WIND SPEED FORECASTING FOR POWER SYSTEM OPERATION

A Dissertation

by

XINXIN ZHU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Huiyan Sang
Co-Chair of Committee,	Marc G. Genton
Committee Members,	Raymond J. Carroll
	Le Xie
Head of Department,	Simon J. Sheather

August 2013

Major Subject: Statistics

ABSTRACT

In order to support large-scale integration of wind power into current electric energy system, accurate wind speed forecasting is essential, because the high variation and limited predictability of wind pose profound challenges to the power system operation in terms of the efficiency of the system. The goal of this dissertation is to develop advanced statistical wind speed predictive models to reduce the uncertainties in wind, especially the short-term future wind speed. Moreover, a criterion is proposed to evaluate the performance of models. Cost reduction in power system operation, as proposed, is more realistic than prevalent criteria, such as, root mean square error (RMSE) and absolute mean error (MAE).

Two advanced space-time statistical models are introduced for short-term wind speed forecasting. One is a modified regime-switching, space-time wind speed forecasting model, which allows the forecast regimes to vary according to the dominant wind direction and seasons. Thus, it avoids a subjective choice of regimes. The other one is a novel model that incorporates a new variable, geostrophic wind, which has strong influence on the surface wind, into one of the advanced space-time statistical forecasting models. This model is motivated by the lack of improvement in forecast accuracy when using air pressure and temperature directly. Using geostrophic wind in the model is not only critical, it also has a meaningful geophysical interpretation.

The importance of model evaluation is emphasized in the dissertation as well. Rather than using RMSE or MAE, the performance of both wind forecasting models mentioned above are assessed by economic benefits with real wind farm data from Pacific Northwest of the U.S and West Texas. Wind forecasts are incorporated into power system economic dispatch models, and the power system operation cost is

used as a loss measure for the performance of the forecasting models. From another perspective, the new criterion leads to cost-effective scheduling of system-wide wind generation with potential economic benefits arising from the system-wide generation of cost savings and ancillary services cost savings.

As an illustration, the integrated forecasts and economic dispatch framework are applied to the Electric Reliability Council of Texas (ERCOT) equivalent 24-bus system. Compared with persistence and autoregressive models, the first model suggests that cost savings from integration of wind power could be on the scale of tens of millions of dollars. For the second model, numerical simulations suggest that the overall generation cost can be reduced by up to 6.6% using look-ahead dispatch coupled with spatio-temporal wind forecast as compared with dispatch with persistent wind forecast model.

ACKNOWLEDGEMENTS

First, I would like to express the deepest appreciation to my advisor Dr. Marc G. Genton, an outstanding statistician in multiple fields, particularly, in spatial-temporal statistics and skew-slliptical sistributions area. As an advisor, he is inspirational, supportive, helpful and patient throughout my four years' PhD study. Without his guidance and persistent help this dissertation would not have been possible. Moreover, with his attitude to research work, he sets a great model for me to follow in my future life.

I also very much appreciate Dr. Huiyan Sang, who agreed to become my advisor for the last year. Her warmth and kindness have been great support for me at the end of my PhD study. As a young statistician of my age, Dr. Sang has already attained her prominent achievements in extreme value study, which inspires me to work harder.

I would like to thank my committee members, Dr. Raymond J. Carroll and Dr. Le Xie. I thank Dr. Xie and his student, Yingzhong Gu, for their great efforts, help, and comments on the join works, for giving me opportunities to share and discuss my researches in their group. Without them this dissertation would not have been finished. I also thank Dr. Kenneth Bowman's great guidance and technical assistance in one important part of my dissertation. I am also very grateful to Dr. Michael Longnecker, the associate Department Head, for his constant help on problems I met through the process of completing my PhD program. In addition, I very much thank my mentor, Dr. Suojin Wang, for his valuable advises on my PhD study and personal life, which I will cherish in my future life.

Finally, I express my special gratitude to my parents, my brother and his wife,

and my friends for their unconditional and endless support. They have been always with me through ups and downs during my PhD study. Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

Driven by environmental concerns, energy independence, and limitation of fossil fuel resource, wind has become one of the most promising alternative energy resources. The main motivation for many countries to use wind energy is to reduce the emission of greenhouse gases, which is threatening the Earth by global warming. 72% of the totally emitted greenhouse gas is carbon dioxide (CO2), which is due to the burning of fossil fuels, e.g. coal, oil, natural gas; they have been the major energy resources for human activities for hundreds of years. In addition, the uneven distribution of such resources is the root that causes a lot of conflicts between countries and regions. The development of wind energy as an alternative is not only needed but necessary.

Wind energy is green, renewable, and rich in resource, and with advanced technologies, it becomes more and more cost-effective. Unlike in old times when wind energy is used to grind grain or pump water, today's techniques allow us to convert wind energy into power or electricity. Due to the uneven heating of the sun, the motion of air or wind rotates the wind turbines, converting the kinetic energy in wind into mechanical energy. Then, electricity is converted from mechanical energy by the rotating magnetic coil in the gearbox, and finally transmitted to community. In the procedure, wind is the only input, which is very rich in nature, and the electricity is generated without any emission of greenhouse gases. Even the wind turbines and towers are recyclable without causing any pollution. Moreover, the innovation in wind turbine design enables us to capture more wind energy than before, e.g. smoother and stronger wind in higher altitude and off-shore. With current sophisticated techniques, the monetary cost per unit of energy produced is similar to the

cost for new coal and natural gas installations.

The wind industry has been growing by leaps and bounds in the last decades. By the end of 2012, 83 countries in the world are using wind power on a commercial basis. The world total installed capacity has reached 282,587 megawatt (MW), compared with 18,084 MW in 2000. Among all countries, China, the U.S, Germany, and Spain have the most rapid development of wind energy. In Chapter 2, more facts about wind energy are introduced.

However, "increased wind power cannot simply be added to the existing grid without transforming the grid in ways that introduce both significant costs and operational inefficiencies", as Haugen and Musser (2012) said. Subject to the high intermittency in wind power, the stability and reliability of the power system are undermined, with costly non-wind backup power, reserve and ancillary services. Unlike other conventional power sources, wind cannot be turned on and off to meet the changes in demand, and the wind power must be used when the wind is actually blowing, due to the lack of feasible energy storage options. In a power system with a high share of wind power, the operation of standby fast power plants is costly, if the wind is not generating electricity as expected.

Therefore, to keep the cost of wind power as low as possible, it is critical to improve the accuracy of wind forecasting, especially, to reduce the uncertainties in the near-term future wind. Compared with long-term (days or longer ahead) wind forecasts, which are hardly accurate, short-term (hours-ahead) ones have higher quality, and are more closely related with power system operation. Through day-ahead, hours-ahead, and even minutes-ahead electricity market, power system dispatch decides the amount of supply of each power plant at a minimum cost to match the demand. The more accurate hours-ahead wind forecasting is, the more efficiency the power system operation has. More importantly, it allows enough time to start

slower but less expensive power plants to make up possible gaps between demand and supply. Thus, accurate short-term wind forecasting is necessary to reduces the cost for reserve and stabilize the power system.

To provide feasible solutions, researchers have devoted time to short-term wind speed forecasting problems. Physical models, statistical models, and the combination of the two areas have been applied to wind prediction. For short-term wind forecasting, statistical models are considered to be the best. And among all statistical models, such as neural networks, fuzzy logic, local regression and time series methods, space-time models are the best because of high forecasting accuracy and friendly model interpretations. The advantages in space-time statistical models are attributed to their abilities to capture the spatial correlation in wind, besides that in time, by incorporating wind information of neighbors. Moreover, the prediction is in the form of probability distribution, instead of point forecasts, providing more information to reduce the uncertainties of wind. A comprehensive review for short-term wind forecasting in power system operation is in Chapter 2, including literature reviews, essential aspects in model development and future work.

The objective of this dissertation includes developing space-time statistical short-term probabilistic wind speed forecasting models, as well as proposing a new criterion to evaluate the model performance. The main contributions of this dissertation are as follow.

- A space-time statistical model is proposed aiming to generalize the regime-switching space-time diurnal (RSTD) model (Gneiting et al. 2006). In the RSTD model, two forecast regimes were predefined according to local prevailing wind direction, and separate forecasting models were developed based on spatial and temporal information. It outperformed persistence and autore-

gressive forecasts in terms of RMSE, in 2-hour-ahead wind speed prediction problems at Vansycle, Oregon.

However, the definition of forecasting regimes in the RSTD model is subjective and not unique. For other situations, under which the winds follow more complicated patterns, the number and position of the forecast regimes are difficult to determine. To eliminate these constraints, the proposed model generalizes the RSTD model by allowing the forecast regimes to vary with the dominant wind direction in each season instead of fixing the forecast regimes based on prior geographic information. By rotating the dividing angles of the regimes with respect to the minimum MAE for each season, the best position of the forecast regimes is detected. The new model is named RRSTD short for rotating RSTD.

- Another proposed space-time statistical model is called TDDGW model. It is constructed to improve the trigonometric direction diurnal (TDD) model, which was developed by Hering and Genton (2010) to generalize the RSTD model by treating wind direction as a circular variable. It obtains equivalent or better forecasting results than those from the RSTD model without requiring prior wind information. However, the forecasting error is not reduced by including additional weather information directly into the model, such as air pressure and temperature, which are considered to be closely related to wind.

  Based on atmospheric dynamics, a new variable, the geostrophic wind, is incorporated into the TDD model. Geostrophic wind is the airflow under geostrophic balance, a balance between the horizontal pressure gradient force and the Coriolis force, which arises from the rotation of the Earth. It is a good approxima-

tion to the actual horizontal wind for large scale atmospheric flows outside the tropics and in the absence of friction. The surface winds, which rotate wind turbines, are strongly affected by friction with the Earth's surface, and thus are generally not in geostrophic balance. However, the low-level winds are driven mainly by the winds at higher levels, which are close to geostrophic balance. Therefore, the geostrophic wind, which can be estimated from surface pressure data, provides useful information about winds close to the ground.

Motivated by eliminating the limitations of the TDD model in incorporating air pressure and temperature information, the TDDGW model contains information of pressure and temperature and has physical interpretability, leading to more accurate forecasts. In addition, the difference in geostrophic wind direction and temperature between current and one day before are also considered in the corresponding models named as TDDGWD, and TDDGWDT, respectively. Simpler and more efficient methods are proposed to fit the diurnal pattern in wind in order to obtain better forecasts.

- A new criterion, the power system operation costs, is proposed to evaluate the performance of models. Model evaluation is important in making decisions on model implementation. The RMSE and MAE are two commonly used loss functions and the model with the smallest loss is considered to be the most advanced. However, for wind forecasting problems, more realistic loss functions are needed, because penalty on underestimates and forecasts for small true values are desired.

  The power curve error (Hering and Genton 2010) was proposed as a loss function. It links prediction of wind speed to wind power by a power curve and evaluates the loss based on the wind power with penalty on underestimates.

In this dissertation, the power system operation cost is used to perform model evaluation from a different perspective. The emphasis is put on the cost-effectiveness of the system-wide power operation with potential economic benefits arising from the system-wide generation of cost savings and ancillary services cost savings. The model that produces forecasts with the most cost savings for the power system operation is the best.

Based on this criterion, space-time wind forecasts from the above two new proposed models are incorporated into a look-ahead economic dispatch framework. Numerical studies in an ERCOT equivalent 24-bus test system are carried out to evaluate the economic benefits with real wind farm data from Pacific Northwest of the U.S and West Texas.

2. LITERATURE REVIEW*

2.1 Introduction

2.1.1 Wind Energy

Environmental concerns and supply uncertainties are driving many countries to rethink their energy mix and develop diverse sources of clean, renewable energy. Cost-effective energy that can be produced without major negative environmental impacts has become the goal worldwide. For example, the European Union (EU), with its ambitious 20/20/20 target, aims to reduce greenhouse gas emissions by 20% (as compared to 1990), to increase the amount of renewable energy to 20% of the energy supply, and to reduce the overall energy consumption by 20% through improved energy efficiency by 2020; see EU (2008).

Wind energy, as a clean and renewable resource, has been under large-scale development around the world in the last decade. World total capacity increased quickly and stably from year 2000 to 2010, more than doubling every third year, as the left panel of Figure 2.1 shows. The total installed capacity reached 196,630 Megawatts (MW) by the end of 2010, out of which 36,864 MW was added in the single year of 2010. The electricity generated from all installed turbines, 430 Terawatthours per annum, is enough to supply the demand of the United Kingdom, the sixth largest economy of the world, according to the World Wind Energy Association (World Wind Energy Association 2010). The average annual growth rate of wind power capacity was about 27% during the years 2000 to 2010, with highest growth rate in year 2001 followed by 2009, as shown in the right panel of Figure 2.1.

*Reprinted with permission from "Short-Term Wind Speed Forecasting for Power System Operation" by Xinxin Zhu and Marc G. Genton, 2012. International Statistical Review, 80, 2-23, Copyright [2013] by John Wiley and Sons.

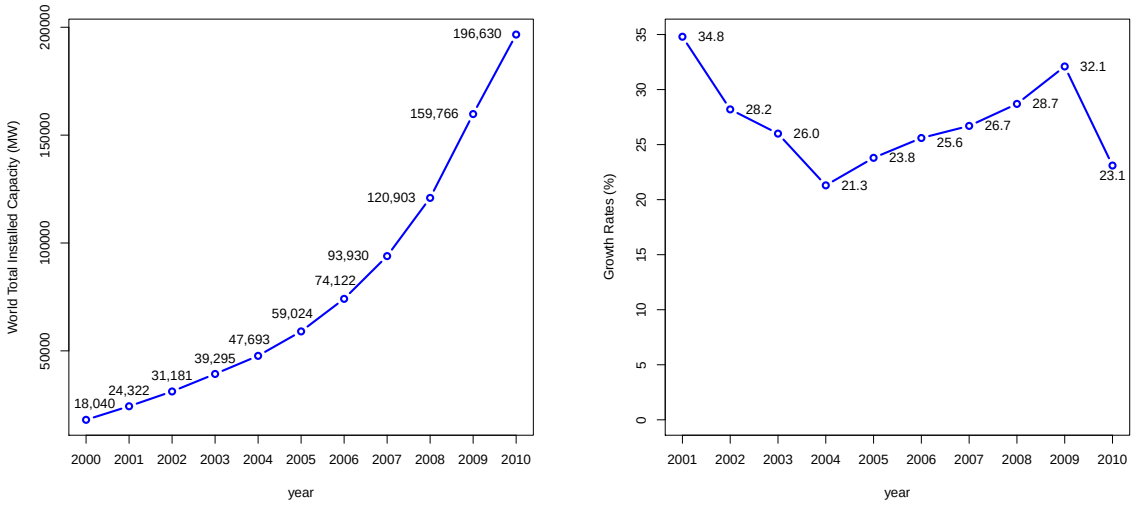

Figure 2.1: Left panel: world total installed wind power capacity from year 2000 to 2010. Right panel: world market growth rate of newly installed wind power capacity from the installed capacity of the previous year during 2000 to 2010.

North America, Europe, and Asia are the top three wind markets, providing 44%, 31%, and 22% respectively of the world total wind capacity in 2010. Asia is contributing the largest amount of new installation, about 55%. This is mainly due to the rapid wind power development in China which became the new leader in 2010 with a total installed wind capacity of over $44,733$ MW, accounting for 23% of the worldwide wind capacity. With the decrease in new capacity in the U.S., North America has fallen to the third position in newly installed turbines with a share of 17%.

Figure 2.2 shows the shares of wind power capacity at the end of 2010 for selected countries, with 74% being accounted for by China, the U.S., Germany, Spain and India.

Wind power is on its way to high level penetration in the electricity supply market. For example, wind has become one of the largest electricity source in a number of

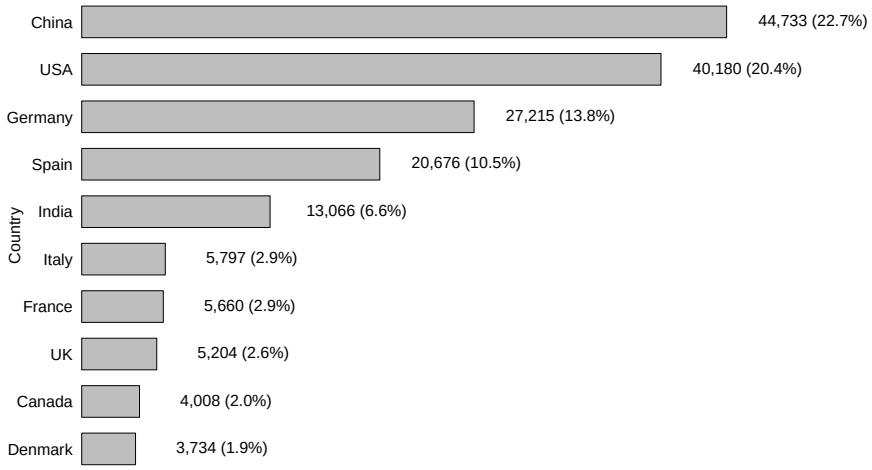

Figure 2.2: Country shares of total installed wind power capacity (in MW and percentage) by the end of 2010.

European countries, such as Denmark, Portugal, and Spain, supplying 16%-21% of the electricity demand in 2010. Worldwide, wind power accounted for 2.5% of the electricity supply in 2010, an increase from 2% in 2009. This value is expected to increase tremendously in future decades. A 2008 report by the U.S. Department of Energy (DOE 2008) described a scenario in which wind energy will provide 20% of the U.S. electricity demand by 2030. China is expecting to develop a total capacity of 150 Gigawatts (GW) by 2020, and 450 GW by 2050 according to a report published by the Chinese Renewable Energy Industries Association (CREIA 2010).

Wind energy has become very attractive due to its renewable and clean nature. First, the wind resource is sustainable and will be available as long as there is uneven heating from the sun on the surface of the earth. Second, wind energy is an emission-free resource. Currently, fossil fuel generation (mainly coal and natural

gas) is the largest electricity source, and fossil fueled power stations are the major emitters of CO_2. Carbon dioxide is the most important greenhouse gas, a major contributor to the global warming observed over the last 100 years. Power generated from wind, in comparison, is green without harmful byproducts produced by other traditional energy sources. Thus, high level penetration of wind power helps to reduce environmental damages from other sources.

In addition, wind power generation is cost-effective. Advanced technologies in wind turbine (or wind generator) design reduce the cost of utilizing wind energy and allow large-scale integration into the current electricity grid. To generate wind power, the only input needed is the wind from nature, which is free. Through turbines, wind energy is converted into mechanical energy which is used to generate electricity. The most modern turbines installed onshore have a capacity between 1.5 MW to 3 MW of electricity each, which means that they can produce 1.5 MW to 3 MW per hour at their maximum rated wind speed. Development of offshore turbine technology allows for effective utilization of stronger and more uniformly blowing wind with rated capacity between 2 MW and 5 MW each (DOE 2010). The cost of a Kilowatt (KW) of wind powered electricity is now nearly the same as that of coal or nuclear energy. Also wind turbines can work 8 to 10 years after installation, and the decommissioning is environmentally friendly by recycling.

2.1.2 Integrating Wind Energy into the Power System

The benefits of wind energy are accompanied by several challenges: high variability, limited predictability, limited dispatchability and non-storability. Unlike fossil fuel generation, wind power is not fully dispatchable. A coal power plant, for example, can be turned on or off, and can adjust its output to the demand. Wind power, however, cannot be controlled by power system operators because wind farms cannot

increase their power generation upon request when there is not sufficient wind. Wind farms can only reduce the output. Also, wind cannot be stored like coal, natural gas or atoms for future power generation. All of these disadvantages of the wind resource pose profound challenges to today's power system operations to integrate large-scale wind power.

The basic function of power system operation is to balance the electricity supply and demand at a minimum cost under the constraints of the transmission network and possible contingencies. At different time scales (day-ahead, hour-ahead, or 5 to 10 minutes-ahead), power system operators decide the output of each power plant to meet the total load forecast and minimize the total cost at the same time. Besides producing electricity, power plants also provide ancillary services, such as frequency regulation and reserve requirement, to help the power system operate in a reliable and secure manner. For the frequency regulation service, the on-line power plants are committed to adjust their outputs to maintain the frequency at the base level (60 hertz in the U.S.) responding to the automatic generation control signal. For the reserve service, some power plants are required to save a certain level of capacity for possible contingencies.

The high uncertainty in wind increases the operation cost and reduces the stability and reliability of power systems. Before integrating variable power resources, such as wind and solar, the main difficulty in power system operation was coming from the uncertainty of the demand. However, when large-scale wind power is integrated into the power system, the variation from the supply brings profound impacts on the operation even on top of the demand uncertainty at different time scales. For example, close to the real-time operation (5 to 10 minutes or hour-ahead), if the wind power generators fail to produce as much electricity as predicted due to the wind slowing down, other fast-responding units, such as gas-fired power plants, which are

very expensive, are needed immediately to balance the load. Otherwise, tremendous losses could be caused by blackouts. Xie et al. (2011a) analyzed the operational challenges due to the high variations and limited predictability in wind and discussed possible solutions in detail.

How to reduce the uncertainty in wind has been the focus of research and new developments in the last decades. To integrate large-scale wind power into power systems smoothly, wind generation forecasting models have emerged rapidly to improve the accuracy of forecasts. They include time series models, numerical weather prediction based models (Giebel 2003) and space-time models. Both short-term (several minutes to hours-ahead) and longer-term (days, weeks to years ahead) wind forecasts are valuable to developing wind power. For instance, Marquis et al. (2011) highlighted the needs of wind forecasts to reach significant penetration levels of wind energy, especially regarding short-term forecasting.

Compared to long-term wind forecasting, short-term forecasting (hours-ahead) is more accurate and reliable. It is critical for effective operation planning with a high penetration level of wind power, in terms of increasing the savings due to reduced committed thermal capacity and savings due to the operation of more efficient units; see Xie et al. (2011a). Long-term wind forecasting is typically based on physics and numerical weather prediction, while statistical models are thought to be more competitive in short-term forecasting problems (Genton and Hering 2007) in terms of forecast accuracy and model interpretation, and it is the main topic of this dissertation.

2.1.3 Outline

Short-term wind prediction has been the focus of extensive research in the last decade. The motivation of this dissertation is to review statistical models for short-

term wind forecasting, to bring up some important issues in evaluating the performance of forecasts, and to describe new challenges in wind forecasting and future research topics.

The article is organized as follows. Section 2.2 describes the relationship between wind speed forecasting and wind power forecasting and the recent trend away from point forecasting to probabilistic forecasting. Section 2.3 summarizes some traditional time series statistical models of wind speed forecasting, including autoregressive models and the Kalman filter method. In Section 2.4, space-time statistical forecasting models are introduced. Evaluation of wind speed forecasting models is discussed in Section 2.5, emphasizing that loss functions should meet the practical requirements in power system operations. Future research topics about ramp events and challenges in offshore wind speed forecasting are discussed in Section 2.6.

2.2 Wind Speed Forecasting

The power system operation balances the supply and demand of power at a minimum cost subject to certain constraints. Given the advanced techniques in load forecasting, the major difficulty of integrating large-scale wind power into the system lies in the uncertainty of wind power generation. Accurate wind power forecasting is the primary motivation, while finding a good way to define the uncertainty so that more information can be provided to the power system operation for efficient decision making is also of great interest.

This section reports on the relationship between wind power forecasting and wind speed forecasting, explains why probabilistic forecasting is a better way to define the uncertainty in wind than just point forecasting, and describes the space-time correlations in wind.

2.2.1 Wind Speed and Power Forecasting

There are two approaches commonly used in wind power forecasting. One approach is to forecast wind power generation directly, and another is to convert wind speed forecasts into wind power based on a certain power curve. A deterministic power curve is usually provided by the wind turbine manufacturer. It maps wind speed into wind power, and it varies with the capacity of the turbine. With the same wind speed, different turbines generate different amounts of energy depending on each turbine's design. Figure 2.3 displays three different types of power curves from 0.3 MW of Nordex (solid), 1.5 MW of GE (dashed) to 2.5 MW of Bonus (dotted). A typical wind power curve has a cut-in speed, a rated speed and a cut-out speed,

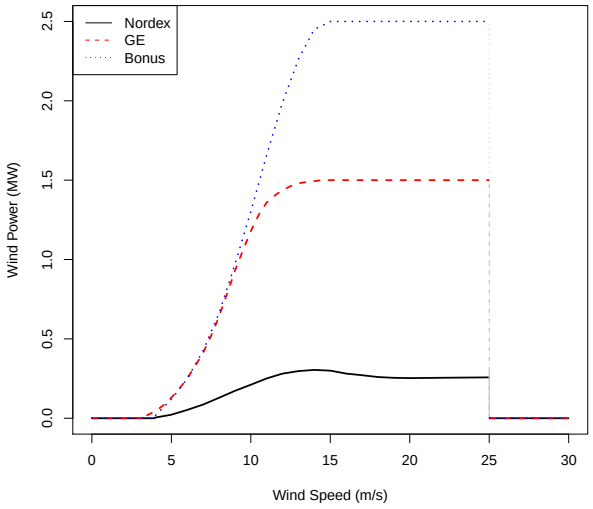

Figure 2.3: Three power curves with different capacity ranges from low to high from three manufacturers: 0.3 MW from Nordex, 1.5 MW from GE, and 2.5 MW from Bonus.

which are speeds at which a turbine starts to work, starts to have a constant maximum output, and stops working to avoid damages. For a 1.5 MW turbine of GE, these speeds are 3.5 m/s, 13.5 m/s and 25 m/s. Recent work by Jeon and Taylor (2011) has however recognized the stochastic nature of the relationship between wind power and wind speed, and has proposed to model it explicitly.

For power system operation, wind power forecasting by converting wind speed forecasts is a better approach than predicting wind power output directly. Neighboring wind farms with different installed wind turbines may share the same wind speed. Instead of requiring separate power forecasts, they can get them by converting the common wind speed forecasts based on their own power curves. Also, wind speed forecasting can be more precise than wind power forecasting due to the spatial correlation of wind. For example, in order to forecast wind power output of a wind farm located downstream of the wind, significant benefits from the upstream wind speed forecasting could be obtained where there is no wind farm or wind power generation available. Therefore, this dissertation focuses on wind speed forecasting.

2.2.2 Point Forecasting Versus Probabilistic Forecasting

There are two major approaches to forecast wind speed: point forecasting and probabilistic forecasting. Point forecasting gives a single value as the forecast of future wind speed, while probabilistic wind speed forecasting models a probability density function for future wind speed.

Probabilistic forecasting is more informative and useful than point forecasting. Though point forecasting is the prime interest of wind speed forecasting, it is not enough for a reliable and secure power system operation. Due to the prediction error, point forecasting has some variability, and it also has no information about how the true value would spread out around the forecast, which is very important for power

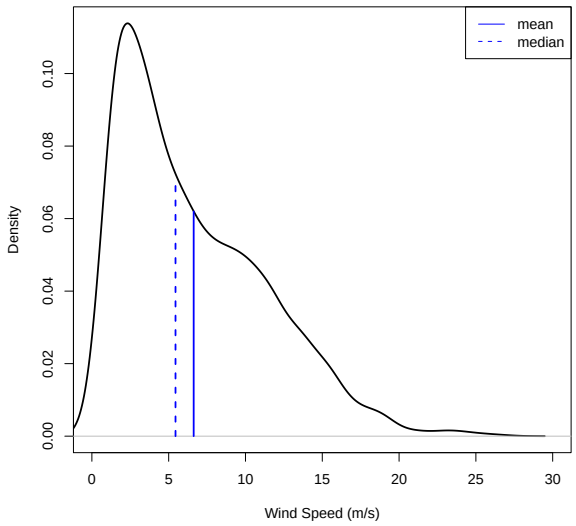

Figure 2.4: Nonparametric density estimation of 2002 hourly wind speed data at Vansycle, Oregon, U.S. The vertical lines represent the sample mean (solid) of 6.6 m/s and the sample median (dashed) of 5.4 m/s. The sample skewness is 0.8.

system operators to make correct decisions. On the other hand, probabilistic forecasting not only gives point forecasts with the mean or quantiles of the distribution, but also provides information about the uncertainty. Confidence intervals of a point forecast, for example, can be calculated and this helps power system operators to make more reliable decisions.

In probabilistic wind speed forecasting, the choice of density functions must be consistent with the wind patterns. Wind speeds are nonnegative valued and usually right skewed due to the low probability of high values; see Figure 2.4 for illustration based on data in Hering and Genton (2010). Some wind regimes can have bimodal rather than unimodal wind speeds, and can also have high percentages of no wind speed or high wind speed. Consequently, densities that are right skewed with nonnegative domain are usually chosen to fit the wind speed distribution. For example,

gamma, Weibull, Rayleigh, truncated normal, and beta distributions have all been used to fit wind speed. Among these distributions, the Weibull distribution is found to be the most accepted for wind energy: it is flexible with a closed form, only has two parameters that are easy to estimate, and has specific goodness-of-fit tests as discussed by Ramírez and Carta (2005) who also pointed out that the data sampling interval has no significant effect on the shape of the density. However, the Weibull distribution cannot represent high percentages of null wind speeds or bimodal cases. The truncated normal distribution was found useful in describing winds with high percentages of null wind speed; see Carta et al. (2009). Mixture distributions with one Weibull and one truncated normal distribution have been fitted to bimodal wind speeds, taking into account null wind speeds as well; see Carta and Ramírez (2007).

The log-normal and square root normal distributions have also been fitted to log and square root transformed wind speed data, but their goodness-of-fits are controversial. Lau and McSharry (2010) applied a logistic transformation to normalized wind power data and fitted a model to the transformed data. They produced 15 minutes to 24 hours ahead probabilistic forecasts that outperformed the forecasts based on a truncated normal distribution with an exponential smoothing method. The latter model was still thought to be a useful alternative in probabilistic forecasting problems due to its robustness and computational efficiency.

Recently, the bivariate skew-t distribution (Azzalini and Genton 2008) has been used by Hering and Genton (2010) in wind speed forecasting problems, after converting wind speed and wind direction into Cartesian components. The space-time forecasting methods in Section 4 are all probabilistic with truncated normal distributions. A multivariate t modeling of wind speed and wind direction would be of interest for wind regimes with a high percentage of high or extreme wind speeds. A valuable review of wind speed probability distributions used in wind energy analyses

can be found in Carta et al. (2009).

2.2.3 Space-Time Wind Correlations

Winds are correlated both in time and space. Wind is driven by the horizontal difference in air pressure, which is caused by uneven heating of the earth's surface by the sun, and as the difference in air pressure takes time to be balanced, wind lasts in time. Therefore, future wind speed is related to current and earlier wind speeds. A windy day at a given location would be expected with high probability if the wind has already been blowing there for several days. Additionally, wind speed and direction are affected by the local geographic features. In flat areas, downstream wind is almost the translation of upstream wind, so the patterns of downstream wind is similar to that of upstream. In areas with mountains, wind speed is slowed down, and air blows in directions that are subject to the constraints of mountain shapes. The correlation in space suggests that information from neighborhoods of the target location could be very useful for accurate wind speed forecasting.

Based on the nature of wind, e.g. correlated in space and time, large amounts of studies have been devoted to developing wind speed forecasting models in the last decades, including physical models and statistical models. Most physical models incorporate output from numerical weather prediction (NWP) models to predict wind speed. However, they are not effective for short-term forecasting due to their computational costs, see Genton and Hering (2007). Statistical models are more competitive for short-term wind speed forecasting. Conventional time series methods, space-time methods, and other techniques (such as neural methods, fuzzy logic methods and hybrid methods) have all been applied to wind speed forecasting. The latter techniques usually use a "black box" approach without good interpretation of the results, while the first two are more interpretable without loss of accuracy of

forecasting and are the main topics of this dissertation. In the next two sections, conventional time series methods and space-time models for short-term wind speed forecasting are reviewed and discussed.

2.3 Time Series Models for Forecasting

2.3.1 Basic Concepts

Let y_1,y_2,...,y_t be the wind speed observations up to time t, and y_{t+k} be the k-step ahead unknown future wind speed to be predicted with $\hat{y}_{t+k}$. Here y_t could be an averaged value at a certain time scale. For example, for hourly average wind speed data, y_t is the average wind speed during hour t, and y_{t+k} is predicted as the average wind speed during the hour $t+k$. Given current and past wind speed observations, a point forecast of wind speed estimates y_{t+k}, and a probabilistic forecast estimates the density of y_{t+k}, denoted by $f(y_{t+k}|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is an unknown parameter vector of the density.

Depending on the engineering and economic goals, there are long-term, medium-term, and short-term wind speed forecasting: long-term (months or years ahead) prediction is of interest for investment planning in generation capacity; medium-term (days ahead) prediction serves for management and maintenance of power system operation; short-term (1-10 hours ahead) prediction is used for effective operations planning.

In this dissertation, short-term wind speed forecasting is considered because it is closely related to power system operations. First, hours ahead forecasting allows conventional power sources to have enough time to start and provide power as demanded in time. Typically, it is between 3 hours to 10 hours, but for quick resources, it can be under 3 hours (Genton and Hering 2007). Second, short-term wind speed forecasting helps power system operations to dispatch more economically. Other

sources with high economical and environmental cost can be down-regulated based on the short-term wind speed predictions, while wind energy can be fully utilized.

Given current and historical wind speed observations, prediction of future wind speed is a classic time series problem. After introducing a reference model, this section mainly reviews some typical statistical time series models used in wind speed forecasting.

2.3.2 Reference Model

Persistence forecasting assumes that the future wind speed is the same as the current one: $\hat{y}_{t+k} = y_t$. This method is reasonable because wind lasts in time. However, due to the high variation of wind, it works better for very short-term forecasting such as 10 minutes ahead.

Often, persistence forecasting is used as a reference for evaluating the performance of advanced forecasting methods. A new method is thought to be advanced and worth implementation when it outperforms the persistence forecasting.

2.3.3 Autoregressive Models

A typical autoregressive (AR) model with p autoregressive terms, denoted by AR(p), is defined as:

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t,$$

where c is a constant, ϕ_i, $i = 1, \ldots, p$ are the autoregressive parameters, ϵ_t is a white noise process, and y_t is wind speed at time t in our case. Here p can be decided with the autocorrelation function or with selection criteria, and parameters can be estimated by the Yule-Walker method under the assumption of stationarity; see Tsay (2010) for more detail.

An AR(p) model can capture the temporal correlation in wind. It assumes that

the future wind speed is a linear combination of current and past wind speed observations with a white noise error. The order p defines the number of previous observations with which the future wind speed correlates, and the parameters ϕ_i, $i = 1, \ldots, p$ describe how strong the correlations are. Since the wind speed distribution is non-Gaussian and seasonal, transformation and modeling the seasonal trend are often necessary. Brown et al. (1984) applied the square root transformation to a series of hourly average wind speed. After fitting and extracting a diurnal trend component, an AR model for the residuals was used.

The AR(p) models have been widely used for short-term wind speed forecasting and they usually outperform persistence forecasting. For example, Schlink and Tetzlaff (1998) used an AR(5) model to forecast wind speed at an airport and found that the AR(5) model produced more precise forecasts. That is the forecast confidence intervals based on the AR(5) model were narrower than those based on the persistence model, permitting a confidence of 97.5% compared to 95% in the persistence model. More recently, Gneiting et al. (2006) used an AR model to fit the center parameter of a truncated normal wind speed distribution after removing the diurnal pattern, and the prediction root mean squared error was reduced by 16% compared to the persistence method.

The AR(p) model is a special case of the autoregressive moving average, ARMA(p, q), model, adding q moving average terms to AR(p): $y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$, where the θ_j's are moving average parameters. ARMA models have also been applied to wind speed forecasting. Tantareanu (1992) found that ARMA models can perform up to 30% better than persistence forecasting for 3 to 10-steps ahead in 4 seconds average of 2.5-Hz sampled data. More generally, autoregressive integrated moving average (ARIMA) models are also used for wind speed simulation and prediction purpose; see Kamal and Jafri (1997) for more detail.

2.3.4 Kalman Filter

The Kalman filter (KF) is another method to predict future wind speed as a linear combination of the current and past observations. Instead of fixing the linear coefficients used in the model, the KF updates them recursively based on the previous data observations and the accuracy of the last forecast, by minimizing mean squared error.

In the KF, wind speed forecasting is described by the following two equations. Here 1-step ahead forecasting is illustrated:

$$y_t = \mathbf{H}_t' \mathbf{A}_t + \nu_t, \tag{2.1}$$

$$\mathbf{A}_{t+1} = \mathbf{\Phi} \mathbf{A}_t + \boldsymbol{\omega}_t. \tag{2.2}$$

Equation (2.1) is the observation equation. It calculates a forecast value y_t at time t as a linear combination of the last N observed wind speed values, denoted by the $N \times 1$ vector $\mathbf{H}_t = (y_{t-1}, y_{t-2}, \ldots, y_{t-N})'$, where N is the order of the filter. The $N \times 1$ state vector $\mathbf{A}_t = (a_{t,1}, a_{t,2}, \ldots, a_{t,N})'$ gives the regression coefficients, and it varies at each time step. The system equation (2.2) defines the time dependent evolution of $\mathbf{A}_t$ and it has covariance matrix $\mathbf{S}_t$ $(N \times N)$. Here $\mathbf{\Phi}$ is a known $N \times N$ transition matrix, and it is usually set to be the identity matrix in applications. The observation noise ν_t is assumed to be normally distributed with mean 0 and variance V_t: $\nu_t \sim N(0, V_t)$. And $\boldsymbol{\omega}_t$ is the system noise, which is also assumed to be normally distributed with mean $\mathbf{0}$ and covariance matrix $\mathbf{W}_t$ $(N \times N)$: $\boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathbf{W}_t)$.

With the new observed value y_t, $\mathbf{A}_t$ is updated as follows:

$$\mathbf{A}_t = \mathbf{A}_{t|t-1} + \mathbf{K}_t(y_t - \mathbf{H}_t' \mathbf{A}_{t|t-1}), \tag{2.3}$$

where $\mathbf{A}_{t|t-1} = \mathbf{\Phi}\mathbf{A}_{t-1}$, $\mathbf{K}_t = \mathbf{S}_{t|t-1}\mathbf{H}_t \big/ (\mathbf{H}_t'\mathbf{S}_{t|t-1}\mathbf{H}_t + V_t)$, and $\mathbf{S}_{t|t-1} = \mathbf{\Phi}\mathbf{S}_{t-1}\mathbf{\Phi}' + \mathbf{W}_{t-1}$. The covariance matrix $\mathbf{S}_t$ is updated as:

$$\mathbf{S}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t')\mathbf{S}_{t|t-1},$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{K}_t$ is the Kalman gain ($N \times 1$). It is related to the uncertainty in the system noise and the observation noise, a weighting factor on the error $y_t - \mathbf{H}_t'\mathbf{A}_{t|t-1}$ in updating $\mathbf{A}_t$ from $\mathbf{A}_{t|t-1}$, as in equation (2.3). The initialization of the KF is simple to do due to its insignificant influence on the final results; see Giebel (2001). The KF can easily adapt to the change in observations, and it does not necessarily require long historical data records. However, it is a problem to estimate the covariance matrix $\mathbf{S}_t$ when the dimension N is high.

Applications of the KF in wind speed forecasting can be found in Bossanyi (1985), Giebel (2001) and Crochet (2004). Bossanyi (1985) found a 10% reduction in root mean squared forecasting error compared to the persistence method in 1-minute-ahead wind speed forecasting problems, but persistence forecasting performed better for hourly data. Geerts (1984) applied both ARMA models and KF to predict wind speed with a forecast horizon of up to 24 hours in hourly time-steps, finding that an ARMA$(2, 1)$ gave better results than KF, but both were better than persistence forecasting up to a 16 hour horizon. Extended from the linear structure of KF, non-linear functions have been developed. Louka et al. (2008) applied polynomial functions to the observation equation in the KF to numerical weather predictions and found significant reduction of the absolute bias with a 4th order polynomial function compared to a linear one.

A space-time KF that includes spatial correlations has been developed and combined with dimension reduction ideas by Wikle and Cressie (1999) for spatial kriging

prediction of near-surface winds over the Pacific ocean. Malmberg et al. (2005) also proposed a space-time KF method to forecast future wind speed over the North Atlantic ocean. However, these space-time KF models are based on large-scale wind datasets collected from a large number of locations. They are not well suited for datasets collected from only a few locations within a neighborhood because there are not enough data to fit an appropriate spatial covariance model which is usually assumed to be stationary and isotropic. Therefore, space-time KF models for small-scale short-term wind speed forecasting would be of interest. Also because the KF updates forecasting results based on new observations and the last forecasting error, if more information from spatial correlations were used in the model then better forecasting results than AR models could be expected.

2.4 Space-Time Statistical Models for Forecasting

2.4.1 Motivation

Wind information from spatial neighborhoods is also very useful for highly accurate short-term wind speed forecasting. Because wind is a horizontal movement in the atmosphere near the surface driven by air pressure, it usually covers a large area. Winds at different locations in that area tend to be positively correlated and share similar characteristics. That is to say that wind speed at a certain location could be predicted from wind speed at adjacent locations.

Taking account of the local topographic information into wind speed forecasting is also highly beneficial. Wind speed and direction are significantly affected by the local terrain, and this is very important in choosing neighborhood information. Flat grounds allow wind to blow uninterrupted, whereas complex terrains can slow down the wind and even change the wind direction. Choosing neighborhoods that bring major contributions in predicting wind speed at a certain location depends on the

local geographic features. For example, wind information observed on one side of a mountain hardly helps to predict wind speed on the other side, while wind at one end of a valley could provide valuable information for the other end in terms of wind speed forecasting.

Extended from traditional time series forecasting models, space-time statistical models take the spatial correlation into account, in addition to the time correlation. They have been the focus of extensive research in recent years. Alexiadis et al. (1999) found that the use of off-site predictors can improve forecast accuracy in forecasts of wind speed and wind power at Thessaloniki Bay, Greece. More recently, de Luna and Genton (2005) provided time-forward predictions with vector autoregressive (VAR) models based on daily averages of wind speeds from 11 synoptic meteorological stations in Ireland. Gneiting et al. (2006) proposed a regime-switching space-time diurnal (RSTD) method, taking into account both spatial and temporal correlations in forecasting wind speed at the Stateline Wind Energy Center in Oregon, U.S.. Hering and Genton (2010) generalized the RSTD model by including wind direction directly into the model. The last two models are discussed in more detail in the following subsections.

2.4.2 Regime-Switching Space-Time Diurnal Model

Gneiting et al. (2006) proposed the Regime-switching Space-Time Diurnal (RSTD) model for predicting the 2-hour ahead average wind speed at the Stateline Wind Energy Center in Vansycle, Oregon, U.S.. Their analysis was based on hourly average wind speed data collected in 2002 and 2003 from Vansycle and two other sites: Goodnoe Hills, WA (146 km west of Vansycle), and Kennewick, WA (39 km northwest of Vansycle); see the map of the locations in Figure 2.5.

These three locations are along the Columbia River Gorge which runs from east

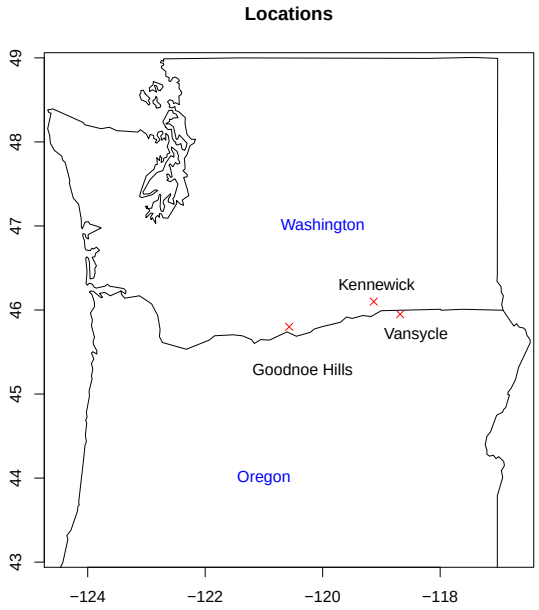

Figure 2.5: Map of the three locations: Vansycle, Kennewick and Goodnoe Hills on the border between Washington and Oregon in the U.S..

to west. Due to the high terrain to the north and south, the airflow runs parallel to the channel of walls, resulting in mostly westerly or easterly winds.

To forecast 2-hour ahead hourly average wind speed at Vansycle, the RSTD model takes advantage of the special landforms of the Columbia River Gorge and chooses Goodnoe Hills, the most westerly station, as the indicator of the forecast regime. Two regimes are defined: a westerly regime and an easterly regime. Then, forecasting models are built separately for each of them.

It is assumed that the 2-hour ahead wind speed at Vansycle, denoted by V_{t+2}, follows a truncated normal distribution on the positive domain, with center parameter μ_{t+2} and scale parameter σ_{t+2}, that is, $V_{t+2} \sim N^{+}(\mu_{t+2}, \sigma^2_{t+2})$. The key is in modeling μ_{t+2} and σ_{t+2}. For the center parameter μ_{t+2}, different models were fitted for each regime. For the westerly regime, $\mu_{t+2} = D_{t+2} + \mu^r_{t+2}$. Here D_s, $s = 1, \ldots, 24$,

are linear combinations of trigonometric functions of the hour of the day, fitting the diurnal pattern of the wind speed:

$$D_s = d_0 + d_1 \sin\left(\frac{2\pi s}{24}\right) + d_2 \cos\left(\frac{2\pi s}{24}\right) + d_3 \sin\left(\frac{4\pi s}{24}\right) + d_4 \cos\left(\frac{4\pi s}{24}\right).$$

After removing a diurnal pattern from the wind speed at Vansycle, the residual, μ^r_{t+2}, is fitted by a linear function of current and past residuals of wind speed at the 3 locations (predictors are selected by Bayesian Information Criteria – BIC):

$$\mu^r_{t+2} = a_0 + a_1 V^r_t + a_2 V^r_{t-1} + a_3 K^r_t + a_4 K^r_{t-1} + a_5 G^r_t, \tag{2.4}$$

where V^r_t, K^r_t, and G^r_t are residual wind speeds at time t at Vansycle, Kennewick and Goodnoe Hills, respectively. For the easterly regime, the center parameter is modeled by a linear function of current and past wind speed of the 3 locations directly, without diurnal component removed, since removing a diurnal pattern did not improve the forecasting results:

$$\mu_{t+2} = a_0 + a_1 V_t + a_2 K_t. \tag{2.5}$$

The scale parameter σ_{t+2} is fitted with the same model in both regimes:

$$\sigma_{t+2} = b_0 + b_1 v_t, \tag{2.6}$$

where the volatility value, v_t, is

$$v_t = \left[\frac{1}{6}\sum_{i=0}^{1}\left\{(V^r_{t-i} - V^r_{t-i-1})^2 + (K^r_{t-i} - K^r_{t-i-1})^2 + (G^r_{t-i} - G^r_{t-i-1})^2\right\}\right]^{1/2}.$$

The coefficients b_0 and b_1 are constrained to be nonnegative. All of the coefficients

in (2.4), (2.5) and (2.6) are estimated by the continuous ranked probability score (CRPS) method; see Gneiting and Raftery (2007).

The RSTD model was trained with data in 2002 and tested with data in 2003. The results were significantly better than univariate time series methods. For example, in July 2003, the RSTD forecasts had a root mean squared prediction error (RMSE) 28% lower than that of the persistence forecasts, while the AR model was 16% lower and the spatial VAR was 27% lower. Moreover, the RSTD model provides a probabilistic forecast, from which uncertainty can be evaluated.

2.4.3 Trigonometric Direction Diurnal Model

The RSTD model relies on the geography of the specific forecasting area, and the decision of the number and position of the forecast regimes can often be far less obvious than the situation in the Columbia River Gorge forecasting region. Hering and Genton (2010) introduced the Trigonometric Direction Diurnal (TDD) model which eliminates the regimes by incorporating wind direction directly into the predictive mean function of the RSTD model. It treats wind direction as a circular variable and uses its sine and cosine, and achieves similar forecast accuracy as the RSTD model. Specifically, Hering and Genton (2010) modeled the residual predictive center, μ^r_{t+2}, of a truncated normal distribution based on the present and past residual wind speed series at all three locations, as:

$$\begin{aligned}\mu^r_{t+2} &= a_0 + a_1 V^r_t + a_2 V^r_{t-1} + a_3 K^r_t + a_4 K^r_{t-1} + a_5 G^r_t + a_6 \sin(\theta^r_{V,t}) + a_7 \cos(\theta^r_{V,t}) \\ &\quad + a_8 \sin(\theta^r_{K,t}) + a_9 \cos(\theta^r_{K,t}) + a_{10} \sin(\theta^r_{G,t}) + a_{11} \cos(\theta^r_{G,t}), \qquad (2.7)\end{aligned}$$

where $\theta^r_{i,t}$, $i \in \{V, K, G\}$, are the residual wind directions at each of the three locations at time t. The coefficients in (2.7) are identified by a forward selection method

with BIC on the dataset of 2002 as training and on the 2003 data as testing. The model for predictive scale, σ_{t+2}, has the same form as for the RSTD model, and CRPS is also used to estimate the coefficients.

The TDD model generalizes the RSTD model while achieving similar forecasting accuracy. The regime definition of the RSTD model is based on the particular geographic features of the target area and the fact that its prevailing winds are westerly or easterly. The TDD model did not need any prior geographic information about the target area, but used the wind direction to help detect the spatial correlation in wind. It is expected that for some areas there are no significant wind patterns or the patterns are too complex to be modeled. Under these circumstances, the TDD model would be more powerful than the RSTD model for wind speed forecasting.

2.4.4 Other Models

There are some other interesting statistical models for short-term wind speed forecasting. For example, Hering and Genton (2010) proposed a model based on the bivariate skew-t distribution as predictive distribution for the first time in wind speed forecasting. They converted wind speed and wind direction data from the three locations into Cartesian coordinates, removed the diurnal trend, and then fitted the residuals with a bivariate skew-t distribution. This method not only took space-time correlations into account, yielded probabilistic forecasts, and achieved similar forecasting accuracy as the RSTD and TDD models, but it also provided forecasts of the wind direction.

Neural networks (NNs), fuzzy logic methods and some hybrid methods have also been applied to short-term wind speed forecasting. Unlike the traditional time series methods and space-time models introduced in this section, they use a "black box" approach and often lack a good interpretation of the model. Still in terms of forecast-

ing accuracy, Sfetsos (2000) compared some of these techniques and ARIMA models. He applied a persistence model, ARIMA models, NN and neuro-fuzzy systems to forecast mean hourly wind speed, and found that NN achieved the best results with a 20-40% average improvement compared to persistence. More studies on NN can be found in Sfetsos (2002) and Cadenas and Rivera (2009). Fuzzy models were applied to wind speed forecasting by Damousis and Dokopoulos (2006) and Damousis et al. (2004), including neighboring locations as well as the target location, and the improvement ranged from 9% to 28%, depending on the forecast horizon, compared to persistence forecasts.

2.5 Evaluation of Forecasts

Evaluating the performance of different models is another important component of wind speed forecasting for power system operation. Before a final decision is made about which forecasting model should be implemented, the loss of each model needs to be evaluated. How to define the loss caused by the forecasts from a model depends on the practical requirements in power system operation. Moreover, the loss of a model should be evaluated based on corresponding forecasts that minimize it. Besides point forecasting, information on the uncertainty of future wind speed is also important to operate power systems efficiently and reliably.

In this section, the importance of matching loss functions and forecasts is emphasized, point out that more realistic loss functions are needed in the problem of wind speed forecasting, propose two relevant loss functions, and describe a numerical experiment. Comparison and uncertainty of forecasts are discussed as well.

2.5.1 Loss Functions and Forecasts

Accurate prediction is one of the most important targets in forecasting uncertain future wind speeds. It is now a common practice to divide the whole data set into

two nonoverlapping parts: training data and testing data. Forecasting models are built based on the training data and evaluated on the testing data. The measure of prediction accuracy depends on how one would evaluate the loss resulting from prediction error, the difference between true value and forecast. Predictors minimizing the loss are preferred. Mean squared error (MSE) and mean absolute error (MAE) are two of the most commonly used loss functions to evaluate predictions. In practice, MSE, MAE or other loss functions are evaluated with point forecasts from models for a certain time period. However, Gneiting (2011a) pointed out that "This can lead to grossly misguided inferences, unless the loss function and the forecasting task are carefully matched." Fildes et al. (2008) also state that "Defining the basic requirements of a good error measure is still a controversial issue."

If the uncertainty of wind speed y_t at time t is modeled by a certain probability distribution function F, and let $\hat{x}_t$ be any predictor with loss $L(y_t, \hat{x}_t)$, then $\hat{y}_t$ is called an optimal forecast if it minimizes the expected loss:

$$\hat{y}_t = \arg\min_{\hat{x}_t} E_F\{L(y_t, \hat{x}_t)\}. \tag{2.8}$$

For MSE, $L(y_t, \hat{y}_t) = (y_t - \hat{y}_t)^2$, and the optimal forecast $\hat{y}_t$ is the mean of the distribution F. For MAE, $L(y_t, \hat{y}_t) = |y_t - \hat{y}_t|$, and the optimal forecast $\hat{y}_t$ is the median of the distribution F.

If the MSE is considered in a wind forecasting problem, then the mean of the predictive distribution should be used. Reciprocally, if the mean of the predictive distribution is the predictor of the true value, then the MSE should be used to evaluate the prediction accuracy. Similarly, when the loss function is MAE, then the median of the predictive distribution should be used. It would be misleading to compare, for example, the MSE of the mean predictor from one forecasting model

with the MSE of the median predictor from another model.

2.5.2 Realistic Loss Functions for Wind

Besides ensuring that the point forecasts and loss functions match, it is still needed to consider an appropriate choice of loss functions for wind speed forecasting. Since short-term wind speed forecasting plays a critical role in system operations of wind power, both underestimation and overestimation of wind speed cause losses in practice. Two properties of loss functions should be taken into account:

1) *Penalization of underestimates.* Underestimates of wind power, resulting from underestimates of wind speed, make power system operators order too much energy in advance from conventional sources to meet the demand. Then down-regulation is needed which is more expensive than up-regulation (when overestimates happen). So underestimates of wind speed should be penalized more strongly than overestimates; see Pinson et al. (2007) for more detail.

2) *Penalization of forecasting errors for small true values.* Because the relative error is larger for small true values than for large ones when the prediction errors are the same, a loss function that penalizes errors for small true values more is preferred in wind speed forecasting. That is, for smaller true values, forecasts with lower relative errors is of the goal.

Neither MSE nor MAE have the above two properties. To evaluate the accuracy of wind speed forecasts, more realistic loss functions are needed. Hering and Genton (2010) proposed a new loss function, the power curve error (PCE). It links prediction of wind speed to wind power by a power curve and evaluates the loss based on the

wind power with penalty on underestimates as follows:

$$L(y,\hat{y}) = \begin{cases} \alpha\{g(y) - g(\hat{y})\}, & \text{if} \quad y \geq \hat{y}, \\ (1-\alpha)\{g(\hat{y}) - g(y)\}, & \text{if} \quad y < \hat{y}, \end{cases}$$

where $g(\cdot)$ is a nondecreasing function linking wind speed to wind power. It has the α-quantile as its optimal forecast (Gneiting 2011b). This loss function puts a penalty on underestimates with weight α, which depends on market rules. Hering and Genton (2010) set the penalty to $\alpha = 0.73$ based on empirical data from the Dutch electricity market in 2002. The PCE can penalize underestimates more heavily than overestimates through the weight α. Errors on small true wind speeds are only partly more penalized through the power curve transformation; see Figure 2.3.

The mean absolute percentage error (MAPE), corresponding to the loss function $L(y_t, \hat{y}_t) = |y_t - \hat{y}_t|/y_t$, is used as a measure of forecast accuracy in time series. MAPE agrees with the two aforementioned properties, namely penalizing underestimates and errors on small true values. Hence it would be a reasonable measure of accuracy for wind speed forecasting. However, its values vary in the interval $[0, \infty)$ for nonnegative wind speed and nonnegative forecasts. And for nonnegative underestimates their losses are less than 1, but for overestimates they can be very large. There is also a problem for true values close to zero. When the actual value is small, it can have large relative errors and make the MAPE meaningless. Both problems were solved by Armstrong (1985) and Flores (1986) with the mean symmetric absolute percentage error (MSAPE) based on the loss function $L(y_t, \hat{y}_t) = 2|y_t - \hat{y}_t|/(y_t + \hat{y}_t)$. Besides satisfying the two above properties, MSAPE has values in $[0, 2]$ for nonnegative wind speed, but there is still a problem when both the forecast value and the actual value are close to zero. To avoid this issue, a modified MSAPE was suggested

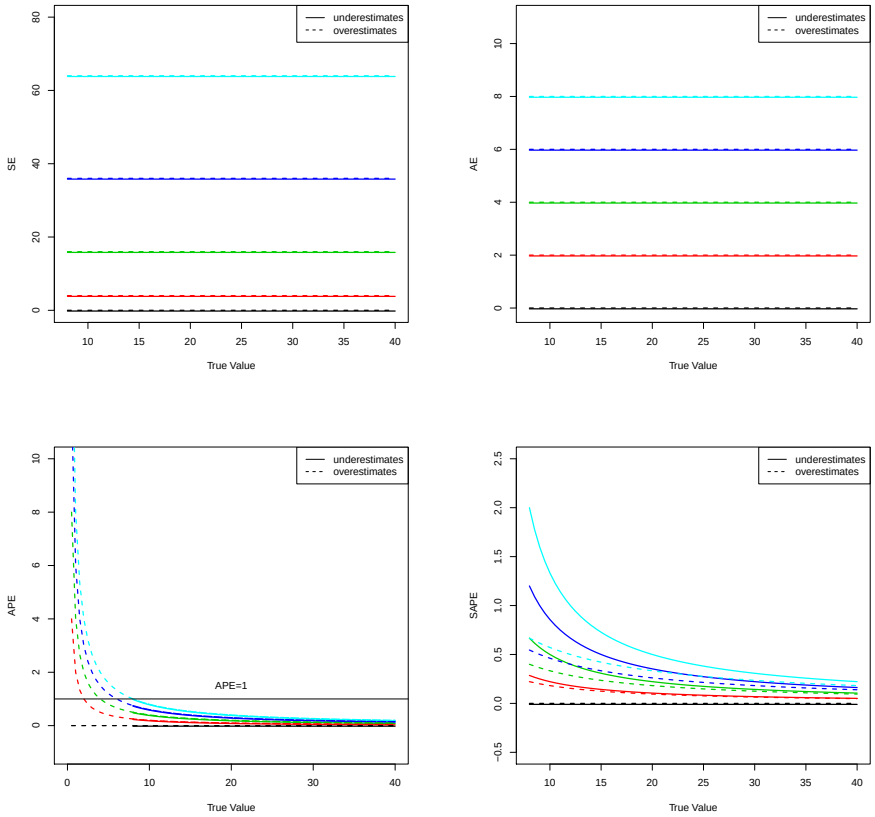

Figure 2.6: Squared error (SE), absolute error (AE), absolute percentage error (APE) and symmetric absolute percentage error (SAPE) for forecasts of true values $y = 8, 8.5, 9, 9.5, \ldots, 40$ with prediction errors 0 (black), 2 (red), 4 (green), 6 (blue), 8 (cyan) for each. Overestimates for extra true values $0, 0.5, 1, \ldots, 7.5$ with prediction errors $0, 2, 4, 6, 8$ are generated for APE.

by Chen and Yang (2004) by adding a nonnegative term to the denominator. Unfortunately, neither MAPE nor MSAPE have closed form for their optimal forecasts, although they could be obtained from (2.8) via simulations from F.

Figure 2.6 illustrates the differences between MSE, MAE, MAPE and MSAPE based on a numerical experiment: five nonnegative forecasts are generated for each true value $y = 8, 8.5, 9, 9.5, \ldots, 40$ with prediction errors $0, 2, 4, 6, 8$. To see that overestimates can result in very large APE values, overestimates for an additional

set of true values $0, 0.5, 1, \ldots, 7.5$ with errors $0, 2, 4, 6, 8$ are generated. The squared error (SE), absolute error (AE), absolute percentage error (APE), and symmetric absolute percentage error (SAPE) are calculated for each forecast. In the plots, the colors represent the five different prediction errors: 0 (black), 2 (red), 4 (green), 6 (blue), 8 (cyan). It is easy to see from the top two plots in Figure 2.6 that the SE and AE are the same for underestimates and overestimates, and for the same prediction error (or the same color), errors on smaller true values have the same loss as on larger ones. Unlike SE and AE, the APE and SAPE in the bottom two plots decrease for the same error when the true value gets larger. Also, for the same true value, the SAPE of underestimates are larger than those of overestimates, and the APE of underestimates are less than 1, while the overestimates can have very large APE values.

With the transition from point forecasting to probabilistic forecasting, verification of probabilistic forecasts has been developed in recent years. Gneiting et al. (2008) introduced methods to assess probabilistic forecasts of multivariate quantities: Box's density ordinate transform works for density forecasts and the multivariate energy score can be used in comparing density forecasts using a proper loss function. CRPS is also used to compare forecasting results in Gneiting et al. (2006) and Pinson and Hagedorn (2011).

2.5.3 Comparison of Forecast Accuracy

Given a certain measure of prediction accuracy or loss L and evaluation period k, one approach to compare forecasting models is by calculating their improvement relative to a reference model, such as persistence forecasting. Specifically,

$$\mathrm{Imp}_L^{\mathrm{ref}}(k) = \frac{L_k^{\mathrm{ref}} - L_k}{L_k^{\mathrm{ref}}},$$

where Imp denotes improvement, L_k^{ref} is the loss of the reference forecasting model during the time period k, L_k is the loss of an advanced model, and L can be any of the above loss functions. Negative values of improvement indicate that the advanced forecasting model performs worse than the reference model with respect to a certain loss, and positive values of improvement mean that the advanced model achieves better results in forecasting than the reference model with respect to that loss; see Monteiro et al. (2009).

Since the improvement may be due to chance, the relative improvement mentioned above is still not sufficient to judge the performance of models, and a statistical test is necessary to see whether the improvement is significant or not under a certain measure of forecasting quality. Meese and Rogoff (1988) tested for equal MSEs. Diebold and Mariano (1995) proposed a statistical test for equal predictive accuracy of two time series models. A robust version of the test was introduced by Dell'Aquila and Ronchetti (2004). Tests for the evaluation of point, interval, probability, and density forecasts were generalized by Giacomini and White (2006). Recently, Hering and Genton (2011) extended the test to compare predictions in spatial fields.

2.5.4 Uncertainty of Forecasts

As explained in Section 2.2.2, probabilistic wind speed forecasting corrects the drawbacks of point forecasting by fitting a probability distribution to describe the uncertainty of future wind speeds, from which prediction intervals can be obtained. Usually, in time series wind speed forecasting models, the prediction errors are assumed to follow a Gaussian distribution, based on which forecast confidence intervals are built. In reality, gaussianity may not hold and wind speed prediction errors may be skewed and heavy-tailed.

Transformations of wind speed, such as logarithm or square root, are commonly used before fitting probabilistic forecasting models with a normality assumption. However, the transformed wind speed may still not necessarily follow a normal distribution. Therefore, more adequate predictive distributions are needed, as discussed in Section 2.2.2. To evaluate the quality of probabilistic forecasts, Gneiting and Raftery (2007) pointed out that the goal of probabilistic forecasting is to maximize the sharpness of the predictive distribution subject to calibration. In terms of prediction intervals, this means that the probabilistic forecasts are better if the intervals are shorter, subject to nominal coverage. More discussion about evaluation of probabilistic forecasts can be found in Monteiro et al. (2009) and Gneiting et al. (2008). Probabilistic wind speed forecasting models based on ensembles are also developed, in which multiple point forecasts are used to generate a probabilistic forecast; see Thorarinsdottir and Gneiting (2010) and Sloughter et al. (2010). In this dissertation, the focus is on building probabilistic forecasting models for the uncertainty in wind based on wind records instead of ensembles. In Section 2.6, another difficult aspect of forecasting the uncertainty in wind due to ramp events is discussed.

2.6 Discussion

2.6.1 Wind Ramp Events

A ramp refers to a phenomenon where a sudden rapid increase or decrease in wind speed occurs. Ramp events can cause severe problems for power system operations with a high proportion of wind power. When a substantial, sudden drop in wind speed occurs in a very short time, another conventional power resource may not be started in time to meet the demand. However, as soon as the wind generation resumes, that conventional power resource is no longer needed. As the installation of capacity of wind power increases, the size of ramp events raises as well. It increases

the cost of power system operations, and challenges the reliability and security of systems that have integrated high amounts of wind power.

A big challenge in integration of wind power is to identify and manage ramp events. Recently, Kamath (2010) described three definitions of ramp events and offered three options about how to count ramp events. Time-of-day and month-of-year ramp event patterns of power generation were examined based on data sets from two locations. It was found that there was no significant difference between the three definitions of ramp events. Although ramp event patterns did depend on the location, it was still difficult to draw any conclusion about the time-of-day and month-of-year patterns of ramp events. Bossavy et al. (2010) gave a new definition of ramp events based on filtered signals of wind power, and proposed two methods to forecast uncertainty related to ramp events. The first method attempted to take into account additional variables of ramp information into probabilistic wind power forecasting, including intensity and the time difference calculated from preliminary point power forecasts. It was found to work well only for the highest quantile forecasts. The second model used ensembles to forecast the uncertainty in ramp events and their timing, and was found to have more skill than the climatology.

Studies about ramp events directly through wind speed are in need. Statistical models that can forecast ramp events would be of great interest.

2.6.2 Offshore Wind Speed Forecasting

Offshore wind energy exploitation is also an important part in wind power development, since the wind offshore is often stronger and less variable. DOE released a comprehensive report on offshore wind power in the U.S. (DOE 2010) and discussed the current status and future plans on developing offshore wind power. Although the current primary challenge of offshore wind energy lies in minimizing the cost of the

turbine installations in the offshore environment, according to the report, wind speed forecasting also plays an important role in the expansion of offshore wind energy.

Moreover, offshore winds have some special conditions due to, for example, the thermal air-sea interaction, dynamic wind-wave interaction, internal stratification of the marine boundary layer, and displaced height of the marine boundary layer. Tambke et al. (2003) and Tambke (2004) tried to adapt onshore short-term wind power predictions to offshore sites by refining the numerical weather prediction with surface roughness and thermal stability, and found that their methods worked very well. Rugbjerg et al. (2006) introduced wave forecasting for offshore wind farms. Pinson and Hagedorn (2011) applied Markov-switching autoregressive (MSAR) models to model and forecast offshore wind power fluctuations and found them to be better than persistence and AR-model-based forecasts on time series wind power data with a 10-minute resolution at Horns and Nysted Rev, Denmark. More research is needed on wind speed forecasting for offshore wind farms.

2.6.3 Final Remarks

this dissertation provided some information about the global overwhelming development of wind energy as a clean, renewable resource with its unmatched benefits, as well as big challenges to current power system operations due to the wind's intermittent and unstable nature. To include a high proportion of wind power in an energy mix, wind power forecasting is often identified as a necessary tool. Accurate wind power forecasting is closely related to power unit commitment and dispatch, and is the key to secure and stable power system operations.

Instead of wind power forecasting, this dissertation has focused on wind speed forecasting, because in power system operations, short-term wind power forecasting is obtained directly from predicted wind speeds through a certain power curve, which

changes with the size of turbines. This means that wind speed forecasting is more general in practice than forecasting wind power. Given historical wind speed data, forecasting future wind speed is a typical time series problem. Some classic time series statistical forecasting models are reviewed, including Kalman filters and ARMA models.

Recently, more advanced forecasting models taking into account spatial correlations by considering other neighboring sites have been developed. The RSTD and TDD models in detail is presented. Space-time forecasting models performed better than traditional time series models, and will be the new trend for developing wind speed forecasting models. Moreover, the RSTD and TDD models deliver probabilistic forecasts by fitting a truncated normal distribution for the future wind speed, which provides much more information about the uncertainty of the forecast, for instance using prediction intervals. Hence, they allow power system operations to work in a more secure manner.

Evaluation of forecasts is also a major issue discussed in this article. As more and more forecasting methods are proposed, an appropriate criteria is important in decision making. Matching between point estimates and loss functions should be emphasized in evaluation. Underestimates of wind speed are not preferred in realistic power system operations, and this should be penalized in evaluation (similarly for errors on small true values). Developing more realistic loss functions is of interest for evaluating forecasting results.

Finally, ramp events create difficulties that cannot be overlooked for accurate wind power forecasting and advanced power system operations. Challenges accompanying offshore wind speed forecasting with the new large installations of offshore wind farms must also be addressed in the near future. Extreme wind speed forecasting is also important to power system operations. Under high wind speeds, wind

turbines must be turned off to avoid possible damages, and power system operations must generate energy from other power resources to compensate. Distributions that allow extreme wind speeds, such as multivariate t, would be of interest for future research. Additionally, Bayesian methods for parameter estimation in wind speed forecasting problems, and kriging methods to forecast wind speed and wind direction at unrecorded locations for planning new wind farms, are also interesting future research topics.

3. ROTATING SPACE-TIME REGIME-SWITCHING WIND SPEED FORECASTING FOR IMPROVED POWER SYSTEM DISPATCH

3.1 Introduction

3.1.1 Wind Energy

Renewable energy, particularly wind energy, is rapidly being integrated into electric power systems throughout the world. In Denmark, wind has become one of the largest sources of electricity, supplying 26% of electricity demand in 2011. In Spain, 15.9% of electricity consumption was generated by wind in 2011, along with 15.6% and 10.6% in Portugal and Germany, respectively, according to the European Wind Energy Association (EWEA 2012). The United States (US) Department of Energy (DOE) published a report in 2008 that described a model-based scenario in which wind energy would provide 20% of US electricity demand by 2030 (DOE 2008). China is pursuing a total capacity of 150 Gigawatts (GW) from wind energy by 2020, 250 GW by 2030 and 450 GW by 2050 (CREIA (2010); see the review by Zhu and Genton (2012) for more information about global wind energy).

Nevertheless, due to the high variation and limited predictability of wind, current power system scheduling methods face profound challenges in integrating large-scale wind power. The basic objective of power system scheduling is to maintain a supply and demand balance at minimum cost, subject to transmission constraints and plausible contingencies. Prior to the introduction of renewable energy sources, such as wind and solar, uncertainty in power system scheduling primarily came from the demand side (Xie et al. 2011a). Now, with the introduction of intermittent wind power, this uncertainty mainly comes from the supply side, given that state-of-the-art load forecasts are highly accurate in the day-ahead stage. The high uncertainty

from wind power generation increases the cost of a power system and reduces its stability and reliability. Highly accurate wind power forecasts are therefore very much needed. Otherwise, any potential savings from wind power would be largely offset by the balancing and ancillary costs provided by fast-response fossil fuel units. An overview of the major technical challenges in power system operations that integrate large-scale wind energy was presented by Xie et al. (2011a).

3.1.2 Wind Speed Forecasting

Highly accurate wind speed prediction is crucial to reducing the uncertainty from the supply side in the power system scheduling. Compared with long-term prediction, short-term forecasting is more accurate and reliable. It is also essential to effective power system operations planning. Hours-ahead wind forecasting gives conventional power sources enough time to start and provide power as demanded on time. Typically, this means between 1 and 10 hours, but it can be under 1 hour if the start-up time for the back-up system is short; see Genton and Hering (2007). Genton and Hering (2007) pointed out that wind power forecasts by converting wind speed forecasts based on a power curve is more general than predicting wind power generation directly. Here, we focus on short-term wind speed forecasting.

Extensive research has been devoted to wind power forecasting problems. Giebel et al. (2011), Kariniotakis et al. (2004), Monteiro et al. (2009) and Zhu and Genton (2012) reviewed approaches to wind power forecasting, including physical methods, statistical models and combined physical-statistical systems. In short-term wind speed forecasting, statistical models have been found to be quite competitive compared with other approaches. Moreover, statistical models that incorporate spatial information are more accurate than the conventional time series models (see Zhu and Genton (2012) for a review). Gneiting et al. (2006) proposed a regime-switching

space-time diurnal (RSTD) model to forecast 2-hour-ahead wind speed at Vansycle, Oregon. Their model outperformed persistence forecasts and autoregressive forecasts by 29% and 13%, respectively, in terms of the root mean squared error (RMSE) in July 2003, for instance. However, the RSTD model relies on local geographic features. To remove these constraints, Hering and Genton (2010) generalized the RSTD model by treating wind direction as a circular variable and including it in their model. They coined it a trigonometric direction diurnal (TDD) model. The TDD model obtained similar or better forecasting results than did the RSTD model without requiring prior geographic information. Tastu et al. (2011) analyzed and modeled short-term wind power forecast errors using spatio-temporal methods, such as regime-switching models based on wind direction and conditional parametric models with regime-switching, substantially reducing variance in the forecast errors. Pinson and Madsen (2012) applied adaptive Markov-switching autoregressive models to offshore wind power forecasting problems in which the regime sequence is not directly observable but follows a first-order Markov chain. Here, a new modification of the RSTD model is proposed to allay its limitations.

Model evaluation is also an important step in making a final decision on which model should be implemented. Usually, a loss function is predefined and the model that can generate forecasts with the smallest loss is considered to be the most advanced. Squared and absolute errors are two commonly used loss functions. However, for wind forecasting problems, more realistic loss functions are needed since penalization on underestimates and forecasts for small true values are desired; see Zhu and Genton (2012) for a more detailed discussion. Hering and Genton (2010) proposed the power curve error as a loss function, which links prediction of wind speed to wind power by a power curve and evaluates the loss based on the wind power with penalty on underestimates. Zhu and Genton (2012) introduced the mean absolute percentage

error and the mean symmetric absolute percentage error as loss functions to penalize both underestimates and forecasts for small true values. In this chapter, a new idea for model evaluation based on the power system operating costs is proposed. Since the ultimate goal is to reduce the cost of the whole power system, it is natural to look for the forecasts generated from a model that produces the most cost savings.

In summary, the main contributions of this chapter are the following:

1) A modified RSTD model for short-term wind speed forecasting is proposed. It generalizes the RSTD model by allowing the forecast regimes to vary with the dominant wind direction in each season instead of fixing the forecast regimes based on prior geographic information. In the original application of the RSTD model, it was straightforward to define west and east forecast regimes due to prevailing westerly winds in the target area. However, for other settings where the winds follow more complicated patterns, the number and position of the forecast regimes are difficult to determine. In the modified model, the best position of the forecast regimes is detected by rotating the dividing angles of the regimes until the minimum MAE for each season is reached. This new model is called RRSTD for rotating RSTD.

2) To evaluate the model, an economic dispatch model is formulated for power systems that incorporates space-time wind forecast information. Numerical simulations are conducted in a representative test system derived in the northwest region of the US and the results demonstrate the economic benefits from improved wind forecasts.

this chapter is organized as follows. In Section 3.2, we first introduce our modified space-time statistical model for short-term wind speed forecasting, the RRSTD model, and then we describe persistence and autoregressive models as references for

later comparisons. In Section 3.3, the newly proposed RRSTD model is then applied to a spatio-temporal wind data set from the northwest region of the US. Its prediction MAE values for each month are compared with reference models. In Section 3.4, an economic dispatch model is proposed that incorporates available short-term, space-time wind power forecasts. An illustrative power system economic dispatch example for the Pacific northwest is presented, which quantifies the potential savings in both generation costs and ancillary services in the proposed dispatch model. Concluding remarks are provided in Section 3.5.

3.2 The Rotating RSTD Model

In this section, the RRSTD model is described in detail while the RSTD model is included as a special case. For more detailed introduction about the RSTD model, readers can refer to Section 2.4.2. Two reference models are also briefly introduced.

3.2.1 RRSTD Model Description

Let $y_{s,t}$ and $\theta_{s,t}$ respectively be wind speed and direction at site s and time t, where $s = 1, \ldots, S$, $t = 1, \ldots, T$, and $\theta_{s,t} \in [0°, 360°)$ with $0°, 90°, 180°, 270°$ indicating southerly, easterly, northerly and westerly winds, respectively. The objective is to predict $y_{s,t+k}$, the k-step-ahead wind speed at site s, where $k = 1, 2, \ldots$. When $k = 1$, for example, depending on the resolution of the wind data, it is a 1-hour-ahead forecasting for hourly wind data, and 10-minute-ahead forecasting for 10-minute wind data. To simplify, the RRSTD model is presented in the setting of forecasting wind speed k-step-ahead at a site s_1, say.

Since wind speed is non-negative and has large values with low probabilities (right-skew distributed), it is assumed that $Y_{s_1,t+k}$ follows a truncated normal distribution (Gneiting et al. 2006) with center and scale parameters $\mu_{s_1,t+k}$ and $\sigma_{s_1,t+k}$: $Y_{s_1,t+k} \sim N^+(\mu_{s_1,t+k}, \sigma_{s_1,t+k})$. To predict $y_{s_1,t+k}$ precisely, the key lies in appropriately

modeling $\mu_{s_1,t+k}$ and $\sigma_{s_1,t+k}$.

Generally, seasonal and diurnal patterns are observed in winds. In the RRSTD model, the diurnal pattern is fitted with two pairs of harmonics as

$$D_{s_1,h} = d_0 + d_1 \sin\left(\frac{2\pi h}{24}\right) + d_2 \cos\left(\frac{2\pi h}{24}\right) + d_3 \sin\left(\frac{4\pi h}{24}\right) + d_4 \cos\left(\frac{4\pi h}{24}\right), \quad (3.1)$$

where h indicates the hour of a day, $h = 1, 2, \ldots, 24$, and the coefficients are estimated by the least squares method. Then, the center parameter is modeled as $\mu_{s_1,t+k} = D_{s_1,t+k} + \mu^r_{s_1,t+k}$, where $\mu^r_{s_1,t+k}$ is the residual wind speed after removing the diurnal pattern.

The residual, $\mu^r_{s_1,t+k}$, is modeled by a linear combination of historical wind speed residuals, up to p-step lags, of itself as well as its neighbors (to take the spatio-temporal correlations in wind into account), allowing the coefficients to vary with the dominant wind direction and season by defining the variable forecast regimes as

$$\mu^r_{s_1,t+k} = \alpha_0 + \sum_{s=1}^{S} \sum_{j=0}^{p} \alpha_{s,j}\left(\theta_{s^*,t}, \boldsymbol{\theta}^*_{m(t+k)}\right) \mu^r_{s,t-j}, \quad (3.2)$$

where α_0 and $\alpha_{s,j}(\cdot,\cdot)$, $s = 1, \ldots, S$ and $j = 0, \ldots, p$ are coefficients, and $\boldsymbol{\theta}^*_{m(t+k)} \subseteq [0°, 360°)$ defines the forecast regimes based on the prevailing wind direction in the season (here month), $m(t+k)$, to which time $t+k$ belongs. Here, $\theta_{s^*,t}$ is the current wind direction at site s^* used to indicate the direction of the nearby future wind. The site $s^* \in \{1, \ldots, S\}$ is located upstream of the wind and indicates the wind source.

The meaning of the above model is that, for a certain season, if the future wind direction at the target site, which is estimated by $\theta_{s^*,t}$, falls into a predefined forecast regime, a particular space-time linear model will be applied to estimate $\mu^r_{s_1,t+k}$, and

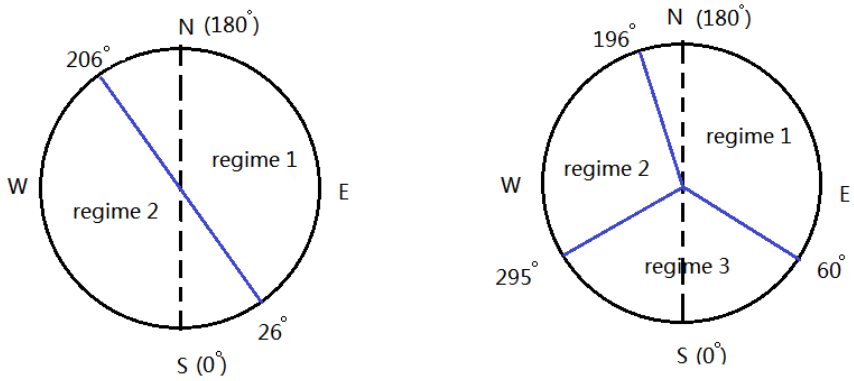

Figure 3.1: Regime dividing plots for $\boldsymbol{\theta}^*_{m(t+k)} = \{26°, 206°\}_{Aug}$ (left) and $\boldsymbol{\theta}^*_{m(t+k)} = \{60°, 196°, 295°\}_{Aug}$ (right). The dashed line connects the south (0°) and the north (180°), with the westerly wind to the left and the easterly wind to the right. Separate models of $\mu^r_{s,t+k}$are built for each regime.

the forecast regimes will be based on the dominant wind in that season. For example, if $\boldsymbol{\theta}^*_{m(t+k)} = \{26°, 206°\}_{Aug}$ and $s^* = s_2$, then, in August, the RRSTD model fits two separate models for the center parameters $\mu^r_{s_1,t+k}$: model 1, when the current wind direction at site s_2 is between 26° and 206°, or $\theta_{s_2,t} \in [26°, 206°)$; model 2, when $\theta_{s_2,t} \in [206°, 360°) \cup [0°, 26°)$; see Figure 3.1 (left panel). The dimension of $\boldsymbol{\theta}^*_{m(t+k)}$ indicates the number of regimes that are defined. For $\boldsymbol{\theta}^*_{m(t+k)} = \{60°, 196°, 295°\}_{Aug}$ and $s^* = s_2$, three separate models are built for the three forecast regimes divided by these angles; see Figure 3.1 (right panel). The scale parameter $\sigma_{s_1,t+k}$ is modeled as

$$\sigma_{s_1,t+k} = b_0 + b_1 v_{s_1,t}, \tag{3.3}$$

where $b_0, b_1 > 0$ and $v_{s_1,t}$ is the volatility value: $v_{s_1,t} = \{\frac{1}{2S}\sum_{s=1}^{S}\sum_{i=0}^{1}(\mu^r_{s,t-i} -$

$\mu^r_{s,t-i-1})^2\}^{1/2}$.

A key point of the RRSTD model is how to decide the number and the position of the regimes. For locations that have significant prevailing wind, this can be determined practically (see Section 3.3.2). The RSTD model is a special case of the RRSTD model with $\boldsymbol{\theta}^* = \{0°, 180°\}$, motivated by the westerly prevailing wind in the northwest region of the US. For other situations, it is proposed that $\boldsymbol{\theta}^*_{m(t+k)}$ be chosen by minimizing the prediction MAE for each season/month after determining the number of regimes. The predictors in (3.2) are selected by the Bayesian Information Criterion as in Hering and Genton (2010). The coefficients in (3.2) along with b_0, b_1 in (3.3) are estimated by means of the continuous ranked probability score method (see Gneiting and Raftery (2007) for more details).

With the estimated predictive distribution, $N^+(\mu_{s_1,t+k}, \sigma_{s_1,t+k})$, the median of the truncated normal distribution is taken as the wind speed forecast k-step-ahead at s_1, defined as

$$z^+_{0.5} = \mu_{s_1,t+k} + \sigma_{s_1,t+k} \cdot \Phi^{-1}\Big\{1 + \Phi\Big(\frac{-\mu_{s_1,t+k}}{\sigma_{s_1,t+k}}\Big)\Big\}\Big/2,$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

3.2.2 Reference Models

To evaluate the performance of the RRSTD model, its forecasts are compared with other models, including the persistence (PSS) and autoregressive (AR) models. The main ideas of the two reference models are introduced briefly in the following. More detailed introductions are presented in Section 2.3.2.

- PSS assumes that the future wind speed is the same as the current one, or $\hat{y}_{s_1,t+k} = y_{s_1,t}$.

- An AR(p) model estimates $\mu^r_{s_1,t+k}$ in (3.2) as a linear combination of the previous p wind speed residuals from the same location only, or $\mu^r_{s_1,t+k} = \alpha_0 + \sum_{j=0}^{p} \alpha_j \mu^r_{s_1,t-j}$. For the scale parameter, a GARCH(1,1) model is used instead of (3.3); see Gneiting et al. (2006).

Due to the high variations in wind, PSS works better for very short-term forecasting, such as 10-minute-ahead predictions. The AR(p) model can capture the temporal correlation in wind patterns and usually outperforms PSS in short-term wind speed forecasting problems.

3.3 Numerical Experiments

3.3.1 Wind Data

The data considered here are 10-minute wind speed (m/s) and direction (degrees) records from three meteorological towers located at Vansycle (Oregon), Kennewick (Washington), and Goodnoe Hills (Washington) in the northwest region of the US. Missing data were imputed by linear interpolation. Detailed information about the data and the three sites can be found in Gneiting et al. (2006).

The training and testing data sets were divided as follows:

- Training set: data from 1 August to 30 November 2002. With the training data, for each month, the regime dividing angles, $\boldsymbol{\theta}^*$, and the wind source indicator, s^*, in (3.2) are learned by minimizing the prediction MAE values. Then, in each forecasting regime, linear models for the center parameter are obtained.

- Testing set: data from 25 February to 30 November 2003. The trained models are evaluated during this period. The parameters in the models are estimated from data that are up to 45 days earlier, as suggested by Gneiting et al. (2006).

3.3.2 Exploratory Data Analysis

An exploratory data analysis was conducted on the relationship between wind speed and wind direction with the aim to determine the number of forecast regimes and how to divide the regimes in the RRSTD model. The wind roses in Figure 3.2 give a view of how wind speed and wind direction are distributed each month from August to November 2002 at Vansycle (left column), Kennewick (middle column), and Goodnoe Hills (right column). In a wind rose, each petal indicates the frequency of winds blowing from a particular direction and the color bands in each petal show the range of wind speeds. The vertical dotted lines give the west-east forecast regimes of the RRSTD (i.e., RSTD) model, while the blue dashed lines give the two forecast regimes of the two-regime RRSTD models with minimum prediction MAE values for each month for 1-hour-ahead forecasting.

As we can see from the wind roses, the wind patterns in this area are quite significant. High frequencies and wide speed ranges are found in the northwest, north and west direction at Vansycle, Kennewick and Goodnoe Hills, respectively, over the four months, followed by winds from the opposite directions. These are consistent with the geographic features in this area, namely that these three locations are along the south, southwest, and north bank of the Columbia River, which runs from east to west along the boundary between Washington and Oregon, with high terrain in both the north and south restricting the air flow.

Based on these wind patterns, using two forecasting regimes is reasonable. In our experiment, two equally divided regimes is used, resulting in a two-regime RRSTD model. More complex RRSTD models with two regimes of different sizes or with multiple (more than two) forecasting regimes could be considered as well. Our exploration of those alternate models revealed insignificant improvements for this

particular data set.

3.3.3 Training Data Results

The two-regime RRSTD model for each month was trained based on the training data set at all three locations for 1-hour-ahead wind speed forecasting. The gains from using the RRSTD model instead of the RSTD are shown in the MAE values in each month in Figure 3.3.

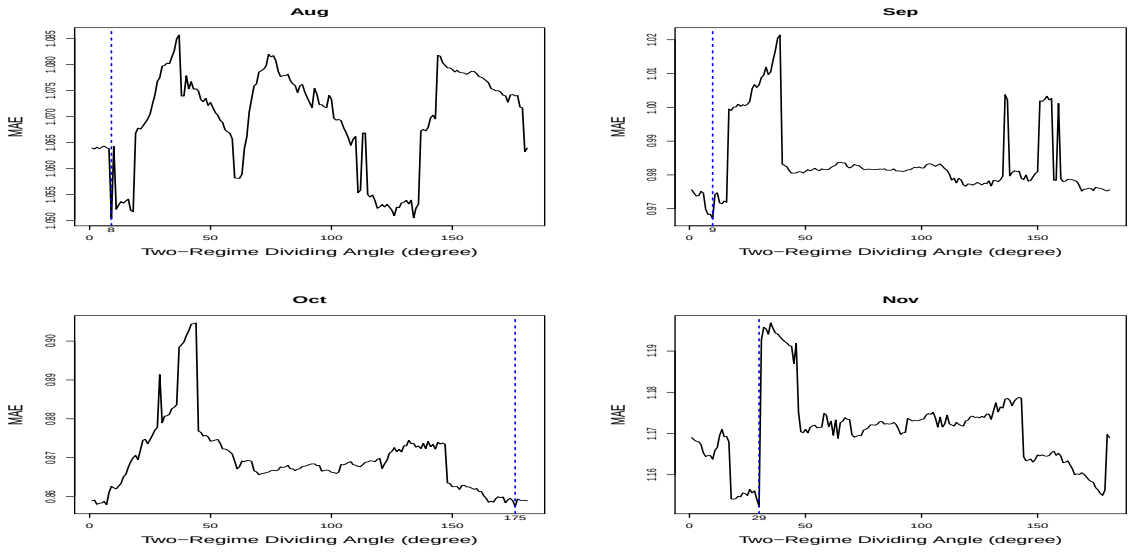

Figure 3.3: Plots of 1-hour-ahead prediction MAE results based on the two-regime RRSTD model with the dividing angle, θ, from 0° to 180° for each month at Vansycle in 2002. The blue dashed line indicates the position of the best two-regime dividing angle, or θ^*, that has the smallest MAE value.

Let $y_{V,t}$, $y_{K,t}$, $y_{G,t}$, $\theta_{V,t}$, $\theta_{K,t}$, and $\theta_{G,t}$ denote the wind speed and direction at Vansycle, Kennewick and Goodnoe Hills. The goal here is to predict $y_{V,t+6}$, $y_{K,t+6}$ and $y_{G,t+6}$ (1-hour-ahead is equal to 6-steps-ahead in 10-minute data). The ob-

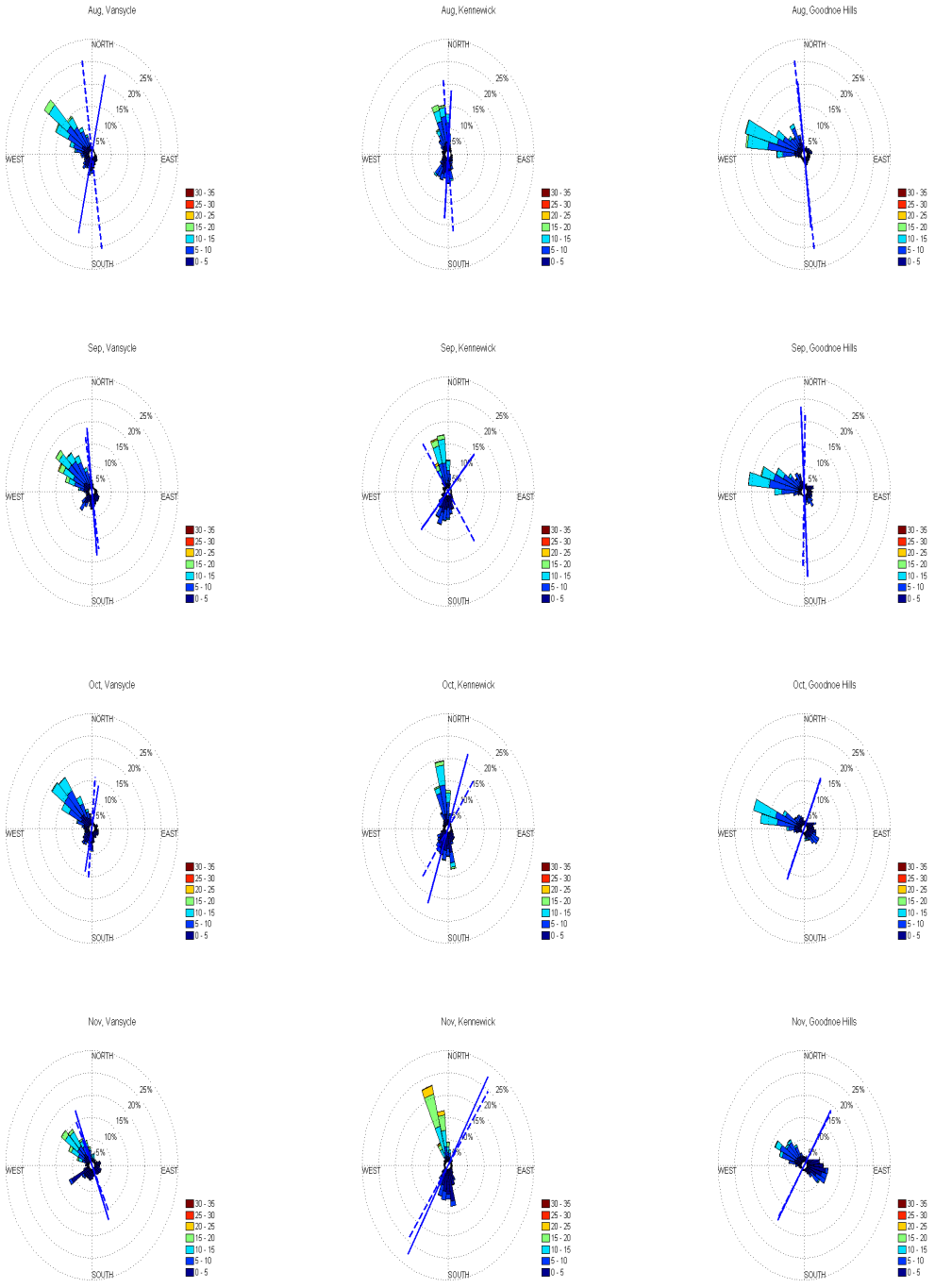

Figure 3.2: Wind roses of data from August to November 2002 at Vansycle (left column), Kennewick (middle column) and Goodnoe Hills (right column).

jective of the training procedure is to find the two-regime dividing angle, $\theta^*_{m(t+6)}$, the wind source indicator, s^*, and the predictors for each forecast regime in (3.2) with minimum prediction MAE value. To do this, a dense number of possible two-regime forecasting designs are tested. Specifically, for a possible dividing angle, $\theta \in \{1^\circ, 2^\circ, \ldots, 180^\circ\}$, a wind source indicator, s^*, is detected with the method used in Gneiting et al. (2006); then, a separate forecasting model is built for each forecasting regime, resulting in prediction MAE values as displayed in Figure 3.3 at Vansycle. The blue dashed line in Figure 3.3 gives the best two-regime dividing angle, θ^*, with the smallest MAE value.

Note that the RSTD model is a special case of the RRSTD model when the regime dividing angle, θ, is equal to 0° or 180°. As shown in Figure 3.3, the minimum MAE value occurs neither at $\theta = 0^\circ$, nor at $\theta = 180^\circ$, while the RRSTD model achieves the smallest MAE value at $\theta = 8^\circ, 9^\circ, 175^\circ$ and 29° for August, September, October, and November, respectively, at Vansycle. This indicates that, although the westerly wind dominates this area, a simple west-east forecast regime is not the best and adjustment is needed for different seasons to achieve more accurate forecasts. The RRSTD model is able to adjust the forecasting regimes to the wind roses shown in Figure 3.2 based on wind direction and season. Similar training results are found for the other two locations.

3.3.4 Testing Data Results

The trained two-regime RRSTD model for each month is applied to forecast 1-hour-ahead wind speed in the same month in the testing data set at all three locations, and the prediction MAE values are compared with the two reference models (see Table 3.1). Due to data limitations, the model for August is implemented to forecast wind speed in May, June and July in the testing data, and a 45-day training

period is used to estimate the coefficients in (3.2). The results show that the RRSTD model outperforms the PSS and AR models as expected. The latter is fitted with a maximum order of 9 based on the Akaike Information Criterion.

Overall, the RRSTD model outperforms PSS, reducing the MAE value by 8.1%, and by 6.6% compared with the AR model at Vansycle. 2-hour-ahead forecasting experiments and experiments with the RSTD and TDD models are conducted, and similar results overall were obtained. In fact, with only four months of training data, the forecasting ability of the RRSTD model is challenged by the assumption that monthly patterns remain similar in the training and testing data, while at least several years of wind data would be needed to model monthly patterns. It is believed that the performance of the RRSTD model would be better if more data were available.

TESTING		MAE							
Sites	Models	May	Jun	Jul	Aug	Sep	Oct	Nov	Overall
Vansycle	PSS	1.27	1.18	1.27	1.23	1.19	1.31	1.23	1.24
	AR	1.25	1.16	1.19	1.21	1.19	1.31	1.23	1.22
	RRSTD	**1.16**	**1.06**	**1.10**	**1.12**	**1.11**	**1.25**	**1.17**	**1.14**
Kennewick	PSS	1.43	1.28	1.39	1.34	1.26	1.44	1.35	1.36
	AR	1.46	1.27	1.36	1.33	1.25	1.44	**1.33**	1.35
	RRSTD	**1.40**	**1.23**	**1.33**	**1.30**	**1.23**	**1.44**	1.35	**1.33**
Goodnoe Hills	PSS	1.17	1.14	1.12	1.15	1.18	1.33	1.28	1.20
	AR	1.13	1.09	1.03	1.12	1.17	1.30	1.26	1.16
	RRSTD	**1.11**	**1.08**	**1.02**	**1.09**	**1.12**	**1.29**	**1.22**	**1.13**

Table 3.1: MAE values of forecasts in the 2003 testing data set based on the two-regime RRSTD models for 1-hour-ahead forecasting at Vansycle, Kennewick and Goodnoe Hills, compared with the PSS and AR models. The smallest MAE values are in bold.

3.4 Integrating Wind Power into a Power System

In this section, space-time wind forecasts are incorporated into electric power system scheduling. The system-wide generation cost savings are compared, as well as the ancillary service cost savings, using the RRSTD, AR, PSS forecasting models. First, a test system based on the Bonneville Power Administration (BPA) system, which covers the area where the wind data were collected, is introduced and studied for power system operation with space-time wind forecasts. Second, the power system dispatch problem that incorporates advanced spatio-temporal correlated wind forecasts are formulated. Finally, a numerical experiment is conducted and analyzed, and the performances of the different forecasting models are compared.

3.4.1 Power System Specification in the BPA Region

A power system economic dispatch model is used by system operators in scheduling power generation. This model determines the power generators' outputs to maintain a balance between supply and demand, as well as to minimize total system operating costs while satisfying security constraints. In this subsection, a detailed power system dispatch procedure is introduced based on the BPA system, which covers the areas of Vansycle, Kennewick and Goodnoe Hills.

Established in 1937, BPA is a nonprofit agency located in the Pacific northwest. About one-third of the electric power used in the northwest comes from BPA, which operates and maintains about 75% of the high-voltage transmission network (15,212 circuit miles) in its service territory (BPA 2010), which includes Idaho, Oregon, Washington, western Montana and small parts of eastern Montana, California, Nevada, Utah and Wyoming.

The major missions of BPA in operating electric energy are to: 1) act as an adequate, efficient, economical and reliable power supply; and to 2) maintain a trans-

mission system capable of integrating different power resources, providing electricity to its customers through inter-regional interconnections and maintaining electrical reliability and stability.

To balance demands for power, the output of every generator in the system has to be scheduled over different time frameworks (i.e., day-ahead, hour-ahead, and 5-10-minutes-ahead). The BPA scheduling procedure (Makarov et al. 2008) is shown in Figure 3.4. In the power generation scheduling process, the system operator at BPA schedules generators to meet the expected demand over several time scales. All the scheduled power generation must be within the output capacity, as well as within the ramping capacity, which refers to the maximum change in power generation output between two consecutive time intervals. For example, a natural gas generator's ramping capacity can be 15% of its maximum output in 10 minutes.

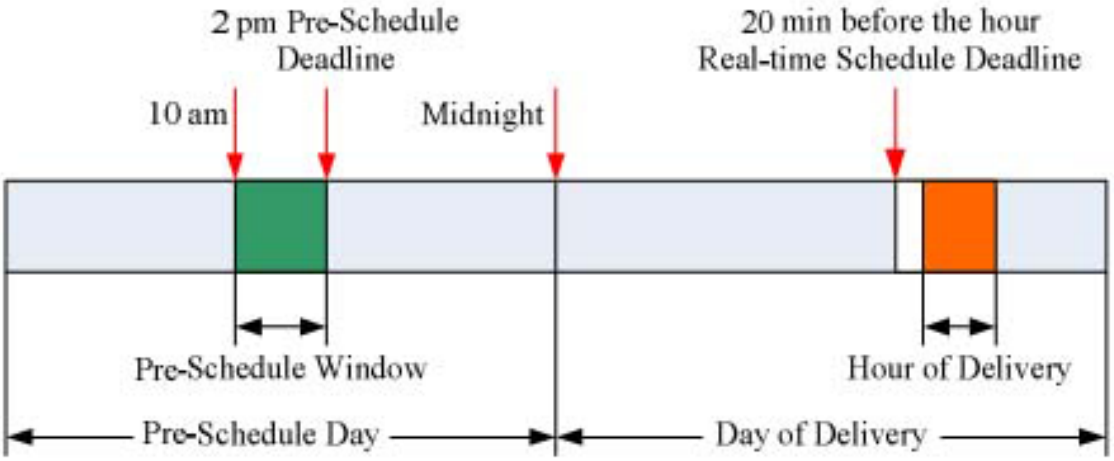

Figure 3.4: BPA's scheduling procedure (Makarov et al. 2008).

Given that it takes several hours to start up or shut down many large generators

(e.g., nuclear, coal), a day-ahead schedule (or pre-schedule) process is required to plan the generators' operations over the next 24 hours. Based on day-ahead forecasts, the pre-schedule is completed before 2:00 pm the day before the day-of-delivery (or the day on which the real-time operation takes place). However, the day-ahead load forecast and day-ahead wind forecast have relatively low accuracies so a real-time schedule that is 1-hour-ahead in BPA is required to discern a mismatch between the near-term forecast and the day-ahead forecast. The real-time schedule is established on the hour-ahead forecast, which has to be completed 20 minutes before the hour-of-delivery (the hour when the real-time operation takes place).

Within each hour, the available wind generation, as well as the electricity demand, still varies from second to second. Such an imbalance between total supply and total demand will cause degradation of the frequency of the electricity, which has very stringent requirements for the safety of many appliances. In order to maintain the system's electrical frequency at 60 Hz, automatic feedback control loops are installed at many generators' speed governors, which is referred to as the automatic generation control mechanism and is very similar in principle to the cruise control in automobiles.

3.4.2 Power System Dispatch with Space-Time Wind Forecasts

Motivated by the development of wind energy, there is now a large body of literature on understanding the impact of wind power on electricity grid operations. Watson et al. (1994) first introduced the numerical weather prediction (NWP) model for power system scheduling and evaluated its benefits (on-line reserve planning) to the England and Wales National Grid. Later, autoregressive moving average (ARMA) models were used for wind forecasting and incorporated into power dispatch models (Tuohy et al. 2009; Soder 2004). Developed from conventional criteria, probabilistic optimal dispatch methods were proposed to quantify the spinning/non-spinning

reserve requirements for integrating wind (Bouffard and Galiana 2008; Doherty and O'Malley 2005). In recent years, many efforts have focused on enhanced day-ahead power system operation using NWP models (Constantinescu et al. 2011; Pappala et al. 2009). To handle potential risks posed by wind generation, advanced dispatch methods such as robust optimization (Zhao and Zeng 2010) and stochastic optimization (Constantinescu et al. 2011; Wang et al. 2008; Wu et al. 2007; Meibom et al. 2011; Papavasiliou et al. 2011) based unit commitment (UC)/economic dispatch (ED) models were proposed and studied. Although there have been many different proposals on what should be an optimal dispatch method in future power systems, actual practice during real-time operations is still a *single-stage* security-constrained economic dispatch (SCED). Our aim is to assess the economic value brought by the RRSTD model using a well-accepted industry model in real-time power system operations. In other words, the power system dispatch model is assumed to be a single-stage SCED. Consequently, we neglect the time step index for decision variables and parameters in the formulation. The mathematical formulation of the single-stage SCED is described as follows with the notation listed in Table 3.2:

$$\min_{P_{G_i}, P_{W_i}, P_{R_i}} : \sum_{i \in G} C_{G_i}(P_{G_i}) + \sum_{i \in W} C_{W_i}(P_{W_i}) + \sum_{i \in G} C_{R_i}(P_{R_i}), \tag{3.4}$$

subject to:

$$\sum_{i \in G} P_{G_i} + \sum_{i \in W} P_{W_i} = \sum_{i \in D} P_{D_i}, \tag{3.5}$$

$$\sum_{i \in G} P_{R_i} \geq R_D + R_W, \tag{3.6}$$

$$|\mathbf{F}| \leqslant \mathbf{F}^{max}, \tag{3.7}$$

$$|P_{G_i} - P^0_{G_i}| \leqslant P^R_i \Delta T, i \in G \cup W, \tag{3.8}$$

$$P^{min}_{G_i} \leqslant P_{G_i} \leqslant P^{max}_{G_i}, \tag{3.9}$$

$$0 \leqslant P_{R_i} \leqslant P^{max}_{G_i}, \tag{3.10}$$

$$P^{min}_{G_i} \leqslant P_{G_i} + P_{R_i} \leqslant P^{max}_{G_i}, \tag{3.11}$$

$$P^{min}_{W_i} \leqslant P_{W_i} \leqslant P^{max}_{W_i}, \tag{3.12}$$

$$P_{W_i} \leqslant \hat{P}_{W_i}. \tag{3.13}$$

G	Set of conventional power plants
D	Set of inelastic loads
W	Set of wind farms
C_{G_i}	Generation cost function of power plant i
C_{W_i}	Generation cost function of wind farm i
C_{R_i}	Reserve cost function of power plant i
P_{G_i}	Scheduled generation of power plant i
P_{W_i}	Scheduled generation of wind farm i
P_{D_i}	Forecasted load level of bus i
P_{R_i}	Scheduled reserve capacity of power plant i
$\mathbf{F}$	Vector of branch flows
$\mathbf{F}^{max}$	Vector of capacity limits of transmission lines
$\triangle T$	Energy Market scheduling interval
P^R_i	Ramping constraints of power plants i
$P^{min}_{G_i}$	Lower operating limit of power plant i
$P^{max}_{G_i}$	Higher operating limit of power plant i
$P^{min}_{W_i}$	Lower operating limit of wind farm i
$P^{max}_{W_i}$	Higher operating limit of wind farm i
$\hat{P}_{W_i}$	Forecasted wind availability for wind farm i

Table 3.2: Notation for the power system dispatch model.

In the proposed formulation, the objective function (3.4) is to minimize the power system's operating costs, which include costs of power generation and costs of providing reserve and regulation services. The decision variables include the dispatched generation output for each generator, P_{G_i}, the dispatched generation output for each wind farm, P_{W_i}, the dispatched generation output for regulation and reserve capacity, P_{R_i}. Constraints on this problem include system and individual unit operating constraints posed by security and reliability. The energy balance equation (3.5) requires that the total power generation always satisfies the total demand in the steady state. The system's reserve and regulation requirements (3.6) are determined by the reliability requirement component of the load, R_D, and the reliability requirement component of wind generation, R_W. The load component is a linear function of actual system-wide load levels for each interval as practiced by major independent system operators (Electric Reliability Council of Texas 2010b). The reserve requirement related to wind generation, R_W, is a linear function of the MAE of the wind generation forecast error. The transmission line capacity limitations (3.7) contribute to network transmission congestion. The ramping constraints of generators are described by (3.8). The upper bounds and lower bounds of conventional generators' outputs are provided by (3.9). The available reserve and regulation capacity constraints are given by (3.10). The wind component is given by the deviation between the actual wind generation production potential and the wind generation forecast. This approach to quantifying system reserve requirements approximates the empirically-based approach to quantifying reserve requirements and serves as a lower bound for the reliability requirement due to wind forecast uncertainty. The capacity constraints of each generator for providing both energy and reserve services are in (3.11). The upper and lower bounds of wind farms' power output are described by (3.12). The wind forecast for each wind farm is provided by (3.13). This

is determined by a space-time forecast model, for instance the RRSTD.

3.4.3 A Realistic Illustrative Example

In this subsection, numerical simulations are performed in a test BPA system. The current real-time operational practice is adopted in the power industry, which is a single-stage, security-constrained economic dispatch model. The wind speed forecasts in Section 3.3 are converted into wind power forecasts with a 2.5 MW Nordex power curve for each wind turbine and scaled up to wind farms based on the BPA system setup. According to the economic dispatch results, different wind forecasting models are compared in terms of potential savings in both generation cost and ancillary services.

Vansycle, Kennewick and Goodnoe Hills are located in the Columbia River Basin. The electric power grid of this area is operated by BPA. Our simulation system is revised from the IEEE Reliability Test System (RTS-24) (Grigg et al. 1999). The generators are categorized as different technology-based power resources, such as hydro, coal, nuclear, natural gas and wind power. The generator capacity portfolio (installed capacity percentage of different technologies) is configured according to the generation portfolio of the practical BPA system (BPA 2010) (see Table 3.3). The network typology of the simulation system is presented in Figure 3.5.

The load profile used in the simulation is scaled from the historical load profile of the BPA system (BPA 2007). Fourteen typical days in seven months of different seasons are selected for simulation. The duration of a simulation is a typical power system operation period of 24 hours ($T = 144$). The different wind forecast methods described in Section 3.2 are implemented in the simulation. The operating interval, ΔT, of generation scheduling is 10 minutes. Wind profiles during the selected fourteen days are collected from the BPA system and scaled to the simu-

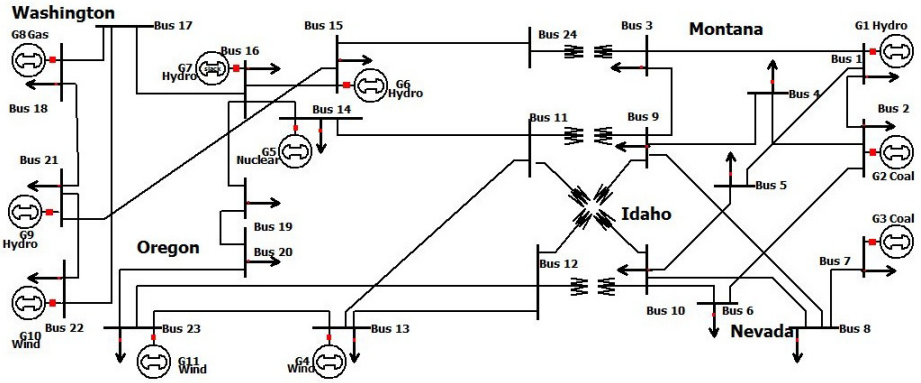

Figure 3.5: A system network diagram with BPA's operation areas. This network has 24 electrical nodes and 11 power generators including hydro, coal, nuclear, natural gas and wind power. The installed generation capacity in the simulation system is configured according to the resources listed in Table 3.3.

Type	Sustained Capacity	Percentage
Hydro	27,142 MW	59.5%
Coal	5,866 MW	12.9%
Combustion turbines	5,526 MW	12.1%
Co-generation	2,938 MW	6.4%
Nuclear	1,150 MW	2.5%
Imports	2,094 MW	4.6%
Non-utility	630 MW	1.4%
Other resources	258 MW	0.6%

Table 3.3: BPA's integrated resources (BPA 2010).

lation system. For example, the wind generation potentials at the three locations ($n = 3$) on 15 August 2003 are presented in Figure 3.6. Wind generation over the maximum generation capability of the wind turbines has to be curtailed for security purposes. The wind component for the reserve requirement is estimated by the MAE of the wind forecast errors. Because over-scheduling (the scheduled wind generation is higher than the actual production capability) requires deployment of additional

No.	Bus	Type	Cap.(MW)	M.C.($/MWh)	RP.(p.u.)
1	1	Hydro	400	6	0.08
2	2	Coal	200	37	0.0081
3	7	Coal	350	35	0.0085
4	13	Wind (GH)	100	3	0.1
5	14	Nuclear	110	21	0.004
6	15	Hydro	700	5	0.074
7	16	Hydro	650	3.7	0.059
8	18	Natural Gas	500	79	0.051
9	21	Hydro	800	3.5	0.081
10	22	Wind (KW)	110	2	0.05
11	23	Wind (VS)	80	1	0.094

Table 3.4: Each generator's configuration including capacity, marginal cost, and ramping rates.

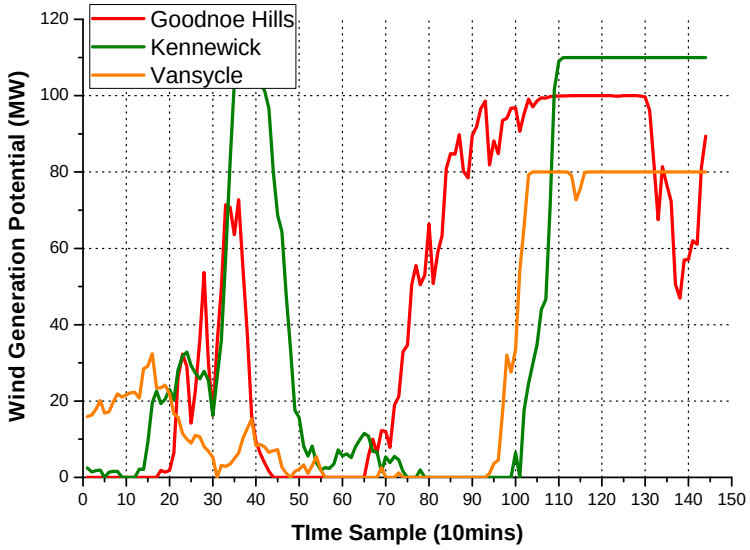

Figure 3.6: The wind generation potential of Vansycle, Kennewick and Goodnoe Hills on 15 August 2003. The horizontal axis indicates different time steps with 10 minutes per interval, and the vertical axis indicates the wind production potential or available wind generation in MW.

reserve capacity, under-forecast errors are not considered in the MAE calculations. A coefficient of 1.2 is multiplied to the MAE value as a reliability margin.

Generator parameters are configured according to Gu and Xie (2010). In the simulation, the minimum output levels of conventional generators, $P_{G_i}^{\min}$, and wind generators, $P_{W_i}^{\min}$, are assumed to be zero. The total installed generation capacity is 4,000 MW. Of this total, the capacity of wind generation is 290 MW, which is about 7.3% (representative of a realistic BPA scenario). Table 3.4 lists the bus number (the number of the electrical node where the generator is located), type (what technology is used), capacity (Cap.: the total power capacity of the generator, in MW), marginal cost (M.C.: the marginal generation cost, which indicates the cost increment due to a power generation increase, in $/MWh), and ramping rate (RP.: the capability of a generator to change its output per minute in normalized per unit value) of each generator.

3.4.4 Analysis of Economic Dispatch Results

In Table 3.5, the performance of the economic dispatch model under different wind forecast models is presented. The wind observation (OB), i.e., the true value, has the lowest system operating cost for all fourteen days. Among different methods, the total operating costs from using PSS are relatively higher over the fourteen days. The AR model, which considers only temporal wind correlations, results in a relatively modest cost-saving performance. The RRSTD model performs better than either PSS or AR models.

For most of the days, the RRSTD has a relatively higher cost savings than the other approaches. Among the fourteen days, 15 August 2003 is selected for a detailed study as reported in the remainder of this section. The operating results for this day are presented in Table 3.6. The row "Energy Market Cost" refers to the generation

Date	1-May	8-May	4-Jun	13-Jun	17-Jul	26-Jul	15-Aug
OB	813,729	783,258	824,637	678,908	832,347	729,972	724,894
PSS	891,771	**895,812**	887,321	874,657	884,629	920,145	892,553
AR	881,738	904,242	891,661	869,866	882,116	908,830	886,935
RRSTD	**870,860**	902,351	**881,143**	**866,084**	**872,777**	**907,764**	**866,633**
Date	28-Aug	1-Sep	15-Sep	3-Oct	31-Oct	17-Nov	20-Nov
OB	822,347	831,509	785,226	834,096	787,220	630,076	694,971
PSS	869,431	879,336	977,584	864,966	906,577	748,126	924,221
AR	878,986	883,951	961,685	**863,474**	**905,447**	**733,469**	914,903
RRSTD	**865,374**	**872,161**	**961,574**	864,207	913,095	751,672	**901,549**

Table 3.5: Economic performance (in $) of wind forecast methods for several days in 2003. The smallest cost is in bold.

	OB	PSS	AR	RRSTD
Total Cost	724,894	892,553	886,935	**866,633**
Energy Market Cost	494,017	519,049	520,625	**517,246**
Regulation Cost	171,619	237,180	237,801	**231,367**
Reserve Cost	59,257	112,748	111,124	**103,982**
Deviation Penalty	0	23,576	17,385	**14,038**
Cost Reduced	18.78%	0.00%	0.63%	**2.90%**

Table 3.6: System operating results (in $) on 15 August 2003. The smallest cost is in bold.

cost from all the generator units in the perspective of the system operator. The rows "Regulation Cost" and "Reserve Cost" give the total costs of providing regulation services and reserve services of all the units. The row "Cost Reduced" refers to the cost savings (in %) from using other forecasting models than the PSS model.

According to the simulation results, the RRSTD model increases the actual wind resource utilization, and reduces the system-wide generation cost, the system's ancillary services (including regulation and reserve services) costs, the wind generation deviation penalty and the total system operating costs.

It can be observed in Table 3.6 that the system-wide operating cost using the

RRSTD model is 2.90% lower than that using the PSS model. One of the advantages of the RRSTD is the reduction in wind generation deviation. As shown in Table 3.6, the wind generation deviation penalty is reduced by almost 60% when the RRSTD model is used compared with the PSS model and by 24% compared with the AR model. The reduction in wind generation deviation is because space-time wind forecasts can increase the wind forecast accuracy (with lower MAE) and reduce the overestimation of available wind generation. In addition, the results of the RRSTD model reveal the advantage in the operating cost of ancillary services. For instance, the total costs of regulation and reserve services are reduced by 2.90% when using the RRSTD model. Given the same wind pressure and system load patterns, the space-time wind forecast model yields higher wind resources utilization and a higher wind generation ratio than the other models. This is because the increased accuracy of the space-time wind forecast model decreases the wind generation that would be wasted by underestimation of available wind generation potentials.

Additionally, a simulation study is carried out to quantify the uncertainty on the 2.90% cost saving on 15 August 2003. For every 10 minutes, 400 realizations are generated from the wind speed predictive distribution for the AR and RRSTD models, respectively, and the median of the realizations is treated as the point forecast. This procedure is carried out 100 times. Hence, 100 wind speed forecasts are simulated for that day for every 10 minutes based on the AR and RRSTD models, respectively. Then the wind speed forecasts are converted to wind power with the 2.5 MW Nordex power curve and put into the power system dispatch model. Figure 3.7 displays the histograms of cost reductions of the system operations with the simulated wind power forecasts from AR and RRSTD, relative to the cost based on the forecasts from PSS. The 95% confidence interval of relative cost savings for using forecasts from the AR instead of those from PSS is [0.61%, 0.66%] and it is [2.84%, 2.95%] for RRSTD.

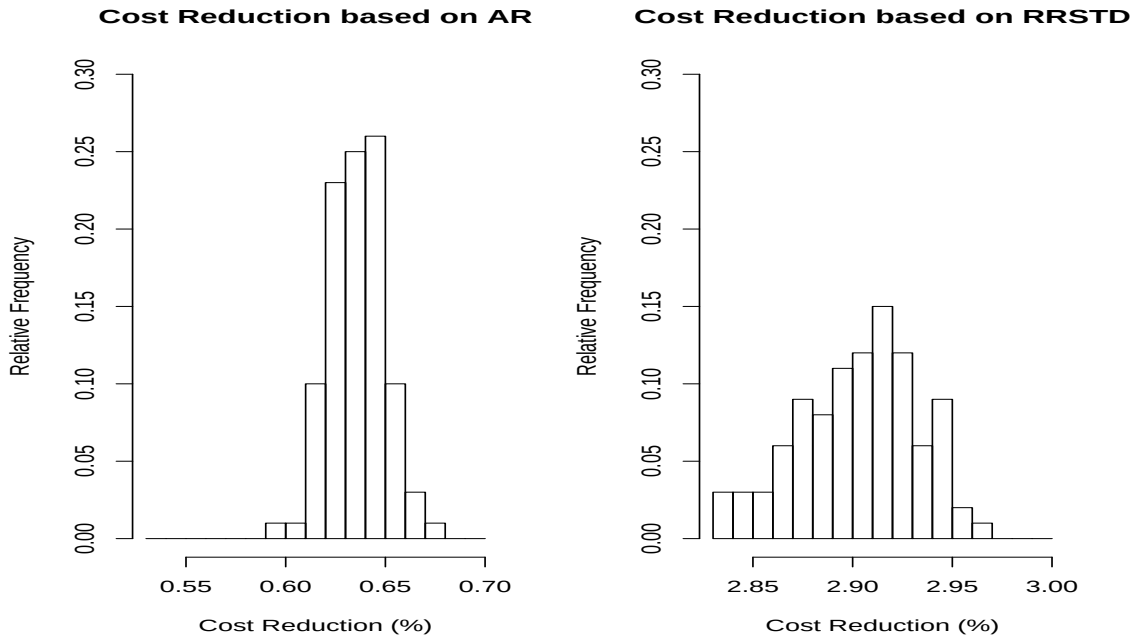

Figure 3.7: Histograms of relative cost savings in percentage based on wind power forecasts from the AR and RRSTD models, compared with the costs based on forecasts from PSS.

The concentrated histograms and narrow confidence intervals demonstrate that the variability of the cost reduction is small. Therefore, it is reliable to conclude that because of the improvement on the wind speed forecasting accuracy by using the RRSTD model, the cost is reduced by around 2.90% instead of using forecasts from PSS, while it is only 0.63% with the AR model, on 15 August 2003.

In Figure 3.8 (top panel), the actual wind generation output at Kennewick is presented. The OB curve depicts the case when there is no wind forecast error, which gives the highest wind generation as well as the best economic dispatch performance. The wind generation profile using the PSS model has the lowest utilization of wind resources, while the wind generation profile of the AR model has the second lowest utilization result. By using the RRSTD model, more wind generation can be inte-

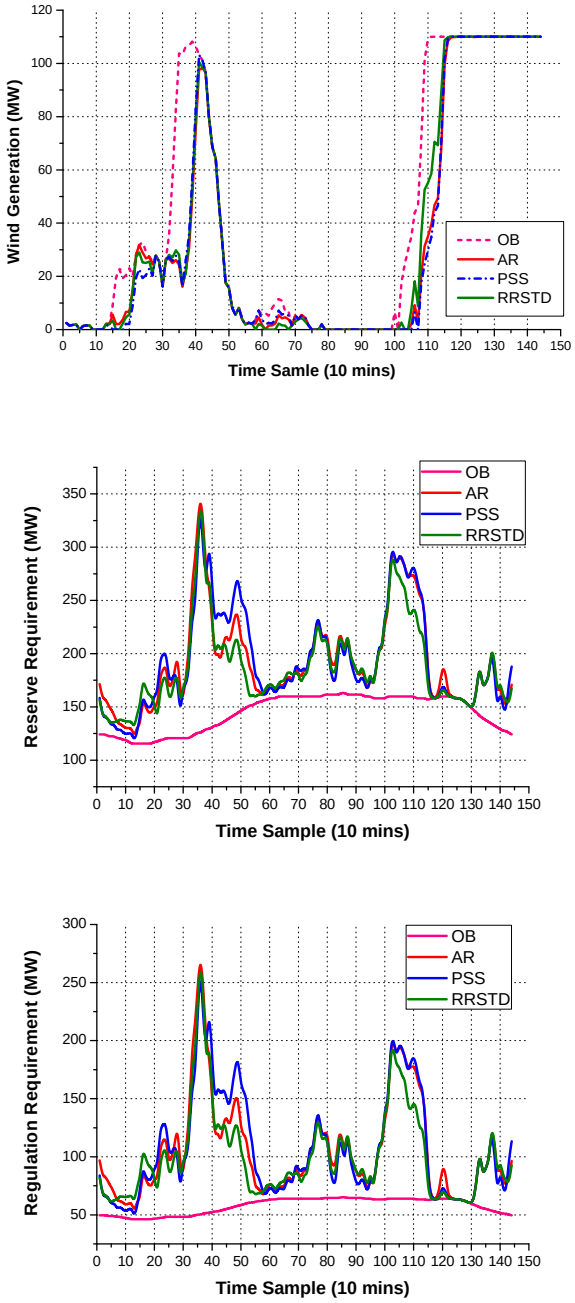

Figure 3.8: Actual wind generation at Kennewick (top panel), total system reserve service requirement (middle panel), and total system regulation service requirement (bottom panel) on 15 August 2003, for different forecasting approaches including OB, PSS, AR, and RRSTD models.

grated into the power system. This is because the underestimation of available wind generation can be avoided during the dispatch process because of the highly accurate wind forecasts by RRSTD and the potential for wind resources not being utilized is reduced.

The system's overall reserve requirement takes into account the uncertainty of both wind generation (mainly forecast errors) and load (demand forecast errors). The selected reserve capacity is used to compensate the energy imbalance within time frameworks of 30 minutes to two hours. In Figure 3.8 (middle panel), the total system reserve requirement for each model is compared. This panel shows that by using the RRSTD model the overall reserve requirement can be reduced due to the improved forecast accuracy.

Regulating energy imbalances in the system keeps the system frequency within a secure range. Unlike reserve services, the capacity for regulation is used to smooth short-term (1 minute to 10 minutes) frequency fluctuations and energy imbalances. Figure 3.8 (bottom panel) shows that the RRSTD model decreases the requirement for regulation capacity and therefore reduces the corresponding regulation cost.

Because of the reduction in both the total power generation cost and the ancillary services costs by using the RRSTD model, the total operating cost is reduced by 2.90% compared with the results from the PSS model. Given that the market for electricity is significant (multi-billions of dollars in regions like Texas and about 240 million for BPA), a 2.90% savings in operating costs means tens of millions of dollars in cost savings due to improved wind forecasts.

3.5 Conclusion

Although wind power is increasingly important to the electricity supply market, the inclusion of wind power is a challenge to power system operations because of the

high variations and limited predictability of wind. Advanced technologies that forecast wind accurately and loss functions that can evaluate forecasts more realistically are needed.

In this chapter, a new space-time model is introduced, the RRSTD model, to solve short-term wind speed forecasting problems. This model generalizes the RSTD model by allowing the forecast regimes to vary with the dominant wind direction and with the season without requiring much prior geographic information. Its forecasts are better than results from the PSS and AR models. It is added that the RRSTD model has the potential to improve forecasts further if more information on monthly wind patterns are available.

Moreover, a new, realistic method is proposed to evaluate forecasts based on power system operating costs through a power system dispatch. To this end, an economic dispatch model is formulated that takes into account the space-time wind forecast information modeled by the RRSTD. Our space-time wind forecasting model reduces the cost of ancillary services, including regulation and reserve costs. These costs were reduced by 2.90% in a realistic illustrative example.

In this chapter, for the RRSTD model, the simplest case of two equally divided regimes based on the local wind patterns is considered. More complex RRSTD models with more than two forecasting regimes or unequally divided regimes could be implemented in more complicated situations. Also, an open challenge for further investigation is to develop an economic dispatch model that makes use of the full spatio-temporal predictive distribution that wind forecasting models can provide.

4. INCORPORATING GEOSTROPHIC WIND INFORMATION FOR IMPROVED SPACE-TIME SHORT-TERM WIND SPEED FORECASTING

4.1 Introduction

Because it is a rich resource that is both green and renewable, wind energy has been developing rapidly worldwide. Although fossil fuels remain the primary energy source for human activities, they are a finite resource that not only threatens the environment through global warming, but also contributes to conflicts between nations. Many countries have aggressive wind energy commitments, although the public is still debating whether wind energy is as green as it is thought to be and whether it is possible to transition so quickly from nonrenewable to renewable energy sources (Busby 2012; Haugen and Musser 2012). For example, the U.S. released a report in 2008 describing a model-based scenario under which wind energy could provide 20% of the electricity demand in the U.S. by 2030 (DOE 2008). The European Union (EU), with its ambitious 20/20/20 target, aims to increase the amount of renewable energy to 20% of the energy supply by 2020 (EU 2008). See the reviews by Genton and Hering (2007) and Zhu and Genton (2012) for more information about wind energy.

One of the biggest challenges in integrating wind power into electricity systems on the large scale comes from the variability of wind. Fluctuations in wind lead to uncertainties in wind power generation, which increases costs and reduces the stability and reliability of a power system. Wind power is not dispatchable, as wind power needed cannot be supplied if the wind is not blowing. In a power system that depends on a high share of wind power, other conventional power resources, such as coal and natural gas, are needed to close the gap between electricity demand

and supply by wind. If the wind is not generating electricity as expected, due to unpredictable variability, the operation of standby and fast power plants is costly. In addition, the higher the share of wind energy in a power system, the more standby power required, thus the higher the costs. An increase in the proportion of wind power in a system requires a corresponding increase of fast but expensive non-wind backup power to balance wind fluctuations. In effect, wind power cannot be simply added into current power systems. Rather, its introduction creates costs and inefficiencies in power systems.

The solution to reducing the uncertainties of wind power generation is accurate wind forecasting. In particular, short-term forecasting up to a few hours ahead is essential. Long-term wind forecasting is less accurate, but high-quality short-term prediction is possible. At the same time, short-term forecasting is closely related to power system dispatch. In a power market, one-day-ahead, hours-ahead, and even minutes-ahead price adjustments are used to determine how much electricity each power plant should generate to meet demand at minimum cost. Precise short-term wind forecasting makes hours-ahead prediction of wind generation possible and allows power systems to utilize wind power more efficiently. Moreover, if there is a gap between the demand and the estimated supply, there is enough time to draw on less expensive back-up power plants. Accurate short-term forecasts reduce the cost for reserves and stabilize the power system.

A number of short-term, statistical, wind forecasting models have been developed (see reviews by Giebel et al. (2011), Kariniotakis et al. (2004), Monteiro et al. (2009) and Zhu and Genton (2012)). Statistical space-time forecasting models that take into account both spatial and temporal correlations in wind have been found to be particularly accurate for short-term forecasting problems. The regime-switching space-time diurnal (RSTD) models, proposed by Gneiting et al. (2006), were found to

outperform persistence, autoregressive, and vector autoregressive models. Since the RSTD models were introduced, researchers have sought to generalize and improve them. For example, Hering and Genton (2010) proposed the trigonometric direction diurnal (TDD) model to generalize the RSTD model by treating wind direction as a circular variable and including it in their model. Zhu et al. (2013b) generalized the RSTD model by allowing forecasting regimes to vary with the prevailing wind and season, obtaining comparable forecasting accuracy. They refer to their model as a rotating RSTD model. Pinson and Madsen (2012) used a first-order Markov chain to determine the regime sequence in offshore wind power forecasting problems and proposed the so-called adaptive Markov-switching autoregressive models.

All of the aforementioned statistical wind forecasting models use only historical wind information – wind speed and direction – to predict future winds. Other atmospheric parameters, such as temperature and pressure, are closely tied to the wind through various physical process and could potentially be included in models to improve prediction accuracy. Directly incorporating temperature and pressure as statistical predictors turns out not to be helpful, however, because winds, for example, are related more closely to horizontal gradients of pressure rather than pressure itself.

In this chapter, a new predictor is introduced to the TDD model, the geostrophic wind (GW), which is the theoretical wind that results from a balance between the atmospheric horizontal pressure gradient force and the Coriolis acceleration, neglecting friction. This new model is named TDDGW. The geostrophic wind not only incorporates air pressure and temperature information, but also has a clear physical interpretation.

The geostrophic wind is a good approximation of the actual horizontal wind in large-scale atmospheric flows outside the tropics and in the absence of friction.

Winds close to the ground are strongly affected by friction with the Earth's surface, and thus are generally not in geostrophic balance; but the surface winds depend to a large degree on winds at higher levels, which generally are nearly in geostrophic balance. In addition, the geostrophic wind can be estimated from surface pressure and temperature data alone, which adds valuable information about winds at higher altitudes that are typically not measured directly by surface anemometers.

Based on numerical experiments using data from West Texas, our TDDGW model achieves higher-quality forecasts than those of the persistence model, the RSTD model and the TDD model. The geostrophic wind direction (D) and the difference in temperature (T) between the current and previous day are also considered, with corresponding models named TDDGWD and TDDGWT, respectively. Additionally, simpler but more efficient methods are proposed to fit the prevailing diurnal wind pattern to obtain better forecasts.

The remainder of this chapter is organized as follows. In Section 4.2, the geostrophic wind is defined and its estimation procedure is introduced in detail. In Section 4.3, the TDDGW model is proposed, along with the TDDGWD and TDDGWT models. Modified diurnal pattern fitting methods are introduced in that section as well. The West Texas data are used as an example in Section 4.4. In Section 4.5, forecast results are evaluated and compared with those from reference models. Finally, Section 4.6 offers final remarks.

4.2 Estimating the Wind in the Free Troposphere

The troposphere is conventionally divided into a lower layer, known as the *planetary boundary layer* (PBL), which interacts strongly with the underlying land or ocean surface through the exchange of heat and momentum, and the overlying *free troposphere*, which is not strongly coupled to the Earth's surface. The depth of the

PBL ranges from a few tens of meters under conditions of calm winds and strong atmospheric stability to several kilometers when the surface is heated intensively by solar radiation. PBL winds often follow a substantial diurnal variation, with weaker winds at night and stronger winds during the day when vertical mixing by convection brings higher momentum air from the free troposphere into the PBL. Wind turbines operate essentially exclusively in the PBL.

Winds in the PBL are generally slower than in the free troposphere due to dissipation by frictional drag with the surface. Because the winds in the free troposphere strongly influence PBL winds, it would be very helpful for wind forecasting purposes to know the winds in the overlying free troposphere. Unfortunately, most wind measurements are made at a standard height of 10 m above the surface, which is obviously not representative of the free troposphere. Various in situ and remote sensing methods are available to measure atmospheric wind profiles, but they can be expensive to install and operate. Given a network of standard surface meteorological observations, however, it is possible to estimate free tropospheric winds using basic dynamical principles (Wallace and Hobbs 2006).

Outside the tropics the large-scale horizontal winds in the free troposphere are given to a good approximation by a balance between the horizontal pressure gradient force and the Coriolis acceleration, which arises from the rotation of the earth. Figure 4.1 illustrates the difference between geostrophic wind (left) and real wind or wind in the PBL (right). The geostrophic wind blows parallel to the isobars (lines of constant pressure). In the PBL friction disrupts the geostrophic balance and the wind tends to blow somewhat across the isobars from higher toward lower pressure. Mathematically this balance can be written as

$$-\frac{1}{\rho}\nabla p = -f\mathbf{k} \times \mathbf{v}_g, \tag{4.1}$$

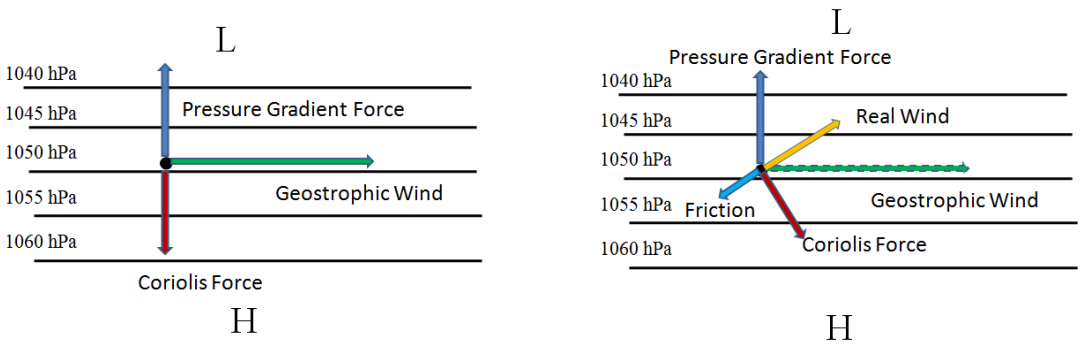

Figure 4.1: The pressure gradient force, Coriolis force, and friction force influence the movement of air parcels near the ground. Geostrophic balance (left) and higher-order balance including friction (right).

where ρ is the atmospheric density, p is the pressure, $f = 2\Omega \sin \phi$ is known as the Coriolis parameter, Ω is the rotation rate of the Earth, ϕ is the latitude, and $\mathbf{k}$ is the local vertical unit vector. Equation (4.1) is known as the geostrophic equation, and the wind velocity that satisfies this equation for a given pressure gradient is known as the geostrophic wind, $\mathbf{v}_g$, which can be decomposed into eastward and northward components, u_g and v_g, respectively. It is convenient to use the hydrostatic equation, which expresses the balance between the vertical pressure gradient force and gravity, to rewrite the pressure gradient force in terms of the horizontal gradient of the geopotential height Z on surfaces of constant pressure

$$-g_0 \nabla Z = -f\mathbf{k} \times \mathbf{v}_g, \tag{4.2}$$

where g_0 is the acceleration of gravity. This substitution eliminates density from the equation.

Solving for the geostrophic wind, $\mathbf{v}_g$, in Equation (4.2) and writing as eastward

and northward components gives

$$\begin{aligned} u_g &= -\frac{g_0}{f}\frac{\partial Z}{\partial y} \quad \text{and} \\ v_g &= \frac{g_0}{f}\frac{\partial Z}{\partial x}, \end{aligned} \tag{4.3}$$

where x and y are local eastward and northward Cartesian coordinates. The actual wind in the free troposphere, $\mathbf{v}$, is given to a good approximation by the geostrophic wind, $\mathbf{v}_g$, which can be computed from the horizontal gradient of the geopotential height, or, equivalently, from the horizontal pressure gradient, which is routinely measured with good precision by barometers at surface meteorological observing stations.

To compute the geostrophic wind components from a network of surface pressure observing stations, the following steps are necessary. First, because the barometers at different stations are typically located at different elevations above sea level, it is necessary to correct the pressure measurements to a standard reference elevation. This can be done with good accuracy through the hydrostatic equation, which in integral form is written as

$$Z = Z_i + \frac{R\bar{T}}{g_0} \ln\left(\frac{p_i}{p_{ref}}\right), \tag{4.4}$$

where Z_i is the geopotential height of barometer i, p_i is the pressure measurement by barometer i, p_{ref} is the desired reference pressure level (e.g., 850 hPa), Z is the geopotential of the reference pressure level, R is the gas constant for air (287 J K^{-1} kg^{-1}), and $\bar{T}$ is the layer-averaged temperature between p_i and p, which can be estimated using surface temperature measurements.

Second, horizontal pressure gradients in the atmosphere are very small. It is

thus important to remove any systematic biases in the pressure measurements at different locations. This can be done to a reasonably good approximation by removing the time-mean pressure and considering only the deviations from the time mean. Although the time-mean geostrophic wind is not zero, it is the variations in the geostrophic wind that are of the greatest importance for forecasting purposes.

Finally, to smooth out small-scale spatial fluctuations in the pressure measurements among the stations in a network, at each time a smooth (planar) surface is fitted to the observed geopotential heights in the form

$$Z(x, y) = a_0 + a_1 x + a_2 y. \tag{4.5}$$

From this, we get

$$\begin{aligned} \frac{\partial Z}{\partial x} &= a_1 \quad \text{and} \\ \frac{\partial Z}{\partial y} &= a_2, \end{aligned} \tag{4.6}$$

which can be substituted into Equation (4.3) to give

$$\begin{aligned} u_g &= -\frac{g_0}{f} a_2 \quad \text{and} \\ v_g &= \frac{g_0}{f} a_1. \end{aligned} \tag{4.7}$$

The geostrophic wind speed and direction are given by $|\mathbf{v}_g| = \sqrt{u_g^2 + v_g^2}$ and $\theta_g = \tan^{-1}(v_g/u_g)$, respectively. For the sake of simplicity, gw and $g\theta$ are used to denote $|\mathbf{v}_g|$ and θ_g, respectively.

4.3 The Trigonometric Direction Diurnal Model with Geostrophic Wind

4.3.1 The TDD Model

The TDD model is an advanced space-time model for short-term wind speed forecasting problems. It generalizes the RSTD model by treating wind direction as a circular variable and including it in the model, such that the non-unique and locally dependent forecasting regimes are eliminated. An detailed introduction about the TDD model is given in Section 2.4.3. In this section, it is introduced briefly and represented slightly differently in the part of center parameter model in order to develop our new model.

For the sake of simplicity, let $y_{s,t}$ and $\theta_{s,t}$, $s = 1, \ldots, S$ and $t = 1, \ldots, T$, be surface wind speed and direction measurements at station s at time t, respectively. The objective is to predict the k-step-ahead wind speed, $y_{i,t+k}$, at one of the stations, $i \in \{1, \ldots, S\}$. For short-term wind speed forecasting problems, the k-step-ahead is from 1 to 6 hours.

In the TDD model, it is assumed that $y_{s,t+k}$ follows a truncated normal distribution, $N^{+}(\mu_{s,t+k}, \sigma_{s,t+k})$, with $\mu_{s,t+k}$ and $\sigma_{s,t+k}$ as the center parameter and the scale parameter, respectively, considering that the density of the wind speed is non-negative and right skewed. If these two parameters are modeled appropriately, accurate probabilistic forecasts can be achieved beyond point forecasts. In the TDD model, these two parameters are modeled as follows, taking $s = 1$ as an example:

a) The center parameter, $\mu_{1,t+k}$, is modeled in two parts:

$$\mu_{1,t+k} = D_{1,t+k} + \mu^{r}_{1,t+k}.$$

The first part, $D_{1,t+k}$, is the diurnal component in the wind speed, which is

fitted by two pairs of trigonometric functions:

$$D_{1,h} = d_0 + d_1 \sin\left(\frac{2\pi h}{24}\right) + d_2 \cos\left(\frac{2\pi h}{24}\right) + d_3 \sin\left(\frac{4\pi h}{24}\right) + d_4 \cos\left(\frac{4\pi h}{24}\right), \tag{4.8}$$

where $h = 1, 2, \ldots, 24$.

Then, the second part, the residual of the wind speed after removing the diurnal component, is modeled as

$$\mu^r_{1,t+k} = \alpha_0 + \sum_{\substack{s=1,\ldots,S \\ j=0,1,\ldots,q_s \\ j'=0,1,\ldots,q'_s}} \alpha_{s,j} y^r_{s,t-j} + \left\{\beta_{s,j'} \cos(\theta^r_{s,t-j'}) + \gamma_{s,j'} \sin(\theta^r_{s,t-j'})\right\}. \tag{4.9}$$

Equation (4.9) models the k-step-ahead wind speed residual as a linear combination of current and past wind speed residuals at all stations up to time lag q_s depending on station s, as well as a pair of trigonometric functions of wind direction residuals whose diurnal patterns are also fitted by the model in (4.8) up to time lag q'_s, which is not necessarily equivalent to q_s. Both q_s and q'_s are determined by the modified Bayesian information criterion (BIC) as described by Hering and Genton (2010).

b) The scale parameter is modeled by a simple linear model of volatility value, v^r_t, in the following form,

$$\sigma^2_{1,t+k} = b_0 + b_1 v^r_t, \tag{4.10}$$

where $v^r_t = \left\{\frac{1}{2S}\sum_{s=1}^{S}\sum_{l=0}^{1}(y^r_{s,t-l} - y^r_{s,t-l-1})^2\right\}^{1/2}$ and b_0, $b_1 > 0$.

The coefficients in the center parameter and scale parameter models are estimated

numerically by minimizing the continuous ranked probability score (CRPS) for a truncated normal distribution, based on a 45-day training period (see Gneiting et al. (2006) and Gneiting and Raftery (2007)).

4.3.2 The TDDGW Model

Based on the discussion of the geostrophic wind in Section 4.2, it is clear that atmospheric pressure and temperature play important roles in wind speed and direction. To reduce uncertainties about winds, an efficient forecasting model should include this critical information. However, the experiments in the next section show that incorporating air pressure and temperature directly into the TDD model does not reduce errors in forecasts. This is because in the TDD model, particularly in the mean structure in Equation (4.9), linearity is assumed between future wind speeds and the covariates. This assumption is invalid when it comes to air pressure and temperature. As a result, no improvement is achieved by incorporating these variables directly into the TDD model.

Instead of seeking nonlinear forms between wind speeds and air pressure and temperature in the mean structure in the TDD model, it is proposed to use the geostrophic wind as a predictor, as this better expresses the physical relationship between temperature, pressure, and wind. In the new TDDGW model, geostrophic wind is incorporated into the TDD model, retaining much of its structure to take account of the spatio-temporal correlations in wind. Specifically, the TDD model is modified by adding geostrophic wind into the center parameter model in Equation (4.9):

$$\mu^r_{1,t+k} = \alpha_0 + \sum_{\substack{s=1,\ldots,S \\ j=0,1,\ldots,q_s \\ j'=0,1,\ldots,q'_s}} \alpha_{s,j} y^r_{s,t-j} + \left\{\beta_{s,j'} \cos(\theta^r_{s,t-j'}) + \gamma_{s,j'} \sin(\theta^r_{s,t-j'})\right\}$$

$$+ c_0 gw^r_{1,t} + c_1 gw^r_{1,t-1} + c_2 gw^r_{1,t-2} + \cdots + c_q gw^r_{1,t-q}, \tag{4.11}$$

where q is the time lag of geostrophic wind depending on the station, s, and is also determined by the aforementioned modified BIC method. Since geostrophic wind is the theoretical wind above the friction layer in the atmosphere, its value for a small area is almost constant. This is why the geostrophic wind is used as a common predictor in Equation (4.11).

In addition to including geostrophic wind in the TDD model, the geostrophic wind direction and the temperature difference between the current and previous day are also considered, because, from the point of view of the atmosphere, these variables are closely related to surface wind. These two modified TDDGW models are named TDDGWD and TDDGWT, and with the two variables simultaneously, TDDGWDT.

Additionally, the diurnal pattern fitting is also modified. Instead of the daily wind pattern in the model in Equation (4.8), the average wind speed of each hour within a certain period is treated as the diurnal pattern. Depending on the period used, there are several versions of the diurnal pattern modeling: MD, a diurnal pattern that takes into account winds in a 45-day sliding window; SMD, a diurnal pattern that is calculated for each season; and YMD, a diurnal pattern based on a whole year's data (or several years' data).

4.3.3 Reference Models

Several models are introduced briefly here as references, including the persistence (PSS), RSTD, and Lasso-based (Tibshirani 1996) modified TDD models:

a) PSS assumes the future wind speed is the same as the current wind speed, $\hat{y}_{s,t+k} = y_{s,t}$.

b) As mentioned above, in the RSTD model, forecasting regimes are defined based on the prevailing wind direction, and for each regime, a separate model is fitted only with historical wind speeds as predictors in Equation (4.9) plus speeds from neighboring stations. For more detailed introduction about the RSTD model, please refer Section 2.4.2.

c) The Lasso-based modified TDD directly includes historical pressure and temperature measurements in the center parameter, Equation (4.9). Rather than using the modified BIC and the CRPS, the Lasso method is applied to select significant predictors and estimate their coefficients simultaneously. Its tuning parameter is chosen by 10-fold cross-validation.

4.4 West Texas Data

4.4.1 Data Description

The wind data used here were collected from mesonet towers at a height of 10 m above the surface in West Texas and Eastern New Mexico (Figure 4.2). The original data archive contains five-minute means of three-second measurements of wind and other atmospheric parameters from more than 60 stations. In the experiment, hourly-averaged data of five-minute means from 1 Jan 2008 to 31 Dec 2010 are used. Although the available data were collected from a height of 10 m, and most wind

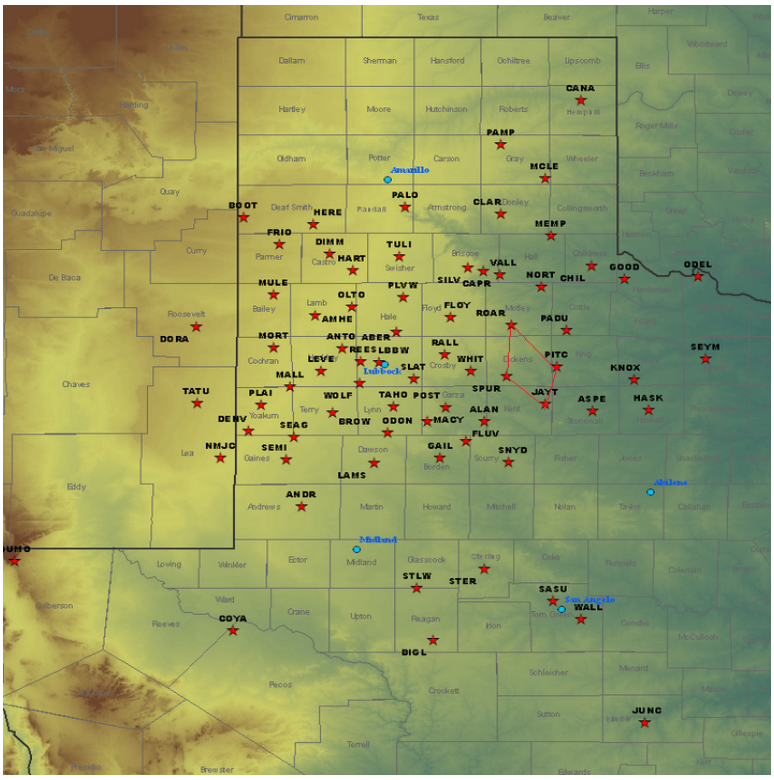

Figure 4.2: The distribution of Mesonet Towers (stars) in West Texas and Eastern New Mexico.

turbine towers today are at least 60 m tall (Busby 2012), winds within the lowest part of the PBL tend to be closely correlated through turbulent mixing of momentum. As a result, 10 m wind data contains significant information about winds at the turbine height.

Texas has a long history of using wind power and it produces the most wind power of any U.S. state, with a total installed capacity of 10,400 MW from over 40 projects established by the end of 2012, providing 9.2% of the electricity consumed in the state during 2012 (Reuters 2013). With a very rich wind resource and broad geographic scope, the potential for wind power in Texas is considerable. West Texas in particular, the area shown in Figure 4.2, has a rich wind resource and many wind

ID	Lat (degree)	Lon (degree)	Elevation (feet)
ASPE	33.16789	-100.19602	1740
FLOY	34.00158	-101.32588	3179
JAYT	**33.23241**	**-100.56778**	**2010**
PADU	33.89053	-100.39886	2021
PITC	**33.56703**	**-100.48061**	**1998**
POST	33.20033	-101.36804	2598
ROAR	**33.93635**	**-100.84540**	**2615**
SLAT	33.45690	-101.61723	3065
SNYD	32.71614	-100.86167	2431
SPUR	**33.48085**	**-100.87636**	**2287**
VALL	34.37896	-100.93175	2450
WHIT	33.52533	-101.16506	2704

Table 4.1: Information on 12 sites in West Texas. The four stations in boldface type provide the data used in the forecast experiments.

farms have been built there. Wind speed forecasting is of great interest in this area.

In our numerical experiment, a small area is chosen with four stations to test the newly proposed model. This area includes PITC, JAYT, SPUR, and ROAR stations in and around Dickens county. These four locations are marked with red lines in Figure 4.2. Our goal was to predict 2-hour-ahead wind speeds at these four locations. The recorded data include wind speed, wind direction, temperature, and pressure. To estimate the geostrophic wind in the TDDGW model, 12 surface stations were selected that surround the four test stations, as listed in Table 4.1. More information is given at http://www.mesonet.ttu.edu/wind.html.

4.4.2 Data Exploration

The area where the four target stations are located in West Texas has both northerly and southerly prevailing winds as shown by the wind roses in Figure 4.3, which were created based on the training data set. In a wind rose, each petal indicates the frequency of winds blowing from a particular direction and the color bands in

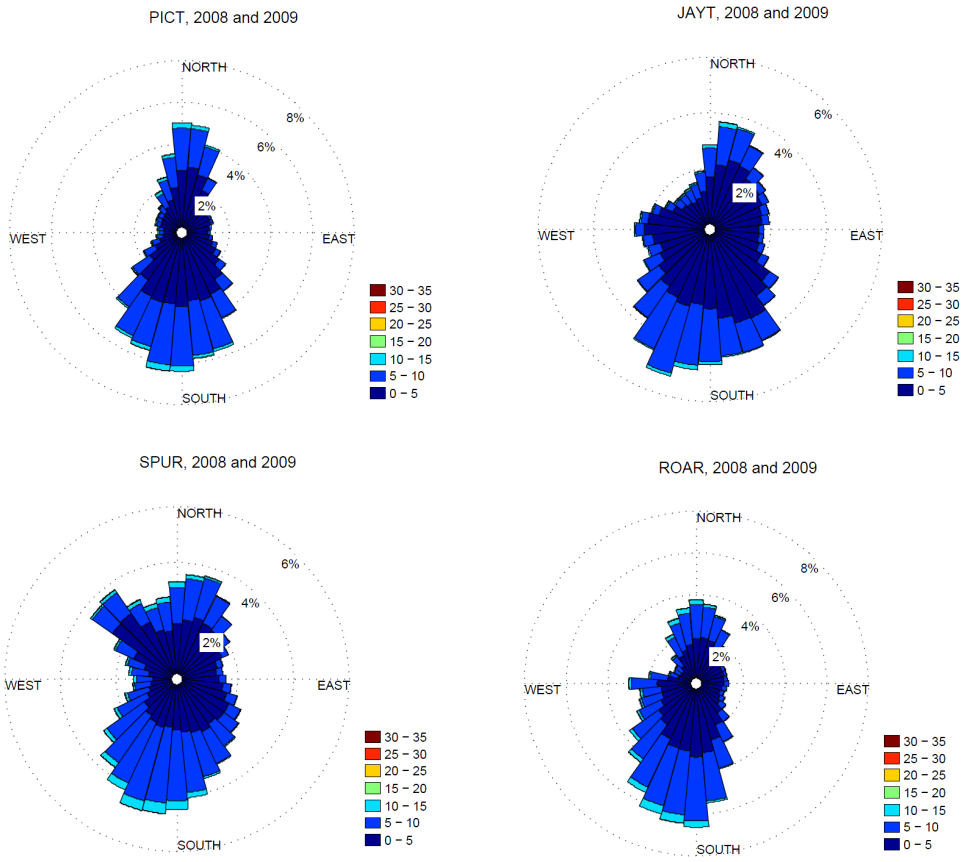

Figure 4.3: Wind roses of wind speeds in 2008-2009 at PICT, JAYT, SPUR, and ROAR.

each petal show the ranges of the wind speed. As the wind roses show, the wind patterns in this area are quite significant. High frequencies and large speed ranges are found in the north and south directions at all four stations. More specifically, the northerly wind dominates this area, with higher wind frequencies blowing from the south than from the north. At station JAYT, the wind direction switches slightly to the northeast. Different from the other three locations, station SPUR has a high frequency in the southeast direction.

The wind speed density plots at the four stations are displayed in Figure 4.4

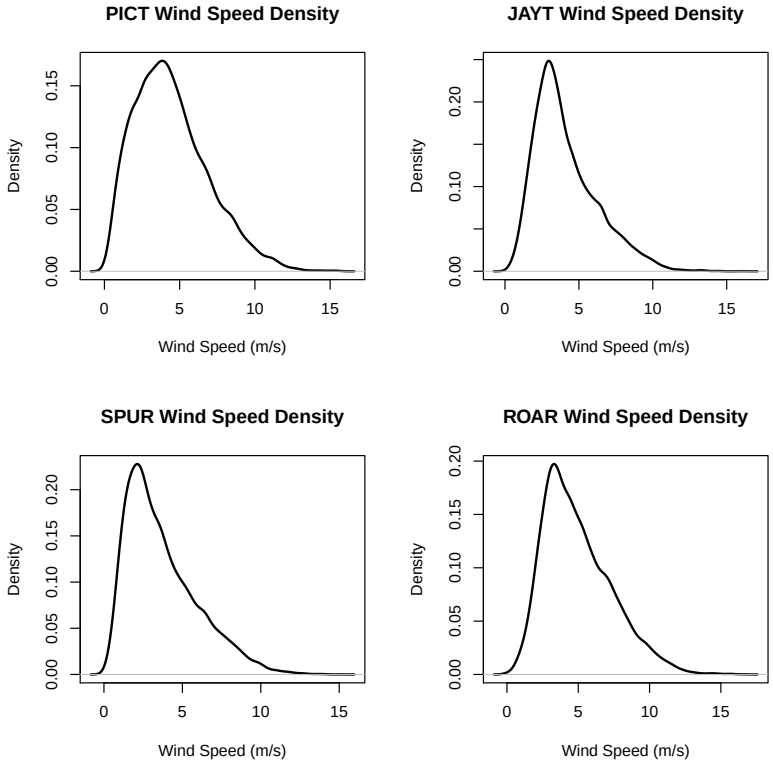

Figure 4.4: Density plots of wind speeds at PICT, JAYT, SPUR, and ROAR in 2008-2009.

based on the wind data from 2008 and 2009. As expected, the densities are skewed right, which means that the assumption in Section 4.3 that the wind speed follows a truncated normal distribution is reasonable.

4.4.3 Geostrophic Wind and Surface Wind

To estimate the geostrophic wind based on surface measurements of air pressure and temperature, the aforementioned two steps in Section 4.2 are carried out. First, for each hour, surface pressure measurements are represented by the geopotential height with Equation (4.4). For the value of $\bar{T}$, the average temperature from the 12 stations listed in Table 4.1 is used and 850 hPa for the reference pressure, p_{ref}.

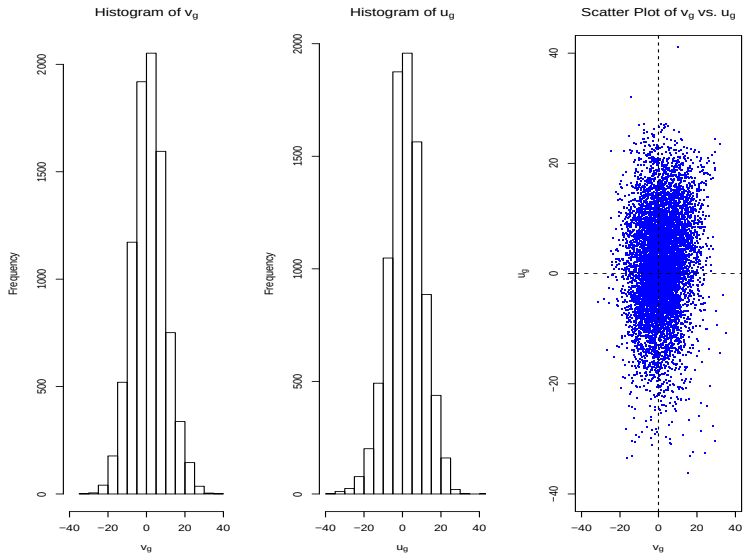

Figure 4.5: Histograms and scatter plots of approximation of geostrophic wind components, u_g and v_g.

Secondly, using the 12 stations' geopotential height data, along with their latitude and longitude data as given in Table 4.1, a geopotential height plane is fitted for each hour, resulting in a geopotential height gradient based on the coefficients of the plane of the x and y horizontal components as shown in Equation (4.7). The monthly average geopotential height is removed before fitting the plane. With these two steps, each hourly surface wind record has a corresponding geostrophic wind estimated from the temperature and pressure information.

In more detail, Figure 4.5 presents histograms of v_g, u_g (in Equation (4.7)), and scatter plots of v_g vs. u_g. Though the prevailing surface wind is southerly as noted, there is no significant pattern found in the geostrophic wind, where positive values of u_g and v_g indicate westerly and southerly wind, respectively.

The four days' hourly geostrophic wind speeds (in black) and surface winds (in red) in 2008 at PICT on the top of Figure 4.6 indicate that the former has larger

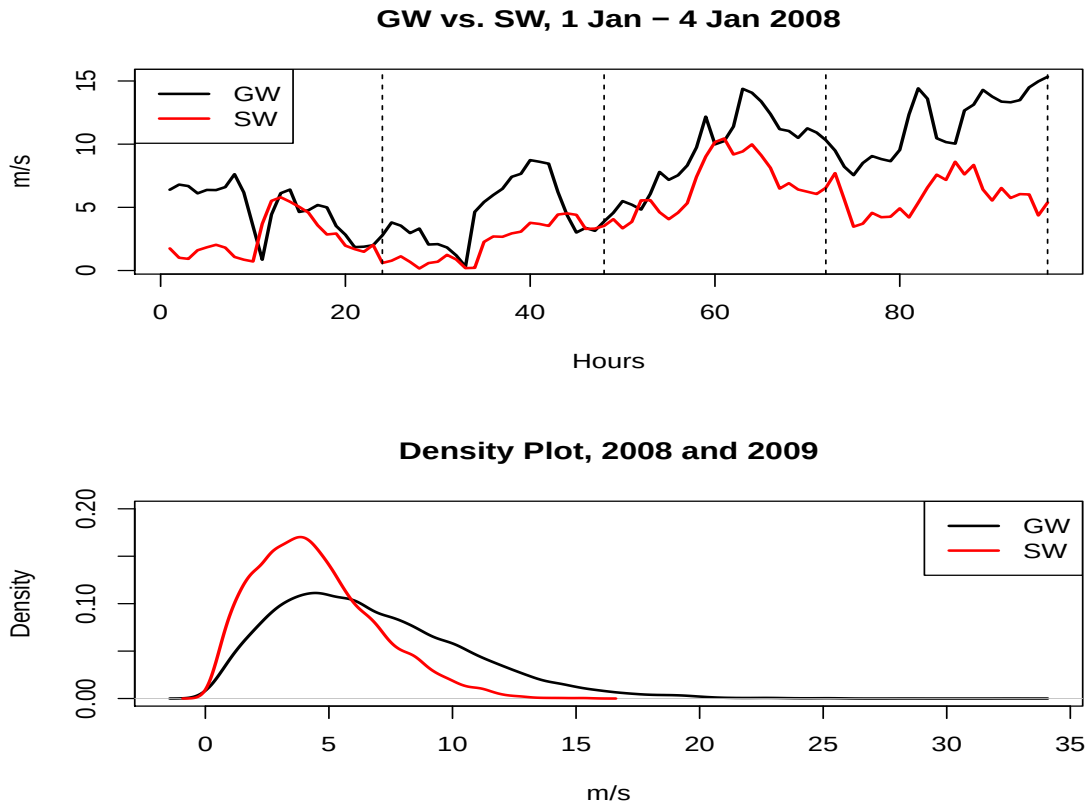

Figure 4.6: Geostrophic wind (GW) vs. surface wind (SW) (top) and density plots of the geostrophic wind and surface wind (bottom).

values than the latter, while the latter has a more wiggly shape. Since the effects of friction forces, which slow down the wind speed and change direction, are ignored in the geostrophic balance, the geostrophic winds blow more strongly and smoothly than do the surface winds. Also, it can be seen in Figure 4.6 that they share similar patterns, which is consistent with the large positive correlation coefficient between the surface wind and the geostrophic wind as listed in Table 4.2. The bottom plot displays the density estimations of the geostrophic wind speed (in black) and the surface wind speed (in red), from which we can see that both densities are skewed right and again the geostrophic wind speed has a larger range than does the surface

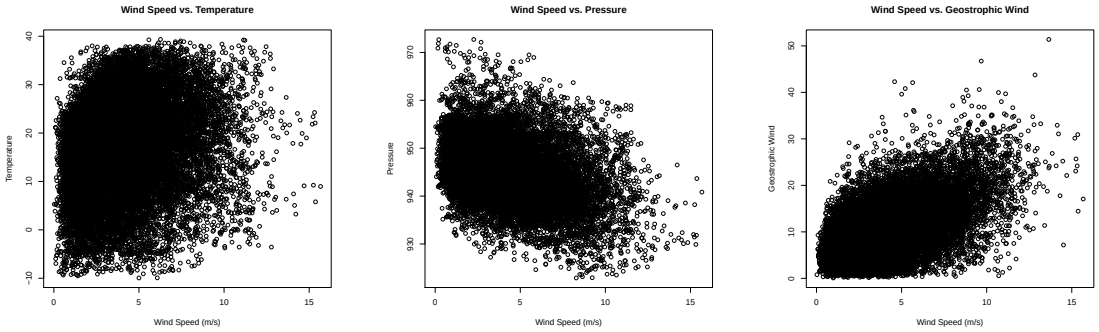

Figure 4.7: Scatter plots of wind speed vs. temperature (left), pressure (middle), and geostrophic wind speed (right).

wind speed.

Figure 4.7 displays scatter plots of wind speed vs. surface temperature (left), pressure (middle) and geostrophic wind speed (right) based on the training data at PICT. From the first plot, we can see that the surface wind speed is positively correlated with temperature but not linearly. This is consistent with the fact that the correlation coefficient between them is 0.19. This means that although it is true that wind speed is closely related to temperature, the linearity assumption in the mean structure of the TDD model in Equation (4.9) is not appropriate. The correlation coefficient of the surface wind speed and pressure is -0.34, indicating a negative weakly linear trend in the scatter plot as well. However, geostrophic wind is more strongly linearly correlated to surface wind, with correlation coefficient equal to 0.53, which is easy to see in the scatter plot. This shows that geostrophic wind not only contains important temperature and pressure information, it also meets the linearity assumption such that it can be integrated into the TDD model. More importantly and interestingly, geostrophic wind has physical interpretability.

Figure 4.8 shows the averaged diurnal pattern of the surface wind speed and geostrophic wind in different seasons as well as for the whole of 2008 (left) and 2009

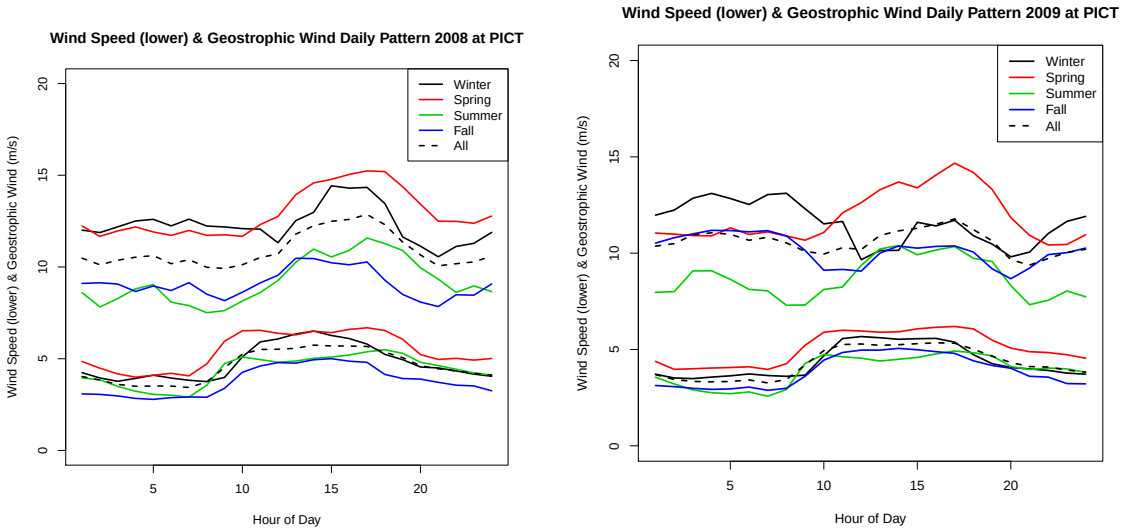

Figure 4.8: Daily pattern of wind speed (lower part in each plot) and geostrophic wind speed (upper part in each plot) in different seasons of 2008 and 2009.

(right) at PICT. The plots show that geostrophic wind has higher speed than surface wind, which is slowed down by the ground friction. Through the hours of the day, the curve of the geostrophic wind is more wiggly and its values are in a scale of 10 to 15 m/s, while the curve of the surface wind is smoother with a range from 2 to 3 m/s. Seasonally, geostrophic wind and surface wind are consistent, having higher speed during winter (Dec to Feb, solid black lines) and spring (Mar to May, solid red lines) than summer (Jun to Aug, solid green lines) and fall (Sep to Nov, solid blue lines).

4.5 Numerical Results

4.5.1 Training Results

In the training procedure, the models for the center parameter are obtained based on the training data set to forecast 2-hour-ahead wind speed at each of the four stations. For example, to predict $y_{P,t+2}$, the 2-hour-ahead wind speed at PICT, the

Variable	t	$t-1$	$t-2$	$t-3$	$t-4$	$t-5$
y_P	0.80	0.70	0.62	0.54	0.47	0.40
gw_P	0.57	0.55	0.53	0.50	0.47	0.43
$\cos(\theta_P)$	-0.06	-0.08	-0.11	-0.13	-0.15	-0.17
$\sin(\theta_P)$	-0.14	-0.16	-0.17	-0.19	-0.20	-0.20
$\cos(g\theta_P)$	0.10	0.09	0.09	0.09	0.10	0.10
$\sin(g\theta_P)$	0.17	0.15	0.13	0.11	0.09	0.07
y_J	0.74	0.66	0.58	0.51	0.45	0.39
$\cos(\theta_J)$	-0.12	-0.14	-0.15	-0.16	-0.17	-0.18
$\sin(\theta_J)$	-0.14	-0.17	-0.19	-0.21	-0.22	-0.23
y_S	0.73	0.64	0.55	0.48	0.40	0.33
$\cos(\theta_S)$	-0.19	-0.20	-0.20	-0.20	-0.20	-0.20
$\sin(\theta_S)$	-0.06	-0.09	-0.11	-0.14	-0.16	-0.18
y_R	0.76	0.70	0.64	0.59	0.53	0.48
$\cos(\theta_R)$	-0.03	-0.04	-0.06	-0.08	-0.10	-0.12
$\sin(\theta_R)$	-0.05	-0.08	-0.11	-0.13	-0.16	-0.17

Table 4.2: Correlation coefficients between $y_{P,t+2}$ and the current and up to 5-step lag surface wind speed (y), direction (θ), geostrophic wind speed (gw) and geostrophic wind direction ($g\theta$) at four stations (P, J, S, and R).

variables listed in Table 4.2, except geostrophic wind direction, are put into the selection pool, and the aforementioned BIC is applied to select significant predictors. The variables in the selection pool include current and up to 10-step lags of wind speed, geostrophic wind speed, and pairs of cosine and sine of the wind direction at all four stations. In the TDDGWD model, the cosine and sine of the geostrophic wind direction are also considered. Different from the cosine and sine of the surface wind direction, which have negative correlations with the 2-hour-ahead wind speed at PICT, the cosine and sine of the geostrophic direction are positively correlated with the 2-hour-ahead wind speed at PICT (see Table 4.2). In the table, the indexes, P, J, S, and R, indicate the four locations.

Site	Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Overall
PICT	TDDGW	0.95	0.81	1.02	0.93	0.91	0.96	0.91	0.91	0.84	0.82	0.97	0.98	0.92
PICT	TDDGW-MD	**0.94**	**0.80**	**0.96**	0.89	**0.92**	**0.91**	**0.83**	**0.86**	0.81	**0.77**	**0.92**	**0.94**	**0.88**
PICT	TDDGW-SMD	**0.94**	0.84	0.98	**0.88**	0.93	0.93	0.85	**0.86**	0.82	0.78	0.94	0.96	0.89
PICT	TDDGW-YMD	0.98	0.81	0.98	0.91	0.95	0.96	0.86	0.88	**0.80**	0.81	0.96	0.98	0.90

Table 4.3: MAE values of 2-hour-ahead forecasts from TDDGW with different diurnal component fitting methods at PICT, 2010. The smallest MAE value of each column is in boldface.

4.5.2 Evaluation of Forecasts

The trained TDDGW, TDDGWT, TDDGWD, and TDDGWDT models are applied to the testing data set with modified diurnal modeling, MD, SMD, and YMD, to predict probabilistically 2-hour-ahead wind speeds at the four stations. Prediction mean absolute errors (MAE) are used to evaluate the performance of the forecasts, which are defined as $\sum_{t=1}^{T} |y_{P,t+2} - \hat{y}_{P,t+2}|$, at station PICT, for example. For $\hat{y}_{P,t+2}$, the median of the predictive truncated normal distribution is used: $\hat{y}_{P,t+2} = \mu_{P,t+2} + \sigma_{P,t+2} \cdot \Phi^{-1}\{0.5 + 0.5 \cdot \Phi(-\mu_{P,t+2}/\sigma_{P,t+2})\}$; see **?** for a discussion of quantiles as optimal point forecasts. A 45-day-sliding window is used to estimate the coefficients in the models with the CRPS method. Forecasts are compared with the reference models listed in Section 4.3.3 in addition to the TDD model.

In Table 4.3, the prediction MAE values of forecasts from the TDDGW model with aforementioned different diurnal modeling methods are listed. Overall the MD method has the smallest MAE values among the four, 0.88 compared with 0.92, 0.89 and 0.90, from TDDGW, TDDGW-SMD, and TDDGW-YMD methods, respectively. The TDDGW-MD model has the smallest MAE value, 10 out of the 12 months, followed by TDDGW-SMD, 3 out of 12 months. This is because in Equation (4.8), the daily pattern of the wind speed is fitted by a continuous smooth function of the time of a day. In this way, any time of the daily wind speed can be obtained

Site	Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Overall
PICT	PSS	1.06	0.87	1.21	1.15	1.15	1.13	1.03	1.05	0.96	0.97	1.17	1.14	1.08
PICT	RSTD	0.93	**0.79**	1.07	0.98	1.02	0.95	0.89	0.90	0.83	0.82	0.99	1.01	0.94
PICT	TDD	0.95	0.81	1.07	0.99	1.00	0.97	0.89	0.93	0.84	0.86	1.01	1.03	0.95
PICT	LASSO-MD	1.04	0.87	1.12	1.05	1.11	1.00	0.92	0.94	0.90	0.88	1.04	1.06	1.00
PICT	TDDGW-MD	0.94	0.80	**0.96**	**0.89**	**0.92**	0.91	**0.83**	**0.86**	0.81	**0.77**	**0.92**	**0.94**	**0.88**
PICT	TDDGWT-MD	0.94	0.82	**0.96**	0.90	**0.92**	0.91	**0.83**	**0.86**	0.81	**0.77**	**0.92**	0.95	**0.88**
PICT	TDDGWD-MD	**0.91**	0.82	0.97	**0.89**	**0.92**	**0.90**	0.84	**0.86**	**0.80**	0.78	**0.92**	**0.94**	**0.88**
PICT	TDDGWDT-MD	**0.91**	0.83	0.97	0.90	**0.92**	**0.90**	0.84	**0.86**	**0.80**	0.78	**0.92**	**0.94**	**0.88**
JAYT	PSS	0.96	0.87	1.18	1.11	1.03	0.96	0.88	0.83	0.83	0.8	1.12	1.04	0.97
JAYT	RSTD	0.87	0.78	0.98	0.94	0.93	0.78	0.76	0.72	0.74	0.67	0.90	0.88	0.83
JAYT	TDD	0.86	0.77	0.97	0.94	0.91	0.79	0.72	0.73	0.73	0.66	0.90	0.87	0.82
JAYT	LASSO-MD	0.89	0.85	1.07	1.03	0.96	0.84	0.77	0.74	0.76	0.7	0.97	0.90	0.87
JAYT	TDDGW-MD	**0.83**	**0.73**	**0.91**	**0.87**	**0.83**	0.74	0.66	**0.68**	**0.68**	**0.62**	**0.84**	0.82	**0.77**
JAYT	TDDGWT-MD	**0.83**	**0.73**	0.92	**0.87**	**0.83**	0.74	**0.65**	**0.68**	**0.68**	**0.62**	**0.84**	**0.81**	**0.77**
JAYT	TDDGWD-MD	**0.83**	**0.73**	0.92	**0.87**	**0.83**	0.74	0.66	**0.68**	**0.68**	**0.62**	0.85	0.82	**0.77**
JAYT	TDDGWDT-MD	**0.83**	**0.73**	0.92	**0.87**	**0.83**	**0.73**	**0.65**	**0.68**	**0.68**	**0.62**	**0.84**	**0.81**	**0.77**
SPUR	PSS	0.98	0.89	1.20	1.13	1.09	1.08	0.87	0.90	0.89	0.88	1.12	1.07	1.01
SPUR	RSTD	**0.84**	0.76	1.05	0.97	0.92	0.89	0.71	0.75	0.75	0.70	0.89	0.92	0.85
SPUR	TDD	0.86	0.76	1.05	0.97	0.94	0.89	0.70	0.77	0.74	0.71	0.90	0.92	0.85
SPUR	LASSO-MD	0.94	0.85	1.06	1.02	1.01	0.90	0.76	0.79	0.76	0.73	0.92	0.93	0.89
SPUR	TDDGW-MD	**0.84**	0.74	**0.95**	**0.88**	**0.85**	**0.78**	**0.64**	0.71	**0.68**	0.66	**0.84**	**0.84**	0.79
SPUR	TDDGWT-MD	**0.84**	0.74	**0.95**	**0.88**	0.86	**0.78**	**0.64**	0.71	**0.68**	0.66	**0.84**	0.85	0.79
SPUR	TDDGWD-MD	**0.84**	**0.73**	**0.95**	0.87	**0.85**	**0.78**	**0.64**	**0.70**	0.67	**0.65**	**0.84**	**0.84**	**0.78**
SPUR	TDDGWDT-MD	**0.84**	**0.73**	**0.95**	0.87	**0.85**	**0.78**	**0.64**	0.71	0.67	**0.65**	**0.84**	0.85	**0.78**
ROAR	PSS	1.03	0.91	1.21	1.22	1.14	1.09	0.91	0.87	1.00	0.89	1.18	1.11	1.05
ROAR	RSTD	1.02	**0.86**	1.10	1.14	1.05	0.97	0.79	0.81	0.91	0.82	1.05	1.04	0.97
ROAR	TDD	**1.01**	**0.86**	1.08	1.09	1.05	0.97	0.78	0.79	0.91	**0.81**	1.04	1.01	0.95
ROAR	LASSO-MD	1.08	0.87	1.09	1.11	1.10	0.97	0.81	0.75	0.90	0.84	1.03	1.04	0.97
ROAR	TDDGW-MD	**1.01**	**0.86**	**0.98**	0.99	**0.98**	**0.89**	**0.75**	0.73	**0.85**	**0.81**	**0.96**	**0.97**	**0.90**
ROAR	TDDGWT-MD	**1.01**	0.87	**0.98**	**0.98**	**0.98**	**0.89**	0.76	**0.72**	0.86	0.82	**0.96**	**0.97**	**0.90**
ROAR	TDDGWD-MD	**1.01**	0.87	**0.98**	0.99	**0.98**	0.90	**0.75**	0.73	**0.85**	0.82	**0.96**	**0.97**	**0.90**
ROAR	TDDGWDT-MD	**1.01**	0.88	**0.98**	**0.98**	**0.98**	**0.89**	0.76	**0.72**	**0.85**	0.82	**0.96**	**0.97**	**0.90**

Table 4.4: MAE values of 2-hour-ahead forecasts from various forecasting models at PICT, JAYT, SPUR, and ROAR, 2010. The smallest MAE value of each column is in boldface.

besides at the hours. As a result, this method would be adjusted to the average wind speed of the day. Differently, MD, SMD, and YMD use exactly (and only provide) the average wind speed on the hours. Since the focus is on hourly ahead forecasting, here using MD, SMD, and YMD is reasonable without losing functionality in practice. Therefore in the following only forecasts from models that use the MD method to fit the diurnal component are displayed.

The MAE values of 2-hour-ahead forecasts from different models at PICT, JAYT, SPUR, and ROAR in 2010 are listed in Table 4.4, and their relative MAE values in percentages with PSS as the reference are presented in Table 4.5, which are cal-

Site	Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Overall
PICT	RSTD	11.7	**8.7**	11.9	14.4	11.4	15.7	13.7	13.9	13.7	16.0	15.6	11.1	13.2
PICT	TDD	9.9	7.1	11.4	14.2	13.1	14.5	13.9	11.3	12.6	12.0	14.1	9.4	12.1
PICT	LASSO-MD	1.4	0.0	7.1	8.3	3.1	11.3	11.2	10.4	6.3	10.0	11.6	6.6	7.5
PICT	TDDGW-MD	11.4	7.4	**20.7**	**22.3**	**19.9**	19.3	**19.5**	17.9	15.7	20.4	21.4	17.2	18.2
PICT	TDDGW-MD-T	11.4	5.9	20.4	21.9	**19.9**	19.6	**19.5**	18.1	15.7	**20.6**	**21.7**	16.9	18.1
PICT	TDDGWD-MD	13.6	5.4	19.9	22.2	19.4	19.9	19.1	18.1	**17.4**	19.7	21.3	**17.6**	**18.3**
PICT	TDDGWDT-MD	**13.8**	4.3	19.8	21.7	19.4	**20.1**	19.1	**18.2**	16.9	19.9	21.4	17.2	18.2
JAYT	RSTD	9.7	11.3	16.8	15.4	10.0	18.5	12.8	13.4	11.1	16.4	20.2	15.3	14.3
JAYT	TDD	10.2	11.5	17.4	15.5	11.8	17.8	17.8	12.5	12.7	17.5	19.8	16.3	15.1
JAYT	LASSO-MD	6.7	3.2	9.4	7.2	6.2	12.4	12.5	11.0	7.9	12.7	13.8	13.6	9.7
JAYT	TDDGW-MD	**13.8**	16.6	**22.3**	**21.6**	**19.5**	23.0	25.3	18.0	**18.4**	23.0	**24.8**	21.5	20.7
JAYT	TDDGW-MD-T	**13.8**	16.6	**22.3**	21.5	**19.5**	23.6	**25.9**	**18.6**	18.3	**23.1**	24.9	**22.4**	**20.9**
JAYT	TDDGWD-MD	**13.8**	**16.8**	22.2	**21.6**	19.0	23.1	25.2	17.8	18.3	22.8	24.7	21.4	20.6
JAYT	TDDGWDT-MD	**13.8**	**16.8**	22.0	21.5	19.0	**23.7**	25.8	18.4	18.2	22.9	**24.8**	22.3	20.8
SPUR	RSTD	14.1	14.2	12.7	14.5	15.6	17.5	18.2	17.0	15.9	20.4	20.7	14.0	16.0
SPUR	TDD	12.0	14.7	12.6	14.7	13.3	16.9	19.2	14.7	16.8	19.7	19.9	14.8	15.5
SPUR	LASSO-MD	3.8	4.6	12.0	10.1	6.8	16.6	13.2	12.3	15.1	16.5	17.8	13.2	11.8
SPUR	TDDGW-MD	14.2	17.3	**20.7**	22.7	21.4	27.2	**26.4**	21.3	24.0	25.2	**25.5**	**21.6**	22.2
SPUR	TDDGW-MD-T	14.2	17.3	20.6	22.4	21.2	27.2	26.3	21.3	23.7	25.2	**25.5**	21.1	22.1
SPUR	TDDGWD-MD	14.2	**17.5**	20.6	**23.3**	**21.6**	**27.4**	26.1	**21.7**	**24.6**	26.0	25.3	**21.6**	**22.4**
SPUR	TDDGWDT-MD	**14.3**	**17.5**	**20.7**	23.1	21.5	27.3	26.1	**21.7**	24.3	**25.9**	25.3	21.2	22.3
ROAR	RSTD	0.7	5.3	8.9	6.7	8.2	11.2	13	6.7	8.6	7.5	11.3	6.3	7.3
ROAR	TDD	2.1	**5.8**	10.4	10.9	7.9	11.6	14.2	8.6	8.2	**9.2**	12.3	8.9	9.1
ROAR	LASSO-MD	-4.7	4.0	10.1	9.5	3.4	11.5	10.2	13.3	9.3	5.0	13.3	5.9	7.4
ROAR	TDDGW-MD	**1.3**	4.9	**19.2**	19.0	**14.7**	18.1	**17.0**	16.2	14.2	8.2	**19.2**	12.6	**14.0**
ROAR	TDDGW-MD-T	1.2	4.4	19.0	19.5	14.3	**18.6**	16.7	**16.7**	14.0	8.0	**19.2**	12.5	**14.0**
ROAR	TDDGWD-MD	**1.3**	4.3	19.1	19.1	14.3	17.9	16.9	16.1	**14.3**	8.0	18.8	**12.9**	13.9
ROAR	TDDGWDT-MD	1.2	3.6	18.9	**19.7**	14.0	18.4	16.5	16.6	14.1	7.9	18.7	12.8	13.9

Table 4.5: Relative (to PSS) MAE values of 2-hour-ahead forecasts from various forecasting models at PICT, JAYT, SPUR, and ROAR, 2010 (%). The largest value of each column is in boldface.

culated from $(MAE_{PSS} - MAE)/MAE_{PSS}$. At PICT, it can be observed that all the space-time models outperform the PSS model as expected, with smaller MAE values. Except for Feb., our new models that incorporated geostrophic wind give more accurate forecasts than the RSTD and TDD models do, with the MAE value 0.88 compared with 0.94 and 0.95. Up to two decimal points, the TDDGW-MD, TDDGWT-MD, TDDGWD-MD and TDDGWDT-MD models have similar MAE values, around 0.88. Looking more closely, the TDDGWD-MD gives the largest reduction in the relative MAE value, around 18.3%. As expected, the models including geostrophic wind are more advanced than is the Lasso-based model, which has only a 7.5% improvement over PSS. They are also better than the other two space-time

models (RSTD and TDD) with 13.2% and 12.1% reductions in MAE values based on PSS. Comparing the results of the TDDGW and TDDGW-MD models, the modified diurnal pattern modeling based on the 45-day-sliding window helps to provide a 3.7% reduction in the MAE value relative to PSS.

Similar results for forecasts of 2-hour-ahead at JAYT, SPUR, and ROAR are found. At JAYT, the largest improvement over PSS is achieved by the TDDGWDT-MD model. It has a 20.9% MAE reduction compared with 9.7%, 14.3% and 15.1% by the LASSO, RSTD and TDD models, respectively, with overall MAE values equal 0.87, 0.82, and 0.83, respectively. At SPUR, the TDDGWD-MD model is the best overall, which reduces the relative MAE value by 22.4%, compared with 11.8%, 16.0% and 15.5% by the LASSO, RSTD and TDD models, respectively. At ROAR, with 14.0% reductions, the TDDGW-MD and TDDGWT-MD models perform the best. Overall, our new method obtains up to 8.2% improvement relative to the best previous space-time methods in this setting.

4.6 Final Remarks

Accurate wind prediction is critical in running power systems that have large shares of wind power. In recent decades many studies have been devoted to the wind forecasting problem for large-scale wind power development around the world.

This chapter develops statistical short-term wind forecasting models based on atmospheric dynamics principles. It proposes the use of the geostrophic wind as a predictor. The geostrophic wind is a good approximation to the winds in the extra-tropical free troposphere and can be estimated using only pressure and temperature data. In terms of the underlying atmospheric physics, the geostrophic wind is much more closely related to the real wind than either temperature or pressure. This is demonstrated by the fact that no improvement is found by directly incorporating

atmospheric temperature and pressure into the most advanced space-time forecasting model to date. The geostrophic wind can be approximated from networks of standard surface meteorological observations. More importantly, it helps to reduce prediction errors significantly when incorporated into space-time models.

In this chapter, more accurate forecasts were achieved by incorporating geostrophic wind information into space-time statistical models and modifying diurnal pattern models in 2-hour-ahead wind speed forecasting. Because directly including pressure and temperature, which are closely related to wind, into the TDD model did not lead to improvements, the geostrophic wind is turned to as a predictor. In addition, trigonometric functions of the geostrophic wind direction and temperature differences between the current and previous day are also considered. Simpler but more efficient methods can be applied to fit the diurnal pattern of wind to obtain better forecasts.

Based on our experiments with wind data from West Texas, for the 2-hour-ahead wind forecasting problem, our model with geostrophic wind achieves up to a 22.4% improvement relative to the persistence method, and up to 8.2% improvement relative to the best previous space-time methods in this setting.

5. SHORT-TERM SPATIO-TEMPORAL WIND POWER FORECAST IN LOOK-AHEAD POWER SYSTEM DISPATCH

5.1 Introduction

Uncertainties and variabilities in renewable generation, such as wind energy, pose significant operational challenges to power system operators Xie et al. (2011a). While conventional wisdom suggests that more spatially dispersed wind farms could be aggregated and "smooth out" total wind generation at any given time, the reality is that wind generation tends to be strongly correlated in many geographical regions Alexiadis et al. (1999). As many regions/states are moving toward renewable portfolio standards (RPS) in the coming decade, the role of *accurate wind prediction* is becoming increasingly important for many regional transmission organizations (RTOs) DOE (2008).

In power systems with high presence of intermittent generation, the main source of uncertainty comes from both demand and supply sides Xie et al. (2011a). State-of-the-art load forecasts could achieve high accuracy in the day-ahead stage Gross and Galiana (1987). Compared with load forecasting, accurate forecast of wind generation still remains an open challenge. There exists a large body of literature on wind power forecasting, and state-of-the-art day-ahead wind forecast based on numerical weather prediction (NWP) models has enabled relatively accurate wind forecast with approximately 15%-20% of wind speed forecast mean absolute error (MAE) Black (2012); IBM (2009); Mahoney et al. (2012); Constantinescu et al. (2011). As the operating time moves closer to the near term (e.g., hour-ahead or 15 minute-ahead), at a high spatial resolution, the computation complexity (in terms of simulation time and memory requirements) often renders NWP models intractable

Constantinescu et al. (2011). In sharp contrast, an alternative data-driven statistical model for near-term wind forecast could potentially provide *accurate and efficient* wind forecasts with MAE reduced to the range of around 5% or less Black (2012). A good set of references can be found in Zhu and Genton (2012). Our proposed spatio-temporal wind forecast model is directly targeted at *computationally efficient near-term* wind forecasts.

Starting from our preliminary work Xie et al. (2011b), the main objective of this chapter is to exploit a novel short-term spatio-temporal wind forecast model and quantify the dispatch benefits from improved short-term wind forecast. Wind generation is driven by wind patterns, which tend to follow certain geographical spatial correlations. For large-region wind farms, the wind generation forecast of the wind could significantly benefit from upstream wind power generation. Enabled by technological advances in sensing, communication, and computation, spatially correlated wind data could be leveraged for accurate system-wide short-term wind forecasts. This is potentially applicable to large-scale wind farms. The performance of such wind forecast model is critically assessed. Furthermore, It is proposed to quantify the economic benefits of improved forecast in a look-ahead economic dispatch model. The suggested contributions of this chapter are:

1. It is proposed to use two spatio-temporal correlated forecast models for short-term wind generation in a region, the TDD (trigonometric direction diurnal Hering and Genton (2010)) and the TDDGW (TDD with geostrophic wind information Zhu et al. (2013a)) models. Both forecasting models take into account local and nearby wind farms' historical wind information. Additionally, based on atmospheric dynamic principles, the latter incorporates geostrophic wind information and has better forecasts than the former one. Both methods

are tested with realistic wind data obtained in Texas, and they demonstrate improved forecast accuracy.

2. The spatio-temporal wind forecasts are incorporated into a look-ahead economic dispatch framework. Numerical study in an ERCOT equivalent 24-bus test system shows improved benefits compared with conventional static dispatch with time-persistent wind forecast models.

this chapter is organized as follows. In Section 5.2 an overview of statistical wind forecast models is provided, which is followed by the introduction of the proposed spatio-temporal wind forecast models. In Section 5.3 we compare the performance of spatio-temporal wind forecasts using realistic wind farm data obtained from West Texas. Section 5.4 presents the look-ahead economic dispatch formulation by incorporating available wind forecast. Numerical illustrations of the economic benefits of incorporating spatio-temporal wind forecast with look-ahead dispatch are presented in Section 5.5. Conclusions and future work are presented in Section 5.6.

5.2 Statistical Wind Forecasting

In this section, an overview and critical assessment of several statistical approaches to short-term wind forecasting is provided. Whereas NWP models play the key role in day-ahead to several hour-ahead wind forecasting, the computational burden and forecasting accuracy of NWP are still challenging in near-term forecasts (minutes-ahead to hour-ahead). As an alternative, data-driven statistical wind forecasting has gained increasing attention for near-term stage forecasts Hering and Genton (2010). A valuable review of short-term wind forecasting for power system operations can be found in Zhu and Genton (2012), including time series models and space-time models.

To assess the performance of several statistical wind forecasting models, West

ID	Location	Area	Latitude	Longitude	Elevation
ROAR	3N Roaring Springs	Roaring S./Motley County	$N33°56'10.86"$	$W100°50'43.38"$	2615 ft.
SPUR	1W Spur	Spur/Dickens County	$N33°28'51.05"$	$W100°52'34.90"$	2287 ft.
JAYT	1SSE Jayton	Jayton (Kent Co. Airport)	$N33°13'56.69"$	$W100°34'03.99"$	2010 ft.
PITC	10WSW Guthrie	Guthrie/King County	$N33°34'01.30"$	$W100°28'50.20"$	1998 ft.

Table 5.1: Site information

Texas is chosen as an example*. West Texas has rich wind resources and wind farms have been constructed on the large scale there. To forecast short-term wind speeds in West Texas, four statistical models are used, including persistence (PSS), autoregressive (AR) models and the *spatio-temporal correlated* TDD and TDDGW models. The first two models only capture the temporal correlation in wind, while the last two take into account both spatial and temporal correlations in wind by including wind information from neighboring wind farms. Better forecasts are thus expected.

5.2.1 Wind Data Source in West Texas

The wind data used here are the 5-minute averages of 3-second measurements of wind speed and direction collected by monitors placed at 10 meters from four sites in West Texas labeled ROAR, SPUR, PICT, and JAYT. Their locations are indicated by the red crosses in Figure 5.1, and more specific geographic information is listed in Table 5.1. The period of the wind data covers three years from January 1, 2008 to December 31, 2010. (The data sets are available at http://www.mesonet.ttu.edu/wind.html).

Winds in this area are mainly from the south or north as discussed in Section 4.4. Given the flatness in this area, the spatial correlation in wind can be captured when a southerly wind is blowing: wind at ROAR will mostly be just a shift from wind at

*While it is difficult to draw any general conclusion independent of geographical locations, the findings from this particular data set will be applicable to similar regions with high penetration of wind power.

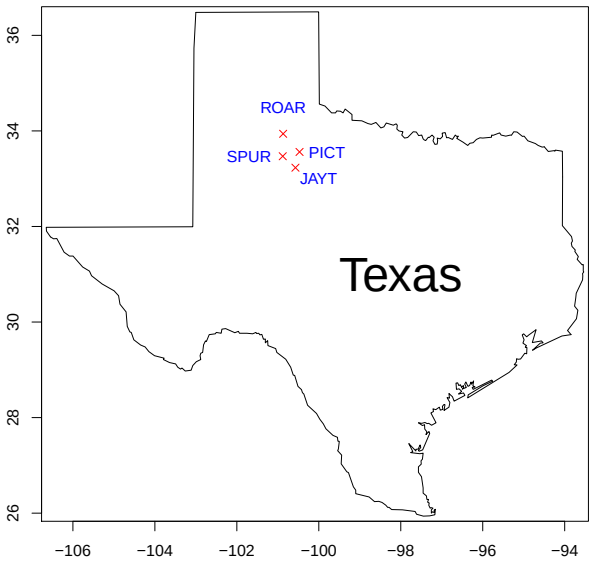

Figure 5.1: Map of the four locations in West Texas

SPUR. This means that to forecast the future wind speed at ROAR, it is definitely helpful to use the current and just past wind information at SPUR. Similarly, when the wind is blowing from the south or southeast, wind information at JAYT and PICT help in predicting the wind speed at ROAR. A good forecasting model should take into account both spatial and temporal correlations in wind.

5.2.2 Space-time Statistical Forecasting Models

Four statistical models, PSS, AR, TDD and TDDGW, are used to forecast short-term wind speed at each of the four sites. In the first two models, only the temporal correlation in wind is considered, while the TDD and TDDGW models utilize wind information from the other three locations so that both spatial and temporal correlations in wind are taken into account. Moreover, the TDDGW model incorporates geostrophic information into the TDD model.

To make it simple, the four models are introduced in the setting of forecasting wind speed at ROAR. Let $y_{R,t}$, $y_{S,t}$, $y_{J,t}$, and $y_{P,t}$ denote the wind speed at time t at ROAR, SPUR, JAYT, and PICT, respectively, and $\theta_{R,t}$, $\theta_{S,t}$, $\theta_{J,t}$, and $\theta_{P,t}$ denote the wind direction at time t. The goal is to estimate $y_{R,t+k}$, or the k-step-ahead wind speed at ROAR, denoted as $\hat{y}_{R,t+k}$, where each step is 5 minutes.

5.2.2.1 Persistent Forecasting

In the PSS model, it is assumed that the future wind speed is the same as the current one. For example, if $y_{R,t}$ is the wind speed at time t at ROAR, then the k-step future wind speed is predicted as $y_{R,t}$, or $\hat{y}_{R,t+k} = y_{R,t}$. PSS works very well for very short-term forecasting, such as 10-minute-ahead. The PSS model is usually treated as a reference and an advanced forecasting model is thought to be good if it outperforms PSS.

5.2.2.2 Autoregressive Models

AR models predict the future wind speed as a linear combination of past wind speeds. In our case, AR is applied to model the center parameter, $\mu_{R,t+k}$, in equation (**??**) (defined in the next part) as follows:

$$\mu^r_{R,t+k} = \alpha_0 + \sum_{i=0}^{p} \alpha_{i+1} \mu^r_{R,t-i}. \tag{5.1}$$

The AR model assumes that future wind speed is related to historical wind information only at the same location, without considering the spatial correlation.

5.2.2.3 Spatio-temporal Trigonometric Direction Diurnal Model

The TDD model is an advanced space-time statistical forecasting model. It generalizes the Regime-Switching Space-Time model Gneiting et al. (2006) by including wind direction in the model. As a probabilistic forecasting model, the TDD model

estimates a predictive distribution for wind speed at time $t+k$, thus providing more information about the uncertainty in wind. The TDDGW model, which incorporates geostrophic wind information into the TDD model, was proposed Zhu et al. (2013a) and more accurate forecasts are obtained than from the TDD model. In Section 2.4.3, an detailed introdcution of the TDD model is presented.

As we know, pressure and temperature also have significant effects on wind speed. If this information could be taken into account in wind forecasting problems, more accurate forecasts would be expected. However, it was found that adding surface pressure and temperature directly into the center parameter model brings no improvement to the forecasting accuracy. This is the motivation of the TDDGW model. It takes geostrophic wind, which extracts information on pressure and temperature, into the TDD model as a predictor. An detailed introdcution of the TDDGW model is given in Section 4.3.2.

5.3 Forecasting Results and Comparison

In this section, the aforementioned four forecasting models are implemented to forecast 10-minute-ahead, 20-minute-ahead and up to 1-hour-ahead wind speed at the four locations in West Texas on one day each month except May 2010. In the AR, TDD and TDDGW models, up to 45 days of historical wind records are used to estimate the coefficients in the trained models, which are based on 2008 and 2009 data. For the diurnal pattern, the averages of 45 days' hourly wind speeds are used. The median of the truncated normal distribution is used as a point forecast:

$$z^{+}_{1/2} = \mu_{t+1} + \sigma_{t+1} \cdot \Phi^{-1}[1/2 + (1/2)\Phi(-\mu_{t+1})/\sigma_{t+1}],$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

To evaluate the performance of the four forecasting models, mean absolute errors

(MAE), defined below, are calculated from the forecasts on the 11 days and listed in Table 5.2:

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |y_{R,t+k} - \hat{y}_{R,t+k}|,$$

where $T = 3168$ for 11 days.

Location	Model	10 min	20 min	30 min	40 min	50 min	60 min
PICT	PSS	0.56	0.72	0.84	0.92	1.00	1.08
	AR	0.55	0.70	0.80	0.87	0.94	1.00
	TDD	**0.54**	**0.68**	**0.77**	0.84	0.90	0.95
	TDDGW	**0.54**	**0.68**	**0.77**	**0.83**	**0.89**	**0.94**
JAYT	PSS	0.50	0.63	0.71	0.78	0.83	0.89
	AR	0.48	0.60	0.68	0.75	0.8	0.86
	TDD	**0.47**	**0.57**	**0.64**	0.69	0.73	0.78
	TDDGW	**0.47**	**0.57**	**0.64**	**0.68**	**0.71**	**0.75**
SPUR	PSS	0.51	0.64	0.73	0.81	0.86	0.92
	AR	0.49	0.61	0.69	0.76	0.80	0.86
	TDD	**0.48**	**0.59**	**0.67**	0.72	0.76	0.81
	TDDGW	0.49	**0.59**	**0.67**	**0.71**	**0.75**	**0.79**
ROAR	PSS	0.55	0.71	0.82	0.92	0.98	1.02
	AR	**0.54**	0.68	0.78	0.86	0.92	0.96
	TDD	**0.54**	**0.67**	0.77	0.85	0.90	0.93
	TDDGW	**0.54**	**0.67**	**0.76**	**0.82**	**0.87**	**0.90**

Table 5.2: MAE values of the 10-minute-ahead, 20-minute-ahead and up to 1-hour-ahead forecasts on 11 days' in 2010 from the PSS, AR, TDD and TDDGW models at the four locations (smallest in bold)

From Table 5.2, we can see that MAE values increase by column, which means that the forecast accuracy reduces when the forecasting horizon, k, gets larger. Among the four models, the AR, TDD, and TDDGW models have smaller MAE values than the PSS and the space-time models, TDD and TDDGW, are more advanced than the PSS and AR models with smaller MAE values. As expected, by incorporating the geostrophic wind information, the TDDGW model increases its

predictive accuracy. Its MAE values are reduced further compared with the TDD model, especially for 40-min-ahead or longer time lead forecasting. Relative to the MAE value of PSS, the TDDGW model obtains 15.7% reduction at JAYT for 1-hour-ahead forecasting, while it is 12.4% for the TDD model. This means that geostrophic wind information helps to improve the forecasting accuracy by 3.3% relative to the PSS model. The computational time for hour-ahead forecast using a laptop PC for one step of the TDDGW model is approximately 1.5 minutes, and the computational time for one step of TDD is approximately 1 minute. In contrast, NWP models for computing one step at hour-ahead forecast takes up to hours Constantinescu et al. (2011). Therefore, data-driven statistical wind forecast models provide computationally feasible solution for near-term operations for system operators. In the next two sections, the economic benefits of improved forecast are quantified in look-ahead dispatch models.

5.4 Power System Dispatch Model

With the spatio-temporal wind forecast models, it is presented in this section a critical assessment of the economic performance for power system operations. The power system scheduling framework formulated in this chapter approximates the existing scheduling practice in industry Electric Reliability Council of Texas (2010a), which is designed with two layers: 1) hourly scheduling and 2) real-time (every 5 minutes) scheduling. The notations are summarized in Table 5.3.

5.4.1 The Two-layer Dispatch Model

The structure of the two-layer dispatch model is described in Fig. 5.2. The models of hourly scheduling and real-time scheduling are presented below[†].

[†]As the main scope of this chapter is to study the short-term benefits (5 minutes to 1 hours) of spatio-temporal wind forecast, the day-ahead unit commitment process is not discussed in this chapter.

G	Set of conventional power plants
D	Set of inelastic loads
W	Set of wind farms
T	Look-ahead window size
n	number of wind farms installed in system
k	Time step in a look-ahead horizon with k_0 as its initial step
C_{G_i}	Generation cost function of power plant i
C_{W_i}	Generation cost function of wind farm i
C_{Rs_i}	Reserve cost function of power plant i
C_{U_i}	Start-up cost of generator i
C_{D_i}	Shut-down cost of generator i
$P_{G_i}^k$	Scheduled generation of power plant i at time k
$P_{W_i}^k$	Scheduled generation of wind farm i at time k
$P_{D_i}^k$	Forecasted load level of bus i at time k
$P_{Rs_i}^k$	Scheduled reserve capacity of power plant i at time k
$\mathbf{F}^k$	Vector of branch flow at time k
$\mathbf{F}^{max}$	Vector of capacity limits of transmission lines
$\triangle T$	Energy Market scheduling interval
P_i^R	Ramping constraints of power plant i
$P_{G_i}^{min}$	Lower operating limit for power plant i
$P_{G_i}^{max}$	Higher operating limit for power plant i
$P_{W_i}^{min}$	Lower operating limit for wind farm i
$P_{W_i}^{max}$	Higher operating limit for wind farm i
$\hat{P}_{W_i}^k$	Forecasted wind availability for wind farm i at time k
$\hat{\mathbf{P}}_W$	The vector of forecasted wind availability
$\tilde{\mathbf{P}}_W$	The vector of historical wind data
x_i^k	on/off status of generator i at time step k
$x_{U_i}^k$	binary indicators of starting-up and shutting down generator i
$x_{D_i}^k$	binary indicators of shutting-down generator i

Table 5.3: Notations

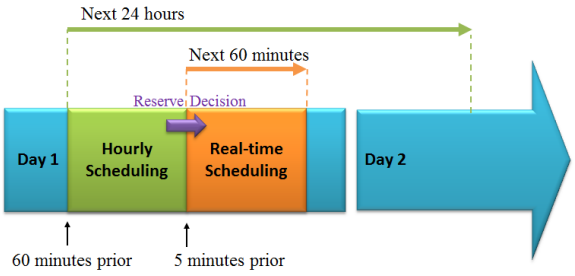

Figure 5.2: Two-layer dispatch model

5.4.1.1 Hourly Scheduling

The hourly scheduling model is a unit commitment model that takes place one hour prior to the real-time operation, as shown in Fig. 5.2. Energy balancing and ancillary services (reserve services) are co-optimized with start-up/shut-down decisions on fast units. Decisions on scheduling and reserves are generated every hour. For each hour, a rolling look-ahead dispatch plan over the next 24 hours is worked out. Only the dispatch results of the first hour will be executed. The generation outputs and reserves of the first two hours serve as inputs to the real-time scheduling level. The model is formalized as follows:

$$\min_{P^k_{G_i},P^k_{W_i},P^k_{Rs_i},x^k_i} : \sum_{k=k_0}^{T}[\sum_{i\in G} C_{G_i}(P^k_{G_i}) + \sum_{i\in W} C_{W_i}(P^k_{W_i}) + \sum_{i\in G} C_{Rs_i}(P^k_{Rs_i}) + \sum_{i\in F} C_{U_i}(x^k_{U_i}) + \sum_{i\in F} C_{D_i}(x^k_{D_i})] \tag{5.2}$$

s.t.

$$\sum_{i\in G} P^k_{G_i} + \sum_{i\in W} P^k_{W_i} = \sum_{i\in D} P^k_{D_i},\ k = k_0,\ldots,T \tag{5.3}$$

$$\sum_{i \in G} P_{Rs_i}^k \geq Rs^k,\ k = k_0, \ldots, T \tag{5.4}$$

$$|\mathbf{F}^k| \leqslant \mathbf{F}^{max},\ k = k_0, \ldots, T \tag{5.5}$$

$$|P_{G_i}^k - P_{G_i}^{k-1}| \leqslant P_i^R \Delta T, i \in G,\ k = k_0, \ldots, T \tag{5.6}$$

$$P_{G_i}^{\min} \leqslant P_{G_i}^k \leqslant P_{G_i}^{max},\ i \in G, k = k_0, \ldots, T \tag{5.7}$$

$$x_i^k P_{G_i}^{\min} \leqslant P_{G_i}^k \leqslant x_i^k P_{G_i}^{\max}, i \in G, k = k_0, \ldots, T \tag{5.8}$$

$$P_{G_i}^k + P_{Rs_i}^k \leqslant P_{G_i}^{max},\ i \in G, k = k_0, \ldots, T \tag{5.9}$$

$$P_{G_i}^k + P_{RS_i}^k \leqslant x_i^k P_{G_i}^{\max}, i \in G, k = k_0, \ldots, T \tag{5.10}$$

$$x_i^k - x_i^{k-1} \leqslant x_{U_i}^k, i \in G, k = k_0, \ldots, T \tag{5.11}$$

$$x_i^{k-1} - x_i^k \leqslant x_{D_i}^k, i \in G, k = k_0, \ldots, T \tag{5.12}$$

$$P_{W_i}^{min} \leqslant P_{W_i}^k \leqslant P_{W_i}^{max},\ i \in W, k = k_0, \ldots, T \tag{5.13}$$

$$P_{W_i}^k \leqslant \hat{P}_{W_i}^k = f(\tilde{\mathbf{P}}_W),\ i \in W, k = k_0, \ldots, T \tag{5.14}$$

$$x_i^k, x_{U_i}^k, x_{D_i}^k \in Binary, i \in G, k = k_0, \ldots, T. \tag{5.15}$$

In the proposed formulation, the objective function (5.2) is to minimize the power system operating costs including generation cost, reserve cost and start-up/shut-down cost of fast units. This scheduling problem is subject to various security constraints. (5.3) is the energy balancing equation. (5.4) is the system reserve requirement, which is often assessed according to system reliability requirement. The details of determination of reserve requirement are discussed in Section 5.4.2. (5.5) is the transmission capacity constraints. (5.6) are the ramping constraints of all generation units. (5.7) and (5.8) are the generators' capacity limits for general units and fast units respectively. (5.9) and (5.10) are the capacity constraints of general units and fast units for providing energy and reserve services. (5.11) and (5.12) are start-up/shut-down indicator constraints. (5.13) is the capacity limit of wind farms.

In this chapter, wind resources are assumed not to participate into ancillary services market providing reserve services. (5.14) is the wind forecast for each wind farm at time k, the details of which are explained in Section 5.2. (5.15) gives the binary constraints to integer decision variables.

5.4.1.2 Real-time Scheduling

Following the hourly scheduling, there is a real-time scheduling level in the dispatch framework, as shown in Fig. 5.2. It takes place every 5 minutes, which considers the scheduling for energy balancing with a look ahead window over the next 60 minutes. Only the dispatch decisions of the first five minutes will be executed. The inputs of reserved capacities, P_{Rs_i}, are obtained from the hourly scheduling level. The optimization problem is formulated as follows:

$$\min_{P^k_{G_i}, P^k_{W_i}} : \sum_{k=k_0}^{T} [\sum_{i \in G} C_{G_i}(P^k_{G_i}) + \sum_{i \in W} C_{W_i}(P^k_{W_i})] \tag{5.16}$$

s.t.

$$\sum_{i \in G} P^k_{G_i} + \sum_{i \in W} P^k_{W_i} = \sum_{i \in D} P^k_{D_i},\ k = k_0, \ldots, T \tag{5.17}$$

$$|\mathbf{F}^k| \leqslant \mathbf{F}^{max},\ k = k_0, \ldots, T \tag{5.18}$$

$$|P^k_{G_i} - P^{k-1}_{G_i}| \leqslant P^R_i \Delta T, i \in G \cup W,\ k = k_0, \ldots, T \tag{5.19}$$

$$P^{min}_{G_i} \leqslant P^k_{G_i} \leqslant P^{max}_{G_i} - P_{Rs_i},\ k = k_0, \ldots, T \tag{5.20}$$

$$P^{min}_{W_i} \leqslant P^k_{W_i} \leqslant P^{max}_{W_i},\ k = k_0, \ldots, T \tag{5.21}$$

$$P^k_{W_i} \leqslant \hat{P}^k_{W_i},\ k = k_0, \ldots, T \tag{5.22}$$

$$\hat{\mathbf{P}}_W = f(\tilde{\mathbf{P}}_W). \tag{5.23}$$

The objective function (5.16) is to minimize the total operating cost for energy balancing. In real-time scheduling, various security constraints are considered. Energy balancing constraints are provided in (5.17). Transmission capacity constraints are given in (5.18). Ramping constraints of generators are presented in (5.19). Capacity constraints of conventional generators and wind farms are described in (5.20) and (5.21), respectively. The dispatch points of wind generation should be no larger than the expected wind production potentials, as is shown in (5.22). The expected wind production potential is generated from wind forecast models, as given in (5.23), and studied in Section 5.2.

5.4.2 Procurement of Operating Reserves

In power system operations, reserve services play an important role to ensure the reliability of the system. Reserve services refer to a certain amount of capacity that is reserved to avoid power shortages due to loss of generation, fluctuations in demand and intermittent resources Wang et al. (2005). Due to the increasing penetration of intermittent resources (i.e., wind), more reserve services are required and deployed to tackle the uncertainty from the generation side. The amount of reserve services required serves an important cost component of integrating renewable generation.

The reserve services in various systems are designed differently. The reserve operation used by ERCOT in this chapter is investigated. There are mainly two kinds of reserve services provided in the ERCOT market: 1) Responsive Reserve Service (RRS) and 2) Non-Spinning Reserve Service (NSRS) Electric Reliability Council of Texas (2010a). RRS are deployed by ERCOT to restore the frequency of the ERCOT System from a contingency that causes a significant deviation from the standard frequency. RRS requires at least 2300 MW for all hours under normal conditions.

NSRS requires generation resources capable of being ramped to a specified output

level within thirty minutes and capable of running at a specified output level for at least one hour Electric Reliability Council of Texas (2011). NSRS are deployed to replace loss of generation, and to compensate for load forecast and/or wind forecast uncertainty.

In ERCOT, the NSRS requirement is determined by 1) the 95th percentile of the observed hourly net load[‡] uncertainty from the previous 30 days from the operation day and from the same month of the previous year and 2) the capacity of the largest unit installed in the system. This reserve requirement R_s at time step k can be formulated as (5.24), where Φ_k is the cumulative distribution function at time step k of observed hourly net load uncertainty from the previous 30 days and from the same month of the previous year, and P_R is the reliability requirement (e.g., 95%):

$$R_s^k = \max\{\Phi_k^{-1}(P_R), \max_{i \in G} P_{G_i}^{max}\}. \tag{5.24}$$

Generally, advanced wind forecasts with improved accuracy can result in the reduction of the necessary reserve requirement to deal with the wind generation uncertainty. However, the actual forecast errors of a wind forecast approach are unknown until the real-time operation. System operators have to estimate reserve requirements as discussed above. The estimated requirement does not necessarily decrease as the accuracy increases in wind generation forecast. The ERCOT methodology of determining reserve requirement is mainly based on historical data. In the long run, improvements in wind generation forecasts can be gradually reflected in their reserve requirements. In the short run, such improvements may be marginal and limited.

Besides the historical data driven approach (e.g., ERCOT), a probabilistic ap-

[‡]Net Load is defined as the ERCOT load minus the estimated total output from wind generation resources Electric Reliability Council of Texas (2011).

proach, which considers spatio-temporal correlation can be used especially for those forecast models that can generate forecast distributions.

It is assumed that the load forecast error and the wind forecast error follow normal distributions and are independent of each other. Given the N-1 (one component outage) criterion, the reserve requirements can be evaluated by solving (5.25):

$$P_{R_s} = \sum_{i \in G} P_{f_i} [\prod_{\substack{j \in G \\ j \neq i}} (1 - P_{f_j})][1 - \Phi(\frac{R_s - P_{G_i}^{\max}}{\sqrt{\sigma_{demand}^2 + \sigma_{wind}^2}})] + [\prod_{i \in G} (1 - P_{f_i})][1 - \Phi(\frac{R_s}{\sqrt{\sigma_{demand}^2 + \sigma_{wind}^2}})]. \quad (5.25)$$

In (5.25), P_{R_s} is the probability of the insufficiency of reserve capacity for coping with unexpected power imbalances from variable generation resources or demand, P_{f_i} is the probability of failure of unit i, σ_{demand} is the standard deviation of load forecast errors, σ_{wind} is the standard deviation of wind forecast errors, when the spatio-temporal correlation is considered. The detailed assessment of σ_{wind} is described in Section 5.2.2. $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution. Given the confidence interval (e.g., 95%), the reserve requirement can be determined by an equation solver such as the optimization toolbox in Matlab.

5.5 Numerical Experiment

In this section, a numerical experiment is conducted on a 24-bus ERCOT equivalent system to critically assess the *operational economic benefits* from improved short-term forecasts.

5.5.1 Simulation Platform Setup

The numerical example is modified from the IEEE Reliability Test System (RTS-24) Grigg et al. (1999). The simulation duration is 24 hours. The operation interval in

real-time scheduling is five minutes. The look-ahead horizon in real-time scheduling is 1 hour. Load profiles for 48 hours are collected from the ERCOT System Giberson (2009). Loads are scaled and factored out according to the portion of different buses. Wind forecasts are generated by various models discussed in Section 5.2 with forecast horizon which ranges from 10 minutes to 60 minutes.

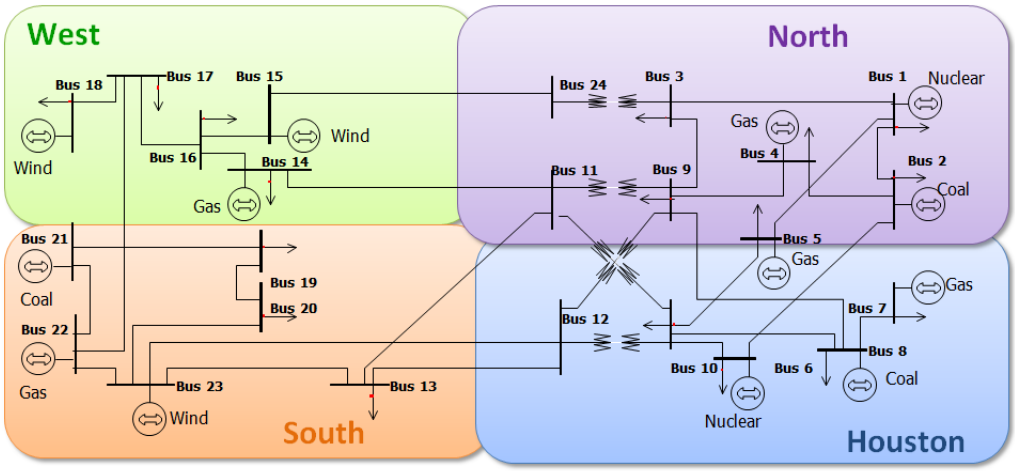

Figure 5.3: The IEEE RTS-24 system (modified)

The generator parameters are scaled according to Electric Reliability Council of Texas (2010). The generator capacity portfolio (the installed capacity percentage of different technologies) is configured and scaled from the real ERCOT system Electric Reliability Council of Texas (2010). The ramping rates and marginal costs are applied as shown in Table 5.4.

In the numerical studies, simulations of twelve sample days[§] are conducted. The

[§]Day 5 for TDDGW model is not available due to the inaccessibility of measurement data. Therefore, for the averaged MAE comparison of wind speed forecasts, only 11 days are considered. For the independent studies of economic benefits in power system operation, Day 5 for models other than TDDGW are presented.

Bus	Type	Cap. (MW)	Cost ($/MWh)	Ramping (MW/min)	Zone
1	Nuclear	140	15	1.12	North
2	Coal	540	20	10.8	North
4	Natural Gas	300	40	15	North
5	Natural Gas	510	37	33.15	North
6	Nuclear	150	11	1.35	Houston
7	Natural Gas	490	39	34.3	Houston
8	Coal	165	23	3.135	Houston
14	Natural Gas	170	38	15.3	West
16	Wind	200	6	18	West
18	Wind	240	4	24	West
21	Coal	300	21	5.4	South
22	Natural Gas	725	36	79.75	South
23	Wind	70	5	7.7	South

Table 5.4: Generator parameters

twelve days are selected as representative days for each month in 2010, as shown in Table 5.5.

Sample	Date	Sample	Date	Sample	Date
Day 1	10-Jan	Day 5	9-May	Day 9	8-Sep
Day 2	27-Feb	Day 6	16-Jun	Day 10	19-Oct
Day 3	12-Mar	Day 7	1-Jul	Day 11	22-Nov
Day 4	21-Apr	Day 8	17-Aug	Day 12	6-Dec

Table 5.5: Sample days in simulation study

The economic benefit from improved wind forecast will also depend on how system reserve is determined. In the simulation, two cases are conducted:

- Case A: The wind forecast quality is considered in determination of reserve requirements.

- Case B: The wind forecast quality is NOT considered in determination of reserve requirements. Only the installed wind capacity is involved.

In Case A, the reserve requirement is impacted by the wind forecast accuracy. An equivalent but simplified approach is taken in the simulation to mimic the methodology that is used by ERCOT to quantify the reserve requirement for those forecast models that do not provide wind forecast errors distribution information (e.g., the PSS model). A probabilistic approach is used for forecast models that provides distribution information of wind forecast errors. For a fair comparison, security factors are introduced to align the reserve requirement to the corresponding MAE performance.

5.5.2 Results and Analysis

In this section, the simulation results of the numerical experiments are presented. The distribution of the forecast errors of the wind generation reveals the accuracy of the forecast approach. The distribution of its errors for the perfect forecast (PF) with 100% accuracy is a concentrated spike at the zero origin of the x-axis. The better the forecast accuracy the closer the distribution pattern is to the central spike. A forecast model with poor accuracy has its errors distributed widely. The probability density distributions of the wind generation forecast errors (for a 200 MW wind farm) using the PSS, AR, TDD and TDDGW models under various simulation days are presented in Fig. 5.4. As we can observe, the distribution of the forecast errors of the PSS model is relatively spread out. The distribution of forecast errors of the TDD model is concentrated and has a higher central spike than do the AR and PSS models. The central spike of the TDDGW is higher than that of any other models. The shape of the forecast error distribution of the TDDGW model is closest to that of the perfect forecast. This is also verified by the wind speed forecast MAE presented in TABLE 5.2.

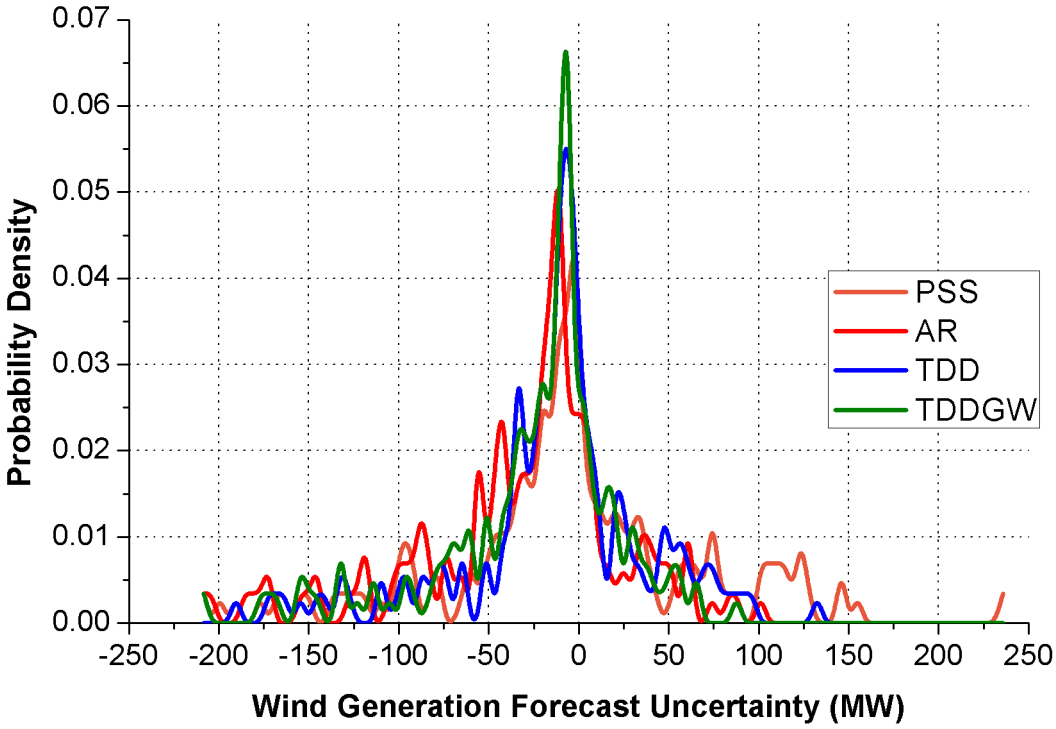

Figure 5.4: Distribution of forecast errors under different forecast models

By incorporating different forecast models into the power system economic dispatch, the economic performance differs. The economic performance results of Case A are presented in Fig. 5.5, which includes the total operating cost of each simulation day. The costs of the perfect forecast, PSS, AR TDD and TDDGW models are represented by the blue bar, the red bar, the green bar, the purple bar and the cyan bar, respectively. As we can see, for most of the cases, the spatio-temporal forecasts (TDD and TDDGW) have lower operating costs than do the PSS and AR models.

Taking the PSS model as a benchmark, the reduction in operating cost by percentage using various forecast models is presented in Fig. 5.6. As we can see, the TDD and TDDGW models, which consider spatio-temporal wind correlation, outperform the AR model and the PSS model in most of the cases. By incorporating

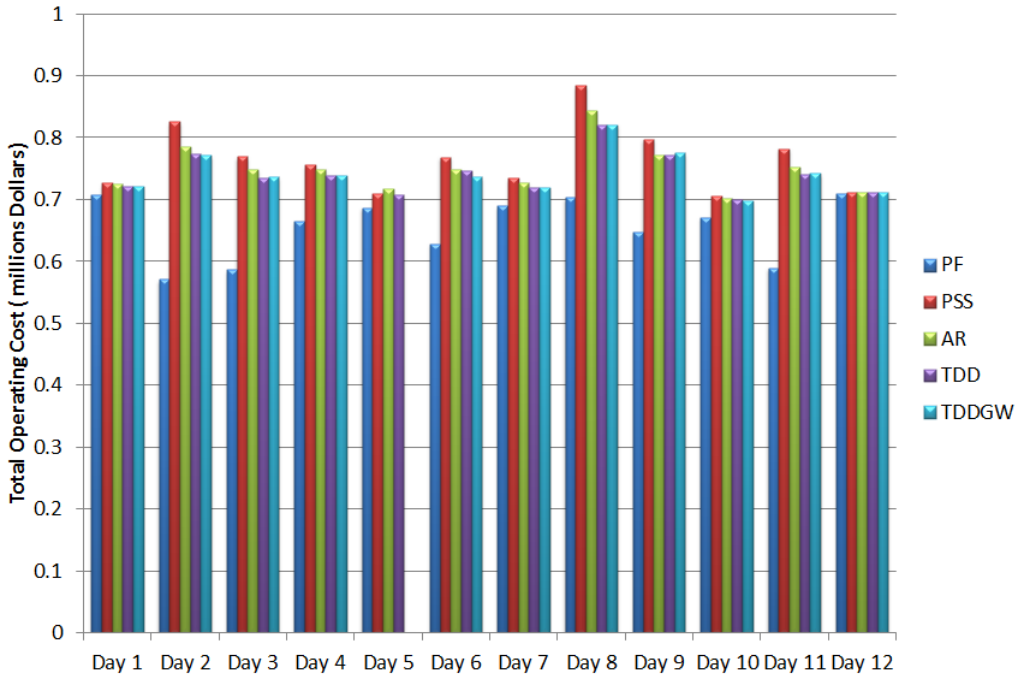

Figure 5.5: Total operating cost using different forecast models (Case A)

the effect of geostrophic wind, the TDDGW model can perform better than the TDD model. For most of the days, the AR model performs better than the PSS model. However, it is observed that for some days (Day 5), the AR model does not produce as good a forecast as does the PSS model. That is the limitation of wind forecast based on purely historical data. In contrast, by incorporating spatial correlations, the TDD model can produce more accurate forecasts than can the PSS model.

Detailed economic performance analysis is presented in Table 5.6¶. Due to the high wind production potential, Day 2 (27-Feb) is selected as a representative day for the study. According to Table 5.6, the TDD model performs better than the AR and PSS models. The reductions in the total operating cost by using the TDD and

¶In Table 5.6, the column "S.T. v.s. PSS" refers to the cost savings from the spatio-temporal models compared with the PSS model in percentage, and the row "Cost Reduced" refers to the overall savings of operating costs compared with the benchmark PSS model.

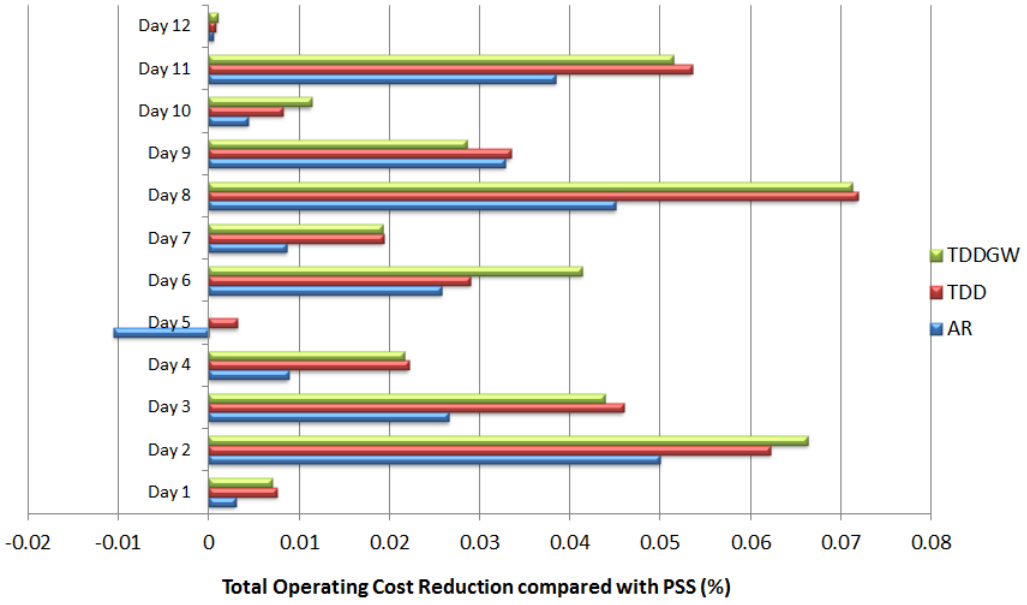

Figure 5.6: Operating cost reduction using different forecast models (Case A)

TDDGW models are about 7% lower than using the PSS model. Given the existing wind penetration (12.8%), a 6-7% cost savings is a considerable amount to the power industry. For energy services, the TDD and TDDGW models outperform the PSS, and AR models. Because of the improved forecast accuracy, less overestimation of wind generation occurs when using spatio-temporal models, which in turn reduces the amount of energy which is compensated by some expensive fast units in the real-time operation. There is an obvious reduction in the reserve cost when we compare the spatio-temporal models with the PSS and AR models. Compared with the PSS model, the AR model achieves a 9.3% reserve cost reduction, the TDDGW model implements a 12.8% reserve cost reduction and the TDD model contributes a 14% cost savings in reserve. By incorporating the spatio-temporal correlation, the improved forecast accuracy for the TDD and TDDGW models results in a decrease in the net load uncertainty which decreases the reserve requirements as well as the

Case A	PSS	AR	TDD	TDDGW	S.T. v.s. PSS
Total Cost	$826,081	$784,798	$774,688	$771,289	6.2% - 6.6%
Energy Balancing Cost	$654,440	$628,934	$627,124	$622,092	4.2% - 4.9%
Reserve Requirement	7435	6759	6403	6471	13.0% - 13.9%
Reserve Cost	$171,641	$155,864	$147,564	$149,196	13.1% - 14.0%
Cost Reduced	0.0%	5.0%	6.2%	6.6%	
Deviation Penalty	$109,797	$72,353	$77,716	$67,837	29.2% - 38.2%
Case B	PSS	AR	TDD	TDDGW	S.T. v.s. PSS
Total Cost	$900,183	$874,676	$872,867	$867,835	3.0% - 3.6%
Energy Balancing Cost	$654,440	$628,934	$627,124	$622,092	4.2% - 4.9%
Reserve Requirement	10684	10684	10684	10684	0.0%
Reserve Cost	$245,743	$245,743	$245,743	$245,743	0.0%
Cost Reduced	0.0%	2.8%	3.0%	3.6%	
Deviation Penalty	$109,797	$72,353	$77,716	$67,837	29.2% - 38.2%

Table 5.6: Economic performapce studies on 27-Feb (case A and case B)

reserve costs. Moreover, the wind deviation penalty Electric Reliability Council of Texas (2010a) is also reduced when using the spatio-temporal model. Instead of using the PSS model, the potential savings for wind farm owners on the wind generation deviation penalty for using the AR, TDD, and TDDGW models are 33.9%, 29.4% and 38.2%, respectively. Therefore, using spatio-temporal forecast not only benefits the power system operators but also benefits the individual wind farm owners.

The simulation of Case A is based on the assumption that forecast accuracy is considered in determining the reserve requirement. However, in many places the reserve requirement is only based on the installed wind capacity and the demand level. The simulation of Case B is thus conducted for comparison.

In Fig. 5.7, the operating cost from using different forecast models for all twelve days are presented. Compared with Fig. 5.5 of Case A, the difference in the total system operating cost for different wind forecast models is much smaller. This is because the reserve requirement stays at almost the same level no matter which

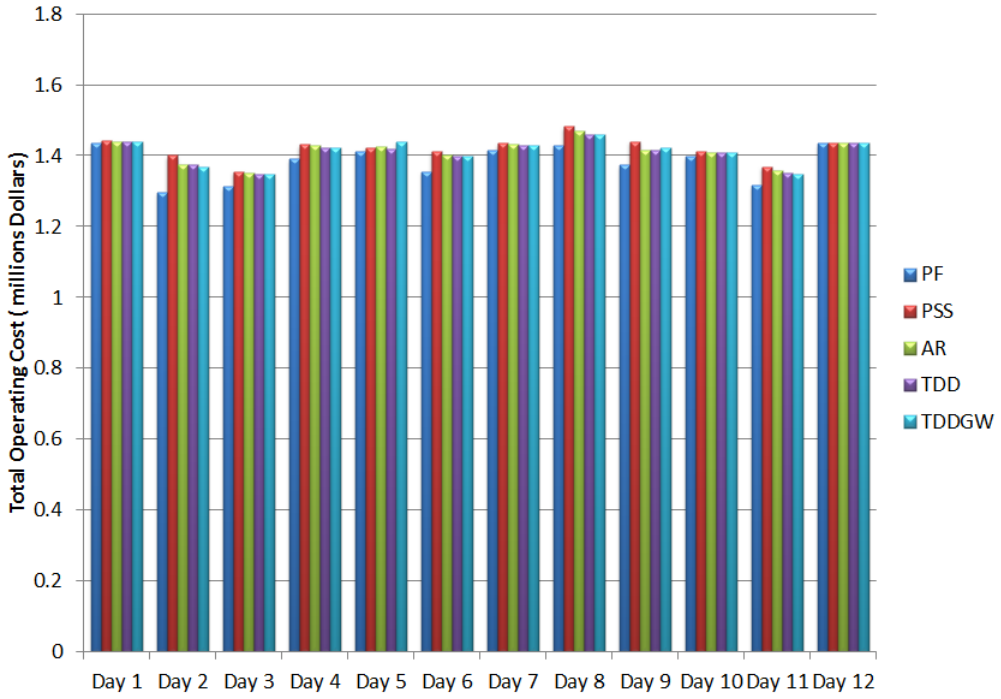

Figure 5.7: Total operating cost using different forecast models (Case B)

wind forecast model is used and no matter how accurate the wind forecast is.

More detailed results of the representative day (27-Feb) are presented in Table 5.6. Different from the results in Case A, the reserve costs are the same under different forecast models. However, the cost savings in the deviation penalty is still attractive because there occur fewer wind generation over-scheduling events. To achieve a more efficient power system with high penetration of wind generation, and to encourage advanced wind forecast technologies, it would be best to consider forecast accuracy in determining the reserve requirement.

5.6 Conclusions

Spatio-temporal wind forecast models (TDD and TDDGW models) are proposed and critically evaluated in this chapter. It is shown that by incorporating spatial

correlations of neighboring wind farms, the forecast quality in the near-term (hours-ahead) could be improved. The TDD and TDDGW models are incorporated into the look-ahead economic dispatch and unit commitment. Compared with conventional temporal-only statistical wind forecast models, such as the PSS models, the spatio-temporal models consider both the local and geographical wind correlations. By leveraging both temporal and spatial wind historical data, more accurate wind forecasts can be obtained. The potential economic benefits of advanced wind forecast are illustrated using an ERCOT equivalent 24 bus system. It is observed that the spatio-temporal model can increase wind resources utilization, and reduce system costs in both energy balancing and ancillary services. Such data-driven statistical methods for short-term wind forecast are also applicable in other similar regions with high wind penetration.

Future work will investigate the applicability of the proposed dispatch model to large-scale wind farms, such as offshore wind farms. Given the more consistent wind pattern over larger geographical areas, the potential benefits of the proposed method could be higher. Another important avenue for future research is to analyze the tradeoff between communication/computation burdens and the improved economic benefits by incorporating more spatially correlated wind data into power system dispatch models.

6. SUMMARY

Motivated by the needs for rapid development of wind power, this dissertation aims to provide accurate wind speed forecasts for power system operation, especially short forecast lead times, as well as seek a new realistic standard to evaluate the performance of forecasts. Specifically, following comprehensive literature review, two advanced space-time statistical probabilistic predictive models are proposed for short-term wind speed forecasting problems, and a new realistic method is introduced to evaluate forecasts based on power system operating costs through a power system dispatch.

The RRSTD model is proposed to generalize the RSTD model by allowing the forecast regimes to vary with the dominant wind direction and with the season without requiring much prior geographic information. Better forecasts are obtained than those from the PSS and AR models. Moreover, the RRSTD model has the potential to improve forecasts further if more information on monthly wind patterns are available, based on the numerical experiments with the wind data in the Pacific Northwest of the U.S.

The TDDGW model and its modifications are constructed by incorporating geostrophic wind into the TDD model. Based on atmospheric dynamic principles, the geostrophic wind is the wind under geostrophic balance. It can be used to approximate the higher level wind, which has significant effect on the lower level wind that rotates the wind turbines. This model is motivated by the fact that no improvement was found by directly incorporating atmospheric pressure and temperature into the TDD model. Based on the West Texas wind data, the TDDGW model achieves up to 22.4% improvement relative to the persistence method, which is significantly better

than the TDD and RSTD models in the 2-hour-ahead wind forecasting problems.

Moreover, a new realistic method is proposed to evaluate forecasts based on power system operating costs through a power system dispatch. The potential economic benefits of advanced wind forecast are illustrated using an ERCOT equivalent 24 bus system. The RRSTD wind forecasting model reduces 2.90% the cost of ancillary services, including regulation and reserve costs, compared with the PSS model. And for the TDDGW model, up to 6.6% overall generation cost can be reduced based on numerical simulation using look-ahead dispatch coupled with spatio-temporal wind forecast, compared with dispatch with persistent wind forecast models.

REFERENCES

M.C. Alexiadis, P.S. Dokopoulos, and H.S. Sahsamanoglou. Wind speed and power forecasting based on spatial correlation models. *IEEE Transactions on Energy Conversion*, 14:836–842, 1999.

S. Armstrong. *Long-range forecasting*. Wiley, 1985.

A. Azzalini and M.G. Genton. Robust likelihood methods based on the skew-t and related distributions. *International Statistical Review*, 76:106–129, 2008.

J. Black. Plans for wind integration in ISO-NE: Progress and challenges. In *8th annual Carnegie Mellon Conference on the electricity industry*, Pittsburgh, 2012.

E.A. Bossanyi. Short-term wind prediction using kalman filters. *Wind Engineering*, 9:1–8, 1985.

A. Bossavy, R. Girard, and G. Kariniotakis. Forecasting uncertainty related to ramps of wind power production. 2010. `http://www.cep.cma.fr/~st/rg/page4/files/ForecastingRampUncertainty.pdf`.

F. Bouffard and F. D. Galiana. Stochastic security for operations planning with significant sind power generation. In *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, pages 1–11, 2008.

BPA. *BPA Total Transmission System Load*. Bonneville Power Administration, 2007.

BPA. *2009 BPA Facts*. Bonneville Power Administration, 2010.

B.G. Brown, R.W. Katz, and A.H. Murphy. Time series models to simulate and forecast wind speed and wind power. *Journal of Climate and Applied Meteorology*, 23:1184–1195, 1984.

R.L. Busby. *Wind Power: the Industry Grows Up.* PennWell Corporation; Tulsa, Okla., 2012.

E. Cadenas and W. Rivera. Short term wind speed forecasting in la venta, oaxaca, mexico, using artificial neural networks. *Renewable Energy*, 34:274–278, 2009.

J.A. Carta and P. Ramírez. Use of finite mixture distribution models in the analysis of wind energy in the canarian archipelago. *Energy Conversion and Management*, 48:281–291, 2007.

J.A. Carta, P. Ramírez, and S. Velazquez. A review of wind speed probability distributions used in wind energy analysis: Case studies in the canary islands. *Renewable and Sustainable Energy Reviews*, 13:933–955, 2009.

Z. Chen and Y. Yang. Assessing forecasting accuracy measures. 2004. Preprint Series.

E.M. Constantinescu, V.M. Zavala, M. Rocklin, Lee Sangmin, and M. Anitescu. A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation. *Power Systems, IEEE Transactions on*, 26(1):431–441, 2011.

CREIA. *2010 China Wind Power Outlook.* China Renewable Energy Industries Association, 2010.

P. Crochet. Adaptive kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in iceland. *Meteor. Appl.*, 11:173–187, 2004.

I.G. Damousis and P. Dokopoulos. A fuzzy model expert system for the forecasting of wind speed and power generation in wind farms. In *Proceedings of the IEEE International Conference on Power Industry Computer Applications PICA*, volume 1, pages 63–69, 2006.

I.G. Damousis, M.C. Alexiadis, J.B. Theocharis, and P. Dokopoulos. A fuzzy model for wind speed prediction and power generation in wind farms using spatial corre-

lation. *IEEE Transactions on Energy Conversion*, 19:352–361, 2004.

X. de Luna and M.G. Genton. Predictive spatio-temporal models for spatially sparse environmental data. *Statistica Sinica*, 15:547–568, 2005.

R. Dell'Aquila and E. Ronchetti. Robust tests of predictive accuracy. *Metron*, 62: 161–184, 2004.

F.X. Diebold and R.S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13:253–263, 1995.

DOE. *20% Wind Energy by 2030: Increasing Wind Energy's Contribution to U.S. Electricity Supply*. The US Department of Energy, 2008.

DOE. Large-scale offshore wind power in the united states. 2010. `www.nrel.gov/wind/pdfs/40745.pdf`.

R. Doherty and M. O'Malley. A new approach to quantify reserve demand in systems with significant installed wind capacity. *Power Systems, IEEE Transactions on*, 20(2):587–595, 2005.

Electric Reliability Council of Texas. ERCOT 2009 Annual Report, 2010. `http://www.ercot.com/news/presentations/2010/index`.

Electric Reliability Council of Texas. ERCOT Nodal Protocals, Section 6: Adjustment Period and Real-Time Operations, July 2010a. `http://www.ercot.com/mktrules/nprotocols/`.

Electric Reliability Council of Texas. *ERCOT Quick Facts*. Electric Reliability Council of Texas, 2010b.

Electric Reliability Council of Texas. ERCOT Methodologies for Determining Ancillary Service Requirements, Dec. 2011. `http://www.ercot.com/content/meetings/rdtf/keydocs/2011/1220/`.

EU. Climate change: Commission welcomes final adoption of europe's climate and energy package. 2008. `http://europa.eu/rapid/pressReleasesAction.`

do?reference=IP/08/199.

EWEA. *Wind in Power: 2011 European Statistics*. The European Wind Energy Association, 2012.

R. Fildes, K. Nikolopoulos, S.F. Crone, and A.A. Syntetos. Forecasting and operational research: A review. *Journal of the Operational Research Society*, 59: 1150–1172, 2008.

B.E. Flores. A pragmatic view of accuracy measurement in forecasting. *Omega (Oxford)*, 14:93–98, 1986.

H. Geerts. Short range prediction of wind speeds: a system-theoretic approach. In *Proceedings of European wind energy conference*, 1984.

M.G. Genton and A.S. Hering. Blowing in the wind. *Significance*, 4:11–14, 2007.

R. Giacomini and H. White. Tests of conditional predictive ability. *Econometrica*, 74:1545–1578, 2006.

M. Giberson. Power prices in ERCOT's west zone: a mix of wind power, natural gas prices, transmission constraints, and (inefficient) congestion management practices. 2009. http://knowledgeproblem.com/2009/07/22/ercot-west-power-prices-2009-jan-jun/.

G. Giebel. *On the benefits distribution generation of wind energy in Europe*. PhD thesis, the Carl von Ossietzky Universitä Oldenburg, Düsseldorf, 2001.

G. Giebel. *The State-Of-The-Art in Short-Term Prediction of Wind Power: A Literature Overview*. RisøNational Laboratory, Denmark, 2003.

G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl. The state-of-the-art in short-term prediction of wind power : A literature overview, 2nd edition. *ANEMOS.plus*, 2011.

T. Gneiting. Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106:746–762, 2011a.

T. Gneiting. Quantiles as optimal point forecasts. *Journal of the American Statistical Association*, 27:197–207, 2011b.

T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 201:359–378, 2007.

T. Gneiting, K. Larson, K. Westrick, M.G. Genton, and E. Aldrich. Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space-time method. *Journal of the American Statistical Association*, 101:968–979, 2006.

T. Gneiting, L.I. Stanberry, E.P. Grimit, L. Held, and N.A. Johnson. Assessing probabilistic forecasts of multivariate quantities, with applications to ensemble predictions of surface winds. *Test*, 17:211–235, 2008.

C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, and C. Singh. The IEEE reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee. *Power Systems, IEEE Transactions on*, 14(3):1010–1020, 1999.

G. Gross and F.D. Galiana. Short-term load forecasting. *Proceedings of the IEEE*, 75:1558 – 1573, 1987.

Y. Gu and L. Xie. Look-ahead coordination of wind energy and electric vehicles: A market-based approach. In *North American Power Symposium*, The University of Texas at Arlington, 2010.

D.M. Haugen and S. Musser. *Renewable Energy.* Greenhaven Press; Detroit, 2012.

A.S. Hering and M.G. Genton. Powering up with space-time wind forecasting. *Journal of the American Statistical Association*, 105(489):92–104, 2010.

A.S. Hering and M.G. Genton. Comparing spatial predictions. *Technometrics*, 2011.

tentatively accepted.

IBM. Deep thunder-precision forecasting for weather-sensitive business operations. 2009. http://www.research.ibm.com/weather/DT.html.

J. Jeon and J. Taylor. Using conditional kernel density estimation for wind power density forecasting. *Journal of the American Statistical Association*, 2011. To appear.

L. Kamal and Y.Z. Jafri. Time series models to simulate and forecast honotey average wind speed in quetta. *Solar Energy*, 61:23–32, 1997.

C. Kamath. Understanding wind ramp events through analysis of historical data. 2010. http://computation.llnl.gov/casc/StarSapphire/pubs/LLNL-CONF-416432.pdf.

G. Kariniotakis, P. Pinson, N. Siebert, G. Giebel, and R. Barthelmie. The-state-of-the-art in short-term prediction of wind power - from an offshore perspective. In *Symposium ADEME, IFREMER*, 2004.

A. Lau and P. McSharry. Approaches for multi-step density forecasts with application to aggregated wind power. *Annals of Applied Statistics*, 4:1311–1341, 2010.

P. Louka, G. Galanisa, N. Siebertd, G. Kariniotakisd, P. Katsafadosa, I. Pytharoulisa, and G. Kallosa. Improvements in wind speed forecasts for wind power prediction purposes using kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics*, 96:2348–2362, 2008.

W.P. Mahoney, K. Parks, G. Wiener, Y. Liu, W.L. Myers, J. Sun, L. Delle Monache, T. Hopson, D. Johnson, and S.E. Haupt. A wind power forecasting system to optimize grid integration. *Sustainable Energy, IEEE Transactions on*, 3(4):670–682, 2012.

Y. Makarov, S. Lu, B. McManus, and J. Pease. The future impact of wind on bpa power system ancillary services. In *Transmission and Distribution Conference and*

Exposition, IEEE/PES, 2008.

A. Malmberg, U. Holst, and J. Holst. Forecasting near-surface ocean winds with kalman filter techniques. *Ocean Engineering*, 32:273–291, 2005.

J. Marquis, M.and Wilczak, M. Ahlstrom, J. Sharp, A. Stern, J.C. Smith, and S. Calvert. Forecasting the wind to reach significant penetration levels of wind energy. *Bulletin of the American Meteorological Society*, 92:1159–1171, 2011.

R.A. Meese and K. Rogoff. Was it real? the exchange rate - interest differential relation over the modern floating-rate period. *Journal of Finance*, 43:933–948, 1988.

P. Meibom, R. Barth, B. Hasche, H. Brand, C. Weber, and M. O'Malley. Stochastic optimization model to study the operational impacts of high wind penetrations in ireland. *Power Systems, IEEE Transactions on*, 26(3):1367–1379, 2011.

C. Monteiro, H. Keko, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann. A quick guide to wind power forecasting: Stat-of-the-art 2009. 2009.

A. Papavasiliou, S.S. Oren, and R.P. O'Neill. Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *Power Systems, IEEE Transactions on*, 26(4):2197–2206, 2011.

V.S. Pappala, I. Erlich, K. Rohrig, and J. Dobschinski. A stochastic model for the optimal operation of a wind-thermal power system. *Power Systems, IEEE Transactions on*, 24(2):940–950, 2009.

P. Pinson and R Hagedorn. Verification of the ecmwf ensemble forecasts of wind speed against analyses and observations. *Meteorological Applications*, 38(10):1089–1099, 2011.

P. Pinson and H. Madsen. Adaptive modeling and forecasting of wind power fluctuations with markov-switching autoregressive models. *Journal of Forecasting*, 31: 281–313, 2012.

P. Pinson, C. Chevallier, and G.N. Kariniotakis. Trading wind generation fom short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems*, 22:1148–1156, 2007.

P. Ramírez and J.A. Carta. Influence of the data sampling interval in the estimation of the parameters of the weibull wind speed probability density distribution: a case study. *Energy Conversion and Management*, 46:2419–2438, 2005.

Reuters. Texas electric grid sets new wind generation record. 2013. Retrieved 2013-03-17.

M. Rugbjerg, O.R. Sorensen, and V. Jacobsen. Wave forecasting for offshore wind farms. In 9^{th} *International Workshop on Wave Hindcasting and Forecasting*, Victoria, B.C. Canada, September 2006.

U. Schlink and G. Tetzlaff. Wind speed forecasting from 1 to 30 minutes. *Theoretical and Applied Climatology*, 60:191–198, 1998.

A. Sfetsos. A comparison of various forecasting techniques applied to mean honotey wind speed time series. *Renewable Energy*, 21:23–35, 2000.

A. Sfetsos. A novel approach for the forecasting of mean honotey wind speed time series. *Renewable Energy*, 27:163–174, 2002.

J.M. Sloughter, T. Gneiting, and A.E. Raftery. Probabilistic wind speed forecasting using ensembles and bayesian model averaging. *Journal of the American Statistical Association*, 105:23–35, 2010.

L. Soder. Simulation of wind speed forecast errors for operation planning of multiarea power systems. In *Probabilistic Methods Applied to Power Systems, 2004 International Conference on*, pages 723 –728, sept. 2004.

J. Tambke. Forecasting offshore wind speeds above the north sea. *Wind Energy*, 8: 3–6, 2004.

J. Tambke, M. Lange, U. Focken, and D. Heineman. Previento meets horns rev

short-term wind power prediction-adaptation to offshore sites. June 2003.

C. Tantareanu. Wind prediction in short-term: a first step for a better wind turbine control. In *Nordvestjysk Folkecenter for Vedvarende Energi*, October 1992.

J. Tastu, P. Pinson, E. Kotwa, H.Aa. Nielsen, and H. Madsen. Spatio-temporal analysis and modeling of wind power forecast errors. *International Statistical Review*, 14:43–60, 2011.

T.L. Thorarinsdottir and T. Gneiting. Probabilistic forecasts of wind speed: Ensemble model output statistics using heteroskedastic censored regression. *Journal of the Royal Statistical Society Series A*, 173:371–388, 2010.

R. Tibshirani. Regression shrinkage and selection via the lasso. *J. Royal. Statist. Soc B*, 58:267–288, 1996.

R.S. Tsay. *Analysis of financial time series*. Wiley, third edition edition, 2010.

A. Tuohy, P. Meibom, E. Denny, and M. O'Malley. Unit commitment for systems with significant wind penetration. *Power Systems, IEEE Transactions on*, 24(2): 592–601, 2009.

J.M. Wallace and P.V. Hobbs. *Atmospheric Science: An Introductory Survey*. Elsevier Academic Press Amsterdam; Boston, 2006.

J. Wang, X. Wang, and Y. Wu. Operating reserve model in the power market. *Power Systems, IEEE Transactions on*, 20(1):223–229, 2005.

J. Wang, M. Shahidehpour, and Z. Li. Security-constrained unit commitment with volatile wind power generation. *Power Systems, IEEE Transactions on*, 23(3): 1319–1327, 2008.

S.J. Watson, L. Landberg, and J.A. Halliday. Application of wind speed forecasting to the integration of wind energy into a large scale power system. *Generation, Transmission and Distribution, IEE Proceedings-*, 141(4):357–362, 1994.

C.K. Wikle and N. Cressie. A dimension-reduced approach to space-time kalman

filtering. *Biometrika*, 86:815–829, 1999.

World Wind Energy Association. World wind energy report, 2010. `http://www.wwindea.org/home/images/stories/pdfs/worldwindenergyreport2010s.pdf`.

L. Wu, M. Shahidehpour, and T. Li. Stochastic security-constrained unit commitment. *Power Systems, IEEE Transactions on*, 22(2):800–811, 2007.

L. Xie, P.M.S. Carvalho, L.A.F.M. Ferreira, J. Liu, B.H. Krogh, N. Popli, and M.D. Ilić. Wind energy integration in power systems: Operational challenges and possible solutions. *Special Issue of the Proceedings of IEEE on Network Systems Engineering for Meeting the Energy and Environment Dream*, 99(1):214–232, 2011a.

L Xie, Y. Gu, X. Zhu, and M.G. Genton. Power system economic dispatch with spatio-temporal wind forecasts. In *Energytech, 2011 IEEE*, pages 1–6, 2011b.

L. Zhao and B. Zeng. Robust unit commitment problem with demand response and wind energy. *University of South Florida, Tech. Rep*, 2010.

X. Zhu and M.G. Genton. Short-term wind speed forecasting for power system operation. *International Statistical Review*, 80:2–23, 2012.

X. Zhu, K. Bowman, and M.G. Genton. Incorporating geostrophic wind information for improved space-time wind speed forecasting. 2013a. Submitted.

X. Zhu, M.G. Genton, Y. Gu, and L. Xie. Short-term wind speed forecasting for improved power system dipatch. 2013b. Submitted.