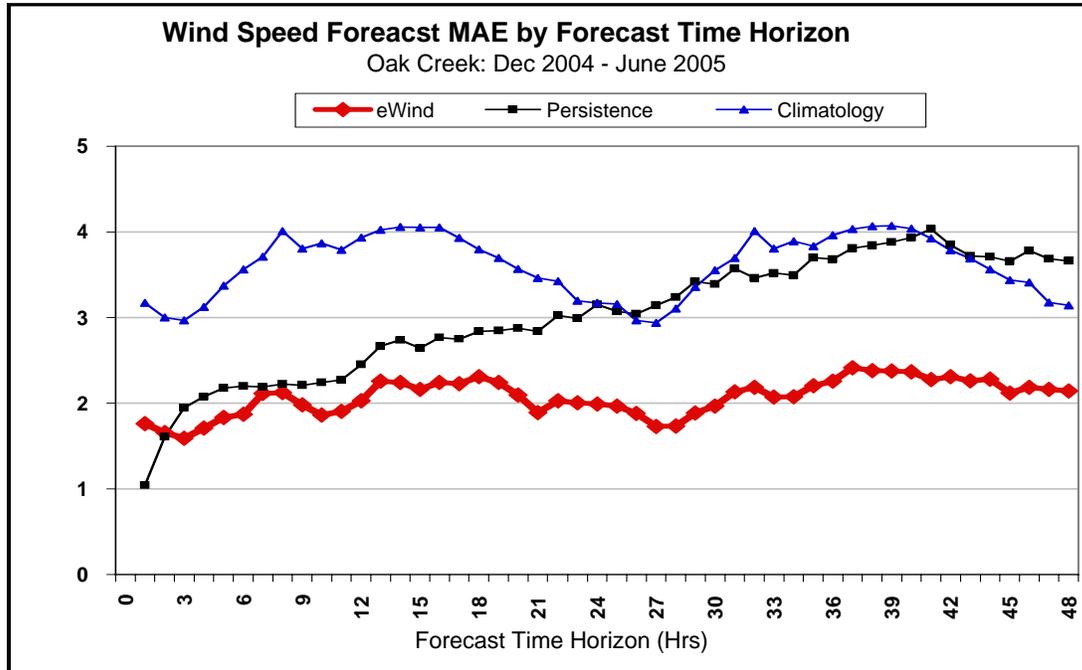


Figure 4-69 MAE of wind-speed forecasts at Oak Creek wind plant vs. forecast time horizon (the look-ahead interval) for seven months of *eWind* (red line with diamond markers), persistence (green line with square markers), and climatology (blue line with triangle markers).

As noted previously, the fact that persistence outperforms the *eWind* forecasts for the first couple of hours is the result of not using the latest data from the wind plant and its vicinity in the forecast process. The MAE of the climatology forecasts is generally in the 20% to 25% range, but it is somewhat noisy because of the limited amount of data used to construct the climatology. A linear fit to the *eWind* curve indicates that the average rate of error growth over the 48-hour period is 0.04% of installed capacity per hour or about 1.0% of installed capacity per day. This is somewhat lower than the MAE growth rate of 1.4% of capacity per day recorded for the Mountain View wind plant. This represents a growth rate of 6.7% per day relative the overall MAE of 15.0% for the 48-hr period.

The MAE of the Oak Creek wind speed forecasts increases from about 1.7 to 2.0 m/s during the first 6 hours to about 2.1 to 2.3 m/s during the last 12 hours of the forecast period. The MAE of the persistence forecasts rises rapidly from about 1 m/s for the first hour to 3 m/s by forecast hour 3. It then rises more slowly and reaches the 3.5 to 4 m/s level near the end of the forecast period. The SMLR2-DP forecasts begin to outperform persistence at hour three, although the difference is modest until about hour 12. After hour 12, the SMLR2-DP forecasts have a substantial and growing advantage over persistence as time increases. The MAE of the climatology forecasts remains between 3 and 4 m/s throughout the forecast period. The MAE exhibits a fairly well-defined diurnal cycle with minima at forecast hours 3 and 27 and what appears to be just after forecast hour 48. All of these minima correspond to about 11:00 a.m. PST. The maxima occur between forecast hours 12 to 15 and 36 to 39, which correspond to 8:00 p.m. to 11:00 p.m. PST. A linear least squares fit to the *eWind*

forecast error data indicates that the average rate of increase in MAE is 0.0085 m/s per hour or about 0.20 m/s per day. This is a growth rate of 9.9% per day relative to the MAE for the entire 48-hour period of 2.03 m/s. This is slightly less, on a percentage basis, than the growth rate of the power production error. This is somewhat different from the results for the Mountain View wind plant, where the percentage power production forecast error growth rate was twice as large as that for the wind speed forecasts.

Physics-Based Model Comparison

The generation of forecasts from three physics-based models provided an opportunity to determine the sensitivity of the forecast performance to the choice of physics-based model and to assess the relative performance of the three models. The performance comparison analyzed the forecast error statistics for forecasts generated by the same MOS method from all three models. For the Oak Creek wind plant, the best performing method, as noted earlier, was the SMLR2-DP method.

Table 4-10 presents the performance statistics for SMLR2-DP forecasts from all three models. The comparison of the monthly forecast performance statistics is complicated by the fact that the forecasts from all three models were not generated for the entire seven-month forecast period. As noted previously, the MASS-6 forecasts were produced for all seven months of the evaluation period and for all 48 hours. The COAMPS forecasts were also produced for all seven months, but these extended out only to 32 hours beyond the forecast delivery time. The WRF forecasts were only available for four of the seven months but did extend throughout the 48-hour forecast period. Since the COAMPS statistics do not include the last 16 hours of the forecast period when errors are expected to be slightly higher, the COAMPS errors would be expected to be lower than those for the other two models if all models performed similarly.

The power production forecasts based on the MASS-6 simulations significantly outperformed those based on the COAMPS simulations. Over the entire seven-month period, the MAE of the MASS-6 power production forecasts was 1.5% of capacity (12.6%) lower than the forecasts generated from the COAMPS output data. In addition, the MASS-6 forecasts exhibited a lower MAE during five of the seven months. However, a close examination of the monthly MAE data in Table 4-10 indicates that there is a well-defined pattern in the relative performance of the two models. The COAMPS-based power production forecasts yielded modestly lower (about 1.0% of rated capacity) during December and January. The MAEs of the two forecasts are fairly similar during February and March. However, the MASS-6 forecasts outperformed the COAMPS-based forecast by a wide margin (3% to 6% of capacity) during April, May, and June. This pattern suggests that there might be seasonal pattern in the performance differences between the two models with the COAMPS model performing better during the cold months and the MASS model performing better during the warm season months.

Table 4-10 Monthly MAE as a percentage of capacity for the 1- to 48-hour Oak Creek power production and wind speed forecasts by the three physics-based models.

Month	Power Production Forecast MAE			Wind Speed Forecast MAE		
	MASS 6	COAMPS	WRF	MASS 6	COAMPS	WRF
Jul-04						
Aug-04						
Sep-04						
Oct-04						
Nov-04						
Dec-04	11.7%	10.7%		2.12	1.65	
Jan-05	12.2%	10.9%		1.80	1.65	
Feb-05	11.7%	12.1%		1.68	1.63	
Mar-05	17.2%	17.3%	18.9%	2.34	2.25	2.51
Apr-05	19.6%	22.4%	20.6%	2.26	2.45	2.13
May-05	17.1%	23.2%	22.4%	2.00	2.42	2.43
Jun-05	15.4%	18.4%	16.9%	1.95	2.17	1.85
Annual	15.0%	16.5%	19.8%	2.03	2.03	2.24
Mar - Jun	17.3%	20.3%	19.8%	2.14	2.32	2.24

Curiously, the overall MAE of the MASS-6 and COAMPS wind speed forecasts for the seven months were the same (2.03 m/s). The COAMPS-based wind speed forecasts significantly outperformed the MASS-based wind speed forecasts during December and January and outperformed the MASS-based forecasts by a smaller margin during February and March. However, the MASS-based wind speed forecasts yielded much lower MAEs during April, May, and June.

A direct comparison with the WRF-based forecasts was possible only during March through June 2005. The MAE of the MASS-6 power production forecasts was 2.5% of rated capacity lower for these four months. This represents an MAE reduction of 13.0%. In addition, the MASS-6 forecasts yielded a lower MAE during all four months.

However, the wind speed forecast performance differences were not as clearly defined. The MAEs of the wind speed forecasts based on the MASS-6 output were 0.1 m/s (4.5%) lower for the entire four-month comparison period, but each model yielded significantly lower MAEs during two of the four months, April and June for the WRF-based forecasts, and March and May for the MASS-6 atmospheric simulations.

Overall, it is reasonable to conclude that there is no significant difference in the wind-speed forecast performance between the two models. But it is still interesting to note that the MASS-6 power production forecasts yielded a significantly lower MAE (1.0% of capacity) even during the month when the WRF-based wind speed forecasts yielded the largest MAE advantage over the MASS-6 forecasts.

The comparison between the COAMPS and the WRF power production forecasts appeared to show a slight advantage for the WRF-based forecasts. For the entire four-month comparison period, the WRF-based forecasts yielded a lower MAE by a margin of only 0.5% of rated capacity. The WRF-based forecasts also yielded a lower MAE during three of the four months. However, the COAMPS forecasts yielded a substantially lower MAE (1.6% of capacity) during March.

The four-month MAEs of the wind speed forecasts also gave a slight advantage to the WRF-based forecasts by 0.08 m/s. However, each model yielded the lower MAE during two of the months, and each model exhibited a substantially lower MAE during one of those months. Overall, it is hard to argue that there is any significant difference in performance between the two models.

The overall conclusion of the physics-based model testing at Oak Creek is that the MASS-6 model showed a fairly well-defined advantage in the power-production forecast performance. However, the wind-speed forecast performance differences were less decisive. The MASS-6 forecasts showed a slight edge in most of the overall statistics, but the month-by-month performance was fairly evenly split.

Forecast Ensembles

An ensemble mean forecast of the power production and the wind speed was generated from the 12 forecasts produced for the Oak Creek wind farm during the four test months (March to June 2005) for which all 12 forecasts were available. The 12 ensemble members included each combination of four MOS methods and three physics-based models.

Table 4-11 presents the MAE for the power production and wind speed forecasts for various combinations of the 12 ensemble members as well as for the ensemble-mean forecast. For the power production forecasts, the MAE for the ensemble-mean forecast was 1.6% of capacity higher than the MAE for both the best overall and the best monthly forecast methods. However, the MAE of the ensemble-mean forecast was 1.8% of capacity lower than the average MAE of all 12 of the individual forecast methods.

The poor performance of the ensemble-mean forecast relative to the best overall and best monthly methods can be traced to the fact that all methods that used an explicit plant output model performed poorly at the Oak Creek plant. Since three of the four MOS methods for each model used the same explicit plant output model, the ensemble mean was heavily weighted by forecasts that shared essentially the same error. An ensemble mean forecast will tend to do poorly when the majority of the forecasts share a significant error (or the errors are correlated).

The situation was somewhat different for the wind speed forecasts. For the four-month period, the ensemble mean forecast yielded an MAE that was 0.07 m/s (3.3%) lower than the best overall method and only 0.02 m/s higher than the composite of the best monthly methods. It was also 0.18 m/s lower than the average MAE of all of the methods. In practice the ensemble mean would have been the best forecast option since one would typically not know what the best monthly method is until a month is over.

The conclusion from the analysis of the Oak Creek ensembles is that the ensemble-mean forecast would have been the best choice for an operational wind speed forecast. However, the ensemble-mean method did not perform well for the power production forecasts. This was because errors associated with the use of the same explicit plant output model degraded the performance of nine of 12 ensemble members (those that depend on the explicit plant output model) for the power production forecasts. The three forecasts that did not use the explicit plant output model (those that employed the SMLR2-DP method) significantly outperformed the nine methods that used the explicit plant output model and the ensemble-mean method.

Table 4-11 Monthly MAE for the March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the Oak Creek wind plant.

Month	Power Production Forecast MAE				Wind Speed Forecast MAE			
	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs
Mar-05	19.2%	17.2%	17.2%	19.8%	2.36	2.34	2.25	2.46
Apr-05	20.6%	19.6%	19.6%	22.5%	2.16	2.26	2.13	2.32
May-05	21.0%	17.1%	17.1%	23.7%	2.03	2.00	2.00	2.24
Jun-05	14.6%	15.4%	15.4%	17.0%	1.73	1.95	1.82	1.97
4-months	18.9%	17.3%	17.3%	20.7%	2.07	2.14	2.05	2.25

Note: The ensemble forecasts include: (1) averaging all the forecasts in a 12-member ensemble ("Ensemble-12 Mean"), (2) the individual method with the lowest MAE for the 4-month period ("Best Overall Method"), (3) a composite of the individual methods with lowest MAE in each month ("Best Monthly Method"); and (4) the average MAE of all 12 members of the ensemble ("Average of MAEs").

PowerWorks Wind Project, Altamont Pass

The PowerWorks wind plant is located in the Altamont Pass of Northern California and has a rated capacity of 90 MW. This plant employs an older wind turbine model, which was installed during the mid- to late-1980s. The US Windpower 56-100 turbine is typical of turbines installed at that time. This turbine is rated at 100 kW and mounted on 60- and 80-foot lattice towers. It is a downwind-facing free-yaw machine with a three-blade, variable-pitch, 56-foot (18 meter) diameter rotor. The power curve for this turbine has a cut-in wind speed of approximately 5 m/s, a rated wind speed of 12 m/s, and a cut-out wind speed of 20 m/s.

At the start of the previous Energy Commission-EPRI forecasting project, PowerWorks, Inc., suggested that, to represent the range of wind conditions that exist at that site, the plant should be divided into 10 distinct groups of wind turbines, each associated with a nearby meteorological tower. The overall rated capacities of the turbine groups range from 17.5 MW (175 turbines associated with tower M723) to 3.4 MW (34 turbines associated with Tower M427). One of the towers, M438, reports ambient temperature and relative humidity in addition to wind speed and direction. In the previous project, PowerWorks, Inc. arranged with the Altamont Infrastructure Company to provide daily files containing the previous day's 30-minute power production, wind speed, and wind direction data. The arrangement

continued throughout the current project. Tower M427 was selected as the site for the wind speed forecasts in this project.

Figures 4-70 and 4-71 show the number of hours available to the forecast verification pool for each month for the power production and Tower M427 wind speed forecasts. The “available hours” are the hours for which both forecast and observational data are available. There were a total of 17,520 possible forecast hours (365 days times 48 forecast hours for each day’s forecast). There were very few hours of missing data for the PowerWorks wind plant. In fact, power production data were unavailable for only 48 hours during the entire one-year period. This represents an unavailable data rate of only 0.3%. The only months with unavailable data hours were October, December and January. The month with the largest percentage of missing data was December, which had a total of 30 unavailable data hours. Therefore, the power production forecast verification statistics are based on 17,472 hours of the possible 17,520 forecast hours. The size of the verification pool is smaller for the forecasts based on the COAMPS and WRF models because forecasts from those models were available for fewer hours than the forecasts from the MASS model.

Although the wind speed data availability was not as high, the wind speed data availability from Tower M427 was also quite good. The wind speed and direction data for M427 were missing for only 620 of the 17,520 forecast hours (3.5%). February 2005 exhibited the highest number of missing data hours with a total of 218 (16.2%). All other months exhibited less than 10% missing data hours and most were less than 5%. Thus, the wind speed verification statistics were based upon a total sample size of 16,900 hours.

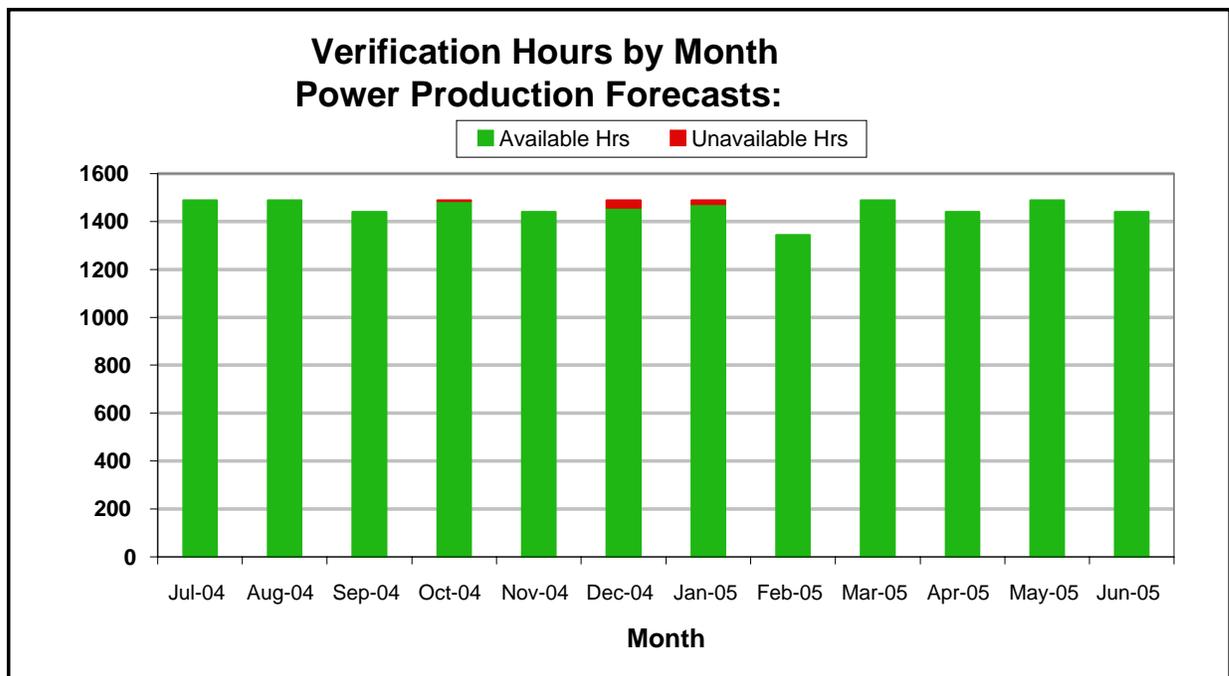


Figure 4-70 Number of hours for which data were available (green) and unavailable (red) for the computation of verification statistics for the PowerWorks power production forecasts produced by the best performing forecast method (MASS-6 SMLR2-ST).

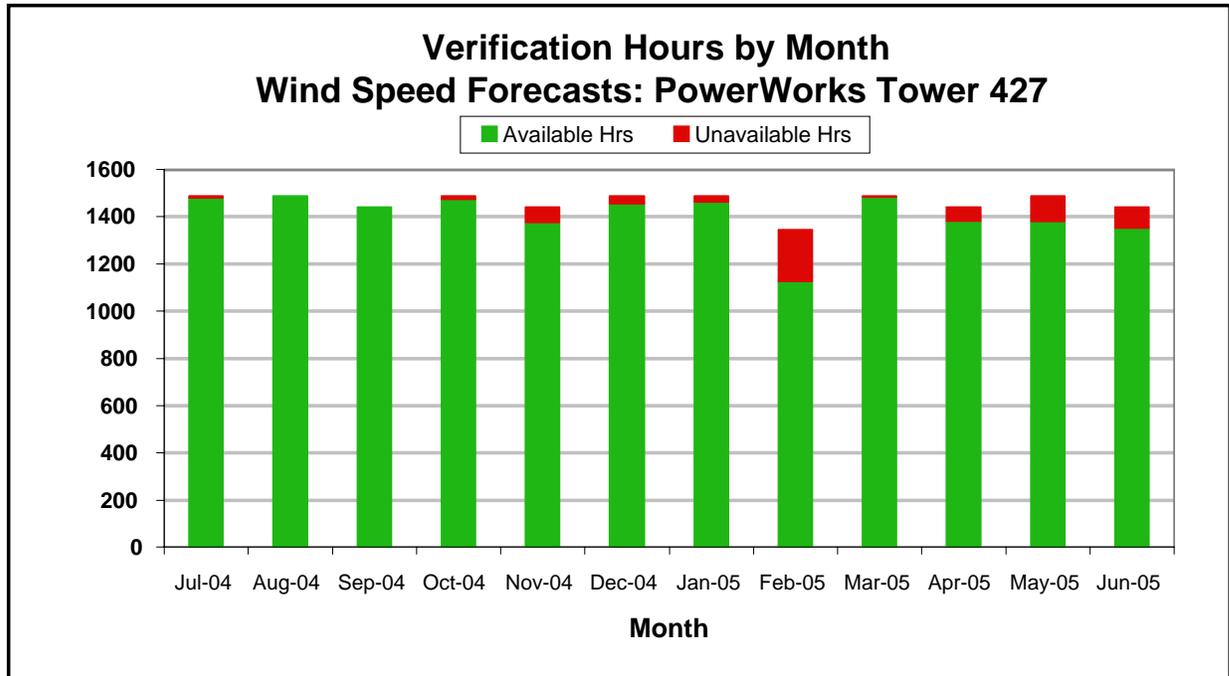


Figure 4-71 Number of hours for which data were available (green) and unavailable (red) for the computation of verification statistics for the wind speed forecasts for PowerWorks Tower M427 by the best performing forecast method (MASS-6 SMLR2-ST).

Overall, the availability and the quality of the data from the PowerWorks sites was excellent throughout the entire one-year period. Thus, the forecast verification statistics are probably among the most representative of all of the sites.

Overall Forecast Performance

Table 4-12 presents the forecast performance statistics by month for the best performing MOS method in combination with the MASS-6 forecast method, which was the SMLR2-ST MOS method. The MAE of the power production forecasts for all 17,472 verifiable forecast hours during the one-year period was 11.9% of the installed capacity. This is approximately 2.2% of capacity lower than the annual MAE of 14.1% achieved by the *eWind* forecast system for the year of forecasting in the previous Energy Commission-EPRI forecasting project (see Table 5-1 of Energy Commission and EPRI, 2003 for the comparable forecast performance statistics from the previous project). The monthly MAE values ranged from 2.0% during January 2005 to a high of 19.5% during August 2004. The annual skill scores were 26.5% and 9% relative to persistence and climatology, respectively.

The annual MAE of the wind speed forecasts for Tower M427 was 2.52 m/s, which is 34.4% of the mean wind speed for the one-year period. The monthly MAE values range from a low of 2.08 m/s for December 2004 to a high of 3.11 m/s for September 2004. The

annual skill score with respect to persistence was 28.2%. The lowest monthly skill score was the 11.5% recorded for November 2004 while the highest monthly skill score was 38.4% for April 2005. The skill score with respect to climatology was 12.4% and ranged from 1.7% for December 2004 to 27.9% for March 2005.

Figure 4-72 shows the error distribution for the one year of power production forecasts. Separate distributions are shown for the first and second 24 hours of the forecast period. The error distributions for the two periods are quite similar. Both are characterized by a strong central peak in the -5% to +5% range with a rapid decrease in the frequency on either side of this peak and a long tail of very low frequencies for medium and large negative and positive errors. The sharp peak in the -5% to 0% bin is a result of the fact that errors of exactly zero are included in the -5% to 0% bin.

Table 4-12 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-ST) for the PowerWorks wind plant

Month	Power Production Forecast				Wind Speed Forecast - Tower 427			
	% MAE(1)	%MAE (2)	Skill-P	Skill-C	MAE (m/s)	% MAE (3)	Skill-P	Skill-C
Jul-04	19.2%	34.8%	37.8%	3.4%	2.60	21.7%	34.8%	2.7%
Aug-04	19.5%	40.0%	38.8%	25.5%	2.68	24.3%	38.4%	22.7%
Sep-04	17.0%	67.0%	20.1%	39.9%	3.11	37.6%	22.3%	12.9%
Oct-04	16.8%	129.4%	-17.9%	16.3%	2.85	44.0%	12.5%	1.7%
Nov-04	3.9%	120.9%	37.5%	62.1%	2.49	61.2%	11.5%	5.7%
Dec-04	4.1%	204.5%	7.3%	60.7%	2.08	54.7%	14.7%	7.1%
Jan-05	2.0%	123.2%	14.2%	79.4%	2.15	49.6%	31.3%	8.9%
Feb-05	2.5%	150.6%	-21.1%	78.5%	2.09	48.4%	20.4%	14.3%
Mar-05	6.3%	88.4%	29.2%	65.9%	2.41	41.8%	31.1%	27.9%
Apr-05	15.5%	82.1%	24.8%	39.3%	2.68	32.7%	31.3%	13.0%
May-05	16.2%	56.4%	24.8%	41.6%	2.53	27.2%	36.6%	14.7%
Jun-05	18.5%	64.4%	36.3%	23.9%	2.51	2.3%	36.7%	8.7%
Annual	11.9%	58.5%	26.5%	38.9%	2.52	34.4%	28.2%	12.4%

(1) MAE as % of rated capacity (2) MAE as a % of production; (3) MAE as % of average wind speed

There are a fairly large number of hours for which the forecast and observations are zero during the low output months during the cold season. This results in an error of exactly zero and thus many more hours fall into the -5% to 0% bin than the 0% to +5% bin. There is a slight asymmetry between the positive and negative sides of the distribution. This is especially apparent for the bins in the 10% to 30% error range on either side of the central peak. The positive errors are substantially more frequent than the negative errors in this range. For example, the frequency of errors in the +10% to +15% bin is about 5% but the frequency of error of a similar magnitude but of the opposite sign is less than 4%. Thus, positive errors of this magnitude occur more than 25% more frequently than negative errors. This is true for the bins up to about 30%. Beyond 30% the differences are slight. The asymmetry is slightly greater for the 25 to 48 hour forecast period than for the 1 to 24 hour period.

Figure 4-73 shows the cumulative error distribution for the 17,472 power production forecast hours that could be evaluated is shown in. This chart shows the frequency of absolute errors that are equal to or smaller than error shown on the horizontal axis. Once again, separate error distributions are shown for the first and second 24 hours of the forecast period. As noted previously the distributions for the first and second half of the forecast period are very similar. This suggests that the performance of the forecasts degrades only slightly from the first to the second day of the forecast period. This chart indicates that the median error (50% frequency) is approximately 5.6% of the capacity for hours 1 to 24 and about 6.3% of capacity for hours 25 to 48. It also shows that 75% of the 1 to 24 hour forecast hours have an absolute error that is less than 18.0% of the capacity while 75% of the 25 to 48 hour forecast hours have an absolute error that is less than 18.6% of capacity.

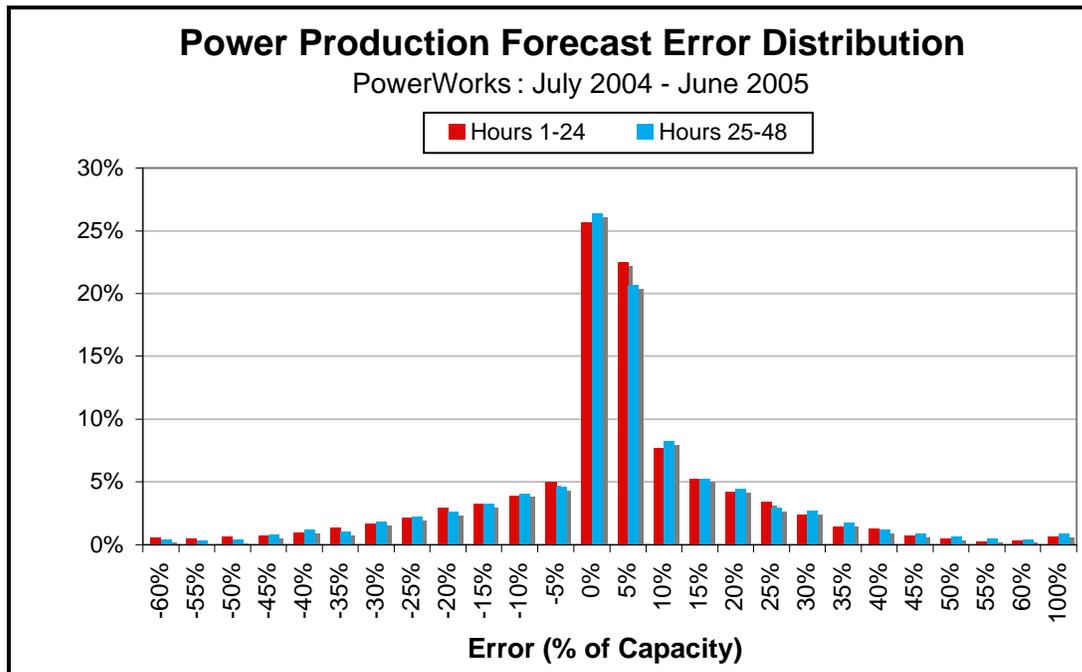


Figure 4-72 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the power production forecasts for the PowerWorks wind plant. The data sample includes 17,472 forecast hours.

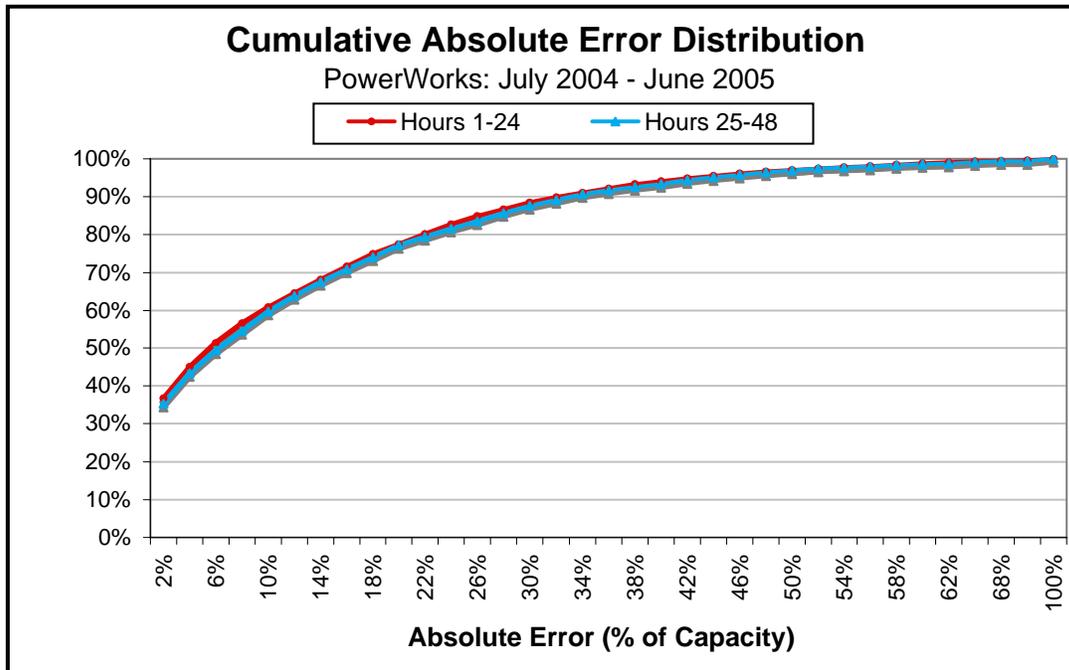


Figure 4-73 Cumulative absolute error frequency distribution (% of hours with an absolute error less than are equal to the value on the horizontal axis) for hours 1 to 48 of the power production forecasts for the PowerWorks wind plant. The data sample includes 17,472 forecast hours.

As noted earlier, there is a strong relationship between the forecasted level of power production and the magnitude of the forecasted error predominantly due to the shape of the turbine power curve. Figure 4-74 illustrates this relationship for the PowerWorks power production forecasts. The MAE is very low for low levels of forecasted production and rises to a peak just under 25% of capacity when the forecasted production is near 40%. The MAEs are generally slightly lower for forecasted production levels above 40%, but they do not decrease as much as they did for some of the other wind plants in this project. In fact, the errors become quite large for forecasted production above 90%. This may be related to the fact that when the forecasted production is near plant capacity, the wind speed will sometimes reach the turbine shut-down speed and the production will decrease to very low levels resulting in a large error. This may happen more frequently for the PowerWorks plant, because the turbines are older and the shut-down speed is lower (20 m/s) than it is for new plants with modern turbines (typically 25 m/s).

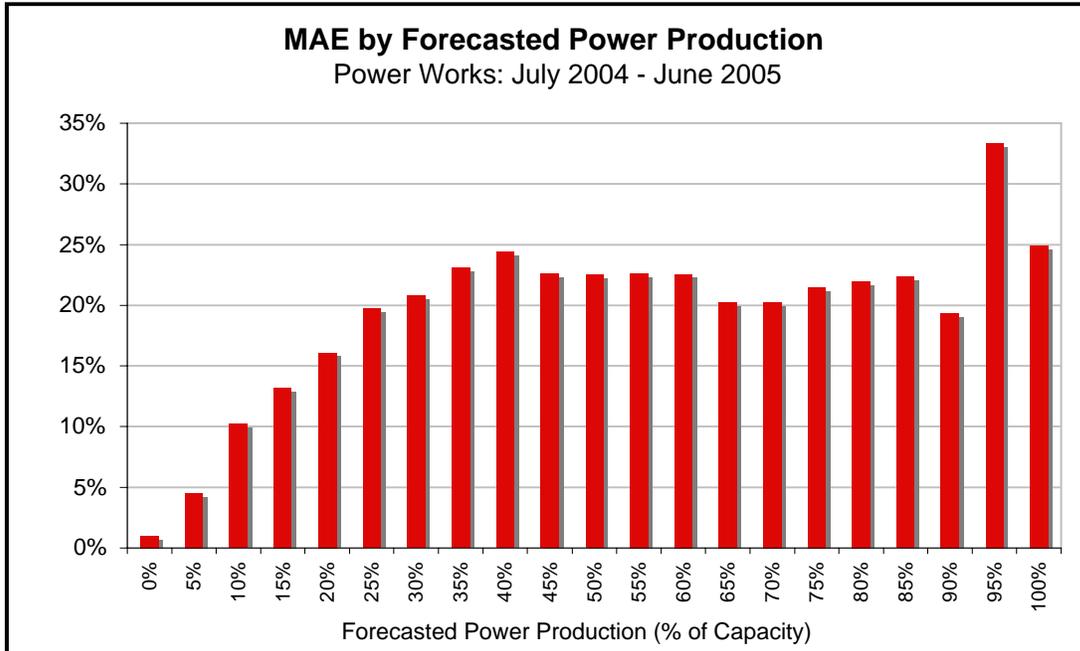


Figure 4-74 MAE of the PowerWorks power production forecasts vs. the forecasted power production. The MAE is shown for 21 bins of forecasted production. Each has a width of 5% of installed capacity and is labeled by the value at the upper end of the bin. Bars with light shading denote bins with a sample size of less than 30 hours.

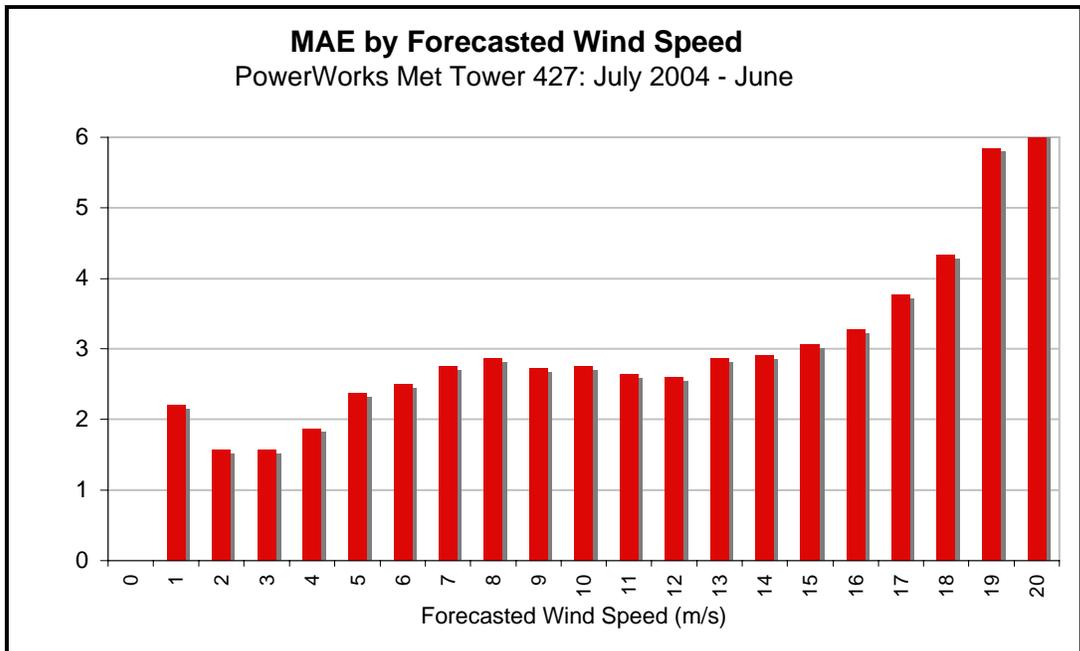


Figure 4-75 MAE of the 1 to 48 hour PowerWorks Tower M427 wind speed forecasts vs. the forecasted wind speed. The MAE is shown for 21 bins of forecasted wind speed. Each has a width of one m/s and is labeled by the value at the upper end of the bin. Bars with light shading denote bins with a sample size of less than 30 hours.

Forecast Performance vs. Time Horizon

An important consideration is the rate of degradation of the forecast performance as the look-ahead period increases. If the performance of the forecasts decreases rapidly with increasing look-ahead period then it is useful and perhaps critical to have frequent forecast updates. On the other hand, if the degradation is slow then the need for frequent updates is reduced. This issue is addressed by constructing MAE vs. forecast time horizon charts for the PowerWorks power production and the M427 wind speed forecasts (Figures 4-76 and 4-77). These charts illustrate the MAE for each look-ahead hour for the best performing *eWind* forecast method (MASS-6 SMLR2-ST) along with the persistence and climatology reference forecasts.

There is a gradual upward trend in the MAE of the *eWind* power production forecasts from near 10%, during the first few hours of the forecast period, to 13% to 14% towards the end of the (Figure 4-76). However this gradual upward trend is somewhat obscured by a pronounced diurnal cycle in the MAE. A linear fit to the *eWind* MAE curve indicates that the average rate of error growth over the 48 hour period is 0.04% of installed capacity per hour or about 1.0 % of installed capacity per day. This is a growth rate of 8.4% per day relative to the overall MAE of 11.9% for the 48 hour period. The MAE of the persistence forecasts increase from about 5% for the first forecast hour to approximately 10% for forecast hour six. The MAE remains near 10% through hour nine and then rises rapidly to a peak near 20% between hours 12 and 15. After 15 hours, the persistence forecast MAE

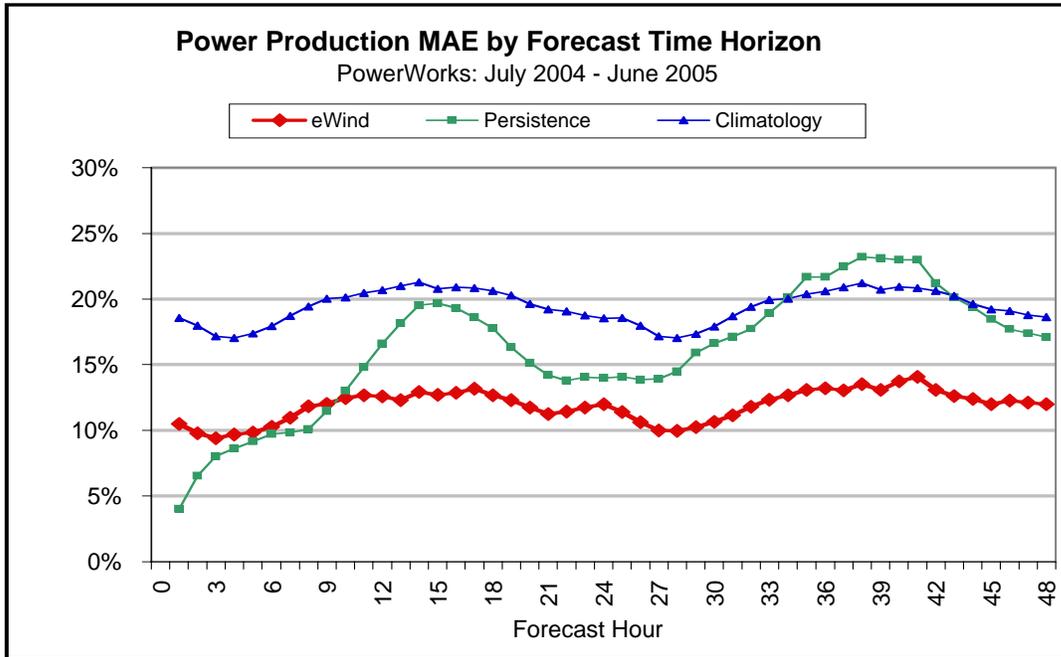


Figure 4-76 MAE of power production forecasts at the PowerWorks wind plant vs. forecast time horizon for 12 months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) forecasts.

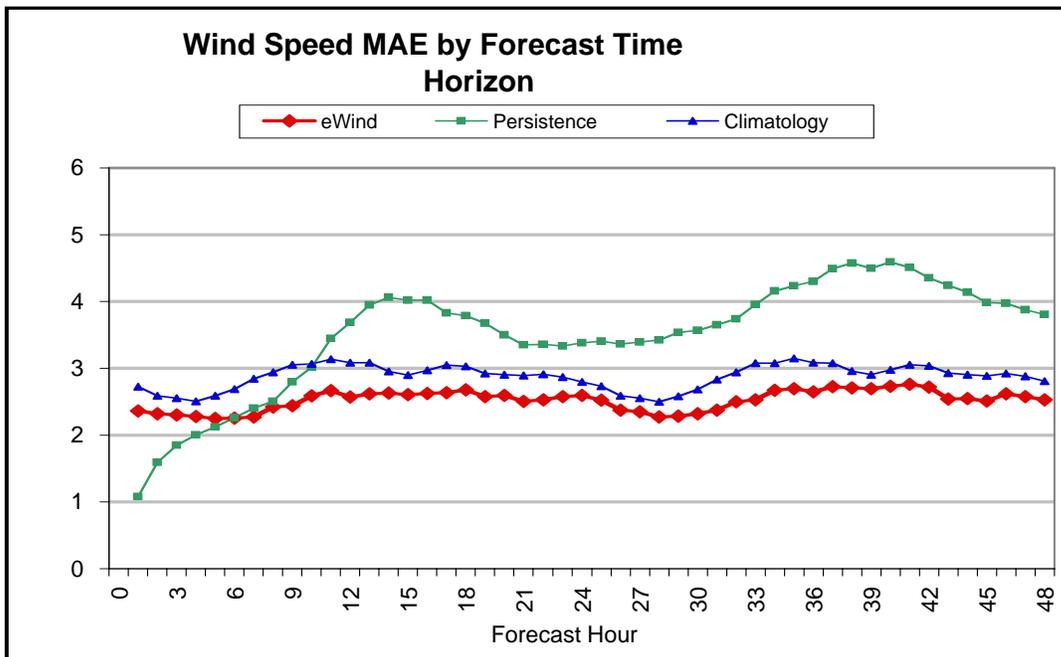


Figure 4-77 MAE of wind-speed forecasts at Tower M427 at the PowerWorks wind plant vs. forecast time horizon for 12 months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) forecasts.

exhibits a pronounced diurnal pattern. The persistence forecasts significantly outperform the *eWind* forecasts for the first three hours of the forecast period. The MAEs for the two forecasts are very similar for hours three through nine. The *eWind* forecasts substantially outperform the persistence forecasts after forecast hour nine. The MAE of the climatology forecasts remains in the 17% to 21% range for most of the forecast period and exhibits a well-defined diurnal pattern. The *eWind* forecasts outperform the climatology forecasts throughout the forecast period by a substantial margin.

The MAE of the M427 wind speed forecasts increases from near 2.2 m/s during the first 12 hours of the forecast period to about 2.8 m/s at the end of the 48 hour period (Figure 4-77). As in the case of the power production forecasts, the MAE exhibits a pronounced diurnal cycle. A linear least squares fit to the *eWind* forecast error data indicates that the average rate of increase of the MAE is 0.0052 m/s per hour or about 0.13 m/s per day. This is a growth rate of 5.0% per day relative to the MAE for the entire 48 hour period of 2.52 m/s. The performance characteristics of the persistence and climatology forecasts are similar to those seen in the corresponding power production forecasts although the amplitude of the diurnal cycles is somewhat less for the wind speed forecasts. This is most likely because the diurnal pattern in the power production errors are magnified by the sensitivity of the steeply sloped portion of the power curve to wind speed forecast errors. In general, the *eWind* forecasts outperform persistence after about nine hours and outperform the climatology forecasts throughout the forecast period.

Overall, the MAE vs. forecast time horizon charts for the PowerWorks plant indicate that the error growth rate for the power production forecasts is very slow (1.0% of capacity per day) and that the *eWind* forecasts outperform persistence by a substantial margin after about nine hours and perform significantly better than climatology throughout the 48-hour forecast period.

Physics-Based Model Comparison

The forecast performance comparison of the three physics-based models was based on the applications using the same MOS method, SMLR2-ST, which was the best-performing MOS method at the PowerWorks plant.

Table 4-13 presents the forecast performance statistics for SMLR2-ST forecasts for all three models. The comparison of the monthly forecast performance statistics is complicated by the fact that the forecasts from all three models were not generated for the entire one-year forecast period, as described previously.

The statistics in Table 4-13 indicate that the performance of the forecasts based upon the MASS-6 and WRF models was, in general, very similar and slightly better than that of the forecasts based on the COAMPS model. For the four months for which a comparison was possible, the MAE of the power production forecasts from both models was 14.1% and the MAE of the wind speed forecasts was also essentially the same, with the MASS-6 forecasts having an insignificantly lower MAE by 0.01 m/s. However, the performance of the forecasts from the two models was substantially different on a monthly basis. Each model

Table 4-13 Monthly mean absolute (MAE) of power production (% of capacity) and wind speed (m/s) for forecasts generated from three different physics-based models for the PowerWorks wind plant.

Month	Power Production Forecast MAE			Wind Speed Forecast MAE		
	MASS 6	COAMPS	WRF	MASS 6	COAMPS	WRF
Jul-04	19.2%	20.2%		2.60	2.53	
Aug-04	19.5%	23.0%		2.68	3.11	
Sep-04	17.0%	20.7%		3.11	3.52	
Oct-04	16.8%	16.4%		2.85	2.97	
Nov-04	3.9%	4.9%		2.49	2.70	
Dec-04	4.1%	3.0%		2.08	1.91	
Jan-05	2.0%	2.1%		2.15	2.14	
Feb-05	2.5%	1.8%		2.09	1.95	
Mar-05	6.3%	5.1%	7.3%	2.41	2.35	2.58
Apr-05	15.5%	16.2%	14.4%	2.68	2.74	2.45
May-05	16.2%	17.5%	15.7%	2.53	2.77	2.53
Jun-05	18.5%	21.1%	18.9%	2.51	2.85	2.58
Annual	11.9%	12.4%	14.0%	2.52	2.61	2.54
Mar - Jun	14.1%	15.0%	14.1%	2.53	2.68	2.54

had a lower MAE for two of the four months. The MASS-6 forecasts yielded a 1.0% lower MAE for March and a 0.5% lower MAE for June while the WRF-based forecasts yielded 1.1% and 0.5% lower MAEs for April and May.

The MAEs of the wind speed forecasts exhibited a similar pattern, with MASS-6 yielding a lower MAE during two months WRF yielding a lower MAE for one month, both models yielding the same MAE out to two decimal places for the fourth month. There was no evidence of a systematic pattern in the relative performance between the two models on a multi-month or seasonal basis. Of course, a four month sample is most likely inadequate to diagnose seasonal patterns in the relative performance of the forecasts. There may be systematic patterns on shorter time scales that are embedded within the four months of forecast data. However, this possibility was not investigated in this project.

Unlike the comparison of the MASS-6 and WRF forecasts, the comparison between the MASS-6 and COAMPS forecasts was available for the entire 12-month period. The forecast verification data indicate that the forecasts based on the MASS-6 numerical simulations outperformed the COAMPS-based forecasts by a modest margin. The actual difference is probably slightly larger since, as noted earlier, the COAMPS forecasts ended at 32 hours into the forecast period and the last 16 hours of the forecast period normally have larger errors than the first 32 hours. However, as noted in the previous section, the rate of error growth with increasing forecast time horizon is fairly slow

The MASS-6 power production forecasts produced a 12-month MAE that was 0.5% of rated capacity (4.0%) lower than that generated by the COAMPS-based forecasts. An examination of the monthly data indicates that the MASS-6 forecasts yielded a lower MAE during eight of the 12 months. There is some indication of a seasonal pattern in the relative

performance of the two sets of forecasts. The MASS-6 forecasts recorded a lower MAE during all six of the months in the warm season period from April through September, even though the months were in two separate years. However, the COAMPS-based forecasts yield a lower MAE for four of the six months during the cold season, extending from October through March. The largest difference in favor of the COAMPS forecasts occurred during March, when the MAE of these forecasts was 1.2% of capacity lower than that achieved by the MASS-6 forecasts.

The MAE pattern of the wind speed forecasts was fairly similar to that for the power production forecasts. For the entire 12-month period, the MAE of the MASS-6 forecasts was 0.09 m/s (3.4%) lower than that of the COAMPS-based forecasts. On a monthly basis, the MASS-6 forecasts had a lower MAE in seven of the 12 months. The monthly MAE pattern is quite similar to that of the power production forecasts with the MASS-6 forecasts yielding a lower MAE during five of the six months during the April through September period and COAMPS recording the lower MAE during four of the six months in the October through March period. The only significant difference between the power production and wind speed forecasts on a monthly basis occurred during July 2004, when the MASS-6 forecasts produced a power production MAE that was 1.0% of capacity lower than the COAMPS-based MAE, while the COAMPS wind speed forecasts generated an MAE that was 0.07 m/s lower.

The comparison of the WRF and COAMPS forecasts was limited to the four-month March through June period. For the entire period, the WRF-based forecasts yielded a power production MAE that was 0.9% of capacity lower than the COAMPS forecasts and a wind speed MAE that was 0.14 m/s lower. The WRF forecasts yielded a lower MAE during three of the four months. However, the pattern is somewhat suggestive of the seasonal MAE pattern seen in the comparison of the MASS-6 and COAMPS forecasts. COAMPS recorded the lower MAE during March while the WRF MAE was lower during April, May and June. That is consistent with the pattern of the COAMPS forecasts performing relatively better during the cold season months that was also observed in the comparison between the MASS-6 and COAMPS forecasts. One possible explanation is that the rate of error growth with increasing forecast time horizon is greater during the cold season. This causes the missing 16 hours at the end of the forecast period in the COAMPS forecasts to have a greater impact on the performance statistics for the cold season.

Forecast Ensembles

An ensemble-mean forecast of the power production and the wind speed was generated from the 12 different forecasts produced for the PowerWorks wind plant for the four months (March to June 2005) for which all 12 forecasts were available. The 12 ensemble members included each combination of the four types of MOS-based forecasts and the three physics-based models.

Table 4-14 presents the MAEs of the power production and wind speed forecasts for various combinations of the 12 ensemble members as well as for the ensemble-mean forecast for each of the four months. The MAE of the ensemble of power production forecasts based on

Table 4-14 Monthly MAE for the March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the PowerWorks wind plant.

Month	Power Production Forecast MAE				Wind Speed Forecast MAE			
	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs
Mar-05	6.2%	6.3%	5.1%	6.4%	2.33	2.29	2.26	2.43
Apr-05	14.6%	15.5%	14.0%	15.1%	2.52	2.65	2.45	2.65
May-05	16.1%	16.2%	15.7%	17.0%	2.40	2.56	2.39	2.59
Jun-05	18.6%	18.5%	18.1%	19.6%	2.40	2.43	2.43	2.62
4-months	13.9%	14.1%	13.2%	14.5%	2.41	2.48	2.38	2.57

Note: The ensemble forecasts include: (1) averaging all the forecasts in a 12-member ensemble (“Ensemble-12 Mean”), (2) the individual method with the lowest MAE for the 4-month period (“Best Overall Method”), (3) a composite of the individual methods with lowest MAE in each month (“Best Monthly Method”); and (4) the mean MAE of all 12 ensemble members (“Average of MAEs”).

the mean of the 12 ensemble members was 0.2% of capacity lower than the MAE of the best performing individual forecast (MASS-6 SMLR2-ST) for the four-month period. This is an MAE reduction (or skill score) of 1.4% relative to the best performing individual method. The MAE of the ensemble-mean forecasts was also 0.6% of capacity lower than the average MAE of all 12 individual forecasts, but it was 0.7% of capacity higher than the composite four-month MAE of the best performing forecast for each month.

The MAE of the ensemble mean wind speed forecasts for the four-month period was 0.07 m/s lower than the MAE of the best performing individual wind speed forecast method. This represents an MAE reduction of 2.8% relative to the best individual forecast. The ensemble-mean forecast had a lower MAE than the best overall individual forecast for three of the four months. The MAE of the ensemble-mean forecast was also 0.16 m/s lower than the average MAE of all 12 ensemble members. However, the ensemble-mean forecast had an MAE that was 0.03 m/s higher than the composite of the best performing method for each month.

In summary, the ensemble-mean forecast outperformed the best performing individual forecast method by 1.4% for the power production forecasts and 2.8% for the wind speed forecasts. The ensemble-mean forecast also produced a lower MAE during three of the four comparison months for both the power production and wind speed forecasts.

The ensemble-mean forecasts did not perform as well as the composite of the best method for each month for either the power production or wind speed forecasts, but it would be difficult to apply the latter approach to real time operations, since it would require a prediction of which method will perform best for the upcoming month.

SMUD Wind Project, Solano

The Sacramento Municipal Utility District (SMUD) operates a 15.2-MW wind plant in the Montezuma Hills of Solano County. The plant consists of 23 660-kW Vestas turbines on land owned by the utility. Sixteen of those turbines have hub heights of 50 meters and were

installed in 2003. The other seven have hub heights of 65 m and were installed in 2004. The plant is located on a series of low 200- to 300-foot hills just to the north of the Sacramento River.

The power production and meteorological data for this wind plant were obtained through the CA ISO PIRP data communication system. The power production data consisted of the energy output at 10-minute intervals. The 10-minute data were aggregated to one-hour intervals for use in the forecast algorithms. Meteorological data available from one meteorological tower included wind speed and direction, temperature and pressure data at a height of 50 meters. The meteorological data were also supplied at 10-minute intervals and aggregated to one-hour intervals for use in the forecasting procedure.

The SMUD facility became a PIRP participant during the summer of 2004. As a result, a full stream of data was not available until July 2004. The July data were used as the initial training set for the MOS algorithms and forecasts began in August 2004. Thus, only 11 months of forecasts could be generated for the SMUD wind plant.

Figures 4-78 and 4-79 show the number of hours in the forecast verification pool when both power production and wind speed data are available at the same time. The forecast data used to compile this chart was from the MASS-6 forecasts, which were available for every possible forecast hour for the verification period. Thus, the unavailable hours in these charts are solely attributable to missing observational data. There were a total of 16,032 possible forecast hours (334 days times 48 forecast hours for each day's forecast) in the 11-month period. Power production data were unavailable for 3,286 (20.5%) of these forecast hours.

Approximately half of the missing power production data hours occurred during October and November. Data were missing during 40.7% of the forecast hours in October and 74.9% of the hours in November. January also yielded substantial unavailable data (32.9%). The unavailable data resulted in a forecast verification pool size of 12,746 forecast hours. The size of the verification pool was smaller for the forecasts based on the COAMPS and WRF models, since output from those models was not available for all of the possible forecast hours.

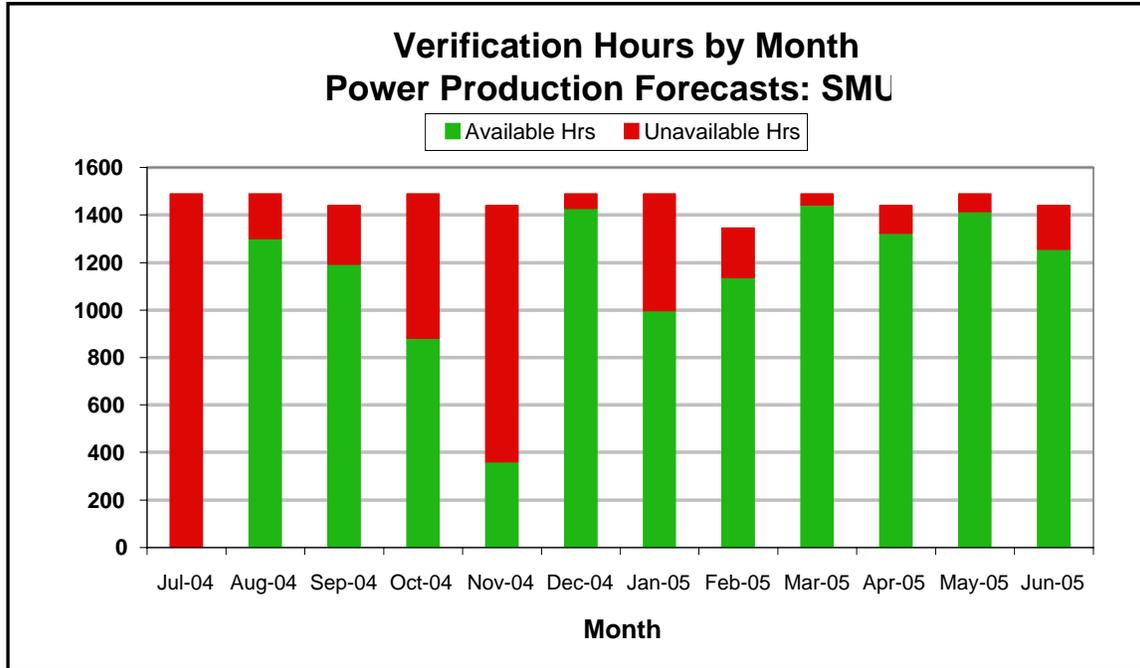


Figure 4-78 Number of hours for which data were available (green) and unavailable (red) for the computation of verification statistics for the SMUD power production forecasts.

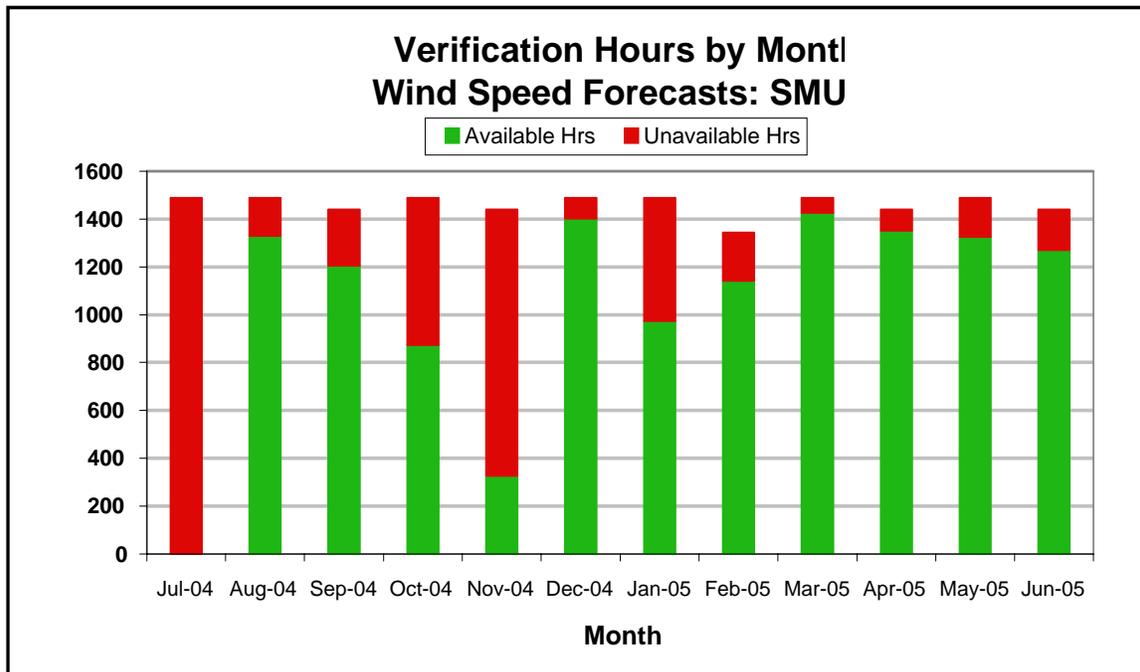


Figure 4-79 Number of hours for which data were available (green) and unavailable (red) for the computation of verification statistics for the wind speed forecasts for the SMUD meteorological tower.

The availability of wind speed data was similar to that of the power production data. In fact, for a large majority of the hours, the availability status of the meteorological and power production data was the same—either both were available or unavailable. The wind speed data was unavailable for 3,406 forecast hours. This represents 21.3% of the forecast hours in the 11-month period. As was the case with the power production data, the months with the most unavailable wind speed data were October, November and January.

Overall, there was a substantial amount (~20%) of unavailable power production and wind speed data during the 11-month period and a large fraction of the missing data hours were in the months of October, November and January. Therefore, the representativeness of the statistics for these months is questionable as it is for months immediately after these months, since training samples for those months will include many of the hours with missing data in the prior months.

Overall Forecast Performance

Table 4-16 presents the forecast performance statistics by month for the best performing MASS-6 forecast and MOS method, which was the SMLR2-ST method at the SMUD wind plant.

Table 4-15 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-ST) for the SMUD wind plant.

Month	Power Production Forecast				Wind Speed Forecast			
	% MAE(1)	%MAE (2)	Skill-P	Skill-C	MAE (m/s)	% MAE (3)	Skill-P	Skill-C
Jul-04								
Aug-04	20.3%	33.7%	42.7%	15.5%	2.00	20.5%	38.9%	14.0%
Sep-04	18.9%	49.9%	43.0%	22.0%	2.07	28.1%	45.0%	17.0%
Oct-04	18.6%	88.7%	11.7%	11.3%	2.22	39.1%	5.6%	9.5%
Nov-04	10.1%	113.6%	61.8%	-2.4%	1.88	33.1%	32.4%	-18.8%
Dec-04	12.8%	102.2%	40.8%	24.8%	2.31	55.9%	27.0%	9.3%
Jan-05	6.5%	94.2%	51.1%	32.6%	1.74	51.5%	15.4%	-5.0%
Feb-05	9.1%	90.8%	37.8%	17.3%	1.85	43.8%	19.0%	2.2%
Mar-05	14.2%	75.4%	35.6%	31.3%	2.06	40.3%	32.8%	25.2%
Apr-05	19.4%	63.4%	33.0%	10.2%	1.93	29.5%	32.7%	15.4%
May-05	19.8%	55.1%	27.0%	3.5%	1.88	25.1%	29.1%	3.2%
Jun-05	19.9%	43.2%	38.4%	5.0%	1.68	20.4%	37.6%	5.9%
Annual	16.0%	60.8%	37.1%	16.1%	1.98	35.9%	31.3%	10.9%

(1) MAE as % of rated capacity (2) MAE as a % of production; (3) MAE as % of average wind speed

The MAE of the power production forecasts for all verifiable forecast hours during the 11-month period (12,746 hours) was 16.0% of the installed capacity or 60.8% of the average production for the 11-month period. The monthly MAE as a percentage of capacity ranged from 6.5% for January 2005 to a high of 20.3% for August 2004. The monthly MAE ranged from 33.7% of rated capacity during August and 113.6% during November 2004. The skill scores relative to persistence and climatology were respectively 37.1% and 16.1%.

The 11-month MAE of the wind speed forecasts for the 50-meter tower was 1.98 m/s, which is 35.9% of the mean wind speed. The monthly MAE values ranged from 1.68 m/s during June 2005 and 2.31 m/s during December 2004. The skill score relative to persistence was 31.3%. The lowest monthly skill score was the 20.4% during June 2005, and the highest skill score was 51.5% during January 2005. The 11-month skill score relative to climatology was 10.9%.

Figure 4-80 presents two error distributions for the 11 months of power production forecasts. One is for the first 24 hours of the forecast period and the other is for second 24 hours. The error distributions for the two periods are quite similar. Both are characterized by a strong central peak in the -5% to 0% bin with a rapid decrease in the frequency on either side of this peak and a long tail of very low frequencies for medium and large negative and positive errors. The sharp peak in the -5% to 0% bin is a result of the fact that errors of exactly zero are included in the -5% to 0% bin. There are a fairly large number of hours for which the forecast and observations are zero during the low output months in the cold season. This results in an error of exactly zero and, thus, there are many more hours that fall into the -5% to 0% bin than the 0% to +5% bin. There is a slight asymmetry between the positive and negative sides of the distribution. There is a slightly greater frequency of negative errors (forecast is too low) of a particular magnitude than there is of positive error of the same magnitude. This is most noticeable in the 10% to 30% of capacity range. Another noteworthy feature of the error distribution is that the frequency of very large errors (greater than 60% of capacity) is much greater on the negative side of the distribution than it is on the positive side. This indicates that the power production is more likely to be close to the rated capacity when it was forecast to be near zero, than the reverse situation, that the actual power is close to zero when it is forecast to be close to rated capacity.

Figure 4-81 presents the cumulative error distribution for the 12,476 power production forecast hours that could be evaluated. This chart shows the frequency of absolute errors that are equal to or smaller than the error shown on the horizontal axis. Separate error distributions are shown for the first and second 24 hours of the forecast period. As noted previously, the distributions for the first and second half of the forecast period are very similar. This indicates that the quality of the forecasts does not degrade very much from the first to the second day of the forecast period. This cumulative error distribution indicates that the median error (50% frequency) is approximately 9.5% of the capacity for hours 1 to 24 and about 10.7% of capacity for hours 25 to 48. It also indicates that 75% of the 1- to 24-hour forecast hours have an absolute error that is less than 23.6% of the capacity while 75% of the 25- to 48- hour forecast hours have an absolute error that is less than 25.2% of capacity.

As noted earlier, there is a strong relationship between the forecasted level of power production and the magnitude of the forecasted error due mostly to the shape of the turbine power curve. Figure 4-82 illustrates this relationship for the 11-months of SMUD power production forecasts. The MAE is about 15% of rated capacity when the forecasted production is relatively low but rises to a broad peak between 20% and 25% when the forecasted production is between approximately 40% and 80% of capacity. This

corresponds to the steeply sloped portion of the power curve. The MAE tends to decrease as the forecasted production level rises above 80% and is near the 15% level when the forecasted production is near the rated capacity (95% to 100%).

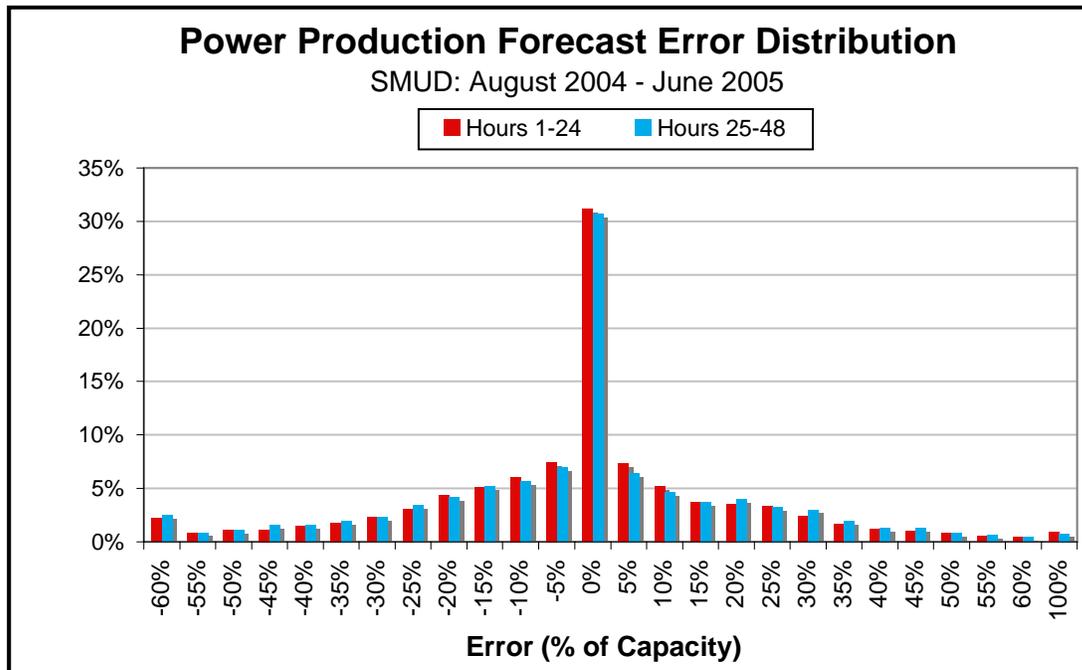


Figure 4-80 Error frequency distribution for hours 1 to 24 (red) and 25 to 48 (blue) of the forecasts for the SMUD wind plant. The data sample includes 12,746 forecast hours.

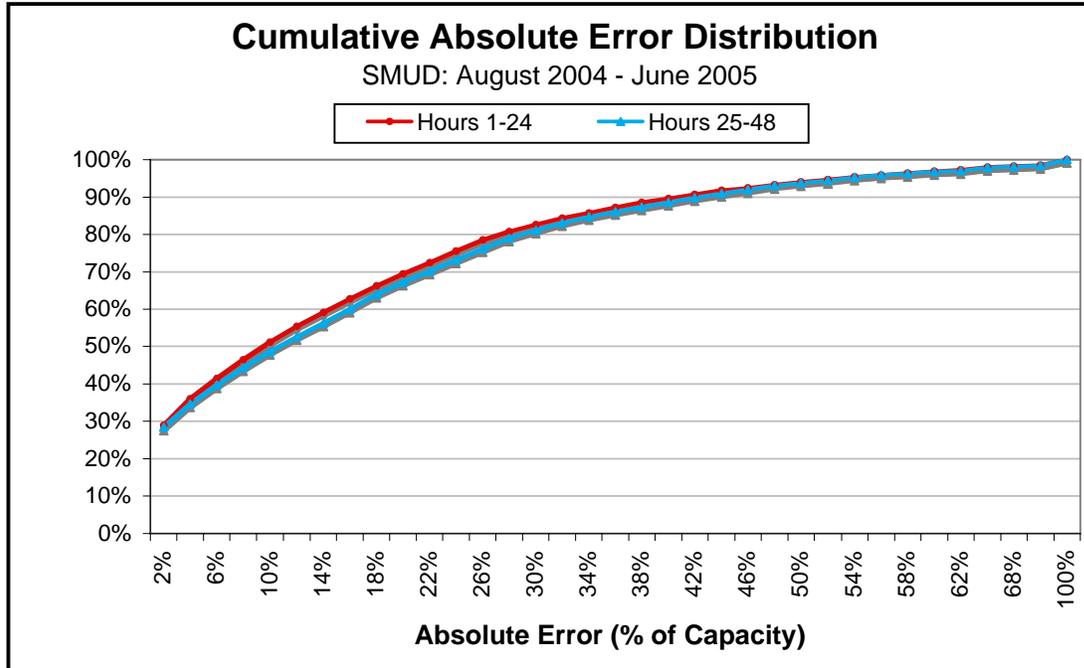


Figure 4-81 Cumulative absolute error frequency distribution (% of hours with an absolute error less than are equal to the value on the horizontal axis) for hours 1 to 48 of the forecasts for the PowerWorks wind plant. The data sample includes 12,746 forecast hours.

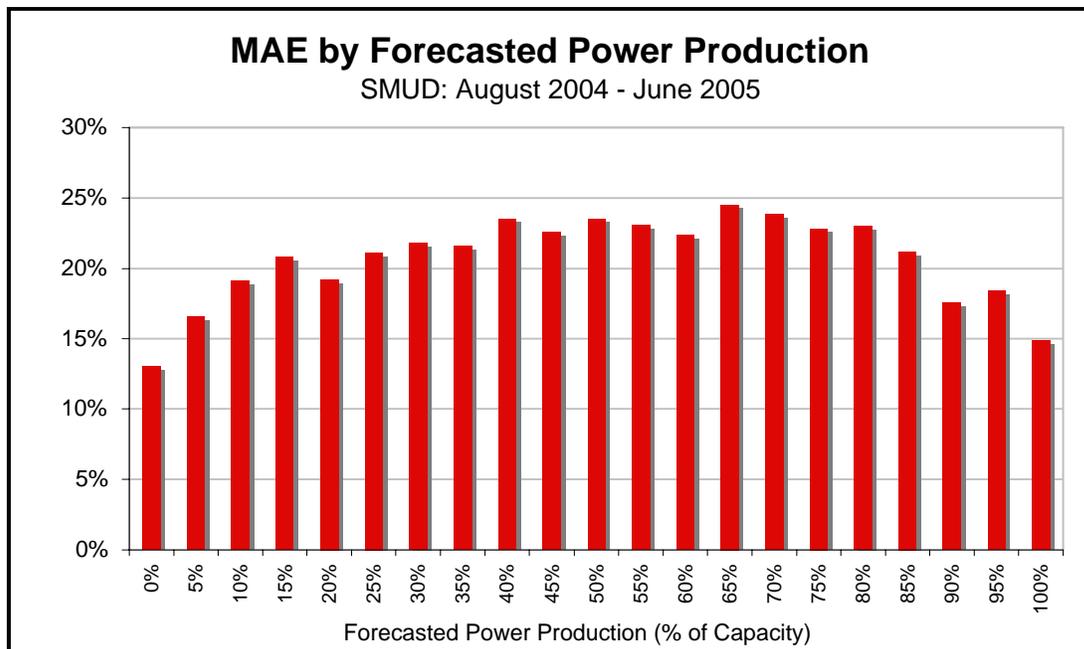


Figure 4-82 MAE of the SMUD 1- to 48-hour power production forecasts vs. the forecasted power production. The MAEs are shown for 21 bins of forecasted production. Each has a width of 5% of installed capacity and is labeled by the value at the upper end of the bin. Bars with light shading denote bins with a sample size of less than 30 hours.

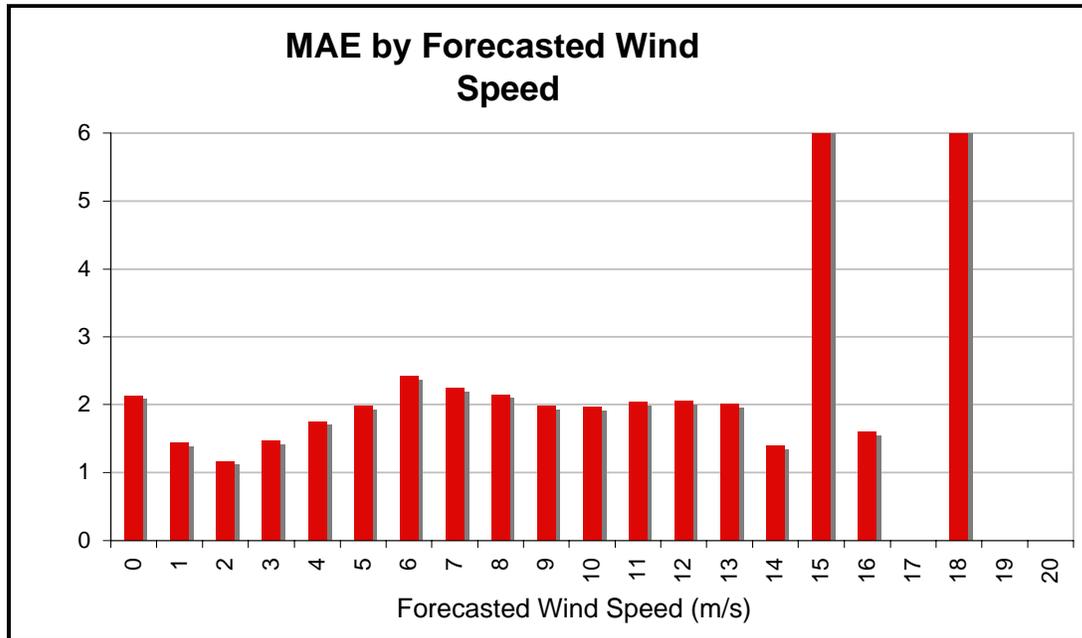


Figure 4-83 MAE of the 1- to 48-hour SMUD Met Tower wind speed forecasts vs. the forecasted wind speed. The MAEs are shown for 21 bins of forecasted wind speed. Each has a width of 1 m/s and is labeled by the value at the upper end of the bin. Bars with light shading denote bins with a sample size of less than 30 hours.

Forecast Performance vs. Time Horizon

An important consideration in the use of wind power production forecasts is the rate of degradation of the forecast performance as the forecast time horizon (the look-ahead period) increases. If the performance of the forecasts decreases rapidly with increasing look-ahead period then it is useful and perhaps critical to have frequent forecast updates. On the other hand, if the degradation is slow then there is less of a need for frequent updates.

This issue was addressed in the forecast performance analysis by constructing MAE vs. forecast time horizon charts for the SMUD power production and wind speed forecasts (Figures 4-84 and 4-85). These charts illustrate the MAE for each look-ahead hour for the best performing *eWind* forecast (MASS-6 SMLR2-ST) and the persistence and climatology reference forecasts. It should be noted again that the climatology forecasts for the SMUD plant were simply the monthly averages for each hour of the day and not a true climatology since a multi-year climatological dataset was not available for this plant. There is a gradual upward trend in the MAE of the *eWind* power production forecasts from near 15% during the first few hours of the forecast period to the 18% to 20% range towards the end of the forecast period. However this gradual upward trend is somewhat obscured by a noticeable diurnal cycle in the MAE. There is a peak in the MAE around forecast hours 18 and 42 (2:00 a.m. PST) and a minimum near forecast hour 30 (2:00 p.m. PST). Thus, the larger errors tend to occur in the middle of the night and the smaller errors during the middle of the day.

A linear fit to the *eWind* MAE curve indicates that the average rate of error growth over the 48-hour period is 0.07% of installed capacity per hour or about 1.7% of installed capacity per day. This is on the high end of the range of error growth rates calculated for the wind plants in this project. The persistence forecasts outperform the *eWind* forecasts for the first two hours of the forecast period and the MAE values are about the same for forecast hour three. The *eWind* forecasts significantly outperform the persistence forecasts after forecast hour three.

The MAE of the wind speed forecasts gradually increases from slightly under 2 m/s during the first 12 hours of the forecast period to just over 2 m/s towards the end of the 48-hour period (Figure 4-85). There is much less evidence of a diurnal cycle in the wind speed forecast MAE than there is in the power production forecast MAE pattern. A linear least squares fit of the hourly *eWind* wind speed forecast MAE data indicates that the average rate of increase in MAE is 0.0044 m/s per hour or about 0.11 m/s per day. The MAE of the persistence forecasts rises rapidly from near 1 m/s for forecast hour one to near 2 m/s by forecast hour five. The persistence forecasts have a lower MAE than the *eWind* forecasts for hours one through three and a similar MAE to the *eWind* forecasts for hours four through six. The *eWind* forecasts decisively outperform the persistence forecasts after forecast hour six.

Overall, the MAE vs. forecast time horizon charts for the SMUD plant indicate that the error growth rate for the day-ahead power production forecasts is higher than for the other plants in this project, but the wind speed forecast error growth rate is near the average of the other wind plants. This suggests that the forecasts of power production for the SMUD plant are more sensitive to wind speed forecast errors than those for most of the other wind plants addressed in this project. Interestingly, the *eWind* power production forecasts began to outperform persistence forecasts earlier in the forecast period than at the other wind plants in this project.

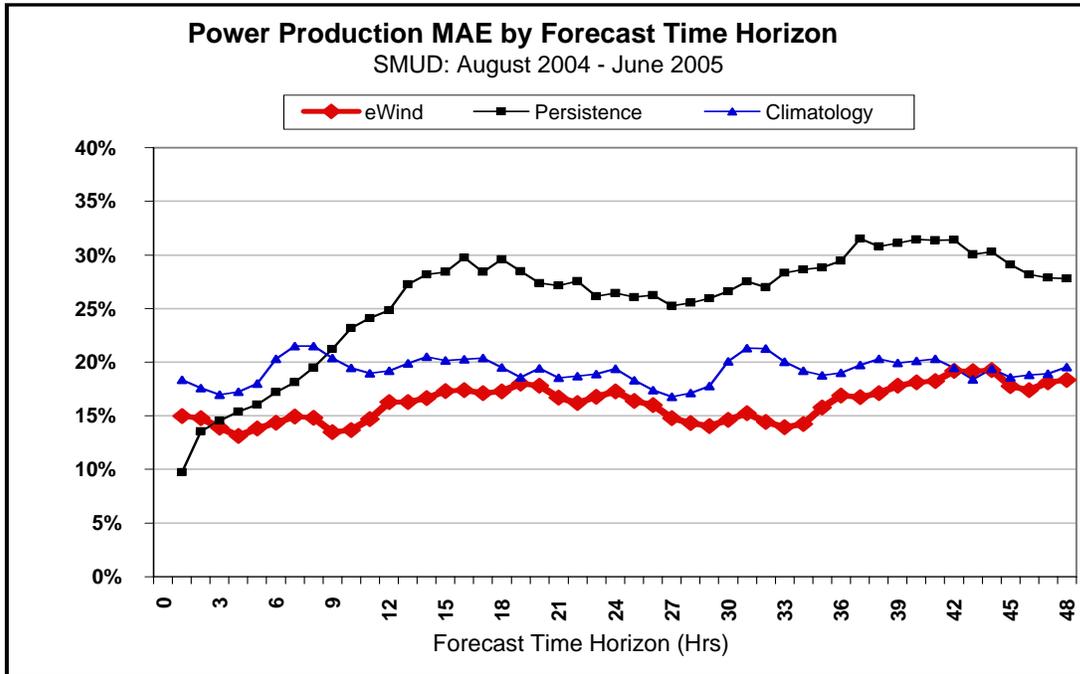


Figure 4-84 MAE by forecast time horizon for 11 months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) power production forecasts for the SMUD wind plant.

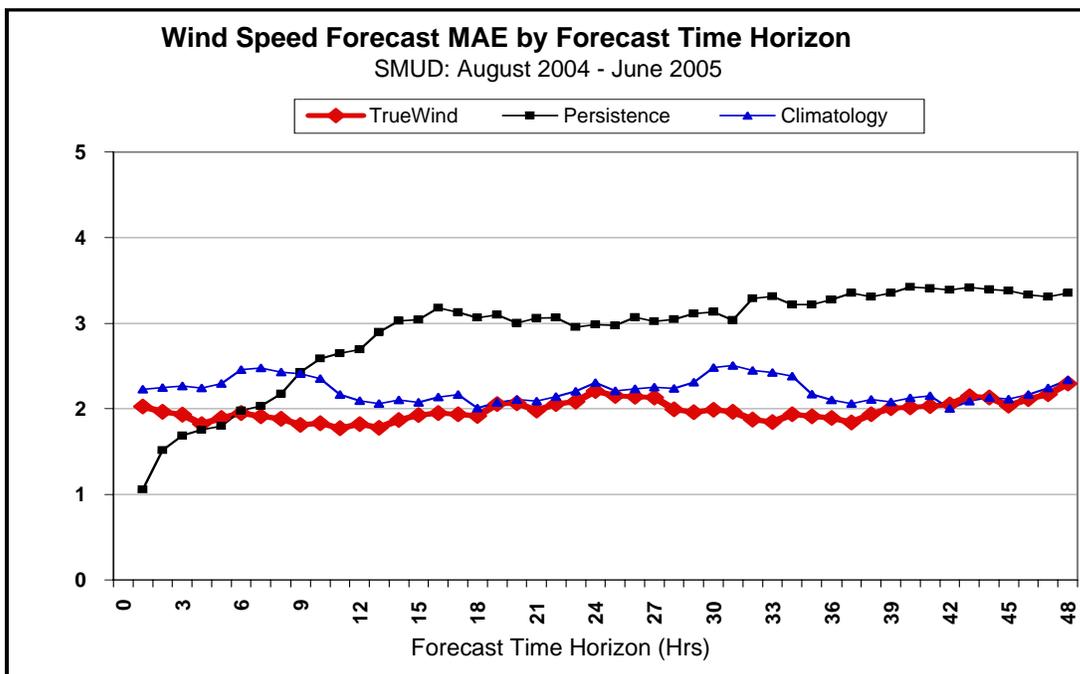


Figure 4-85 MAE by forecast time horizon for 11 months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) wind speed forecasts for the 50 m level at SMUD Meteorological Tower.

Physics-Based Model Comparison

A performance comparison among the forecasts generated by the three physics-based models employed in this project was done by analyzing the forecast error statistics for forecasts generated by the same MOS method applied to the output from all three physics-based models. The best performing MOS method for all three models for the SMUD wind plant was the SMLR2-ST scheme. Table 4-16 presents the forecast performance statistics for the SMLR2-ST MOS-adjusted forecasts and the three physics-based models. Comparison of the monthly forecast performance statistics is complicated by the fact that the forecasts from all three models were not generated for the entire one-year forecast period, as described previously.

The statistics in Table 4-16 indicate that the overall performance of the forecasts based upon the MASS-6, WRF and COAMPS models was fairly similar with perhaps a slight advantage for the MASS-6 and WRF forecasts over the COAMPS forecasts. For the four months for which a comparison was possible, the MAE of the power production forecasts from the MASS-6 model was 18.3% while the MAE for the WRF-based forecasts was 19.0%. The difference in overall performance was mostly attributable to the performance during March when the MASS-6 forecasts yielded an MAE that was 2.2% of capacity lower than that of the WRF-based forecasts. The MASS-6 forecasts had a smaller advantage during April and the MAEs were very similar during May and June, with the WRF-based forecasts having a slight advantage for May and the MASS-6 forecasts during June.

The MAEs of the wind speed forecasts from the two models exhibited a similar pattern to that of the power production MAEs. The overall MAE of the wind speed forecasts was essentially the same although MASS-6 showed a slight numerical advantage with an MAE of 1.89 m/s versus a 1.90 m/s MAE for the WRF-based forecasts. As in the case of the power production forecasts, the MASS-6 wind speed forecasts yielded a much lower MAE (0.26 m/s lower) than the WRF-based forecasts for the month during March. However, the wind speed MAE values during April and June were essentially the same, while the WRF-based forecasts yielded a significantly lower MAE (0.19 m/s lower) during May. Overall, it is hard to argue that there is a substantial overall difference in performance between the two models. However, there may be some seasons when there are significant differences, but it was not possible to identify them using the four-month sample..

A comparison between the MASS-6 and COAMPS forecasts was available for the entire 11-month period. The forecasts based on the MASS-6 output data generally outperformed the COAMPS-based forecasts by a modest amount. The actual difference is probably slightly larger since, as noted earlier, the COAMPS forecasts ended at 32 hours into the forecast period and the last 16 hours of the forecast period typically have larger errors than the first 32 hours. However, as noted in the previous section, the rate of error growth with increasing forecast time horizon is fairly low so the impact of this difference on the MAE comparison is most likely quite modest. The MASS-6 power production forecasts produced an 11-month MAE that was 0.4 % of rated capacity (2.5%) lower than that generated by the COAMPS-based forecasts. An analysis of the monthly data indicates that the MASS-6 forecasts yielded a lower MAE during 7 of the 11 months. In addition, there

were two other months (February and May) in which the COAMPS-based forecasts yielded a lower MAE by just 0.1% of capacity. If the MAE values were adjusted to account for the shorter COAMPS look-ahead period, the MAE of the COAMPS-based forecasts would most likely have been slightly higher than that of the MASS-6 forecasts. There is no suggestion of a seasonal pattern in the relative performance of the MASS-6 and COAMPS forecasts.

The data in Table 4-16 indicate that the MASS-6 wind speed forecasts had a slightly greater advantage over the COAMPS-based forecasts than the power production forecasts. The MASS-6 wind speed forecast MAE was lower by 0.12 m/s (5.7%) over the 11-month period. In addition, the MASS-6 forecasts yielded a lower MAE during 9 of the 11 months. MASS-6 had a substantially greater advantage in the MAE values for the raw wind speed forecasts that were interpolated directly from the three-dimensional grid of output data from the two models. For the entire 11-month period, the raw MASS-6 wind speed forecasts exhibited an MAE of 2.52 m/s while the raw COAMPS forecasts exhibited an MAE of 3.18 m/s.

The comparison of the WRF and COAMPS forecasts was limited to the four-month March through June period. For the entire period, the COAMPS-based forecasts yielded a power production MAE that was 0.3% of rated capacity lower than that for the WRF-based forecasts. Each model yielded a lower MAE during two of the four months.

Although the WRF-based power production forecasts yielded a slightly higher overall MAE than the COAMPS-based forecasts, the WRF-based wind speed forecasts actually yielded a lower overall MAE by 0.08 m/s (4.2%). Each model yielded a lower wind speed MAE during two of the four months. The main reason was that the WRF-based forecasts yielded a significantly lower MAE during May, when its wind speed forecast MAE was 0.21 m/s (11.1%) lower.

The overall forecast performance by the three physics-based models for the SMUD wind plant was fairly similar. The performance statistics suggest that the forecasts based on the MASS-6 simulations had a slight and probably insignificant edge in performance for both the power production and wind speed forecasts. There was no significant difference in forecast performance between the forecasts based on COAMPS and WRF physics-based models.

Table 4-16 Monthly MAE of power production (% of capacity) and wind speed (m/s) for forecasts generated by the three physics-based models for the SMUD wind plant.

Month	Power Production Forecast MAE			Wind Speed Forecast MAE		
	MASS 6	COAMPS	WRF	MASS 6	COAMPS	WRF
Jul-04						
Aug-04	20.3%	21.4%		2.00	2.11	
Sep-04	18.9%	21.4%		2.07	2.19	
Oct-04	18.6%	20.0%		2.22	2.48	
Nov-04	10.1%	15.2%		1.88	2.52	
Dec-04	12.8%	12.0%		2.31	2.24	
Jan-05	6.5%	6.6%		1.74	2.26	
Feb-05	9.1%	9.0%		1.85	1.87	
Mar-05	14.2%	15.3%	16.4%	2.06	2.26	2.32
Apr-05	19.4%	21.2%	20.0%	1.93	2.15	1.94
May-05	19.8%	19.7%	19.5%	1.88	1.90	1.69
Jun-05	19.9%	18.6%	20.2%	1.68	1.62	1.66
Annual	16.0%	16.4%	18.9%	1.98	2.10	1.92
Mar - Jun	18.3%	18.7%	19.0%	1.89	1.98	1.90

Forecast Ensembles

An ensemble-mean power production forecast and the wind speed was generated from the 12 different forecasts produced for the SMUD wind plant for the four months (March 2005 through June 2005) for which all 12 forecasts were available. The 12 ensemble members consisted of each combination of the four MOS methods and the three physics-based models.

Table 4-17 presents the MAEs for the power production and wind speed forecasts for several combinations of the 12 ensemble members as well as for the ensemble-mean forecast. For the power production forecasts, the MAE of the ensemble-mean forecast was 0.3% of rated capacity lower (1.7%) than the MAE of the best performing individual forecast method (MASS-6 SMLR2-ST) for the entire four-month period. It was also 1.5% of capacity lower than the average MAE of all 12 of the individual forecasts. The MAE of the ensemble-mean power production forecast was even 0.2% of capacity lower than the composite of the best performing forecast for each month.

The MAE of the ensemble-mean wind speed forecasts for the four-month period was 0.05 m/s lower than the MAE of the best performing individual wind speed forecast method, which is an MAE reduction (or skill score) of 2.6%. The MAE of the ensemble-mean forecast was also 0.12 m/s lower than the average MAE of the 12 ensemble members. However, in contrast to the power production forecasts, the ensemble-mean wind speed forecast yielded an MAE that was slightly higher (0.02 m/s) than the best monthly forecast composite.

In summary, the ensemble mean yielded a skill score of 1.7%, relative to the best performing individual power production forecast method, and a skill score of 2.6%, relative

Table 4-17 March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the SMUD wind plant

Month	Power Production Forecast MAE				Wind Speed Forecast MAE			
	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs
Mar-05	14.4%	14.6%	14.2%	15.5%	2.16	2.32	2.06	2.29
Apr-05	18.4%	17.7%	17.7%	19.7%	1.88	1.95	1.92	2.02
May-05	17.4%	18.6%	18.6%	19.5%	1.66	1.68	1.66	1.81
Jun-05	18.3%	18.7%	18.6%	19.6%	1.57	1.54	1.54	1.66
4-months	17.1%	17.4%	17.3%	18.6%	1.82	1.87	1.80	1.94

Note: The ensemble forecasts include: (1) averaging all the forecasts in a 12-member ensemble ("Ensemble-12 Mean"), (2) the individual method with the lowest MAE for the 4-month period ("Best Overall Method"), (3) a composite of the individual methods with lowest MAE in each month ("Best Monthly Method"); and (4) the mean MAE of all 12 ensemble members ("Average of MAEs").

to the best wind speed forecast method for the four-month test period. This means that, for the SMUD wind plant and the specific four-month period, the best forecast performance would be achieved by using the ensemble-mean forecast over any of the individual forecasts.

The ensemble-mean forecasts yielded a slightly higher MAEs than those generated using the best method for each month. However, there is currently no technique that can reliably determine in advance which method will perform best during a particular month, and this approach cannot currently be used in practice. However, the results indicate that further improvements in performance might be achieved if an effective technique to select or more heavily weight the best performing forecast for a given day, week, month or other time period could be formulated.

High Winds Wind Project, Solano

The High Winds Energy Center is located in the Montezuma Hills of Solano County. The plant is built on a series of low (200 to 300 foot) hills just to the north of the Sacramento River and is adjacent to the SMUD wind plant. The plant was constructed in 2003. It consists of 90 Vestas 1.8 MW turbines with a hub height of 65 meter, which yields a rated capacity of 162 MW.

The power production and meteorological data for this wind plant were obtained through the CA ISO PIRP data communication system. The power production data consisted of the energy output for the entire wind plant in 10-minute intervals. The 10-minute data were aggregated to one-hour intervals for use in the forecast algorithms. Data were available from six meteorological towers. The meteorological towers provided wind speed and direction, temperature and pressure data at a height of 65 meter. The meteorological data were also supplied at 10-minute intervals and aggregated to one-hour averages. The wind data from the six towers were also aggregated into a single composite wind speed and direction for the wind plant. The wind speed and direction forecasts were made for the composite wind speed and direction data for the wind plant.

The wind speed and direction forecasts were made for the composite wind speed and direction and not for the individual towers. This approach was employed because no information about the specific locations of the meteorological towers and turbines was available for this wind plant. It would be reasonable to expect that some improvement in power production forecast performance could be obtained by assigning clusters of turbines to individual meteorological towers and then forecasting the wind and power production for each cluster in a manner similar to what was done for the PowerWorks wind plant in Altamont Pass.

Figures 4-86 and 4-87 show the number of hours available to the forecast evaluation pool. The “available hours” are the hours for which both forecast and observational data are available. MASS-6 forecasts were available for all forecast hours during the 12-month period. Thus, any unavailable hours were attributable to missing observational data. There were a total of 17,520 possible forecast hours (365 days times 48 forecast hours for each day’s forecast) during the 12-month period. Power production data were unavailable for only 764 (4.4%) of these forecast hours. Most of the missing data hours occurred during December when power production data were not available for 608 hours (40.8% of the December hours). The percentage of unavailable data hours for each of the other 11 months was 3.3% or less. Thus, 16,756 hours were available to the power production forecast evaluation pool. The size of the evaluation pool is smaller for the forecasts based on the COAMPS and WRF models since output from those models was not available for all of the possible forecast hours.

The availability of wind speed data was not quite as high as that of the power production data. Wind data was unavailable for 1,526 (8.7%) of the forecast hours. The majority of the unavailable wind speed data was for the months of September and December, which together accounted for 1,086 of the unavailable data hours. However, 808 of these hours occurred during December. The fraction unavailable wind data was quite low for the other 10 months, with no month having an unavailable data percentage greater than 6%.

In summary, the availability of power production and wind data for the High Winds plant was very good with the exception of December, which had a large fraction of unavailable power production and wind speed data.

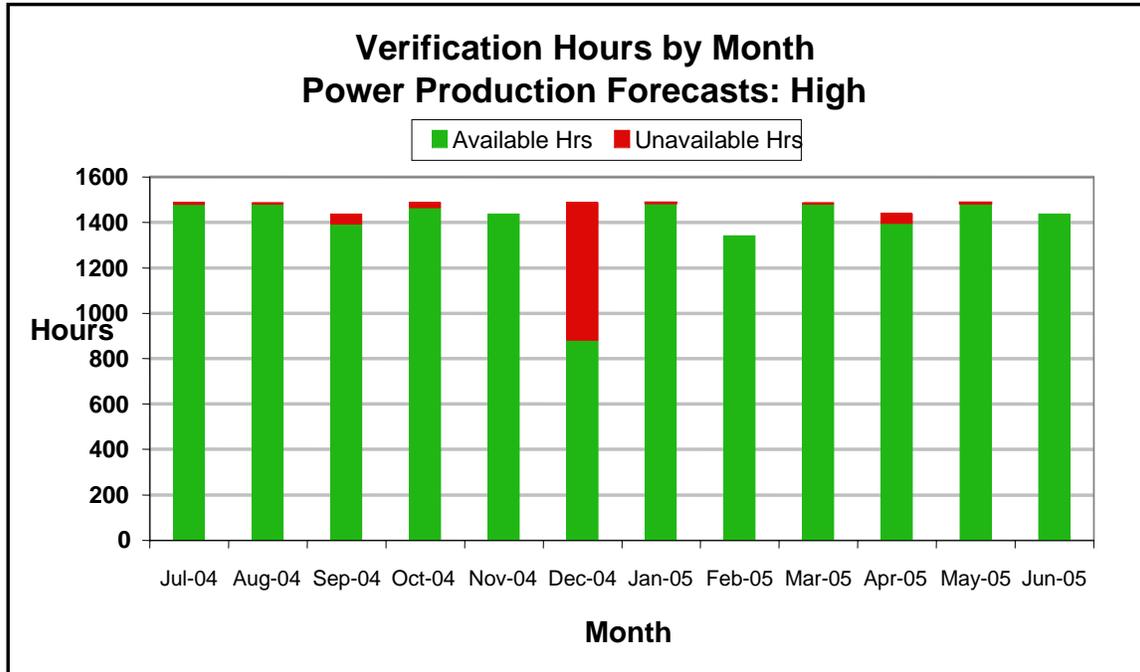


Figure 4-86 Number of hours with available and unavailable data for the computation of power production forecast verification statistics for the High Winds wind plant for the July 2004 to June 2005 period.

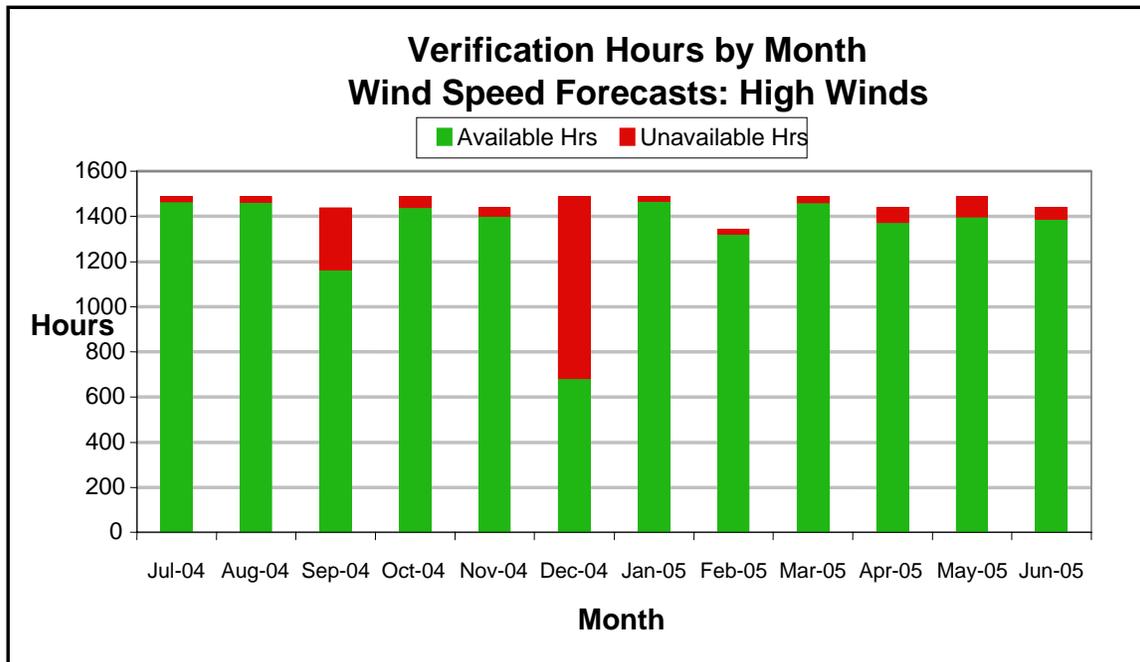


Figure 4-87 Number of hours with available and unavailable data for the computation of wind speed forecast verification statistics for the High Winds wind farm for the July 2004 to June 2005 period.

Overall Forecast Performance

Table 4-18 presents the forecast performance statistics by month for the best performing MASS-6 forecast MOS method (SMLR2-ST) at the High Winds plant in the Solano area.

The MAE of the power production forecasts was 16.8% of the installed capacity or 50.7% of the average production for all verifiable forecast hours during the one-year period (16,756

Table 4-18 Monthly MAE and skill scores for the overall best performing power production and wind speed forecast method (MASS-6 SMLR2-ST) for the High Winds wind plant

Month	Power Production Forecast				Wind Speed Forecast			
	% MAE(1)	%MAE (2)	Skill-P	Skill-C	MAE (m/s)	% MAE (3)	Skill-P	Skill-C
Jul-04	17.6%	24.5%	25.9%	-14.3%	1.75	16.0%	23.8%	-5.6%
Aug-04	18.2%	29.1%	42.1%	18.8%	1.93	19.0%	37.3%	12.9%
Sep-04	18.7%	50.8%	36.0%	23.9%	2.34	31.1%	32.3%	12.2%
Oct-04	19.4%	70.9%	21.0%	25.3%	2.54	39.0%	23.4%	14.1%
Nov-04	11.9%	86.3%	26.0%	33.1%	2.55	60.2%	19.5%	16.3%
Dec-04	17.7%	100.9%	35.9%	12.7%	2.69	53.3%	35.9%	12.6%
Jan-05	8.8%	90.2%	15.5%	34.1%	2.12	60.9%	17.4%	-6.2%
Feb-05	12.1%	89.0%	18.9%	19.0%	2.14	50.1%	17.6%	1.5%
Mar-05	17.9%	74.1%	20.7%	28.3%	2.36	43.6%	32.0%	23.8%
Apr-05	19.3%	61.3%	28.7%	13.4%	2.07	30.2%	27.1%	14.4%
May-05	20.9%	53.8%	19.1%	-1.3%	2.03	25.7%	25.2%	-1.3%
Jun-05	19.3%	39.0%	37.1%	11.1%	1.66	18.8%	36.1%	8.0%
Annual	16.8%	50.7%	28.6%	17.5%	2.16	31.9%	27.1%	9.9%

(1) MAE as % of rated capacity (2) MAE as a % of production; (3) MAE as % of average wind speed

hours). The monthly MAE ranged from a low of 8.8% of rated capacity during January 2005 and a high of 20.9% during May 2005. As a percentage of the average monthly production, the MAE ranged from 24.5% during August 2004 to 100.9% during December 2004. The skill scores relative to persistence and climatology were 28.6% and 17.5%, respectively. It should be noted again that, since a multi-year climatological dataset was not available, the climatology forecasts for the High Winds plant were simply the monthly averages for each hour of the day and do not represent true climatology data.

The annual MAE of the wind speed forecasts was 2.16 m/s, which is 31.9% of the mean wind speed for the 12-month period. The monthly MAE values range from a low of 1.66 m/s during June 2005 to a high of 2.69 m/s during December 2004. The skill score of wind speed forecasts relative to persistence was 27.1%. The lowest monthly skill score was the 17.4% recorded during January 2005, while the highest was 37.3% during August 2004. The skill score vs. climatology was 9.9%, but was not based on a true climatology as noted in the discussion of the power production forecast performance.

Figure 4-88 shows the error distribution for the one-year of power production forecasts. The chart shows two error distributions. One is for the first 24 hours of the forecast period, and the other for the second 24 hours. The error distributions for the two periods are quite

similar. Both are characterized by a strong central peak in the -5% to 0% bin, with a rapid decrease in the frequency on either side of this peak and a long tail of very low frequencies for medium and large negative and positive errors. The sharp peak in the -5% to 0% bin occurs because errors of exactly zero are included in the -5% to 0% bin.

There are a fairly large number of hours for which the forecast and observations are zero during the low output months in the cold season. This results in an error of exactly zero and, thus, there are many more hours that fall into the -5% to 0% bin than the 0% to +5% bin. There is a noticeable asymmetry between the positive and negative sides of the distribution. There is a greater frequency of negative errors (the forecast is too low) of a particular magnitude than there is of positive errors of the same magnitude.

Figure 4-89 illustrates the cumulative error distribution for the 16,756 power production forecast hours that could be verified. This chart shows the frequency of absolute errors that are equal to or smaller than the error shown on the horizontal axis. Separate error distributions are shown for the first and second 24 hours of the forecast period. As noted previously, the distributions for the first and second half of the forecast period are very similar. This indicates that the quality of the forecasts degrades only slightly from the first to the second day of the forecast period. The median forecast error (50% frequency) is approximately 10.6% of the rated capacity for hours 1 to 24, and about 11.7% of capacity for hours 25 to 48. The chart also shows that 75% of the forecast hours during the first 24-hour period yield absolute errors that are less than 24.1% of the rated capacity, while 75% of the forecast hours for the second 24-hour period have absolute errors that are less than 25.4% of capacity.

As noted earlier, there is a strong relationship between the forecasted level of power production and the magnitude of the forecasted error due primarily to the shape of the turbine power curve. Figure 4-90 illustrates this relationship for the 12 months of High Winds power production forecasts. The MAE is below 15% when the forecasted production is 5% of capacity or lower, but the MAE increases immediately to 20% when the forecasted production rises above 5% and remains near 20% in a very broad peak, until the forecasted production rises above 65%. Above 65%, the MAE generally decreases as the forecasted production increases. The MAE is in the 10% to 15% range when the forecasted production is between 80% and 95% and the MAE is well below 5% when the forecasted production is above 95%. This implies that forecasts that anticipate the production will be near full capacity or near zero have a high degree of reliability for the High Winds plant.

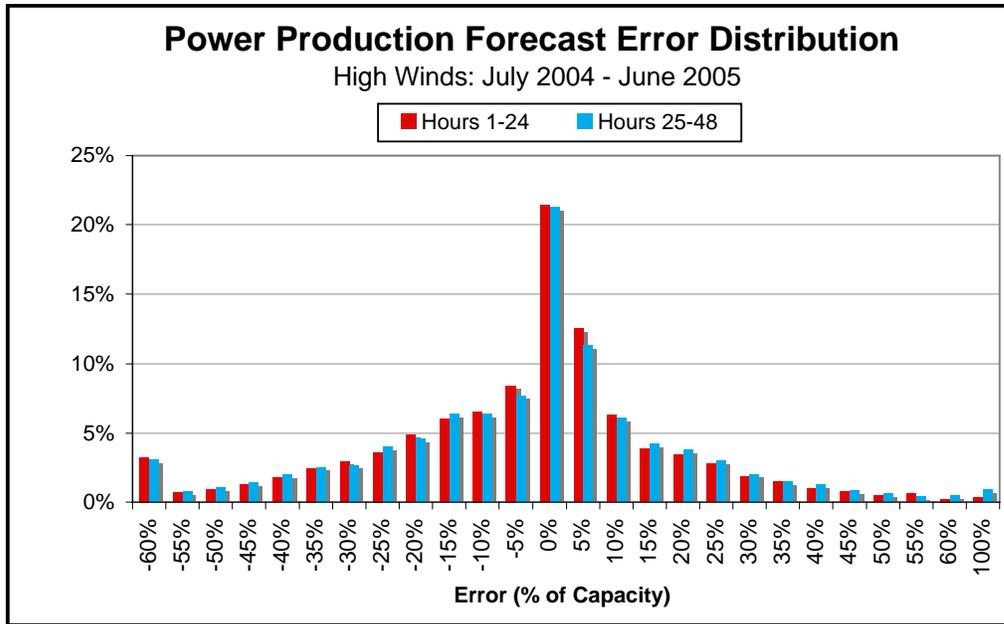


Figure 4-88 Error frequency distribution for hours one to 24 (red) and 25 to 48 (blue) of the forecasts for the High Winds wind plant. The data sample included 16,756 forecast hours during the July 2004 to June 2005 period.

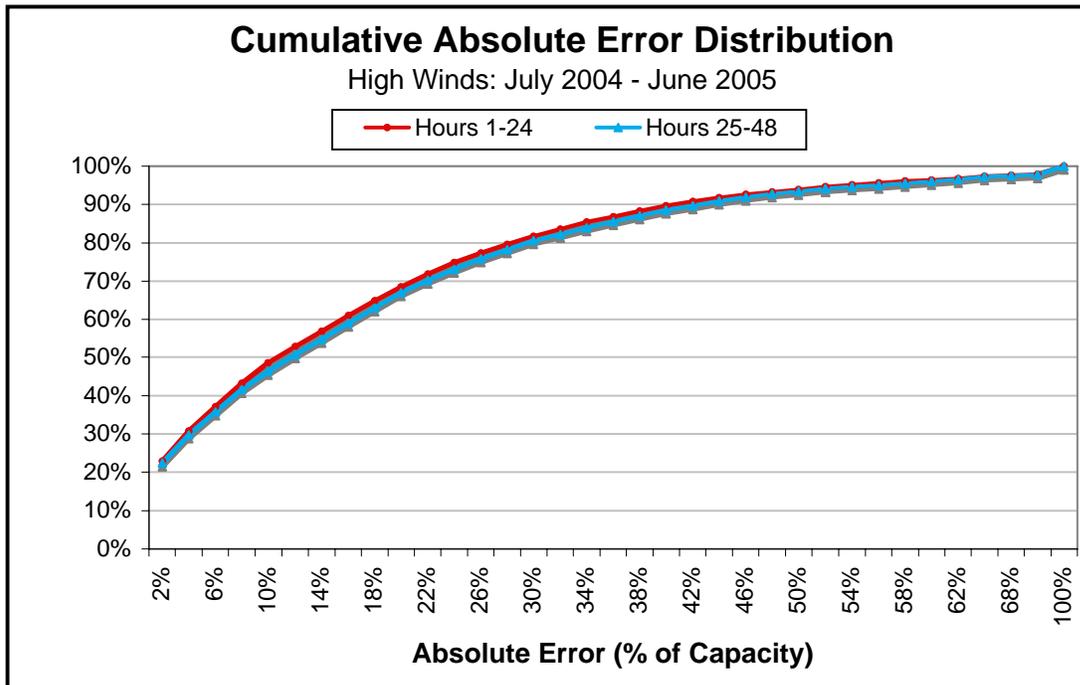


Figure 4-89 Cumulative absolute error frequency distribution (% of hours with an absolute error less than are equal to the value on the horizontal axis) for hours 1 to 48 of the forecasts for the High Winds wind plant. The data sample included 16,756 forecast hours.

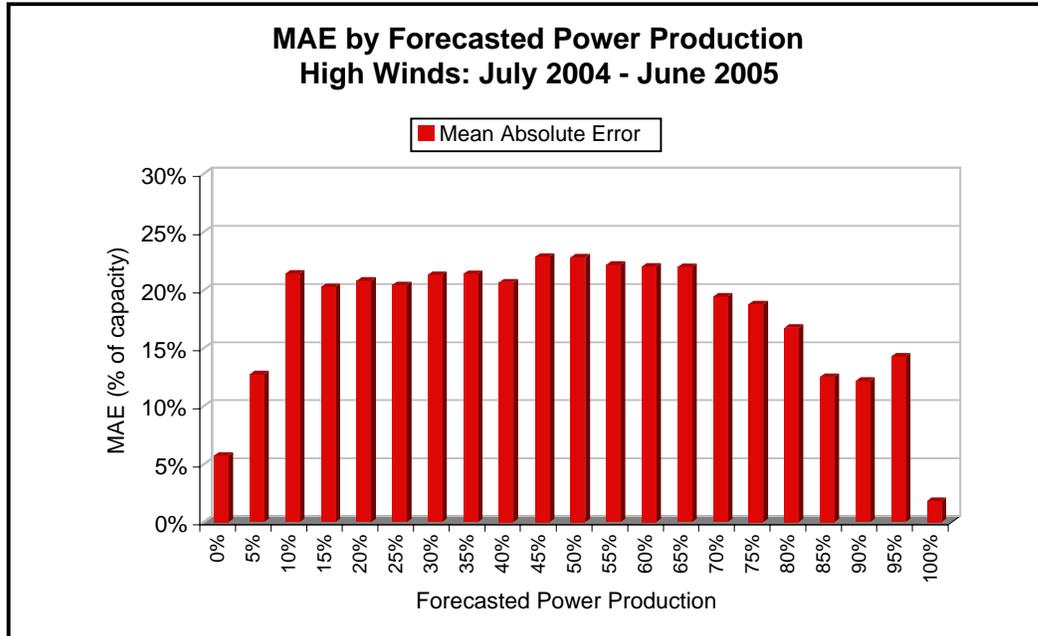


Figure 4-90 MAE of the High Winds 1- to 48-hour power production forecasts vs. the forecasted power production. The MAE is shown for 21 bins of forecasted production. Each has a width of 5% of installed capacity and is labeled by the value at the upper end of the bin. Bars with light shading denote bins with a sample size of less than 30 hours.

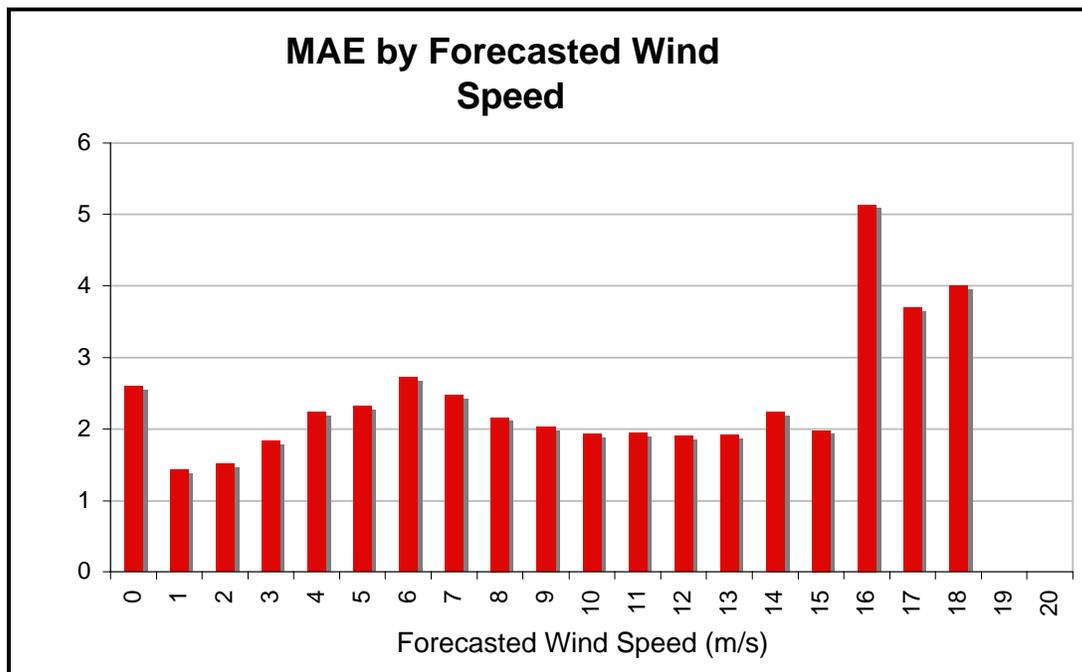


Figure 4-91 MAEs of the 1- to 48-hour High Winds wind speed forecasts vs. the forecasted wind speed. The MAE is shown for 21 bins of forecasted wind speed. Each has a width of 1 m/s and is labeled by the value at the upper end of the bin. Bars with light shading denote bins with a sample size of less than 30 hours.

The magnitude of the wind speed forecast error also has some dependence on the forecasted wind speed, but not as much as the magnitude of the power production forecast error has on the forecasted level of production. Figure 4-91 illustrates this relationship for the High Winds plant. In general, the chart is quite similar to the pattern shown for the SMUD plant in Figure 4-83.

The MAE is generally below 2 m/s for wind speeds under 5 m/s, except for forecasts of exactly zero wind speed, which yield MAEs slightly above 2 m/s. However, this MAE value is based on a very small sample size, since there were only 10 hours with a forecasted speed of exactly zero during the 12-month period. The MAE values rise as the forecasted wind speeds increase and reach a peak near 2.7 m/s for forecasted wind speeds between 5 and 6 m/s. The MAE is generally lower when the forecasted wind speed is above this range. In fact, the MAE is close to 2 m/s throughout the 8 to 15 m/s range of forecasted wind speeds. The wind speed forecast error peak near 6 m/s is towards the lower end of the steeply sloped portion of the power curve where the sensitivity to wind speed forecast error is greatest and thus contributes to the rapid rise in error as the forecasted production rises above 5%.

Forecast Performance vs. Time Horizon

An important consideration of using wind power production forecasts is the rate of degradation of the forecast performance as the forecast time horizon increases the look-ahead period. If the performance of the forecasts decreases rapidly with increasing look-ahead period, then it is useful and perhaps critical to frequently update the forecasts. On the other hand, if the degradation rate is low, then the need for frequent updates is reduced.

This issue is addressed in the analysis of the forecast performance by the MAE vs. forecast time horizon charts for the High Winds power production and wind speed forecasts, shown in Figures 4-92 and 4-93. These charts show the MAE for each look-ahead hour for the MASS-6 SMLR2-ST forecasts (labeled *eWind*) and the persistence and climatology reference forecasts.

There is a gradual upward trend in the MAE of the power production forecasts. The MAE rises from near 15% during the first few hours of the forecast period to the 17% to 19% range towards the end of the forecast period. There is very little evidence of a diurnal pattern in the MAE as there was for some of the other wind plants. A linear fit to the *eWind* MAE curve indicates that the average rate of error growth over the 48-hour period is 0.07% of installed capacity per hour or about 1.7% of installed capacity per day. This is on the high end of range of error growth rates calculated for the other wind plants and is almost exactly the same as the error growth rate calculated for the nearby SMUD wind plant. The MAE of the persistence forecasts rise rapidly from about 7% for forecast hour one to near 15% by forecast hour four. The persistence forecasts outperform the SMLR2-ST forecasts for the first three hours of the forecast period and the MAE values are about the same for forecast hours four through eight. The *eWind* forecasts significantly outperform the persistence forecasts after forecast hour eight. The *eWind* forecasts outperform the climatology forecasts throughout the forecast period although, as is typical, the skill relative

to a climatology forecast (or the gap in the MAE between the *eWind* and the climatology forecast) gradually decreases as the forecast time horizon increases.

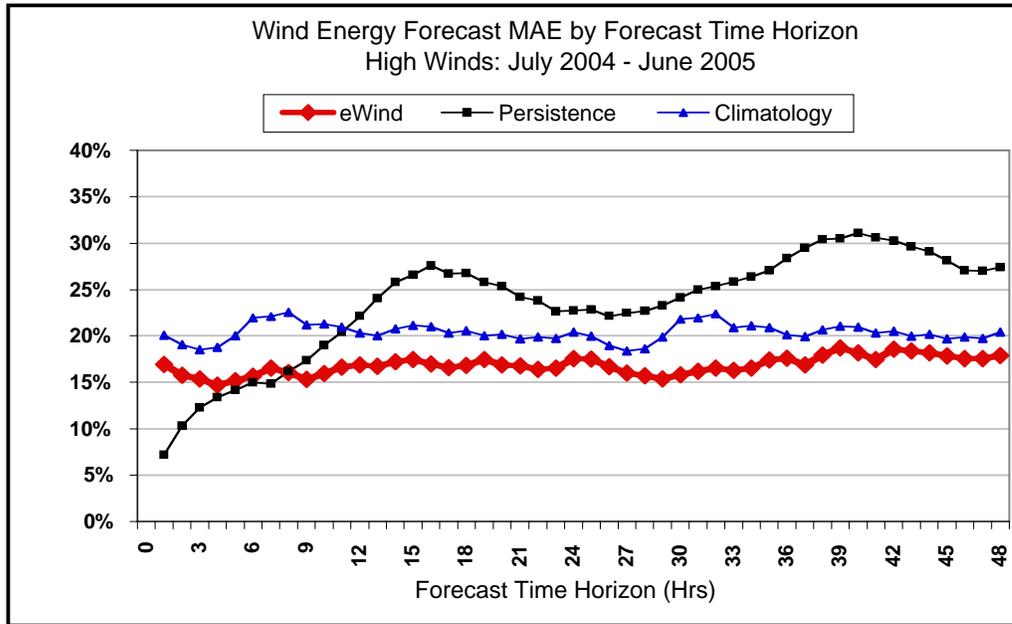


Figure 4-92 MAEs of power production forecasts at High Winds wind plant vs. forecast time horizon for 12 months of *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) forecasts.

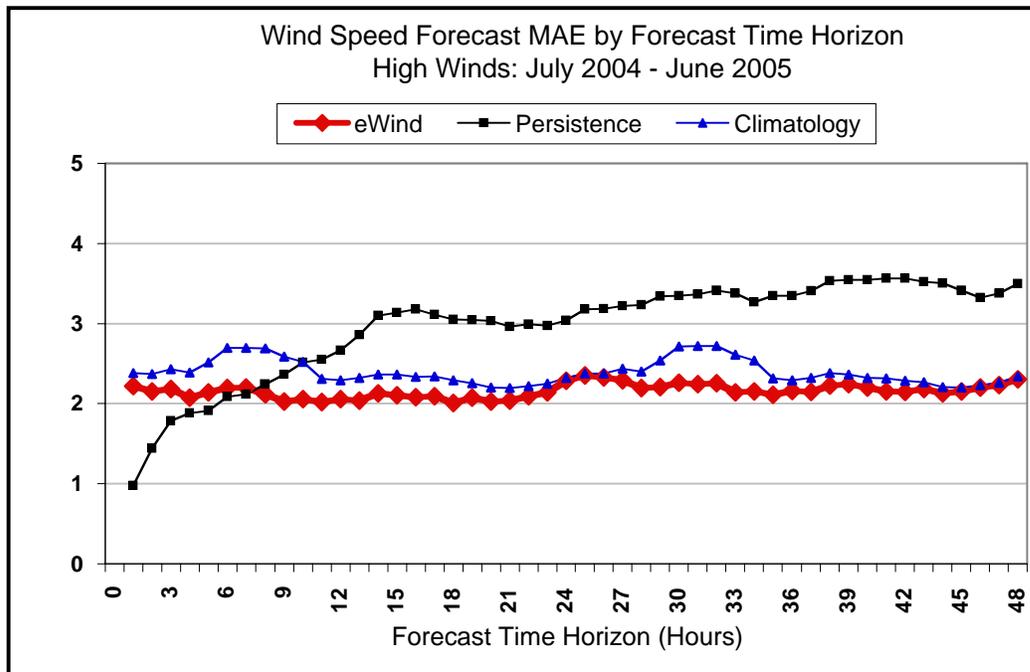


Figure 4-93 MAEs of wind speed forecasts at High Winds wind plant based on *eWind* (red line with diamond markers), persistence (green line with square markers) and climatology (blue line with triangle markers) forecasts.

The MAE of the wind speed forecasts increases very slowly from just over two m/s, during the first 12 hours of the forecast period, to near 2.2 towards the end of the 48-hour period (Figure 4-94). There is no indication of the diurnal cycle in the wind speed forecast MAE pattern that was fairly prominent in the MAE pattern for some of the other wind plants. A linear least squares fit to the *eWind* wind speed forecast MAE data indicates that the average rate of increase in MAE is 0.0024 m/s per hour or about 0.06 m/s per day. The MAE of the persistence forecasts rises rapidly from near one m/s for forecast hour one to near two m/s by forecast hour five. The persistence forecasts yield a lower MAE than the *eWind* forecasts during forecast hours one through four and a similar MAE to the *eWind* forecasts for forecast hours five through eight. The *eWind* forecasts decisively outperform the persistence forecasts after forecast hour eight.

Overall, the MAE vs. forecast time horizon charts for the High Winds plant indicate that the error growth rate for the power production forecasts is similar to that calculated for the nearby SMUD plant and on the high end of the error growth rates for the other wind plants participating in this project. However, the wind speed forecast error growth rate is much lower than that of the SMUD plant and on the low end of the distribution of values for the wind plants in this project.

Physics-Based Model Comparison

A performance comparison among the forecasts generated by the three physics-based models employed in this project was done by analyzing the forecast error statistics for forecasts generated by the same MOS method from the output of all three models.

Table 4-19 presents the forecast performance statistics for SMLR2-ST forecasts from all three models. As noted previously, the comparison of the forecast performance statistics is complicated by the fact that the forecasts from all three models were not generated for the entire one-year forecast period.

The statistics in Table 4-19 indicate that the overall performance of the MASS-6, WRF and COAMPS forecast models was fairly similar, with perhaps a slight advantage for the MASS-6 forecasts. The relative performance of the three models was quite similar to that at the SMUD wind plant. For the four months for which a comparison was possible, the MAE of the MASS-6 power production forecasts was 18.6% vs. 19.0% for the WRF forecasts. As in the case of the SMUD plant, the net difference in overall forecast performance was due to 2% for capacity lower MAE of the MASS-6 vs. WRF forecasts. The forecast performance differences for each of the other three months were less than 0.3% of rated capacity, with MASS 6 having a slight advantage during April and WRF during May and June. Thus, there was virtually no difference in the performance for 3 of the 4 months. The MAEs of the wind speed forecasts from the two models exhibited a similar pattern.

Table 4-19 Monthly MAEs of power production (% of capacity) and wind speed (m/s) for forecasts generated by the three physics-based models for the High Winds wind plant.

Month	Power Production Forecast MAE			Wind Speed Forecast MAE		
	MASS 6	COAMPS	WRF	MASS 6	COAMPS	WRF
Jul-04	17.0%	16.5%		1.75	1.63	
Aug-04	17.6%	19.5%		1.93	1.97	
Sep-04	18.1%	18.8%		2.34	2.50	
Oct-04	18.7%	18.1%		2.54	2.57	
Nov-04	11.5%	12.2%		2.55	2.75	
Dec-04	17.1%	16.0%		2.69	2.87	
Jan-05	8.4%	7.1%		2.12	2.25	
Feb-05	11.7%	11.1%		2.14	2.12	
Mar-05	17.2%	19.1%	19.2%	2.36	2.70	2.68
Apr-05	18.6%	18.8%	18.7%	2.07	2.17	2.07
May-05	20.2%	20.4%	19.9%	2.03	2.07	1.91
Jun-05	18.6%	19.0%	18.4%	1.66	1.70	1.61
Annual	16.2%	16.3%	19.0%	2.16	2.25	2.08
Mar - Jun	18.6%	19.3%	19.0%	2.03	2.16	2.08

For the entire four-month period, the MAEs of the MASS-6 forecasts were 0.05 m/s (2.4%) lower than those of the WRF-based forecasts. However, as in the case of the power production forecasts, this net difference was largely attributable to the fact that the MASS-6 MAE was 0.32 m/s (11.9%) lower than the MAE of the WRF-based forecasts during March. The differences were much smaller during the other three months, with the WRF-based forecasts showing an advantage during May and June.

Unlike the case of the MASS-6 and WRF forecasts, it was possible to compare the performance of the MASS-6 and COAMPS forecasts during the entire 12-month period. The forecast performance of the two models was similar, with a very slight edge to the MASS-6 forecasts. The slight edge of the MASS-6 model is probably slightly larger than indicated by the data in Table 4-19 because, as noted earlier, the COAMPS forecasts ended at 32 hours into the forecast period and the last 16 hours of the forecast period normally yield larger errors than the first 32 hours. However, as noted in the previous section, the rate of error growth with increasing forecast time horizon is fairly low so the impact of this difference in forecast periods is most likely modest.

The MASS-6 power production forecasts yielded a 12-month MAE that was 0.1% of rated capacity (0.6%) lower than the MAE of the COAMPS forecasts. The monthly data indicate that the MASS-6 forecasts yielded a lower MAE during seven of the 12 months. There is no discernable pattern in the relative performance. The MAE differences often switch signs from month to month and there is no suggestion of a seasonal pattern. All of this suggests that there is very little overall difference in performance of the power production forecasts between these two models for the High Winds wind plant.

The performance difference was slightly greater for the wind speed forecasts. The MASS-6 wind speed forecasts yielded a 12-month MAE that was lower by 0.09 m/s (4%), and the

monthly data show that the MASS-6 forecasts yielded the lowest MAE during 10 of the 12 months. MASS-6 also yielded a substantially lower MAE values for the raw wind speed forecasts that were interpolated directly from grid point output data from the two models. For the entire 12-month period, the raw MASS-6 wind speed forecasts yielded an MAE of 2.79 m/s vs. 3.8 m/s for the COAMPS forecasts. This indicates that the COAMPS forecasts received a greater benefit from the MOS procedure than the MASS-6 forecasts.

The comparison of the WRF and COAMPS forecasts was limited to the four-month March through June period. During the four-month period, the MAE of the WRF-based forecasts was 0.3% of capacity lower than that of the COAMPS-based forecasts. The WRF-based forecasts yielded a lower MAE during three of the four months, although none of the differences were greater than 0.5% of capacity.

The WRF-based wind speed forecasts also yielded a lower overall MAE by 0.08 m/s (3.7%). The MAE of the WRF-based forecasts was lower for all four months although, during three of the four months, the difference was 0.1 m/s or less.

Figure 4-94 illustrates the dependence of the relative forecast performance of the three models on the forecast time horizon. As one might expect, the overall pattern is fairly similar for the three models but there are some systematic differences. The COAMPS-based forecasts yield much larger errors during hour three to hour nine of the forecast period. Another observation is that the amplitude of the diurnal MAE pattern was greater for the WRF-based forecasts than for the other two models. The WRF-based forecasts produced a substantially lower MAE during forecast hours 0 to 6 and 24 to 30. This corresponds to the 8:00 a.m. to 2:00 p.m. period of the first and second days of the forecast period. However, the WRF-based forecasts produced larger errors for the period extending from late afternoon to late evening, especially on the second day of the forecast pattern. There is potential to improve forecast performance by exploiting patterns such as those described above via an intelligent ensemble forecast that weights forecasts that are likely to perform better during specific time windows more heavily.

In summary, the overall forecast performance by the three physics-based models for the High Winds wind plant was fairly similar. The performances statistics suggested that the forecasts based on the MASS-6 simulations had a slight edge in performance for both power

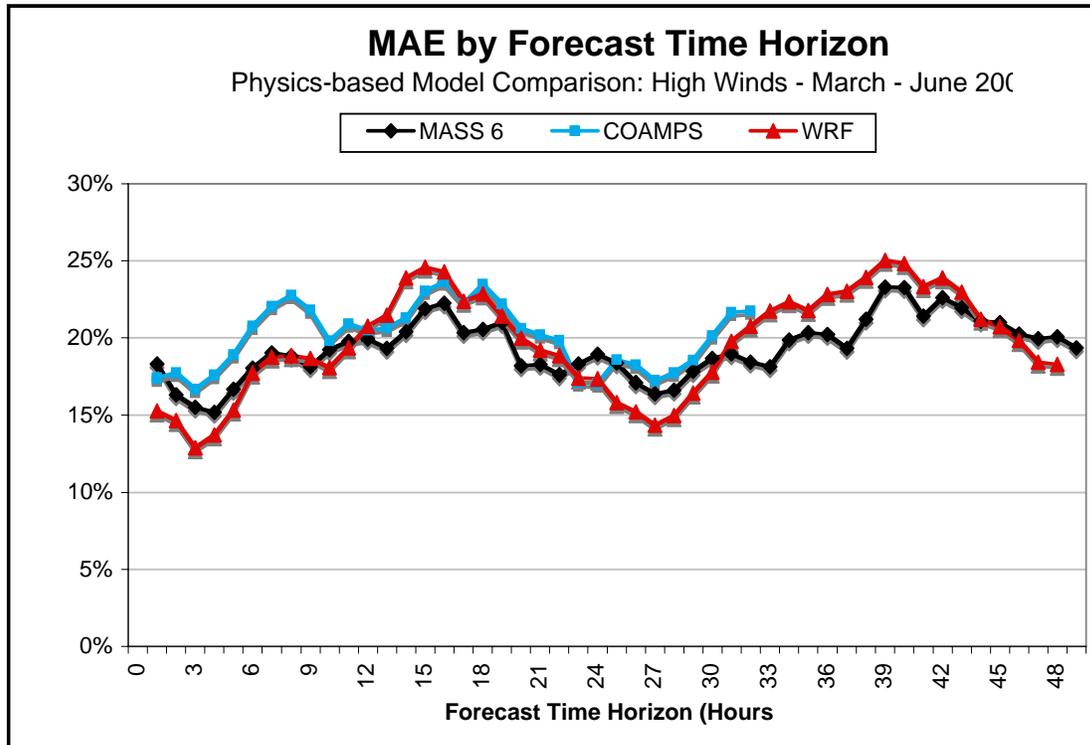


Figure 4-94 MAEs of power production forecasts for the High Winds wind plant based on the MASS-6, COAMPS and WRF physics-based models vs. forecast time horizon the look-ahead interval) for four months (March to June 2005).

production and wind speed forecasts, and that the WRF-based forecasts had a slight edge over the COAMPS-based forecasts for the four-month period.

Forecast Ensembles

An ensemble-mean forecast of the power production and the wind speed was generated from the 12 different forecasts produced for the High Winds wind plant for the four months (March 2005 through June 2005) for which all 12 forecasts were available. The 12 ensemble members consisted of each combination of the four different MOS methods and the three physics-based models.

Table 4-20 presents the MAEs of the power production and wind speed forecasts for several combinations of the 12 ensemble members as well as the ensemble-mean forecast.

For the power production forecasts, the MAE of the ensemble-mean forecast was 0.1% of capacity (0.5%) higher than the MAE of the best performing individual forecast (MASS-6 SMLR2-ST) during the entire four-month period and 0.8% of capacity (4.5%) higher than the composite of the best performing forecast for each month. The best overall method yielded a lower MAE than the ensemble-mean during three of the four months. However, the ensemble-mean yielded a substantially lower MAE (1% of capacity) during the one month that it did achieve a lower MAE. This was the largest absolute difference between

the best overall method and the ensemble mean among the four months for which a comparison could be made.

The best monthly method produced a lower power production forecast MAE than the ensemble-mean method during all four of the comparison months. Thus, the performance of the ensemble-mean forecast for the High Winds wind plant was not quite as good as it was for some of the other wind plants. However, the MAE of the ensemble-mean was 0.9% of capacity lower than the average MAE of all 12 of the individual forecasts.

Table 4-20 Monthly MAE for the March to June 2005 period for ensemble forecasts of the power production (% of capacity) and wind speed (m/s) for the High Winds wind plant.

Month	Power Production Forecast MAE				Wind Speed Forecast MAE			
	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs
Mar-05	18.8%	18.6%	17.9%	19.1%	2.47	2.33	2.33	2.63
Apr-05	18.0%	17.1%	17.1%	18.7%	1.96	2.06	2.01	2.12
May-05	18.5%	18.3%	18.3%	20.2%	1.80	2.01	1.85	2.01
Jun-05	18.4%	19.4%	17.1%	19.1%	1.56	1.69	1.61	1.68
4-months	18.4%	18.3%	17.6%	19.3%	1.95	2.02	1.95	2.11

Note: The ensemble forecasts include: (1) averaging all the forecasts in a 12-member ensemble ("Ensemble-12 Mean"); (2) the individual method with the lowest MAE for the four -month period ("Best Overall Method"); (3) a composite of the individual methods with lowest MAE in each month ("Best Monthly Method"); and (4) the mean MAE of all 12 ensemble members ("Average of MAEs").

The MAE of the ensemble-mean wind speed forecasts for the four-month period was 0.07 m/s (3.5%) lower than the MAE of the best performing individual wind speed forecast method. In addition, the ensemble-mean forecast yielded a lower MAE during three of the four comparison months. During those months the ensemble-mean method yielded skill scores of 4.8%, 5.5% and 7.7% with respect to the best overall method. In addition, the MAE of the ensemble-mean forecast was also the same (1.95 m/s) as that of the composite forecast created by using the best performing forecast for each month. However, the ensemble-mean method yielded a lower wind speed forecast MAE during three of the four comparison months. The ensemble-mean forecast was 0.16 m/s lower than the average MAE of all of ensemble members.

In summary, the ensemble-mean forecast method yielded a skill score of -0.5% relative to the best performing individual power production forecast method, and a skill score of 3.5% relative to the best performing wind speed forecast method during this four-month period.

Thus, the ensemble-mean was clearly the best choice for a wind speed forecast as it clearly outperformed all 12 ensemble members. However, that was not the case for the power production forecasts since the best performing individual method slightly outperformed the ensemble-mean method. In addition, the composite of the best method for each month had a substantial advantage over the ensemble-mean. There is currently no method to reliably determine in advance which method will perform best in a particular month. Therefore, this

approach cannot currently be used in practice. However, the results indicate that further forecast performance improvements on the order of 5% to 10% might be achieved in some cases if an effective technique to select the best performing forecast for a given day, week, month or other time period could be formulated.

Summary

The next-day wind energy forecasting algorithm research and testing was performed in two phases. The first phase conducted a series of experiments to screen potential methods to improve forecast performance, as measured by the reduction in forecast error relative to that in the previous Energy Commission-EPRI wind energy forecasting project (Energy Commission and EPRI, 2003). The second phase applied and tested the most promising methods at five wind projects in California for periods ranging from several months to a full year, depending on data availability. In both phases, the forecast performance was measured by the mean errors, mean absolute errors, and the skill scores of the 48-hour forecasts of hourly wind speed and energy generation vs. the observed data. The results of each phase are summarized below.

Screening of Improved Data and Forecast Methodologies

The improved data and forecast methodology screening experiments were designed to test potential methods of reducing forecast error and were chosen according to two criteria: (1) those that had the greatest promise of improving the forecasts; and (2) those that could be tested with the data available for this project and within the resource limitations of the project.

Table 4-21 describes the six potential enhancements of the forecast algorithm tested in experiments and summarizes the impact of the forecast system modifications within each focus area on the performance of the wind power production forecasts.

The MAE reductions in Table 4-21 represent the combined net reductions for each of the forecast system modifications at both the Mountain View and PowerWorks wind plants during a six-month test period relative to the MAEs in the previous Energy Commission-EPRI wind forecasting evaluation (EPRI-Energy Commission PIER 2003). For a variety of reasons, the baseline method are not the same for each focus area, as noted in Table 4-21.

The most significant reductions of the mean absolute error (MAE) resulted from Focus Areas 1, 3, 4, and 5. In order of decreasing improvement of MAE, the four modifications are:

- Stratified 2-stage Screening Multiple Linear Regression (SMLR) Model Output Statistics (MOS) method (15.8%)
- MODIS and Pathfinder Water Surface Temperature (WST) data (12.3%)
- Plant scale power curve based on deviations from power curve (4.6%).

Table 4-21 Summary of the net reduction in the power production forecast MAE associated with the forecast system modifications in each focus area.

Focus Area	Forecast System Modification	MAE Reduction (%)
1	MODIS and Pathfinder WST data (4 km) BASELINE: NCEP OI WST (110 km)	12.3%
2	1 km physics-based model grid BASELINE: 10 km physics-based model grid	-4.7%
3	WRF with 40 km grid as the physics-based model BASELINE: MASS with 40 km grid	4.0%
4	Stratified 2-stage SMLR scheme BASELINE: Screening Multiple Linear Regression	15.8%
5	Model deviations from power curve BASELINE: Speed-based plant-scale power curve	4.6%
6	Mean of an ensemble of forecasts BASELINE: "Best" single forecast method	0.8%

- Physics-based WRF atmospheric model using the 40-km grid (4.0%).

Focus Area 2, using a higher resolution grid (smaller grid cells) for the physics-based model, did not produce a net improvement in forecast performance for the six test months. Indeed, the MAE of the power production forecast *increased* by 4.7% when the grid size was reduced to 1 km from the 10 km used in the previous Energy Commission-EPRI project.

However, there is more to the story. Before applying the Model Output Statistics (MOS) adjustments, the MAEs of the raw unadjusted forecasts substantially decreased by reducing the grid size from 10 km to 1 km. But after applying the MOS adjustments, the MOS-adjusted forecasts based on the 10-km and 1-km physics-based models were very similar. In effect, the MOS procedure compensated for the higher error of the raw 10-km vs. the 1-km physics-based models and tended to equalize the performance of the 10-km and 1-km grids. Thus there was no net benefit realized from reducing the grid size.

Focus Area 6, using an ensemble-mean forecast in place of the best individual method, yielded an MAE reduction of only 0.8%. However, the performance of the ensemble-mean approach was most likely reduced by limiting the number of ensemble members to two (a MASS-based a WRF-based forecast) in this screening phase experiment. Other investigations (for non-wind energy applications) reported in the meteorological literature indicate that a fairly large number of ensemble members is needed to realize a more substantial potential impact of the ensemble-mean approach. The second phase of this task conducted a more meaningful test of the ensemble approach, in which 12 different forecast methods were employed for a portion of the one-year forecast evaluation period for five wind plants.

It should also be noted that the MAE reductions listed in Table 4-21 are not likely to be additive for at least two reasons. First, the MAE changes associated with the forecast system modifications are not likely to be highly correlated with each other. Therefore, in

general, they will frequently offset each other. Second, some of modifications may provide alternate ways to eliminate the same systematic error. For example, systematic errors associated with mis-specifying water surface temperatures may be reduced by using an improved water surface temperature database or perhaps by using a more sophisticated MOS formulation that can statistically correct physics-based model errors attributable to the poor WST data.

Forecast Evaluation at Five California Wind Projects

Hourly power production and wind speed forecasts were generated twice daily for a 48-hour look-ahead period and for five wind plants in California over the 12-month period, July 1, 2004 to June 30, 2005. The wind plants were: (1) the 66.6 MW Mountain View I and II wind plant in the San Geronio Pass; (2) the 34.5 MW Oak Creek wind plant in Tehachapi Pass; (3) the 90 MW PowerWorks wind plant in Altamont Pass; (4) the 15.18 MW SMUD wind plant in Solano County; and (5) the 162 MW High Winds wind plant also in Solano County.

Forecasting Protocol

The forecasts generated predictions of the hourly power production, wind speed and wind direction for a 48-hour period beginning at 8:00 a.m. PST and 8:00 p.m. PST each day. Three different physics-based models and four different Model Output Statistics (MOS) procedures were used to produce an ensemble of 12 different forecasts for each of the wind plants. The forecasts were generated in a mixture of real-time and historical modes although all forecasts only used data that was or would have been available to the forecast system in an operational real-time mode.

To permit a more detailed analysis of the forecast performance a comprehensive analysis of the forecast performance was conducted only for the forecasts generated during the morning forecast cycle the forecasts delivered at 8:00 a.m. each day). A review of the performance statistics indicates that the differences in forecast performance between the morning and afternoon cycles were not significant and that almost all of the conclusions from the analysis of the performance of the morning forecast cycle would apply to the combined pool of both the afternoon and morning forecast cycles.

Overall Forecast Performance Results

The annual mean absolute errors (MAEs) and skill scores of the forecast for all 48 forecast hours for both the power production and wind speed forecasts for each of the wind plants are presented in Table 4-22. In each case, the results are for the forecast method that gave the lowest forecast error over the entire one-year period.

One's impression of the forecast performance is tied somewhat to which performance parameters are considered. The average MAE as a percentage of a plant's installed capacity for all five wind plants was 14.5%. It ranged from a low of 11.9% for the

PowerWorks plant to a high of 16.8% for the High Winds plant. However, the average MAE as a percentage of the actual plant production was 52.7% and ranged from a low of 36.6% for the Mountain View plant to a high of 60.8% for the SMUD plant. The average skill score relative to a persistence forecast was 33.2% and ranged from a high of 40.5% for the Mountain View plant to a low of 26.5% for the PowerWorks plant. The average skill

Table 4-22 Mean absolute error of the best performing individual power production and wind speed forecasts for the entire 48-hour forecast period for each of the five participating wind plants.

Site	Power Production Forecast MAE				Wind Speed Forecast MAE			
	MAE % of Capacity	MAE % of Prod	Skill vs Persistence	Skill vs Climatology	MAE m/s	MAE % of Speed	Skill vs Persistence	Skill vs Climatology
Mountain View	13.0%	36.6%	40.5%	47.7%	2.65	27.5%	40.1%	27.0%
Oak Creek	15.0%	57.1%	33.2%	27.1%	2.03	40.6%	32.8%	43.8%
PowerWorks	11.9%	58.5%	26.5%	38.9%	2.52	34.4%	28.2%	12.4%
SMUD	16.0%	60.8%	37.1%	16.1%	1.98	35.9%	31.3%	10.9%
High Winds	16.8%	50.7%	28.6%	17.5%	2.16	31.9%	27.1%	9.9%
Overall	14.5%	52.7%	33.2%	29.5%	2.27	34.1%	31.9%	20.8%

score with respect to a climatology forecast was 29.5% and ranged from 47.7% for the Mountain View plant to 16.1% for the SMUD wind plant.

For the Mountain View wind plant, the wind energy forecast MAE decreased from 16.6% of rated capacity in the previous Energy Commission-EPRI project to 13.0%, and the wind speed MAE for the Catellus Tower decreased from 3.05 m/s to 2.65 m/s. In addition, the skill score vs. persistence increased from 37.5% to 40.5% and the skill score vs. climatology increased from 36.4% to 47.7%.

For the PowerWorks wind plant, the wind energy forecast MAE decreased from 14.1% of rated capacity to 11.9%, while the wind speed forecast MAE for PowerWorks Tower M438 decreased from 1.93 m/s to 1.78 m/s. In addition, the skill score of the power production forecasts vs. climatology increased from 30.9% to 38.9%. These statistics indicate that the forecast system enhancements substantially improved forecast performance relative to the previous Energy Commission-EPRI project.

The average MAE of the power production forecasts for all five participating wind plants was 14.5% of installed capacity and 52.7% of the average production. The skill scores of the power production forecasts were 33.2% vs. persistence and 29.5% vs. climatology.

It should be noted that the skill scores relative to climatology are most likely too low. This is because no climatological data were available for three of the wind plants and the actual monthly mean wind energy generation for each hour of the day was assumed for the climatology value instead of independently-collected climatology data. Therefore, the forecast error of the resulting climatology forecast is lower than would be expected if actual climatology data had been used. The average mean absolute error of the wind speed forecast was 2.27 m/s, which is 34.1% of the average wind speed.

Variability of Power Generation Forecast Error

The forecast performance results raise a number of issues and questions about the nature of the variability in forecast performance. For example, the annual power-generation forecast MAE exhibits considerable variability between the wind plants, ranging from 11.9% of rated capacity at the PowerWorks plant to 16.8% at the High Winds plant.

Furthermore, the performance statistics indicate that the wind speed forecast MAEs are not highly correlated with the wind power production forecast MAEs. For example, the SMUD wind plant had the lowest wind speed MAE but the second highest power production MAE, while the Mountain View wind plant had the highest wind speed MAE but the second lowest wind energy forecast MAE. Clearly, the magnitude of the wind speed MAE is not the controlling factor for the wind power production MAE.

An obvious question is: what factors are responsible for this variability? It is clearly of critical importance to understand the factors that contribute to forecast performance variability in order to determine the direction of future efforts to improve forecast performance.

Therefore, a detailed analysis of the forecast errors for three wind plants was conducted. By eliminating the differences in wind speed errors between the wind plants, it was possible to gain considerable insight. This was done by assuming a constant two meter/sec wind speed error, randomly distributed between positive and negative deviations from the observed values for all hours of the year. The resulting wind energy forecast MAEs were 8.8% of rated capacity at the Mountain View 1 and 2 wind plant and 13.1% and 14.7% at the High Winds and SMUD wind plants, respectively. Thus, even with virtually identical overall wind speed error magnitudes, the wind power production MAEs still vary significantly between the wind plants.

The principal reason for the energy-forecast MAE variation between plants appears to be related to differences in the maximum slopes of the plant-scale power curves and the wind speed frequency distributions between the plants.

Figure 4-71 shows empirical plant-scale power curves for each of the three wind plants. The general shapes of the curves imply that the sensitivity of wind energy forecast error to wind speed forecast error varies with wind speed. When the wind speed is in the range corresponding to the steeply-sloped middle section of the plant-scale power curve, the wind energy forecast error should be most-sensitive to wind speed forecast errors. Thus, the maximum wind energy forecast errors should occur at sites where the wind speed spends many hours in the middle range of the power curve.

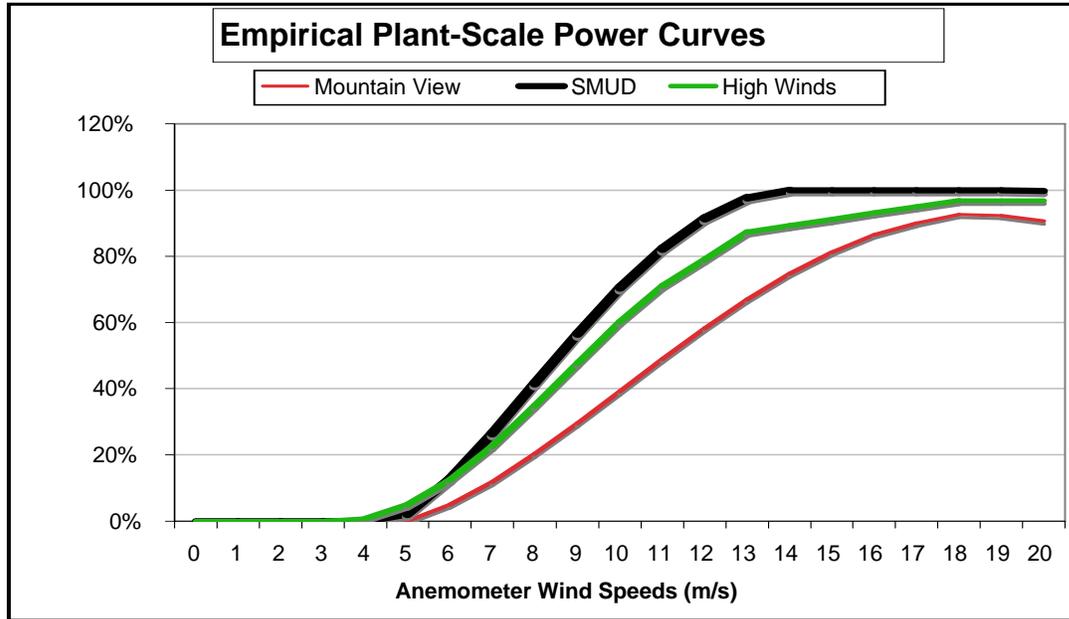


Figure 4-95 Empirical median plant-scale power curves derived from measured wind energy generation and wind speed data for three of the participating wind plants

However, the maximum slopes of the power curves also vary between the wind plants, and plants with higher maximum slopes should also exhibit higher wind energy forecast errors. The slopes are mostly determined by the correlation between the wind speeds at the individual wind turbine locations. This was verified by the “constant 2 m/s error experiment”. The SMUD wind plant with the steepest-slope power curve also exhibited the highest wind energy forecast MAE when a constant wind speed forecast error of 2 m/s was assumed. The Mountain View plant with the lowest maximum power curve slope exhibited the lowest energy forecast MAE.

Ensemble Forecasts

Applications and testing of ensembles of individual forecast methods have demonstrated improved forecast performance in other meteorological forecast applications. To assess the potential of ensemble forecasting of wind power production, a composite forecast was generated from the ensemble of the 12 individual forecasts that were generated in this project. The 12 individual forecasts consisted of each combination of four MOS procedures and three physics-based models. The full suite of methods was only available for the four-month period, March 2005 through June 2005, so the evaluation of the ensemble approach was limited to that period.

Table 4-24 summarizes the wind speed and energy forecast MAEs for four combinations of power generation and wind speed forecasts from the 12-member ensemble. The first columns of the power production and wind speed sections list the MAEs of the ensemble-mean forecasts. The forecasts were constructed by calculating the average of all available members of the 12-member ensemble for each hour of every forecast cycle.

The next columns to the right (labeled “Best Overall Method”) list the MAEs of the method that yielded the lowest MAE over the entire four-month period for each wind plant. This is the method that would typically be used in an operational forecast environment, since it would most likely be classified as the best method when reviewing performance statistics compiled over a long period.

Table 4-23 Mean absolute errors of ensemble power production (% of rated capacity) and wind speed (m/s) forecasts for the five participating wind plants

Month	Power Production Forecast MAE				Wind Speed Forecast MAE			
	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs	Ensemble-12 Mean	Best Overall Method	Best Monthly Method	Average of MAEs
Mountain View	16.4%	16.0%	15.7%	17.9%	2.62	2.66	2.63	2.89
Oak Creek	18.9%	17.3%	17.3%	20.7%	2.07	2.14	2.05	2.25
PowerWorks	13.9%	14.1%	13.2%	14.5%	2.41	2.48	2.38	2.57
SMUD	17.1%	17.4%	17.3%	18.6%	1.82	1.87	1.80	1.94
HighWinds	18.4%	18.3%	17.6%	19.3%	1.95	2.02	1.95	2.11
4-months	17.0%	16.6%	16.2%	18.2%	2.17	2.24	2.16	2.35

Notes: The ensemble forecasts include: (1) averaging all the forecasts in a 12-member ensemble (“Ensemble-12 Mean”), (2) the individual method with the lowest MAE for the 4-month period (“Best Overall Method”), (3) a composite of the individual methods with lowest MAE in each month (“Best Monthly Method”); and (4) the average MAE of all 12 members of the ensemble (“Average of MAEs”).

The next columns to the right, labeled “Best Monthly Method,” list the composite MAEs of the forecasts created by using the methods with the lowest MAEs for each of the four months. It would be difficult to use this approach in an operational environment, since one would have to identify which method was going to perform best during a particular month before the month began. However, it may be possible to develop an “intelligent ensemble composite” that varies the weight placed on different members of the ensemble in the ensemble-composite forecast as a function of parameters which indicate which ensemble members are likely to perform better for a particular forecast cycle. This is a possible area for future research.

The rightmost columns, labeled “Average of MAEs,” represent the average MAEs of all 12 ensemble members. These are the forecast MAEs that are most likely if one had no knowledge about the relative performance of the forecast methods and randomly selected a forecast method each day.

The results indicate that the performance of the ensemble mean forecasts of wind speed was substantially better than that of power production.

For the wind speed forecasts, the ensemble-mean method clearly outperformed the best overall method. The five-plant average MAE for the ensemble-mean wind speed forecasts was 3.1% lower than the composite MAE of the best overall forecast for each wind plant. In fact, the MAE of the ensemble-mean was only very slightly higher than the MAE of the “Best Monthly Forecast”. Furthermore, the ensemble-mean forecast had a lower MAE than the best overall method for all five of the wind plants. The ensemble-mean forecast also had an MAE that was almost 8% lower than that of the average MAE of all of the ensemble members.

However, for the power production forecasts, ensemble mean method had a slightly higher four-month average MAE (17.0%) than the composite of the best overall method for each plant (16.6%). Much of the difference is attributable to very poor forecast performance at the Oak Creek plant, because one plant output model significantly outperformed all others, due to unusual systematic variations in the plant-scale power curve. Excluding the Oak Creek plant, the performance of the ensemble-mean power production forecast was about the same as that of the best overall method for each plant. This is still not as good as the performance of the ensemble-mean for the wind speed forecasts. The reason for this is not clear but may be related to the fact that there was less diversity in plant output models than in other parts of the system. This will require further investigation.

These limited tests indicate that the ensemble approach appears to have at least modest potential to improve wind speed forecast performance.

The simple approach of constructing an average wind-speed forecast from all ensemble members could reduce wind speed MAE by 3% to 5% relative to the best overall forecast method. The improvement could increase further if an intelligent ensemble composite can be constructed by weighting certain ensemble members more heavily when key parameters indicate they are more likely to produce better forecasts.

For power production forecasts, further refinements of the ensemble forecasting approach are needed. One possibility is to explore ways in which the uncertainty in the plant-scale power curve relationship can be addressed by the forecast ensemble. Another approach is to apply adjustments to the plant-scale power curve to account for variations in atmospheric stability. One way to do this would be to use rapid-update, high-resolution simulations of the wind flow field around the wind turbines to estimate deviations from a reference plant-scale power curve based on atmospheric stability and other parameters. The next section discusses an initial investigation of this concept.

Conclusions

The results of the screening-phase next-day forecast experiments indicated that:

1. The largest positive impacts on forecast performance occurred using improved higher-resolution water surface temperature values as input to the physics-based model and a more sophisticated statistical procedure in Model Output Statistics (MOS) component of the forecast system;
2. Using a higher resolution grid for the physics-based model generally improved the performance of the raw physics-based model forecast, but did not significantly improve the performance of forecasts after the MOS procedure was applied;
3. The use of a next generation physics-based model (the WRF) and a more sophisticated plant output model yielded modest improvements in forecast performance in the experimental sample, but the significance of these improvements for a larger sample was questionable; and

4. Forecasts generated by computing the mean from an ensemble of two forecasts from different physics-based models produced an insignificant improvement in performance, but this disappointing performance may have been attributable to the small size and limited diversity of the ensemble in the screening phase of the project.

The results of the one-year forecast evaluation experiment at five wind projects indicated that:

1. The composite annual mean absolute error (MAE) of the 48-hour power production forecasts was 14.5% of installed capacity and 52.7% of average production for the entire one-year period and the five participating wind plants;
2. The composite annual MAE of the 48-hour wind speed forecasts was 2.27 m/s or 34.1% of the mean wind speed for the one-year period at the meteorological tower sites within each wind plant;
3. There was a considerable improvement in forecast performance between the forecast evaluation conducted in the previous (2001-02 period) and current (2004-05 period) projects for the two wind projects that participated in both projects; the MAE of the annual power production forecast MAE decreased from 16.6% to 13.0% at the Mountain View 1 and 2 plant at San Geronio, and from 14.1% to 11.9% at the PowerWorks plant at Altamont Pass;
4. Ensemble forecasting yielded a 2% to 5% reduction in the MAEs of the wind speed forecasts, but there was no significant improvement in the MAE of the power production forecasts.
5. The annual power production forecast MAEs varied significantly among the wind plants, ranging from 11.9% to 16.8%;
6. The annual power production MAE was not well-correlated with the wind speed forecast MAE, with low power production MAEs often occurring with high wind speed MAEs, and vice versa; and
7. A substantial portion of the power production forecast MAE variability between wind plants is due to differences between the wind speed frequency distributions and the maximum slopes of the wind plant-scale power curves.
8. In addition, the high-resolution wind flow simulations described in Section 5 indicate that atmospheric stability in the surface boundary layer affects the relationship of wind speeds and power generation between individual wind turbines and thus causes temporal variations in the plant-scale power curve. It may be possible to improve forecast performance by modeling the impact of anticipated changes in atmospheric stability on the plant-scale power curve.

Recommendations

Although significant progress was made in the improvement of the accuracy of power production forecasts and the understanding of the characteristics of forecast errors in this project, there are still many promising paths to pursue to further improve day-ahead power production forecast performance.

Ideally, the development of power production forecasting technology should be viewed as an ongoing cyclic process. The first step in each development cycle is a forecast evaluation experiment for a set of wind plants similar to the one performed in the second phase of this task. Next the data from this evaluation experiment should be thoroughly analyzed to understand the error ranges and characteristics of the current state-of-the-art technology. The combination of an understanding of the error characteristics of the current state-of-the-art forecast systems and awareness of new data or modeling technology will enable modifications to the forecast system that can be tested in the evaluation phase of a new development cycle.

At present, the most promising opportunity for further improvement of forecast performance is the emergence of higher-resolution and more accurate measurements of atmospheric and ground and water surface variables by satellite-based and ground-based remote sensing systems. The challenge will be to effectively use the high volumes of data produced by these systems to obtain the maximum possible improvement in the performance of power production forecasts on the day-ahead or other look-ahead time scales. It is recommended that the next phase of wind energy forecasting research continue in California focus developing techniques to effectively utilize the next generation of remotely-sensed data.

5

NUMERICAL MODELING OF WIND FLOW AND WIND PLANT POWER CURVES OVER COMPLEX TERRAIN

Introduction

A significant source of uncertainty and error in wind power production forecasts is attributable to the scatter in the relationship between the average wind speed over a prescribed time interval (for example, an hour) measured at one or more meteorological towers within a wind plant and the average plant power production over that interval. The scatter in this relationship means that even if a perfect forecast is made for the wind speed and direction at the meteorological tower(s), there will still be a substantial error in the power production forecasts.

Typically, the mean absolute error of the power production forecasts for a wind plant given a perfect wind forecast for the plant's meteorological towers is about 5% of a plant's installed capacity. This is a substantial portion of the typical mean absolute error of state-of-the-art next-day power production forecasts of about 15% to 20% of installed capacity. Some of the power production error associated with the perfect meteorological tower wind speed forecast is associated with non-meteorological factors such as variations in turbine availability and turbine performance. However, most of this error is attributable to the variation in wind speed and direction within the plant's domain, which causes the wind speed experienced by each turbine to be different from that measured at the meteorological tower.

The objective of the numerical modeling of wind flow over complex terrain task is to improve the accuracy of plant-scale wind power production forecasts by constructing generic (not case specific) relationships between the wind speed and direction at a plant's meteorological tower and the wind speed and direction at each turbine location within the farm based on data from very high resolution physics-based numerical simulations of the wind flow within and in the vicinity of the wind plant for a representative sample of cases. The vision is that once relationships have been constructed, they will be used to estimate the wind speed at each turbine location from the measured or forecasted speed at the meteorological tower. The wind speed at each turbine location can then be used to calculate the power production of the turbine through the use of the manufacturer's power curve. Ultimately, the calculated power production for each turbine can be aggregated to yield a power production estimate for the entire wind plant.

The approach that is typically used to construct a prediction of a plant's power production from a forecast of the wind speed at a plant's meteorological tower is the empirical plant-scale power curve method. In this approach, a plant-scale power curve is created by statistically fitting a

curve to a sample of wind speed measurements from a meteorological tower within the wind plant and the corresponding measured power production values. The resulting plant-scale power curve is then used to calculate the power production from a forecasted wind speed for the meteorological tower site.

Figure 5-1 shows an example of such an empirical plant-scale power curve and a sample of hourly wind speed and power production data. This example is for PowerWorks, Inc. meteorological tower M127 in the Altamont Pass and a cluster of turbines in the vicinity of this tower. The black markers depict the hourly data from July 2002, and the red line depicts the predicted power production for the same set of July 2002 hours based on an empirical power curve constructed from June 2002 power production data and the wind speed measurements at Tower M127 (or a perfect forecast of the hourly wind speeds at M127).

Figure 5-1 indicates that there is a fair amount of scatter in the hourly data relative to values predicted by the empirical power curve. This scatter represents error in the power production forecast. As noted previously, much of this scatter is attributable to the fact that winds near the ground can be highly variable with time and over small distances, particularly in complex terrain such as in the Altamont Pass. Therefore, the actual wind speeds experienced simultaneously at individual turbines in the cluster may differ significantly from each other and that measured by the anemometer on the meteorological tower. The plant-scale power curve implicitly accounts for the average relationship between the wind speeds at the meteorological tower and those at the turbine sites, but power production forecast error will be caused by situations in which the distribution of wind speed over the cluster departs from the average conditions represented by the plant-scale power curve. The challenge is to construct a more accurate model of the relationship between the wind speed measurements at the meteorological tower and the speeds at the individual turbines.

Researchers at the University of California at Davis (UC Davis) used a boundary layer wind tunnel to model the airflow over the cluster of wind turbines associated with Tower M127. They calculated the ratio of wind speeds at each turbine to the speed at the meteorological tower by measuring the simulated steady-state wind speed in the boundary layer wind tunnel at the meteorological tower and each of the 87 turbine locations. The UC Davis group measured the wind speed ratios for four different wind directions by rotating the terrain model within the wind tunnel.

In this task, the wind speed ratios will be inferred from a set of high-resolution, physics-based, atmospheric numerical model simulations for the same conditions simulated in the boundary layer wind tunnel. The question is whether a numerical model can correctly simulate wind speed differences between individual turbines, and whether the simulated differences calculated from a few detailed simulations can improve the prediction of the plant power output over a longer period of time.

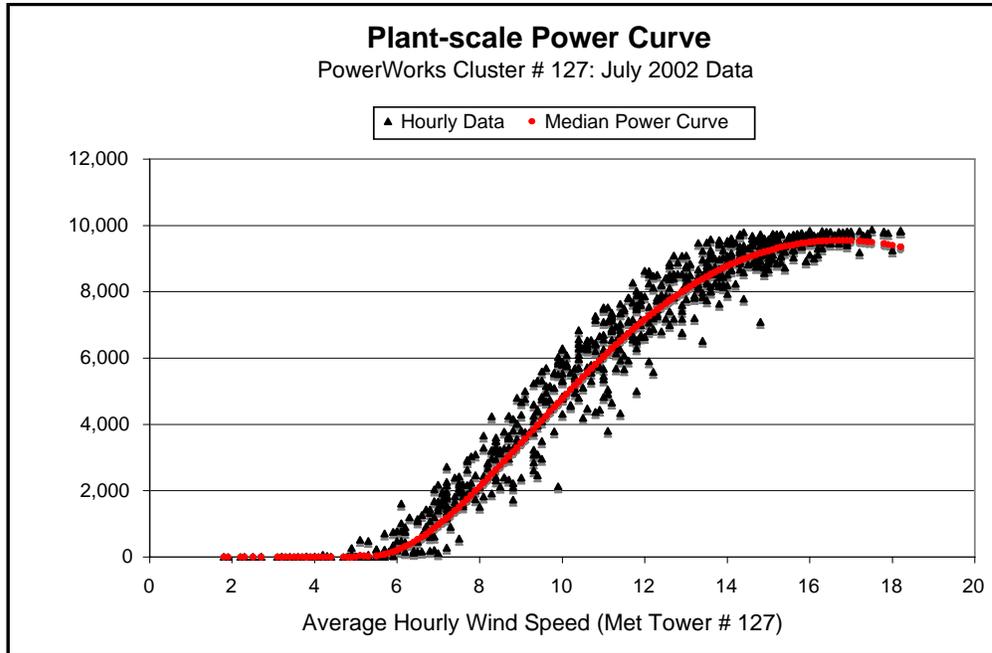


Figure 5-1 Hourly power output vs. measured hourly average wind speed (black triangles) for the PowerWorks cluster 127 for July 2002 and hourly power output estimates (red circles) using a plant-scale power curve derived from hourly data from June 2002.

High-Resolution Numerical Simulations of the Airflow over the Altamont Pass

In recent years, AWS Truewind has routinely used a physics-based atmospheric numerical model to forecast power production for a large fraction of California wind farms. Typically the model is configured to begin with a coarse outer grid with a 30- to 40-km resolution, then to nest down through higher resolution grids, to a final grid resolution of about 5 km. The output from the highest resolution grid is then used to drive a statistical technique, usually some version of a Model Output Statistics (MOS) scheme. The final result is a statistical forecast of wind speed over time at one or more locations for a given wind farm, which is then used to predict the wind plant power output through the use of an empirical plant-scale power curve.

For operational forecasting, the grid size of the highest resolution grid is limited by the computational resources required to execute the simulations in the time required to produce a useful forecast. The model time step decreases as the grid size decreases, so more time steps and more computational time are required for higher resolution simulations to simulate the same time period. Because of this limitation, it is difficult to run a model in real time with grid cells small enough (less than 1 km) to fully resolve the complex terrain of the Altamont Pass, which is only a few kilometers wide. To use an atmospheric numerical model to explicitly simulate wind speeds at each of the 93 turbines in the Tower M127 cluster, an even higher resolution (perhaps 25 or 50 m) would be required. As available computer power continues to increase, it may be feasible to run those kinds of simulations within a few years, but for the current project, a limited set of experimental high resolution simulations were executed in an attempt to understand the variability of the wind speeds in the M127 cluster of turbines, and evaluate whether the

knowledge gained enables us to predict the plant power output better than the typical plant power curve approach.

The MASS Model

Version 6.4 of the Mesoscale Atmospheric Simulation System (MASS) was used in this investigation. MASS was originally developed in the 1980's (Kaplan *et al.* 1982) as a research model. It has since been used for a wide range of research and commercial applications, and in recent years some of the model's databases and physical parameterizations have been optimized for a number of wind energy applications. MASS is a three-dimensional physics-based model which uses a set of mathematical equations that represent the basic physical principles of conservation of mass, momentum and energy and the equation of state for moist air.

Model Configuration

A set of five sequential nested model grids were used to model the Tower M127 area. Each grid consisted of a matrix of 125x125 horizontal grid points and 25 vertical levels extending from the ground up to 100 mb (about 16 km above sea level). The outermost (coarsest resolution) grid employed a grid cell size of 30 km, followed by nested grids with grid cell sizes of 8 km, 2 km, 500 m and 100 m. Figure 5-2 shows the terrain field for the highest resolution 100 m grid 2; the high terrain feature near the middle of the grid is Brushy Peak, the 519 m hill where the Tower M127 cluster is located. Because the terrain dataset also has a 100-m resolution, the height of the hill on the model grid is a little lower, between 450 and 500 m in the model. Figure 5-3 shows a wider view of the Altamont Pass terrain and also depicts the locations of each of the 11 PowerWorks, Inc. meteorological towers.

Each 30-km simulation was initialized with archived grid point data from the National Center for Environmental Prediction's (NCEP's) "Aviation" or "AVN" model, which has since been designated the Global Forecast System (GFS). Rawinsonde and surface observations were assimilated with an optimum interpolation analysis scheme to produce an initial state at 1200 UTC (0400 PST) on the first day of the simulation period. A 24-hr, 30-km simulation ending at 1200 UTC the next day was executed. A simulation using a grid with 8-km grid cells employed the output from the parent 30-km simulation for initial and boundary conditions, and was run over the same 24-hr period. Next, the 2-km simulation used the output from the 8-km simulation for initial and boundary conditions. The 2-km simulation extended for a 12-hr period beginning at 0000 UTC (1600 PST) on the second day. The output from the 2-km simulation was then used to supply the initial and boundary conditions for a nested simulation with 500-m grid size. The 500-m simulation was run for the same 12-hr period as the 2-km simulation. Finally a 100-m simulation was run for six hours, beginning at either 0000 UTC or 0600 UTC (2200 PST). The 100-m simulation used the output from the 500-m simulation to supply its initial and boundary conditions. Figure 5-2 shows the geographical area covered by the 100-m grid and the terrain elevation of the grid cells.

For each grid, a 4-km satellite-derived water surface temperature (WST) dataset was used to initialize water temperatures. The 4-km WST data is valuable because it has sufficient resolution to represent the water temperature of the San Francisco Bay reasonably well. The temperature of

the water in San Francisco Bay is an important factor the development of winds in the Altamont Pass. Especially during the warm season, the winds in the Bay Area are primarily forced by the difference in air temperature between the cool marine areas to the west of the Pass and the much warmer land areas in the valley to the east of the Pass. The December 2001 and May 2002 simulations used the AVHRR Pathfinder WST dataset, which extends back into the mid-1980s. For the July 2002 simulations, the WST dataset from the MODIS Aqua satellite was used. This satellite was launched just prior to July 2002 and the data from this sensor did not become available until July 2002. The MODIS WST data are thought to be somewhat better than the Pathfinder data due to advances in instrumentation and processing.

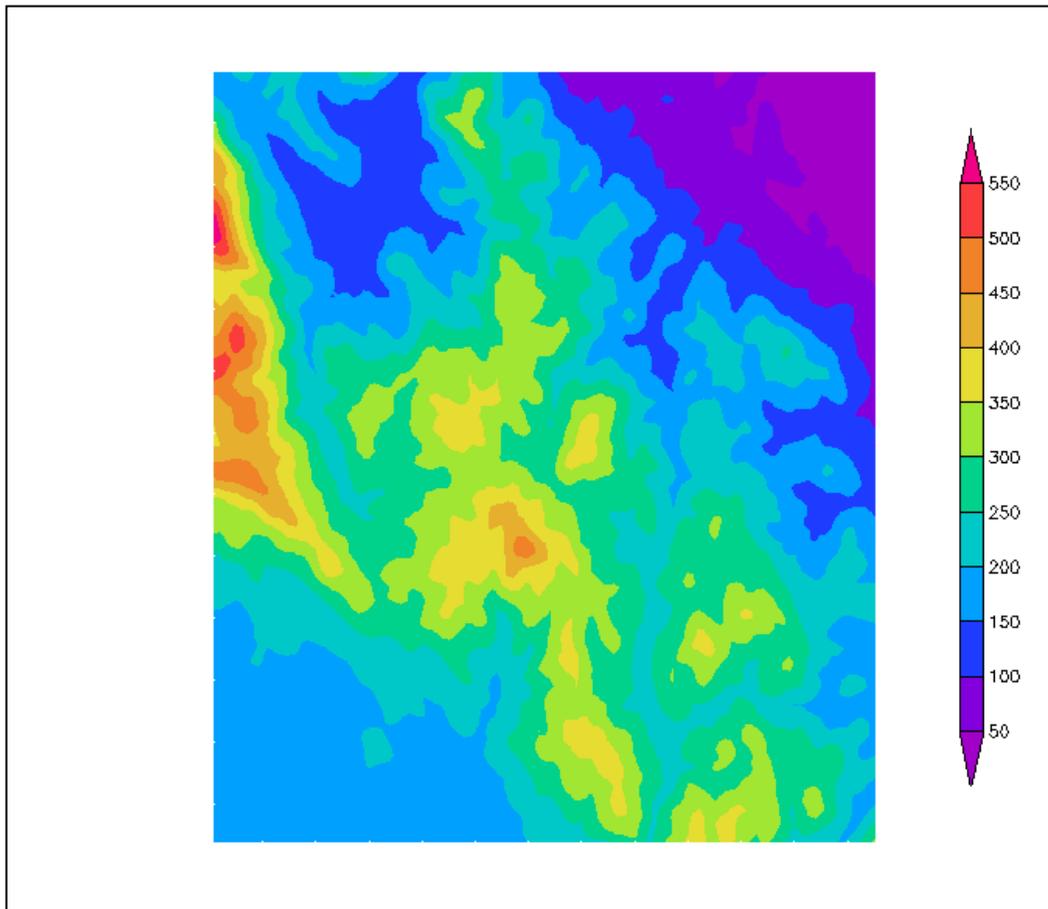


Figure 5-2 The area encompassed by the domain and the associated terrain representation for the 100-m MASS model grid. The color shading indicates the terrain elevation above sea level in meters.

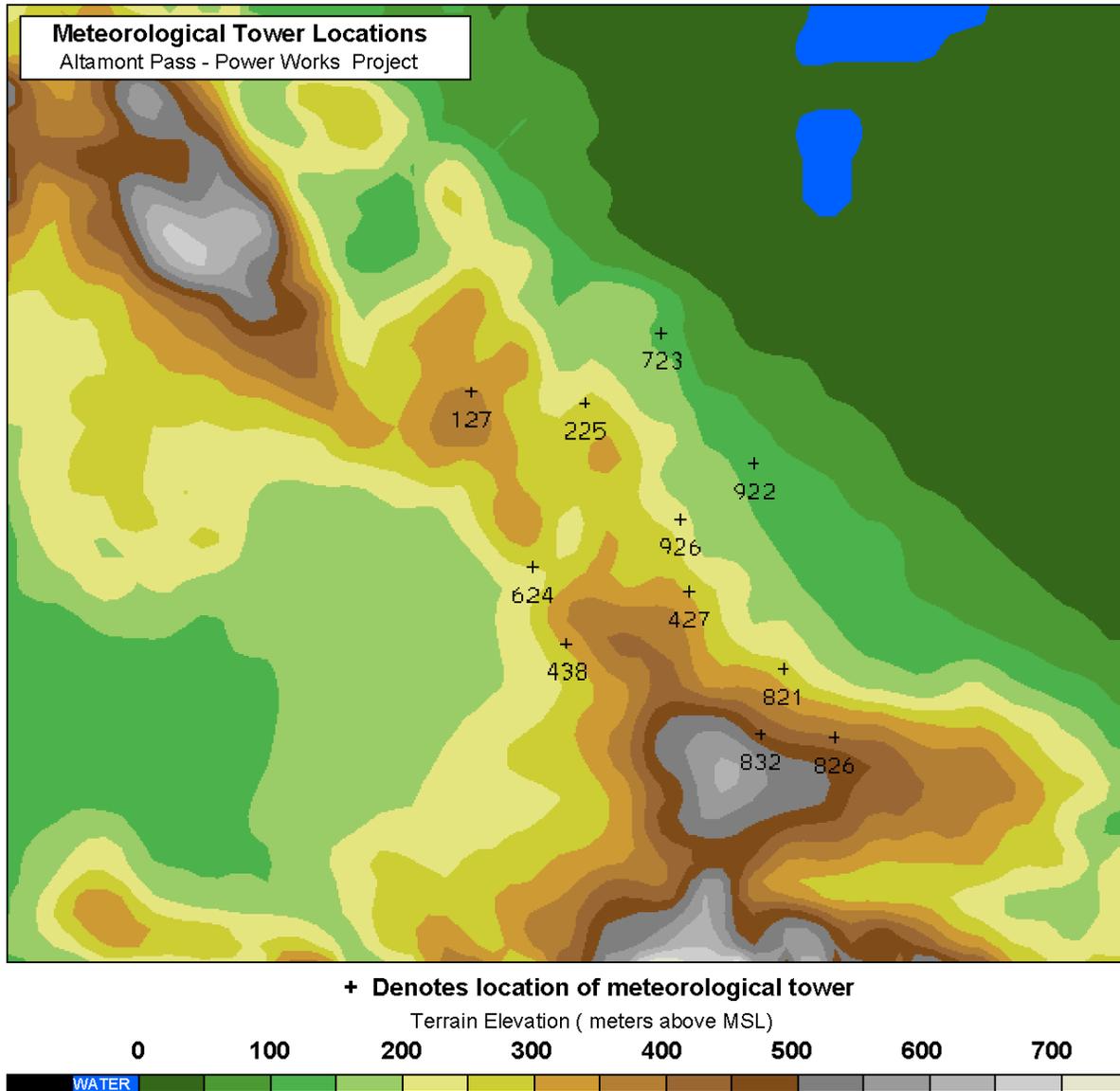


Figure 5-3 A broader view of the terrain of the Altamont Pass. The terrain is shaded in meters, and the locations of each of the 11 PowerWorks, Inc. meteorological towers are denoted by the “+” symbol and the associated 3-digit tower identification number.

Ten Case Studies

Ideally, 100-meter simulations would be executed over a long time period, but these very high-resolution simulations require very short model time steps (about 0.5 second), so they take a long time to execute, even on the relatively fast Intel or AMD processors used by AWS Truewind. A single 6-hour 100 meter simulation took about a day and a half to complete. Therefore, a set of ten case studies were chosen from the model runs executed in Subtask 6.1 for three months: December 2001, May 2002 and July 2002. It was decided to limit the 100-meter simulations to the nighttime hours (1600 PST-0400 PST) because the winds through the Pass are usually strongest at night, and therefore most relevant to power production. Table 5-1 lists the ten cases

chosen from the three months. Cases were chosen to represent a range of conditions, including high, moderate and low wind speed cases. Some of the cases were well-simulated by the 1-km simulations performed in Subtask 6.1 of this project, while others were not handled as well by the model. Five cases include 100 meter simulations that begin at 0000 UTC (1600 PST) and five begin at 0600 UTC (2200 PST). The initialization time of the 100 meter simulations was assigned randomly so that each month had at least one simulation beginning at each time.

Table 5-1 Ten Cases Chosen for Numerical Simulations Using a 100-m Grid.

Date	Period of 100 m simulation	Comments
2 December 2001	0000-0600 UTC	Storm event - 14 m/s winds at Tower M427; fairly good simulation
15 December 2001	0600-1200 UTC	Low wind speed day (about 2-3 m/s in model and observations all day)
20 December 2001	0600-1200 UTC	Weaker storm event than Dec 2: pretty good simulation of winds with the 1 km model
6 May 2002	0000-0600 UTC	Very high wind speeds (22 m/s at tower M427), 1 km simulation has a large low bias
19 May 2002	0000-0600 UTC	Low wind speeds for May; 1 km simulation was very good
30 May 2002	0600-1200 UTC	Moderate wind speeds: 1 km simulation was pretty good
16 July 2002	0000-0600 UTC	Average diurnal wind speed day; pretty well simulated at 1 km
19 July 2002	0000-0600 UTC	Noticeable diurnal cycle but the wind speeds are well below the typical July day (peak of about 10 m/s at tower M427). Well simulated at 1 km
24 July 2002	0600-1200 UTC	One of the windiest days of the month (20 m/s at tower M427), not that well simulated (wind speeds are 5 m/s or a little more too low at the peak)
25 July 2002	0600-1200 UTC	Average wind speed and diurnal cycle; quite well simulated at 1 km

Figure 5-4 shows simulated wind vectors at turbine hub height (18 meters above the ground) from the 100 meter simulation, several hours into the 16 July 2002 simulation. Some basic features of high wind events in the Altamont Pass can be seen – the sluggish flow just upstream of the Pass, speed-up over high terrain such as at the Brushy Peak location of the Tower M127 cluster and the higher terrain to the northwest.

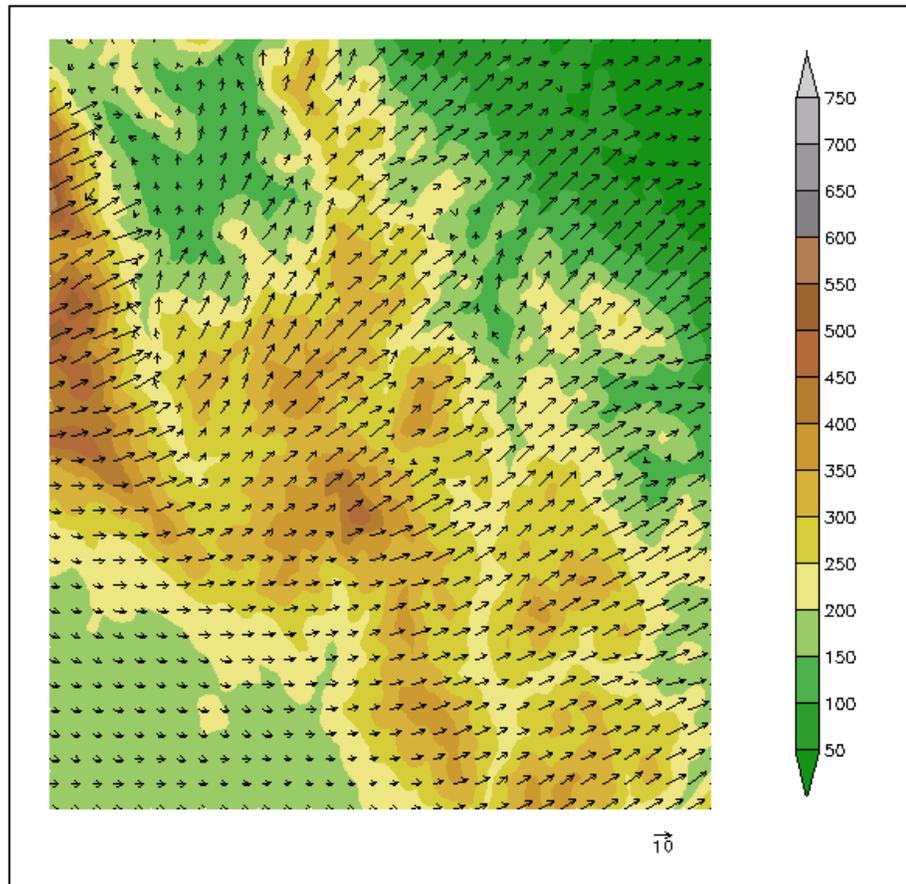


Figure 5-4 The 18-m wind field at 0430 UTC 16 July 2002 (8:30 p.m. PST), several hours into the 100 m simulation. Wind vectors are shown against a shaded terrain elevation background. The length of the vectors is proportional to the wind speed; a 10 m/s scale arrow is shown at the bottom. The maximum winds at this time are about 25 m/s. For readability, a vector is shown for every fourth grid point 400 m between points instead of the actual 100 m spacing of the computational grid).

Predicting Plant Power Production from Inferred Turbine Wind Speeds

Turbine Wind Speed Ratios From Simulations

Since almost all of the 93 turbines associated with Tower M127 have a hub height of 18 meter, the 18 meter wind speed was extracted from each of the ten 100-m simulations listed in Table 5-1. Wind speed data were extracted every 15 minutes during the 6-hr simulations. The ratio of the simulated wind speed at each turbine location to the Tower M127 simulated wind speed was calculated for each turbine site. Figure 5-5 shows how these ratios vary with time during the course of the simulation for the 16 July 2002 case. The figures include curves for only 33 of the 93 turbine locations, because the turbines are close together and several 100-meter cells contain more than one turbine.

Each of the wind speed ratios begin with a value of 1 at the beginning of the simulation because the 100 meter simulation was initialized at the same time as its 500-m parent simulation. Therefore, the initial spatial differences are from the 2-km simulation that served as the parent for the 500-m simulation, and the entire M127 cluster is contained within one 2-km grid cell. Differences in neighboring 100 meter grid cells develop as smaller-scale wind features develop during the simulation. Most of the turbines have lower wind speeds than Tower M127, indicated by ratios less than 1, because the tower is on a higher-elevation, more favorable position for strong winds than a majority of the surrounding turbines. The mean ratio for all of the turbines declines during the course of the simulation, and the wind speeds at some individual turbine locations decline to as little as 60% of the M127 wind speed.

The evolution of the turbine wind speed ratios for different cases indicates the complexity of the wind flow through the Altamont Pass. Figure 5-6 shows the evolution of the wind speed ratios vs. time for the 19 May 2002 case. In this case, the ratios are very steady and spread over a smaller range of values than the 16 July case. In contrast, Figure 5-7 shows the results for the December 6, 2001, case – the ratios are highly variable throughout the simulation as the wind speeds at M127 oscillate between 1 and 10 m/s. At some turbines, the wind speeds increase to as much as four times those at Tower M127.

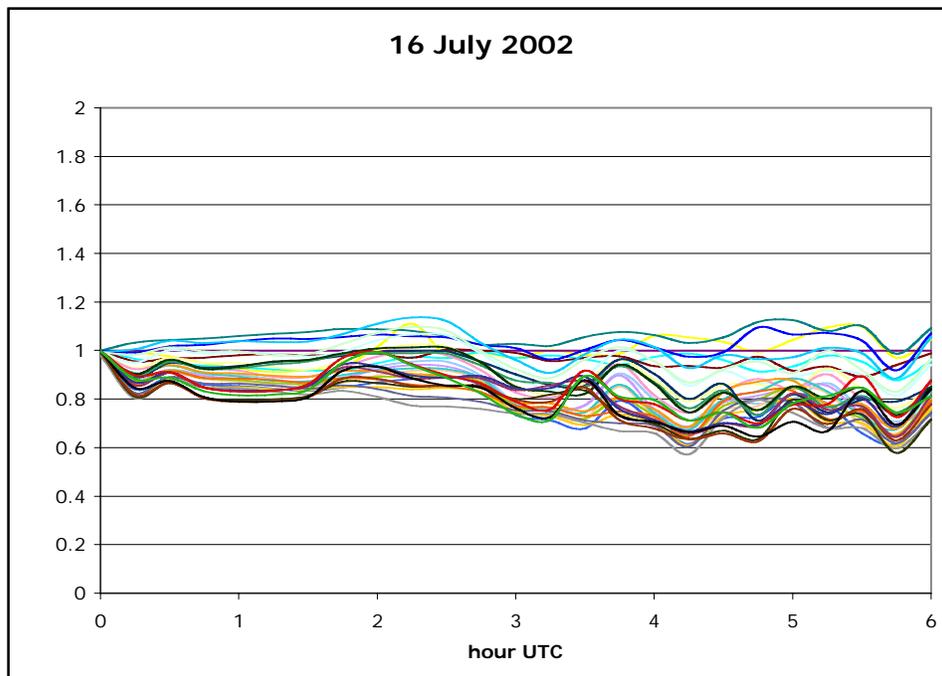


Figure 5-5 Evolution of turbine wind speed ratios over the 6-hr 100 m simulation beginning at 0000 UTC 16 July 2002.

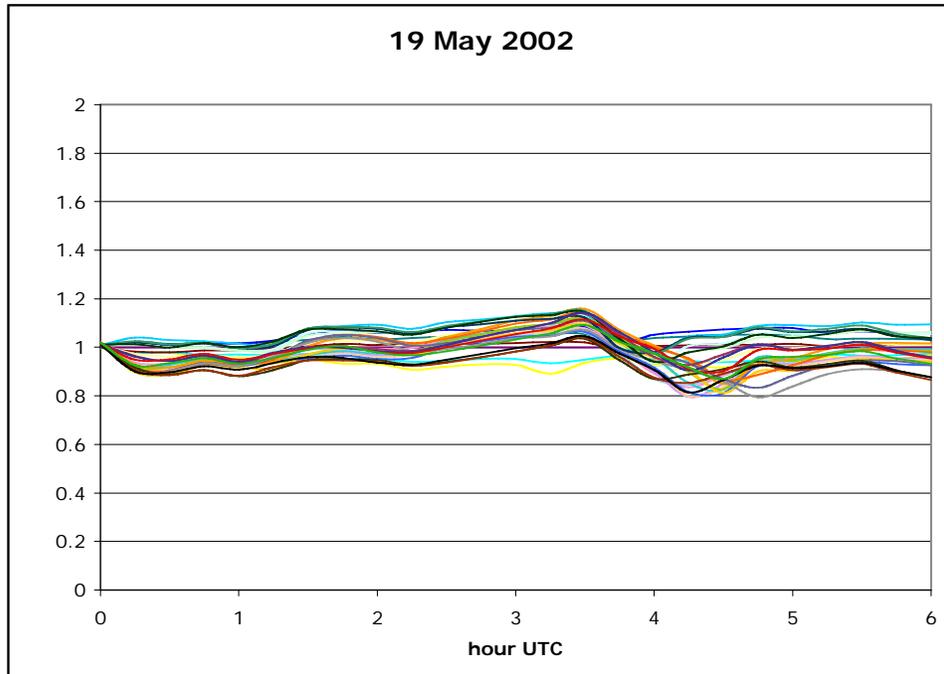


Figure 5-6 Evolution of turbine wind speed ratios over the 6-hr 100 m simulation beginning at 0000 UTC 19 May 2002.

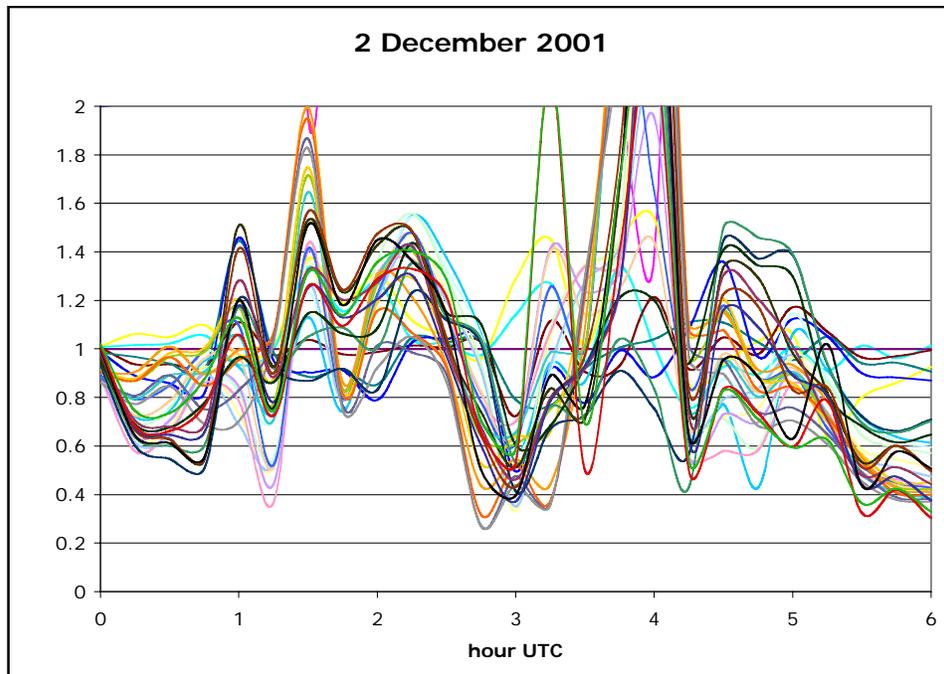


Figure 5-7 Evolution of turbine wind speed ratios over the 6-hr 100 m simulation beginning at 0000 UTC 2 December 2001.

The wild oscillations of Figure 5-7 might suggest that the turbine-to-turbine wind speed variations are too unpredictable to use as the basis for plant power production forecasting, but meaningful patterns emerge when the combined behavior of all of the turbines in the cluster is examined. To calculate a plant power output based on inferred wind speeds at individual turbine locations, an average ratio at each of the turbine locations was calculated by averaging ratios every 15 minutes over the six hours of all ten cases, a total of 250 values at each turbine location. The estimated latitudes and longitudes for the meteorological tower and each of the turbines that were used for the extraction were kindly provided by Dave Lubitz at UC Davis (Lubitz 2005).

Figure 5-8 compares these mean ratios to the ratios calculated from the data gathered from the flow simulations in the UC Davis wind tunnel when the wind direction was set to the climatologically dominant direction of 240 degrees. There is no apparent correlation between the two sets of ratios. When the wind tunnel ratios are plotted against mean ratios from individual cases, some show a modest correlation, with linear correlation coefficients no higher than 0.3. The weak to non-existent correlation between the ratios derived from numerical simulations and those derived from wind tunnel simulations ratios probably reflects the complexity of the wind flow in the area, and the very different approaches used to simulate the flow. The most significant difference is probably due to the fact that the boundary layer wind tunnel always simulates a flow regime with neutral static stability, while the numerical model simulates flow for a wide range of stabilities.

Figure 5-9 shows the mean wind speed ratios calculated for each of the ten cases and compares the results to those generated using the wind tunnel data. The two approaches produce significantly different mean turbine wind speed ratios for each of the ten simulated cases, the mean ratio of the ten cases, and for two wind directions from the wind tunnel data. Although some individual turbine locations exhibit higher mean wind speeds than the meteorological tower, the mean wind speed of the entire cluster is less than that of the meteorological tower in every case, because Tower M127 is located in one of the windier ridge top locations.

When the wind tunnel models flow from the southwest direction (“Wind Tunnel-240” in Figure 5-9), the mean turbine wind speed ratio is 0.97. For southeast winds (“Wind Tunnel-150”), the mean ratio decrease to 0.91, because wind farms in the Altamont Pass are oriented along ridgelines to optimize production when the wind is from the dominant southwest direction. Ratios for the ten model-simulated cases range from 0.85 to 0.99, even though nine of the ten cases experienced wind directions from the southwest. The single case with winds from the southeast instead of southwest, the December 2, 2001 case, actually yielded the highest mean wind-speed ratio of any case, 0.99. The differences between the ratios from the numerical model and wind tunnel simulations cannot be explained by wind direction. The analysis presented in the following sections indicates that the boundary layer stability appears to be the key factor.

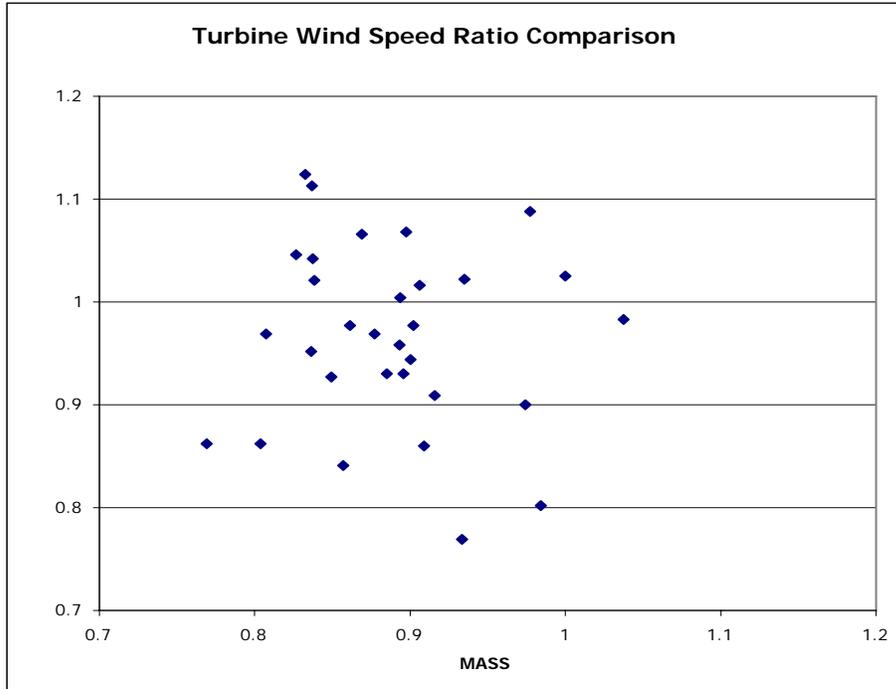


Figure 5-8 Comparison of mean turbine wind speed ratios calculated from the 100-m simulations to those calculated from the UC Davis wind tunnel data for the dominant wind direction of 240 degrees. Points representing 31 turbine locations are shown, which are the 33 locations resolved by the 100-m grid, minus two turbine locations which were outside the wind tunnel domain.

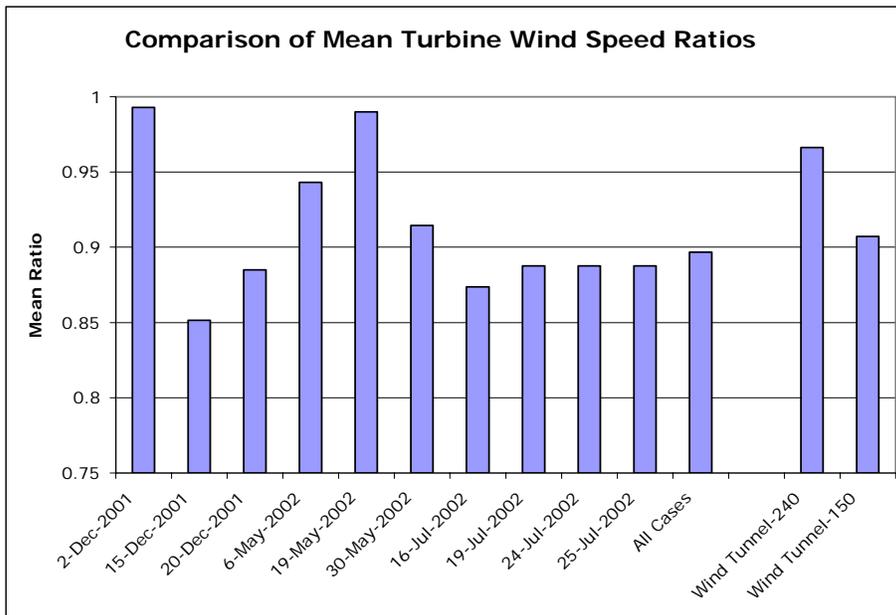


Figure 5-9 The mean wind speed ratio over all turbine locations for each of the ten 100 m numerical simulations and for simulations of two wind directions in the UC Davis wind tunnel.

The Importance of Stability

Averaged over all seasons and times of day, the lower atmosphere is stably stratified, meaning that vertical motion is minimized and horizontal wind speeds are almost always large compared to vertical wind speeds. The atmosphere tends to become less stable during the day as incoming solar radiation warms the surface, and more stable at night as the surface cools. Stability can be difficult to quantify, but a common approach is to examine the vertical profile of potential temperature. With strong daytime surface heating, the lowest few hundred meters of the atmosphere become well-mixed and the potential temperature is nearly constant with height, indicating near neutral (sometimes referred to as low) stability. With strong surface cooling on a clear night, or when the overlying air mass is cool and stable, the potential temperature will increase with height.

Figure 5-10 shows the 100-meter numerical simulation of potential temperature profiles at the meteorological tower location for the 100 meter simulation on December 2, 2001, the case with the highest mean turbine wind speed ratio. Even though the atmosphere tends to be more stable during the winter, these profiles are nearly neutral, with only small increases of potential temperature with height.

Under neutral stability, an air parcel experiences little resistance to uplift and flow around hills and other obstacles, and higher winds from aloft easily mix all the way down to the surface. So it can be expected that wind speeds at less favorable turbine locations and lower elevations may still be comparable to those at more favorable locations such as M127. Although the wind speed ratios for this case were highly variable (Figure 5-7), when averaged over the entire simulation, the average wind speed ratios at the turbine locations (0.99) are close to unity, and the average wind speeds at the turbine locations are very close to those at met tower, as expected for low stability.

In contrast, the December 15, 2001 case had the lowest mean turbine wind speed ratio, and the potential temperature profiles demonstrate that it is far more stable (Figure 5-11). The strong increase of potential temperature with height in the lowest 200 meter indicates the presence of a shallow stable layer which there is resistance to uplift and flow over obstacles, resulting in flow stagnation or “blocking” at lower elevation locations. Also, the stable boundary layer will resist mixing of stronger, higher-level winds down to the surface. In stable flow in complex terrain, cool stable air will tend to pool in the lower elevations, allowing the more “exposed” higher-elevation locations to experience the strongest winds. The mean turbine wind speed ratio of 0.85 for this case suggests that the model is responding appropriately to stable conditions, by reducing the wind speed at many turbine locations.

The stability in the lowest 200 meter for the 16 July 2002 case increases significantly during the 6-hour simulation (Figure 5-12), which extends from the late afternoon (4:00 p.m. PST) into the evening (10:00 p.m. PST). The evolution of turbine wind speed ratios during the simulation (Figure 5-5) mirrors the stability increase, as the ratios for many turbines decrease steadily. Stability often increases during the evening because of the radiational cooling of the earth’s surface.

However, for a midsummer evening, high-wind speed case at Altamont Pass, the wind speed ratios are likely to be affected by stabilization of the lower levels of the atmosphere due to the intrusion of the cooler marine layer from the ocean. The marine layer is transported across the Pacific coast and the Berkeley Hills by the large-scale sea breeze, accumulated on the Livermore plateau, and finally drained down to the much warmer Central Valley through the Altamont Pass.

Under the hypothesis that stability has a key, systematic influence on the turbine-to-turbine variation in wind speeds across the M127 turbine cluster, an effort was made to quantify that relationship. Stability is a difficult parameter to precisely quantify because it varies vertically as well as horizontally, and there is no standard way to measure it. Several parameters were tested without success, partly because near-surface stability is sometimes inferred from parameters that depend on a correlation between stable conditions and low wind speeds and turbulence levels in stable boundary layers. But the wind flow pattern in the Altamont Pass often involves the unusual combination of strong winds and high stability.

Stull (1988) defines several stability parameters that may be calculated from only the potential temperature profile. One of these parameters, the *cooling integral depth scale* is defined as:

$$H_{\Delta\theta} = \frac{\int_0^h [\theta_0 - \theta(z)] dz}{(\theta_0 - \theta_s)} \quad (5-1)$$

where θ_0 is the potential temperature at h , the top of the layer being considered, and θ_s is the potential temperature at the surface. The cooling integrated depth scale can be used to calculate the *bulk turbulence scale* from

$$B = \frac{H_{\Delta\theta}}{(\theta_0 - \theta_s)} \quad (5-2)$$

Stull (1988) notes that “Large values of B correspond to deep SBLs (stable boundary layers) with small surface temperature change, while small values correspond to shallow $H_{\Delta\theta}$ and large surface cooling.” For the ten simulated cases, larger values of the bulk turbulence scale were found for near-neutral cases, while smaller values were found for the more stable cases.

Figure 5-13 plots the mean turbine wind speed ratio plotted vs. the median bulk turbulence scale (B) for each case. The layer from the surface to 400 meter was used for these calculations ($h = 400\text{ m}$) on the assumption that the depth of the layer should be approximately the same as the difference in elevation between the Livermore plateau (about 120 meter) and the hills in Altamont Pass (about 500 meter). A larger set of simulations might yield a better general stability parameter than the bulk turbulence scale, but the parameter provides a reasonable initial indicator to parameterize the influence of stability. The biggest problem is that both the bulk turbulence scale and turbine wind speed ratios vary strongly with time during the transition from a daytime

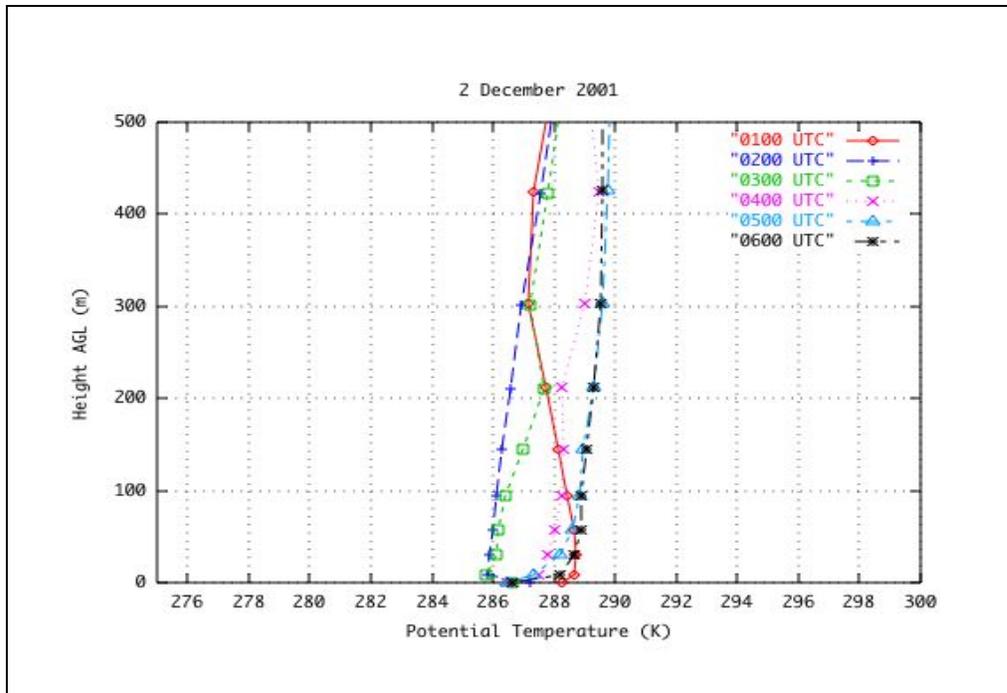


Figure 5-10 Hourly vertical profiles of potential temperature (in degrees Kelvin) from the 100 m simulation at Tower M127 from 0100 to 0600 UTC 2 December 2001. The vertical coordinate is height above ground level (AGL) in meters.

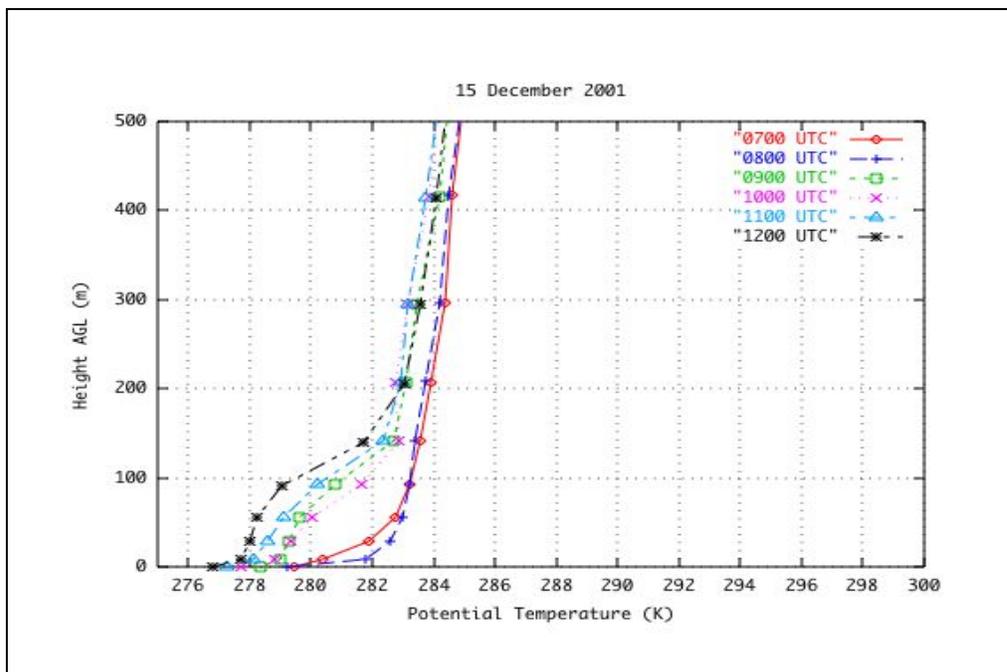


Figure 5-11 Hourly vertical profiles of potential temperature (in degrees Kelvin) from the 100 m simulation at Tower M127 from 0700 to 1200 UTC 15 December 2001. The vertical coordinate is height above ground level (AGL) in meters.

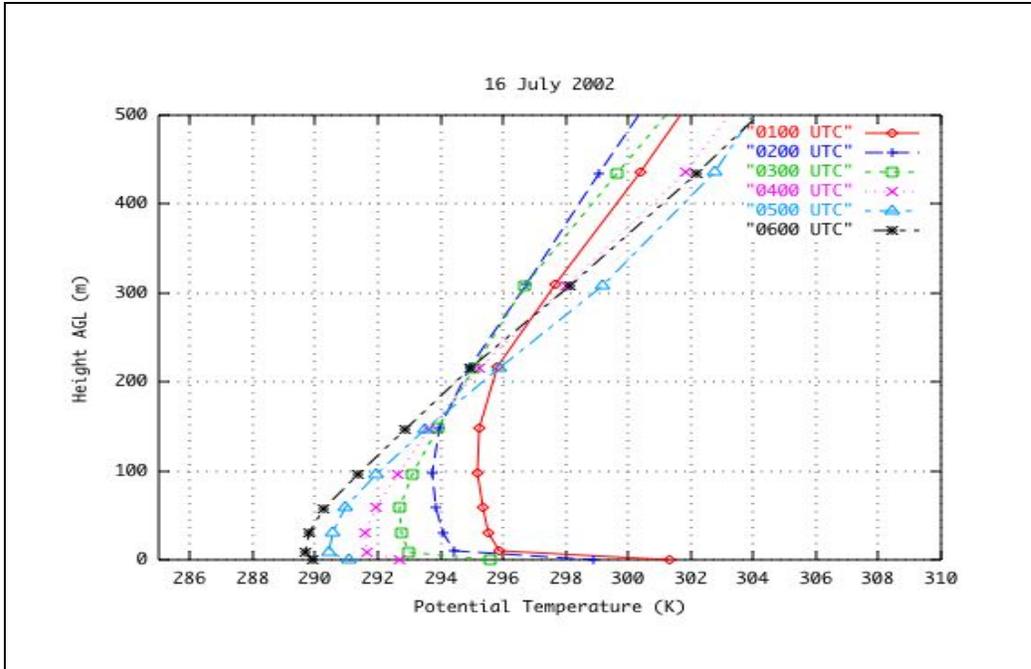


Figure 5-12 Hourly vertical profiles of potential temperature (in degrees Kelvin) from the 100 m simulation at Met Tower 127 from 0100 to 0600 UTC 16 July 2002. The vertical coordinate is height above ground level in meters.

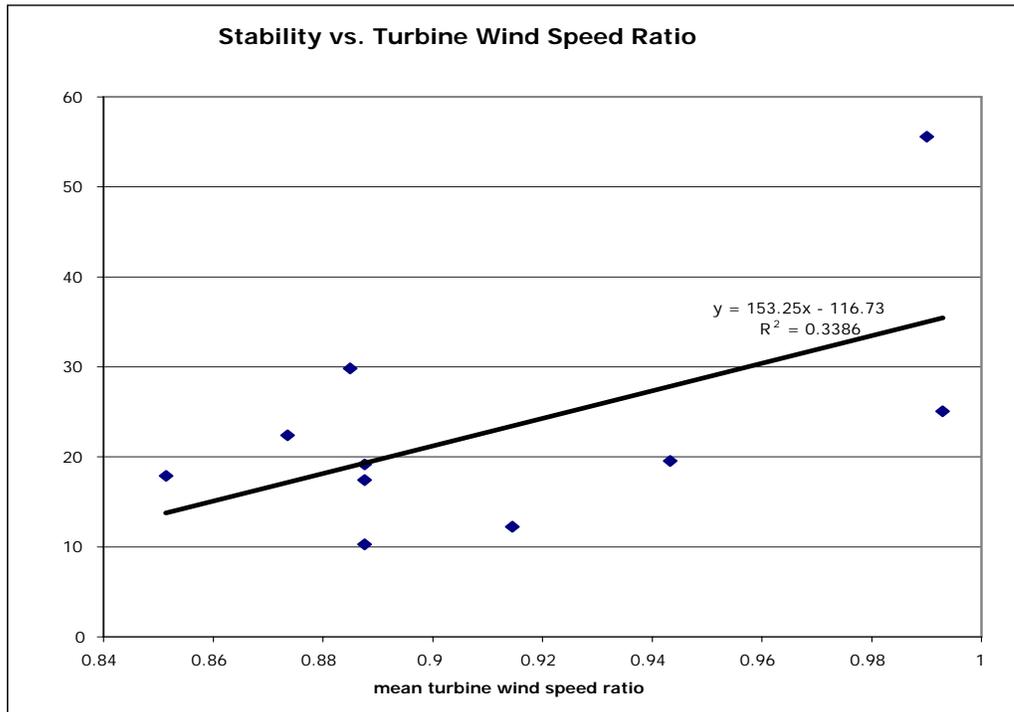


Figure 5-13 The median bulk turbulence scale (defined by Eqn 5-2) vs. the mean turbine wind speed ratio for each of the ten high-resolution numerical simulations. Each data point represents one of the ten simulated cases listed in Table 5-1.

near-neutral boundary layer to a more stable nocturnal boundary layer, necessitating averaging over many time periods.

Calculating Plant Power Production

Given a set of turbine wind speed ratios and either an observed or forecasted wind speed at M127, the prediction of plant power proceeds exactly as it does for the wind tunnel approach. The turbine wind speed ratios are multiplied by the meteorological tower wind speed to infer the wind speed at each turbine.

The 93 wind turbines were installed during the 1980's and are identical Kenetech 56-100 wind turbines, except that some have 18-m (60-ft) heights and some have 24-m (80-ft) hub heights. The manufacturer's power curve is used to calculate a power generation for each turbine (P_{turb} , in kW) for the wind speed (U_{turb} , in m/s):

$$P_{turb} = \begin{cases} 0, & U_{turb} \leq 4.8319; \\ -0.0224U_{turb}^3 + 1.8448U_{turb}^2 - 11.559U_{turb} + 15.308, & 4.8319 < U_{turb} < 12.1704; \\ 107.5, & U_{turb} \geq 12.1704; \end{cases} \quad (5-3)$$

The calculated power production for each turbine is then aggregated to calculate the total power production for the M127 cluster of turbines. Dave Lubitz at UC Davis provided Perl code that he used for the wind tunnel calculations (Lubitz 2005). Portions of this code were incorporated into the AWST code to perform the similar task.

Figure 5-14 shows the plant power production predicted by the turbine ratios from the high-resolution numerical simulations and the observed wind speeds for July 2002, compared to the observed hourly power production. It should be compared to Figure 5-1, which shows a statistical plant-scale power curve for the same month. This chart indicates that the turbine ratios derived from the numerical simulations appear to systematically under-predict power output for wind speeds in the middle of the steep part of the curve, and over-predict power for the higher wind speeds.

Evaluating Plant Power Predictions

The set of turbine wind speed ratios derived from the ten high-resolution numerical simulations were used to produce plant-scale power production predictions from a set of either measured or forecasted M127 wind speeds. These predictions were compared to the predictions based on the empirical plant-scale power curve (similar to the one shown in Figure 5-1) and the turbine ratios derived from the UC Davis boundary layer wind tunnel data. First the turbine wind speed ratios averaged over all 60 hours of the high-resolution simulations were tested. These mean ratios were then modified to account for stability variations and tested again.

Power Predictions Using Observed Wind Speeds

Two datasets were used to evaluate the plant power production predictions: (1) a collection of 10 months of hourly M127 observed wind speed, wind direction and total cluster 127 power production, along with forecasts of the same parameters from numerical model simulations produced for an earlier EPRI project (EPRI 2002). It will be referred to as the “Dec 2001-Sep 2002” dataset; and (2) a collection of over three years of half-hourly M127 observed wind speed, wind direction and plant power production data, where an attempt has been made to correct the power to 100% turbine availability, using information on offline status. It will be referred to as

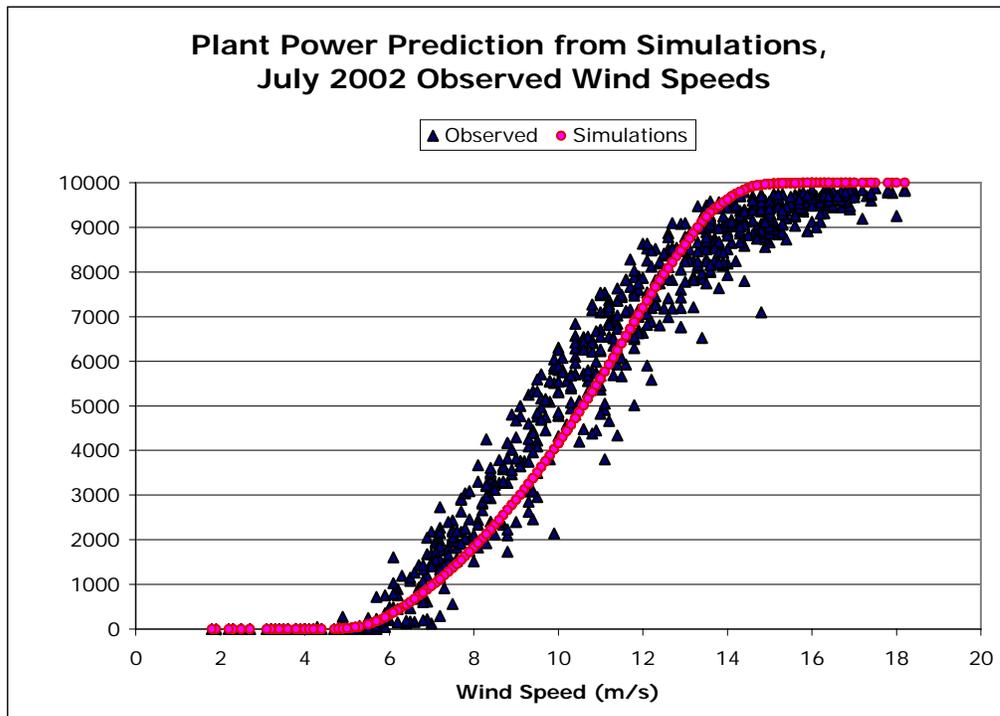


Figure 5-14 Hourly power output vs. measured hourly average wind speed (black triangles) for the PowerWorks M127 turbine cluster for July 2002, and hourly power output predictions (red circles) using the numerically-simulated mean turbine wind speed ratios.

the “Jun 2001-Jul 2004” dataset. For the latter dataset, only the hourly values were used, and the power values were multiplied by two in order to consistently use energy units of kWh.

For operational wind forecasting, AWS Truewind frequently recalculates the empirical plant-scale power curve as new data become available. The plant-scale power curve used for the experiments in this project was:

$$P_{127} = \begin{cases} 0, & U_{127} \leq 5.2 \text{ || } U_{127} > 25; \\ -0.009U_{127}^5 + 1.1427U_{127}^4 - 47.393U_{127}^3 + 786.41U_{127}^2 - & 5.2 < U_{127} \leq 22; \\ 4418.5U_{127} + 7588.5, & \\ -2600U_{127} + 65000, & 22 < U_{127} \leq 25; \end{cases} \quad (5-4)$$

where U_{127} is the wind speed at M127 and P_{127} is the power output for the associated cluster.

For the Dec 2001 to Sep 2002 dataset, the measured M127 wind speeds were used to predict the hourly plant power output in kWh by three methods: (1) the plant-scale power curve given in Equation 5-4; (2) the mean turbine wind speed ratios calculated over the 60 hours of high-resolution model simulations; and (3) the turbine wind speed ratios calculated in the UC Davis wind tunnel with the prevailing southwest wind direction (240 deg). Using the wind tunnel ratios, power predictions at the 87 of 93 wind turbines included in the wind tunnel were accumulated; there was no increase to account for the omitted turbines and no density correction was made.

Before spending a good deal of effort to evaluate those three methods, it may be useful to first compare the plant power curve method to a much simpler approach (Figure 5-15). If the observed wind speed at M127 is assumed to prevail at each of the 93 turbine locations, and the manufacturer’s power curve given by Equation 3-3 is used to calculate the power production for each turbine, the total plant power production is severely over-predicted, as shown by the large positive mean error (ME) and larger mean absolute error (MAE) of the “Manufacturer’s Power Curve” on the right side of the chart. The MAE and ME are expressed as a percentage of the hourly plant capacity, which is 9997.5 kWh for the M127 cluster. The positive bias is high, because most of the turbines frequently experience significantly lower wind speeds than Tower M127. The plant power curve accounts for this tendency because the data used to construct the curve account for its effect. Both the numerical model and wind tunnel approaches predict that lower wind speeds than an most of the turbine locations will typically have lower wind speeds than the at the M127 tower.

Figure 5-16 compares MEs and MAEs for each of the methods. The empirical plant power curve technique yielded the lowest MAE of the three methods, and the MAEs of the wind tunnel and numerical simulation methods were only 0.1% higher and 9.4% higher. The numerical simulation method yielded a significantly lower ME at 0.4%. Figure 5-17 shows the same comparison for the lengthier June 2001 to July 2004 dataset. The results are almost identical for both datasets, so the relative performance seems to be robust, and unaffected by the attempt to account for offline turbines in the June 2001 to July 2004 dataset.

Figure 5-18 compares the three methods for each of the ten months of the December 2001 to September 2002 dataset. The relative performance shows clear seasonal trends – both the empirical plant power curve and wind tunnel methods yielded significantly lower MAE’s (by as much as 2.5%) than the numerical simulation method for the five months from April through August, with the empirical plant power curve being the best. For the other five months, the numerical simulation method outperforms the other two, with MAE’s as much as 1.2% lower. In most cases, the wind tunnel method errors are between the other two. The reason for this seasonal variation in relative performance is not clear. The numerical simulations were executed

for cases in months when the simulation method performed both better and worse than the other methods. Figure 5-19 shows the monthly ME values for the three methods. The numerical simulation method has smaller absolute ME's than the other two in eight out of ten months. All three methods have relatively large positive biases in January and February.

The coherent seasonal differences in the performance of the methods suggest that some improvement might be achieved by predicting the power with the best method for each month. Figure 5-20 shows a modest improvement in both MAE and ME if the power is predicted using the empirical plant power curve from April through August and with the average turbine ratios from the numerical simulations in the other five months.

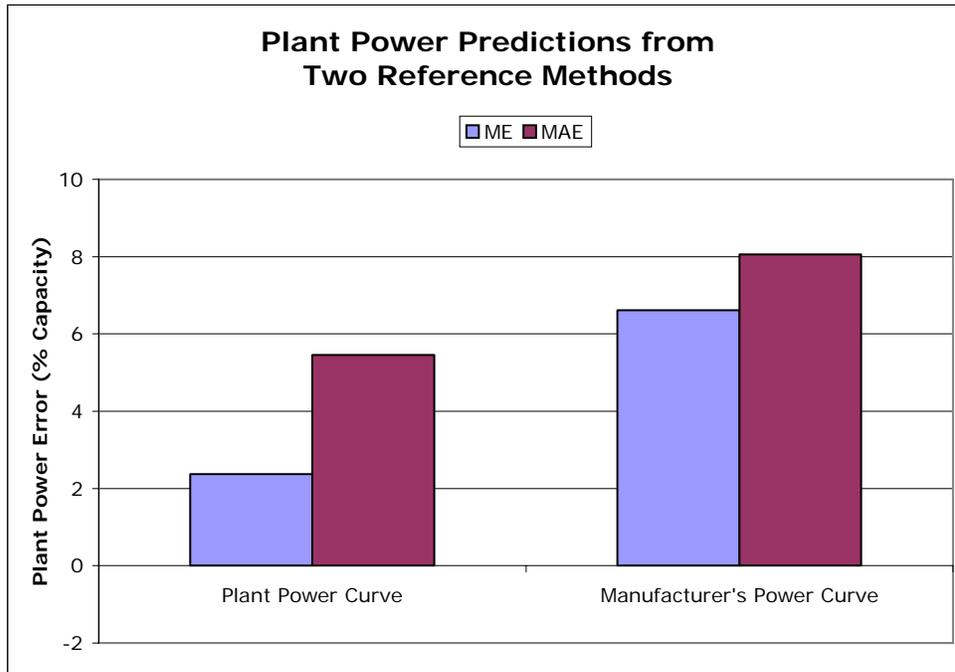


Figure 5-15 Comparison of two reference methods of predicting the plant power output using measured M127 wind speeds from the Dec 2001 to Sep 2002 dataset. The “Plant Power Curve” method uses Equation 4-1, and the “Manufacturer’s Power Curve” method assumes that the M127 wind speed is uniform across all of the turbines and the manufacturer’s power curve (Equation 3-3) can be used to calculate the power output of each turbine. The mean error (ME) and mean absolute error (MAE) of the hourly power output (kWh) are expressed as a percentage of hourly capacity of the M127 cluster, which is 9997.5 kWh.

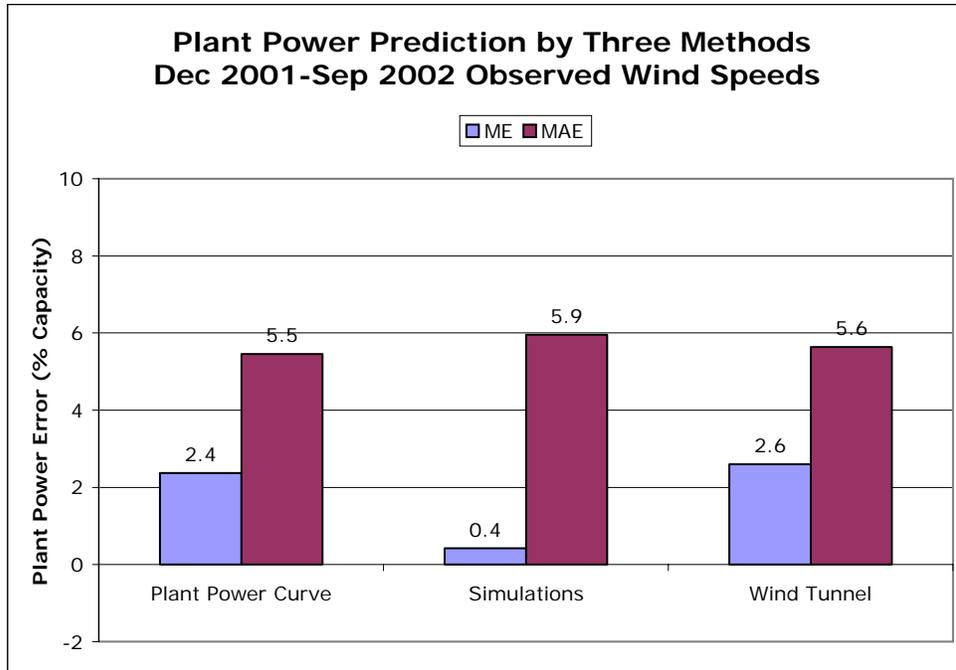


Figure 5-16 Comparison of mean error (ME) and mean absolute error (MAE) for three methods of predicting the plant power output using observed M127 wind speeds from the Dec 2001 to Sep 2002 dataset.

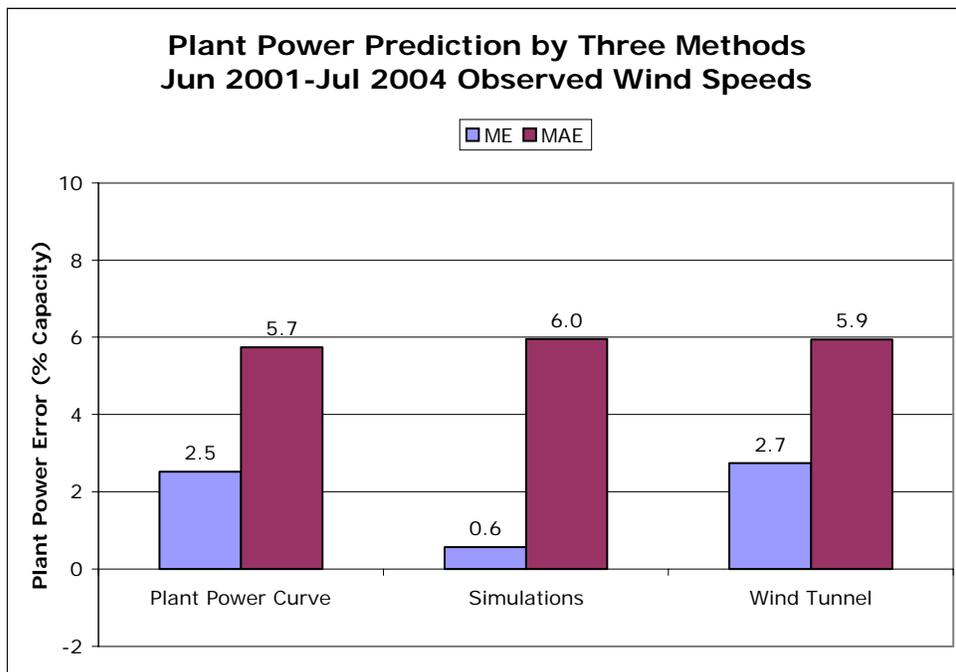


Figure 5-17 Comparison of mean error (ME) and mean absolute error (MAE) for three methods of predicting the plant power output using observed M127 wind speeds from the June 2001 to July 2004 dataset.

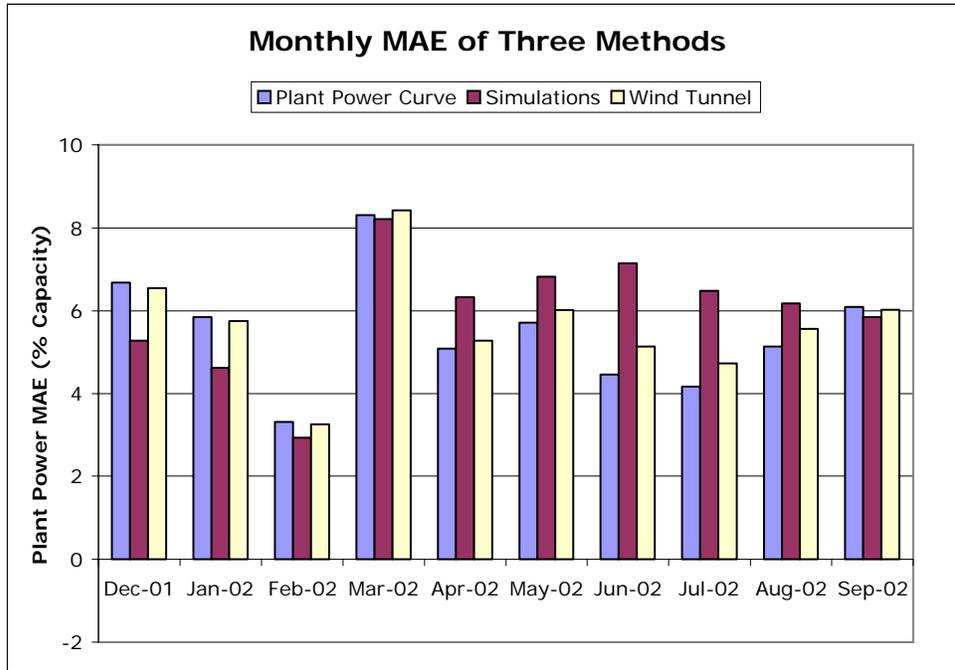


Figure 5-18 Monthly comparison of the mean absolute error (MAE) for the three methods of predicting the plant power output using observed M127 wind speeds from the December 2001 to September 2002 dataset.

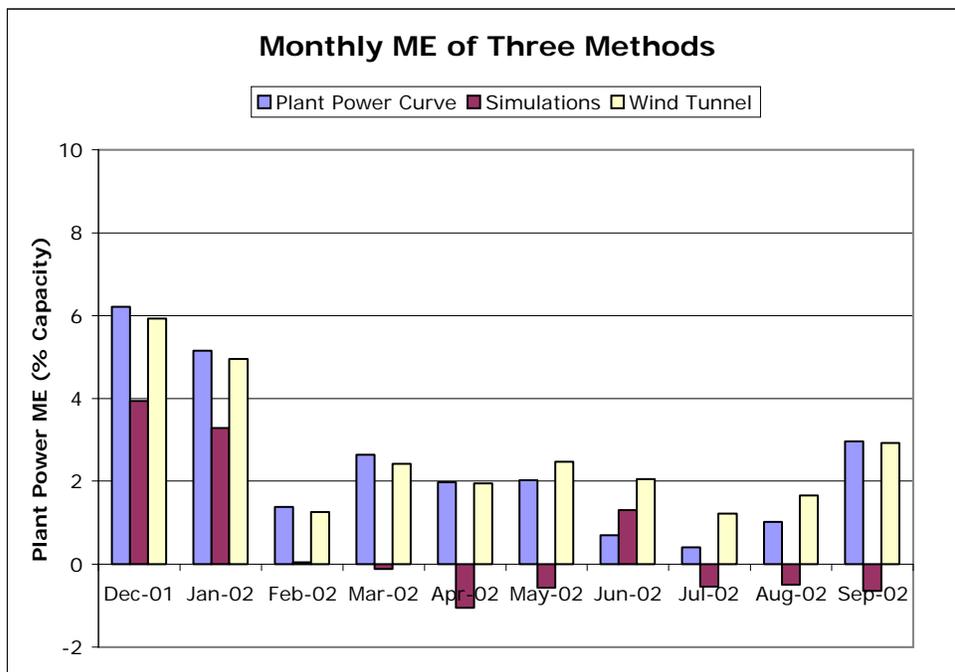


Figure 5-19 Monthly comparison of the mean error (ME) for the three methods of predicting the plant power output using observed M127 wind speeds from the Dec 2001 to Sep 2002 dataset.

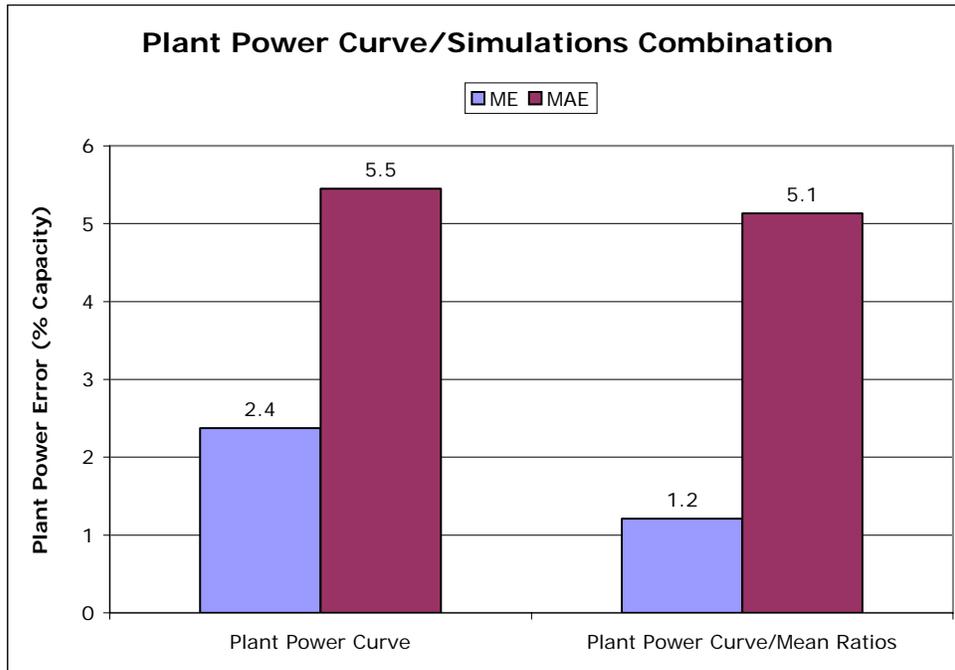


Figure 5-20 Comparison of the mean error (ME) and mean absolute error (MAE) produced by the empirical plant power curve for all months and those produced by employing the plant power curve for its best-performing months (April-August) and the mean turbine ratios from the numerical simulation method for the other five months.

Power Predictions Using Forecasted Wind Speeds

Since the objective is to ultimately improve the performance of real power production forecasts, it is important to see how the methods perform using forecasted rather than observed winds. To test the performance of the methods in a forecasting application, forecasts were produced from all three methods using forecasted M127 wind speed data for the December 2001 to September 2002 period. The forecasted wind speed data was obtained from the *eWind* forecasts produced in the previous Energy Commission-EPRI wind forecasting project (EPRI 2002). These predictions are based upon a MOS-adjustment to the data from a 10 km physics-based model simulation. The MOS (statistical) forecasts were derived from physics-based model data for a period 18 to 41 hours after the initialization of the simulation. Therefore, the set of statistical forecasts may be of varying quality.

Figure 5-21 compares the three methods over the entire December 2001 to September 2002 dataset. The empirical plant power curve again has the lowest MAE, but only by a very small margin. But this time the plant power curve also has a lower ME, and the numerical simulation method has a significant negative bias that the other two methods do not have. Figures 5-22 and 5-23 show the monthly MAE and ME values, respectively. The relative performance of the three methods is similar to what was seen in the forecasts based on the observed wind speed data, but all three methods showed different seasonal performance trends. Using observed wind speeds,

January and February were the most biased months for all three methods (Figure 5-19), but March, April and September were the most biased when using forecasted winds. Merging the plant power curve and numerical simulation method by month as in the previous section again improved the MAE by a few tenths of a percent of capacity, but the merged ME was not as good as the plant power curve ME.

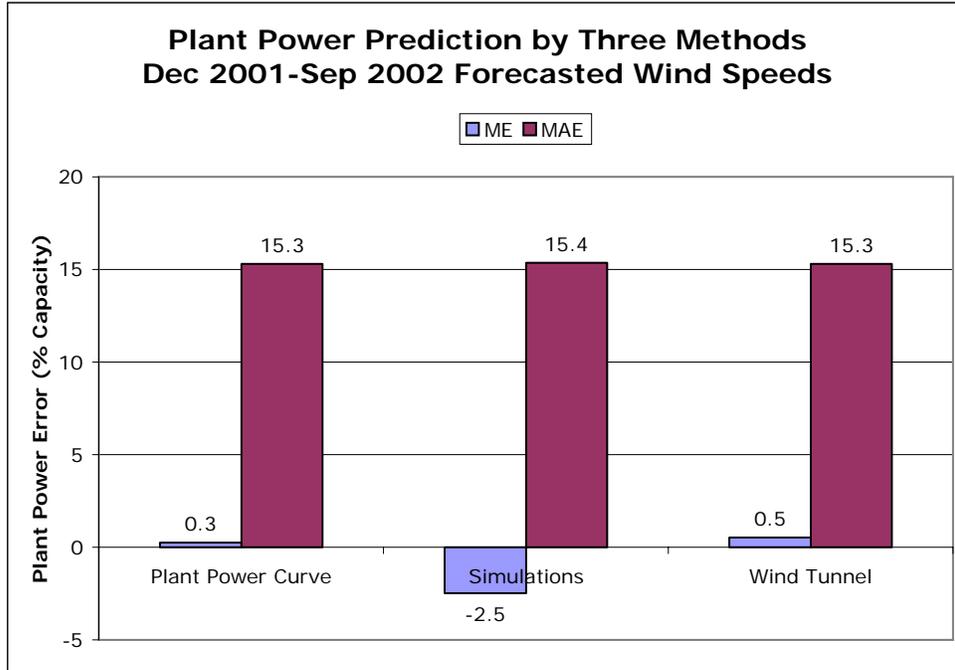


Figure 5-21 Comparison of three methods of predicting the plant power output using forecasted M127 wind speeds from the December 2001 to September 2002 dataset.

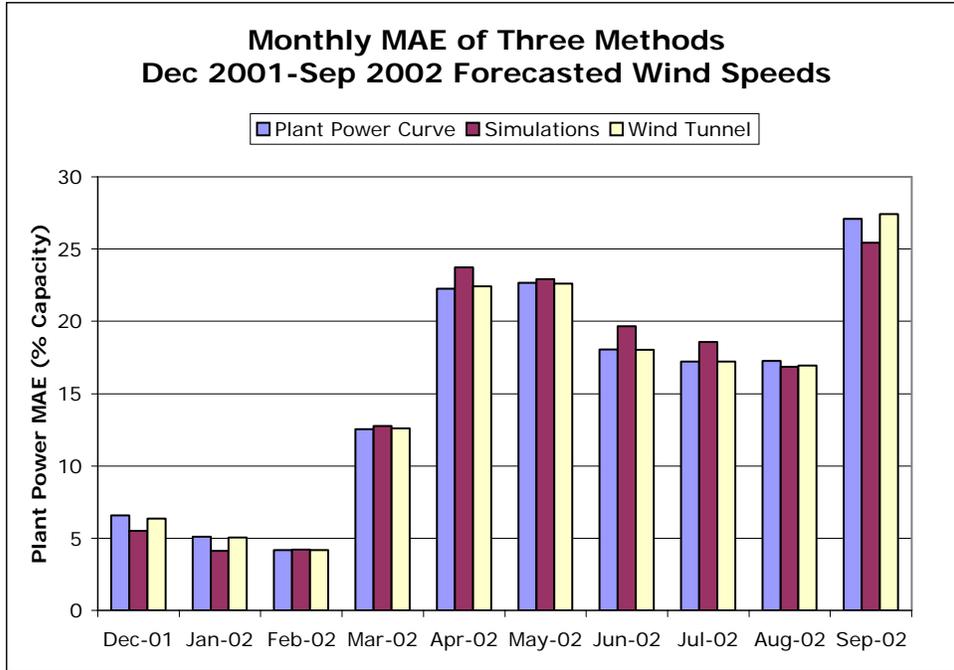


Figure 5-22 Monthly comparison of the mean absolute error (MAE) for the three methods of predicting the plant power output using forecasted M127 wind speeds from the Dec 2001-Sep 2002 dataset.

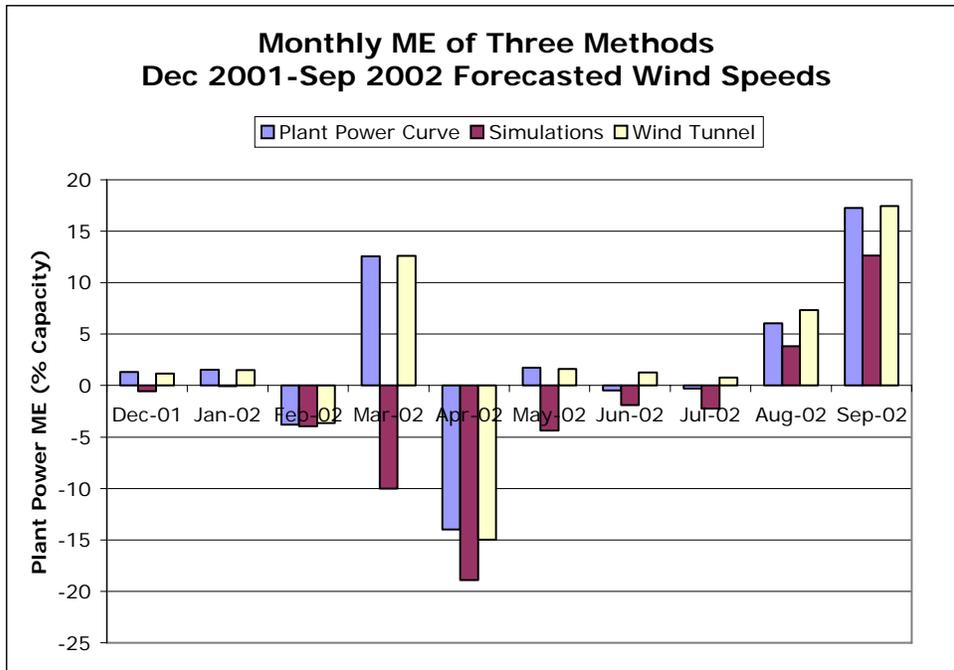


Figure 5-23 Monthly comparison of the mean error (ME) for the three methods of predicting the plant power output using forecasted Met Tower 127 wind speeds from the Dec 2001-Sep 2002 dataset.

Power Predictions with Stability Modification

As noted previously, the static stability of the lowest few hundred meters of the atmosphere over the Altamont Pass clearly influenced the turbine wind speed ratios, with more stable conditions leading to smaller mean ratios as the wind speed difference widened between the better exposed locations such as Tower M127 and many of the less favorably-sited turbines. An obvious question is whether we can use stability information to make better turbine wind speed inferences and therefore better plant power production predictions.

From the ten high-resolution simulations, the bulk turbulence scale was extracted at the Tower M127 site every hour through the 6-hr simulations. At these 70 times, the mean turbine wind speed ratio was calculated from the bulk turbulence scale using the simple linear relationship shown in Figure 5-13:

$$\bar{R}_{stab} = \frac{(B + 116.73)}{153.25} \quad (5-5)$$

A zero or negative value of B indicates neutral or unstable conditions, so the largest mean turbine ratio found in the ten cases (0.99) was used, and \bar{R}_{stab} was not allowed to be smaller than 0.85, the smallest mean ratio found for the most stable simulated case. Once this stability-adjusted mean turbine wind speed ratio was calculated for a given time, the wind speed ratios at individual turbines were adjusted:

$$R_{stab} = \left(\frac{\bar{R}_{stab}}{\bar{R}} \right) R \quad (5-6)$$

where R_{stab} is the stability-adjusted wind speed ratio at a turbine, \bar{R} is the mean turbine ratio over all cases, and R is an unadjusted wind speed ratio at the turbine.

Figure 5-24 shows the effect of applying the stability correction on the mean absolute forecast errors. Even without the stability correction, the “Mean Ratios” forecast yielded significantly lower MEs and MAEs than the empirical plant power curve and wind tunnel methods. This seems logical since these are the exact time periods used for the calculation of the mean ratios. The addition of the stability correction (“Stability-modified” on the plot) further reduces both the MAE and ME, a very encouraging result. This suggests that the model is able to resolve important stability differences and their effect on the complex flow, even using the very simple stability function in Equation 5-5.

However, the value of the stability correction doesn't seem to extend outside the 60 hours of the high-resolution simulations. When the same bulk turbulence scale parameter was extracted from the much coarser 8 km simulations which cover the ten 24-hour simulations, the adjustment of turbine wind speed ratios results in a lower ME but the MAE increases from 8.6 to 9.0% (Figure 5-25).

The reason for this could be that the coarser simulations do not resolve the terrain well enough to correctly simulate the near-surface stability in a complex area such as the Altamont Pass. It is also likely that using high-resolution 100 meter simulations at night (either from 4:00 p.m. to 10:00 p.m. or 10:00 p.m. to 4:00 a.m. PST) means that the stability relationship derived from them doesn't handle the daytime hours as well – in other words, it's unavoidably biased toward the nighttime hours. A larger sample of simulations is needed to investigate the impact of broadening the range of daytime hours it is possible that an improved stability parameter might be found.

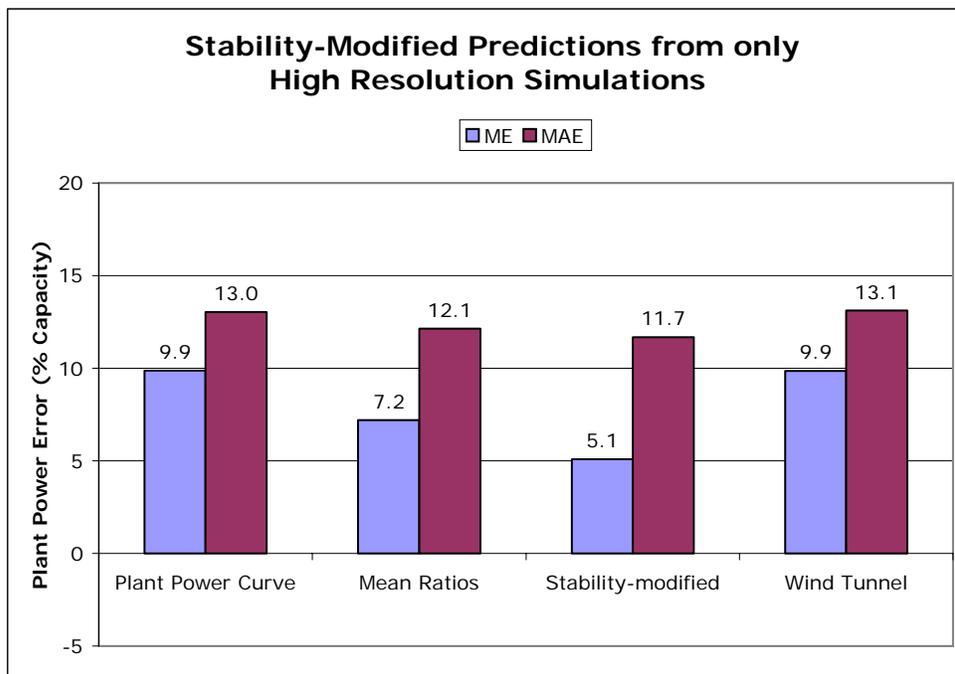


Figure 5-24 Comparison of four methods of predicting the plant power output using observed M127 wind speeds from the ten 6-hr periods simulated by the high-resolution numerical model. The “Plant Power Curve” and “Wind Tunnel” methods are the same as in previous plots. The “Mean Ratios” method is the same as “Simulations” in previous plots, and “Stability-modified” refers to predictions in which the turbine wind speed ratios were adjusted for stability.

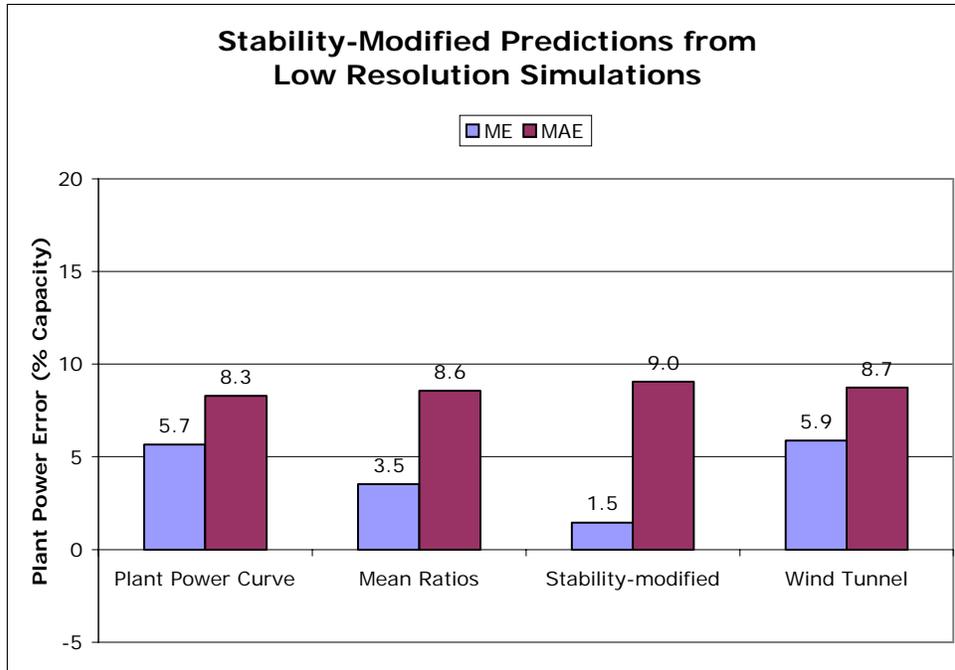


Figure 5-25 Comparison of four methods of predicting the plant power output using observed M127 wind speeds only from the ten 24-hr periods simulated by the model. The four methods are the same as in Figure 5-24.

Summary and Conclusions

A set of high-resolution, physics-based numerical simulations were used to simulate the wind speed distribution down to the level of individual turbines for the Tower M127 turbine cluster within the PowerWorks wind plant in the Altamont Pass of California. The 100 meter simulations contained 33 grid cells in the cluster, with one to five turbines falling in each grid cell. For a set of ten 6-hour simulations, the ratio of the wind speed at each turbine location to the speed at M127 was calculated at each time and averaged together to obtain a set of mean turbine wind speed ratios. These ratios were found to be quite different from those obtained from the UC Davis boundary layer wind tunnel. These numerical simulation-derived ratios were used to infer wind speeds at turbine locations given a wind speed at M127. The power production for each turbine was calculated from the manufacturer’s power curve, and the power at all the turbines in the cluster was combined to predict the total plant power output.

When this method was driven with two datasets of observed wind speeds covering either ten months or 3 years, its predictions yielded mean absolute errors (MAE) from 0.3 to 0.4% (of plant capacity) higher than those obtained through the use of an empirical (statistical) plant-scale power curve, and 0.1 to 0.3% higher than those obtained using the wind tunnel method. The mean error (ME) of the numerical simulation method however, was about 2% lower than either of the other two methods. When using ten months of forecasted wind speeds, the MAE’s of the three methods were almost indistinguishable, but the ME of the numerical simulation method was about 3% worse than the other two methods. In general, the overall performance of the

simulation and wind tunnel methods were very comparable to the empirical plant power curve, but the empirical plant power curve almost always had a slightly lower MAE.

The ten simulated cases were characterized by a wide range of stability in the lowest few hundred meters of the atmosphere. The high wind speed cases in July had relatively high stability. Some of the December cases were also very stable, probably a result of the presence of a wintertime cold stable air mass, and strong radiational cooling at the surface when skies were clear. One of the December cases and two of the May cases were not very stable, and under this near-neutral stability, it would be expected that the numerical simulations would behave more like the neutral stability flow simulated in the wind tunnel.

Averaging the turbine wind speed ratios over all cases has the effect of choosing a moderate value of stability. These mean ratios (mean ratio of 0.90) were much smaller than those obtained in the neutral wind tunnel (mean ratio of 0.97), because the inclusion of stable cases tends to lower the wind speed at many of the turbines due to terrain blocking and increased resistance to mixing down higher winds from aloft. Since the wind tunnel method always assumes neutral stability, one would expect it to systematically over-predict the wind speed at the turbine locations for cases that are on the stable side of neutral. However, 6 of the 93 turbines were not modeled in the wind tunnel domain and were neglected in the power production predictions produced by the wind tunnel method. This amounts to an implicit lowering of the turbine wind speed ratios. With that implicit adjustment, the wind tunnel method performed similarly to the other two methods, even though the density correction and wind direction factors were not included.

It was demonstrated that modifying the turbine wind speed ratios in a very simple way using a stability parameter calculated from the 100 meter simulation data improved the power prediction. The MAE of the stability-corrected power prediction was 1.3% of plant capacity lower than the other two methods, and the ME was 4% lower. This encouraging result is tempered by the fact that the same stability correction didn't help when applied outside the time frame of the 100 m simulations. In that test, the stability at the met tower was extracted from a much coarser 8 km simulations. The lack of improvement suggested that the 8 km runs do not resolve the meteorology of the Pass well enough to provide a useful source of stability information. One strategy would be to run operational simulations to as high a resolution as possible, perhaps 2 or 1 km – the stability from these higher-resolution runs may be sufficiently dependable to improve the power forecasting by applying a stability correction to the turbine wind speed ratios.

Another approach would be to use the understanding of the effect of stability on the M127 wind speed-plant power relationship to improve the empirical plant power curve. Perhaps a set of stability classes could be defined, and an empirical plant power curve could be derived for each class. Alternately, a more sophisticated statistical approach that included a relevant stability parameter as an input could be utilized.

6

HIGH-RESOLUTION WEATHER AND WIND FLOW FORECASTING

Introduction

Accurate forecasting of wind energy generation begins with an accurate forecast of the weather conditions in the region and accounting for local terrain, bodies of water, and other surface features that affect the wind speed and direction at the location of each wind turbine. The National Atmospheric Release Advisory Center (NARAC) at the Lawrence Livermore National Laboratory (LLNL) has the capability to generate accurate weather and wind speed and direction forecasts at very fine grid resolution using the latest meteorological models. As a result, LLNL joined the project team to generate high resolution wind speed and direction forecasts for use in the project.

The principal objectives of LLNL's involvement was to provide real-time wind speed and direction forecasts for use in development of improved wind energy forecasting algorithms and to support wind tunnel and numerical modeling of wind flow over complex terrain.

This section of the report describes LLNL's three-dimensional real-time weather forecast model, the design of the experiment to be conducted at the location with the complex terrain. The assessment of weather forecast accuracy would help quantify the source of wind energy forecast errors from the atmospheric forecast model and/or wind-tunnel module for further improvement in the wind energy forecasting system.

COAMPS Model and Experiment Design

A modified version of the Naval Research Laboratory's (NRL's) three-dimensional Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS), Version 2.0.15 was used in this study (Chin et al. 2000, 2001 and 2005). COAMPS consists of a data assimilation system, a nonhydrostatic atmospheric forecast model, and a hydrostatic ocean model.

In this study, we use only the data assimilation and the atmospheric model to provide real-time forecasts. The atmospheric forecast model is composed of a compressible form of the dynamics, nest-grid capability, and parameterizations of subgrid-scale turbulence, surface momentum and heat fluxes, explicit ice microphysics, subgrid-scale cumulus clouds, shortwave and longwave radiation, and urban canopy physics. The terrain-following vertical coordinate is also used to simulate airflow over an irregular surface. The model terrain is given from one-km resolution

terrain database using Silhouette average method with filter. Hodur (1997) provides additional details of the COAMPS model, and Chin et al. (2005) address newly-implemented physics.

The model domain contains 31 grid points in the vertical direction, with the grid size varied to maximize resolution at lower levels. The vertical consists of nine grid points below 127-meter elevation, with the grid spacings of 4, 4, 4, 6, 10, 16, 24, 34 and 50 meters, starting at ground level. The grid spacing aloft gradually increases to 800 meter at 3.152-km altitude. Above this level, the grid size is uniformly set at 800 m up to 7.592-km altitude. Then, the grid size gradually increases to 5.0 km at 24.352-km altitude.

In the horizontal direction, a total of three nested domains are used. Both zonal and meridional coordinates have 61 grid points for all nested grids. A uniform grid size of 36 km is used for the outer coarse grid (nest_1), the grid sizes of the inner grids are each one third of the size of the previous grid, for example, the middle grid size is 12 km (nest_2), and the inner grid size is 4 km (nest_3). with a constant size ratio of three to define the inner nest grids.

The time series of the forecasts use time steps of 90 and 45 seconds for non-sound and sound wave calculations, respectively. The time steps for the finer-grid domains are reduced in proportion to the nest-grid size ratios. The rigid boundary condition is imposed at the vertical boundary. A sponge-damping layer is placed above 10.052 km to minimize the reflection of internal gravity waves off the rigid upper boundary. The Davies (1976) boundary condition is applied to the lateral boundaries with a nudging zone of seven grid points at each lateral boundary. A time filter with a coefficient of 0.2 is applied to control computational instability associated with the leapfrog time approximation in the model.

In this study, two watches of 48-hour forecasts (00Z and 12 Z, respectively) are performed daily over California for a grid centered at the Altamont Pass and during the 12 months from July 2004 to June 2005. However, due to the size limit of huge forecast data storage, only the nested-grid data for the first week of each month are stored and used to assess the forecast errors with respect to the measurements at 11 available tower observations. Nonetheless, the yearly forecast data of the finest grids were stored at UC Davis for a separate study to evaluate the wind energy forecast errors. In this study, the forecast errors are measured by the mean absolute errors of the wind speed and direction forecast vs. the observed data for each forecast hour, which avoids the self-canceling effect of under-and over-prediction present in the mean forecast error.

Measurements for Model Validation

A total of 11 meteorological tower stations at the Altamont Pass are used to calculate the mean absolute forecast errors for wind speed and direction. These observations are available in 30-minute averages and Table 2-1 shows the coordinates and the heights of the met tower measurements for the 11 towers. All tower measurements occur at 18-meter elevation, except those for towers 225 and 438 at 24 meters, and tower 832 at 30 meters. The model forecast is interpolated to the same location as the observations in both the horizontal and vertical (height above the ground) directions to compute the absolute forecast errors.

Quality assurance of tower measurements shows great uncertainty in the wind direction observations. In particular, the measured wind direction at station 624 reported a constant value of $0.3 \sim 0.5$ throughout the year of interest. In the mean time, five other stations, Nos. 427, 821, 826, 832, and 922, exhibit a small range of fluctuation (few degrees) with respect to a wind angle around a few tens of degrees. This category of wind direction measurements may arise from the channeling effect of the local terrain, and does not necessarily imply poor quality. All of six met towers are designated as uncertain (UNC) stations, and the remaining five stations with reliable wind direction measurements are referred to as REL stations in this study. The separation of reliable and uncertain stations is also used to gauge its impact on the forecast errors.

Results

The mean absolute errors of forecast wind speed and direction were derived for each month using the forecast and observed data for each forecast hour throughout the 48-hour forecast period, and averaging over the available met towers and 14 weekly forecasts (two forecasts per day, seven days per week). Normally, the mean absolute errors were calculated using the forecasts for the first week of each month. Occasionally, the forecast period was shifted several days to accommodate the availability of station measurements.

The forecast errors for a warm month (June 2005) clearly exhibit the dependence of wind speed error on the grid resolution using the data from all stations (Figure 6-2a), but the wind direction error does not show the same dependence (Figure 6-2b). The observed station wind speed of the warm month sometimes reaches 20 m/s. The absolute error of forecast wind speed is about six m/s with the coarser grid resolution (36 km) and can decrease to four m/s in the higher resolution forecast (4 km). With the separation of UNC and REL stations, the impact of grid resolution on wind speed errors remains unchanged (Figures 6-2c and 6-2e). In contrast, the wind direction error is significantly reduced with increasing grid resolution using REL stations, while the error is fairly large at UNC stations and shows an opposite dependence on the grid resolution (Figures 6-2d and 6-2f).

In contrast, Figure 6-3 presents the forecast errors for a cold month (December 2004), which show different patterns relative to that of the warm-months. Unlike the warm months, there is no clear dependence of forecast errors on grid resolution in the cold month, even with the separation of REL and UNC stations. In addition, the magnitude of wind speed errors is noticeably reduced in the cold month as a result of weaker winds (< 10 m/s, except for the storm periods). Although the resolution impact is weak during the cold months, the separation of UNC and REL stations still exhibits qualitative improvement in the forecast error for both wind speed and direction. Therefore, only the results from the REL stations are shown for the rest of months under investigation to gauge a complete seasonal variation of forecast errors.

Figures 6-4, 6-5, 6-6, and 6-7 illustrate the detailed month-to-month variations of the forecast errors using the measurements from only REL stations. A noticeable change of grid resolution impact on the wind speed error appears at two transition times; (1) during October and November, and (2) during March and April. However, this pattern is not strong for the wind direction errors during the warm months before October 2004, most likely because the COAMPS physics changed in September 2004 with the consideration of soil moisture impact on the re-

distribution of surface latent and sensible heat fluxes, and the resulting development of planetary boundary layer. As a result, this new physics act to improve the surface energy budget and its related surface temperature and wind forecast. Therefore, a clear grid resolution impact on the wind direction starts to appear in the warm months after October 2004. This new physics also reduces the forecast errors of wind speed in the warm months.

As shown in Table 6-1, the finest resolution (4-km in the nest_3 domain) used in the simulations is still not fine enough as the resolved model terrain heights in the nest_3 domain at most of the measurement stations substantially differ from the actual terrain heights. The LLNL terrain database indicates that the model needs at least one more nest to properly represent the local terrain forcing to reduce the forecast errors.

To illustrate the impact of inaccurately modeled terrain forcing, Figure 6-8 shows the vertical profiles of forecast horizontal winds during warm and cold months. The evolution of these forecast winds exhibits strong nocturnal winds at the Altamont Pass. The vertical wind shear in the lowest 100 meters is also large, particularly during the warm months. To estimate the quantitative impact of the modeled terrain on forecast winds, Figure 2-9 compares the forecast errors for June 2005 and December 2004 with the terrain calibration, i.e., comparing forecast vs. measured wind at the actual station height. As expected, the main differences appear during the warm months, particularly when the wind speed error exceeds one m/s for each grid resolution. This suggests that simply improving the modeled terrain geometry is not sufficient to reduce the forecast errors.

The yearly forecast results clearly indicate that the forecast errors over the Altamont Pass area can be characterized by a semi-annual variation. In the warm months when the synoptic-scale front activity is weak, the model exhibits a strong grid resolution impact on the forecast accuracy as a result of the improved representation of local terrain with increasing grid resolution. In contrast, the cold months coincide with the prevailing frontal activity so that the local terrain becomes a secondary forcing for the model forecast. Therefore, there is no clear dependence of forecast errors on grid resolution during the cold months. In addition, the large forecast error of wind direction in the cold months may arise from the improper model terrain geometry.

Summary

NARAC's real-time wind prediction system, COAMPS, was used to generate real-time 48-hour forecasts of wind speed and direction at three grid resolutions, 36, 12, and 4 km, and each grid was centered at a location near the complex terrain of Altamont Pass. The forecasts were generated twice daily at 00Z and 12Z) for one year starting from July 1, 2004. The model outputs were used to support development of the next day wind energy forecasting algorithms and numerical modeling of wind flow over the wind turbines associated with specific met towers and the resulting wind plant power curve (described in Chapters 4 and 5 of this report, Volume 2), and wind tunnel modeling and wind energy forecasting at University of California at Davis (described in Volume 3). However, only the validation of COAMPS forecast is presented in this report. The evaluation of the overall wind energy forecasting system is covered by UC Davis in a separate report.

The month-to-month variation of wind forecast errors clearly exhibits a semi-annual fluctuation with prominent dependence on the grid resolution in the warm months (i.e., strong wind power period) when the frontal activity is weak; the large forecast errors are systematically reduced with increasing grid resolution for both wind speed and direction. However, this dependence diminishes when synoptic-scale frontal activity prevails in the cold months.

The remaining question to be addressed from this research outcome is whether the grid resolution dependence would continue with decreasing grid size or if this dependence tendency converges at a certain grid size in the strong wind power period. Although the increasing grid resolution can resolve a better representation of model terrain geometry (magnitude and shape), further study of Silhouette terrain representation and the use of finer resolution terrain database are highly recommended to improve the forecast accuracy for the airflow over the complex terrain.

Due to the computational limitation of a shared memory model, it becomes infeasible to meet the real-time forecast requirement for simulations with further increased resolutions. The implication of this research strongly supports the value of high-performance computing to further improve the wind energy forecast. To this end, NARAC is in transition to convert COAMPS to the distributed memory code to allow faster and higher resolution real-time forecast. During recent testing of MPI (message passing interface), COAMPS demonstrated its readiness for the operational use later in 2005.

Acknowledgments

The author wishes to thank Electric Power Research Institute for providing tower measurements to validate the accuracy of model forecast. This work was supported by the California Energy Commission Wind Energy Forecasting Project (Contract No. E2IP219C7619) and the U.S. Department of Energy (DOE) National Advisory Release Assessment Program at LLNL, and conducted under the auspices of the DOE by the University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Table 6-1 Station information for meteorological towers at Altamont Pass. Station and modeled terrain heights are also shown for different grid resolutions.

Station	Sta_ID	Sta_lon(°)	Sta_lat(°)	Sta_H(m)	36km_H	12km_H	4km_H	1.33km_H	0.44km_H
1	127	-121.6927	37.764416	335.28	102.23	175.04	175.81	295.98	335.33
2	225	-121.6673	37.767766	312.42	101.32	179.34	159.22	254.75	256.50
3	427	-121.6285	37.717583	342.90	108.99	239.48	238.27	276.08	302.95
4	438	-121.6734	37.711416	415.75	110.25	225.73	223.38	272.39	304.52
5	624	-121.6836	37.729388	347.90	107.70	203.52	190.67	260.13	269.08
6	723	-121.6357	37.788916	198.12	98.47	166.03	119.07	122.31	147.58
7	821	-121.5997	37.701777	237.74	112.20	264.60	278.24	289.95	278.00
8	826	-121.5798	37.675527	419.10	116.72	300.48	349.88	394.08	460.91
9	832	-121.6141	37.680555	441.96	114.75	288.87	340.08	434.68	536.57
10	922	-121.6073	37.759777	121.92	103.84	193.14	145.17	117.46	125.04
11	926	-121.6316	37.738555	242.32	105.69	213.53	191.03	221.19	217.46

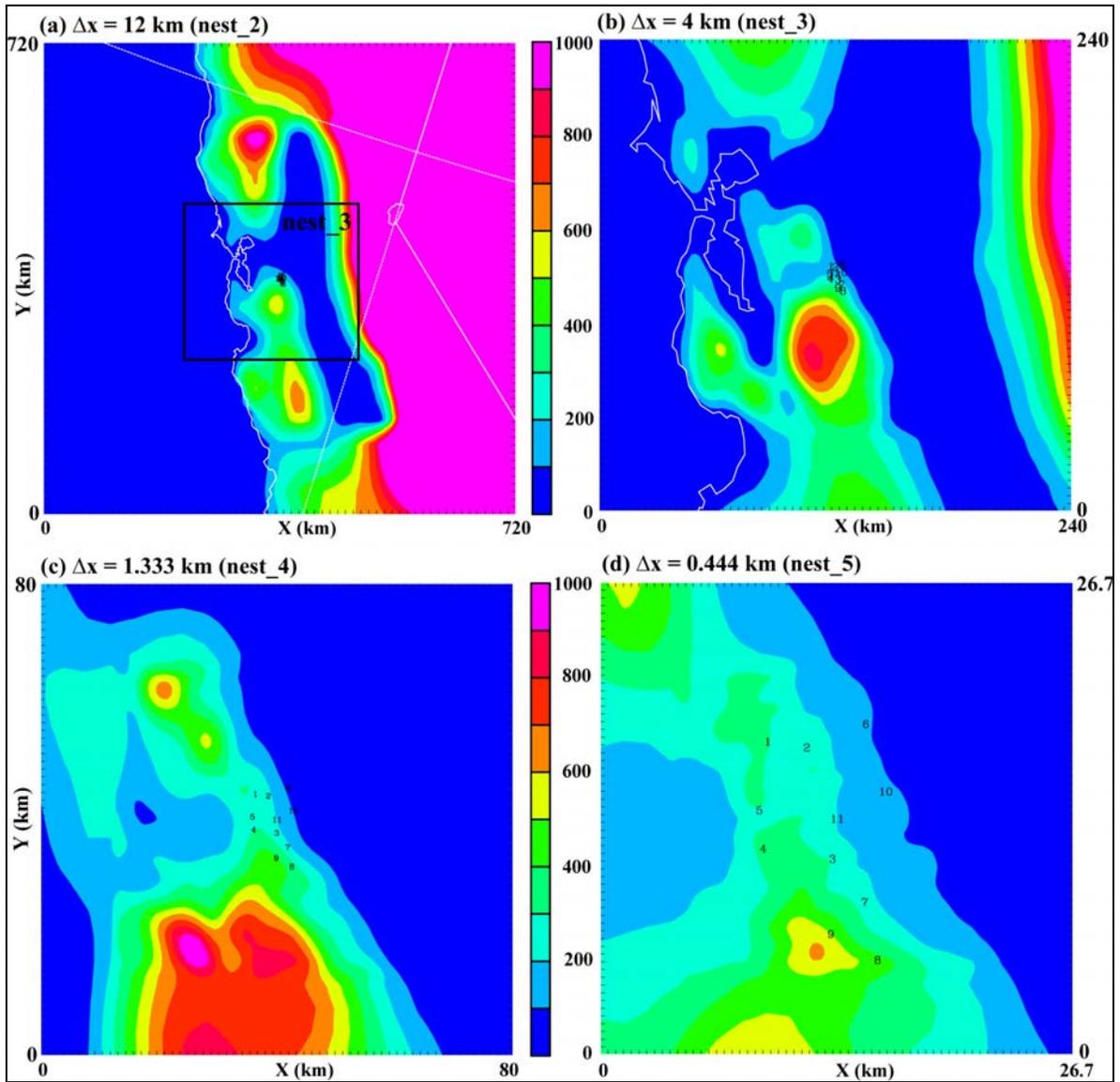


Figure 6-1 Resolution of terrain elevation (meters) vs. grid size (km) of the nested domain: (a) $\Delta x = 12$ km (nest_2), (b) $\Delta x = 4$ km (nest_3), (c) $\Delta x = 1.333$ km (nest_4), and (d) $\Delta x = 0.444$ km (nest_5). The letters mark the locations of the met towers used.

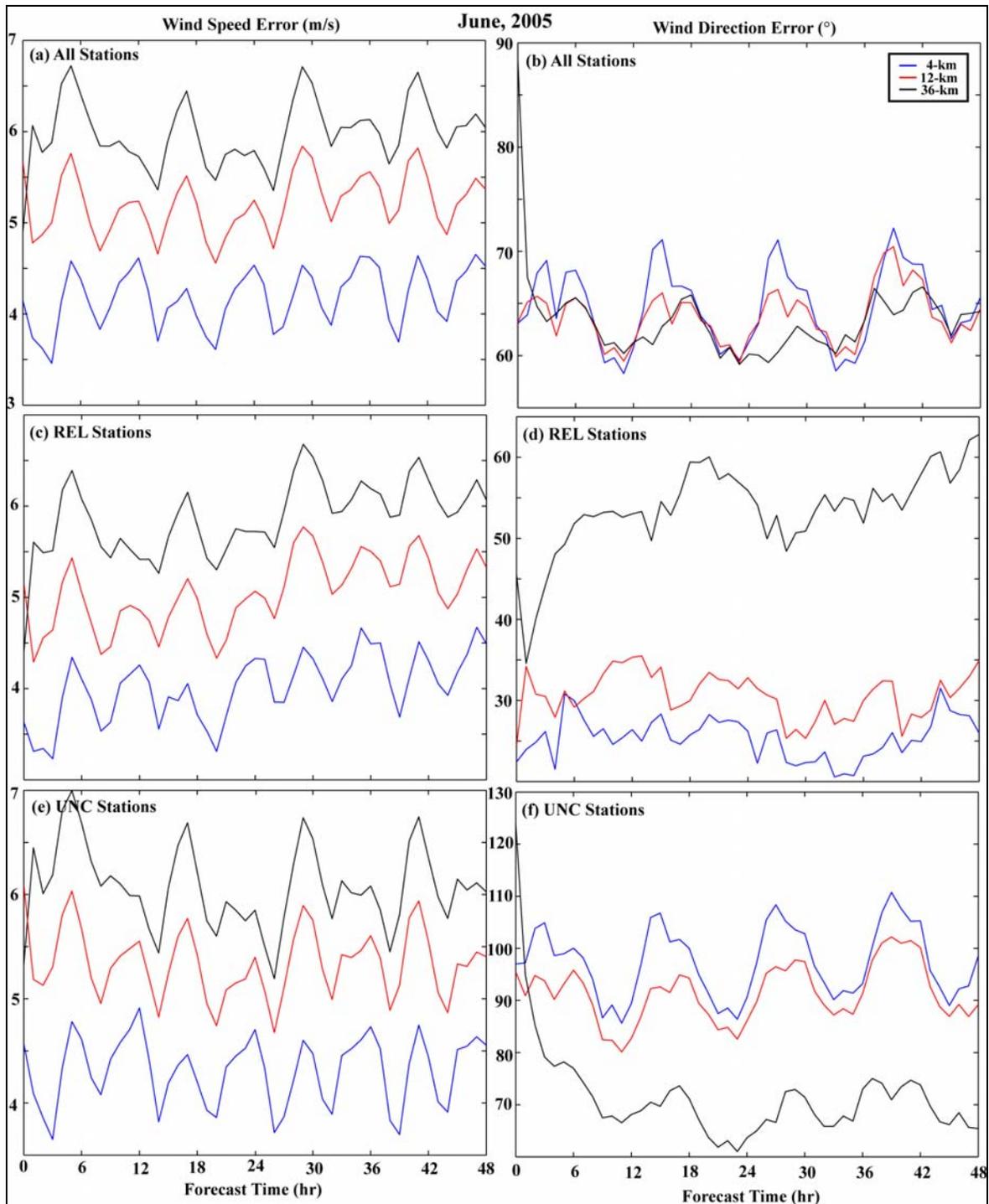


Figure 6-2 Weekly mean absolute errors of NARAC forecasts averaged over the selected stations for June 1-7, 2005. The colored lines represent the results for different horizontal resolutions (36, 12, and 4km, respectively). The left panels present MAEs for wind speed forecasts, and the right panels for wind direction. The top panels (a and b) are the forecast errors using the measurements from all stations, the middle ones (c and d) using 5 reliable stations, and the bottom plots (e and f) using uncertain measurements from the remaining 6 stations.

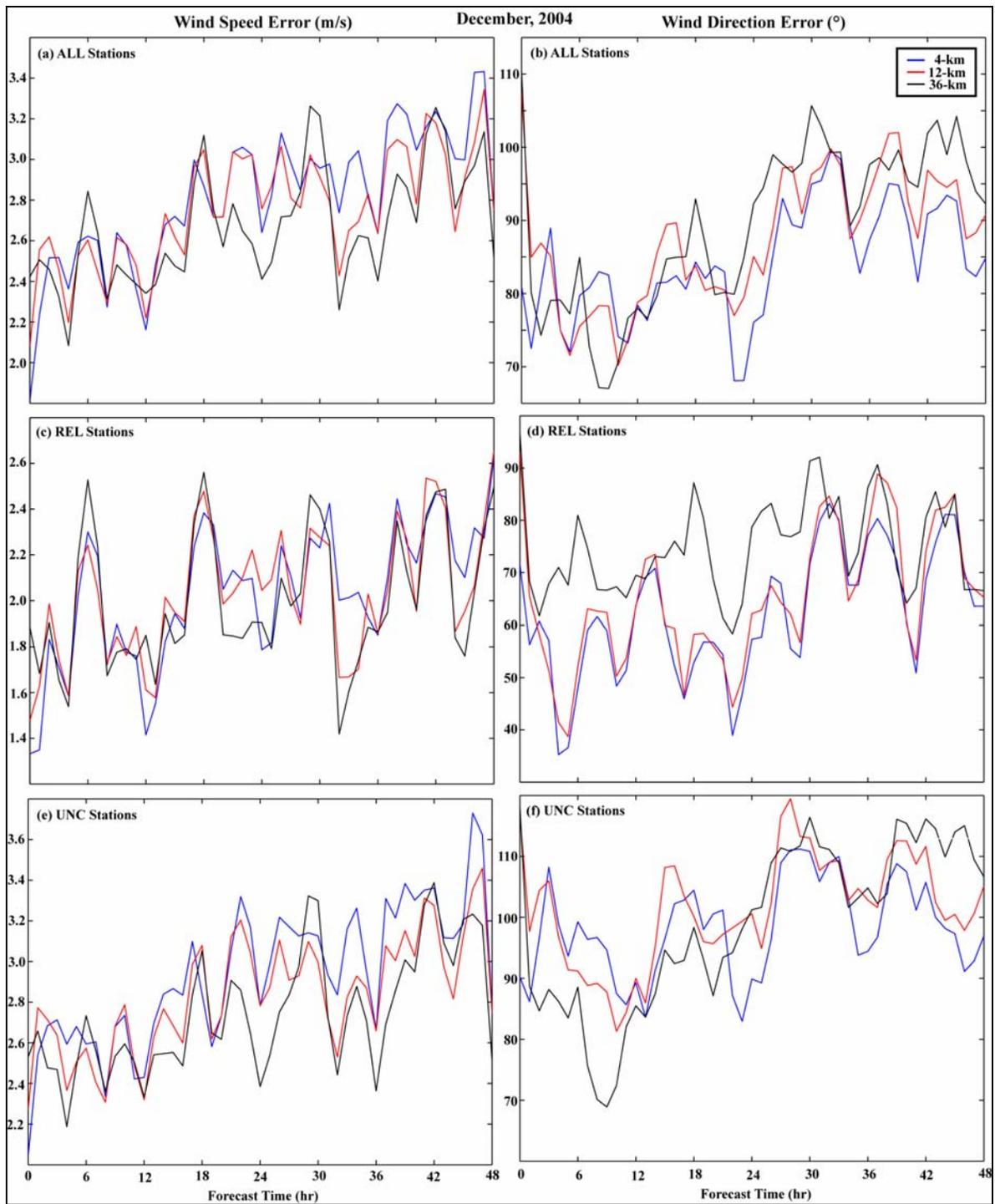


Figure 6-3 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts averaged over the selected stations for December 2004.

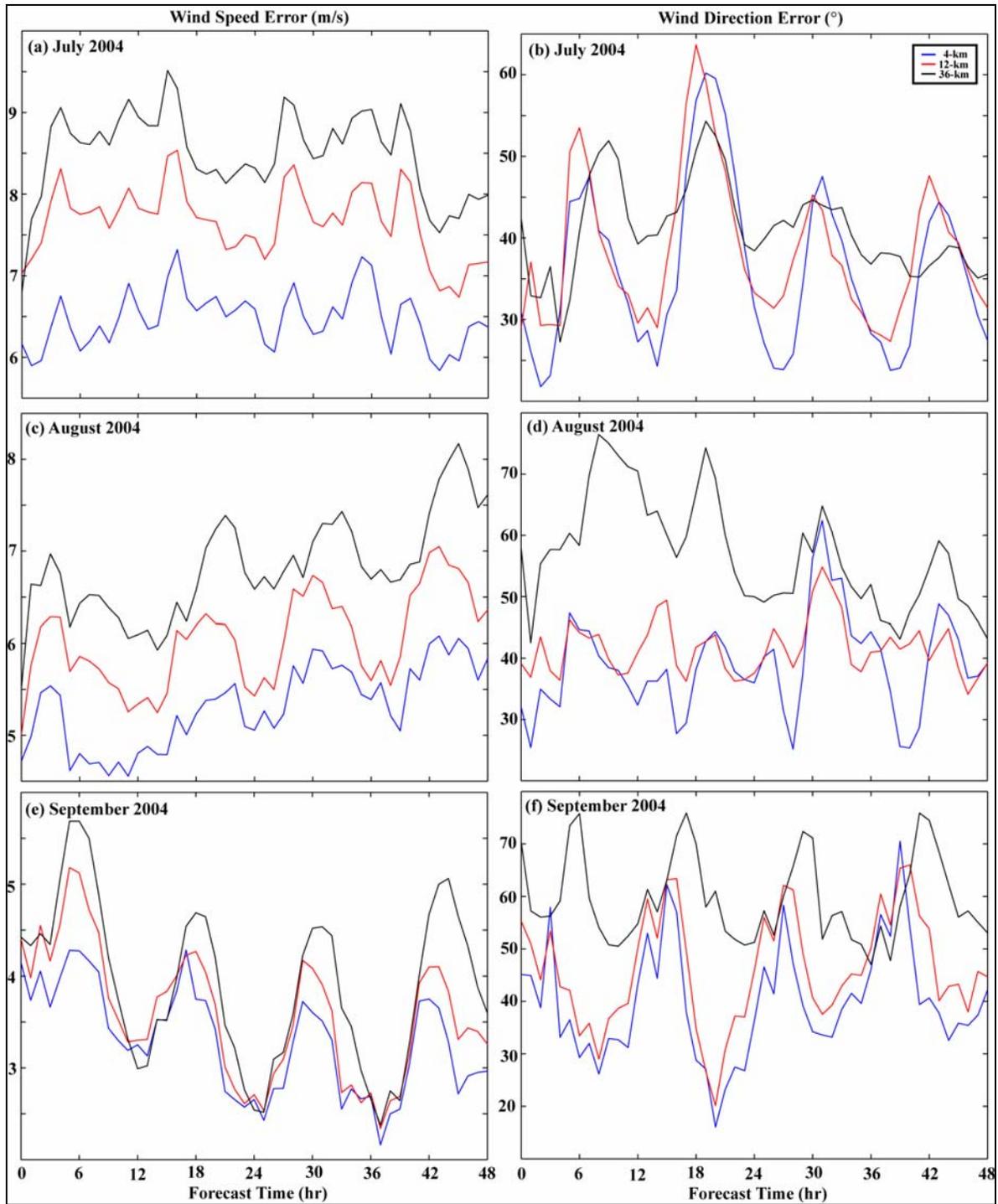


Figure 6-4 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts for July 2004 (a and b), August 2004 (c and d), and September 2004 (e and f), respectively.

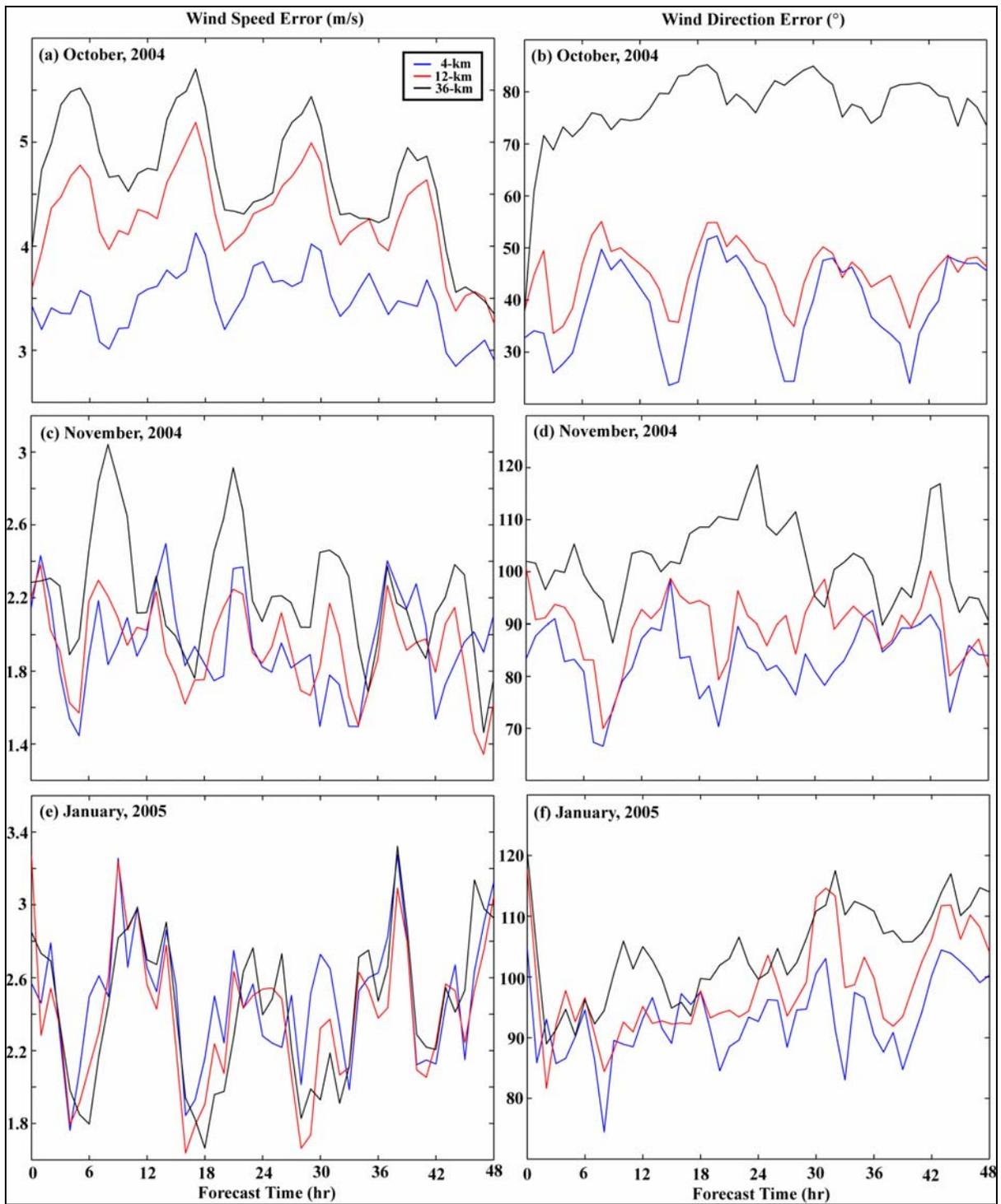


Figure 6-5 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts for October 2004 (a and b), November 2004 (c and d), and January 2005 (e and f), respectively.

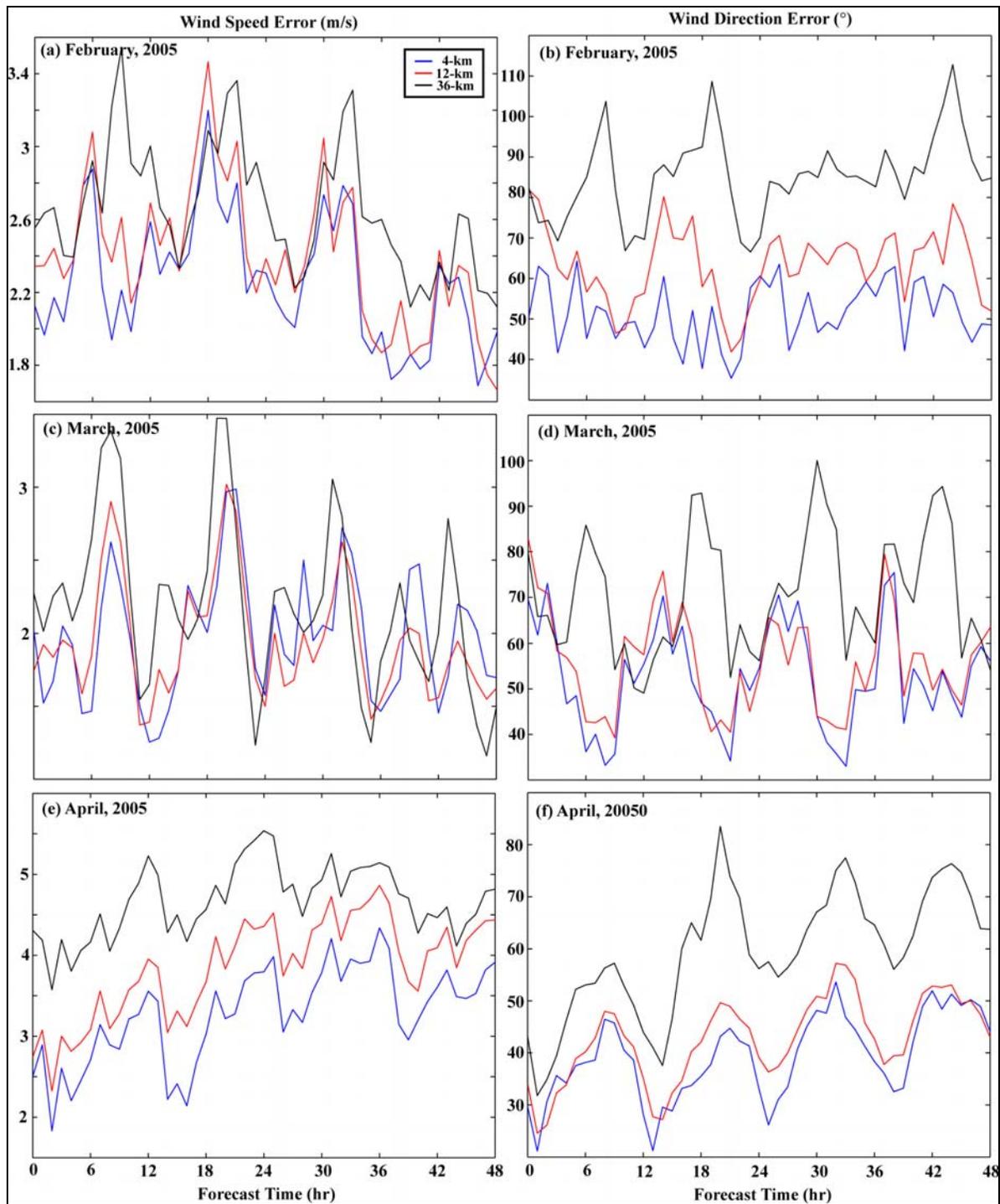


Figure 6-6 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts for February 2005 (a and b), March 2005 (c and d), and April 2005 (e and f), respectively.

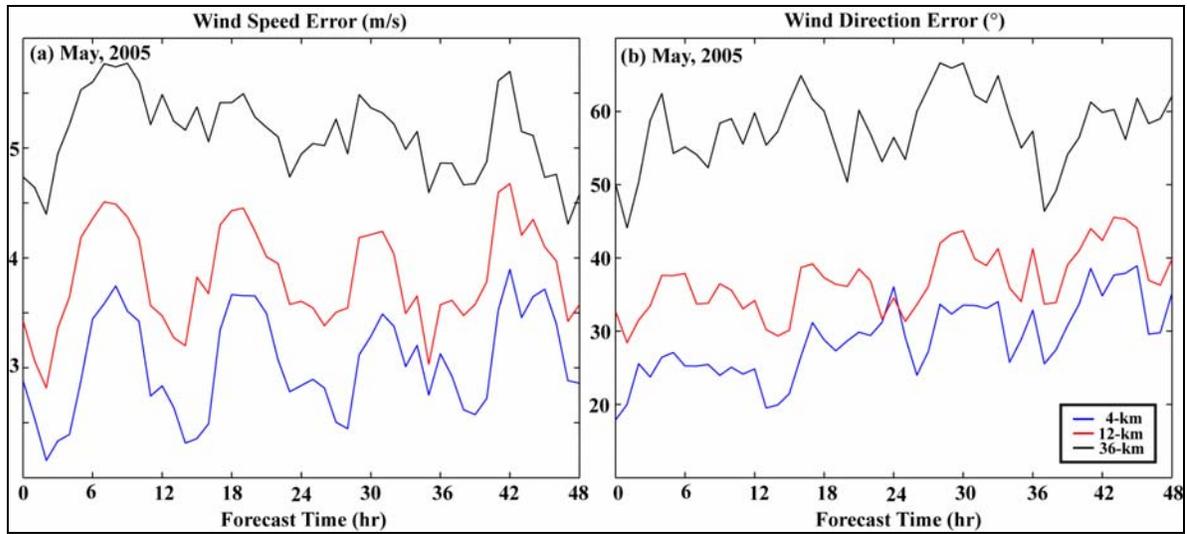


Figure 6-7 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts for May 2005.

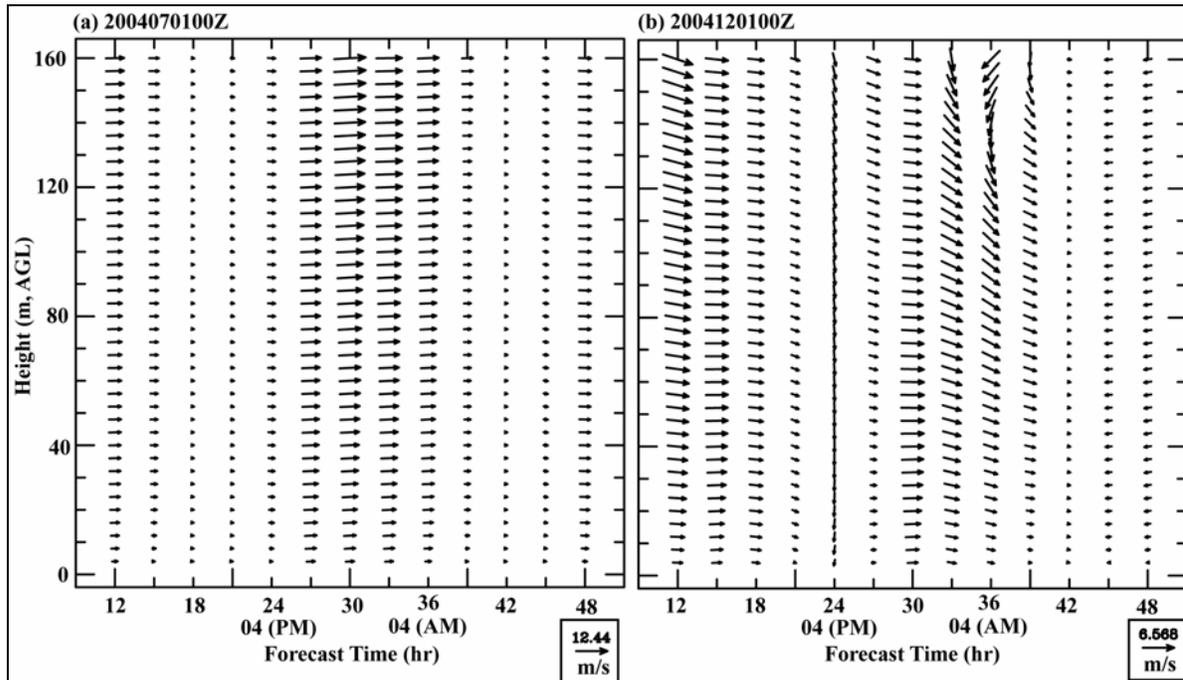


Figure 6-8 Vertical profiles of forecast horizontal winds at station 127 from the nest_3 domain ($\Delta x = 4$ km), (a) from 00Z of July 01, 2004, and (b) from 00Z of December 01, 2004.

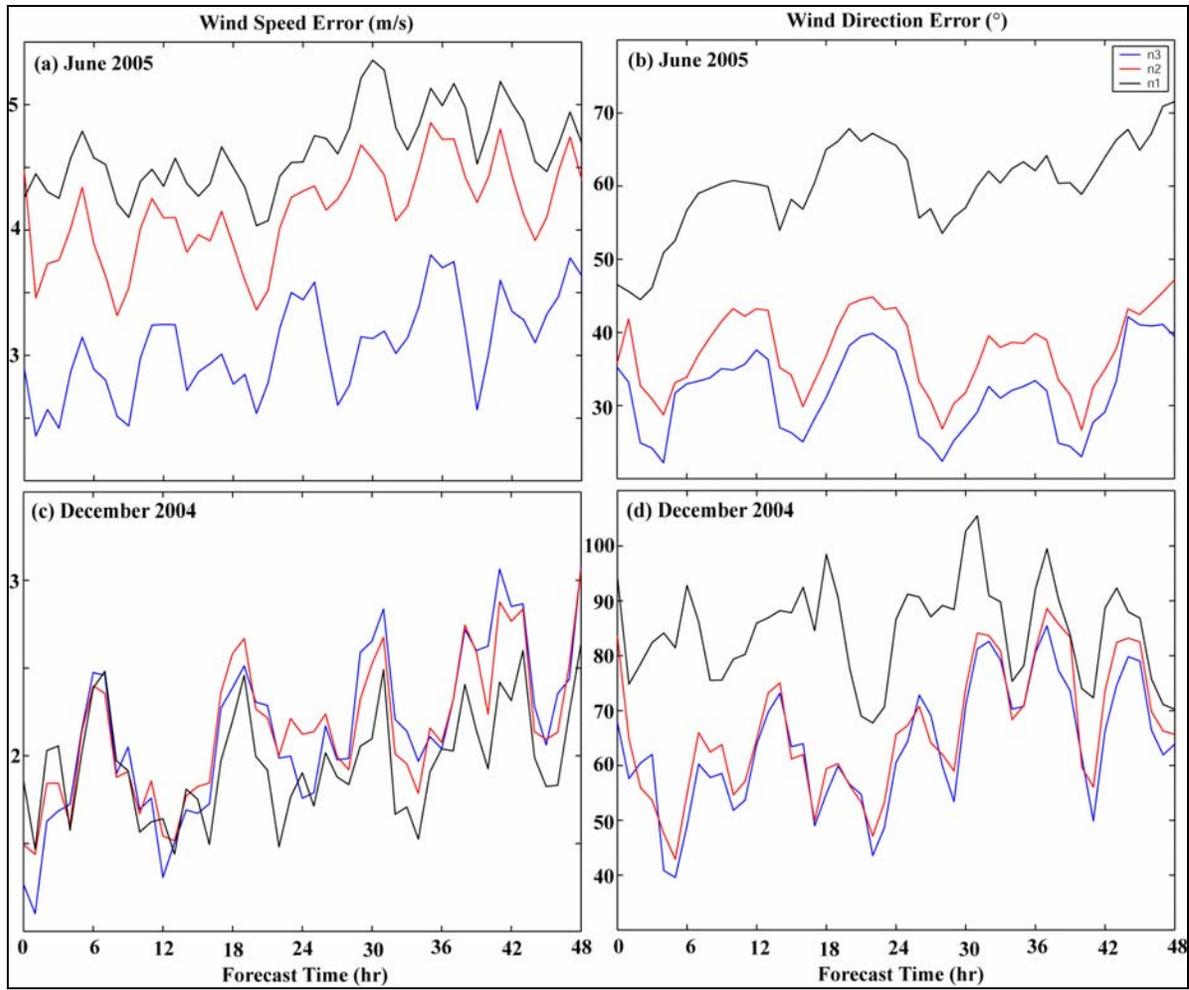


Figure 6-9 Weekly mean absolute errors of NARAC wind speed (left) and wind direction (right) forecasts with calibration of modeled terrain height for June 2005 (a and b) and December 2004 (c and d).

7

CONCLUSIONS AND RECOMMENDATIONS

The overall project objective was to further develop technology that would ultimately improve the accuracy of wind power production forecasts on both the short-term and day-ahead time scales.

Within this broad objective, the specific project objectives were:

1. Design a short-term forecast system that could produce 5-minute forecasts of wind power production for the next three hours with an update frequency of 5 minutes and conduct an initial performance assessment of the system;
2. Assess the potential improvement in day-ahead power production forecasts that could be achieved using new datasets, forecast techniques, and other enhancements, and test the performance of a modified forecast system over a one-year period for a set of wind plants representing the major wind power generation regions of California;
3. Evaluate the potential forecast benefit of improved plant-scale power output models constructed from data generated by very high resolution numerical simulations and simulations in a boundary layer wind tunnel; and
4. Create a California wind generation Research Database (CARD) that consists of one year of simulated values of meteorological variables relevant to wind generation research and planning on a high resolution grid.

Conclusions

All four of these objectives were achieved, and significant advances were made in the understanding of both short-term and day-ahead forecasting issues.

Short-Term Forecast System Design and Evaluation

The first objective was addressed by designing a two-stage forecast system based upon the artificial neural network (ANN) method.

The first stage of the system consists of a mini-ensemble of three different forecast methods that exploit different input datasets and predictive tools. The second stage of the system is composed of an artificial neural network that weights each of the three forecasts from the first stage according to their recent performance characteristics and creates an “optimal” composite forecast.

A limited version of this forecast system was tested with regional power production data for 2004 provided by the CA ISO. This version utilized only one method, an autoregressive approach, in the first stage of the system. This limited version of the system demonstrated considerable skill relative to a persistence forecast during the warm season, but virtually no skill during the cold season. The lack of skill during the cold season was expected, and it is anticipated that the other two forecast methods included in the first stage of the system will provide most of the forecast skill during the cold season.

Next-Day Forecast Model Improvements

The second objective was addressed in two phases.

The first phase conducted a screening evaluation of the potential of six enhancements of the forecast algorithm to improve forecast performance via a series of forecast experiments using a subset of the data from the previous Energy Commission-EPRI forecasting project. Several of the modifications produced substantial improvement in forecast performance and others appeared to have little impact on forecast performance. However, in some of the experiments, there were circumstances that raised questions about the representativeness of the results. Therefore, a broad set of the enhancements was implemented in the revised *eWind* forecast system.

The second phase evaluated the performance of the revised *eWind* forecast system based on one-year of daily 1- to 48-hour forecasts at each of five California wind plants. Two of the five wind plants had also participated in the previous Energy Commission-EPRI project, which enabled a direct comparison of forecast performance for the same wind plants. Almost all of the annual statistics for these two plants showed substantial improvement between the previous to the current project. In addition, an analysis of the forecast errors in this project provided considerable insight into the factors which are responsible for the variations in forecast performance vs. wind plant, season, weather regime and other factors. This analysis provides excellent guidance for decisions about the future direction of research and development efforts to improve wind power production forecast performance.

Improved Plant-Scale Power Curve Models

The third objective was addressed by executing a set of high-resolution, physics-based numerical simulations to model the wind speed distribution down to the level of individual turbines. The modeling was based on the Tower M127 turbine cluster at the PowerWorks wind plant in the Altamont Pass of California.

For a set of ten 6-hour simulations, the ratio of the wind speed at each turbine location to the speed at M127 was calculated at each time and averaged together to obtain a set of mean turbine wind speed ratios. These numerical simulation-derived ratios were used to infer wind speeds at turbine locations given a wind speed at M127. The power production for each turbine was calculated from the manufacturer's power curve, and the power at all the turbines in the cluster was combined to predict the total plant power output.

In general, the overall performance of these simulation-derived plant power predictions were very comparable to those produced by the empirical plant power curve, but the empirical plant power curve almost always had a slightly lower mean absolute error. Similar predictions based on results from the UC Davis boundary layer wind tunnel were also quite comparable.

The ten simulated cases represented wide range of atmospheric stabilities in the lowest few hundred meters of the atmosphere. It was demonstrated that modifying the turbine wind speed ratios in a very simple way using a stability parameter calculated from the 100 meter simulation data improved the power production forecast performance.

The resulting mean absolute error of the stability-corrected power prediction was 1.3% of plant capacity lower than the baseline method, and the mean error was 4% lower. It may be feasible to improve plant power predictions by running operational simulations to as high a resolution as possible, perhaps at one- or two-kilometer grid size. The atmospheric stability levels indicated by these higher-resolution runs may be sufficiently dependable to apply the stability correction to the turbine wind speed ratios and improve the power-production Forecast performance.

California Wind Generation Research Dataset (CARD)

The fourth objective was addressed by compiling the California wind generation Research Dataset (CARD). The CARD database was generated from the physics-based numerical simulations that generated the year of daily 48-hour power production forecasts related to the second objective.

The CARD database consists of hourly data for hourly wind speed, direction, power density, ambient temperature, and water/air mixing ratio at multiple levels for the one-year period, July 1, 2004 to June 30, 2005, over two grids with grid cell sizes of 5 km. One grid is centered near the Altamont and Solano wind resource areas of Northern California. The other grid is centered to the east of the Los Angeles Basin and includes the Tehachapi and San Geronio wind generation regions.

Recommendations

The project accomplishments provide a foundation for both further improvements of wind power production forecast systems and implementation of the improved forecast methods and the resulting forecasts by system operators and other decision-makers in an effective manner.

Recommendations for further research and implementation:

1. Implement and conduct a one-year test of the full two-stage short-term (0 to 3 hour) forecast system concept developed in the project, including the two forecast methods and the second-stage ANN model that were not implemented and tested in current project.

It is anticipated that the full system will provide significantly better forecast performance than the single forecast method tested in this project.

2. Address the issues associated with the operational implementation of the short-term forecast system at CA ISO.

The sequential steps might include (1) obtain feedback about the needs and desires of the anticipated operational users of the short-term forecast products at CA ISO via a meeting with the operational users to inform them of the characteristics of the short-term forecast products that could be made available and to get their feedback on how, where and when they should be interfaced into the operational environment; (2) develop and implement a real-time operational test in which potential users would be able to access the forecast data within the operational environment; (3) after the test period, obtain feedback from the users to provide bases for further refinement of the system; (4) design and implement a revised system that better meets the needs of the user.

This entire process can be conducted in parallel with item (1) above.

3. Conduct another development-evaluation cycle for the day-ahead (1- to 48-hour) forecast algorithm.

There are a number of new remotely-sensed data that will soon become available and have the potential to improve day-ahead as well as short-term forecasting. The use of these datasets should be explored in the next forecast system development-evaluation cycle in conjunction with the results of an extended analysis (beyond what was done in this project) of the errors in the forecasts produced in this project.

4. Formulate forecast performance metrics that provide a better measure of the power production forecast performance characteristics that are most critical to users of the forecast data.

This will likely vary from user to user. However, it is important to define the ranges of forecast characteristics that are important to users and to design metrics to evaluate them. For example, recent discussion in wind forecasting sessions at wind industry meetings has indicated the forecasts for periods in which there are large upward ramps in the wind power production are particularly critical to many users. A metric that isolates the performance during those periods may be of greater value to some users than generic metric such as the overall MAE or skill score used in this project. A broad investigation of this issue is needed.

8

REFERENCES

1. American Wind Energy Association (AWEA), Global Market Report 2004a. <http://www.awea.org/pubs/documents/globalmerket2004.pdf>, 2004.
2. American Wind Energy Association (AWEA), Wind Energy potential 2004b. http://www.awea.org/faq/tutorial/wwt_potential.html, 2004.
3. Archer C. L., and M. Z. Jacobson, 2005: Evaluation of global wind power. *J. Geophys. Res.*, (inpress).
4. Bacon, D. P., N. N. Ahmad, Z. Boybeyi, T. J. Dunn, M. S. Hall, P. C. S. Lee, R. A. Sarma, M. D. Turner, K. T. Waight, S. H. Young, and J. W. Zack, 2000: A Dynamically Adapting Weather and Dispersion Model: The Operational Multiscale Environment model with Grid Adaptivity (OMEGA). *Mon. Wea. Rev.*, 128, 2044-2076
5. Box, G.E.P and F.M. Jenkins. *Time Series Analysis: Forecasting and Control*, second edition. Oakland, CA: Holden-Day, 1976
6. Brown, O.B., and P.J. Minnett, 1999, MODIS Infrared Sea Surface Temperature Algorithm Theoretical Basis Document, Ver 2.0. Available at: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod25.pdf
7. CA ISO, 2005, Hawkins, David Wind Generation Forecasting: A Balancing Authority View, presentation to Utility Wind Integration Group Technical Workshop, Sacramento, CA, November 2005.
8. California Energy Commission and Electric Power Research Institute (EPRI), 2003a: California Wind Energy Forecasting System Development and Testing Phase 1: Initial Testing, EPRI Palo Alto, CA, California Energy Commission, Sacramento, CA: 2003. 1003778.
9. California Energy Commission and Electric Power Research Institute (EPRI), 2003b: California Wind Energy Forecasting System Development and Testing Phase 2: 12-Month Testing, EPRI Palo Alto, CA, California Energy Commission, Sacramento, CA: 2003. 1003779.
10. California Energy Commission and Electric Power Research Institute (EPRI), 2006a: California Regional Wind Energy Forecasting System Development, Volume 1: Executive

References

- Summary, EPRI Palo Alto, CA, California Energy Commission, Sacramento, CA: 2006. 1013262.
11. California Energy Commission and Electric Power Research Institute (EPRI), 2006b: California Regional Wind Energy Forecasting System Development, Volume 2: Wind Energy Forecasting System Development and Testing and Numerical Modeling of Wind Flow over Complex Terrain, EPRI Palo Alto, CA, California Energy Commission, Sacramento, CA: 2006. 1013263.
 12. California Energy Commission and Electric Power Research Institute (EPRI), 2006c: California Regional Wind Energy Forecasting System Development, Volume 3: Wind Tunnel Modeling of Wind Flow over Complex Terrain, EPRI Palo Alto, CA, California Energy Commission, Sacramento, CA: 2006. 1013264.
 13. California Energy Commission and Electric Power Research Institute (EPRI), 2006d: California Regional Wind Energy Forecasting System Development, Volume 4: California Wind Generation Research Dataset (CARD), EPRI Palo Alto, CA, California Energy Commission, Sacramento, CA: 2006. 1013265.
 14. Chin, H.-N. S., M. J. Leach, and M. J. Brown, 2000: A preliminary study of the urban canopy effects on a regional-scale model: Sensitivity assessment of an idealized case. Proceedings of the third Symposium on the Urban Environment, July 14-18, 2000, Davis, California, Amer. Meteor. Soc., 76-77.
 15. Chin, H.-N. S., M. J. Leach, G. Sugiyama, and F. J. Aluzzi, 2001: A Preliminary Study of Surface Temperature Cold Bias in COAMPS. Proceedings of the Ninth Conference on Mesoscale Processes, 30 July -2 August 2001, Fort Lauderdale, Florida, Amer. Meteor. Soc., 249-254.
 16. Chin, H.-N. S., M. J. Leach, G. A. Sugiyama, J. M. Leone Jr., H. Walker, J. S. Nasstrom, and M. J. Brown, 2005: Evaluation of an urban canopy parameterization in a meso-scale model using VTMX and URBAN 2000 data. *Mon. Wea. Rev.*, 133, 2043-2068.
 17. E.ON Netz, 2004, Wind Report 2004, E.ON Netz, Bayreuth, Germany 2004.
 18. EPRI, 2003: Wind Energy Forecasting Applications in Texas and California: EPRI - California Energy Commission - U.S. Department of Energy Wind Energy Forecasting Program, Electric Power Research Institute, Palo Alto, CA: 2003. 1004038.
 19. EPRI, 2004: Wind Power Integration Technology Assessment and Case Studies, Electric Power Research Institute, Palo Alto, CA: 2004. 1004806.
 20. EPRI, 2005, Wind Power Integration: Smoothing Short-Term Power Fluctuations, Electric Power Research Institute, Palo Alto, CA: 2004. 1004806.

21. EPRI, 2005, Wind Power Integration: Energy Storage for Firming and Shaping, Electric Power Research Institute, Palo Alto, CA: 2004. 1004806.
22. Hamill T.M., C. Snyder, D.P. Baumhefner, Z. Toth, and S.L. Mullen, 2000: Ensemble forecasting in the short to medium range: report from a workshop. *Bull. Amer. Meteor. Soc.*, 81, 2653-2664.
23. Hodur, R., 1997: The Naval Research Laboratory's coupled ocean-atmospheric meso-scale prediction system (COAMPS). *Mon. Wea. Rev.*, 125, 1414-1430.
24. Kaplan, M. L., J. W. Zack, V. C. Wong and J. J. Tuccillo, 1982: Initial results from a meso-scale atmospheric simulation system and comparisons with an AVE-SESAME I data set., *Mon. Wea. Rev.*, 110, 1564-1590.
25. Kilpatrick, K. A., G. P. Podesta, and R. H. Evans, 2001: Overview of the NOAA/NASA AVHRR Pathfinder algorithm for sea surface temperature and associated matchup database, *J. Geophys. Res.*, 106, C5, 9179-9197.
26. Lubitz, W.D., 2005: Personal communication, draft of PhD. thesis, University of California at Davis.
27. Michalakes, J., J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, and W. Wang, 2004: "The Weather Research and Forecast Model: Software Architecture and Performance," to appear in proceedings of the 11th ECMWF Workshop on the Use of High Performance Computing In Meteorology, 25-29 October 2004, Reading U.K. Ed. George Mozdzyński.
28. Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses. *J. Climate*, 7, 929-948.
29. Roulston, M. S., D.T. Kaplan, J. Hardenberg, and L. A. Smith, 2003: Using Medium Range Weather Forecasts to improve the value of wind energy production. *Renewable Energy*, 28, 585-602.
30. Sivillo, J.K., J.E. Ahlquist, and Z. Toth, 1997: An ensemble forecasting primer. *Wea. Forecasting*, 12, 809-818.
31. Stull, R.B., 1988: *An Introduction to Boundary Layer Meteorology*, 666 pp., Kluwer Academic Publishers, Dordrecht, Boston.
32. Torrence, C. and G.P. Compo, 1998: A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.*, 79, 61-78.
33. Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, 125, 3297-3319.

References

34. Vapnik, V. and A. Chervonenkis. Theory of Pattern Recognition (in Russian). Nauka, Moscow, 1974. (German Translation: W. Wapnik & A. Tschervonenkis, Theorie der Zeichenerkennung, Akademie-Verlag, Berlin, 1979).
35. Vapnik, V., S. Golowich, and A. Smola. Support vector method for function approximation, regression estimation, and signal processing. In M. C. Mozer, M. I. Jordan, and T. Petsche, editors, *Advances in Neural Information Processing Systems 9*, pages 281–287, Cambridge, MA, 1997. MIT Press.
36. Whitaker, J.S. and A.F. Louge, 1998: The relationship between ensemble spread and ensemble mean skill. *Mon. Wea. Rev.*, 126, 3292-3302.