# Short-Term Wind Speed Forecasting for Power System Operations

Xinxin Zhu and Marc G. Genton[1]

November 24, 2011

## Abstract

The emphasis on renewable energy and concerns about the environment have led to large-scale wind energy penetration worldwide. However, there are also significant challenges associated with the use of wind energy due to the intermittent and unstable nature of wind. High-quality short-term wind speed forecasting is critical to reliable and secure power system operations. This article begins with an overview of the current status of worldwide wind power developments and future trends. It then reviews some statistical short-term wind speed forecasting models, including traditional time series approaches and more advanced space-time statistical models. It also discusses the evaluation of forecast accuracy, in particular the need for realistic loss functions. New challenges in wind speed forecasting regarding ramp events and offshore wind farms are also presented.

**Key words:** Evaluation; Forecasting; Loss function; Ramp event; Space-time model; Statistical model; Time series model; Wind speed; Wind power.

**Short title**: Short-Term Wind Speed Forecasting

[1]Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA.
E-mail: {xzhu, genton}@stat.tamu.edu

# 1 Introduction

## 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, 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 1 shows. The total installed capacity reached $196,630$ Megawatts (MW) by the end of 2010, out of

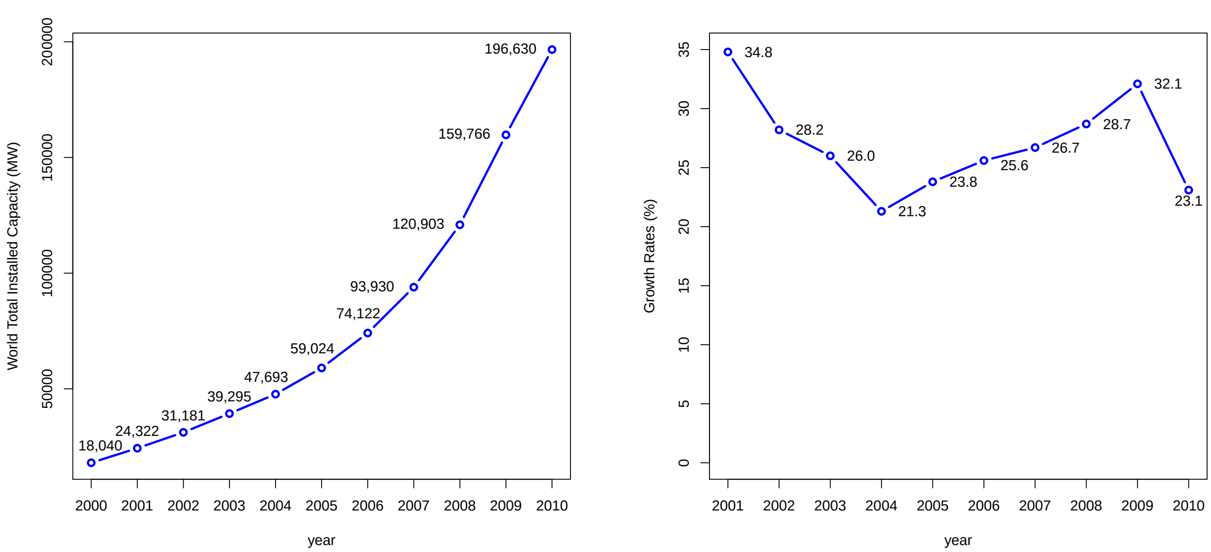


Figure 1: Left panel: world total installed wind power capacity from year 2000 to 2010. Right panel: world market growth rate of newly installed wind power capacity from the installed capacity of the previous year during 2000 to 2010.

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 (WWEA, 2009). 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 1.

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 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.,

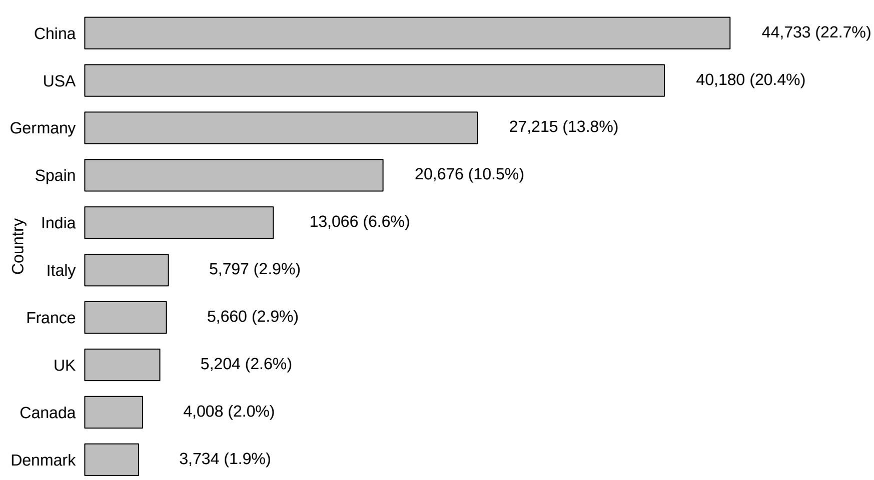


Figure 2: Country shares of total installed wind power capacity (in MW and percentage) by the end of 2010.

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 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.

## 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. (2011) 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 et al., 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. (2011). 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 & Hering, 2007) in terms of forecast accuracy and model interpretation, and it is the main topic of this article.

### 1.3 Outline

Short-term wind prediction has been the focus of extensive research in the last decade. The motivation of this article 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 describes the relationship between wind speed forecasting and wind power forecasting and the recent trend away from point forecasting to probabilistic forecasting. Section 3 summarizes some traditional time series statistical models of wind speed forecasting, including autoregressive models and the Kalman filter method. In Section 4, space-time statistical forecasting models are introduced. Evaluation of wind speed forecasting models is discussed in Section 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 6.

## 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.

## 2.1 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 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, 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 & Taylor (2011) has however

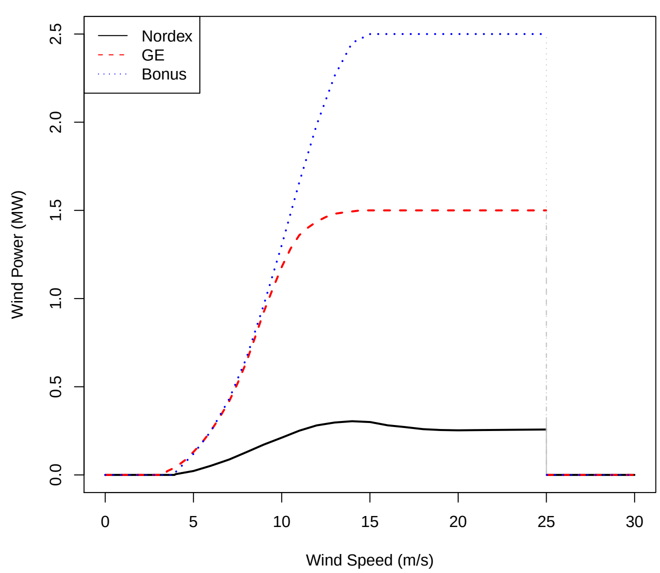


Figure 3: Three power curves with different capacity ranges from low to high from three manufacturers: 0.3 MW from Nordex, 1.5 MW from GE, and 2.5 MW from Bonus.

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 article focuses on wind speed forecasting.

## 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 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 4 for illustration based on data in Hering & 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 & 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. (2008). Mixture distributions with one

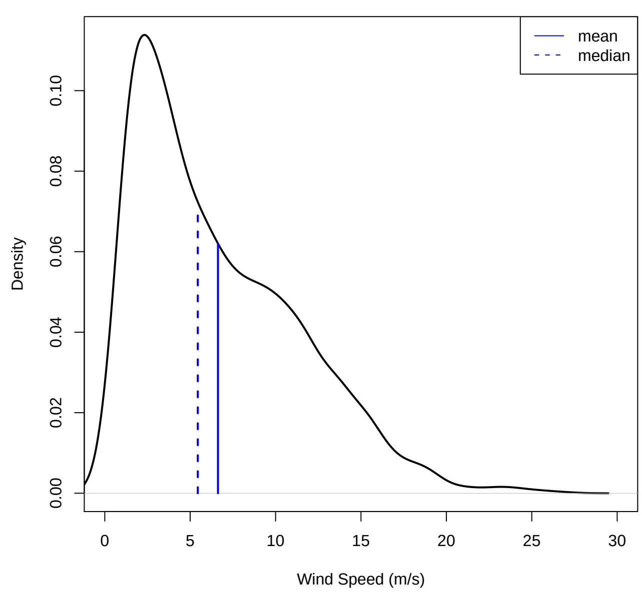


Figure 4: Nonparametric density estimation of 2002 hourly wind speed data at Vansycle, Oregon, U.S. The vertical lines represent the sample mean (solid) of 6.6 m/s and the sample median (dashed) of 5.4 m/s. The sample skewness is 0.8.

Weibull and one truncated normal distribution have been fitted to bimodal wind speeds, taking into account null wind speeds as well; see Carta & 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 & 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 & Genton, 2008, and references therein) has been used by Hering & 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).

## 2.3 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 & 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 article. In the next two sections, conventional time series methods and space-time models for short-term wind speed forecasting are reviewed and discussed.

# 3 Time Series Models for Forecasting

## 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 article, we consider short-term wind speed forecasting 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 & 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.

## 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.

## 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, $\phi_i$, $i = 1, \ldots, p$ are the autoregressive parameters, $\epsilon_t$ 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 $\phi_i$, $i = 1, \ldots, p$ 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 & 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 $\theta_j$'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 & Jafri

(1997) for more detail.

## 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 we illustrate 1-step ahead forecasting:

$$y_t = \mathbf{H}_t'\mathbf{A}_t + \nu_t, \tag{1}$$

$$\mathbf{A}_{t+1} = \mathbf{\Phi}\mathbf{A}_t + \boldsymbol{\omega}_t. \tag{2}$$

Equation (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) 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 $\nu_t$ 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:

$$\mathbf{A}_t = \mathbf{A}_{t|t-1} + \mathbf{K}_t(y_t - \mathbf{H}_t'\mathbf{A}_{t|t-1}), \tag{3}$$

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 (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. Geert (1985) 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. (2006) 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 & 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.

# 4 Space-Time Statistical Models for Forecasting

## 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 & 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 & 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.

## 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 5. These three locations are along the Columbia River Gorge which runs from east to west. Due to the high terrain to the north and south, the airflow runs parallel

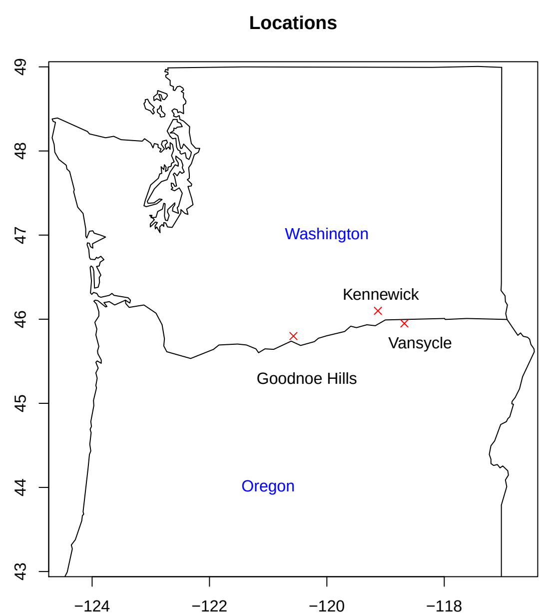


Figure 5: Map of the three locations: Vansycle, Kennewick and Goodnoe Hills on the border between Washington and Oregon in the U.S..

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 $\mu_{t+2}$ and scale parameter $\sigma_{t+2}$, that is, $V_{t+2} \sim N^+(\mu_{t+2}, \sigma^2_{t+2})$. The key is in modeling $\mu_{t+2}$ and $\sigma_{t+2}$. For the center parameter $\mu_{t+2}$, 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, $\mu^r_{t+2}$, 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):

$$\mu^r_{t+2} = a_0 + a_1 V^r_t + a_2 V^r_{t-1} + a_3 K^r_t + a_4 K^r_{t-1} + a_5 G^r_t, \tag{4}$$

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:

$$\mu_{t+2} = a_0 + a_1 V_t + a_2 K_t. \tag{5}$$

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

$$\sigma_{t+2} = b_0 + b_1 v_t, \tag{6}$$

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 (4), (5) and (6) are estimated by the continuous ranked probability score (CRPS) method; see Gneiting & 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.

### 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 & 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 & Genton (2010) modeled the residual predictive center, $\mu^r_{t+2}$, of a truncated normal distribution based on the present and past residual wind speed series at all three locations, as:

$$\begin{aligned}\mu^r_{t+2} &= a_0 + a_1 V^r_t + a_2 V^r_{t-1} + a_3 K^r_t + a_4 K^r_{t-1} + a_5 G^r_t + a_6 \sin(\theta^r_{V,t}) + a_7 \cos(\theta^r_{V,t}) \\ &\quad + a_8 \sin(\theta^r_{K,t}) + a_9 \cos(\theta^r_{K,t}) + a_{10} \sin(\theta^r_{G,t}) + a_{11} \cos(\theta^r_{G,t}), \end{aligned} \tag{7}$$

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 (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, $\sigma_{t+2}$, 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.

## 4.4 Other Models

There are some other interesting statistical models for short-term wind speed forecasting. For example, Hering & 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 forecasting 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 & Rivera (2000). Fuzzy models were applied to wind speed forecasting by Damousis & Dokopoulos (2001) 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.

# 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, we emphasize the importance of matching loss functions and forecasts, 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.

## 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 we 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:

$$\hat{y}_t = \arg\min_{\hat{x}_t} E_F\{L(y_t, \hat{x}_t)\}. \tag{8}$$

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.

## 5.2 Realistic Loss Functions for Wind

Besides ensuring that the point forecasts and loss functions match, we still need 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, we want the forecasts to have lower relative errors.

Neither MSE nor MAE have the above two properties. To evaluate the accuracy of wind speed forecasts, more realistic loss functions are needed. Hering & 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 $\alpha$-quantile as its optimal forecast (Gneiting, 2011b). This loss function puts a penalty on underestimates with weight $\alpha$, which depends on market rules. Hering & 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 $\alpha$. Errors on small true wind speeds are only partly more penalized through the power curve transformation; see Figure 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 by Chen & 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 (8) via simulations from $F$.

Figure 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 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

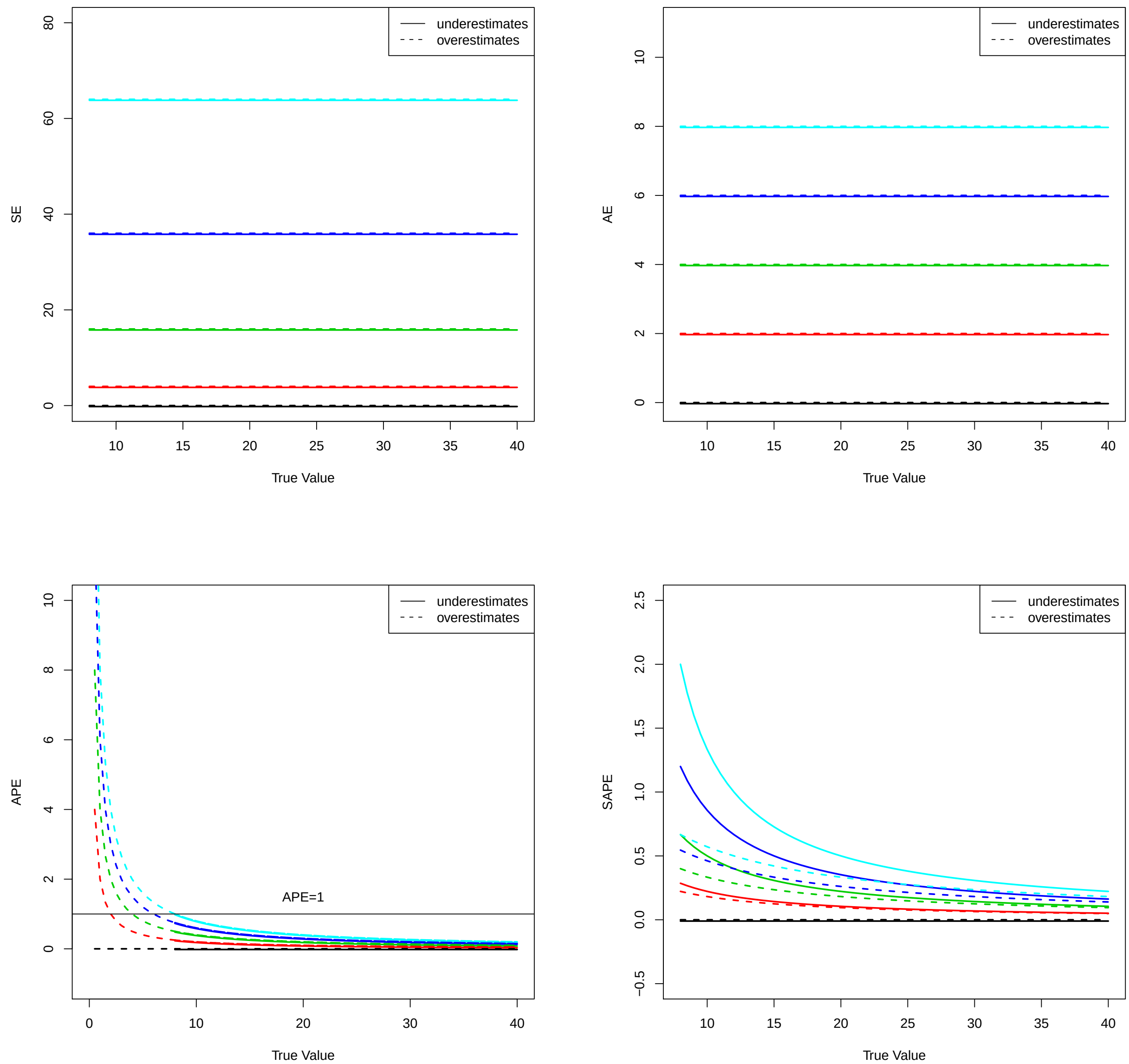


Figure 6: Squared error (SE), absolute error (AE), absolute percentage error (APE) and symmetric absolute percentage error (SAPE) for forecasts of true values $y = 8, 8.5, 9, 9.5, \ldots, 40$ with prediction errors 0 (black), 2 (red), 4 (green), 6 (blue), 8 (cyan) for each. Overestimates for extra true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

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 & Hagedorn (2011).

## 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, $L_k^{\mathrm{ref}}$ 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 ANL/DIS (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 & Rogoff (1988) tested for equal MSEs. Diebold & 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 & Ronchetti (2004). Tests for the evaluation of point, interval, probability, and density forecasts were generalized by Giacomini & White (2006). Recently, Hering & Genton (2011) extended the test to compare predictions in spatial fields.

## 5.4 Uncertainty of Forecasts

As explained in Section 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. To evaluate the quality of probabilistic forecasts, Gneiting et al. (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 Monterio et al. (2009) and Gneiting (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 article, we focused on building probabilistic forecasting models for the uncertainty in wind based on wind records instead of ensembles. In Section 6, another difficult aspect of forecasting the uncertainty in wind due to ramp events is discussed.

# 6 Discussion

## 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.

## 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 & Madsen (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.

## 6.3 Final Remarks

This article 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 article 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. We reviewed some classic time series statistical forecasting models, including Kalman filters and ARMA models.

Recently, more advanced forecasting models taking into account spatial correlations by considering other neighboring sites have been developed. We presented the RSTD and TDD models in detail. 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.

# Acknowledgments

This research was partially supported by NSF grant DMS-1007504 and by Award No. KUS-C1-016-04 made by King Abdullah University of Science and Technology (KAUST). The authors thank the editor, an associate editor, four reviewers, Amanda Hering and Le Xie for valuable comments that have improved this article.

# References


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

ANL/DIS (2009). Wind power forecasting: state-of-the-art 2009. (Available at: http://www.dis.anl.gov/pubs/65613.pdf).

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

AWEA. (2009). U.S. Wind industry annual market report year ending 2009. (Available at: http://www.awea.org/)

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

Bofinger, S., Luig, A. & Beyer, H.G. (2002). Qualification of wind power forecasts. *Poster P-GWP093 on the Global Windpower Conference and Exhibition*, Paris, France, 2-5.

Bossanyi, E.A. (1985). Short-term wind prediction using Kalman filters. *Wind Engineering*, **9**, 1-8.

Bossavy, A., Girard, R. & Kariniotakis, G. (2010). Forecasting uncertainty related to ramps of wind power production.
(available at http://www.cep.cma.fr/ st/rg/page4/files/ForecastingRampUncertainty.pdf)

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

Cadenas, E. & Rivera, W. (2009). Short term wind speed forecasting in La Venta, Oaxaca, Mexico, using artificial neural networks. *Renewable Energy*, **34**, 274-278.

Carta, J.A. & Ramírez, P. (2007). Use of finite mixture distribution models in the analysis of wind energy in the Canarian Archipelago. *Energy Conversion and Management*, **48**, 281-291.

Carta, J.A., Ramírez, P. & Velazquez, S. (2009). 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.

Carta, J.A., Ramírez, P. & Velazquez, S. (2009). Influence of the level of fit a density probability function to wind-speed data on the WECS mean power output estimation. *Energy Conversion and Management*, **49**, 2647-2655.

Chen, Z. & Yang, Y. (2004). Assessing forecasting accuracy measures. Preprint Series, Department of Economics, Iowa State University.

CREIA (2010). 2010 China wind power outlook.
(Available at: http://www.greenpeace.org/raw/content/eastasia/press/reports/wind-power-report-english-2010.pdf)

Crochet, P. (2004). Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland. *Meteor. Appl.*, **11**, 173-187.

Damousis, I.G., Alexiadis, M.C., Theocharis, J.B. & Dokopoulos, P. (2004). A fuzzy model for wind speed prediction and power generation in wind farms using spatial correlation. *IEEE Transactions on Energy Conversion*, **19**, 352-361.

Damousis, I.G. & Dokopoulos, P. (2001). A fuzzy model expert system for the forecasting of wind speed and power generation in wind farms. *Proceedings of the IEEE International Conference on Power Industry Computer Applications PICA*, **01**, 63-69.

Dell'Aquila, R. & Ronchetti, E. (2004). Robust tests of predictive accuracy. *Metron*, **62**, 161-184.

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

Diebold, F.X. & Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, **13**, 253-263.

DOE (2010). Large-scale offshore wind power in the United States.
(available at: http://www.nrel.gov/wind/pdfs/40745.pdf).

European Union (2008). Climate change: Commission welcomes final adoption of Europe's climate and energy package. Press Release, EU, Dec. 17, 2008.

(Available at: http://europa.eu/rapid/pressReleasesAction.do?reference=IP/08/1998.)

Fildes, R., Nikolopoulos, K., Crone, S.F. & Syntetos, A.A. (2008). Forecasting and operational research: A Review. *Journal of the Operational Research Society*, **59**, 1150-1172.

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

Fukuda, H., Tamaki, S., Nakamura, M., Nagai, H., Shijo, F., Asato, S. & Onaga, K. (2001). The development of wind velocity prediction method based on a data-mining type autoregressive model. *Proceedings of the European Wind Energy Conference*, Copenhagen, Denmark, 741-744.

Geerts, H. (1984). Short range prediction of wind speeds: a system-theoretic approach. *Proceedings of European wind energy conference*, Hamburg, Germany, 594-599.

Genton, M.G. & Hering, A.S. (2007). Blowing in the Wind. *Significance*, **4**, 11-14.

Giacomini, R. & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, **74**, 1545-1578.

Giebel, G. (2001). On the benefits of distributed generation of wind energy in Europe. PhD thesis from the Carl von Ossietzky Universitä Oldenburg, Düsseldorf.

Giebel, G., Brownsword, R. & Kariniotakis, G. (2003). The state of the art in short term prediction of wind power: A literature overview. *ANEMOS Project*.

Gneiting, T. (2008). Editorial: Probabilistic forecasting. *Journal of the Royal Statistical Society Series A*, **171**, 319-321.

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

Gneiting, T. (2011b). Quantiles as optimal point forecasts. *International Journal of Forecasting*, **27**, 197-207.

Gneiting, T., Balabdaoui, F. & Raftery, A.E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society Series B*, **69**, 243-268.

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

Gneiting, T. & Raftery, A.E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, **102**, 359-378.

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

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

Hering, A.S. & Genton, M.G. (2011). Comparing spatial predictions. *Technometrics*, **53**, 414-425.

Jeon, J. & Taylor, J. (2011). Using conditional kernel density estimation for wind power density forecasting. *Journal of the American Statistical Association*, to appear.

Kamal, L. & Jafri, Y.Z. (1997). Time series models to simulate and forecast hourly average wind speed in Quetta. *Solar Energy*, **61**, 23-32.

Kamath, C. (2010). Understanding wind ramp events through analysis of historical data. (available at: https://computation.llnl.gov/casc/StarSapphire/pubs/LLNL-CONF-416432.pdf)

Lange, M. & Heinemann, D. (2002). Accuracy of short term wind power predictions depending on meteorological conditions. *Poster P-GWP091 on the Global Wind power Conference and Exhibition, Paris, France, 2-5 April 2002.*

Lange, M. & Waldl, H.P. (2001). Assessing the uncertainty of wind power predictions with regard to specific weather situations. *Proceedings of the European Wind Energy Conference, Copenhagen, Denmark, 2-6 June 2001*, 695-698. (Note: accessible by following the link provided from their university homepage.)

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

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

Luig, A., Bofinger, S. & Beyer, H.G., (2001). Analysis of confidence intervals for the prediction of regional wind power output. *Proceedings of the European Wind Energy Conference, Copenhagen, Denmark, 2-6 June 2001*, 725-728.

Malmberg, A., Holst, U. & Holst, J. (2005). Forecasting near-surface ocean winds with Kalman filter techniques. *Ocean Engineering*, **32**, 273-291.

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

Meese, R.A. & Rogoff, K. (1988). Was it Real? The exchange rate - interest differential relation over the modern floating-rate period. *Journal of Finance*, **43**, 933-948.

Monteiro et al. (2009). Wind power forecasting: state-of-the-art 2009. (Available at: http://anemos.cma.fr/download/ANEMOS_D1.1_StateOfTheArt_v1.1.pdf)

Nielsen, T.S., Joensen, A., Madsen, H., Landberg, L. & Giebel, G. (1998). A new reference model for wind power forecasting. *Wind Energy*, **1**, 29-34.

Pinson, P., Chevallier, C. & Kariniotakis, G.N. (2007). Trading wind generation from short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems*, **22**, 1148-1156.

Pinson, P. & Hagedorn, R. (2011). Verification of the ECMWF ensemble forecasts of wind speed against analyses and observations. *Meteorological Applications*, accepted

Pinson, P. & Madsen, H. (2011). Adaptive modelling and forecasting of offshore wind power fluctuations with Markov-switching autoregressive models. *Journal of Forecasting*, available online.

Ramírez P. & Carta J.A. (2005). 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.

Rugbjerg, M., Sorensen, O.R. & Jacobsen, V. (2006). Wave forecasting for offshore wind farms. $9^{th}$ *International Workshop on Wave Hindcasting and Forecasting, Victoria, B.C. Canada, September 24-29, 2006.*

Schlink, U. & Tetzlaff, G. (1998). Wind speed forecasting from 1 to 30 minutes. *Theoretical and Applied Climatology*, **60**, 191-198.

Sfetsos, A. (2000). A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renewable Energy*, **21**, 23-35.

Sfetsos, A. (2002). A novel approach for the forecasting of mean hourly wind speed time series. *Renewable Energy*, **27**, 163-174.

Sloughter, J.M., Gneiting, T. & Raftery, A.E. (2010). Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. *Journal of the American Statistical Association*, **105**, 25-35.

Tambke, J. (2004). Forecasting offshore wind speeds above the North Sea. *Wind Energy*, **8**, 3-6.

Tambke, J., Lange, M., Focken, U. & Heineman, D. (2003). Previento meets horns rev short-term wind power prediction-adaptation to offshore sites. *Proceedings of the European Wind Energy Conference EWEC in Madrid, Spain June 2003*.

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

Tsay, R.S. (2010). Analysis of financial time series (Third edition). *Wiley*.

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

Wikle, C.K., & Cressie, N. (1999). A dimension-reduced approach to space-time Kalman filtering. *Biometrika*, **86**, 815-829.

WWEA (2009). World wind energy report 2009. (Available at: http://www.wwindea.org/home/images/stories/worldwindenergyreport2009_s.pdf).

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