IMPROVEMENT OF WIND FORECASTING ACCURACY AND ITS IMPACTS ON BIDDING STRATEGY OPTIMIZATION FOR WIND GENERATION COMPANIES

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Gong Li

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Industrial and Manufacturing Engineering

April 2012

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Improvement of Wind Forecasting Accuracy and its Impacts on Bidding Strategy Optimization for Wind Generation Companies

By

Gong Li

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Shi, Jing
Chair

Billen-green, Canan

Yadav, Om Prakash

Yuvarajan, Subbaraya

Gou, Bei

Approved:

April 17, 2013
Date

Billen-green, Canan
Department Chair

ABSTRACT

One major issue of wind generation is its intermittence and uncertainty due to the highly volatile nature of wind resource, and it affects both the economy and the operation of the wind farms and the distribution networks. It is thus urgently needed to develop modeling methods for accurate and reliable forecasts on wind power generation. Meanwhile, along with the ongoing electricity market deregulation and liberalization, wind energy is expected to be directly auctioned in the wholesale market. This brings the wind generation companies another issue of particular importance, i.e., how to maximize the profits by optimizing the bids in the gradually deregulated electricity market based on the improved wind forecasts. As such, the main objective of this dissertation research is to investigate and develop reliable modeling methods for tackling the two issues.

To reach the objective, three main research tasks are identified and accomplished. Task 1 is about testing forecasting models for wind speed and power. After a thorough investigation into currently available forecasting methods, several representative models including autoregressive integrated moving average (ARIMA) and artificial neural networks (ANN) are developed for short-term wind forecasting. The forecasting performances are evaluated and compared in terms of mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). The results reveal that no single model can outperform others universally. This indicates the need of generating a single robust and reliable forecast by applying a post-processing method. As such, a reliable and adaptive model for short-term forecasting the wind power is developed via adaptive Bayesian model averaging algorithms in Task 2. Experiments are performed for both long-term wind assessment and short-term wind forecasting. The results show that the proposed BMA-based model can always provide adaptive, reliable, and

comparatively accurate forecast results in terms of MAE, RMSE, and MAPE. It also provides a unified approach to tackle the challenging model selection issue in wind forecasting applications. Task 3 is about developing a modeling method for optimizing the wind power bidding process in the deregulated electricity wholesale market. The optimal bids on wind power must take into account the uncertainty in wind forecasts and wind power generation. This research investigates the application of combining improved wind forecasts with agent-based models to optimize the bid and maximize the net earnings. The WSCC 9-bus 3-machine power system network and the IEEE 30-bus 9-GenCo power system network are adopted. Both single-sided and double-sided auctions are considered. The results demonstrate that improving wind forecasting accuracy helps increase the net earnings of wind generation companies, and that the implementation of agent learning algorithms further improves the earnings. The results also verify that agent-based simulation is a viable modeling tool for providing realistic insights about the complex interactions among different market participants and various market factors.

ACKNOWLEGEMENTS

This dissertation is impossible to be finished without the help and support from people and organizations that are gratefully acknowledged here.

First and foremost, my greatest debt belongs to my dedicated advisor, Dr. Jing Shi, who has offered me valuable topics and inspiring suggestions. I cannot imagine how I could have reached this stage without his professional support in seeking research topics, tackling experiment problems, analyzing the results, revising paper drafts, etc. His creative ideas, strict research philosophies, efficient instructions, and patient explanations are greatly appreciated and worthy of emulation, which will continue benefiting me in my career throughout my life. He has also spared countless time and provided valuable suggestions on life challenges. His positive life philosophy and active life attitude have impressed and influenced me greatly.

Meanwhile, I would like to acknowledge my doctoral committee members, Dr. Canan Bilen-Green, Dr. Om Prakash Yadav, Dr. Subbaraya Yuvarajan, and Dr. Bei Gou. I appreciate their willingness to share their time and expertise, their professional insights, and their support throughout this process. Besides, I want to thank Dr. Terry Pilling for his professional suggestions on research topics and help with my data collection.

I would also like to acknowledge the role of Dr. John R. Cook, Dr. Jun Zhang, Dr. Kambiz Farahmand, and the rest of the Industrial and Manufacturing Engineering (IME) faculty and staff. Their professional support and kind help have greatly benefited me during the past years as well.

I must thank Junyi Zhou, Yonghua Yang, Jun Yang, Guojie Wang, Jianfei Wu, Yidong Peng, and many others friends, for their help, support, and friendship. Their invaluable advice, as well as countless happy time we shared together, will be cherished forever.

Lastly, I would like to acknowledge my family members for their love and support. I especially thank my beloved wife, Ms. Xiaoyan Yin, and our adorable son, Kunshu Li. Ms. Yin's patience, understanding, endless support and encouragement during this educational journey have motivated me to reach my goals. The relaxed presence and happy time that Kunshu brings to my family helps me not to take life too seriously. I also want to thank my parents, parents-in-law, brothers, sisters, and all the rest family members for their never-ending support and motivation throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

With the deterioration of the environment and the gradual depletion of conventional energy sources, wind energy has become the world's fastest growing source of renewable energy. Wind-powered generation, or shortly wind power, is one typical way to utilize the energy carried in wind. The world has seen an average annual growth rate of 27.6% in terms of installed wind power capacity from 2006 and 2010 (Global Wind Energy Association, 2011). Correspondingly, the penetration level of wind power into electricity market is also rising rapidly. The penetration is expected to reach 8% by the year of 2018 globally (Renewableenergyworld.com, 2009). Especially, the U.S. Department of Energy (2008) sets an ambitious goal that wind energy contributes to 20% U.S. electricity consumption by 2030.

The major part of current wind power generation is still sold with a long term power purchase agreement (PPA). Nevertheless, along with the fast development of wind power and its increasing penetration into the electricity markets, wind generation companies (WGenCos) are encouraged to directly participate into the wholesale electricity market by presenting supply offers and committing the delivery of the settled amount of wind energy at given or agreed moments. Meanwhile, the electricity markets are undergoing reconstruction or liberalization, which aims to develop a deregulated but secure environment to enable fair competition, unbundling electricity services, and open access to the network (Liu and Wu, 2006). For example, one commonly adopted market structure is the day-ahead power exchange, in which an independent system operator (ISO) matches the supply offers and the demand bids, both in the form of simple quantity-price blocks for a given period. In such electricity markets, the sellers mainly compete by submitting supply offers and guaranteeing the delivery of the settled amount of energy at each specified moment of the given period. Especially, only limited information about other participants and various market uncertainties such as load variations, competitor's behavior, and power system contingencies (Hu, Grozev, and Batten, 2005) is available for the sellers to make their decisions. Therefore, it is of great interest and importance for the generation companies (GenCos) to strategically optimize their supply offers so as to maximize their benefits by participating in the deregulated electricity markets (Krause, Beck, Cherkaoui, Germond, Andersson, and Ernst, 2006).

1.1. Motivation

The bidding strategy optimization issue is especially critical for WGenCos. Compared with the traditional GenCos, WGenCos participating in the electricity market have to deal with the large variations and intermittence of the distributed wind power production in addition to load swings and possible outages of production capacity (Georgilakis, 2008). Participation into deregulated electricity markets implies that the WGenCos need to present offers and commit the delivery of the settled amount of wind power at an agreed moment. Reliable delivery of electrical energy to load centers entails a continuous process of scheduling and adjusting electricity generation in response to the constantly changing demand. If the actual delivered energy by one generator is larger or less than the committed, the generator will be charged with a cost for other generators' rescheduling their generations to maintain the balance between the generation and the load. However, the intermittent and stochastic nature of wind makes it critical and yet very difficult for the WGenCos to accurately forecast their wind-powered generation in the following weeks, days, hours, and even minutes (Usaola and Angarita, 2007). It is estimated that the cost impact of wind's variability can reach about 10% of the wholesale value of the wind energy (Demeo, Grant, Milligan, and Schuerger, 2005).

In short, one major challenge of wind power is its intermittence and uncertainty due to the highly volatile nature of wind resource, and this affects both the economy and the operation of the WGenCos and the distribution networks. Novel modeling methods are thus urgently needed for generating accurate and reliable forecasts of wind power. Another issue of particular importance is how to maximize WGenCos' net earnings by optimizing their bids in the gradually deregulated electricity market. Accordingly, research is also urgently needed to develop methods for optimizing the WGenCo's bidding strategy in the deregulated wholesale market. However, the existing research efforts towards the above-mentioned challenges are still far from sufficient. This is the direct motivation of the dissertation research.

1.2. Objectives, Tasks and Framework

1.2.1. Research objectives

Many techniques and models have been developed for short-term wind forecasts (Ma, Luan, Jiang, Liu, and Zhang, 2009), but their performances vary with model settings, application scenarios, and evaluation metrics. The industry usually prefers a final single forecast that is always better than or at least close to the individual best forecasts from various models for all evaluation metrics. Meanwhile, the studies on bidding wind power in the electricity market have just started recently, and thus the relevant literature is very scarce. In order to maximize the revenue from selling wind power in the wholesale market, the WGenCos must take into account the uncertainty in wind generation in addition to other constraints in preparing their bidding strategies. As such, the objectives of this dissertation research are (1) building a flexible and reliable forecasting model to enhance the wind forecasting accuracy, and then, (2) investigating and developing methods for optimizing the WGenCos bidding in the electricity wholesale market by integrating the uncertainty of wind forecasts.

1.2.2. Research tasks and significances

To reach the above-mentioned objectives, three main tasks are identified and their significances are described as follows:

Task 1: *Building and implementing forecasting models for wind speed and power.* Forecasting plays a key role in improving energy markets efficiency, reducing the amount of reserves while maintaining system security. After a thorough investigation into currently available forecasting methods, several representative models are developed, implemented, and evaluated. This builds a solid base for the following research tasks; it also provides a good reference for other related research.

Task 2: *Developing adaptive and reliable methods for wind power forecasting*. Based on the findings obtained above, a reliable and adaptive model for forecasting the wind power under different forecasting horizons is developed via adaptive Bayesian model averaging.

Task 3: *Investigating and developing methods for optimizing the bidding process of wind power in the deregulated electricity wholesale market.* Nowadays, all the GenCos with wind power in their generation mix, bidding to the liberalized electricity market, need wind forecasts to support their bidding decision. Regulation power must be reserved for dealing with the fast load variations and unforeseen problems with production capacity. The regulation power needed will further increase with a higher penetration of wind generated electricity. Therefore, the optimal bids in the electricity market must take into account this uncertainty in order to obtain the maximum net earnings from selling wind energy. This research investigates the application of combining improved wind forecasts with agent-based modeling algorithms to optimize the bid and maximize the net earnings.

1.2.3. Research framework

Figure 1 shows the schematic framework for the thesis research. Based on the above-mentioned tasks, the research work can be divided into the following major phases.

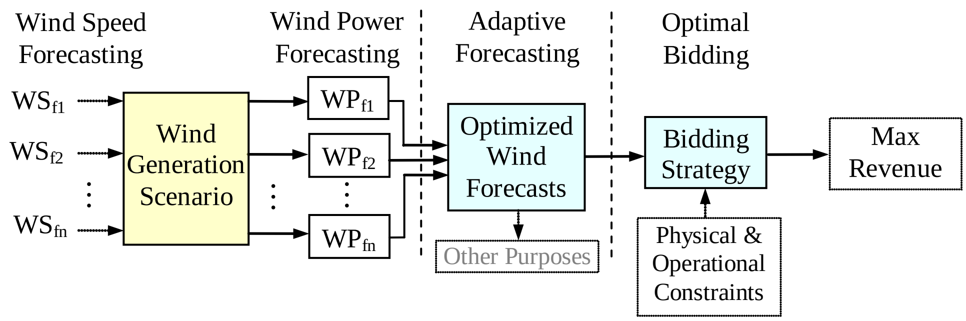


Figure 1. Research framework.

Firstly, a thorough investigation into currently available forecasting methods is performed. Based on this, several representative models, including autoregressive integrated moving average (ARIMA) and artificial neural networks (ANN), are developed, implemented, evaluated, and compared. The results reveal that no single model can outperform others universally in terms of multiple performance evaluation metrics. In other words, the best one among the models studied will be different based on different performance metrics, even for the same wind dataset. Moreover, the selection of the best model is also affected by, or dependent upon, the data sources. These findings indicate the need of generating a single robust and reliable forecast by applying a post-processing method.

Secondly, after an investigation into the literature, the Bayesian model averaging algorithm, which is an emerging approach in other research areas, is creatively introduced to the field of wind forecasting. Experiments are performed for both long-term and short-term wind forecasting. The results show that the proposed BMA-based model can always provide adaptive, reliable and comparatively accurate forecast results. It provides a unified approach to tackle the challenging model selection issue in wind forecasting applications. The BMA model is also characterized by its ability to generate probabilistic forecasts, which can correspondingly provide more information than those point estimation models.

Thirdly, agent-based simulation approach is selected, for the first time, to investigate the bidding optimization of a wind GenCo in the deregulated day-ahead electricity wholesale markets, by considering the effect of short-term forecasting accuracies of wind power generation. The results clearly demonstrate that improving wind forecasting accuracy helps increase the net earnings of WGenCos and that agent-based simulation is a viable modeling tool for providing realistic insights about the complex interactions among different market participants and various market factors.

1.3. Dissertation Outline

The remainder of this dissertation is structured as follows. In Chapter 2 an extensive review is conducted on the existing literature related to this dissertation research, which includes the topics of long-term wind prediction, short-term wind forecasting, and bidding strategy analysis. In Chapter 3, several ARIMA and ANN models are developed and implemented for short-term wind forecasting, mainly based on the philosophy of point estimation. The forecasting performances are evaluated and compared in terms of root mean square errors (RMSE), mean absolute errors (MAE), and mean absolute percentage errors (MAPE), respectively. In Chapter 4, the adaptive BMA algorithm is firstly tried on long-term wind potential estimation, and then employed to perform short-term wind forecasting. In Chapter 5, the agent-based modeling method is employed to analyze the issues of WGenCos bidding strategy optimization in the deregulated day-ahead electricity wholesale markets, especially by considering the probabilistic wind forecasting errors. The effects of improving wind forecasting accuracy and the utilization of agent-based simulation are analyzed. Finally, conclusive remarks are drawn and several possible directions for further research in the near future are pointed out in Chapter 6.

2. LITERATURE REVIEW

This chapter presents an extensive review on the existing literature related to the proposed research. For the issue of wind forecasting, various methods and tools have been investigated and applied, which are already reviewed in several publications (Wu and Hong, 2007; Ma, Luan, Jiang, Liu, and Zhang, 2009; Soman, Zareipour, Malik, and Mandal, 2010). Therefore, this chapter first provides a brief summary of the above-mentioned several review papers and presents a concise review on the typical short-term wind forecasting techniques in Section 2.1. The methods for estimating long-term wind distribution are then reviewed in Section 2.2. After that, the fundamental knowledge about electricity markets and their deregulations is introduced in Section 2.3. For the issue of bidding strategy optimization, many methods, models, and tools have also been introduced in the literature, but no comprehensive review on GenCos bidding strategy analysis is available. In view of this, a thorough review on literature related to this issue is finally provided in Section 2.4.

2.1. Short-Term Wind Forecasting

Short-term wind forecasting can provide useful information for the operations of wind generation systems (Monfared Rastegar, and Kojabadi, 2009; Pinson, Chevallier, and Kariniotakis, 2007) and the integration of wind power with the power grids (Ma et al., 2009). For instance, short-term forecasts can help with the daily or intraday wholesale market, system management and maintenance scheduling, which are usually of great importance to system operators, electricity companies, and wind farm promoters (Sfetsos et al., 2000; Costa et al., 2008). Accurate short-term forecasts of wind power are vital for the efficiency of wind power generation systems (Monfared, Rastegar, and Kojabadi, 2009) as well as for the integration of wind energy into the power systems. As aforementioned, wind power generation has been growing at an unprecedented rate in recent years, and this exponential increase is expected to continue in the following decade. However, a specific challenging problem in wind power generation is its intermittence and uncertainty due to the highly volatile wind resource. It is estimated that the cost impact of this variability is around 10% of the wholesale value of the wind energy (Demeo et al., 2005). The intermittence and uncertainty of the wind power production affect the operations and management of both wind farms and power system (Georgilakis et al., 2008; Durán et al., 2007; Lerner et al., 2009) and the influence exists on all time scales (Giebel et al., 2003).

In the past decades, research efforts have been made to develop sound short-term forecasting methods. In this regard, Giebel et al. (2003) and Costa et al. (2008) provide a fairly good summary of the developments pertaining to short-term wind power forecasting methods and techniques, including physical models, conventional statistical models, hybrid physical-statistical models, the artificial intelligence based models, and others. Statistical models describe the problem mathematically based on the random time series of historical data by pattern identification, parameter estimation, and model checking methods. Box-Jenkins models (Box et al., 1994), often called auto regressive integrated moving average (ARIMA) models, are widely used in this area because of the model robustness and explicit model expression. Artificial neural network (ANN) technology is also widely used. ANNs can learn from past data, recognize hidden patterns or relationships in historical observations and use them to forecast future values (More et al., 2003). The ANN models trained with time series have the ability to model arbitrarily linear and nonlinear functions. Support vector machine (SVM) is essentially a kernel-based learning algorithm for solving nonlinear classification and regression problems, and has been applied in a wide variety of areas. The SVM model for wind forecasting demonstrates better performance than multilayer perceptron (MLP) models (Ji et al., 2007; Mohandes et al., 2004).

Wu and Hong (2007) explain the characteristics of wind speed time series, i.e., high degree of volatility and deviation firstly. The wind forecasting models published in 2006 and earlier are grouped into four types: persistence models, numeric weather prediction (NWP), statistical and ANN models, and hybrid models. This may not be a very good classification nowadays, but it does provide a concise overview of those earlier wind forecasting publications. Besides, the difference between wind speed forecasting and wind power forecasting is pointed out, and the non-linear complex relationship between wind speed and wind turbine generation is emphasized.

Ma et al. (2009) perform a bibliographical survey on the general background of research and developments in the fields of wind forecasting. Various wind forecasting methods are divided into two broad groups: the physical methods, which usually employ many physical considerations to reach a good forecasting precision, and the statistical methods, which focus on revealing and utilizing the hidden relationship of the real-time wind power data. It is noted that in typical wind forecasting process, both physical and statistical models are utilized simultaneously. In view of this, the existing forecasting models are further classified into four categories: the physical model, the conventional statistical model, the spatial correlation model, and the artificial intelligence and other new methods. It is concluded that further study on artificial intelligence methods and their training algorithms should be one promising research topic, that combining different forecasting models should be a good way to achieve satisfactory results in both long-term wind prediction and short-term wind forecasting.

Soman et al. (2010) provide a simple classification based on the time scale of forecasting horizon, as shown in Table 1. Note that this classification is somehow vague. For example, the "mid-term" in the table can be regarded as short-term in this study and other studies, whereas the "long-term wind" prediction can refer to several months or even several years, as interpreted in this study as well as other studies.

Table 1. Classification of wind forecasting based on horizons.

Type	Horizon	Main applications
Long-term	One day to one week	Unit commitment decisions Reserve requirement decisions Maintenance scheduling to obtain optimal operating cost
Mid-term	Six hours to one day	Generator online/offline decisions Operational security in day-ahead electricity market
Short-term	Half an hour to six hours	Economic load dispatch planning Load-changing decisions
Very short-term	Several seconds to half an hour	Electricity market clearing; Regulation actions

The dilemma in choosing the forecasting model for best performance is that no particular models are universally superior to other models for all types of applications and under all conditions (Burnham et al., 2002; Li and Shi, 2010a). The performances of the short-term forecasts from various models can vary with model settings, application scenarios, and performance metrics. Also, it is not unusual that more than one model might provide plausible forecasts. Thus, a final single forecast is often expected which should take advantage of all plausible forecasts. It will be ideal that this final single forecast is better than the individual ones, or at least always close to the best available forecast (Sanchez, 2008). In this regard, an adaptive combination procedure is necessary for generating such an efficient single forecast.

2.2. Long-Term Wind Speed Distribution

Wind speed is the most important parameter to be considered in the design and operations of wind power systems. Especially, its probability density distribution can provide abundant information for and affect the operational decisions and performance of the wind energy systems. Therefore, a large number of studies have been published on modeling the wind speed frequency distributions with probability density functions (PDFs). There are a few parametric distributions that work well for this purpose, such as lognormal (Luna and Church, 1974), inverse Gaussian (Bardsley, 1980), Weibull and Rayleigh (Van der Auwera, Meyer, and Malet, 1980; Lun and Lam, 2000), and generalized extreme value (Bauer, 1996). Besides these univariate models, some multivariate distributions are also evaluated while considering both wind speed and wind direction such as anisotropic distribution and angular-linear distribution (Weber, 1997; Carta, Ramírez, and Bueno, 2008) Recently, the concept of maximum entropy principle (MEP) was also introduced to derive the PDF of wind speed distribution (Li and Li, 2005; Akpinar and Akpinar, 2007). A detailed review on modeling the probability distributions in wind energy analysis can be found in literature (Carta, Ramírez, and Velázquez, 2009).

Although extensive efforts have been made on modeling wind speed distributions, the conventional statistical approach faces challenges. Typically, when a candidate PDF is proposed, one needs to know whether the proposed distribution is viable in describing the wind speed data. This is often performed by either comparing with the popular Weibull distribution and/or evaluating the goodness-of-fit by some statistic metrics. The reality is that Weibull distribution might not be the most appropriate model for benchmark. Also, heavily relying on particular evaluation metrics could make the results biased and misleading since different goodness-of-fit statistics may generate different preference ranks among the PDF models tested. Moreover, these approaches mainly focus on the uncertainty in the parameters of a specific model by assuming a fixed model structure, and a “best” model is usually determined based on some model selection criteria with the uncertainty between the models being neglected (Burnham and Anderson, 2002).The uncertainty between models may be important in making inference especially in the cases where more than one candidate distribution is considered plausible while those candidate distributions differ in predictions. These problematic situations reflect the urgent need for new methodologies.

Recently, Bayesian model averaging (BMA) has gained popularity in various fields, such as management science, medicine, and meteorology, because it can produce more accurate and reliable predictions than other techniques (Hoeting, Madigan, Raftery and Volinsky, 1999; Viallefont, Raftery, and Richardson, 2001). The output of BMA is a weighted average of the model PDFs centered on the bias-corrected predictions, where the weights reflect the relative contributions of the component models to the ensemble over the sample data. The variance of the BMA PDF contains two components, the within-model error variance and the between-model variance, both estimated from the sample data. BMA is originally developed as an approach of combining inferences and predictions from multiple statistical models, and applied to statistical linear regression and related models (Kass and Raftery, 1995). Raftery, Gneiting, Balabdaoui, and Polakowski (2005) apply BMA as a statistical post-processing method to generate probabilistic forecasts in the form of PDF, and this yields calibrated and sharp predictive PDFs of the surface temperature and sea level pressure whose true distributions are approximately normal. Sloughter, Raftery, Gneiting, and Fraley (2007) and Duan, Ajami, Gao, and Sorooshian (2007), respectively, modify this method and apply it to forecast the quantitative precipitation which has a skewed distribution. These efforts indicate that BMA can provide a more reliable description of the total predictive uncertainty than the original element, leading to a sharper and better calibrated PDF for the probabilistic predictions. As such, it is appealing to apply BMA method to derive the predictive model of long-term wind speed distribution.

2.3. Deregulated Electricity Wholesale Markets

In a traditional monopolistic or vertically integrated electricity market, power providers mainly aim to minimize the expected costs while maintaining an adequate security of supply (Liu, Jiang, and Yan, 2008). Since 1980s, however, the electricity markets have been gradually evolving towards liberalized or deregulated structures, which are characterized by open competitive energy markets, unbundling electricity services, open access to the network, etc. To establish a competitive electricity market and improve its efficiency, the restructured market allows for gaming on the market power and tends to stimulate the emergence of new technologies (Liu and Wu, 2006). The electricity wholesale market can be regarded as a market where the electricity can be sold or purchased at varying prices and delivered either immediately or at a given moment. Participants have to make decisions independently under complicated situations with insufficient information about their rivals and various uncertainties in the market such as load variations, competitor's behavior, and power system contingencies.

The deregulated electricity market behaves more like an imperfect competition or oligopoly market due to the special characteristics of the actual electricity market, such as a limited number of suppliers, long construction periods of power plants, large capital investment sizes, transmission constraints, and transmission losses (Wen and David, 2001). Typically, only a few dominating generation companies (GenCos) serve a given geographic region. In such an oligopolistic market, an individual GenCo has its market power, that is, it can affect and manipulate market price via its strategic bidding behavior (David and Wen, 2000). This indicates an opportunity for the GenCos to increase their profits through strategic bidding. Therefore, it is possible for the GenCos to maximize their profits by optimizing the bidding strategy in the deregulated electricity wholesale market while minimizing the associated risks. Also, the electricity market is a complex dynamic system with complicated interactions among physical structures, market rules and participants. As such, each participating agent faces risks and uncertainties while pursuing profit maximization (Vahidinasab, Jadid, 2010; Hu, Grozev, and Batten, 2005).

Meanwhile, renewable energies such as wind, solar, and biomass are regarded as a key factor in tackling global climate change and energy shortage crisis (Guler, 2009). For example, wind energy has globally experienced fast growth during the past decade (Li and Shi, 2010c; Saidur, Islam, Rahim, and Solangi, 2010). Along with the ongoing worldwide utilization of wind power and its increasing penetration into electricity markets, more wind GenCos (WGenCos) are encouraged and expected to participate in the electricity market by presenting offers and committing the delivery of the agreed amount of wind power at a given moment (Giabardo, Zugno, Pinson, and Madsen, 2010). However, regulation power must be reserved to deal with the possible fast load variations and unforeseen problems with production capacity (Amjady and Keynia, 2009). With a high penetration of such intermittent generations, the regulation power needed will further increase in that such generations are characterized by large variations in addition to the load swings and outages of production capacity (Cormack, Hollis, Zareipour and Rosehart, 2010). When bidding in the electricity markets, wind GenCos must pay for energy production deviations resulted from the prediction error, which can be as much as 10% of the total generator energy incomes (Fabbri, Gómez, Rivier, and Méndez, 2005). The optimal bidding strategy, especially in the deregulated electricity markets with higher penetration of intermittent renewable energy, must consider the cost paid for the energy imbalance in relation with the above-mentioned uncertainties and constraints. Therefore, one critical problem faced by the WGenCos is how to optimize their supply offers in the electricity markets to maximize their net earnings according to not only the available information of the markets and participants but also the forecasts on wind energy production.

It should be noted that the intermittent nature of wind energy production with large variations could affect various aspects of the power system as well, such as the transmission and distribution grids (Ostergaard, 2003; Garcés, Conejo, García-Bertrand, and Romero, 2009; Burke and O'Malley, 2010; Baringo and Conejo, 2011), the ancillary services (Ostergaard, 2006), the system operation and development (Strbac, Shakoor, Black, Pudjianto, and Bopp, 2007; Klobasa, 2010), the generation technology mix (De Jonghe, Delarue, Belmans, and D'Haeseleer, 2011), and the costs (Dale, 2004; Denny, and O'Malley, 2007).

Haas and Auer (2006) emphasize six prerequisites for effective competition in reformed wholesale electricity markets: (1) separation of the grid from generation and supply; (2) wholesale price deregulation; (3) sufficient transmission capacity for a competitive market and non-discriminating grid access; (4) excess generation capacity from many competing generators; (5) an equilibrium relationship between short-term wholesale markets and long-term financial instruments for marketers to manage wholesale-market price volatility; and (6) an essentially hands-off government policy that encompasses reduced oversight and privatization.

A variety of electricity market reconstruction models have been proposed, which generally can be categorized into three types: (a) market pools (PoolCo), (b) bilateral contract (BC) markets, and (c) hybrid markets (HM) (Yucekaya, Valenzuela, and Dozier, 2009). PoolCos are popular among the approaches to organizing electricity trading. Essentially, a PoolCo is a more centralized marketplace where an Independent System Operator (ISO) clears the market according to the bids from both the sellers and the buyers, operates and manages the entire system, and maintains its reliability. In a PoolCo, different GenCos compete not for specific customers but for the right to supply energy to the grid. A bilateral contract model defines a flexible market where the participants can specify and negotiate the terms and conditions of trading agreements independent from the ISO. The ISO mainly ensures enough transmission capacity and security.

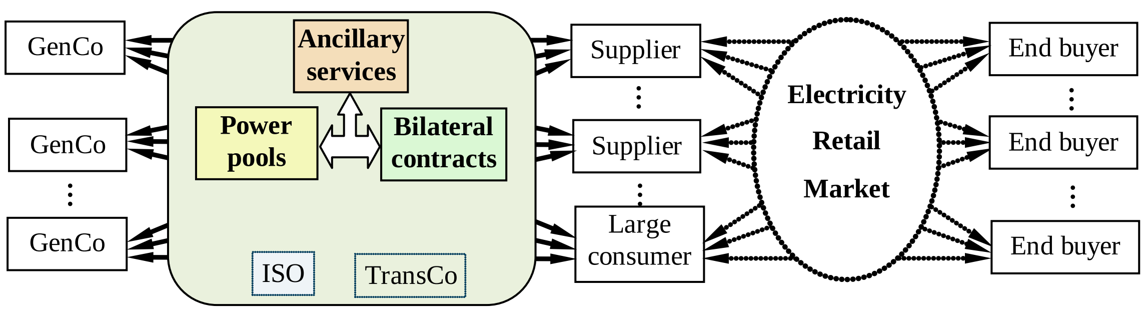


Figure 2. The general market structure of deregulated electricity markets.

Generally, the market structure of a fully deregulated hybrid electricity market can be illustrated by Figure 2. It may contain not only power pools and bilateral contracts but also ancillary services (AS) such as frequency and voltage controls, load following, energy imbalance, spinning reserve, and supplementary reserve and standby reserve (Foley, Gallachóiró, Hur, Baldick, and McKeogh, 2010). Competing generators offer their electricity output to retailers in a wholesale electricity market. The retailers then re-price the electricity and take it to the retail market. Although the wholesale pricing used to be the exclusive domain of large retailing suppliers, the markets are increasingly opening up to large end-buyers as well. The participants can directly sign bilateral contracts with others and/or bid in the PoolCos. Also, a fully deregulated hybrid market should allow the end buyers to choose among different competing suppliers in the electricity retail market. It should be mentioned that this study mainly focuses on the electricity wholesale market in view that the GenCos usually do not interact with the end consumers directly.

Participating in electricity market implies presenting bids and committing the delivery of the agreed amount of energy at a given moment. Reliable delivery of electrical energy to load centers entails a continuous process of scheduling and adjusting electricity generation in response to constantly changing demand (Hu, Grozev, and Batten, 2005). If the actual energy delivered by one generator is greater or smaller than what is committed, the GenCo will pay for additional cost of maintaining the balance between the generation and the load. In deregulated electricity markets, a GenCo is usually an entity owning generating facilities and participating in the market with the sole objective of maximizing its benefit (Kang, Kim, and Hur, 2007). Most evidently, individual bidding strategies are of essence to the interactions where the participants' actions affect others' possible outcomes.

Generally, a bid may include several energy price segments together with the corresponding quantity of electricity. As the most common structure, the pool-based market is an auction center in which all competitive participants are required to submit quantity-price pairwise bids that they commit to receive from or pay to the pool. As illustrated in Figure 3, once the bidding period ends, an ISO ranks and then matches the selling offers with buying bids so that the buying bids of the highest price are matched with the selling offers of the lowest price (Ott, 2003). When all the demands are met, the price of either the last accepted offer or the first rejected offer will be set as the market clearing price (MCP). The pricing mechanism for all the rest dispatched GenCos can be either uniform pricing (UP) or pay-as-bid (PAB). The UP auction indicates that all the winning suppliers are paid at an MCP. The PAB auction means that each winning supplier bidder is paid at its bidding price of the committed amount of electricity.

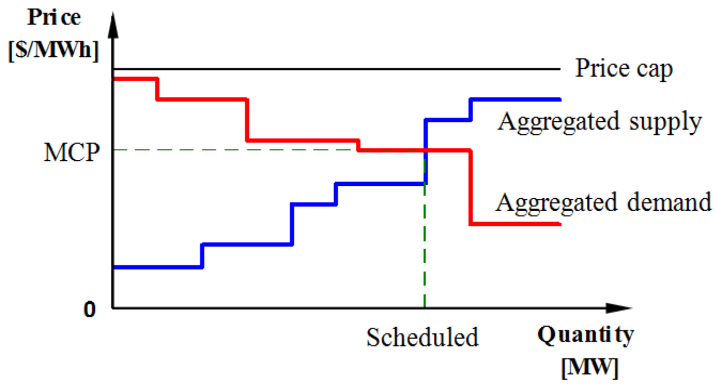


Figure 3. Market clearing mechanism.

Each different market structure has its own auction rules and bidding protocols. Various auction rules can be categorized into two types: static or dynamic. In static auctions, the bidders submit sealed bids, whereas dynamic auctions allow the bidders to observe others' bids so that they can revise their bids sequentially (David and Wen, 2000). The static auction may work based on either the UP or PAB rule. Bidding strategy in PAB-based market is more complex and potentially more important than that in UP-based market. In this case, the GenCos should estimate the uncertain MCP and bid slightly less than it. The PAB rule represents the future trend in the deregulated electricity markets and is expected to lower the market prices and reduce the price volatility (Xiong, Okuma, and Fujita, 2004). Although the majority of operating electricity markets currently still employs the sealed bid auction with the UP auction rule, extensive research has been carried out on the applications of PAB rule as well (Rahimiyan and Rajabimashhad, 2008). According to different market designs, the bidding protocols can be divided into two types: single-part bids or multipart bids (David and Wen, 2000). Under the single-part bidding protocol, as adopted in the California type power exchange (PX) market, the

GenCos bid independent prices for each hour; the winning bid and the schedules for each hour are determined via a simple market clearing process according to the intersection of supply and demand bid curves. This decentralized approach does not require the ISO to make unit commitment decisions. Instead, the GenCos have to consider all involved costs and constraints in preparing their bids. Therefore, whenever a significant physical or technical constraint occurs in a generation unit, a modification mechanism should be applied to the schedule, e.g., via short-term balancing market. In contrast, a multipart bid, as addressed in the England–Wales or British type electricity market, may include separate prices for ramps, start-up costs, shut-down costs, no-load operations, and energy. Although this type of bidding protocol can reflect the cost structure and the technical constraints of generation units, the non-convex Unit Commitment (UC) problem might not converge to a global optimal solution for large scale systems, possibly resulting in inequitable dispatches for different GenCos.

2.4. Bidding Electricity in Wholesale Markets

In 1988, Schweppe (1988) firstly noticed that some electricity utilities changed their price structures from rigid fixed prices to a wholesale price marketplace. Later, David (1993) formally addressed the strategic bidding issue for competitive power suppliers and developed a conceptual optimal bidding model and a dynamic programming method for England–Wales type electricity markets, in which each GenCo bids a fixed price for each generation block. From then on, the strategic bidding problem for competitive GenCos has attracted more attention, and various modeling approaches have thus been proposed to generate strategic bidding strategies (Song, Ni, Wenb, Hou, and Wu, 2003). Based on different modeling bases, various modeling methods for bidding strategy analysis in the electricity wholesale market can be divided into four general groups: (1) single GenCo optimization models, (2) game theory based models, (3) agent-based models, and (4) hybrid or other models, as illustrated in Figure 4.

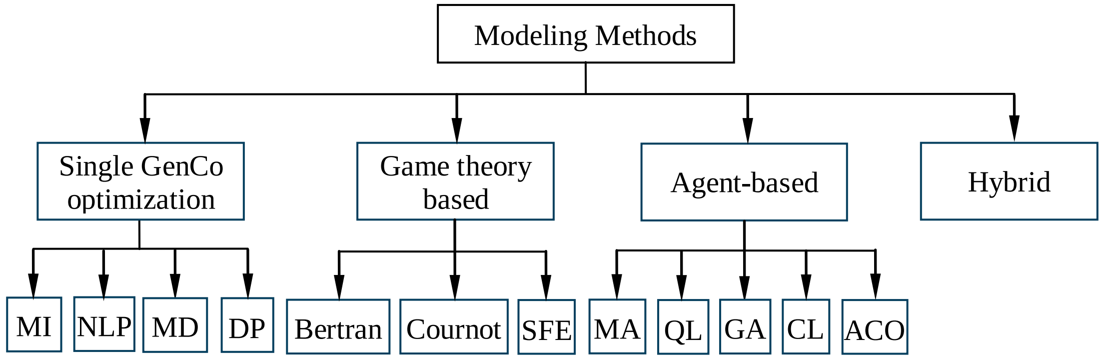


Figure 4. Modeling methods for bidding electricity in the wholesale market.

Each group of models may be further divided into small subgroups according to the model formulation and solution algorithms. For example, the single GenCo optimization models include many mathematical programming methods such as Mixed Integer Programming (MIP), Nonlinear Programming (NLP), and Dynamic Programming (DP); the game theory models might adopt different competition rules: Bertrand competition, Cournot competition, Supply Function Equilibrium (SFE), and some other newly proposed competition rules; the agent-based models can be categorized in terms of different learning algorithms such as model-based adaptation algorithms (MA), genetic algorithms (GA), Q-Learning (QL), computational learning (CL), Ant Colony Optimization (ACO), etc.

Generally, the single GenCo optimization models focus on only one specific player while simplifying other players and the influencing factors as a set of deterministic or stochastic independent variables, whereas game theory and agent-based approaches deal with the situation of more than one player in the market. Game theory equilibrium models investigate the bidding strategies from the perspective of players' mutual interactions. Agent-based models tend to mimic human behaviors and simulate optimal bidding strategies (Gao and Sheble 2010).

Table 2. Characteristics of three types of models.

Models	Characteristics
Single GenCo optimization	Developing optimization models to describe the entities in the electricity market with the objective of finding an optimal solution: • Well-established and solid mathematical foundation • Generally focusing on one specific player in the system by simplifying the rest of the system as a set of exogenous variables • Usually modeling no aspects of players' intelligent behaviors • Difficult to model the complex, uncertain and dynamic systems or analytically derive the optimal bidding strategy for the GenCos in the deregulated electricity markets
Game theory	Modeling the market as a game and mathematically capturing the players' behavior in the game where one player's success in making choices depends on the others' choices • Usually mathematically well-defined, involving a set of game players, a set of bidding strategies, and a specification of payoffs for each possible combination of bidding strategies • Analyzing the economic equilibria of the electricity market by focusing on the players' interactions • Capable of providing analytical rationale and explanation on how strategic bidding behaviors affect the GenCos' market power and profits • All players are assumed to be rational, which does not generally hold in the reality • The frustrating issue of multiple equilibria often occurs in solving the model
Agent-based	Modeling the complex electricity market as collections of rule-based agents interacting with one another dynamically and intelligently, simulating human beings' behavior to make optimal bidding strategies • Only a few simple rules are specified for and followed by various agents that situated in the network and behave intelligently in the system • Agents usually have and only require imperfect, local information and visibility • No centralized control or planning is required although random elements often exist either among variable agents or in the system • Agents can interact with each other directly or through the environment, resulting in complex emergent global behavior of dynamic-equilibrium and adaptation • More flexible, robust, and easily implemented compared with analytical approaches • Capable of capturing the details about agents behaviors, which is helpful in figuring out the relationships between individual decisions and system behavior • Capable of modeling the dynamics of systems that are not in equilibrium as well • The underlying mathematical foundation is still not well developed Requiring computation-intensive procedures

The main characteristics of the three modeling approaches are provided in Table 2. Recently, some hybrid and non-conventional methods have also been proposed. The representative publications in the four groups are reviewed and summarized in the following sections, respectively.

2.4.1. Single GenCo optimization models

In earlier publications, the issue of optimal bidding strategy selection is often addressed as a cost-minimization problem and solved via traditional cost-based UC algorithms (Borghetti, Frangioni, Lacalandra, Nucci, and Pelacchi, 2003). More recently, under the assumption that the MCP could be regarded as an exogenous variable (Valenzuela and Mazumdar, 2003), many mathematical programming approaches have been applied to address the problem of optimal bidding strategy selection. The majority of model formulations incorporate stochastic probabilistic elements, either in the problem data (e.g., the objective function and the constraints), or in the algorithm (through random parameter values, random choices, etc.), or in both (Spall, 2003). An insightful discussion on the application of stochastic programming methods to the energy market can be found in (Wallace and Fleten, 2003). The typical optimization methods adopted in the literature include Integer Linear Programming (ILP), Mixed Integer Programming (MIP), Multi-Objective Linear Programming (MOLP), Nonlinear Programming (NLP), Dynamic Programming (DP) (Foley et al., 2010), newsvendor, and Markov decision process (MDP) models. As mentioned before, the literature adopting these models typically optimizes the bidding strategy for a single market participant while ignoring or simplifying the behavior aspects of other players.

As reviewed by Conejo and Prieto (2001), many mathematical programming problems in a competitive electric energy framework can be modeled as mixed integer linear programming

(MILP) models. De la Torre, Arroyo, Conejo, and Contreras (2002) formulate an MILP model for a price-maker GenCo to solve the self-scheduling problem and maximize the profit in a pool-based electricity market. Conejo, Nogales, and Arroyo (2002) propose several mathematical programming models for a price-taking thermal GenCo to derive the optimal bidding strategy in a pool-based market with highly uncertain MCPs. The problem is first modeled as a stochastic mixed-integer linear programming (SILP) model and then transformed to two MILP models, one of which could be easily solved using a commercial ILP solver. Then a simple but informative bidding rule is derived from the solution. Similarly, from the perspective of a price-taking hydropower GenCo (HGenCo) participating in the PX of Nord Pool, a day-ahead power market, Fleten and Kristoffersen (2007) transform a two-stage SILP model to an MILP model for determining optimal bidding strategies by taking into account the discrete market price scenarios.

The effect of market price uncertainty on bidding optimization is explicitly explored by comparing the stochastic approach to the deterministic counterpart. Angarita, Usaola, and Martínez-Crespo (2009) present the application of stochastic optimization technique in maximizing the joint profit of hydro and wind generators in a pool-based electricity market. To handle the uncertainty of wind prediction, the hourly wind power is regarded as a discrete random variable in the optimization problem. Compared to other bid strategies that make use of the expected wind power value, the combined bidding strategy gives rise to significant improvements. Also, it is a useful tool for GenCos to avoid penalty costs or income reduction.

Sen, Yu, and Genc (2006) propose a multi-stage SILP model for scheduling and hedging in wholesale electricity markets. This SILP model captures stochastic electricity demand, electricity forward price, gas forward price, and wholesale price of electricity. Based on the structure of the SILP model, a nested column generation decomposition strategy is proposed to decompose the model into three interrelated sub-problems. The experimental results demonstrate that the proposed approach could provide robust decisions for scheduling and hedging problems. Besides, Ni, Luh, and Rourke (2004) develop a stochastic mixed-integer program (SMIP) to systematically handle the MCP uncertainties, bidding risk management, and self-scheduling requirements for a hydrothermal GenCo to maximize profits under a deregulated market. The model is solved by the proposed algorithm combining Lagrangian relaxation and stochastic dynamic programming (SDP) method.

De Ladurantaye, Gendreau, and Potvin (2007) introduce an SILP model to maximize the profits for a hydropower GenCo of multiple power plants along a river in a deregulated market. The proposed model aims to support the price-taking GenCo in its day-ahead bidding decisions. Morales, Conejo, and Perez-Ruiz (2010) address a multi-stage SILP problem which determines the best bidding strategy for a WGenCo in an electricity market including various trading floors. In the SILP problem, four uncertain sources are considered: wind power generation, day-ahead market price, adjustment market price and imbalance energy price. The multi-stage SILP problem is formulated as a linear programming (LP) model, and a case study for a WGenCo in Kansas is conducted for the illustration of solving the LP model.

Gross, Finlay, and Deltas (1999) develop a nonlinear programming (NLP) model to optimize strategic bids of a GenCo in a multi-period auction market under the assumption of perfect competition, and propose a Lagrangian relaxation (LR) method to solve the NLP model. Similarly, to deal with the GenCo’s bidding optimization and self-scheduling problem, Zhang, Wang, and Luh (2000) develop an NLP model by considering uncertain bidding information of other participants, and solve this model by an LR method.

Yucekaya, Valenzuela, and Dozier (2009) present two particle swarm optimization (PSO) algorithms to determine bid prices and quantities under the rules of the Pennsylvania, New Jersey, and Maryland market. The first one employs a conventional PSO technique whereas the second uses a decomposition technique in conjunction with the PSO approach. It is found that the latter algorithm can dramatically outperform the former. Also, it is shown that for nonlinear cost functions, PSO solutions provide higher expected profits than marginal cost-based bidding.

Similarly, in considering the non-convex operating cost functions of thermal generating units and minimum up/down time constraints, Boonchuay and Ongsakul (2011) propose an optimal risky bidding strategy for a GenCo by self-organizing hierarchical particle swarm optimization with time-varying acceleration coefficients. With rivals' behavior in competitive environment being simulated via the Monte Carlo method, the significant risk index based on mean–standard deviation ratio (MSR) is maximized to generate the optimal bid. The proposed approach is concluded to be capable of providing a higher MSR than other PSO methods.

Wen and David (2001a) propose a stochastic nonlinear programming (SNLP) model for deciding optimal bidding strategies for competitive power suppliers in a sealed bid auction based electricity market. The model assumes that the power supplier bids a linear supply function and is paid at the MCP with the system dispatch levels being stipulated by a market operator to minimize customers' payments. It is shown that the MCP can be significantly higher than the competitive levels if the suppliers bid strategically, and that the market power of the suppliers will be reduced if the load is elastic to the price of electricity. Similarly, they build more SNLP models and propose a genetic algorithm based method to build bidding strategies for power suppliers in the California-type day-ahead energy market in which power suppliers are required to simultaneously bid 24 linear energy supply functions, one for each hour, and the system dispatch levels are stipulated separately for each hour by utilizing the UP pricing rule. The method is believed to be especially suitable for those suppliers with marginal or near-marginal generating units (Wen and David, 2001b).

Ma, Wen, Ni, and Liu (2005) develop an SNLP model for optimizing bidding strategies by considering the risks for the GenCos participating in a pool-based single-buyer electricity market. Each GenCo is assumed to bid a linear supply function and the system is dispatched to minimize the total purchasing cost of the single-buyer. Each GenCo chooses the coefficients in the linear supply function for making tradeoff between two conflicting objectives: profit maximization and risk minimization.

Guan, Ho, and La (2001) develop a bidding strategy based on the theory of ordinal optimization. The basic idea is using an approximate model to describe the influence of bidding strategies on the MCP. A nominal bid curve is obtained by solving optimal power generation for a given set of MCPs via Lagrangian relaxation. The best bid is then selected by solving full hydrothermal scheduling or unit commitment problems.

Usaola, and Angarita (2007) formulate a stochastic linear programming model by including the probability density of wind forecasting and analyze the optimal bidding strategy for a WGenCo in a wholesale market. It is found that the most accurate prediction is achieved when bids are updated in intraday markets by using more recent predictions. However, the most accurate prediction cannot ensure the highest revenues due to the different prices of spilled and bought energy and the bias of the prediction programs. In order to generate maximum revenue, the uncertainty of the power prediction must be considered.

Zhang (2009) uses a dynamic random effect ordered probit model to analyze the GenCo’s bidding behavior in the NYISO day-ahead wholesale market. The results show that the generators in higher-priced groups tend to withhold their capacity strategically to push up market prices. It is also verified that demand conditions might greatly affect market prices. Rahimiyan and Rajabimashhadi (2007) formulate the bidding decision-making problem from a supplier's viewpoint in a PAB auction wholesale market by assuming a normally distributed MCP, and analyze the effect of some risk factors on the supplier-expected benefit and selling amount.

Song, Liu, Lawarrée, and Dahlgren (2000) propose a bidding decision-making strategy in which the impacts of production limit and market share on the optimal bidding strategies are considered. It is concluded that the Markov decision process (MDP) model is able to optimize the decision over a planning horizon. However, the model makes a few strong assumptions such as ignoring the operational constraints of power systems and giving no provision for incorporating risk attitude in an ordinary MDP.

Gajjar, Khaparde, Nagaraju, and Soman (2003) formulate the GenCo optimal bidding in a deregulated power market in the framework of MDP. An optimal strategy is devised to maximize the profit by employing the temporal difference technique and actor-critic learning algorithm. The method is concluded to be especially useful for long-term profit maximization under stochastic risks.

Bathurst, Bathurst, Weatherill, Weatherill, Strbac, and Strbac (2002) present a strategy for bidding a few hours before the operation time for the wind producers under the New Electricity Trading Arrangements (NETA), changed to British Electricity Trading Transmission Arrangements since 2005, by using Markov processes for simulating a wind farm. The method demonstrates substantial reductions in the imbalance costs as well as the effect of market closure delays and window lengths of wind forecasting.

Pinson, Chevalier, and Kariniotakis (2007) study daily bidding by using the rules of the Dutch APX electricity market, where bids are presented only once per day and not updated in intraday markets. It is found that as a result of reduced regulating market costs from better hourly predictions to the market, wind power producer could obtain up to 8% more profit if the time between market bids and delivery is shortened from the day-ahead Elwholesale market. They also formulate a general methodology for deriving optimal bidding strategies based on probabilistic forecasts of wind generation in the form of predictive distribution. This flexible methodology can be tailored based on the needs of a specific market participant (Pinson et al. 2007).

Besides, bi-level optimization is often applied either to represent the strategic interaction among suppliers or in hybrid markets where electrical energy and spinning reserve are simultaneously traded (Haghighat, Seifi, and Kian, 2007; Soleymani, Ranjbar, and Shirani, 2007) or in the presence of future contracts (Yuan, Liu, and Jiang, 2007) and bilateral contracts (Badri, Jadid, Rashidinejad, and Moghaddam, 2008). Also, the influence of extra objectives such as the minimization of supplier emission of pollutants (Vahidinasab and Jadid, 2009), or the influence of unit reliability (Soleymani, Ranjbar, and Shirani, 2008) has been analyzed using optimization models. The competition process can be represented as a dynamic feedback system as well (Liu and Wu, 2006). Attaviriyanupap, Kita, Tanaka, and Hasegawa (2005) propose an algorithm for determining the optimal bidding strategy for a GenCo in the deregulated day-ahead power and reserve markets. The optimal bidding parameters for both markets are determined by solving an optimization problem, which considers unit commitment constraints such as generating limits and unit minimum up/down time constraints. In the study, evolutionary programming (EM) technique is used to solve the problem.

2.4.2. Game theory models

Game theory models, also called equilibrium models, optimize the bidding strategies by investigating players' interactions and analyzing economic equilibria of the system. Typically, in a game, each player chooses the strategy from its own strategy set; then a payoff will be assigned to each player by the payoff function; as a result the optimal solution can be reached via Nash equilibrium. Nash equilibrium is a strategy combination of all players in which no player can increase its payoff by changing its own strategy alone so that every player will finally choose its strategy exactly as the equilibrium strategy combination. Game theory models provide analytical rationale and explanation on how market power can be exercised via strategic bidding behavior, but the assumption that all players are rational usually does not hold in practice (Gao and Sheble, 2010). Also, it is limited by the requirement of common knowledge on all GenCos' actual production costs (Song , Ni, Wenb, Hou, and Wu, 2003).

One major criterion for classifying game theory based methods is the level of competition: cooperative and non-cooperative (Shahidehpour, Yamin, and Li, 2002). According to the competition level in the liberalized electricity market, three general types of game models in imperfect competition, namely, Bertrand, Cournot, and SFE, have been proposed in the recent literature. As the most competitive model, a Bertrand competition model is an oligopolistic framework where the GenCos compete with one another by using prices as the strategic variables and ignore their capacity constraints. In classic Cournot models, however, the GenCos compete by using quantities as strategy choices, under the assumptions of homogenous products, price-responsive demand, and an MCP is determined by the intersection of aggregated supply and market demand curves. In the SFE models, the GenCos compete through the simultaneous choice of supply functions (Soleymani, Ranjbar, and Shirani, 2008). The competition level as well as the derived price equilibria generally lies between the Bertrand and Cournot model (Younes and Ilic, 1999). Recently, some other methods are also used to model and analyze the strategic behavior in deregulated electricity markets.

In a standard Bertrand competition model, identical sellers are assumed to have constant unit costs and no capacity constraints in competing on the price offers to consumers, and this will inexorably cause the identical sellers to price at marginal cost (von der Fehr, and Harbord, 2002). However, the electricity wholesale market deviates from the standard Bertrand price game. It is a capacity-constrained oligopoly and the marginal cost pricing is unlikely to be an optimal bidding strategy (Armstrong, Cowan, and Vickers, 1994). Therefore, only a few relevant references can be found in recent publications.

Federico and Rahman (2003) analyze the effects of changing auction rule from UP in the wholesale market to PAB under two polar market structures (i.e., a perfect competition or *Bertrand* structure and a perfect collusion or monopoly bidding) with demand uncertainty. It is found that under *Bertrand* structure there is a trade-off between efficiency and consumer surplus while changing to the PAB rule. Also, a move from UP to PAB under monopoly conditions has a negative impact on profits and output (weakly), a positive impact on consumer surplus, and ambiguous implications for welfare and average prices. Based on generalized Bertrand game, Ernst, Minoia, and Marija (2004) propose to optimize the profit and obtain a strategic bid by assuming that each GenCo has a constant marginal cost over its domain of generation and its rivals do not change their bids from the last round of market to the next one. They further adopt quadratic cost functions for GenCos and include a supply function in the bids instead of playing a Bertrand game (Minoia, Ernst, and Marija, 2004). Hu, Kapuscinski, and Lovejoy (2010) define Bertrand-Edgeworth (B-E) auctions as a modified version of Bertrand-Edgeworth games where the demand is inelastic and a price cap is set exogenously. B-E auctions are motivated by the discriminatory procurement auctions used in some wholesale electricity markets. They characterize the equilibrium structure for B-E auctions with multiple asymmetric bidding suppliers. Based on a proposed numerical algorithm, it is numerically illustrated that a weak (low-capacity) bidder does not necessarily price more aggressively in an oligopoly market. Bunn and Oliveira (2003) develop a stylized Bertrand game with constraints to explore the main strategic decisions faced by the GenCos in the England and Wales (E&W) market based on the proposed NETA.

Compared with Bertrand price-setting strategies, the quantity-setting equilibrium is more realistic for the electricity market. The Bertrand equilibrium assumes that by providing a lower price than others, a firm can capture the entire market demand and then meet it by expanding its output. However, such assumption is not tenable in view of the increasing marginal cost of electricity generation at a time point and the generation capacity constraints. Because the GenCos provide a homogeneous product, the Cournot assumption that firms make strategic decisions by quantity-setting behavior is considered to be a better approximation to reality than the price-setting assumption (Borenstein, Bushnell and Knittel, 1998; Hobbs, 1999). The representative Cournot competition based models are summarized as follows.

Borenstein and Bushnell (1999) use demand and plant-level cost data to simulate the competition in a restructured California electricity market. This approach recognizes that firms might have an incentive to restrict output in order to raise price, and it enables the explicit analysis on each firm's ability to do so. It is found that, while the results make the deregulation of generation less attractive than where there is no market power, they do not suggest that deregulation is a mistake. It is argued that policies promoting the responsiveness of both consumers and producers to price fluctuations can significantly affect the reduction of market power. Willems (2002) studies a Cournot competition based game with two GenCos who share one transmission line with a limited capacity to supply price-taking consumers. In the model, three different rules for the network operator to allocate transmission capacity are investigated, namely, all-or-nothing, proportional, and efficient rationing. Their effects on the GenCos competition and revenue are analyzed. Kian, Cruz, and Thomas (2005) use feedback Nash-Cournot strategies for the market participants to bid in dynamic electricity double-sided auctions. The simulation results show that compared with single-sided auctions, double-sided auctions are more efficient and lead to more stable and competitive MCPs. Tamaschke, Docwra, and Stillman (2005) focus on measuring the extent to which market power has been exercised in a deregulated electricity generation sector, and emphasize the need to consider the concept of market power in a long-term dynamic context. A market power index is constructed by considering the differences between the actual market returns and long-term competitive returns, and it is estimated by using a mathematical optimization model. The results suggest that generators have exercised significant market power from deregulation. To investigate whether individual participants can increase profits by withholding generation from the market, Ahn and Niemeyer (2007) develop a Cournot-based model of Korean power system for a set of loads representing the load duration curve for Korea's system loads in 2002. The results indicate a strong possibility for the exercise of market power to increase market price in Korean market.

Krause, Beck, Cherkaoui, Germond, Andersson, and Ernst (2006) perform a Nash equilibrium analysis by defining a pool market as a repeatedly played matrix game and compare it with an agent-based model. It is concluded that GenCos may act strategically by bidding above their marginal production costs. Kang, Kim, and Hur (2007) also propose a bidding model by using a two-player static game theory and analyze both demand-fixed and demand-under-uncertainty scenarios. The objective is to exploit a methodology to build the optimal bidding strategies for competitive power suppliers in day-ahead auction-based electricity markets with only the information on their possible future profits estimated from forecast system demand. It is indicated that a precise forecast of the demand can help the player to gain the advantage in the game. Under the Cournot assumption, Park, Ki, Kim, Jung, and Park (2001) analyze the power transactions in a deregulated energy marketplace such as PoolCo by modeling it as a non-cooperative game with complete information and determining the solution in a continuous strategy domain. A new hybrid solution approach employing a two-dimensional graphical approach as well as an analytical method is proposed to provide more apprehensible analysis.

Supply function equilibrium (SFE), originally introduced by Klemperer and Meyer (1989), is a way of describing how competitors could maximize profits in the competitive market of a single product with uncertain demands. In such a market, the participants prefer to set supply functions rather than compete in prices (Bertrand competition) or quantities (Cournot competition). Green and Newbery (1992) further advance the SFE theory by considering capacity constraints in analyzing the competition in the British electricity wholesale market, and they develop a model for the market of privatization. It is verified that supply curve bidding (SCB) can better benefit the GenCos compared to fixed quantity-price bids.

The advantages of SFE model compared with other models are also discussed by Baldick, Grant, and Kahn (2004). The SFE model, which constitutes a good compromise between the Cournot and Bertrand models, is believed to most accurately reflect the actual behavior of players in the real power markets. Also, it is more appropriate for the centralized markets where each GenCo bids in terms of a supply curve (Vahidinasab and Jadid, 2010). Therefore, the worldwide on-going deregulation of the electric market has been stimulating the SFE based modeling analysis of strategic bidding behavior. Representative SFE based models are summarized and discussed below.

Li and Shahidehpour (2005) propose an SFE based modeling method for analyzing the competition and bidding strategy among GenCos with incomplete information while taking into account transmission constraints. The competition is modeled as a bi-level problem in which the upper subproblem maximizes individual GenCos' payoffs and the lower subproblem clears the market. Sensitivity functions are developed for each GenCo's payoff with respect to its bidding strategies in order to solve the bi-level problem. An eight-bus system is employed to illustrate the proposed method, and the numerical results show the impact of transfer capability on GenCos' bidding strategies.

In the Standard Market Design (SMD) setup for electricity markets, Al-Agtash (2010) presents an SCB approach that iteratively alters the SFE model solutions and selects the best bid based on both the market-clearing locational marginal prices (LMP) and network conditions. This enables the GenCos to derive their best offering strategy in both the DAM and the long-term contractual markets. However, the results could vary significantly from one system to another according to the system characteristics such as network topology and suppliers' associated generation capacities.

It is observed that the SFE models are difficult in embracing congestion conditions, capacity constraints, and large systems with significant number of generators, unless strong restrictions are placed (Haghighat, Seifi, and Rahimikian, 2008). The supply curve bids obtained from the SFE models may not lead to a maximum profit, especially when the network is highly constrained. In view of this, Alagtash and Yamin (2004) reformulate the supply function equilibrium model for a GenCo owning a number of generators and present a new approach for optimal supply curve bidding (OSCB) using Benders decomposition. In their model, the offers of individual generators are simultaneously optimized to maximize the GenCo's total profit by taking into account physical constraints as well as transmission security constraints.

Many existing applications found in literature are limited to small systems due to the difficulty of analytically calculating the SFE for large systems. In view of this, Bompard et al. (2010) present an analytical approach to represent the strategic bidding behavior of the GenCos using an SFE model for large systems, in which the decision variables are the slopes of the supply function. The proposed approach proves to be rather precise in determining the SFE and it can consider the GenCos' capacity limits. It is also shown that generation capacity constraints may contribute to GenCos' market power. One GenCo may be pivotal if its rivals are capacity constrained. It could substantially raise the market price by unilaterally withholding the output.

To fully consider the impact of capacity constraints and pivotal firms on equilibrium predictions, Genc and Reynolds (2011) characterize the set of symmetric SFEs for capacity-constrained GenCos in the UP-auction wholesale market. It is shown that the rise of GenCos' capacities could lead to the increase of this set of equilibria. Holmberg (2009) also numerically studies the asymmetric supply function equilibrium with capacity constraints and shows that in this case, the valid SFE can be calculated by means of an algorithm that combines numerical integration with an optimization procedure that searches for an end condition.

To develop optimal bidding strategies for the GenCos of oligopolistic energy markets, Vahidinasab and Jadid (2010) study the impacts of GenCos' pollutant emission on their bidding strategies. By neglecting demand side bidding and bilateral contracts, the GenCos are assumed to submit quadratic bidding curves to the market and bid the quadratic coefficient of bidding curves under the locational marginal pricing mechanism. The model employs SFE to represent each supplier's strategic behavior. Normal boundary intersection approach is used for generating Pareto optimal set and fuzzy decision making is employed to select the best compromise solution. The MCP is obtained via the multi-objective optimal power flow method. The optimal bidding strategies are mathematically developed by using a bi-level optimization problem solution.

In view that most previous research mainly focuses on a single-period, single-market model, Gao and Sheble (2010) develop an SFE model for GenCos' bidding optimization in a multiple-period and multi-market scenario. For the proposed SFE model, they obtain the equilibrium condition by using discrete time optimal control which considers fuel resource constraints. The multiple-period optimal bid strategy is analytically derived by manipulating the intercept parameters of the SFE model. Both electricity and fuel markets are simultaneously considered, and the SFEs with resource constraint and transmission constraint are investigated, respectively.

Sahraei-Ardakani and Rahimi-Kian (2009) propose an *n*-player dynamic game model to analyze the bidding strategy in an oligopolistic electricity market with either fixed or variable demands. In this analytical short-term model, all players first choose their own strategy, and then, define an improvement matrix to improve the bidding strategy after seeing one another's strategy and the resulting payoff. A state-space approach is used to obtain the supply function equilibrium strategy for the model under many realistic constraints such as production bounds, ramping limits, up/down times, system contingencies, and transmission congestions.

Generally, in a centralized market, where the power pool collects suppliers' bids and loads and determines the dispatching schedule, the competition level as well as the model type is dependent upon the bidding procedure and the pricing rule. The GenCos may bid on prices without worrying about quantities (Bertrand competition), but they may also bid on their production quantity as a function of the prices to be received from the equilibrium prices (SFE). In a decentralized market, the transactions are performed in bilateral or multilateral markets and the type of short-term competition is endogenous. The GenCos may compete by choosing the quantity they are willing to put on the market and the MCP is to be determined by an ISO (Cournot competition). The supply function model is a good compromise between Cournot and Bertrand competitions in a highly decentralized market (Younes and Ilic, 1999).

2.4.3. Agent-based models

The restructured electricity markets typically involve price-quantity pairwise bids for the sale of large amounts of electricity by a small number of GenCos, resulting in extremely complex market processes in which traditional analytical and statistical tools are difficult to be applied (Tesfatsion, 2002). In an agent-based model, market participants are modeled as adaptive agents with different bidding preferences and strategies, and the suppliers are enabled to utilize their past experiences to improve their behaviors in the market. Each agent may develop the optimal bidding strategy by learning from its past experiences obtained from the direct interaction with environment. This brings a new type of numerical analysis theory to deal with complex trading issues in the restructured electricity market (Tesfatsion, 2006). Generally, the agent-based modeling procedure can be described as follows: (1) define the research questions to be resolved; (2) construct a model comprising an initial population of agents; (3) specify the initial model state by defining the agents' attributes and the structural and institutional framework of the electricity market within which the agents operate; (4) have the model evolve over time without further intervention; (5) analyze simulation results and evaluate the regularities observed in the data (Weidlich and Veit, 2008)

Agent-Based Computational Economics (ACE) is a fairly young research paradigm that offers methods for realistic electricity market modeling to overcome some shortcomings of the other methods discussed above (Weidlich and Veit (2008). A growing number of researchers have developed many agent-based models for simulating electricity markets (Sueyoshi and Tadiparthi, 2005; 2007). However, compared with single GenCo optimization models and game theory based models, agent-based modeling studies for bidding strategy analysis are much less in literature. The representative publications are reviewed as follows.

Rahimiyan and Rajabimashhadi (2008) compare the Q-learning (QL) approach and the model-based (MB) approach in optimizing supplier's bidding strategy under electricity PAB auction rule. The suppliers' behaviors are modeled in a multi-agent system, and the simulation results show that the Q-learning algorithm can enable the suppliers to find the optimal bidding strategy in the PoolCo market. For different PDFs, the QL algorithm can always converge to the optimal solution obtained using the model-based approach. The results also show that the suppliers could adopt the bidding strategy according to their rivals' behavior and other effective factors in power system operation using the QL algorithm. Sheble (2001) proposes a genetic algorithm (GA) and an ACE framework for optimizing sellers' bidding strategies in a double-sided auction market where some players attempt to benefit from applying different strategies to cause economic instabilities and intentionally drive market prices. It is demonstrated that the market power can be significantly experienced by some players under the UP rule. Naghibi-Sistani, Akbarzadehtootoonchi, Javididashtebayaz, and Rajabimashhadi (2006) propose a Q-learning algorithm for the participants to find the optimal bidding strategy in the PoolCo electricity market. In this method, each bidder, independent of the others, learns about its state and the wholesale price. The results show that with the temperature variation reinforcement learning, the suppliers can learn the optimal policy found by game theory and can be adaptive to the parameter variations in the market

Wen and David (2001d) adopt Monte Carlo simulation and a refined genetic algorithm to build optimally coordinated bidding strategies for competitive suppliers in energy and spinning reserve markets. Under the assumptions that each supplier bids a linear supply function into the energy and spinning reserve markets, respectively, and that the two markets are dispatched separately, each supplier chooses the coefficients in the two linear supply functions to maximize the total benefits, considering the rivals' possible bidding policy. They also apply a similar GA-based method to develop an overall bidding strategy for the suppliers in the day-ahead market (Wen and David, 2001).

Gountis and Bakirtzis (2004) propose a GA approach to optimize the profits of individual GenCos with multiple generating units. The model uses Monte Carlo simulation to calculate the expected profit and GA to find the optimal strategy. It is assumed that each supplier bids a linear supply function and considers the other bidders' bidding behaviors in the forms of PDF. However, such an assumption is usually not realistic since the rational generator can be expected in a competitive bidding environment. Therefore, the profits estimated by the proposed algorithm are not realizable. Earlier agent-based simulations employing GA algorithm for bidding strategy analysis can be found in (Richter and Sheble, 1998; Richter and Sheble, 2000).

Besides, some other learning algorithms have also been investigated. For example, Fujii, Okamura, Inagaki, and Yamaji (2004) apply a multi-agent model, which learns a bidding strategy autonomously through trial-and-error search action, to numerically analyze the price formation process of an open electric market. The model is believed to be helpful in analyzing more general electricity markets which have several different types of power plants with unit commitment costs, seasonal and hourly demand fluctuation, real-time regulation market and operating reserve market. Bunn et al. (2010) develop an agent-based simulation model by using computational learning (CL) algorithm to investigate the impact of vertical integration between electricity generators and retailers on market power in a competitive wholesale market setting. It is found that in various cases, whilst vertical integration generally reduces wholesale prices, it can increase or decrease the market power of other generators, depending upon the market share and the technology segment of the market.

Azadeh, Skandari, and Maleki-Shoja (2010) propose an agent based simulation model based on Ant Colony Optimization (ACO) algorithm to compare three wholesale electricity market clearing strategies: UP, PAB, and generalized Vickrey rules. The proposed model is suitable for high-dimension bidding functions and enables modelers to avoid “curse of dimensionality”. They investigate step-wise discrete bid functions and their impacts on the efficiency of the market under different available market settlement rules. The assumptions in the proposed algorithm include: inelastic and fixed demand; call market with different price settlement rules (UP, PAB, and generalized Vickrey rules); no transmission constraint; and a step-wise discrete bid function. The method can solve dynamic and static combinatorial optimization problems of market strategy optimization.

Besides, there are a few other studies on agent-based modeling applications in literature. Xiong, Okuma, and Fujita (2004) compare the UP and PAB auction rules by using a multi-agent approach, where each adaptive GenCo develops the bidding prices based on Q-learning algorithm. The experimental results show that the PAB auction can result in lower expected market prices and price volatility. It is also shown that the demand-side response has less effect in reducing market prices under the PAB auction rule since in this auction the bidders bid as close to the market prices as possible, which makes the aggregate supply curve more flattened than that under the UP rule. Xiong, Hashiyama, and Okuma (2002) propose an evolutionary daily bidding strategy for the supplier in a perfectly competitive day-ahead electricity auction market under the PAB auction rule. The feasibility of the proposed bidding strategy is verified by agent-based simulations. Walter and Gomide (2003) present a GA-based approach to obtain evolutionary GenCo bidding strategies in the power market. Simulation results illustrate that the strategies derived by this approach is superior in enhancing the GenCo's profitability over the commonly adopted marginal cost-based approaches. They also introduce a co-evolutionary algorithm to obtain a profitable bidding strategy for the participant by using information commonly available in a much dynamic environment. The results demonstrate that this approach can further improve the GenCo's profits (Walter and Gomide, 2008). Gao, Gutierrez-Alcaraz, and Sheble (2006) build an adaptive multi-agent model to analyze and compare the application performances of genetic algorithm, evolutionary programming, and particle swarm in simulating participants' bidding behaviors. Sueyoshi and Tadiparthi (2008) develop an agent-based decision support system (DSS) for analyzing the dynamic price change in the competitive electricity wholesale environment. The proposed DSS is effective in assessing new trading strategies in the electricity PoolCo market.

Agent-based models can mimic human behaviors and simulate optimal bidding procedures. In such models, market participants are handled as adaptive agents with different bidding preferences and strategies, and their bidding decisions are influenced by many uncontrollable factors. This group of models is more flexible, robust, and easily implemented compared with the previous two groups of approaches, and thus it opens up a new type of modeling analysis to deal with business complexity of electricity trading. However, the underlying mathematical foundation for agent-based modeling has not yet been clearly verified to date (Gao and Sheble, 2010).

2.4.4. Hybrid models and other modeling methods

Besides the above three major groups of modeling approaches, there are a few other innovative methods developed recently for strategic bidding analysis. In particular, the hybrid approach that combines multiple modeling methods stimulates enormous interests among the researchers.

Yamin and Shahidehpour (2003) develop a hybrid model combining LR algorithm and GA for the GenCos to generate a proper unit commitment scheduling and derive the optimal supply curves. It is shown that the proposed hybrid model is better than the LR approach and the traditional unit commitment approach in terms of helping the GenCos to increase profits. They also adopt the augmented Lagrangian relaxation algorithm to solve self-scheduling and energy bidding problems in competitive electricity markets constrained by transmission congestions, fuel, and emission. The supply curve is derived as a function of generation schedule to achieve the maximum profit. The slope of the supply curve is dependent upon the forecast price and the power output obtained in the self-scheduling result (Yamin and Shahidehpour, 2004).

Sueyoshi (2009) proposes an agent-based approach equipped with game theory for analyzing the strategic collaboration among learning agents during a dynamic market change in the 2000-2001 California electricity crisis. The concept of partial reinforcement learning is incorporated into trading agents who can learn from both the dynamic market change and the collaboration with other traders. It is found that the learning speed of traders becomes slow when a large fluctuation occurs in the power exchange market. Azevedo and Correia (2006) propose a model by combining Bayes' rule and the game theory for the participants in the first stage of an electric energy bilateral contract auction in the Brazilian market. The model can help one agent to attribute bids to the other agents and observe the consequences.

In addition, Song, Ni, Wenb, Hou, and Wu (2003) present the concept of conjectural variation (CV) and its applications to the strategic bidding of GenCo in the oligopolistic electricity wholesale market. The conjecture of a firm is defined as its belief or expectation of how its rivals will react to the change of its output. It is verified that, in a real wholesale market containing multiple players, CV based bidding strategy (CVBS) enables a GenCo to integrate its rivals' responses into one pseudo-competitor's response and to make optimal decision accordingly based on available imperfect information announced in the market. It is also demonstrated that classical game theoretical bidding strategies (GTBS), such as Bertrand, Cournot, Stackelberg and monopoly (collusion), are actually the special cases in the CVBS solution family.

3. TIME SERIES BASED SHORT-TERM WIND FORECASTING

In this chapter, two kinds of typical forecasting models, autoregressive integrated moving average (ARIMA) and artificial neural networks (ANN) are built, implemented, evaluated, compared, and comprehensively investigated in performing short-term wind forecasting. The time series based modeling methods of ARIMA and three types of ANNs are concisely presented in Section 3.1 first. After a brief introduction to the performance metrics is given in Section 3.2, the study on the application of ARIMA and different ANN modeling methods for one-hour-ahead wind speed forecasting is performed and presented in Sections 3.3 and 3.4, respectively. Section 3.5 presents a short discussion on the results.

3.1. Time Series Models

A time series is a set of measurements or observations collected at successive points or over successive periods. Time series methods, e.g., ARIMA and ANN, can generate forecasts based solely on the hidden patterns in the historical data by using time as independent variable. Especially, the ANN trained with time series have the ability to model arbitrarily linear and nonlinear functions. Having been widely utilized in various different fields including transient detection, pattern recognition, approximation, and time series forecast, artificial neural network (ANN) is a promising technology in the field of wind forecasting applications. For example, Alexiadis, Dokopoulos, Sahsamanoglou, and Manousaridis (1998) claim that their ANN predictor is about 10% better than persistence model for one-step-ahead forecast.

3.1.1. Box & Jenkins models

The mathematical description of the wind pattern recognition or forecast problem aims to find an estimate $f(t+k)$ of the wind vector $y(t+k)$ based on the previous n measurements $y(t)$, $y(t$-

1),..., *y*(*t*-*m*+*1*). In order to have accurate wind speed forecast, *k* is chosen to be small and this is called short-term wind speed forecast.

Box and Jenkins (1994) propose an interactive approach for fitting such ARIMA models to time series. The linear expression of a random time series is as follows:

$$y_t = \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \phi_j e_{t-j} + e_t, \tag{1}$$

where φ_i is the *i*th autoregressive parameter, ϕ_j is the *j*th moving average parameter, e_t is the error term at time *t*, and y_t is the value of wind speed at time *t*.

This is a typical ARMA model, denoted as ARMA(*p*, *q*). If *q* is assumed to be zero, i.e., the model only involves autoregressive terms, it is usually referred as an AR model, denoted as AR(*p*). Similarly, the model only involving moving average terms is usually referred to as an MA model, MA(*q*). Especially, the persistence model is the simplest ARMA model which applies the present value as the prediction value of next time. If the time series is not stationary or a linear trend exists in the data, an ordered differential transformation can be applied and an ARIMA model can be obtained correspondingly, usually denoted as ARIMA(*p*,*d*,*q*), in which *d* represents the order of the differential transformation (Ma et al., 2009). The application of such an approach usually involves identifying the possible model structure, estimating the model parameters, checking model adequacy, and forecasting the objective of interest, which are explained in the following context when the application results are presented and discussed.

3.1.2. BP neural networks

As one of the most popular ANN techniques, feed forward back-propagation (BP) neural network is a kind of supervised learning neural network. It is usually composed of one input layer, one or more hidden layers, and one output layer. The source nodes in the input layer of the network supply respective elements of the activation pattern or input vector, which constitute the input signals applied to the neurons in the hidden layer. The hidden neurons function to intervene between the external input and the network output. The output signals of the hidden layer are used as inputs to the output layer. The output signals of the neurons in the output layer of the network constitute the overall response of the network to the activation patterns applied by the input-layer neurons (More and Deo, 2003). Figure 5 illustrates the common topology of one three-layer BP neural network model, which is adopted in this dissertation study. A BP network with biases, a sigmoid layer, and a linear output layer is capable of approximating any function with a finite number of discontinuities.

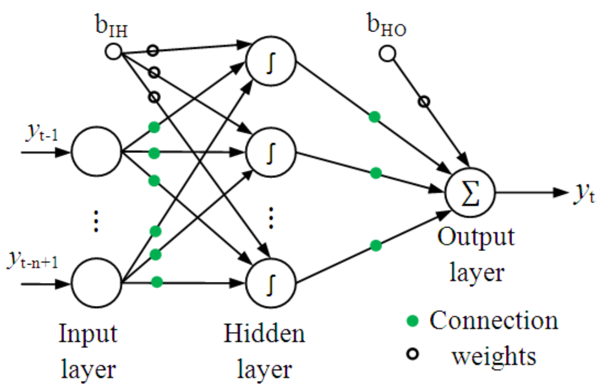


Figure 5. Topology of a BP neural network.

With the network weights and biases initialized, the neural network must be trained before it can help solve any particular problem. Generally, the BP learning algorithms iteratively perform two phases, error propagation and weight updating, until the network performance is satisfactory. During the process of propagation, a training sample's inputs are first forward propagated through the network to generate the output activations, which are then compared with the training pattern's target and backward propagated through the network to generate the delta or "errors" of all output and hidden neurons. During the weight-updating phase, the output delta are multiplied by the input activation to obtain the gradient of the weight; and the weight is updated or changed in the opposite direction of the gradient by subtracting a ratio of it (usually called as learning rate) from the weight. With *n* input neurons, *m* hidden neurons, a brief description of general training process of BP neural networks with *n* input neurons, *m* hidden neurons, and one output neuron could be found in (Li and Shi, 2010a). The general input-output relationship of neuron *j* in layer *l* can be expressed as:

$$y_j^l = f_j^{(l)}\left[\sum_{i=1}^{l-1} \omega_{ij}^{(l)} y_i^{(l-1)} - b_j^{(l)}\right] \quad (j = 1, 2, \ldots, n_l; l = 1, 2, \ldots, m), \tag{2}$$

where $f_j^{(l)}$ is the activation function of neuron *j*, $\omega_{ij}^{(l)}$ is the connection weight between neuron *i* in layer (*l*-1) and neuron *j* in layer *k*, $b_j^{(l)}$ is the bias or threshold value of neuron *j*, and n_l is the number of neurons in layer *l*.

The activation function of the node in the hidden layer is usually a sigmoid function, as given in Eq. (3), whereas a linear activation function is usually used for the output layer,

$$f_j^{(k)}(y) = \frac{1}{1+\exp(-y)}. \tag{3}$$

The Levenberg-Marquardt algorithm is adopted as the training methods in this study in view that it is usually the fastest method for BP models. The Levenberg-Marquardt algorithm updates the weights and bias as follows:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left[\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}\right]^{-1}\mathbf{J}^T\mathbf{e}, \tag{4}$$

where $\mathbf{x}_k$ is a vector of current weights and biases, **J** is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, **e** is a vector of network errors, and μ is a scalar. The detailed description of the algorithm can be found in (Press, Flannery, Teukolsky, and Vetterling, 1992).

3.1.3. RBF neural networks

The RBF neural network is a multi-input, single-output forward network, which is composed of an input layer, a hidden layer, and an output layer, as illustrated in Figure 6.

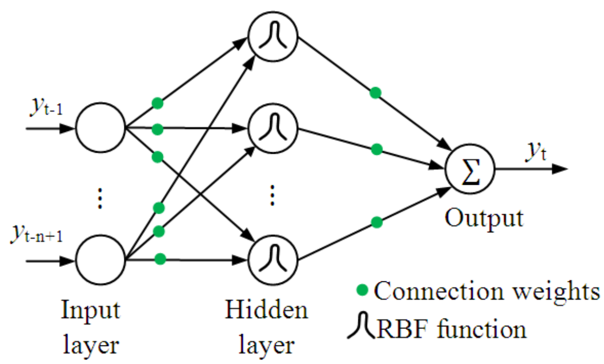


Figure 6. Topology of RBF neural network.

RBF consists of two layers whose output nodes form a linear combination of the basis functions in the hidden layer produce a significant nonzero response to input stimulus only when the input falls within a small localized region of the input space. Transformation of the inputs is essential for fighting the curse of dimensionality in empirical modeling. The type of input transformation of the RBF is the local nonlinear projection using a radial fixed shape basis function. After nonlinearly squashing the multidimensional inputs without considering the output space, the radial basis functions play a role as regressors. The weights of the regressors can therefore be determined using the linear least square method, which gives an important advantage for convergence (Kisi, 2007).

It is assumed that, given N n-dimension different points $\{x_i \in R^n, i = 1, 2, \cdots, N\}$ and N real numbers $\{y_i \in R, i = 1, 2, \cdots, N\}$, a nonlinear function $f(x)$ can be found satisfying $f(x_i) = y_i, i = 1, 2, \cdots, N$. This function is called an RBF when it depends only on the radial distance $r = \|x - t\|$, where t refers to the centre of point x.

The RBF approach consists in choosing f from a linear space of dimension *N*, depending on the data points $\{x_i \in R^n, i = 1, 2, \cdots, N\}$. The basis of this space is chosen to be the set of functions

$$\{h(\|x - x_i\|), i = 1, 2, \cdots, N\}, \tag{5}$$

where $\|\cdot\|$ is the Euclidean norm on R^n. Therefore, the solution of the above-mentioned interpolation problem has the following form:

$$f(x) = \sum_{i=1}^{N} c_i h(\|x - x_i\|), \tag{6}$$

where coefficients c_i can be obtained by imposing the interpolation conditions $f(x_i) = y_i, i = 1, 2, \cdots, N$ on the above equation. Thus, the following solution can be derived.

$$f(x) = \sum_{i=1}^{N} c_i h(\|x_j - x_i\|), \ j = 1, 2, \cdots, N \ . \tag{7}$$

By defining the vectors *y*, *c* and the symmetric matrix *H* as $(\mathbf{y})_j = y_j$, $(\mathbf{c})_j = c_j$, $(\mathbf{H})_{ij} = h(\|x_j - x_i\|)$, the coefficients c_i can be obtained from $\mathbf{c} = \mathbf{H}^{-1}\mathbf{y}$.

3.1.4. Adaptive linear element networks (ADALINE)

The structure of a simple ADALINE used in this paper is illustrated in Figure 7, where the weight matrix of W has only one row, corresponding to the one-column input vector p. The network output is

$$f = \mathbf{Wp} + b = \sum_{i=1}^{n} w_i y_i + b = \sum_{i=0}^{n} w_i y_i \ , \tag{8}$$

where y_0 represents the threshold of bias b with a weight of $w_0 = 1$.

First developed by Widrow and Lehr (2002), the ADALINE networks can only solve linearly separable problems. However, the least mean squared (LMS) or Widrow-Hoff learning rule can minimize the mean squared error (MSE) and search for the global minimum point in space, thus moving the decision boundaries as far as it can from the training patterns. During the learning process, the LMS rule diminishes MSE, a mathematical function is defined in the multi-dimension space of weights for a set of given training patterns as follows,

$$MSE = \frac{1}{2L}\sum_{l=1}^{L}\varepsilon_l^2 = \frac{1}{2L}\sum_{l=1}^{L}(d_l - f_l)^2\,, \tag{9}$$

where L is the number of patterns in the training dataset, d_l and f_l represent the desired value and the forecast one of the network, respectively.

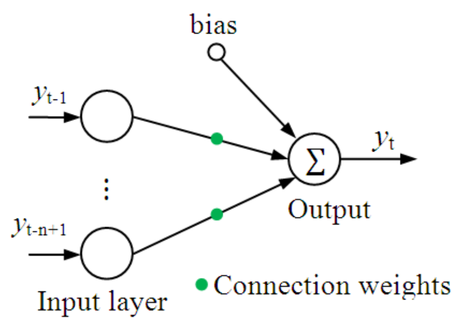


Figure 7. Topology of a simple ADALINE network.

The changes in weights at time (t+1) and time t are proportional to the descendent gradient of the error function, which is commonly defined as the learning rate α

$$w_i(t+1) = w_i(t) + \alpha(d_l - f_l)y_{l_i}\,. \tag{10}$$

3.2. Performance Metrics

The performance metrics adopted in this dissertation study include mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). MAE is a linear score with all the individual differences being weighted equally in the average. RMSE is a quadratic scoring rule which measures the average magnitude of the error. Since the errors are squared before being averaged, the RMSE gives a comparatively high weight to large errors.

This makes it most useful when large errors are particularly undesirable. MAPE is measure of accuracy in a fitted time series value in statistics, specifically trending.

Their calculation equations are expressed as follows:

$$MAE_j = \frac{1}{T}\sum_{t=1}^{T}\left|y_t - f_{j,t}\right|, \tag{11}$$

$$RMSE_j = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(y_t - f_{j,t}\right)^2}, \tag{12}$$

$$MAPE_j = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{y_t - f_{j,t}}{y_t}\right|, \tag{13}$$

where y_t and $f_{j,t}$ denote the measurements or observations and the forecast value from model j, for a given time point t. respectively, T is the length of testing samples of data employed for performance evaluation and comparison.

3.3. Box & Jenkins Modeling for One-Hour-Ahead Wind Speed Forecasting

In order to quantitatively investigate, evaluate and compare the performances of the above-mentioned time series based forecasting methods, different Box & Jenkins models are first developed to perform one-hour-ahead wind speed forecasting.

The hourly mean wind speed collected at an observation site, Hannaford (Hann) in North Dakota, is adopted. Table 3 gives the geographical information of the site and its wind speed characteristics during the whole year of 2002. It should be noted that the wind speeds at the height of 10 meter are used as recommended by the World Meteorological Organization (WMO) (Mathew, 2006).

As afore-introduced, the application of this modeling method involves identifying the plausible models, estimating the parameters, checking model adequacy, and forecasting the objective of interest, here, the next hourly wind speed. The plausible models can be identified by examining the time series plot, the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots, as shown in Figures 8 – 10, respectively.

Table 3. Information of the observation site and its wind speed characteristics.

Site	Latitude (North)	Longitude (West)	Elevation (m)	Wind speed at 10 m above ground level (m/s)		
				Mean	Min	Max
Hannaford (Hann)	47º19'39"	98º12'34"	448	7.404	0.362	25.562

Figure 8 indicates that the hourly wind speed varies randomly within a limit. The trend that large autocorrelations die out individually can be observed in Figure 9, indicating that a constant mean might exist. Thus, the data might be regarded as stationary approximately. However, in order to verify this judgment, and more importantly, to evaluate and compare the forecasting performances of different Box & Jenkins models, the 1-order differencing model is also investigated in the following forecasting process.

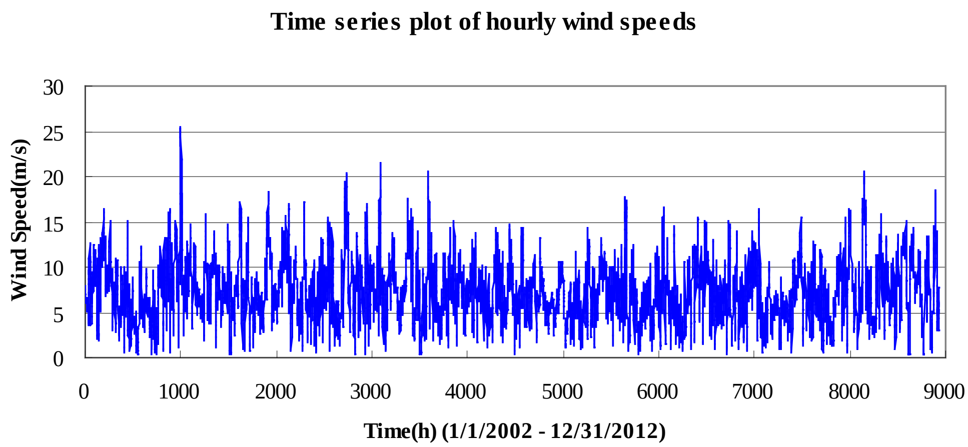


Figure 8. Time series plot of hourly wind speed data.

In order to identify proper autoregressive or moving average components to be included in the Box-Jenkins model, the ACF and PACF plots of the stationary data should be further examined. Generally, for a stationary time series, an ACF with large spikes at initial lags that decay to zero or a PACF with a large spike at the first and possibly at the second lag indicates an autoregressive process, whereas an ACF with a large spike at the first lag and a PACF with large spikes at initial lags that finally die out indicates a moving average process. If the ACF and the PACF both exhibit large spikes gradually decaying to zero, both autoregressive and moving averages processes should be considered (Brockwell and Davis, 2002). It can be observed from Figures 9 and 10 that the autoregressive process is significant compared with the moving average component for the selected dataset. In this study, however, besides AR models, ARMA models and ARIMA models are also investigated in the following forecasting process for purpose of double-checking and comparison.

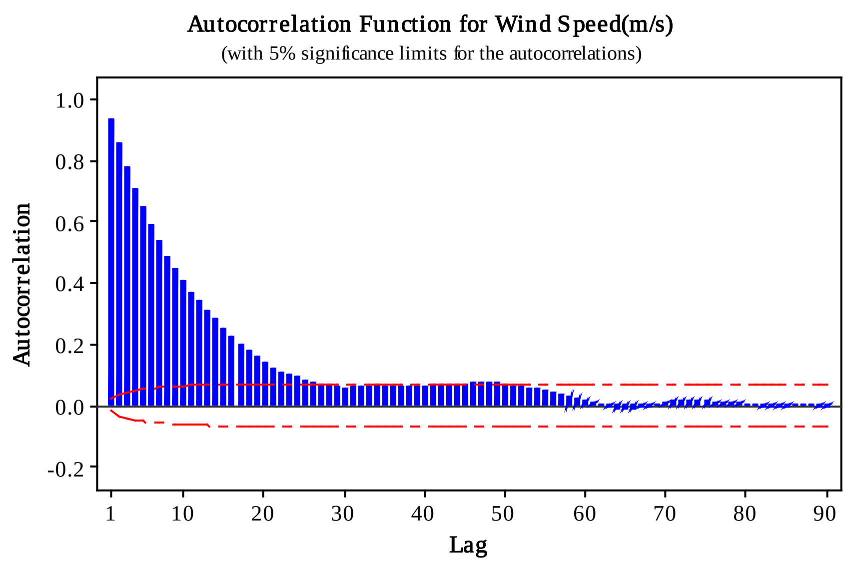


Figure 9. ACF plot of hourly wind speed data.

Based on the process, different plausible AR, ARMA, and ARIMA models are fitted with the significance of their model parameters is examined. Table 4 gives the testing results for AR models for purpose of demonstration. AR(2) is tested first and its parameters are estimated. Based on the *p*-values of *t*-tests, it can be seen that besides the constant parameter, both the two AR components are significant (*p*-values <0.05). In this case, the AR(3) model is further investigated and its parameters are estimated as shown in Table 4. It can be easily figured out that the third parameters can be ignored since its *p*-value is greater than 0.05.

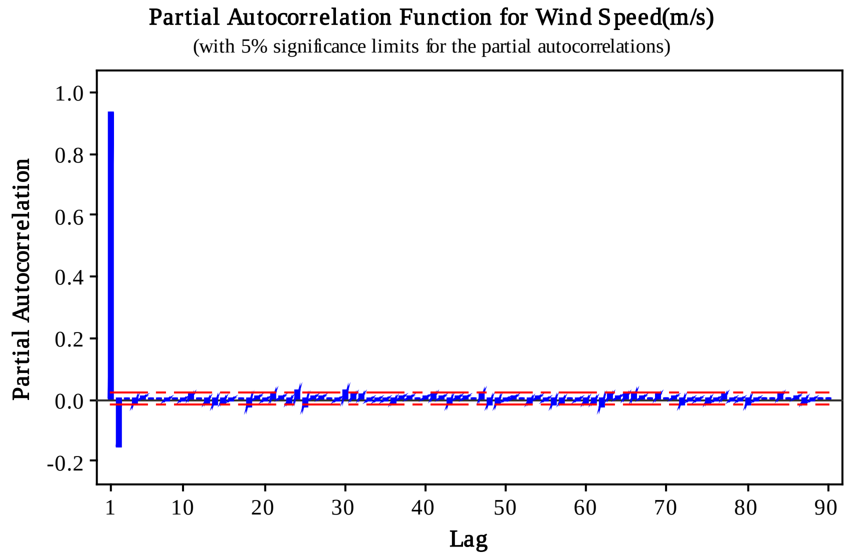


Figure 10. PACF plot of hourly wind speed data.

Table 4. Estimates of model parameters of AR(3) and AR(2).

Type	AR(2)		AR(3)	
	Parameter	P-value	Parameter	P-value
AR1	1.0788	0.000	1.0795	0.000
AR2	-0.1532	0.000	-0.1582	0.000
AR3	/	/	0.0046	*0.665*
Constant	0.54542	0.000	0.54286	0.000

Similarly, different ARMA models and ARIMA models are tested and examined. As a result, three models, AR(2), ARMA(1,1), and ARIMA(1,1,1), are determined as the plausible models for fitting the data adopted in this study. These models are then used to perform one-hour-ahead forecasts for the observed wind speeds during the last five days of December, 2012 (120 points). The MAE, RMSE and MAPE values of their forecasting results are calculated as the performance indices, which are shown in Table 5. The time series plots of their forecasts are demonstrated in Figure 11 as well.

Table 5. Performances of different BOX-JENKINS forecasting models.

BOX-JENKINS Models	MAE	RMSE	MAPE
AR(2)	0.974	1.277	0.211
ARMA(1,1)	0.972	1.273	0.209
ARIMA(1,1,1)	0.985	1.282	0.197

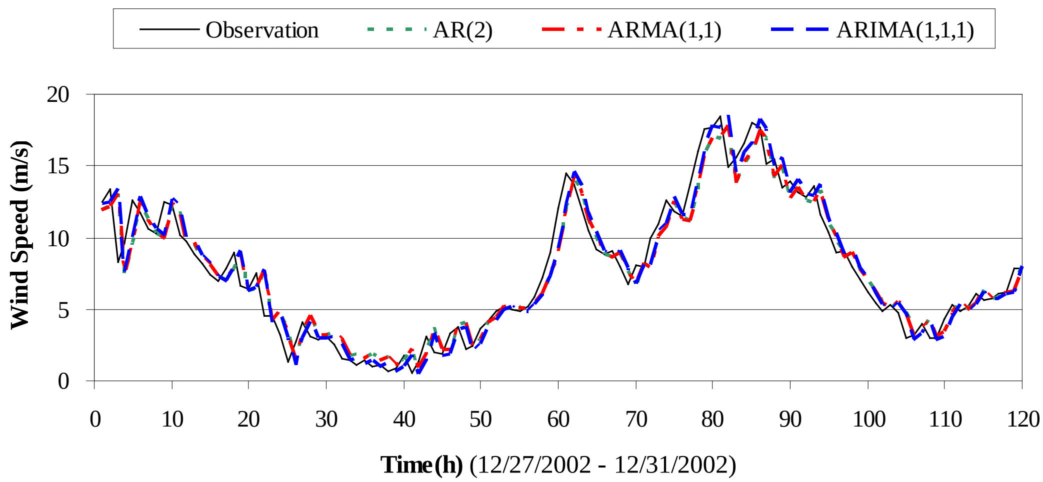


Figure 11. Time series plots of BOX-JENKINS model forecasts.

It can be observed from both Table 5 and Figure 11 that all the three models can provide satisfactory one-hour-ahead forecasts on wind speed. There is no significant difference in their forecasting performances. Especially, depending on the performance indices used, different models could be deemed as the 'best' one. For example, the forecasting model of ARMA(1,1) has slightly better performance than AR(2) and ARIMA(1,1,1) in terms of both MAE and RMSE, but it is slightly 'worse' than ARIMA(1,1,1) in terms of MAPE.

3.4. ANN Based Modeling for One-Hour-Ahead Wind Speed Forecasting

A comprehensive comparison study is further performed to quantitatively evaluate the performance of the above-mentioned three different ANN models (BP, RBF, and ADLINE) in their short-term wind forecasting applications. The results further verify that even for the same wind dataset, no single neural network model outperforms others universally in terms of different performance indices. Moreover, the selection of the type of neural network for best performance is also dependent upon the data sources. Among the optimal models obtained, the relative difference in terms of one particular performance metric could reach as much as 20%. This indicates the need of generating a single robust and reliable forecast by applying some post-processing algorithms.

For purpose of comparison, the same dataset collected from site Hannaford (Hann) is adopted. Besides, another observation site, Killdear (Kill) in North Dakota is also employed in order to further investigate the performance consistency of ANN models with different wind datasets. Their geographical information and its wind speed characteristics during the entire year of 2002 are summarized in Table 6.

Table 6. Information of two ND sites and their wind speed characteristics.

Site	Latitude (North)	Longitude (West)	Elevation (m)	Wind speed at 10 m above ground level (m/s)		
				Mean	Min	Max
Hannaford (Hann)	47º19'39"	98º12'34"	448	7.404	0.362	25.562
Killdear (Kill)	47º22'48"	102º45'36"	788	7.736	0.349	21.373

3.4.1. ANN model building

The models adopted here have the same structures as illustrated in Section 3.1. Similarly, for the input layer, the ACF and PACF plots could provide useful information. For purpose of comparison, however, the inputs of the ANN models varies from previous 1 to 8 observations for all models. Correspondingly, each previous *n* (n=1, 2, …, 8) observations are preprocessed and converted into one input vector of the format for ANN-based models. The preprocessed data are further divided into three subsets: training, evaluation, and testing datasets, respectively. With the last 120 (five days) input vectors being separated as the testing dataset, 5000 other input vectors are randomly selected as the training dataset, and the rest is used to construct the validation dataset.

For the hidden layer, as previously noted, gradually increasing the number of hidden neurons might be a good trial-and-error solution to the difficulty in creating a sufficiently accurate BP neural network. However, according to Wanas, Auda, Kamel, and Karray (2002), the optimal number of hidden-layer neurons for BP and ADALINE models can be selected as the integer number close to $\log(T)$, where T is the number of training vectors. In this study, their numbers of hidden-layer neurons are thus selected as 4. As for RBF model, the number will be optimized iteratively during training process.

The output layer of all three types of ANN models only contains one neuron representing the forecast value of next hourly average wind speed. For BP and ADALINE models, different learning rates are examined in the study. For BP, it is from 0.025 to 0.5 with 0.025 increments, while for ADALINE, it is from maximum learning rate (MLR) to one tenth of MLR with one tenth of MLR decrement. For RBF models, different spread constants are examined in the study, namely, from 0.5 to 1.5 with 0.1 increments.

The data collected at each site are preprocessed and transformed into the specific input format for ANN models according to n, the number of previous hourly wind speeds in each input vector. The formatted data are further divided into training dataset D_T, validation dataset D_V, and testing or evaluation dataset D_E. During the NN process, all the models of ADALINE, BP, and RBF networks are trained and evaluated with the corresponding datasets at various learning rates as mentioned in Chapter 3. After that, each model is applied to forecast the next hour wind speeds corresponding to the testing data and the values of MAE, RMSE, and MAPE are calculated. The models of each type, which generate the smallest MAE, RMSE, or MAPE, are selected. They are regarded as the best model among the tested models of the same type.

3.4.2. Forecasting results for the site of Hann

Through extensive calculation, the corresponding MAE, RMSE, and MAPE values are obtained according to the forecasting results from different ANN models. For the purpose of brevity, the details of each possible combination of parameters are not provided. Table 7 only gives the results about the optimal ANN models that generate the smallest MAE, RMSE, and/or MAPE values, for the dataset collected from the site of Hann.

Table 7. Optimal ANN models (Hann).

Best model	MAE	RMSE	MAPE	CT(s)
LNN_Obs_2_lr_8.9e-8	0.965	1.271	0.196	1.482
LNN_Obs_4_lr_4.5e-8	0.980	1.290	0.194	1.529
BP_Obs_8_LR_0.075	0.945	1.269	0.206	3.323
BP_Obs_6_LR_0.1	*0.951*	*1.254*	0.211	*2.777*
RBF_Obs_2_spread_0.5	0.989	1.297	0.232	0.047
RBF_Obs_3_spread_0.5	0.997	1.294	0.234	0.047
RBF_Obs_7_spread_0.7	1.058	1.390	0.221	0.046

Among all the 160 BP models tested, the combination of using 6 previous observations and learning at a rate of 0.1 generates the smallest RMSE value (1.254), while using eight previous observations and a 0.075 learning rate resulted into the smallest MAE (0.945) and MAPE (0.206). For all the ADALINE models (denoted as LNN in the table), using 4 previous observations and a learning rate of 4.49e-8 gives the smallest MAPE (0.194), while using 2 previous observations and a learning rate of 8.86e-8 gives the smallest MAE (0.965) and RMSE (1.271). The situation is more complicated with the RBF models, in which each smallest MAE, RMSE, or MAPE metric corresponds to a different combination of parameters.

More importantly, Table 7 indicates that different types of ANN models could demonstrate different performances. The BP model using previous 6 observations and trained at a learning rate of 0.1 seems to be the best in terms of both MAE and RMSE, while the ADALINE model using 4 previous observations and learning at 4.49e-8 learning rate appears to be the best on in term of MAPE. The best ADALINE model outperforms the best BP model by 4.8% and the best RBF model by 14.0% in terms of MAPE.

Besides, the convergence times (CT) for the seven models are also summarized in Table 7. It can be seen that the RBF models generally converge most swiftly whereas the BP models take the longest time for convergence.

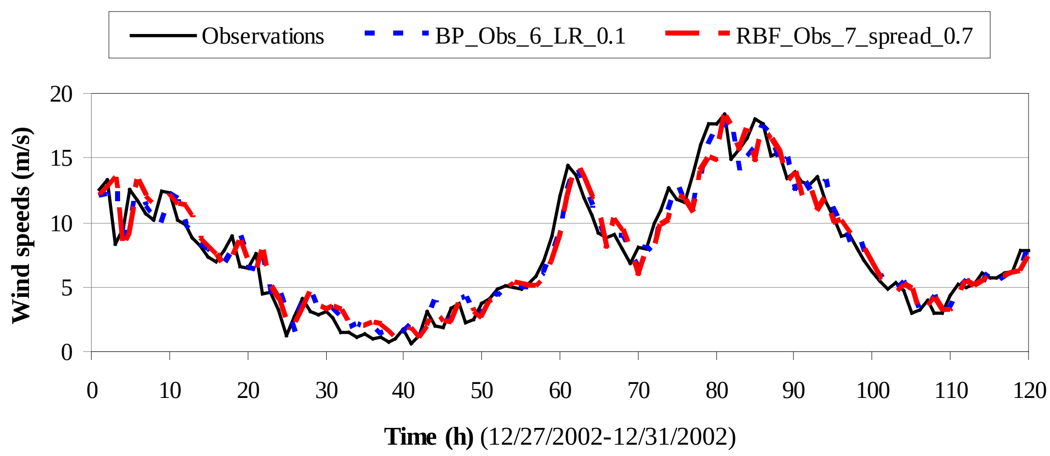


Figure 12. Forecasts of the models generating the smallest/largest RMSE (Hann).

For purpose of illustration, the forecasts time series from the BP and RBF models that generate the smallest and the largest RMSE values, respectively, are plotted in Figure 12. In the meantime, Figure 13 provides the time series plots of the corresponding forecast errors for the two specific models. It can be observed from both figures that the BP model performs better than the RBF model in terms of RMSE (as well as MAE/MAPE) for the total 120-time-point forecasts. The forecasts from the BP model are more consistent with the observations compared to the forecasts from the RBF model. However, at some specific time points, the RBF model is still able to generate smaller forecast errors, e.g., around time point 85. This verifies that the forecasting performance of ANN models changes with not only time but also the selected prediction lengths.

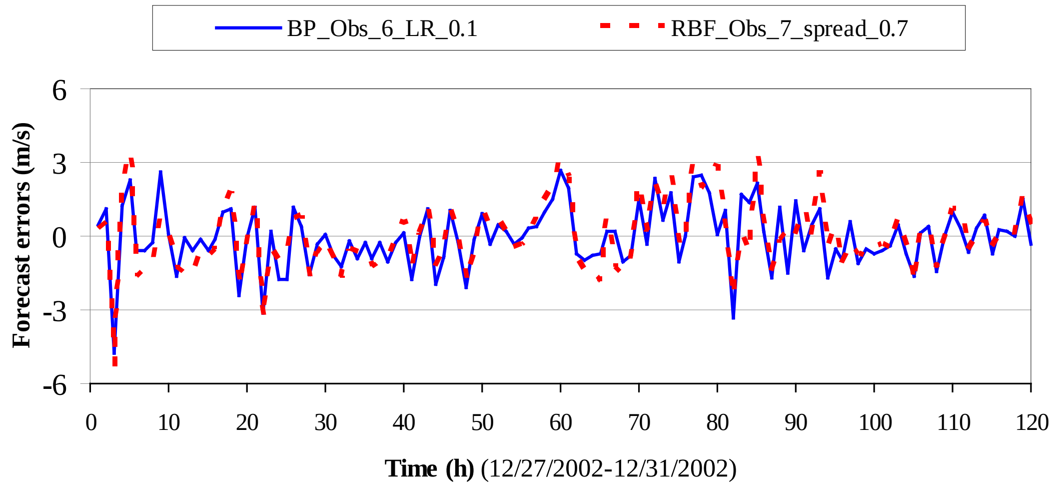
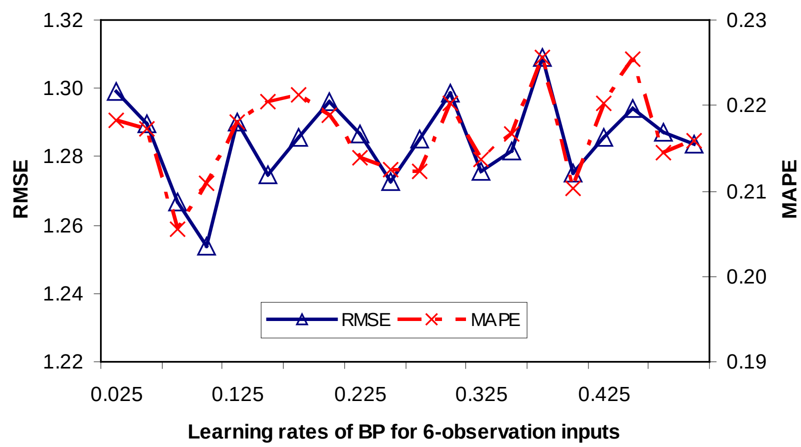


Figure 13. Forecast errors of the models having the smallest/largest RMSE (Hann).

Besides, the forecasting results confirm that different learning rates or spread constants, as well as different number of inputs, affect the forecasting performance in terms of MAE, RMSE or MAPE. Figure 14 only shows the influences of learning rates and observation number on the performances of BP models. In view that MAE follows the same trend as MAPE, only RMSE and MAPE are presented in the figure. It can be observed that with 6-observation inputs, the effect of learning rate on RMSE is not as significant as that on MAPE. The relative variation of MAPE is close to 10% in this case study. On the other hand, while adopting the same fixed learning rate of 0.1, the influences of using different number of observations as inputs on both RMSE and MAPE are relatively less significant. These observations indicate that while building the ANN forecasting models, their model parameters should be cautiously determined.



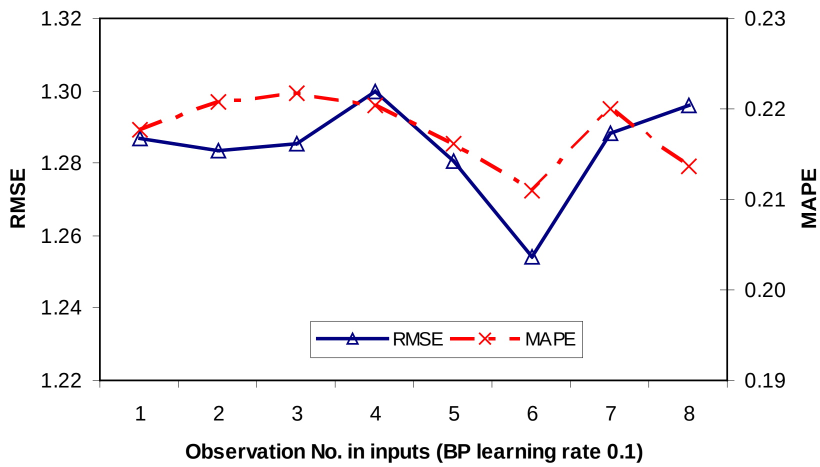


Figure 14. Effects of learning rates and input numbers on forecasting accuracy (Hann).

3.4.3. Forecasting results for the site of Kill

The forecasting results as well as the model convergence time are summarized in Table 8. Similarly, for the purpose of brevity, only the ANN models that generated either the smallest MAE, or RMSE, or MAPE values are presented.

The forecasting results for the site of Kill further confirmed the findings concluded in last subsection. It can be seen from Table 8 that different types of models demonstrate different forecast performances for site Kill. The RBF model trained with 5 previous observations at a 0.6 spread performs the best among the seven optimal models in term of RMSE (1.519), whereas the performance of BP model trained with 6-previous-observation inputs at a learning speed of 0.175 can be deemed as the best one in terms of both MAE(1.137) and MAPE(0.180). The best two of all tested ADALINE models, however, perform not so well as the best BP and RBF models. Meanwhile, it can also be observed that, among all the tested models of each specific ANN type, the best performing ones are not consistent with respect to different performance metrics. For example, the BP model adopting 7 previous observations as inputs and trained at a 0.545 learning rate performs the best in terms of RMSE, whereas the model adopting 6 previous observations and trained at a learning rate of 0.175 can be deemed as the best in terms of MAE.

Table 8. Optimal ANN models (Kill).

Best model	MAE	RMSE	MAPE	CT(s)
BP_Obs_6_LR_0.175	*1.137*	1.530	*0.180*	3.412
BP_Obs_7_LR_0.475	1.140	1.525	0.180	2.336
RBF_Obs_1_spread_0.7	1.157	1.534	0.185	0.047
RBF_Obs_5_spread_0.6	*1.175*	*1.519*	0.186	0.045
RBF_Obs_5_spread_0.7	1.169	1.522	0.184	0.046
LNN_Obs_1_lr_1.8e-6	1.157	1.538	0.186	*1.405*
LNN_Obs_2_lr_9.2e-8	1.168	*1.557*	0.185	1.436

The forecast time series from the RBF and ADALINE models that generated the smallest and the largest RMSE values, respectively, are shown in Figure 15. In this case, the difference is really minor almost across the entire time span. Similarly, Figure 16 shows the time series plots of corresponding forecast errors. Again, it can be observed that the performance advantage of the RBF model is not that significant. Although it performs slightly better than the ADALINE model in terms of RMSE as well as MAE/MAPE for the total 120-time-point forecasts, the ADALINE model still can produce smaller forecast errors at some specific time points, e.g., at time points 72 and 73.

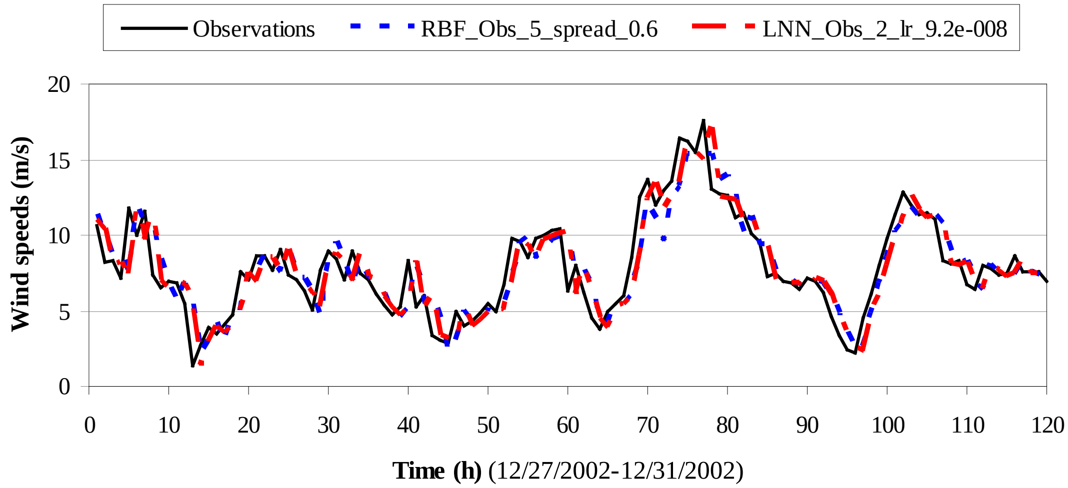


Figure 15. Forecasts of the models generating the smallest/largest RMSE (Kill).

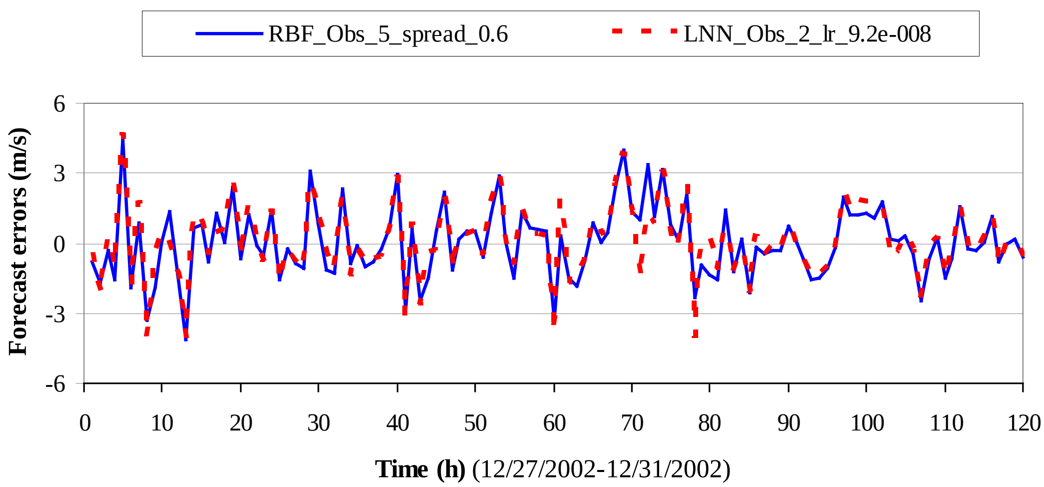


Figure 16. Forecast errors of the models having smallest/largest RMSE (Kill).

3.5. Discussion

The forecasting results presented above verify that the time series based models investigated in this dissertation, both the ARIMA models and the ANN based models, can provide plausible forecasts in their one-hour-ahead wind forecasting applications. Especially, the results demonstrate the truth of no universally 'best' model and the necessity of post-processing the forecasts from different plausible models.

By comparing Table 6 and Table 7, it can be seen that for the same dataset and application case, the ANN based models have demonstrated no significant advantage over the ARIMA models. Considering the complexity of modeling process, the ARIMA model could be a good choice in such application cases. However, ANN methods have more flexibility and robustness. For example, it can be directly applied without considering the linear or nonlinear trend of the time series.

The forecasting results also verify that artificial neural networks trained with different inputs or at different learning rates demonstrate varying accuracy in performing one-hour-ahead wind speed forecasting. For instance, for site Kill, the RBF models (shown in Table 8) trained with 1 and 5 previous observations produce inconsistent MAE, RMSE and MAPE values, although both models are trained with the same spread of 0.7. On the other hand, the two RBF models trained with same inputs (5 previous observations) but with different spread rates also demonstrate slightly different performances. Therefore, while building the ANN based forecasting models, factors such as model inputs and learning/spread rates, should be properly determined since this decision could directly affect the forecasting accuracy.

By observing both Table 7 and Table 8, it can be observed that different types of ANN models could demonstrate significantly different forecasting accuracies. This observation confirms that multiple types of ANN models should be evaluated and compared before the most suitable type can be determined. Nevertheless, the complicated issue is that different evaluation metrics often give inconsistent ranking among candidate models. This presents a major challenge on which metric(s) to be adopted in practice. Some publications can be found in developing a single general performance index by combining different evaluation criteria for measuring the forecasting accuracy so that the above-mentioned confusion can be avoided in the process of model selection. Meanwhile, it can be seen that, among all tested models, the optimal one selected for a specific site might not be suitable for another site. It is thus not recommended to employ only one type of ANN model in forecasting the wind speed at different sites.

Actually, as argued by Sanchez (2008), a final single forecast that could take advantage of a set of plausible forecasts has to be produced in many practical situations of the wind energy industry. For example, forecasts from alternative forecast agencies should be used since there is not a superior agency. Meanwhile, the forecast agencies themselves should perform forecasting tasks for the client by adopting alternative models or procedures. Therefore, it is apparent that an efficient forecast combination procedure might be of great importance for wind speed forecast. In this case, the Bayesian model averaging (BMA) method (Wasserman, 2000), an adaptive model combination method, might be used for post-processing different plausible forecasts to form a single forecast. This is the motivation for the research of the next chapter.

4. BMA-BASED WIND DISTRIBUTION AND SHORT-TERM FORECASTING

As suggested in the previous chapter, one solution to the challenge of forecasting model selection could be adopting some advanced algorithms to utilize the information from different forecast time series to generate one single forecast time series. With its demonstrated success in many other fields, BMA is determined to be a promising choice and thus investigated in this chapter. The BMA algorithm is firstly introduced in Section 4.1, and it is then investigated in estimating the long-term wind distribution in Section 4.2. This method is further employed for short-term wind speed forecasting practice, and a two-step adaptive forecasting method is thus proposed and tested in Section 4.3. The forecasting results demonstrate that the proposed methodology not only is an enhancement for reliable wind speed forecast using ANN models, but also opens up the opportunity for utilizing the information from any other types of forecasting models.

4.1. Bayesian Model Averaging

4.1.1. BMA algorithms

As a statistical procedure to infer consensus forecasts, the BMA method weighs individual forecasts based on their posterior model probabilities, with the better-performing forecasts receiving larger weights than the worse-performing ones. As the result, it generates an averaged single model, especially in cases where more than one models have a non-negligible posterior probability (Wasserman, 2000; Congdon, 2006).

A model space M composed of J distributional models M_j $(j=1, 2, \ldots, J)$ is considered for forecasting y. Let D denote the training data of observations and f_j the forecast value of model j. The probability density function (PDF) of the BMA probabilistic forecasts of y can be represented as follows,

$$p(y \mid D) = \sum_{j=1}^{J} w_j \, p(y \mid f_j, D). \tag{14}$$

For a particular model space, this is actually an average on the posterior distributions, $p(y \mid f_j, D)$, of the component models, weighted by their posterior probabilities, $w_j = P(f_j \mid D)$, where $\sum_{j=1}^{J} w_j = 1$ always holds. The posterior model probability $P(M_j|D)$, which is also known as the likelihood of M_j being the correct forecasting model given the observations data D, reflects how well this particular model matches the observations of D. For the given dataset D, the posterior probability for any component model M_k in the model space is calculated by

$$P(M_k|D) = \frac{P(M_k)P(D|M_k)}{\sum_{j=1}^{J} P(M_j)P(D|M_j)}, \tag{15}$$

where the marginal likelihood of each model is calculated by

$$P(D|M_k) = \int p(D \mid \theta_k, M_k)\, p(\theta_k \mid M_k) d\theta_k, \tag{16}$$

where θ_k is the vector of parameters in model M_k.

The posterior mean of the BMA forecasts can then be calculated by

$$E[y \mid D)] = \sum_{j=1}^{J} p(f_j \mid D) \cdot E[p_j(y \mid f_j, D)] = \sum_{j=1}^{J} w_j f_j, \tag{17}$$

and correspondingly, the variance of the BMA forecasts can be calculated by

$$Var[p(y \mid D)] = \sum_{j=1}^{J} w_j \left(f_j - \sum_{i=1}^{J} w_i f_i \right)^2 + \sum_{j=1}^{J} w_j \sigma_j^2, \tag{18}$$

where σ_j^2 is the variance associated with the model forecast f_j given the observation data of D.

4.1.2. Applications procedures of BMA

Generally, a proper set of candidate models must be selected first. This is then followed by a reliable and adaptive combination as explained above. The basic idea and general application procedures of BMA method is illustrated in Figure 17.

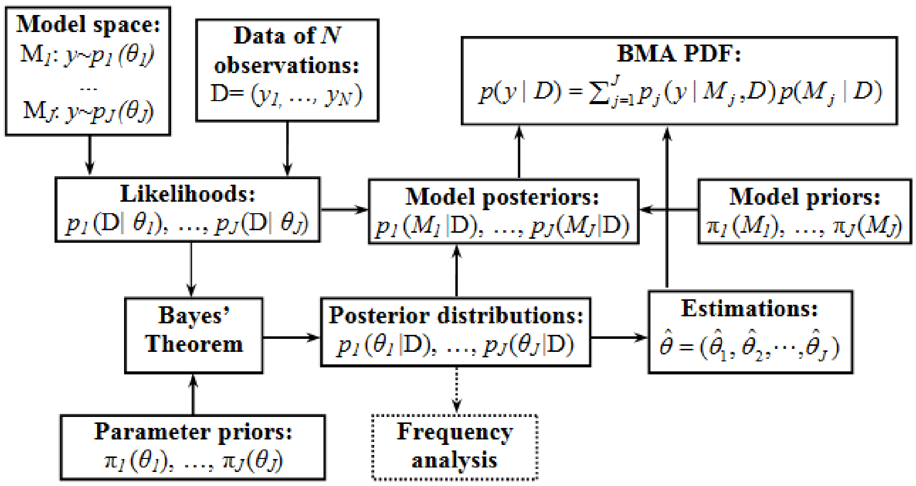


Figure 17. Application algorithm of BMA.

While the BMA method is theoretically attractive, two technical challenges exist for its practical application. The first challenge is how to properly select a set of candidate models. One simple approach is employing the complete set of models when the number of component models is limited and their structures are not too complex. An alternative approach is selecting a subset of plausible models to avoid using obviously poorly performing models (Gibbons et al., 2008). The criteria such as Akaike Information Criterion (AIC) (Akaike, 1974), Bayesian Information Criterion (BIC) (Kass and Wasserman, 1995) can be used in this case. The second challenge is the difficulty in calculating the marginal model likelihood. The marginal model likelihood may be analytically intractable, especially in many cases where no closed form integral is available. Therefore, a number of approximation methods have been proposed such as Laplace approximation, AIC, and BIC. These approximations can be obtained either from the output of common statistical routines (Laplace) or by hands (AIC and BIC). Recently, several Monte Carlo (MC) numerical methods have been developed for computing either marginal likelihoods or ratios of marginal likelihoods such as importance sampling, harmonic mean estimator, and CHIB's method (Gelman, Carlin, Stern, and Rubin, 2003). Monte Carlo integration draws samples from the required distribution and then generates sample averages to approximate expectations. The popular Markov Chain Monte Carlo (MCMC) approach draws these samples by running a properly constructed Markov chain.

4.1.3. MCMC sampling

The basic idea of MCMC is constructing a sampler which simulates a Markov chain converging to the posterior distribution. After a large number of iteration steps, the chain converges and its state can then be used as a sample from the desired distribution. Among many available MCMC algorithms, the Gibbs sampler and the Metropolis–Hastings algorithm are two simplest and yet most popular methods.

Gelfand and Smith (1990) suggest the Gibbs sampling approach for Bayesian computation. Let $p(\theta)$ denote the probability density of unknown parameter vector θ of interests, the key idea of Gibbs sampling is to generate a Markov chain and update the component of (θ_1, …, θ_k) in turn by drawing from the following full conditionals,

$$\theta_k^{(t)} \sim f_{\theta_k|\theta_{-k}}(\cdot \,|\, \theta_1^{(t)}, \cdots, \theta_{k-1}^{(t)}, \theta_{k+1}^{(t-1)}, \cdots, \theta_K^{(t-1)}). \tag{19}$$

This scheme is a Markov chain, with equilibrium distribution $p(\theta)$. Thus, for a large enough t, $\theta^{(t)}$ can be viewed as a simulated observation from $p(\theta)$. That is, provided a suitable "break-in" time for convergence, $\theta^{(t)}$, $\theta^{(t+1)}$, $\theta^{(t+2)}$, … can be handled as dependent samples from $p(\theta)$.

Metropolis–Hastings algorithm is a generalized form of Gibbs sampler. At time t, a candidate point β is firstly sampled from a proposal distribution $q(\cdot|\theta_t)$. This candidate β might then be accepted as $\theta^{(t+1)}$ with probability

$$\alpha(\theta, \beta) = \min\left[\frac{p(\beta)\, q(\theta_t \mid \beta)}{p(\theta_t)\, q(\beta \mid \theta_t)}, 1\right]. \tag{20}$$

If the candidate β is accepted, the next state becomes $\theta^{(t+1)} = \beta$. Otherwise, $\theta^{(t+1)} = \theta^{(t)}$. The proposal distribution is arbitrary and the equilibrium distribution of the chain will be $p(\theta)$ given the chain is irreducible and aperiodic.

Constructing a Markov chain with the desired properties is usually not difficult. It is hard, however, to determine how many steps are needed for the chain to converge to the stationary distribution within an acceptable error range. A simple method for convergence diagnostics is observing the history plotting of the time series for each quantity of interest or the plotting of auto-correlation functions. An alternative is using the convergence measures such as the Brooks-Gelman-Rubin (BGR) statistic R (Brooks and Gelman, 1998), which can be calculated as follows

$$\hat{R} = \frac{m+1}{m}\left(\frac{n-1}{n} + \frac{B}{W}\right) - \frac{n-1}{mn}, \tag{21}$$

where W denotes the mean of the empirical variance for each of multiple (m) MCMC chains (each with n different samples) with the “break-in” period excluded and the chain thinned, B denotes the variance of the mean across all the chains. If the chains converge, both estimates should be unbiased, i.e., $B=W$, and thus R will be equal to 1.

4.1.4. Marginal likelihood calculation

MCMC method enables one to draw samples from any complex unknown distributions. Once the samples of each model parameters are drawn from their posterior distributions, the marginal likelihood can then be calculated by many different methods. In view that the models tested are not complex and that the observations are large, we adopt the following harmonic mean of the likelihood values, originally proposed by Newton and Raftery(1994).

$$P(D|M_k) \approx \left\{ \frac{1}{G} \sum_{i=1}^{G} \frac{1}{p(D \mid \theta_{ik}, M_k)} \right\}^{-1}, \tag{22}$$

where the values of parameter vector θ_{ik} are randomly sampled from the posterior density $p(\theta_k \mid D, M_k)$ via MCMC sampling and G is the total number of samples θ_{ik} for model M_k.

By drawing enough samples from the Monte Carlo Markov chain, the integral for the desired distribution could then be approximated. In the actual computation where a large number of observations are considered, the following strategy is adopted to avoid numerical overflows or underflows in computation. Denote

$$S_k = \frac{1}{G} \sum_{i=1}^{G} \frac{1}{p(D \mid \theta_{ik}, M_k)} = \frac{1}{G} \sum_{i=1}^{G} R_{ik}, \tag{23}$$

The values of $\log(R_{ik}) = \log\left\{ \left[p(D \mid \theta_{ik}, M_k) \right]^{-1} \right\}$ are stored separately for each sampled value of θ_{ik}. By introducing a dummy variable, $c_{\max}$, we have

$$S_k = \frac{1}{G} \sum_{i=1}^{G} R_{ik} = \frac{1}{G} \sum_{i=1}^{G} \exp[\log(R_{ik}) - c_{\max}] \exp(c_{\max}), \tag{24}$$

$$\log(S_k) = \log\left\{ \frac{1}{G} \sum_{i=1}^{G} \exp[\log(R_{ik}) - c_{\max}] \right\} + c_{\max}, \tag{25}$$

$$\log[\hat{p}(D \mid M_k)] = -\log(S_k). \tag{26}$$

Given dataset D, the posterior probability of each element M_k can then be calculated by

$$P(M_k|D) = \frac{\exp(-\log(S_k))}{\sum_{j=1}^{J} \exp(-\log(S_j))} = \frac{\exp(-\log(S_k) - c_{\max})}{\sum_{j=1}^{J} \exp(-\log(S_j) - c_{\max})}, \tag{27}$$

where $c_{\max}$ denotes the maximum of $-\log(S_j)$, namely, $c_{\max} = \max\{-\log(S_j),\ j = 1,\cdots J\}$.

4.2. BMA-Based Long-Term Wind Probabilistic Prediction

By combining BMA and Markov Chain Monte Carlo (MCMC) sampling methods, a new approach is proposed for deriving a reliable and robust approximation of wind speed. The derived BMA probability density function (PDF) of the wind speed is an average of the model PDFs included in the model space weighted by their posterior probabilities over the sample data. The MCMC method provides an effective way for numerically computing marginal likelihoods, which are essential for obtaining the posterior model probabilities. The approach is applied to multiple sites with high wind power potentials in North Dakota, as listed in Table 9. The mean hourly wind speed data at these sites are collected over two years (2001 and 2002).

Table 9. Geographical information of observation sites and their wind speeds.

Site	Latitude (North)	Longitude (West)	Elevation (m)	Wind speed at 10 m above ground level (m/s)		
				Mean	Min	Max
Alfred (Alf)	46º35'15"	99º00'46"	631	6.248	0.005	26.030
Green River (Gre)	47º04'05"	102º55'38"	818	5.437	0.005	21.977
Olga (Olg)	48º46'48"	98º02'16"	475	4.954	0.005	19.125
Ray/Wheelock (Ray)	48º15'57"	103º11'52"	750	5.872	0.005	21.488
Powers Lake (Powe)	48º23'32"	102º37'20"	777	8.008	0.362	26.442
Hannaford (Hann)	47º19'39"	98º12'34"	448	7.404	0.362	25.562
Kulm (Kulm)	46º17'56"	98º51'58"	600	8.786	0.362	23.859
Luverne (Luve)	47º19'47"	97º54'27"	460	7.632	0.362	23.412
West Finley (West)	47º30'02"	97º52'22"	469	7.777	0.362	23.769

4.2.1. Selection of candidate models

In order to implement BMA method, a proper set of candidate models must be selected at the beginning. In the case of estimating wind speed distribution, many conventional statistical distribution models are shown to be viable options according to literature. These model forms are listed in Table 10.

Table 10. Candidate distribution models and the prior of their parameters.

Model	Probability density function	Priors of model parameters
Weibull (Weib-3)	$\frac{\alpha(x-\gamma)^{\alpha-1}}{\beta^{\alpha}}\exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)$, γ<$x$<∞; 0 otherwise	$\alpha, \beta \sim \Gamma(0.5, 0.5)$ $\gamma \sim uniform(-20, \min(x))$
Weibull (Weib-2)	$\frac{\alpha(x)^{\alpha-1}}{\beta^{\alpha}}\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)$, γ<$x$<∞; 0 otherwise	$\alpha, \beta \sim \Gamma(0.5, 0.5)$
Rayleigh (Rayl)	$\frac{x}{\beta^2}\exp\left(-\frac{1}{2}\left(\frac{x}{\beta}\right)^2\right)$, 0<$x$<∞; 0 otherwise	$\beta \sim \Gamma(0.5, 0.5)$
Lognormal (LogN-3)	$\frac{1}{(x-\gamma)\sigma\sqrt{2\pi}}\exp\left(\left(-\frac{1}{2}\right)\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)^2\right)$ γ<x<∞; 0 otherwise	$\mu \sim N(0, 20)$ $1/\sigma^2 \sim \chi_1^2$ $\gamma \sim uniform(-20, \min(x))$
Inverse Gaussian (IG-3)	$\left(\frac{\lambda}{2\pi(x-\gamma)^3}\right)^{1/2}\exp\left(-\frac{\lambda(x-\gamma-\mu)^2}{2\mu^2(x-\gamma)}\right)$ γ<x<∞; 0 otherwise	$\mu \sim N(0, 20)$ $\lambda \sim \Gamma(0.5, 0.5)$ $\gamma \sim uniform(-20, \min(x))$
Gamma (Gam-3)	$\frac{1}{\beta\Gamma(\alpha)}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\exp\left(-\frac{x-\gamma}{\beta}\right)$ γ<x<∞; 0 otherwise	$\alpha, \beta \sim \Gamma(0.5, 0.5)$ $\gamma \sim uniform(-20, \min(x))$
Inverse gamma (PS5)	$\frac{1}{\beta\Gamma(\alpha)}\exp\left(-\frac{\beta}{x}\right)\Big/\left(\frac{x}{\beta}\right)^{\alpha+1}$, 0<$x$<∞; 0 otherwise	$\alpha, \beta \sim \Gamma(0.5, 0.5)$
Erlang (EL)	$\frac{1}{\beta(\alpha-1)!}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\exp\left(-\frac{x-\gamma}{\beta}\right)$ γ<x<∞; 0 otherwise	$\alpha, \beta \sim \Gamma(0.5, 0.5)$ $\gamma \sim uniform(-20, \min(x))$
Gumbel-maximum (EV)	$1/\beta\exp\left(\frac{x-\alpha}{\beta}+\exp\left(-\frac{x-\alpha}{\beta}\right)\right)$ -∞<x< ∞	$\alpha, \beta \sim \Gamma(0.5, 0.5)$

It can be seen that this list is comprehensive and covers most typical models for estimating wind speed distribution, and thus it can be regarded as a good model space. Certainly, this list still cannot be deemed as exhaustive since other potential distribution models may exist. Two-parameter Weibull distribution (Weib-2) is a special case of the general Weibull distribution which uses the default location parameter. Weib-2 is included because it is by far the most popular statistical distribution for wind speed. Rayleigh distribution is a special case of two-parameter Weibull distribution, and it is included for the same reason (Li and Shi, 2010c).

4.2.2. Determination of model priors and model parameter priors

For a given set of models, the effectiveness of the BMA approach heavily relies on the specification of the model priors and the model parameter priors. As for the model priors, it is ideal that we can quantify and include prior knowledge in model development. When little prior information is available about the relative plausibility of the models, however, a simple and reasonable method is adopting the uniform distribution as the non-informative prior to favor all models equally (Hoeting and Madigan, 1999). This method is adopted in this study considering that all models ever demonstrated their good performance in describing the distribution of wind speed and that no prior information is available before applying them for any site.

A prior distribution of the model parameter is considered proper if it satisfies the requirements for the probability distribution. It should be noted that, however, not all priors need to be proper in order to yield a proper posterior distribution for the model parameters (Congdon, 2006). Especially, to reduce the influence of the priors on the parameters, the most common and practical approach for prior specification in this context is to specify non-informative priors that allow the posterior to accumulate probability at or near the actual data-generated model. Such a posterior can serve as a heuristic device to identify promising models for BMA applications. The prior distributions for each model tested in this research are listed in the right column of Table

10. Each model parameter prior can be regarded as uninformative since little information about the parameters is available.

4.2.3. MCMC sampling and posterior probability calculation

To estimate the marginal likelihood, model parameters are sampled from their posterior distributions using the coded WinBUGS programs. WinBUGS is statistical software used for MCMC-based Bayesian modeling (http://www.mrc-bsu.cam.ac.uk/bugs/). The application procedures of using WinBUGS for MCMC sampling are as follows. To perform convergence diagnostics, each distribution model is simulated by two chains with 20,000 iterations for each chain. The convergence diagnostics can be performed by observing the history plotting and/or the plotting of Brooks-Gelman-Rubin (BGR) convergence statistic. After the convergence, another 5,000 samples are drawn for each parameter of each model from their posterior distributions, respectively. Once enough parameter samples for each model have been drawn from the posterior distributions, the marginal likelihood is calculated according to the harmonic mean estimator proposed by Newton and Raftery (1994). The advantage of this estimator is that no knowledge on the form of the posterior distribution is required.

The BGR convergence diagnosis of the MCMC simulation via WinBUGS is performed for each model and the wind speed data from each site. It is observed that all models converge approximately after 15,000 iterations. As an example, Figure 18 illustrates the convergence of simulations on the parameters of the 3-parameter Weibull distribution for site Alfred. In this figure, the dotted (blue) line represents the width of the central 80% interval of the pooled runs, the dashed (green) line represents the average width of the 80% intervals within the individual runs, and the solid (red) line is their ratio R. Convergence is obtained if R is close to 1 and with both the two widths converging to stability. It can be observed that the MCMC simulation for the parameters of 3-parameter Weibull distribution for site Alfred clearly converges after 15,000 iterations.

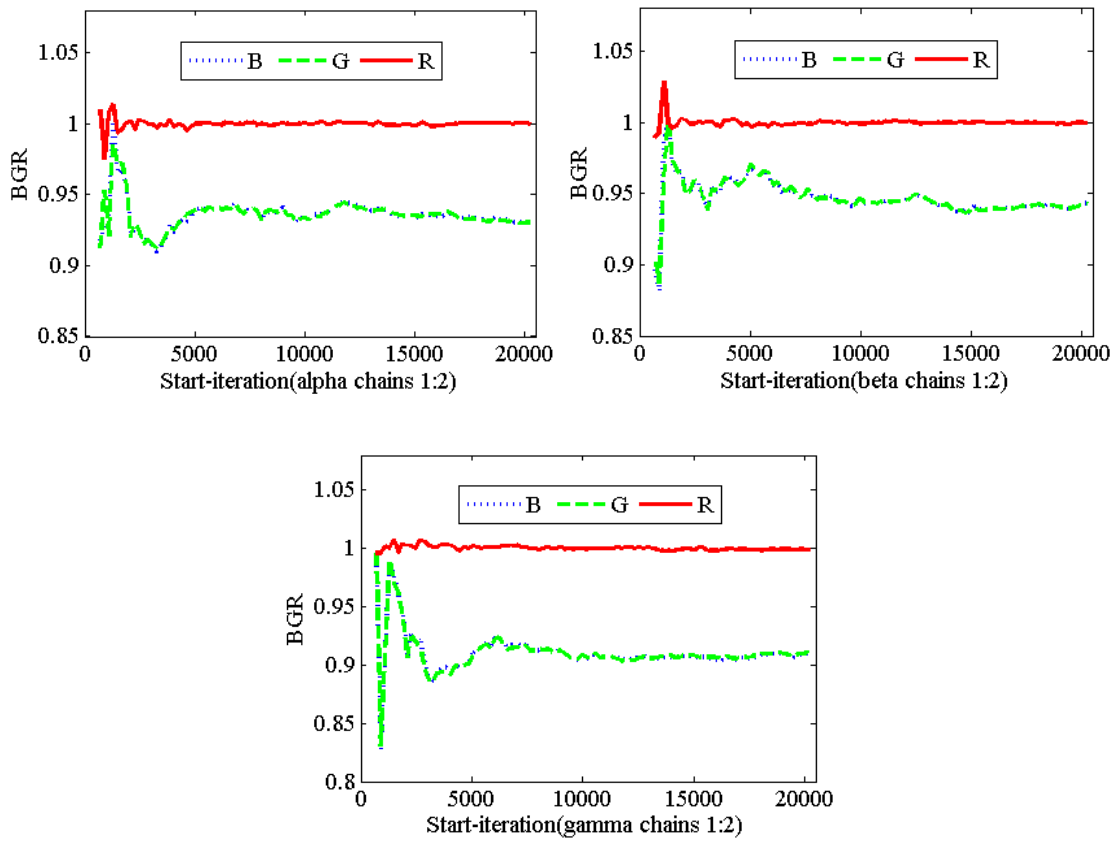


Figure 18. BGR convergence diagnostics plot of Weib-3 model parameters (Alf).

Once convergence is achieved, the simulation is run for a further number of iterations to obtain samples that can be used for posterior inference. In MCMC simulation for each model, after the 15,000 “break-in” iterations, the simulation converges and then another 5,000 samples are drawn for each parameter from their posterior distributions. Usually, the more samples saved, the more accurate is the posterior estimates. The Monte Carlo error for each parameter is one statistic to assess the accuracy of the posterior estimates. It is an estimate of the difference between the estimated posterior mean of the sampled values and the true posterior mean. The simulation should be run until the Monte Carlo error for each parameter of interest is less than about 5% of the sample standard deviation. It is found that this rule is satisfied by the 5,000 samples for each model parameter. For purpose of illustration, Table 11 shows the statistical results of the model parameters of the 3-parameter Weibull distribution for site Alfred. It can be seen that the standard deviations of parameter estimates are small compared with their means. Besides, the small MC errors indicate that the parameter estimates should have been stabilized. The quantity reported in the MC error column gives an estimate of the Monte Carlo standard error of the mean, which is an indicator that the chains have mixed and further simulation will not change inferences appreciably.

Table 11. Simulation results of model parameters of Weibull-3 distribution (Alfred).

Parameter	Mean	S.D.	MC error	2.5%	Median	97.5%	Sample
α	1.973	0.01272	2.636E-4	1.949	1.973	1.998	5000
β	7.094	0.03336	6.764E-4	7.03	7.094	7.158	5000
γ	-0.03281	0.01241	2.394E-4	-0.06061	-0.03172	-0.01089	5000

The kernel density estimation is a non-parametric way of estimating the probability density function of a random variable, which makes it possible to extrapolate the data to the entire population. Figure 19 illustrates the kernel density plotting of the model parameter samples for the same distribution and site. For a symmetric distribution, the mode usually equals the mean. For a skewed distribution, however, the mode is usually better than the mean or the median in capturing important information about the random variable or the population. It can be seen that, with each parameter being regarding as random variable, the posterior distribution of the location parameter is skewed to the right although the shape and scale parameter densities seem to be approximately symmetric. Therefore, it is more appropriate to use the mode as the model parameter to estimate the distributional density of each model for each site.

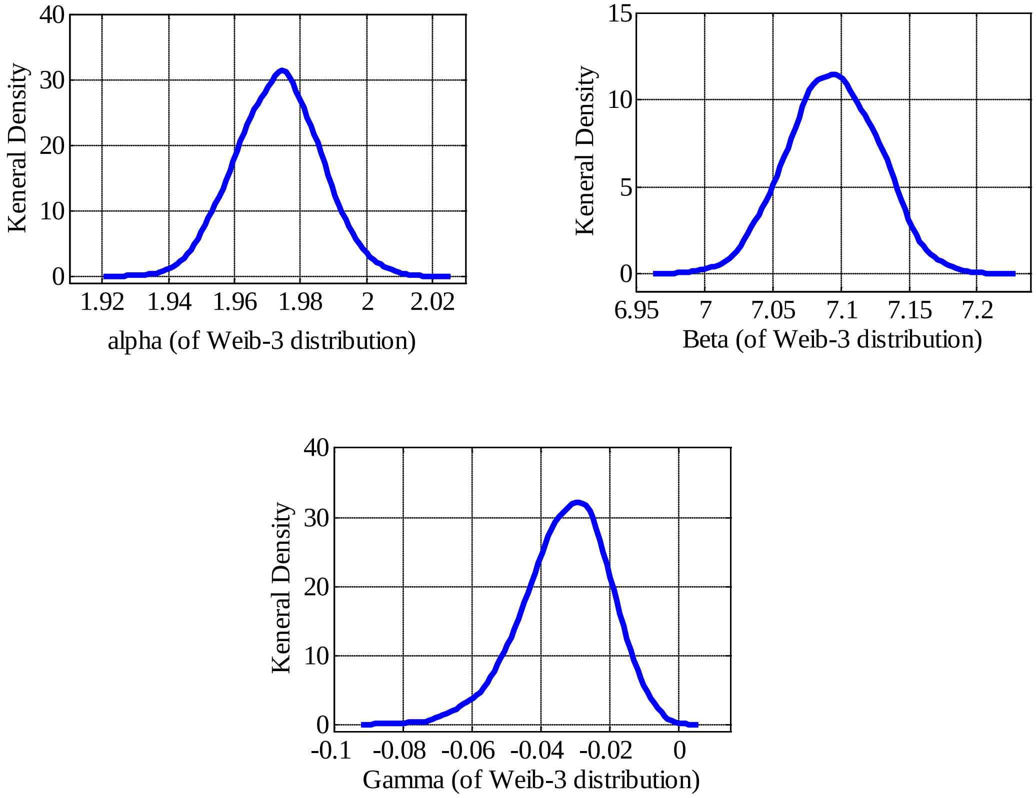


Figure 19. Posterior densities of parameter samples of Weibull-3 distribution (Alfred).

4.2.4. BMA-based prediction results

Once all these posterior samples have been generated, the marginal likelihood and the corresponding posterior probability for each model can then be calculated as explained in Section 4.1. Table 12 shows the posterior probability of each distribution model for the nine sites. The posterior probability of zero in the table means that the value is extremely small, less than 0.000001. Since a larger posterior probability means a better fit for distribution models, the results can serve the purpose of model selection - the best model should be the one with the highest posterior probability value. For example, lognormal distribution can be regarded as the best model for modeling the wind speed distribution at site Green River since its posterior probability is the highest one. The traditional χ^2 statistic is 273.5 for this numerically derived lognormal distribution, while it is larger than 400 for any other candidate distributions. This further proves lognormal distribution is the predominating distribution for describing the wind data at site Green River.

Table 12. Posterior probabilities of distribution models for selected sites.

Site	Distribution								
	Weib-2	Rayl	Weib-3	LogN-3	IG-3	Gam-3	PS5	EL	EV
Alf	0	0	0	0	0	0.941733	0	0.058267	0
Gre	0	0	0	0.999992	0.000003	0	0	0	0.000005
Olg	0	0	0	0	0	0.122127	0	0.877873	0
Ray	0	0	0	0.991982	0.005647	0.002130	0	0	0.000241
Powe	0.000017	0	0.999983	0	0	0	0	0	0
Hann	0.023673	0	0.046151	0.000002	0	0.930174	0	0	0
Kill	0.001711	0	0.998289	0	0	0	0	0	0
Luve	0	0	0	0.746703	0	0.253297	0	0	0
West	0.579722	0	0.420278	0	0	0	0	0	0

However, the table also shows that more than one model demonstrates a non-ignorable posterior probability for almost half of the sites. A good example is site Luverne. For this site, lognormal distribution has a posterior probability of 0.746703 and Gamma distribution has a value of 0.253297, indicating that both models are acceptable for describing the wind speed distribution. This can again be confirmed by the χ^2 statistics. The χ^2-square values are 198.0 and 219.8 for lognormal and Gamma distributions, respectively, while it is larger than 400 for the rest distributions. Here, the difference between lognormal and Gamma distributions is insignificant. In such cases, we must be cautious about interpreting lognormal as the best model since Gamma distribution might be preferred based on a different criterion.

The BMA PDF can now be derived by using the derived posterior model probability and the estimated posterior likelihood of each model. For purpose of brevity, Figure 20 illustrates the histograms and estimated densities for sites Alf and Hann only. In the figure, the dotted curve (blue) represents one of the insignificant models, the dashed curve (green) is the probability density curve of the "best" model, and the solid line (red) is the derived BMA PDF.

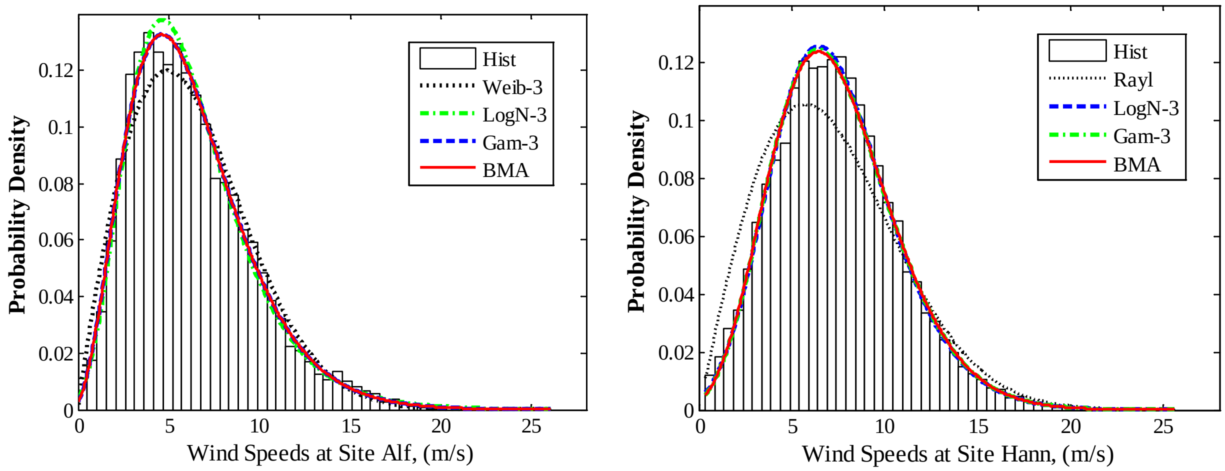


Figure 20. Histograms and estimated probability density plots for sites Alf and Hann.

It can be seen that the previous numerically-derived 'best' model for each site can describe the distribution of the real observations well. However, it can also be observed that the wind speed distributions are skewed at various sites with different 'best' models, but the derived BMA PDF always agrees well with the wind speed distribution for each site. More importantly, the BMA PDF always completely overlaps with the distributions with non-negligible posterior probabilities, which makes them nearly indistinguishable in the plots. The overlapping in the plots seems to be an issue for visualization purpose. However, it indicates that the derived BMA PDF is accurate, and that the BMA approach is universally reliable and robust in modeling the wind speed distributions.

The results also indicate that no distribution universally outperforms others for all North Dakota sites considered in this study. For instance, for site Alfred, Gamma distribution, with a 0.941733 posterior probability, is superior over all the rest models while for site Green River, it is substituted by the log-normal distribution which has a posterior probability of 0.99992. On the other hand, BMA provides a unified approach to solve the dilemma. The BMA model derived from weighing different models with their posterior probabilities can always match the real wind speed distribution well since it automatically considers the variability between different models.

As found in literature, the studies of estimating wind speed distribution usually follow two approaches. One is that a statistical distribution form is assumed, and the parameters of the distribution are obtained and the goodness-of-fit is evaluated. The other more general approach is via selecting several possible distribution forms. The best model is determined based on some goodness-of-fit (GOF) metrics (Celik, 2004; Rogersa, Rogersb, and Manwell, 2005). However, different GOF metrics often produce contradicting ranking of fitting among the candidate distributions (Zhou, Erdem, Li, and Shi, 2010). Actually, the uncertainty of the parameters of each model can be regarded as within-model error variance, and the uncertainty of selecting the best model can be regarded as the between-model error variance. The common GOF test statistics can only evaluate the within-model variance with the between-model one being ignored. However, the variance of the BMA PDF always includes both the between-model variance and the within-model variance, both estimated from the sample data. This is a typical advantage of BMA estimation compared with any model element in the model space.

Based on the posterior probabilities of all the models, the traditional model selection is made possible and the best model for each site can be determined. Actually, once the elements of the model space are determined, there is no need to consider whether the performance of any individual candidate model is good or not. This is because its performance will be automatically considered and weighted as its posterior probability in deriving the BMA PDF. It is also unnecessary for the set of models under consideration to have common structure. Different distributions in the model space can be derived from entirely different principles. The only requirement is that they predict the same quantities of interest. To this point, any type of model forms can be included in the model space. This is another advantage of applying the BMA approach to estimate the distribution of wind speeds.

4.3. Two - Step Method for Short - Term Wind Speed Forecasting

In view that BMA algorithm has been successfully applied and demonstrated its advantages in estimating the long-term wind speed distribution, it is of great interest to further investigate into its feasibility and reliability in performing probabilistic short-term wind forecasting. Based on the BMA algorithm and the three aforementioned ANN models (BP, RBF, and ADALINE), a robust two-step methodology for forecasting hourly wind speeds is thus proposed. For purpose of comparison, the hourly wind speed data adopted in Chapter 3 are used to demonstrate the effectiveness of the proposed modeling method. The three performance metrics adopted in Chapter 3, namely, MAE, RMSE, and MAPE, are again employed to evaluate the forecasting accuracy. The results indicate that, although the performances of different ANN-based models are different in performing one-hour-ahead wind speed forecasting for different sites or under different evaluation metrics, the proposed Bayesian combination method can always provide adaptive, reliable and comparatively accurate forecast results. It also provides a unified approach to tackle the challenging model selection issue in the field of wind speed forecasting.

4.3.1. Procedures of two-step wind speed forecasting

The application procedures of the entire analysis process are shown in Figure 21. The first major step is to obtain best time series forecasts from each type of ANN models in terms of each of the three performance metrics. The second step is to combine the multiple time series forecasts into one single time series by adopting Bayesian combination algorithm.

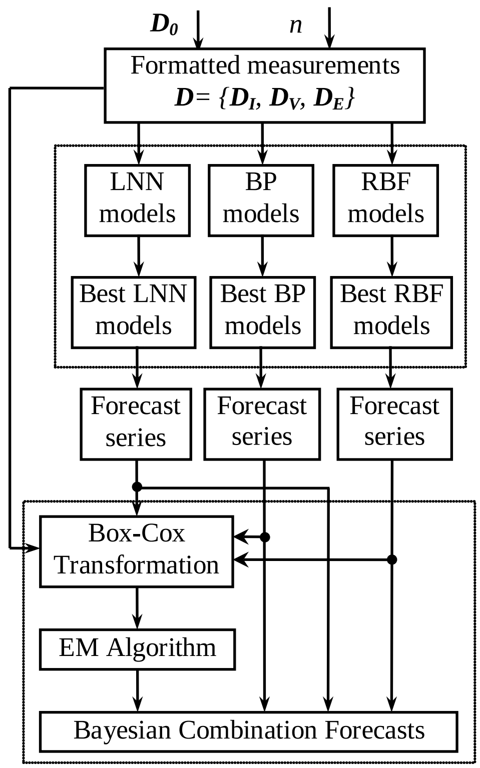


Figure 21. BMA-based two-step wind forecasting procedure.

The first step is the same as mentioned in Chapter 3, the data collected at each site, D_0 are first preprocessed and transformed into the specific format according to n, the number of previous hourly wind speeds in each input vector. The formatted data are further divided into training dataset D_T, validation dataset D_V, and testing or evaluation dataset D_E. All the models of ADALINE, BP, and RBF network models are then trained and evaluated with the corresponding datasets at various learning rates. After that, each model is applied to forecast the next hour wind speeds corresponding to the testing data and the values of MAE, RMSE, and MAPE are calculated. The models of each type, which generate the smallest MAE, RMSE, or MAPE, are selected. They are regarded as the best model among the tested models of the same type.

In the second step, the forecast time series of each best model, together with the observation time series in the testing dataset, are first post-processed via Box-Cox transformation so that each of these transformed time series would be approximately normally distributed. After that, the EM algorithm is applied to each observation-forecast combined time series for estimating the corresponding model parameter vector $\theta = [w_j, \sigma^2{}_j]$ (j=1,2,…, J), where w_j represents the posterior probability of the jth forecasting model being the best one among the model space, and σ_j^2 is the estimated average variance of the jth model forecast, as presented and explained in Chapter 3. The Bayesian combination postprocessing is thereafter performed and the corresponding Bayesian forecasting values, forecasting variances as well as the values of three performance metrics are calculated, respectively.

4.3.2. Expectation-maximization and Box-Cox transformation

As indicated in Chapter 3, although the Markov Chain Monte Carlo (MCMC) method can simulate any complex probability distribution, it usually involves high computational complexity that is comparatively time-consuming (Li and Shi, 2010c). As an alternative, the Expectation-Maximization (EM) algorithm is also powerful in estimating the posterior probabilities of component models. Especially, compared with the MCMC method, the EM algorithm is significantly more efficient although it functions well mainly for normal or approximately normal distributions (McLachlan and Krishnan, 1997; Liu, 2001). In view of the advantage of computational efficiency, the EM method is employed in this study to estimate the model parameter vector $\theta = [w_j, \sigma^2{}_j]\,(j = 1, \cdots J)$ via the following log-likelihood function,

$$\ell(\theta) = \log\left(\sum_{j=1}^{J} w_j p_j (y \mid f_k, D) \right). \tag{28}$$

The EM iterative algorithm alternates between the Expectation (E) step and Maximization (M) step repeatedly until certain convergence criteria are satisfied. EM algorithm employs an unobservable or latent variable, $z_{j,t}$, which will be equal to 1 if the *j*th model forecast is the best one at time point *t*; otherwise zero. At any random time point *t*, only one $z_{j,t}$ can be 1 and the rest should be zero (McLachlan and Krishnan, 1997).

In the E step of the (*k*+1)th iteration, the $z_{j,t}$ are estimated as given in Eq. (29), based on current estimate of parameter vector θ,

$$\hat{z}_{j,t}^{(k+1)} = \frac{w_j^{(k)} p_j^k (y_t \mid f_{j,t}, \sigma_j^k)}{\sum_{m=1}^{J} w_m^{(k)} p_m^k (y_t \mid f_{m,t}, \sigma_m^k)}. \tag{29}$$

In the M step of the (*k*+1)th iteration, the parameter vector θ, the weights and the variances, are estimated as follows based on current estimates of $z_{j,t}$,

$$w_j^{(k+1)} = \frac{1}{T} \sum_{t=1}^{T} \hat{z}_{j,t}^{(k+1)}, \tag{30}$$

$$\sigma_j^{2(k+1)} = \frac{\sum_{t=1}^{T} \left[\hat{z}_{j,t}^{(k)} \cdot (y_t - f_{j,t})^2 \right]}{\sum_{t=1}^{T} \hat{z}_{j,t}^{(k)}}. \tag{31}$$

The detailed description of the EM algorithm can be found in (McLachlan and Krishnan, 1997). However, since the distribution of wind speed is usually skewed, as verified in Section 4.2, the transformation of wind speed data to a Gaussian distribution is required before the EM algorithm can be applied. In this case, the Box-Cox transformation is a straightforward and simple procedure for non-normality correction (Carta, Ramírez, and Velázquez, 2009), and it is thus adopted in this study. The Box-Cox transformation algorithm can be briefly presented by the following equations:

$$\begin{cases} y^{(\lambda)} = \dfrac{y^{\lambda} - 1}{\lambda}, (\lambda \neq 0) \\ y^{(\lambda)} = \ln(y), \ (\lambda = 0) \end{cases}, \tag{32}$$

where y is the wind speed data in this study, and the optimal value of λ can be determined by choosing the value that maximizes the following log-likelihood function for each time series,

$$f(y,\lambda) = -\frac{n}{2}\cdot\ln\left[\sum_{i=1}^{n}\frac{(y_i(\lambda)-\bar{y}(\lambda))^2}{n}\right] + (\lambda-1)\cdot\sum_{i=1}^{n}\ln(y_i) \quad . \tag{33}$$

4.3.3. Forecasting results for the site of Kill

The three types of ANN forecasting models are built in the same way as explained in Section 3.3. The model structures are exactly the same as those employed previously. For each type of neural networks (i.e., BP, RBF, and ADALINE), the models are first trained with specific *n*-observation input vectors at a specific learning rate or spread constant, and the final performances are calculated in terms of the metrics (i.e. MAE, RMSE, and MAPE). For the purpose of brevity, only the ANN models of each type that generated either the smallest MAE, or RMSE, or MAPE values are adopted. Moreover, these best NN models are the component models for the sequential BMA approach. The weight and variance of each component model are obtained, and then the performance of the BMA time series forecast is evaluated. The results are summarized and shown in Table 13.

It can be seen that different types of models demonstrate different forecast performances for site Kill. The RBF model trained with 5 previous observations at a 0.6 spread performs the best in term of RMSE (1.519), whereas the performance of BP model trained with 6 previous observations at a learning speed of 0.175 can be deemed as the best one in terms of both MAE(1.137) and MAPE(0.180). The best two of all tested ADALINE models, however, perform not so well as the best BP and RBF models. Meanwhile, it can also be seen that, among all the models of each ANN type, the best performing ones are not consistent with respect to different performance metrics. For instance, the BP model using 7 previous observations and trained at a

0.545 learning rate performs the best in terms of RMSE, whereas the model trained with 6 previous observations at a learning rate of 0.175 can be deemed as the best in terms of MAE.

Table 13. Best ANN models and their parameters in combined forecast (Kill).

Best model	MAE	RMSE	MAPE	Weight	Variance
BP_Obs_6_LR_0.175	*1.137*	1.530	*0.180*	0.1336	0.2367
BP_Obs_7_LR_0.475	1.140	*1.525*	0.180	0.1669	1.2335
RBF_Obs_1_spread_0.7	1.157	1.534	0.185	0.1302	0.2724
RBF_Obs_5_spread_0.6	1.175	*1.519*	0.186	0.1594	0.9093
RBF_Obs_5_spread_0.7	1.169	1.522	0.184	0.1558	0.8806
LNN_Obs_1_lr_1.8e-6	1.157	1.538	0.186	0.1275	0.2239
LNN_Obs_2_lr_9.2e-8	1.168	*1.557*	0.185	0.1267	0.1919
BMA	1.137	1.508	0.181	/	/

Regarding the BMA results from EM iterations, it can be seen that the estimated posterior probabilities (i.e., weights) are different for the seven models. The BP model trained with the inputs of previous 7 observations at a 0.475 learning rate is assigned the largest weight value of 0.1669. However, the weights of the component models do not deviate from each other too much. This indicates that all the 7 component models contribute to the final BMA forecast to certain extent. Actually, the weight estimated by EM algorithm via maximum likelihood estimation can be regarded as another model selection or performance evaluation criterion if needed, but this is not the main task of the BMA method. With the forecasts and the posterior probabilities of all the seven component models, the Bayesian combination forecast could be readily obtained. The corresponding performances of the combination forecast in terms of MAE, RMSE, and MAPE are listed in the last row of Table 13. It can be seen that the BMA model forecasts with the closest accuracy to the best model in terms of MAPE, the same accuracy in terms of MAE, even more accurate than the best model in terms of RMSE, confirming the adaptive, reliable, and comparatively accurate nature of Bayesian combination forecasts.

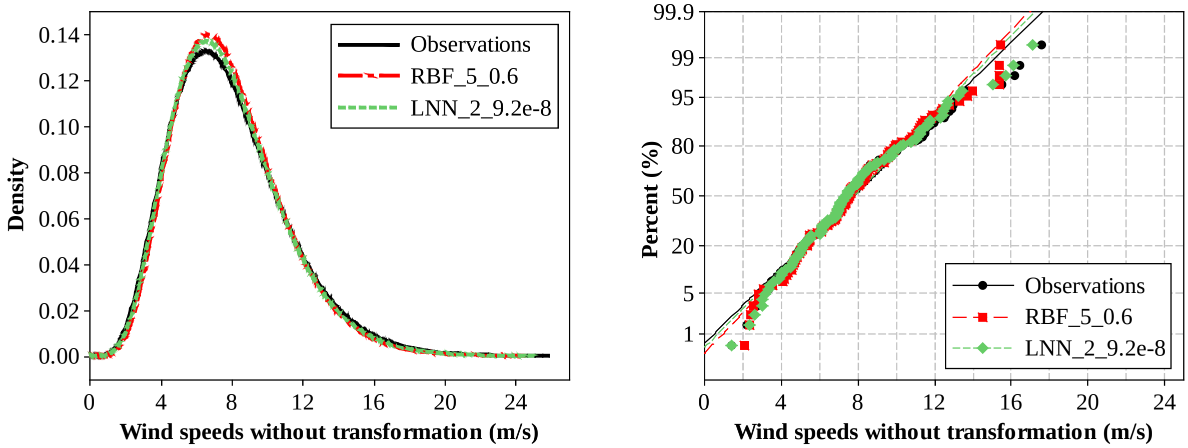


Figure 22. Distribution and normal probability plots of observations and forecasts.

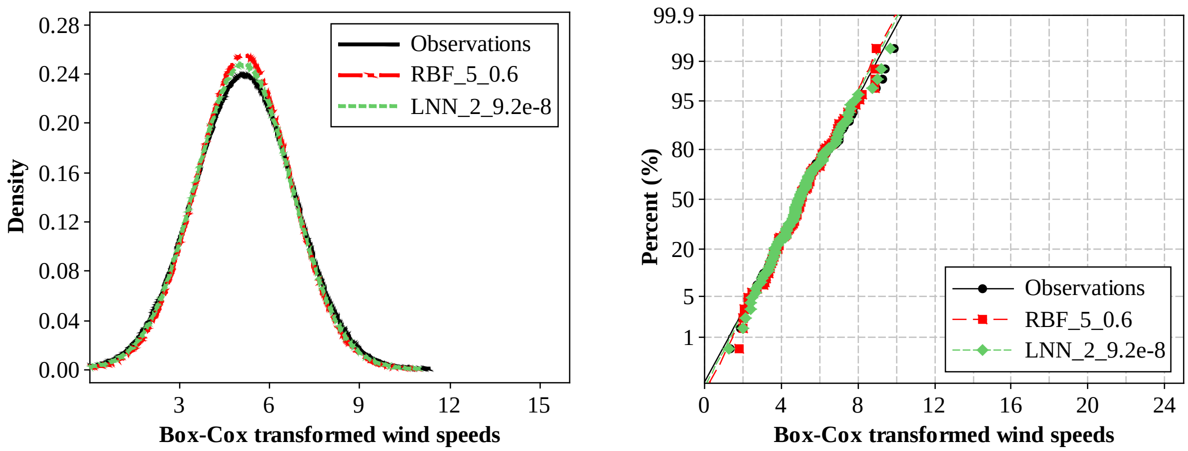


Figure 23. Distribution and normal probability plots of transformed data and forecasts.

As afore-mentioned, due to the non-Gaussian nature of wind speeds at site Kill, Box-Cox transformation is needed for each forecast time series before the BMA step. For consistency, the data transformation of the same parameters is applied to all the time series including the original observations and the forecast ones. Figures 22 and 23 demonstrate the distributions and the normal probability plots of the observation time series and the two forecast time series with the lowest and highest RMSE values, before and after the Box-Cox transformation respectively. It can be seen that the transformed data satisfactorily follow normal distribution although the original observation is skewed or non-Gaussian.

Figure 24 shows the forecasting error time series for the 120 consecutive hours from the RBF model (generating the smallest RMSE value, as shown in Table 13) and the ADALINE model (generating the largest RMSE value, as shown in Table 13), respectively. Especially, the forecasting errors of the corresponding Bayesian combination forecasts are plotted for purpose of comparison. It can be observed that, although RBF model forecasts are generally more accurate than the ADALINE model in terms of RMSE, ADALINE model still has smaller forecast errors at some specific time points, e.g., time points 11 and 12.

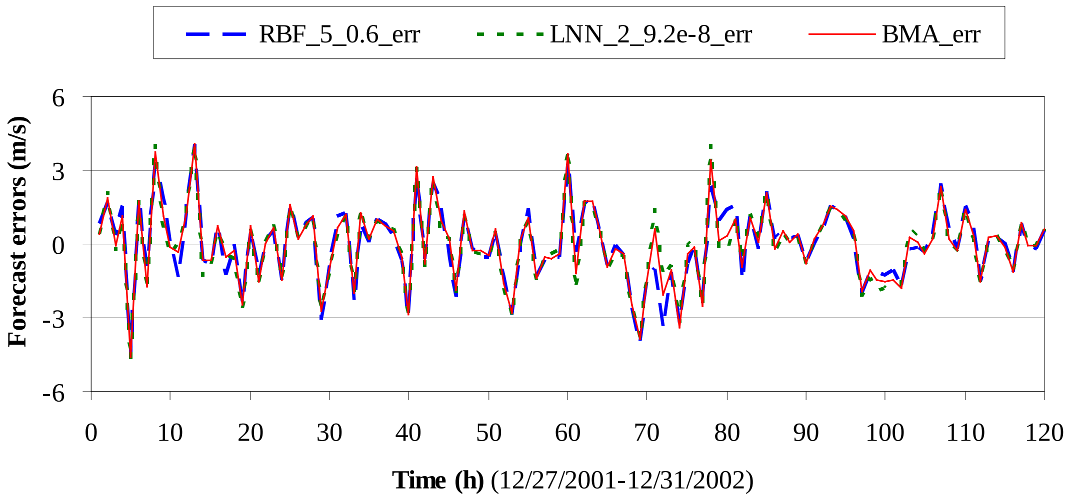


Figure 24. Forecast errors of BMA model and the two ‘best’ component models (Kill).

More importantly, the BMA forecasting errors are always close to, if not less than, the smallest error of the best-performing component model at any time point. This indicates that the performance of each forecasting model varies from time to time, but the BMA forecast error stays within the variation range at any time point. Due to the performance inconsistency along the time scale among component forecasting models, in long-run BMA forecasting can always manage to achieve an overall performance close to, or even better than, that of the best model. This confirms the adaptive and reliable nature of the Bayesian combination forecasts.

Besides, another advantage of the BMA forecast is that it can also provide the probability information corresponding to the forecast. After the EM iterations optimize the model parameter vector $\theta = [w_j, \sigma_j^2]$($j$=1,2,…, J), the variance or the standard deviation of each specific Bayesian combination forecast can be calculated in that all variables in the equation are known. Therefore, the preferred probability information of the Bayesian combination forecast can be obtained. For example, the 95% confidence interval of the Bayesian forecast can be calculated as $f_{BMA} \pm 1.96\sigma_{BMA}$.

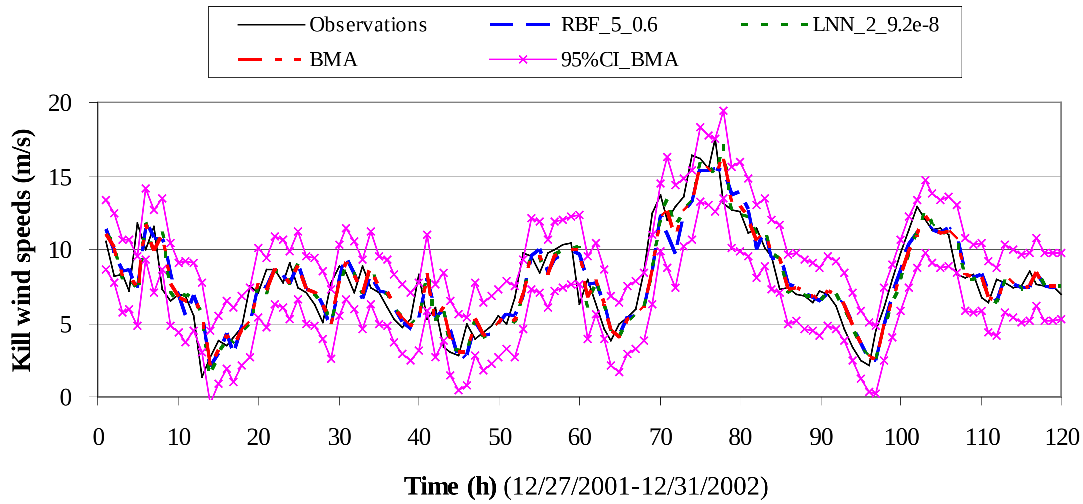


Figure 25. Time series plots of forecasts, observations, and 95% CI of BMA forecasts.

Figure 25 shows the forecast time series from the above-mentioned RBF and ADALINE models that generate the smallest and the largest RMSE values for site Kill, respectively, and the Bayesian combinations forecasts as well as the corresponding 95% confidence intervals. Encompassing the range of values in which the hourly averaged observation falls in a given percentage of the time, such an interval can be incorporated into the decision making process for production scheduling and control. The financial implications of failing to use wind power forecasts is obvious. For example, imbalance charges resulting from deviations in scheduled output will increase project operating costs. Fortunately, Bayesian combination forecasts can help to reduce or minimize these penalties. Such information can also reduce the significant opportunity costs of being too conservative in bidding output into a forward market, due to uncertainty of availability.

4.3.4. Forecasting results for the site of Hann

Similarly, Table 14 presents the results about the best ANN models that generate the smallest MAE, RMSE, and/or MAPE values, for the wind dataset from the site of Hann. Some similar findings can be observed for this site as compared with site Kill. Take the RBF models as an example, each smallest MAE, RMSE, or MAPE corresponds to a different combination of model parameters, indicating that different learning rates or spread constants, as well as different number of inputs, can affect the forecast accuracy of some specific NN model.

Table 14. Best component NN models and their weights and variances (Hann).

Best model	MAE	RMSE	MAPE	Weight	Variance
LNN_Obs_2_lr_8.9e-8	0.965	1.271	0.196	0.1336	0.028
LNN_Obs_4_lr_4.5e-8	0.980	1.290	0.194	0.1669	0.033
BP_Obs_8_LR_0.075	0.945	1.269	0.204	0.1302	0.049
BP_Obs_6_LR_0.1	0.951	1.254	0.211	0.1594	0.047
RBF_Obs_2_spread_0.5	0.989	1.297	0.232	0.1558	0.030
RBF_Obs_3_spread_0.5	0.997	1.294	0.234	0.1275	0.024
RBF_Obs_7_spread_0.7	1.058	1.390	0.221	0.1267	0.061
BMA	0.954	1.257	0.207	/	/

It is also verified by Table 14 that different types of NN models usually forecast with different performances. For instance, in terms of RMSE, the BP model using previous 6 observations and trained at a learning rate of 0.1 seems to be the best, whereas the RBF model trained with 7 previous observations at a learning spread of 0.7 appears to be the worst in term of

RSME. Compared with Table 13, it can be seen that the performance of NN models also varies from site to site. This further verifies the necessity of developing a robust and accurate combination method for short-term wind speed forecasts.

Similarly, all the time series are transformed into approximately normally distributed ones, and the posterior probabilities are thereafter estimated via EM iterations, in term of weight. After the BMA combination forecasts are generated for the testing data, the model performances are calculated and listed in the last row of Table 14. Again, it can be seen that the performance of BMA model is always close to the best ones in terms of all performance metrics.

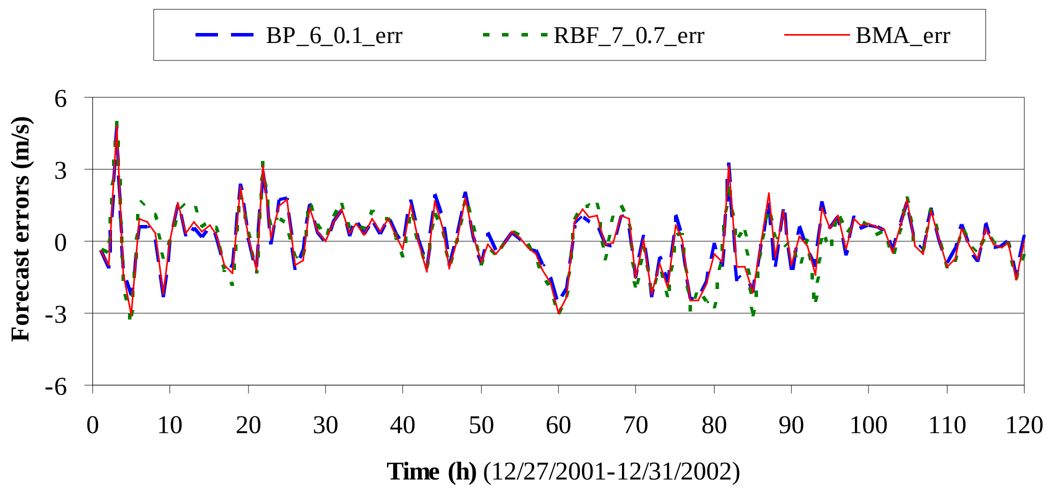


Figure 26. Forecast errors of BMA model and the two component models (Hann).

Figure 26 illustrates the plots of the forecast error time series from the BP and RBF models that generate the smallest and largest RMSE values, respectively, as well as the Bayesian combination forecast errors. It can be observed that the best BP model in term of RMSE does not always perform better than RBF model. The BMA model can always stay within the inconsistent discrepancy among the component models, and this in long run helps produce the performance close to (or even better than) the best component model in terms of any metrics. Meanwhile, after the estimation of model parameter vectors by EM algorithm, the probability information of the Bayesian combination forecast also becomes available.

Figure 27 shows the forecast time series from the above-mentioned three models as well as the 95% confidence interval for Bayesian combinations forecasts. As mentioned previously, this information is very useful for decision making in wind energy production.

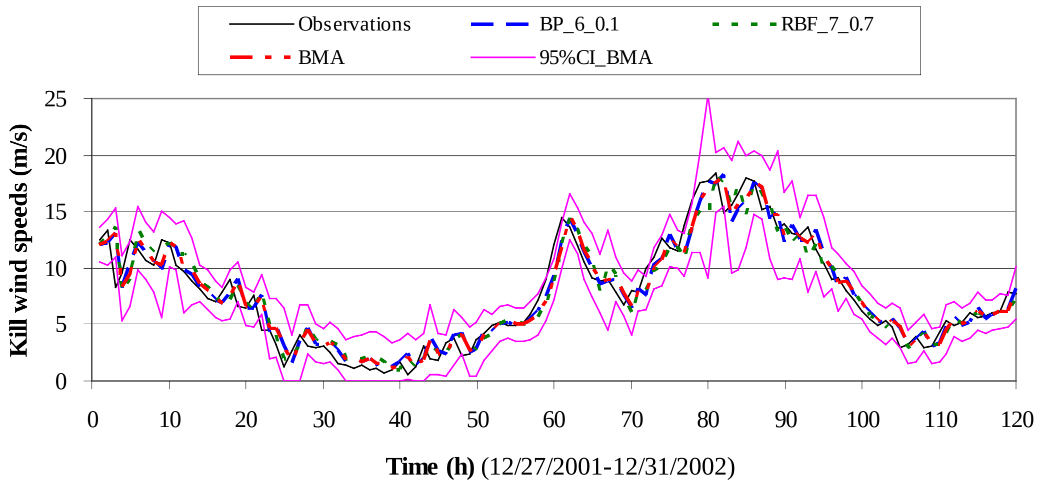


Figure 27. Plots of forecasts, observations, and 95% CI of BMA forecasts (Hann).

4.3.5. Discussion

The results indicate that learning rates or spread constants, as well as number of inputs, affect the forecast performance of NN models. Therefore, in building the NN models for forecasting wind speed, these factors should be properly determined. More importantly, the choice of best NN model varies between sites Kill and Hann and among the evaluation criteria MAE, RMSE, and MAPE. This confirms the needs of evaluating multiple types of neural network based on multiple performance metrics. Otherwise, the results can be misleading and biased. However, it poses a challenge to the practical application of the NN forecasting models in industry – if only one time series forecast can be adopted, which one should we use? This is very real for wind energy industry, where the direct operation relies on one clear and definite forecast instead of multiple choices. A traditional model selection process among the NN models may not be the good answer to the question.

The alternative is to generate a final single forecast that could take advantage of a set of plausible forecasts. For example, the forecasts from alternative forecast agencies should be considered before making a decision. Each forecast agency might make forecasts for the client by adopting alternative models or procedures such as different NN models as introduced in this paper. In such cases, the agency also has to combine all the available information in order to provide a single forecast (Sánchez, 2008). Therefore, it is apparent that an efficient forecast combination procedure should be of great importance for short-term wind speed forecasts. Ideally, this single time series forecast should have close-to-optimal performance under each evaluation metric.

The proposed two-step methodology adds the step of Bayesian combination analysis on top of the extensive evaluation of various NN models. The applications of the BMA algorithm into the datasets collected from two North Dakota sites verify that, by weighting the forecast from each specific model with the corresponding posterior probability, the algorithm can generate adaptive, reliable and comparatively high-accuracy forecasts. The robustness and reliability lie in that, once the candidate models are included in the model space, the Bayesian combination analysis is universal and thus avoids the confusion and ambiguity of model selection. A significantly plausible model will be weighted highly whereas the influence on the Bayesian combination forecast of whichever insignificant model could be automatically minimized to the lowest degree. Besides, the BMA algorithm can also provide the confidence interval on the combined forecast value, supplying abundant helpful information for related decision-makings for wind energy conversion. Clearly, this methodology not only is an enhancement for reliable wind speed forecast using NN models, but also provides the opportunity for other types of forecasting models – the same methodology should be applicable to ARIMA and other models.

5. AGENT-BASED SIMULATION ON WGENCO'S BIDDING WIND POWER

In this chapter, agent-based simulation models are employed to analyze the issue of WGenCos' bidding strategy optimization while considering the uncertainty in wind generation or wind forecasting accuracies. Section 5.1 briefly explains the market structure and auction rules as well as the fundamental agent-based simulation modeling method adopted in this study. Section 5.2 then presents the details of the agent-based simulation on the single-sided bidding of wind power in the day-ahead electricity market which is connected to a 9-bus 3-generator power grid, by using the BP-based point estimates of wind power. To evaluate the effects of forecasting accuracy on the WGenCos net earnings, bidding with the forecasts from the persistence model is also simulated as the benchmark for comparison. Considering that the future deregulated electricity market should be a double-sided one, the agent-based simulations are further employed to study the bidding of wind power in a double-sided day-ahead electricity market assumed to be connected to a modified IEEE 30-bus 6-generator power system. The results are presented and analyzed in Section 5.3. It should be noted that emphasis is placed on the economic benefits of combining and improving the short-term forecasting accuracy of wind generation in the day-ahead electricity markets in all the case studies. The results clearly demonstrate that improving wind forecasting accuracy helps to increase the net earnings of the wind generation company. Also, it is demonstrated that agent-based simulation is a viable modeling tool which can provide realistic insights for the complex interactions among different participants and various market factors.

5.1. Agent-Based Simulation and Deregulated Electricity Markets

In an agent-based model, market participants are modeled as adaptive agents with different bidding preferences and strategies. Each agent may develop the optimal bidding strategy by learning from its past experiences obtained from the direct interaction with environment (Tesfatsion, 2006). Generally, the agent-based modeling procedures include the following steps (Weidlich and Veit, 2008): (1) defining the research questions to be resolved; (2) constructing a model comprising an initial population of agents; (3) specifying the initial model state by defining the agents' attributes and the structural and institutional framework of the electricity market; (4) having the model evolve over time; and (5) analyzing simulation results and evaluating the regularities observed in the data.

5.1.1. Market structure and auction procedures

The Western System Coordinating Council (WSCC) 9-bus 3-machine power system network (Anderson and Fouad, 2003) is adopted to represent the transmission grid since it has been widely used as a benchmark system in power system related research (Babulal and Kannan, 2006; Gallardo and Ledesma, 2008; Nwohu, 2010). Figure 28 illustrates the market structure and its auctions procedures, which will be adopted in Section 5.2. As can be seen in Figure 28, three GenCos (sellers) are located at buses 1, 2, and 3, and three loads (buyers) are at buses 5, 7, and 9, respectively. The generator at bus 2 is a WGenCo, and the other two generators at buses 1 and 3 are traditional GenCos. The market structure and the logical flow of daily market operations are depicted as follows.

- The day-ahead electricity market is operated by an ISO over the 9-bus 3-generator AC transmission grid. The GenCos and the buyers intensively maximize their net earnings via optimally bidding actions, whereas the ISO mainly ensures the operational reliability and efficiency of the market while maximizing the total net surplus under constraints.
- The simulation starts from day 1 to the user-specified last day (i.e., day 31 in this study), with each day consisting of 24 hours (0, 1, …, 23). At the beginning of each day D

(H00), the hourly wind generations are forecasted via the persistence model and the BP neural network model [Li and Shi, 2011], respectively for each hour of the next day, and scaled into two corresponding penetration levels, namely, 5% and 20%, respectively.

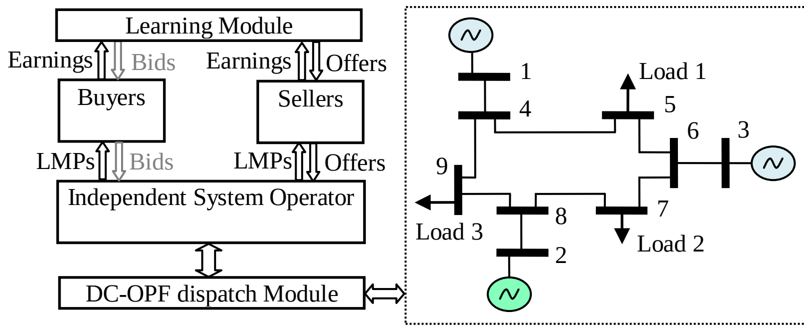


Figure 28. Framework of the day-ahead electricity market and grid network.

- At the beginning of each day D (H00), each GenCo reports a supply offer consisting of a linear marginal cost function defined within the reported capacity interval for day D+1. The GenCos have learning capabilities enabling them to choose different supply offers according to their current supply choice probabilities. Meanwhile, each buyer presents to the ISO a demand bid for day D+1, which is assumed to be a fixed 24-hour load profile for every day during the simulation period.
- After receiving the supply offers and the demand bids during the early morning of day D, the ISO determines and publicly declares the hourly locational marginal prices (LMPs) and power supply settlement for day D+1 via solving the hourly bid/offer-based direct current optimal power flow (DC-OPF) problem. It should be noted that LMP is better than MCP especially when transmission congestions exist. MCP mainly reflects the cost of serving energy demand without considering transmission, whereas LMP is a pricing mechanism that reflects the costs of re-dispatching out-of-merit energy and delivering energy to the location with transmission congestions. LMP can reflects the actual cost for

buyers and sellers to deliver energy at their locations on the transmission system. Therefore, LMP encourages the efficient use of the transmission system by assigning costs to users based on the way energy is actually delivered.

- At the end of each day D, according to the LMPs and the commitments settled by the ISO for day D+1 of the day-ahead market. Each GenCo updates its action choice probability based on its net earnings from the settlement of day D+1.

Necessary assumptions are made in this study, which include: (1) only the day-ahead electricity markets are considered, that is, the GenCos do not update their supply offers by participating in the real-time markets; (2) no GenCo adopts the control strategy, e.g., combining energy storage strategy or cooperating with other participants; (3) no system disturbances (e.g., weather changes) or shocks (e.g., line or generation outages) exist during the simulation period; (4) neither the number of the participants nor the participants themselves will decrease, increase, or be substituted during the simulation period.

5.1.2. Agent-based modeling methodology

Each GenCo is modeled as an agent who reports at the beginning of day D a single supply offer $s_i^R = (a_i^R, b_i^R, Cap_i^{RL}, Cap_i^{RU})$ for use in each hour H of day D+1 in the day-ahead market, representing its reported marginal cost function:

$$MC_i^R(p_{Gi}) = a_i^R + 2 \cdot b_i^R \cdot p_{Gi}^R, \tag{34}$$

which is defined over a reported power generation interval of GenCo i, (Cap_i^{RL}, Cap_i^{RU}), that is,

$$Cap_i^{RL} \le p_{Gi}^R \le Cap_i^{RU}. \tag{35}$$

Generally, the total cost is composed of two parts, namely, total fixed cost and total variable cost. The total fixed cost of a GenCo is the unavoidable production cost independent of its generation level, whereas the total variable cost refers to the GenCo's cost that could be avoidable and varies with the generation level (Li and Tesfatsion, 2009).

The true variable cost of GenCo *i* for each hour H can be calculated as the integral of its marginal cost function over its real-time settled generation p_{Gi} of hour H.

$$TVC_i(p_{Gi}) = \int_0^{p_{Gi}} MC_i(p)dp = a_i \cdot p_{Gi} + b_i \cdot (p_{Gi})^2. \tag{36}$$

The net earnings are defined as the revenue minus the true total avoidable cost. For example, if GenCo *i* located at bus *k*(*i*) is dispatched at a generation level p_{Gi}^D at price $LMP_{k(i)}$ for hour H of day D+1 in the day-ahead market, the net earnings of GenCo *i* for hour H of day D+1, won at the end of day D, is calculated as

$$NE_i(H,D) = LMP_{k(i)} \cdot p_{Gi}^D - TVC_i(p_{Gi}^D), \tag{37}$$

The total net earnings of GenCo *i* over all 24 hours of day D+1, won at the end of day D, can then be calculated as

$$NE_i(D) = \sum_{H=00}^{23} NE_i(H,D). \tag{38}$$

All GenCos should ensure the delivery of the amount committed to and settled by the ISO. For the traditional GenCos, they can ensure that the real generation is equal to the dispatched, that is, $p_{Gi} = p_{Gi}^D$. However, more calculation is needed for the WGenCo because of the uncertainty of wind generation. In this study, it is assumed that the difference between the real wind generation and the dispatched one could be balanced in the following way. If $p_{Gi} > p_{Gi}^D$, the access will be sold at the system selling price $C_{k(i)}^{sel}$ of bus *k(i)*, the net earnings for hour H of day D+1 is calculated as

$$NE_i(H,D) = LMP_{k(i)} \cdot p_{Gi}^D - TVC_i(p_{Gi}^D) + (p_{Gi} - p_{Gi}^D) \cdot C_{k(i)}^{sel}. \tag{39}$$

Otherwise, if $p_{Gi} < p_{Gi}^{D}$, the shortfall will be filled in by purchasing power at the system buying price, $C_{k(i)}^{buy}$, and the net earnings for hour H of day D+1 will be calculated as

$$NE_i(H,D) = LMP_{k(i)} \cdot p_{Gi}^{D} - TVC_i(p_{Gi}^{D}) - (p_{Gi}^{D} - p_{Gi}) \cdot C_{k(i)}^{buy}, \tag{40}$$

where the imbalance prices of $C_{k(i)}^{sel}$ and $C_{k(i)}^{buy}$ are assumed to equal to a certain proportion of the $LMP_{k(i)}$, e.g., 90% and 110%, respectively, as recommended by the Federal Energy Regulatory Commission (FERC) under order number 890. The total net earnings can then be calculated by summing up the benefits for 24 hours of day D+1.

As aforementioned, at the beginning of each day D, each GenCo *i* should choose a supply offer $s_i^R = (a_i^R, b_i^R, Cap_i^{RU})$ for each hour H of day D+1 from its action domain AD$_i$ and report them to the ISO of the day-ahead market. For purpose of earning profits, the reported marginal cost should be above or at least on the real marginal cost function; also, it is usually in an upward-sloping format, i.e., $b_i^R > 0$. Besides, the reported capacity interval should be within the real capacity interval (Cap_i^L, Cap_i^U), that is,

$$0 \le Cap_i^L \le Cap_i^{RL} \le p_{Gi} \le Cap_i^{RU} \le Cap_i^U . \tag{41}$$

Under these assumptions, the action domain matrix AD$_i$ with finite positive cardinality M*i*, defined as the supply offer choice set of GenCo *i*, is constructed as follows (Li, Sun, and Tesfatsion, 2009).

For any GenCo *i*, given any positive value of base-slope parameter bs_i, three integer-value density-control parameters (M1$_i$, M2$_i$, and M3$_i$; $\mathrm{M}_i = \mathrm{M1}_i \cdot \mathrm{M2}_i \cdot \mathrm{M3}_i$) are specified to control the number of possible ordinate values a_i^R, slope values b_i^R, and upper capacity limits Cap_i^{RU}, respectively. Meanwhile, three percentage-format range-index parameters ($RIMax_i^L$,

$RIMax_i^U$, and $RIMin_i^C$) are defined to control the ranges of a_i^R, b_i^R, and Cap_i^{RU}, respectively. As verified in (Sun and Tesfatsion, 2007), each supply offer $s_i^R = (a_i^R, b_i^R, Cap_i^{RL}, Cap_i^{RU})$ corresponds to and can be represented by a percentage-format vector $s_i^A = (RI_i^L, RI_i^U, RCap_i^L, RCap_i^U)$, where $RI_i^L \in [0, RIMax_i^L]$, $RI_i^U \in [0, RIMax_i^U]$, $RCap_i^U \in [RIMin_i^C, 1]$, $RCap_i^L \in [0, RCap_i^U)$, and $RCap_i^L = Cap_i^L / RCap_i^U$. Therefore, once bs_i, M1$_i$, M2$_i$, M3$_i$, $RIMax_i^L$, $RIMax_i^U$, and $RIMin_i^C$ are specified, AD$_i$ is specifically defined for GenCo *i* to adaptively choose its offer $s_i^R = (a_i^R, b_i^R, Cap_i^{RU})$ by learning from its own past daily net earnings.

The learning process is realized via Variant Roth-Erev reinforcement learning (VRE-RL), a variant of a stochastic reinforcement learning algorithms proposed by Sun and Tesfatsion (2007). Compared with the simplest RE-RL, the VRE-RL algorithm can ensure the updating of propensities even with zero-valued net earnings and it can handle negative propensity values as well. The learning procedure is briefly described below.

Step 1: Specify the initial propensity and the initial selected action. At the beginning of day D=1, the current propensity of GenCo *i*'s offer $s_{im_i}^R \in$AD$_i$ is given by $IP_{im_i}(D)$ for m_i = 1, ..., M$_i$. In this study, GenCo *i*'s M$_i$ initial propensities are all assumed to be $IP_{im_0}(1)$, a user-specified real number, that is, under this assumption, the initial action $s_{im_0}^R$ can thus be randomly specified.

Step 2: Calculate the daily net earnings of day D. The auction market is run and the actually daily net earnings $NE_{im_0}(D)$ can then be updated accordingly.

Step 3: Update the propensities of M$_i$ possible actions. At the end of day D, the current propensity $IP_{im_i}(D)$ of each supply offer $s_{im_i}^R \in$AD$_i$ is updated via the following equations,

$$IP_{im_i}(D+1) = (1-r_i)\cdot IP_{im_i}(D) + R_{im_i}(D), \tag{42}$$

$$R_{im_i}(D) = \begin{cases} (1-e_i)\cdot NE_{im_0}(D) & if\ m_i = m_0 \\ e_i \cdot IP_{im_i}(D)/(M_i - 1) & if\ m_i \neq m_0 \end{cases}, \tag{43}$$

where $R_{im_i}(D)$ refers to the response component, r_i is the recency parameter acting as a damper on the growth of the propensities over time, e_i is the experimentation parameter encouraging the continued experimentation of reinforcement with various offers at the beginning of the learning process.

Step 4: Update the choice probabilities for GenCo *i* to select an offer for day D+1,

$$CP_{im_i}(D+1) = \frac{\exp[IP_{im_i}(D+1)/T_i]}{\sum_{m_j=1}^{M_i}\exp[IP_{im_j}(D+1)/T_i]}, \tag{44}$$

where T_i is a cooling parameter that affects the degree to which GenCo *i* makes use of propensity values in determining its choice probabilities.

Step 5: If the maximum day is not reached, let D=D+1 and repeat Steps 2-4.

5.2. Agent-Based Single-Sided Bidding of Wind Power

In order to evaluate the effects of wind forecasting accuracy on the net earnings of WGenCo participating in the deregulated electricity wholesale market as described above, the agent-based simulation approach is employed. The simulation models, especially the agent-learning algorithms are built based on the Agent-based Modeling of Electricity Systems of AMES Market package V2.05, an open-source Wholesale Power Market Test Bed developed in the Java environment (Li, Sun, and Tesfatsion, 2009).

Four simulation scenarios of WGenCo bidding in the day-ahead electricity market are investigated, which are reflected by two wind penetration levels (5% and 20%) and two algorithms (with or without learning). The forecasts of wind power generation are obtained by two time series forecasting models (i.e., persistence, BP neural network), which generates different forecasting accuracies.

5.2.1. Modeling procedures and parameters

This study includes the following experimental modules: wind power forecasting, learning algorithms programming, integration with the DC-OPF dispatch module, and data post-processing. The experiments are performed as mapped in Figure 29. Basically, the ISO calculates the optimal power flow and decides on the LMPs. After the offers and bids are collected, the ISO solves this optimization problem via calling DC-OPF dispatch module and declares the quantities and LMPs; the sellers and the buyers then call the learning module to generate their optimal strategies for the next day's offers or bids.

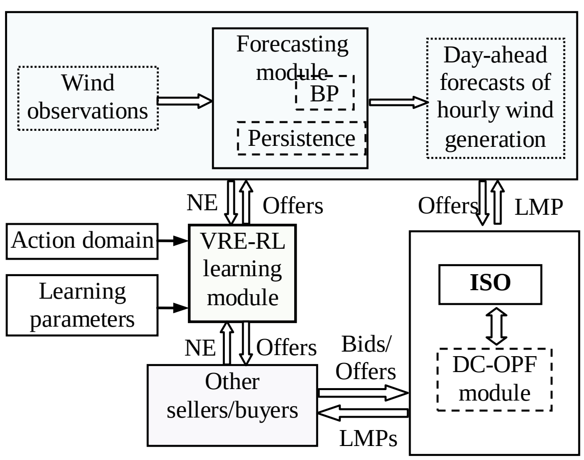


Figure 29. Schematic of the experimental design and auction process.

In view that no wind generation data are available for the observation sites adopted in previous study, the hourly wind generation data employed in this study are collected from a real wind farm located in North Dakota, U.S. during November 1, 2008 – May 31, 2009, among which the dataset during May 1 - 31, 2009 are reserved as for forecasting purpose. It should be noted that, as has been verified in many literatures, the short-term forecasting methods and models previously presented are also applicable to short-term wind power forecasting.

As shown in Figure 29, the BP forecasting model and the persistence model are adopted to generate forecasts of different accuracies in this study. For BP neural network forecasting, all the earlier observations are used for training and testing the BP models. Wind energy penetration level is calculated as the total amount of wind energy produced divided by gross electricity demand during the same period (Wind energy the facts, 2011). Milligan, Lew, Corbus, Piwko, Miller, Clarke, et al. (2009) further define three wind power penetration levels in the U.S. power systems, and the three levels are 10%, 20%, and 30% for low, high, and even higher penetrations, respectively. In this study, by directly scaling the total wind generations, we obtain the wind energy penetration levels of 5% and 20%. The two levels are used to approximate low and high penetration levels, respectively.

The VRE-RL learning process is realized via programming in the Matlab environment. To construct the action domains, only bs_i, $M1_i$, $M2_i$, $M3_i$, $RIMax_i^L$, $RIMax_i^U$, and $RIMin_i^C$ need to be specified. These parameters are selected by referring to (Li et al., 2009), as shown in Table 15, with the densities of ordinates and slopes being increased.

Table 15. Action domain parameters.

GenCo i	bs_i	$M1_i$	$M2_i$	$M3_i$	$RIMax_i^L$	$RIMax_i^U$	$RIMin_i^C$
1	0.001	5	5	1	0.75	0.75	1
2(WGenCo)	0.001	5	5	1	0.75	0.75	1
3	0.001	5	5	1	0.75	0.75	1

The initial propensity of each action indicates the GenCo's expectation on the net earnings from adopting this action. Too high an initial propensity may lead to the GenCo's cycling through various admissible actions. Too low an initial propensity could result in the early fixation of the agent on some non-optimal action at the beginning stage, thus reducing the probability of choosing other actions. The cooling parameter can help control the degree to which each GenCo *i* makes use of propensity values in determining its choice probabilities. The initial propensity $q_i(1)$ and cooling parameters T_i are selected according to the best combined parameter $q_i(1)/T_i$ proved in (Li et al., 2009), that is, each GenCo's initial propensity $q_i(1)$ is set to equal its estimated maximum daily net earnings (MaxDNE_i) and the cooling parameter T_i is set as one percent of the initial propensity.

However, in order to verify the possible impacts on the learning results of recency parameter r_i, and experimental parameter e_i, different values of these parameters are tested in this study, as shown in Table 16. It can be seen that three representative values, namely, 0.05, 0.5, and 0.95, for both r_i and e_i where $r_i \in [0, 1]$ and $e_i \in [0, 1)$, are used to cover the entire study range (Pentapalli, 2008). Besides, the fixed 24-hour loads from three buses are adopted from the case study of (Shahidehpour, Yamin, and Li, 2002), as shown in Table 17.

Table 16. Different values of learning parameters.

GenCo *i*	r_i	e_i	$q_{i(1)}/\text{MaxDNE}_i$	$q_i(1)/\text{T}_i$
1	(0.05, 0.5, 0.95)	(0.05, 0.5, 0.95)	1	100
2(WGenCo)	(0.05, 0.5, 0.95)	(0.05, 0.5, 0.95)	1	100
3	(0.05, 0.5, 0.95)	(0.05, 0.5, 0.95)	1	100

The power flow data of the network built on the WSCC 9-bus 3-generator power system are shown in Table 18 (Zimmerman, Murillo-Sanchez, and Thomas, 2011). It should be noted that, to reflect the fixed 24-hour daily loads, the values of *Pd* at the corresponding buses are dynamically changed and updated during the simulation. Besides, three buyers are modeled as fake GenCos with negative generations and negative marginal cost functions. Also, by referring to literature (Li et al., 2009), the true marginal costs of three GenCos are selected and illustrated in Figure 30, and the corresponding parameters are summarized in Table 19.

Table 17. Profiles of three fixed daily loads.

Pd (MW)	H00	H01	H02	H03	H04	H05	H06	H07	H08	H09	H10	H11
1(bus 5)	350	322.93	305.04	296.02	287.16	291.59	296.02	314.07	358.86	394.8	403.82	408.25
2(bus 7)	300	276.8	261.47	253.73	246.13	249.93	253.73	269.2	307.6	338.4	346.13	349.93
3(bus 9)	250	230.66	217.89	211.44	205.11	208.28	211.44	224.33	256.33	282	288.44	291.61
Pd (MW)	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23
1(bus 5)	403.82	394.8	390.37	390.37	408.25	448.62	430.73	426.14	421.71	412.69	390.37	363.46
2(bus 7)	346.13	338.4	334.6	334.6	349.93	384.53	369.2	365.26	361.47	353.73	334.6	311.53
3(bus 9)	288.44	282	278.83	278.83	291.61	320.44	307.67	304.39	301.22	294.78	278.83	259.61

The recency parameter acts as a damper on the growth of the propensities over time and it allows the agent to ignore the past learning effect by the factor of recency parameter, with larger recency parameter indicating higher forgetting effect. The experimentation parameter controls the degree of exploration during the online learning process, with the larger parameter value increasing the propensity of those actions that are not chosen currently, thus encouraging continued experimentation of reinforcement with various offers at the beginning days of the learning process (Pentapalli, 2008).

To compare the effects of recency parameters and experimentation parameters on the simulation results, the scenario of 20% wind penetration level is examined and the total net earnings are summarized in Table 20 for simulation cases with different recency parameters and experimentation parameters.

Table 18. Power flow data of WSCC 9-bus, 3-generator test bed.

System MVA Base

Base MVA = 100

Bus Data

Bus	type	Pd	Qd	Gs	Bs	area	Vm	Va	baseKV	zone	Vmax	Vmin
1	3	0	0	0	0	1	1	0	345	1	1.1	0.9
2	2	0	0	0	0	1	1	0	345	1	1.1	0.9
3	2	0	0	0	0	1	1	0	345	1	1.1	0.9
4	1	0	0	0	0	1	1	0	345	1	1.1	0.9
5	1	448.62	30	0	0	1	1	0	345	1	1.1	0.9
6	1	0	0	0	0	1	1	0	345	1	1.1	0.9
7	1	384.53	35	0	0	1	1	0	345	1	1.1	0.9
8	1	0	0	0	0	1	1	0	345	1	1.1	0.9
9	1	320.44	50	0	0	1	1	0	345	1	1.1	0.9

Branch data

Fbus	Tbus	r	x	b	rateA	rateB	rateC	ratio	angle	status
1	4	0	0.0576	0	10000	250	250	0	0	1
4	5	0.017	0.092	0.158	10000	250	250	0	0	1
5	6	0.039	0.17	0.358	10000	150	150	0	0	1
3	6	0	0.0586	0	10000	300	300	0	0	1
6	7	0.0119	0.1008	0.209	10000	150	150	0	0	1
7	8	0.0085	0.072	0.149	10000	250	250	0	0	1
8	2	0	0.0625	0	10000	250	250	0	0	1
8	9	0.032	0.161	0.306	10000	250	250	0	0	1
9	4	0.01	0.085	0.176	10000	250	250	0	0	1

Area Data

area	refbus
1	5

Generator Data

Bus	Pg	Qg	Qmax	Qmin	Vg	mBase	status	Pmax	Pmin
1	550.26	0	100	-100	1	100	1	600	100
2	Forecast	0	100	-100	1	100	1	600	0
3	650.53	0	100	-100	1	100	1	600	100
5	*-30*	*-15*	*0*	*-15*	*1*	*100*	*1*	*0*	*-30*
7	*-30*	*-12*	*0*	*-12*	*1*	*100*	*1*	*0*	*-30*
9	*-30*	*-7.5*	*0*	*-7.5*	*1*	*100*	*1*	*0*	*-30*

Generator Cost Data

Type	start	shut	n	x1	y1	x2	y2	x3	y3	x4	y4
1	0	0	4	0	0	100	3140	350	12215	600	23040
1	0	0	4	0	0	100	2680	400	10955	600	21480
1	0	0	4	0	0	100	2160	400	8960	600	17760
1	*0*	*0*	*4*	*-30*	*-3000*	*-20*	*-2000*	*-10*	*-1000*	*0*	*0*
1	*0*	*0*	*4*	*-30*	*-3000*	*-20*	*-2000*	*-10*	*-1000*	*0*	*0*
1	*0*	*0*	*4*	*-30*	*-3000*	*-20*	*-2000*	*-10*	*-1000*	*0*	*0*

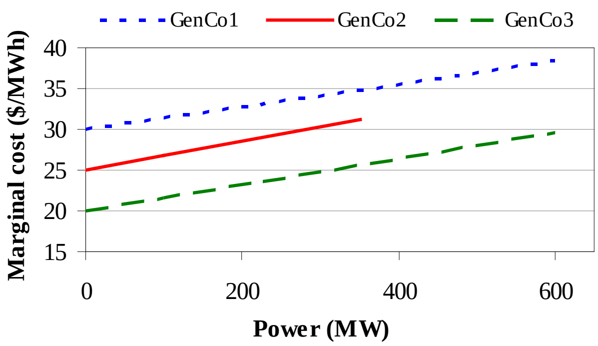


Figure 30. Real marginal costs of three GenCos in the network.

Table 19. True marginal cost functions and capacity limits of GenCos.

GenCo	a_i ($/MWh)	b_i ($/MW^{2}h)	Cap^L (MW)	Cap^U (MW)
1	30	0.006	100	600
2(WGenCo)	25	0.005	0	Forecast values
3	20	0.007	100	600

Table 20. Total net earnings of WGenCo vs. learning parameters (20% wind).

Total net earnings ($10^3$$)		Experimentation parameter e_i		
		0.05	0.5	0.95
Recency parameter r_i	0.05	1266.203	1337.677	1389.162
	0.5	1317.676	1348.456	1340.178
	0.95	1351.342	1342.869	1283.078

It can be observed that among the tested cases, the one with a recency parameter r_i of 0.05 and an experimental parameter e_i of 0.95 seems to be the best one, which leads to a total net earnings of $1,389,162. This is about 9.7% more than the output of the worst combination of e_i=0.05 and r_i=0.05. In addition, it can also be seen that the impact of recency parameter is less significant when the experimentation parameter is 0.5, and similarly the impact of experimentation parameter is less significant when the recency parameter is 0.5. To provide a common comparison platform, all the following simulation results are obtained with the recency parameter of 0.05 and an experimentation parameter of 0.95, without losing the generalization.

5.2.2. Forecasting results

Recall that we use two approaches to obtain forecasts of wind power generation, namely, persistence method, and BP neural network model. The persistence model simply uses the last observation of current day as the possible observations of each of the following 24 hours in the next day. For BP models, different input parameters and learning rates are tested, and the model generating the smallest mean absolute errors (MAE) and root mean squared errors (RMSE) is selected to compare with the persistence model.

The forecasting results of the persistence and BP models are summarized and compared as in Table 21. The forecasting errors of the two models are plotted in Figure 31 as well.

Table 21. Performances of two forecasting models.

Model	MAE	RMSE	MAPE
BP	193.632	229.726	559.690
Persistence	232.362	300.849	662.447

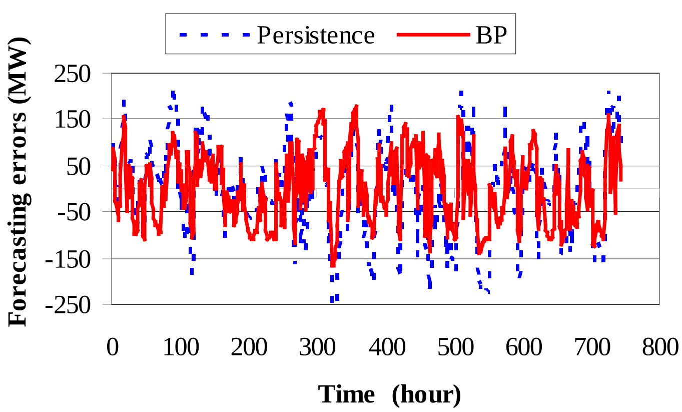


Figure 31. Wind forecasting errors of persistence and BP models.

It can be seen from both Table 21 and Figure 31 that the BP model is more accurate than the persistence model in terms of all three performance measures. To be exact, the BP model is 16.7%, 23.6%, and 15.5% more accurate compared with the persistence model in terms of MAE, RMSE, and MAPE, respectively, even though the persistence model is often regarded as a reliable and robust method for short-term forecasting.

5.2.3. Bidding for scenarios with 5% wind power

The wind penetration level of 5% approximates the case of low-medium wind power penetration level. The simulated auction results for the scenarios at this penetration level are summarized in Table 22. It can be seen that the total net earnings of WGenCo using BP forecasting results are higher than that using persistence forecasting results, no matter if GenCo learning is adopted or not. Without the adoption of learning algorithm, the net earnings of WGenCo by using BP forecasts in the day-ahead market is $495,030, about 5.5% more than that of using the forecasts of persistence model ($469,420). When learning capability is empowered, this percentage of extra earnings is also around 4.5%. This verifies that improving the forecasting accuracy can help increase the net earnings for the WGenCo participating in the day-ahead electricity market under the same other conditions.

Table 22. Auction results for the WGenCo at 5% penetration level.

Scenarios		Qty_Sel(MWh)	Ave_LMP($/MWh)	Tot_NE($10^3$$)
W/O learning	BP	47850.178	39.244	495.030
	Persistence	48732.784	39.412	469.420
W/ learning	BP	47850.181	44.151	741.704
	Persistence	49057.782	44.274	710.018

Meanwhile, it can also be observed that the cases of empowering GenCos with learning ability could obtain more earnings compared with the cases without learning capability. When the BP forecasts are used, the net earning increases by 49.8%, from $495,030 (without learning) to $741,704 (with learning). When the persistence model forecasts are used, the net earnings increase by 51.3%, from $469,420 (without learning) to $710,018 (with learning). The advantage of GenCo's adopting reinforcement learning algorithms in optimizing its bidding strategy in the day-ahead electricity market is also directly reflected by the average locational market price. For instance, based on BP forecasts, the case with GenCos' learning generates a high average market clearing price of $44.151/MWh whereas the case of without GenCos' learning leads to a low price of $39.244/MWh.

5.2.4. Bidding for scenarios with 20% wind power

The simulated auction results for the scenarios at the wind penetration level of 20% are summarized in Table 23. Again, it can be seen that more accurate forecasting results can increase the net earnings for the WGenCo. For the cases with learning capability, using BP forecasts brings the WGenCo a total net earning of $2,383,078, about 2.2% more than that earned by using the persistence model forecasts, i.e., $2,330,971. For the cases without learning, the improvement is 3.7%. Both improvements demonstrate the benefit of improving forecasting accuracy of wind power generation.

Table 23. Auction results for the WGenCo at 20% penetration level.

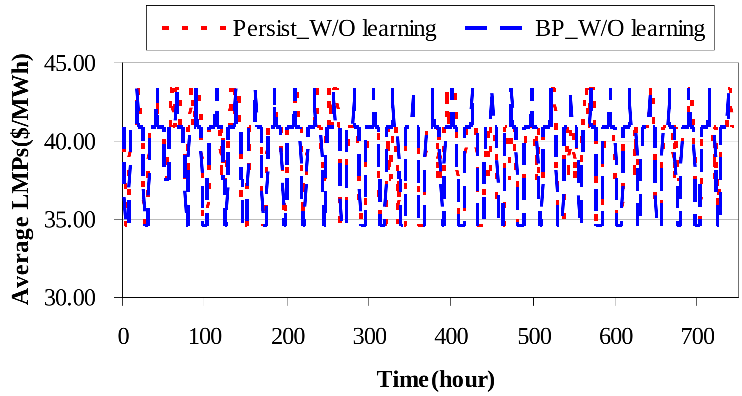
Scenarios		Tot_Qty(MWh)	Ave_LMP($/MWh)	Tot_NE($10^3$$)
W/O learning	BP	173666.203	37.677	1389.162
	Persistence	141971.143	38.094	1340.188
W/ learning	BP	178472.676	42.456	2383.078
	Persistence	143513.342	42.869	2330.971

Similarly, it can be seen that adopting reinforcement learning is beneficial to GenCos. For instance, by using BP forecasts and without adopting learning algorithm, the WGenCo earns $1,389,162, whereas with learning the WGenCo obtains the net earnings of $2,383,078. This is an increase of around 71.5%. The advantage of GenCo learning in optimizing bidding strategy in the day-ahead electricity market is also reflected by the average locational market price. For instance, the average clearing prices for the WGenCo using BP forecasting results are $42.456/MWh and $37.677/MWh for using learning and not using learning, respectively.

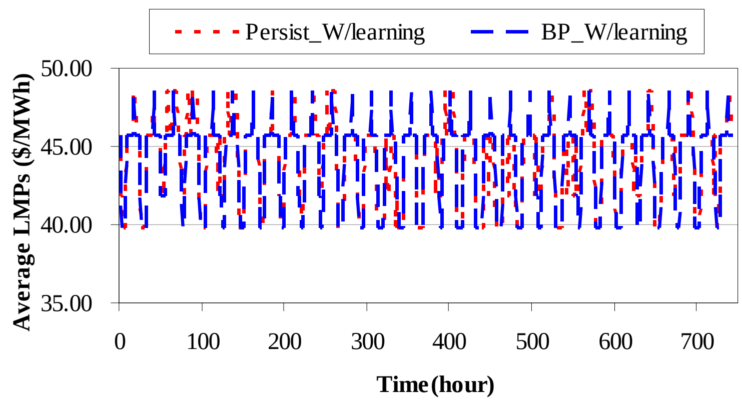
5.2.5. Comparison and discussion

Agent-based modeling approach can also provide detailed information about the complex interactions among different market participants and influencing factors, such as the hourly LMPs for each day, the supply and demand curves at some specific hour and/or for some specific day, etc. Based on the simulation results, more insights about the market structure as well as bidding optimization can be derived. For instance, Figure 32 illustrates the LMPs for the cases with and without GenCo learning at a 5% wind penetration level at bus 2 where the WGenCo is located, respectively. It can also be observed that with GenCo learning, the average clearing price is significantly higher, which means that the GenCos can increase their net earnings by adopting reinforcement learning in optimizing their bidding strategy.

A further comparison of Tables 22 and 23 reveals that the wind penetration level of 20% generally leads to comparatively low clearing prices compared with the lower level of 5%. For instance, with GenCo learning and using BP forecasts for the WGenCo, the average locational market price changes from $44.148/MWh at the level of 5% to $42.456/MWh at the level of 20%. This observation can be repeated for all the other cases as well. Moreover, more details can be revealed by checking the LMP evolution with respect to time. Figure 33 illustrates how LMP changes at bus 2 during the last day of the simulation period. It can be seen that, for the majority of the 24 hours, the LMPs at the higher wind penetration level are lower than those at the lower level, while for the remaining hours, the LMPs of the two penetration levels are equal. This explains why the average LMPs at 20% wind penetration level is higher than those at 5%. Therefore, it can be cautiously concluded that high wind penetration levels can generally help bring down the market clearing price of the day-ahead electricity market.



(a) Average LMPs without GenCo learning



(b) Average LMPs without GenCo learning

Figure 32. LMP evolutions at bus 2 for the scenarios at 5% wind penetration level.

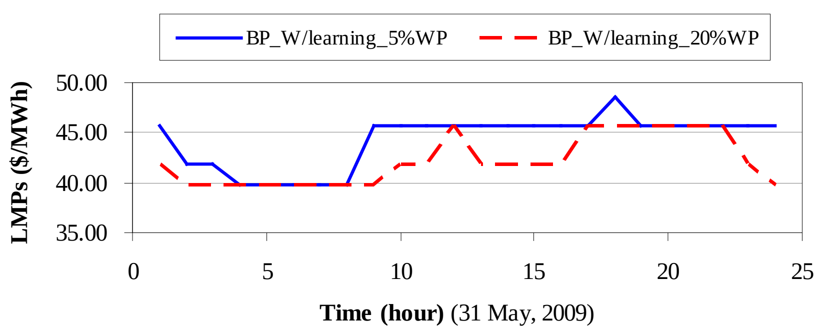


Figure 33. LMPs for the 24 hours of last auction day.

It should be mentioned that the LMPs at all nine buses are the same according to the simulation results. One main reason is that the transmission capacity of the grids is high enough so that no physical constraint exists. In this case, the MCP is the same as the LMP.

5.3. Agent-Based Double-Sided Bidding of Wind Power

Section 5.2 demonstrates the impacts of forecasting accuracy on WGenCo bidding optimization and provides some important insightful information. However, only the simplest electricity network with single-sided auction rule is considered. Further research is conducted in this section. The modified IEEE 30-bus 6-generator power system network (Anderson and Fouad, 2003) is adopted to represent the transmission grid in view that it is widely used in power system related research and more complex than the simple 9-bus system adopted in the previous section. The scenarios investigated also consider the influences of short-term wind power forecasts at two different wind penetration levels, 5% and 20%. The agent-based models are built for the double-sided auction market with limited upper price, based on the modified IEEE 30-bus grid system with constrained transmission capacities. One main intension of this study is to verify if the previous findings can repeat with this more complex and more realistic set-up.

5.3.1. Scenarios and simulation environment

Similarly, the WGenCo bids in the day-ahead electricity market by using the BP-based point estimates data, same as the one used in Section 5.2. However, to better represent the low and high penetration levels, the wind power capacity are scaled into 5% and 20% penetration levels, respectively. Also, two algorithms, with or without GenCos learning, are considered. So, there are totally four scenarios.

The auction procedures and pricing rules of the day-ahead electricity market, the agents' action domain construction and learning procedures, the calculation of costs and net earnings are all similar to those introduced in Section 5.2. The differences include (1) the price-sensitive demand bids are considered besides the fixed daily demand; (2) a modified IEEE 30-bus 9-GenCo power system network is assumed to be connected with the market operations.

Figure 34 gives the single ling diagram of the IEEE 30-bus 9-GenCo power system network, which has been widely used in power system related research (Li and Jiang, 2011; Liao and Wu, 2011). The generator at bus 30 is assumed to be the WGenCo, and the rest eight sellers are regarded as traditional GenCos. The GenCos' marginal cost function, the price-sensitive demand bid function and the power flow data of the modified IEEE 30-bus 9-generator 21-buyer power system are adopted from the case study of (Li et al., 2009).

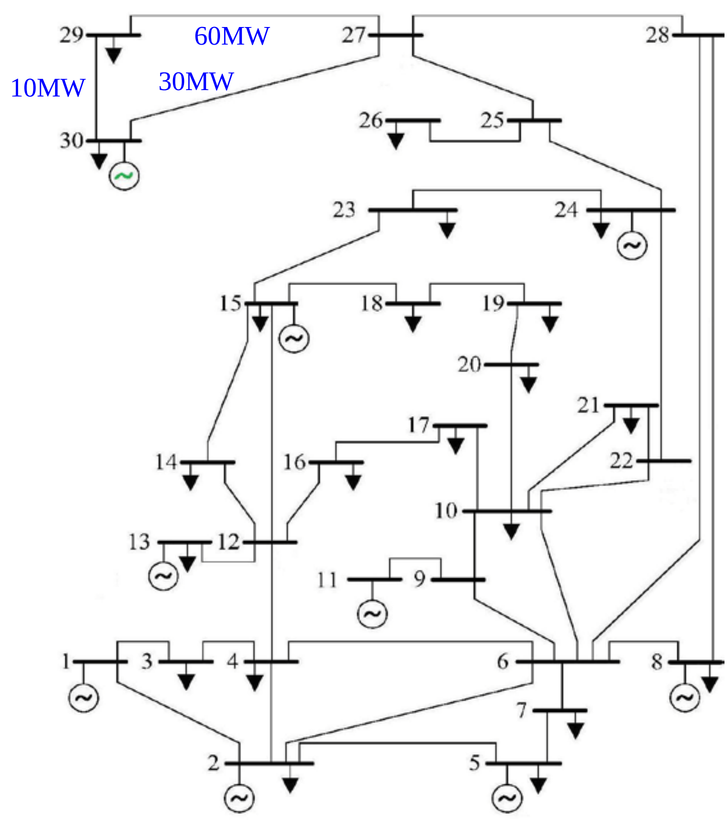


Figure 34. Single line diagram of the IEEE 30-bus network.

The upper capacity of WGenCo at bus 30 is specified according to the average of daily forecast wind power, as summarized in Table 24, where 47.153MW is the average hourly capacity of the last simulation day at the 20% penetration level. For purpose of illustration, the maximum demands and the original parameter c for the price-sensitive bid function of buyer 21 at bus 30 on the last simulation day are plotted in Figure 35, where its original parameter *d* for the price-sensitive bid function is set to be constant \$0.04/MW^{2}h under the assumption of linear marginal cost function.

Table 24. True marginal cost functions and capacity limits of the nine GenCos.

GenCo	Bus No.	a_i (\$/MWh)	b_i (\$/MW^{2}h)	Cap^L (MW)	Cap^U (MW)
GenCo1	1	11.295	0.005	0	90.000
GenCo2	2	17.212	0.007	0	90.000
GenCo3	5	13.123	0.008	0	60.000
GenCo4	8	13.254	0.009	0	40.000
GenCo5	11	37.859	0.013	0	30.000
GenCo6	13	19.326	0.012	0	60.000
GenCo7	15	18.325	0.007	0	50.000
GenCo8	24	38.895	0.016	0	30.000
WGenCo	30	20.158	0.007	0	47.153

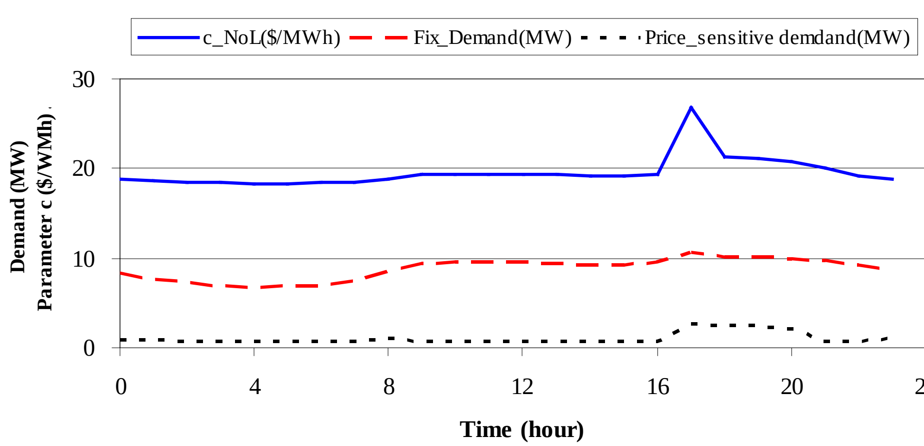


Figure 35. Demands and demand bid function parameters of buyer 21 at bus 30.

The VRE-RL learning algorithms is realized and combined with the DC-OPF dispatch module for all market participants. The action domains are constructed by bs_i, $M1_i$, $M2_i$, $M3_i$, $RIMax_i^L$, $RIMax_i^U$, and $RIMin_i^C$ which are selected as shown in Table 25. Besides, for the purpose of a fair comparison, all the GenCos learning processes are simulated with a recency parameter of 0.05 and an experimentation parameter of 0.95. Also, each GenCo's initial propensity $q_i(1)$ is set to equal its estimated maximum daily net earnings (MaxDNEi) and the cooling parameter T_i is set as one percent of the initial propensity.

Table 25. Parameters of action domain and learning parameters.

GenCo *i*	bs_i	$M1_i$	$M2_i$	$M3_i$	$RIMax_i^L$	$RIMax_i^U$	$RIMin_i^C$
1-9	0.001	10	10	1	0.75	0.75	1

5.3.2. Forecasting results

For BP-based forecasting, different input parameters and learning rates are tested to select the best combination in terms of mean absolute errors (MAE) and root mean squared errors (RMSE), and the model outputs are selected to generate the scaled wind power forecasts at 5% and 20% penetration levels. The average observed value, the average BP forecast, and the forecasting errors of wind power are summarized in Table 26, together with two average wind power values scaled to represent two penetration levels. In order to evaluate the influence of forecasting errors on the net earnings of the WGenCo, the total net earnings in the "perfect" case that all real hourly wind generations could be sold at the LMPs of bus 30 are also calculated for each scenario in the following results.

Figure 36 illustrates the real observations of wind power and the corresponding day-ahead forecasts from the best-performing BP model. It can be seen that for the 24-hour-ahead forecasting application, the forecasting errors of the BP model are not ignorable although they are still acceptable. In order to evaluate the above-mentioned influence of forecasting errors on the net earnings of the WGenCo, the total net earnings in the "perfect" case that all real hourly wind generations are assumed to be sold at the LMPs of bus 30 are also calculated for each scenario in the following results.

Table 26. BP forecasting results of wind power and average scaled generation values.

Ave_Obs	Ave_Forecast	MAE	RMSE	Ave_5%	Ave_20%
345.468KW	347.290KW	188.135	230.269	11.788MW	47.153MW

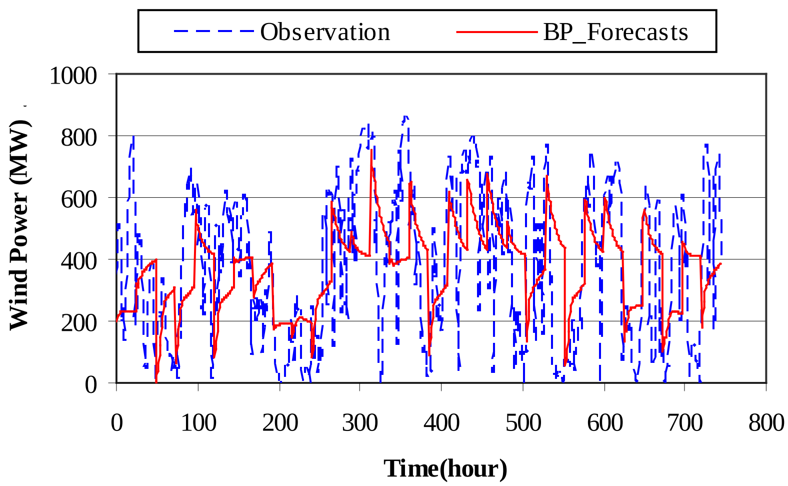


Figure 36. Time series plots of observations and BP forecasts.

5.3.3. Bidding for scenarios with 5% wind power

For the scenarios with the low (5%) wind penetration level, the simulated auction results for the WGenCo are summarized in Table 27. It can be seen that based on the same BP forecasting results, the total net earning of the WGenCo with the adoption of learning algorithm is $397,945, which is around 1.54 times more than the net earnings without GenCo learning ($156,629). This clearly shows the advantage of GenCo learning. It should be noted that the total quantity settled by the ISO for the case with learning is less than that without learning.

Table 27. Auction results for the WGenCo at 5% wind penetration level.

Scenarios	Qty_Gen (MWh)	Qty_Set(MWh)	Ave_LMP ($/MWh)	Tot_NE ($10^3$$)	Tot_NE_P ($10^3$$)
No_Learning	8724.241	6337.20	19.476	156.629	169.701
Learning		5037.200	49.556	397.945	429.587

Figure 37 illustrates the quantities of wind power settled by the ISO during the last five simulation days, for the cases both with and without learning. It is observed that without learning, the wind power could be either sold at a high amount or no power could be settled by the ISO. However, the LMPs determined for the case with learning are comparatively higher than those without learning. For example, in terms of the average value, the LMP at bus 30 for WGenCo with learning is $49.556/MWh, which is around 1.54 times more than the one obtained without learning ($19.476/MWh). This explains why the total net earnings with the adoption of learning are still much more than those obtained without learning.

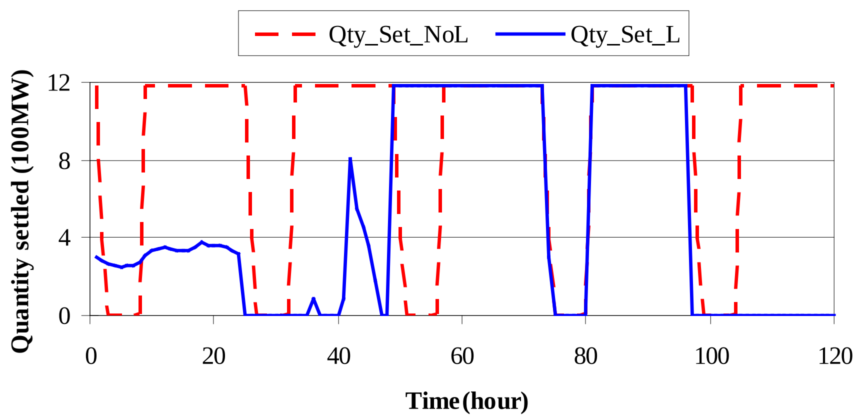


Figure 37. Quantities settled for last five simulation days.

Besides, it can be seen that the net earnings based on the wind power forecast with forecasting errors is less than the one that could be earned in the "perfect" case. To be exact, if the future wind generation were perfectly available, without learning, the WGenCo's total net earnings could be $169,701, which is about 8.3% more than current one ($156,629). For the cases with the adoption of GenCos learning, the improvement could be around 8.0% as well. Both indicate that the forecasting errors could bring negative impacts on the maximization of the net earnings for the WGenCo.

5.3.4. Bidding for scenarios with 20% wind power

The simulated auction results for the scenarios at the 20% wind penetration level are summarized in Table 28. Again, the influences of GenCo learning and forecast errors can be observed from this table. For the case with learning, the WGenCo obtains a total net earning of $1,397,618, which is about 1.24 times more than that obtained from the case without learning ($622,865), both based on the same BP forecast values. This advantage of GenCo's adopting learning algorithms in optimizing its bidding strategy in the day-ahead electricity market can be directly observed from the average LMP as well. For instance, the average LMP of the case with GenCo learning reaches $43.454/MWh, about 1.24 times higher than that obtained in the case without GenCo learning ($19.405/MWh). Similarly, compared with the "perfect" case, the case based on the wind power forecast with errors, generates obviously less net earnings, either with or without the adoption of learning, indicating the negative impacts of forecasting errors. If the future real wind generation were perfectly available, without learning, the WGenCo's total net earnings could be $676,499, which is about 8.6% more than current one ($622,865). For the cases with GenCos learning, this improvement could be around 8.2% as well.

Table 28. Auction results for the WGenCo at 20% wind penetration level.

Scenarios	Qty_Gen (MWh)	Qty_Set (MWh)	Ave_LMP ($/MWh)	Tot_NE ($10^3$$)	Tot_NE_P ($10^3$$)
W/O learning	34896.963	21912.040	19.405	622.865	676.499
W/ learning		17780.130	43.454	1397.618	1511.737

Besides, it could be observed that the quantities of wind power settled in all scenarios are less than the real wind generations. One possible reason lies in the simplification of using the daily average to represent the upper capacity of the WGenCo. It could be better to use the real-time forecast value as the reference of the reported upper capacity instead of using the daily average value. Besides, after checking the physical constraint and local demand, it can be seen that the maximum capacity of the two branches, from bus 30 to bus 29 and bus 27, respectively, are both 30MW. This indicates that whenever the wind power is larger than 60MW, the surplus definitely cannot go through the power grid. Therefore, with the increased wind power installation, the WGenCos need to either broaden of the grid network or combine their wind generation with some other energy storage or conversion systems.

Figure 38 illustrates the evolution of LMPs at bus 30 determined for the last day of each simulation scenario. It can be seen that without the adoption of learning algorithm, the WGenCo just simply bids on its real marginal cost function; and correspondingly, the LMPs at two penetration levels are almost the same.

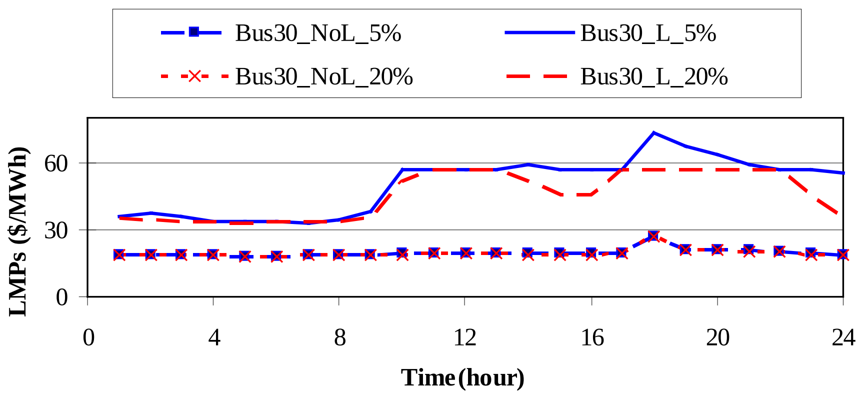


Figure 38. Comparison of LMPs at bus 30 during the last simulation day.

However, at each specific penetration level, with the adoption of GenCo learning, the corresponding LMP values increase significantly. In particular, it can be observed that overall, smaller LMPs are achieved at the higher penetration level. Without learning, the average LMP at bus 30 (for WGenCo) with 20% wind penetration is only about 0.4% less than the one with 5% penetration level; with learning, the decrease will be slightly large, around 12.3%. This implies that a higher wind penetration level may help reduce the LMPs. However, it should also be noted that with a higher wind penetration level, other costs as well as the security concern could increase, which is not considered in current simulation scenarios though. The optimal penetration level by considering the balance of benefits and negative effects is an interesting topic worth further investigation.

Besides, it should be noted that based on the agent-based models, the GenCos can learn from not only their own experiences but also their competitors and the entire market, even though part of such information is absent or not publicly announced. This is one of the typical advantages of agent-based modeling methods. During the bidding process, the WGenCo actually does not know the competitors' reported capacity or marginal price information. However, with learning ability, the WGenCo can dynamically change and select the best actions within the action domains, Based on the newest prosperities, which are dynamically affected by the competitor's bidding behavior as well as other market factors, the action of the highest probability to increase the WGenCo's net earnings will be automatically selected. In other words, the learning module can automatically take into account the changes of the bidding environment which are affected by all market components including both the sellers and their competitors.

6. CONCLUSIONS AND FUTURE WORK

This chapter mainly summarizes the findings and contributions of this dissertation. Besides, some meaningful topics and directions are suggested for future research.

6.1. Conclusions

Wind energy is becoming the world's fastest growing source of clean and renewable energy. The predictability of wind generation is essential for both wind farm operations management and the integration of wind energy into the power system. This is highly related to accurate and reliable short-term wind forecasting, which remains a challenge due to the continuous temporal and spatial fluctuations of the wind resource. Therefore, adaptive and reliable modeling methods are urgently needed for accurate and reliable forecasts of wind speed and wind power. Meanwhile, with the ongoing electricity market liberalization as well as the increasing wind power capacity, wind generation companies are expected to directly trade the wind energy in the electricity markets. Another issue of particular importance for wind generation companies is how to maximize their net earnings by optimizing their bids in the gradually deregulated electricity market. However, the existing research efforts towards the above-mentioned two issues are still far from sufficient. This has directly stimulated this dissertation research. The research tasks conducted in this dissertation study, as well as the corresponding findings and contributions, are summarized as follows.

First, after a thorough review on currently available forecasting methods, two types of time series forecasting models, Box-Jenkins models and artificial neural networks are investigated, implemented, and evaluated, respectively. A comprehensive study is firstly performed to evaluate and compare the forecasting accuracies of different Box-Jenkins models including AR, ARMA, and ARIMA, followed by a further study on the applications of three typical kinds of artificial neural networks, namely, BP, RBF, and ADALINE, in performing short-term wind speed forecasting. The results indicate that for one-hour-ahead wind forecasting, no significant difference in the forecasting accuracies among the models studied. Especially, the results show that the forecasting performances of different models change from site to site, from time to time. This confirms the need of this study – multiple types of ANN models should be evaluated before the most suitable type can be determined. Also, it provides a useful reference for future research.

Meanwhile, in many practical situations of the wind energy industry, a final single forecast that could take advantage of a set of plausible forecasts needs to be produced. For example, the forecasts from alternative forecast agencies should be used since there is not a superior agency. Meanwhile, the forecast agencies themselves might make forecasts for the client by adopting alternative models or procedures such as different ANN models. In order to provide a single forecast, the agency needs to combine all the available information. Therefore, it is apparent that an efficient forecast combination procedure might be of great importance for wind speed forecast. In view of this, after a survey over various methods applied in other research fields, the adaptive Bayesian model averaging based modeling method is adopted and investigated.

The BMA algorithm is, for the first time, adopted for modeling long-term wind speed distribution. The results reveal that while no single candidate distribution in the model space universally outperforms others in estimating the wind distributions, whereas the BMA PDFs, derived by averaging many candidate models with their posterior model probabilities, are always suitable for describing the wind speed distributions with high accuracy. In this way, the BMA approach could provide a unified solution for statistical modeling of long term wind speed distribution with high reliability and robustness. By applying this approach, the confusion of using traditional goodness-of-fit metrics for model evaluation can also be avoided. If more than one distribution is considered plausible, the BMA model will include all the plausible models into its PDF. In the cases where only one outperforming model exists according to the posterior probability, the BMA PDF will overlap with that distributional PDF. The findings together with the proposed model could benefit the users greatly in wind farm siting, long-term wind potential estimation, etc.

Motivated by this success, we further apply the BMA method to improve the short-term wind speed forecasts of typical neural network models. It is discovered that the preferred choices of ANN models (in terms of type and optimal parameters) are inconsistent with different performance metrics and vary from site to site. In other words, none of the models are universally superior to others. However, by applying the proposed two-step forecasting method to the candidate models, one single time series of forecasts are derived. Especially, the BMA model always performs well for different sites, demonstrating its capability of being highly adaptive and robust. This contribution could benefit the industry in the operation of wind farms and the integration of wind energy into the power systems. As a result, risks due to forecast error can be minimized.

Finally, to analyze the effects of the forecasting accuracies of wind power on the bidding strategy optimization or the maximization of selling wind power in the electricity markets, we conduct two agent-based simulation studies on bidding wind power in a day–ahead electricity market with either single-sided or double-sided auction protocols, by using the point estimation of hourly wind generation. Two wind penetration levels, low (5%) and high (20%), are also investigated and compared, respectively. Under the assumed market and system settings, it is found that (1) generally, higher forecasting accuracy of wind power generation can lead to higher net earnings for the WGenCo; (2) applying learning algorithms can help increase the net earnings of the day-ahead electricity market participants; (3) increasing wind penetration level could help reduce the market clearing price; and (4) the agent-based modeling approach can provide insights for the interactions among different market structures and influence factors.

6.2. Future Work

While the wind installation capacity keeps increasing and the electricity market continuously experiences reconstruction or liberalization, both issues studied in this dissertation still remain as hot challenging research topics.

As for the first issue on how to improve the short-term wind forecasting accuracy, the following directions and areas are of high interest and importance.

- Applying BMA and other Bayesian methods: As reviewed in (Li and Shi, 2012), Bayesian methods provide powerful tools for short-term wind forecasting as well as for solving other issues during the wind energy utilization process. However, the attention paid to this direction is far from enough. One possible reason is that Bayesian methods usually involve large complexity and computations. Along with the development of high-speed computation technology, this will not be the obstacle any longer. Meanwhile, it is also attractive to include different types of component models in applying BMA method. Currently, the model space only contains time series based models. However, other types of models, e.g., physic models can also be included if possible.
- Developing advanced artificial intelligence models and novel training algorithms: Currently, some advanced artificial intelligence models have been proposed and

successfully employed in other application areas. Therefore, it is attractive to introduce them into the area of wind forecasting so as to further improve the forecasting accuracy. Meanwhile, for ANN models currently applied in this area, some advanced training algorithms are expected to improve models' convergence speeds or deal with such problems as over-fitting, non-convergence, etc.

- Developing new hybrid models: One way is to integrate currently available models, e.g., using the output from the physical models as the inputs of statistical models so as to improve the forecasting accuracy. Another method is borrowing ideas from other research areas.
- Forecasting wind power over a large area: Examining larger areas might result in better overall forecasts since the wind forecasting errors or the wind power variations between different wind turbines distributed over a large area could usually compensate or cancel out to some degree. This has been demonstrated in several studies (Focken et al., 2001; Holttinen, 2005; Ostergaard, 2008). Along with the rapid development of wind power, this area is to be further investigated.

As for the second issue on how to optimize the WGenCo's bidding strategy in the electricity market, since the electricity wholesale market is still under reconstruction, many interesting and meaningful issues related to bidding optimization under various possible market designs should be investigated and tackled, which include, but are not limited to, the following topics:

- Bidding with probabilistic wind forecasts: The application of probabilistic wind forecasts has demonstrated its effectiveness in optimizing the bidding strategy in some literature. However, such researches are expected to be further carried out.

Due to the limitation of resources, this part has not been investigated. A preliminary cast study based on the BMA forecast indicates that this direction can be a great research topic in the near future.

- Bidding under double-sided auction mechanism: Allowing the buyers as well as the sellers to submit their competitive bids, double-sided auctions are commonly regarded as a better setting for deregulating the electricity markets. However, bidding under such settings may also face with increased uncertainty and risks, and this calls for quantitative modeling analysis.
- Bidding under different market prototypes: Various electricity market structures are being designed and tested to ensure free access, fair competition, high efficiency, and systems security and reliability. The analysis on GenCos' strategic bidding or market power is important for the market design efforts.
- Bidding under PAB auction rule: Bidding strategy under the discriminatory pricing rule has its advantages and disadvantages in deregulating the market. However, the relevant research efforts are still far from enough.
- Bidding in hybrid or combined markets: The hybrid market often exists in which electrical energy is traded together with spinning reserves simultaneously or in the presence of future contracts and bilateral contracts. Also, with the combination of day-ahead, hour-ahead, and real-time markets, many reconstructed electricity markets allow the participants to update day-ahead supply offers and purchasing bids before the actual delivery. Besides, the future electricity markets, by combining the wholesale and retailing markets, should allow the end users to bid in

the markets directly. Bidding optimization in the combined markets will be a significant topic for research.

- Advance in modeling methods and algorithms: To better represent the transaction behaviors in complex electricity markets, the shortcomings of each method should be overcome by continuously pushing for new theoretical developments.
- Bidding cooperatively or with internal control strategy: Without affecting the oligopolistic nature of the deregulated electricity market, locational GenCos, especially for those owning different generation resources, could bid in the market cooperatively. For example, a WGenCo can cooperate with neighboring WGenCos or other type of GenCos, or the WGenCo can adopt some internal balancing strategies, e.g., integrating the wind generation with storage technology. The effects of such cooperative bidding strategies or internal controls should be further investigated.
- Risk management under uncertainty: To deal with the risks of bidding under the uncertainty in demand, price and production, different risk control measures and operational scenarios need to be investigated. This is especially important for those renewable GenCos with distributed and intermittent productions. For instance, besides the uncertain load and electricity price, the wind generators should also consider the uncertain wind generations before submitting their supply offers to the electricity markets.

REFERENCES

Aggarwal S, Saini L, and Kumar A, 2009. Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power and Energy Systems* 31(1):13–22.

Ahn N, and Niemeyer V, 2007. Modeling market power in Korea's emerging power market. *Energy Policy* 35(2):899–906.

Akaike H, 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19(6):716–23.

Al-Agtash SY, 2010. Supply curve bidding of electricity in constrained power networks. *Energy* 35(7):2886–92.

Alagtash S, and Yamin HY, 2004. Optimal supply curve bidding using benders decomposition in competitive electricity markets. *Electric Power Systems Research* 71(3):245–55.

Alexiadis M, Dokopoulos PS, Sahsamanoglou HS, and Manousaridis IM, 1998. Short-term forecasting of wind speed and related electrical power. *Sol. Energy* 63(1):61–8.

Al-Hamadi HM, and Soliman SA, 2004. Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model, *Electr. Power Syst. Res.* 68(1):47–59.

Amjady N, 2006. Day-ahead price forecasting of electricity markets by a new fuzzy neural network. *IEEE Trans. Power Syst.* 21(2):887–96.

Amjady N, and Keynia F, 2008. Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. *Int. J. Electr. Power Energy Syst*. 30(9):533–46.

Amjady N, and Keynia F, 2009. Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm. *IEEE Trans. Power Syst.* 24(1):306–18.

Amjady N, and Daraeepour A, 2009. Mixed price and load forecasting of electricity markets by a new iterative prediction method. *Electric Power Systems Research* 79(9):1329–36.

Amjady N, and Keynia F, 2009. Day-ahead price forecasting of electricity markets by a new feature selection algorithm and cascaded neural network technique. *Energy Conversion and Management* 50(12):2976–82.

Anderson JA, 2009. Electricity restructuring: a review of efforts around the world and the consumer response. *The Electricity Journal* 22(3):70–86.

Angarita JL, Usaola J, and Martínez-Crespo J, 2009. Combined hydro-wind generation bids in a pool-based electricity market. *Electric Power Systems Research* 79(7):1038–46.

Armstrong M, Cowan S, and Vickers J, 1994. Regulatory Reform - Economic Analysis and British Experience, Oxford University Press: Oxford.

Attaviriyanupap P, Kita H, Tanaka E, and Hasegawa J, 2005. New bidding strategy formulation for day-ahead energy and reserve markets based on evolutionary programming. *International Journal of Electrical Power & Energy Systems* 27(3):157–67.

Azadeh A, Skandari MR, and Maleki-Shoja B, 2010. An integrated ant colony optimization approach to compare strategies of clearing market in electricity markets: Agent-based simulation. *Energy Policy* 38(10):6307–19.

Azevedo EM, and Correia PB, 2006. Bidding strategies in Brazilian electricity auctions. *International Journal of Electrical Power & Energy Systems* 28(5):309–14.

Badri A, Jadid S, Rashidinejad M, and Moghaddam MP, 2008. Optimal bidding strategies in oligopoly markets considering bilateral contracts and transmission constraints. *Electric Power Systems Research* 78(6):1089–98.

Baldick R, Grant R, and Kahn E, 2004. Theory and application of linear supply function equilibrium in electricity markets. *Journal of Regulatory Economics* 25(2):143–67.

Bathurst GN, Bathurst GN, Weatherill J, Weatherill J, Strbac G, and Strbac G, 2002. Trading wind generation in short term energy markets. *IEEE Transactions on Power Systems* 17(3):782–9.

Bofinger S, Luig A, and Beyer HG, 2002. Qualification of wind power forecasts. *Proc. Global Wind Power Conf.*, Paris, France, April 2–5, 2002.

Bompard E, Lu W, Napoli R, and Jiang X, 2010. A supply function model for representing the strategic bidding of the producers in constrained electricity markets. *International Journal of Electrical Power & Energy Systems* 32(6):678–87.

Boonchuay C, and Ongsakul W, 2011. Optimal risky bidding strategy for a generating company by self-organising hierarchical particle swarm optimisation. *Energy Conversion and Management* 52(2):1047–53.

Borenstein S, Bushnell J, and Knittel CR, 1998. Market power in electricity markets: beyond concentration measures. PWP-059, University of California Energy Institute, Berkeley, CA.

Borenstein S, and Bushnell J, 1999. An empirical analysis of the potential for market power in California's electricity industry. *Journal of Industrial Economics* 47(3):285–323.

Borghetti A, Frangioni A, Lacalandra F, Nucci CA, and Pelacchi P, 2003. Using of a cost-based unit commitment algorithm to assist bidding strategy decisions. In: *Proc. of 2003 IEEE Bologna Power Tech Conference* Bologna, Italy, June 23–26, 2003.

Borghetti A, Massucco S, and Silvestro F, 2009. Influence of feasibility constrains on the bidding strategy selection in a day-ahead electricity market session. *Electric Power Systems Research* 79(12):1727–37.

Bourry F, Juban J, Costa LM, and Kariniotakis G, 2008. Advanced strategies for wind power trading in short-term electricity markets. In: *Proc. of European Wind Energy Conference & Exhibition* (EWEC 2008), Belgium. Brussels, March 31–April 3, 2008.

Bourry F, and Kariniotakis G, 2009. Strategies for wind power trading in sequential short-term electricity markets. In: *Proc. of European Wind Energy Conference & Exhibition* (EWEC 2009), Marseille, April 16–19, 2009.

Box GEP, and Jenkins GM, 1994. Time Series Analysis: Forecasting and Control, 3rd Edition. Prentice Hall.

Brockwell PJ, and Davis RA, 2002. Introduction to time series and forecasting. New York: Springer.

Brooks S, and Gelman A, 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7:434–55.

Brunekreeft G, 2001. A multiple-unit, multiple-period auction in the British electricity spot market. *Energy Economics* 23(1):99–118.

Bunn DW, Martoccia M, Ochoa P, Kim H, Ahn NS, and Yoon YB, 2010. Vertical integration and market power: A model-based analysis of restructuring in the Korean electricity market. *Energy Policy* 38(7):3710–16.

Bunn DW, and Oliveira FS, 2003. Evaluating individual market power in electricity markets via agent-based simulation. *Annals of Operations Research* 121(1–4):57–77.

Bunn DW, 2000. Forecasting loads and prices in competitive power markets, *Proc. IEEE* 88 (2):163–9.

Burnham KP, and Anderson DR, 2002. Model Selection and Multimodel Inference, second ed. New York: Springer.

Carta JA, Ramírez P, and Velázquez S, 2009. A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands. *Renewable and Sustainable Energy Rev.* 13(5):933–55.

Celik AN, 2004. On the distributional parameters used in assessment of the suitability of wind speed probability density functions. *Energy Convers. Manage* 45:1735–47.

Christensen JF, Sorensen AG, Hatziargyriou N, and Donnelly M, 1998. Methods and models for evaluating the impact of decentralized generation. CIGRE (*International Conference on large High Voltage Electric Systems*), Paris, France.

Conejo AJ, Plazas MA, Espinola R, and Molina AB, 2005. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Trans. Power Syst.* 20(2):1035–42.

Congdon P, 2006. Bayesian statistical modeling. Hoboken: John Wiley & Sons.

Conejo AJ, and Prieto FJ, 2001. Mathematical programming and electricity markets. Sociedad de Estad\ística e Investigaci\ón Operativa TOP 9(1):1–53.

Conejo AJ, Nogales FJ, and Arroyo JM, 2002. Price-taker bidding strategy under price uncertainty. *IEEE Trans. Power Syst.* 17(4):1081–8.

Cormack J, Hollis A, Zareipour H, and Rosehart W, 2010. The large-scale integration of wind generation: impacts on price, reliability and dispatchable conventional suppliers. *Energy Policy* 38(7):3837–46.

Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, and Feitosa E, 2008. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews* 12(6):1725–44.

Crawford G, Crespo J, and Tauchen H, 2007. Bidding asymmetries in multi-unit auctions: Implications of bid function equilibria in the British spot market for electricity. *International Journal of Industrial Organization* 25(6):1233–68.

Dale L, Milborrow D, Slark R, and Strbac G, 2004. Total cost estimates for large-scale wind scenarios in uk. *Energy Policy* 32(17):1949–56.

David AK, and Wen F, 2000. Strategic bidding in competitive electricity markets: a literature survey. *Proc. of IEEE 2000 Power Engineering Society Summer Meeting* 4:2168–73.

David AK, 1993. Competitive bidding in electricity supply. *IEE Proc.: Generation, Transmiss. Distrib.* 140(5):421–6.

DeCesaro J, Porter K, and Milligan M, 2009. Wind energy and power system operations: A review of wind integration studies to date. *The Electricity Journal* 22(10):34–43.

De la Torre S, Arroyo JM, Conejo AJ, and Contreras J, 2002. Price maker self-scheduling in a pool-based electricity market: a mixed-integer LP approach. *IEEE Trans. Power Syst.* 17(4):1037–42.

De la Torre S, Conejo AJ, and Contreras J, 2004. Finding multiperiod Nash equilibria in pool-based electricity markets. *IEEE Transactions on Power Systems* 19(1):643–51.

De Ladurantaye D, Gendreau M, and Potvin JY, 2007. Strategic bidding for price-taker hydroelectricity producers. *IEEE Transactions on Power Systems* 22(4):2187–203.

Delarue ED, Luickx PJ, and D'haeseleer WD, 2009. The actual effect of wind power on overall electricity generation costs and co2 emissions. *Energy Conversion and Management*

50(6):1450–56.

Demeo EA, Grant W, Milligan MR, and Schuerger MJ, 2005. Wind plant integration. Costs, status, and issues. *IEEE Power &Energy Magazine* 3(6):38–46.

Douglas AP, Breipohl AM, Lee FN, and Adapa R, 1998. Risk due to load forecast uncertainty in short term power system planning, *IEEE Trans. Power Syst.* 13(4):1493–99.

Dragoon K, 2003. Assessing wind integration costs with dispatch models: a case study of PacifiCorp. In: *Proc. of WINDPOWER 2003* Austin, Texas, May 18–21, 2003.

Duan Q, Ajami N, Gao X, and Sorooshian S, 2007. Multi-model ensemble hydrologic prediction using bayesian model averaging. *Advances in Water Resources* 30(5):1371–86.

Durán MJ, Cros D, and Riquelme J, 2007. Short-term wind power forecast based on ARX models. *Journal of Energy Engineering* 133(3):172–80.

Erasmo C., and Wilfrido R, 2007. Wind speed forecasting in the South Coast of Oaxaca, Mexico. *Renewable Energy* 32(12):2116–28.

Ernst D, Minoia A, and Marija I, 2004. Market dynamics driven by the decision-making of power producers. In: *Proc. of 2004 IREP Symposium - Bulk Power System Dynamics and Control – VI*, 1–6.

Fabbri A, Gómez T, Rivier J, and Méndez V, 2005. Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market. *IEEE Trans on Power Systems* 20(3):1440–46.

Fan S, and Chen L, 2006. Short-term load forecasting based on an adaptive hybrid method, *IEEE Trans. Power Syst.* 21(1):392–401.

Fan S, Mao C, and Chen L, 2007. Next-day electricity-price forecasting using a hybrid network, *IET Gen. Transm. Distrib.* 1(1):176–82.

Federico G, and Rahman D, 2003. Bidding in an electricity pay-as-bid auction. *Journal of Regulatory Economics* 24(2):175–211.

Fleten S, and Kristoffersen T, 2007. Stochastic programming for optimizing bidding strategies of a Nordic hydropower producer. *European Journal of Operational Research* 181(2):916–28.

Focken U, Lange M, and Waldl HM, 2001. Previento - a wind power prediction system with an innovative upscaling algorithm. In *Proc. of European Wind Energy Conference*, Copenhagen, July 1–5, 2001.

Focken U, Lange M, Mo¨nnich K, Waldl H-P, Georg Beyer H, and Luig A, 2002. Short-term prediction of the aggregated power output of wind farms – a statistical analysis of the reduction of the prediction error by spatial smoothing effects. *Journal of Wind Engineering and Industrial Aerodynamics* 90(3):231–46.

Foley AM, Gallachóir BPó, Hur J, Baldick R, and McKeogh EJ, 2010. A strategic review of electricity systems models. *Energy* 35(12):4522–30.

Fraley C, Raftery AE, Gneiting T. and Sloughter JM, 2007. ensembleBMA: An R Package for Probabilistic Forecasting using Ensembles and Bayesian Model Averaging, *Technical Report No. 516*, Department of Statistics, University of Washington.

Fujii Y, Okamura T, Inagaki K, and Yamaji K, 2004. Basic analysis of the pricing processes in modeled electricity market with multi-agent simulation. *IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies*, Hong Kong, 1:257–61.

Gajjar GR, Khaparde SA, Nagaraju P, and Soman SA, 2003. Application of actor-critic learning algorithm for optimal bidding problem of a GENCO. *IEEE Trans. Power Syst.* 18(1):11–8.

Galloway S, Bell G, Burt G, McDonald J, and Siewierski T, 2006. Managing the risk of trading wind energy in a competitive market. *IEE Proc.: Generation, Transmission and Distribution* 153(1):106–14.

Gao C, Bompard E, Napoli R, Wan Q, and Zhou J, 2008. Bidding strategy with forecast technology based on support vector machine in the electricity market. *Physica A: Statistical Mechanics and its Applications* 387(15):3874–81.

Gao F, and Sheble GB, 2010. Electricity market equilibrium model with resource constraint and transmission congestion. *Electric Power Systems Research* 80(1):9–18.

Gao F, Gutierrez-Alcaraz G, and Sheble GB, 2006. Comparison of artificial life techniques for market simulation. In: *Proc. of 39th Hawaii International Conference on System Sciences* (HICSS'2006) 10:243a.

Garcia-Martos C, Rodriguez J, and Sanchez MJ, 2007. Mixed models for short-run forecasting of electricity prices: application for the Spanish market, *IEEE Trans. Power Syst.* 22(2):544–51.

Garcia RC, Contreras J, van Akkeren M, and Garcia JBC, 2005. A GARCH forecasting model to predict day-ahead electricity prices, *IEEE Trans. Power Syst.* 20(2):867–74.

Gelfand AE, and Smith AFM, 1990. Sampling based approach to calculating marginal densities. *J. Amer. Statistical Assoc.* 85:398–409.

Gelman A, Carlin JB, Stern HS, and Rubin DB, 2003. Bayesian data analysis (2nd Edition). Boca Raton:Chapman Hall and CRC Press.

Genc TS, and Reynolds SS, 2011. Supply function equilibria with capacity constraints and pivotal suppliers. *International Journal of Industrial Organization* 29(4):432–42.

Georgilakis P, 2008. Technical challenges associated with the integration of wind power into power systems. *Renewable and Sustainable Energy Reviews* 12(3):852–63.

Giabardo P, Zugno M, Pinson P, and Madsen H, 2010. Feedback, competition and stochasticity in a day ahead electricity market. *Energy Economics* 32(2):292–301.

Gibbons J, Cox G, Wood A, Craigon J, Ramsden S, Tarsitano D, and Crout N, 2008. Applying bayesian model averaging to mechanistic models: An example and comparison of methods. *Environmental Modelling & Software* 23(8):973–85.

Giebel G., Brownsword R., and Kariniotakis G, 2003. The state-of-the-art in short term prediction of wind power: a literature overview. *Project ANEMOS, Deliverable Report D1.1, 2003.*

Global Wind Energy Association (GWEA), 2011. World Wind Energy Report 2010. URL:http://www.wwindea.org/home/images/stories/pdfs/worldwindenergyreport2010_s.pdf. Accessed on January 10, 2012.

Gountis VP, and Bakirtzis AG, 2004. Bidding strategies for electricity producers in a competitive electricity marketplace. *IEEE Transactions on Power Systems* 19(1):356–65.

Green R, and Newbery DM, 1992. Competition in the British electricity spot market. *J. Political Econ.* 100(5):929–53.

Gross G, Finlay DJ, and Deltas G, 1999. Strategic bidding in electric generation supply markets. In: *Proc. of IEEE Power Engineering Society Winter Meeting*, 309–15.

Gross G, and Galiana FD, 1987. Short term load forecasting, *Proc. IEEE* 75(12):1558–73.

Guan X, Ho Y, and Lai F, 2001. An ordinal optimization based bidding strategy for electric power suppliers in the daily energy market. *IEEE Transactions on Power Systems* 16(4):788–97.

Guler O, 2009. Wind energy status in electrical energy production of turkey. *Renewable and Sustainable Energy Reviews* 13(2):473–8.

Haas R, and Auer H, 2006. The prerequisites for effective competition in restructured wholesale electricity markets. *Energy* 31(6–7):857–64.

Haghighat H, Seifi H, and Kian AR, 2007. Gaming analysis in joint energy and spinning reserve markets. *IEEE Transactions on Power Systems* 22(4):2074–85.

Haghighat H, Seifi H, and Rahimikian A, 2008. The role of market pricing mechanism under imperfect competition. *Decision Support Systems* 45(2):267–77.

Hao S, 2000. A study of basic bidding strategy in clearing pricing auctions. *IEEE Transactions on Power Systems* 15(3):975–80.

Hobbs BF, 1999. Models of Cournot competition in POOLCO and bilateral markets. In: *Proc. of*

American Power Conference, Chicago, USA.

Hobbs BF, Metzler CB, and Pang J, 2000. Strategic gaming analysis for electric power systems: an MPEC approach. *IEEE Trans. Power Syst.* 15(2):638–45.

Hoeting JA, Madigan D, Raftery AE, and Volinsky CT, 1999. Bayesian Model Averaging: A Tutorial. *Statistical Science* 14:382–417.

Holmberg P, 2009. Numerical calculation of an asymmetric supply function equilibrium with capacity constraints. *European Journal of Operational Research* 199(1):285–95.

Holttinen H, 2005. Optimal electricity market for wind power. *Energy Policy* 33(16):2052–63.

Hong W, 2009. Electric load forecasting by support vector model. *Applied Mathematical Modelling* 33(5):2444–54.

Hu S, Kapuscinski R, and Lovejoy WS, 2010. Bertrand-edgeworth auction with multiple asymmetric bidders. *Social Science Research Network Working Paper Series* (August 16, 2010). Available at SSRN: http://ssrn.com/abstract=1659996. Accessed on August 31, 2011.

Hu X, Grozev G, and Batten D, 2005. Empirical observations of bidding patterns in Australia's national electricity market. *Energy Policy* 33(16):2075–86.

Huang SJ, and Shih KR, 2003. Short-term load forecasting via ARMA model identification including non-Gaussian process considerations, *IEEE Trans. Power Syst.* 18(2):673–9.

Ji GR, Han P, and Zhai YJ, 2007. Wind speed forecasting based on support vector machine with forecasting error estimation. *Machine Learning and Cybernetics, 2007 International Conference on*, 2735–39.

Kaelbling LP, Littman ML, and Moore AW, 1996. Reinforcement learning: a survey. *J. Artif. Intell. Res.* 4(5):237–85.

Kang D, Kim B, and Hur D, 2007. Supplier bidding strategy based on non-cooperative game theory concepts in single auction power pools. *Electric Power Systems Research* 77(5–6):630–6.

Kariniotakis G, Stavrakakis G, and Nogaret E, 1996. Wind power forecasting using advanced neural network models. *IEEE Trans Energy Conversion* 11(4):762–7.

Kariniotakis G. et al., 2004. What performance can be expected by short term wind power prediction models depending on site characteristics? *Proc. Of EWEC 2004*, London.

Kariniotakis G et al., 2006. Evaluation of Advanced Wind Power Forecasting Models –Results of the Anemos *Project. EWEC 2006*. Athens.

Kass RE, and Raftery AE, 1995. Bayes factors. *Journal of the American Statistical Association*

90(430):773–95.

Kass RE, and Wasserman L, 1995. A reference Bayesian test for nested hypotheses with large samples. *Journal of the American Statistical Association* 90(431):928–34.

Kian A, Cruz JB, and Thomas RJ, 2005. Bidding strategies in oligopolistic dynamic electricity double-sided auctions, *IEEE Transactions on Power Systems* 20 (1):50–8.

Kisi O, 2007. Comparison of different ANN techniques in river flow prediction. *Civ. Eng. Environ. Syst.* (24)3:211–31

Klemperer PD, and Meyer MA, 1989. Supply function equilibrium in oligopoly under uncertainty. *Econometrica* 57(6):1243–77.

Krause T, Beck E, Cherkaoui R, Germond A, Andersson G, and Ernst D, 2006. A comparison of nash equilibria analysis and agent-based modelling for power markets. *International Journal of Electrical Power & Energy Systems.* 28(9):599–607.

Lerner J, Grundmeyer M, and Garvert M, 2009. The importance of wind forecasting. *Renewable Energy Focus* 10(2): 64–6.

Li, G, Liu CC, Mattson C, and Lawarree J, 2007. Day-ahead electricity price forecasting in a grid environment, *IEEE Trans. Power Syst.* 22(1):266–74.

Li G, and Shi J, 2010a. On comparing three artificial neural networks for wind speed forecasting. *Applied Energy* 87(7):2313–20.

Li G, and Shi J, 2010b. Comparison of different time series methods for short-term forecasting of wind power production. In *Proc. of ASME 2010 4th International Conference on Energy Sustainability*, Phoenix, Arizona, USA, May 17–22, 2010.

Li G, and Shi J, 2010c. Application of Bayesian model averaging in modeling long-term wind speed distributions. *Renewable Energy* 35(6):1192–202.

Li G, Shi J, and Zhou J, 2011. Bayesian adaptive combination of short-term wind speed forecasts from neural network models. *Renewable Energy* 36(1):352–9.

Li G, Shi J, and Zhou J, 2010. Effect of time horizon on short-term wind speed forecasting. In *Proc. of the 2010 Industrial Engineering Research Conference,* A. Johnson and J. Miller, eds., Cancun, MEXICO, June 5–9, 2010.

Li G, and Shi J, 2011. Agent-based Simulation on WGenCo Bidding in Day-Ahead Electricity Markets. In: *Proc. of the 2011 Industrial Engineering Research Conference*, May 21–25, 2011, Reno, Nevada, paper 713.

Li G, Shi J, and Qu X, 2011. Modeling methods for GenCo bidding strategy optimization in the liberalized electricity spot market–A state-of-the-art review. *Energy* 36(8):4686–00.

Li G, and Shi J, 2012. Agent-based modeling on double-sided auction of wind power. In: *Proc. of the 2012 Industrial and Systems Engineering Research Conference*, A. Johnson and J. Miller, eds., Orlando, Florida, USA, May 19–23, 2012.

Li G, Shi J, and Zhou J, 2011. Agent-based simulation on WGenCo bidding strategies in day-ahead electricity markets. In: *Proc. of the 2011 Industrial Engineering Research Conference*, A. Johnson and J. Miller, eds., Reno, Nevada, USA, May 21–25, 2011.

Li G, Erdem E, and Shi J, 2009. Deriving wind speed distributions based on Bayesian model averaging method. In: *Proc. of the 2009 Industrial Engineering Research Conference*. Miami, USA, May 30–June 3, 2009.

Li X, and Jiang C, 2011. Short-term operation model and risk management for wind power penetrated system in electricity market. *IEEE Transactions on Power Systems* 26(2):932–39.

Liao HL, and Wu QH, 2011. Optimal power flow in wind power integrated systems using function optimization by learning automata. In: *2011 IEEE Power and Energy Society General Meeting*, 1–8.

Li T, and Shahidehpour M, 2005. Strategic bidding of transmission - constrained GENCOs with incomplete information. *IEEE Trans. Power Syst.* 20(1):437–47.

Liu H, Jiang C, and Yan Z, 2008. A review on risk-constrained hydropower scheduling in deregulated power market. *Renewable and Sustainable Energy Reviews* 12(5):1465–75.

Liu H, Yuan B, Dai H, Lin J, Wu FF, et al, 2001. Framework design of a general-purpose power market simulator based on multi-agent technology, *IEEE Power Eng. Soc. Summer Meeting* 3, 1478–82.

Liu JS, 2001. Monte carlo strategies in scientific computing. New York: Springer-Verlag.

Liu Y, and Wu FF, 2006. Generator bidding in oligopolistic electricity markets using optimal control: fundamentals and application. *IEEE Transactions on Power Systems* 21(3):1050–61.

Liu Z, Zhang X, and Lieu J, 2010. Design of the incentive mechanism in electricity auction market based on the signaling game theory. *Energy* 35(4):1813–19.

Lora AT, Santos JMR, Exposito AG, Ramos JLM, and Santos JCR, 2007. Electricity market price forecasting based onweighted nearest neighbors techniques. *IEEE Trans. Power Syst.* 22(3):1294–301.

Ma L, Luan S, Jiang C, Liu H, and Zhang Y, 2009. A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Review* 13(4):915–20.

Ma X, Wen F, Ni Y, and Liu J, 2005. Towards the development of risk-constrained optimal bidding strategies for generation companies in electricity markets. *Electric Power Systems Research* 73(3):305–12.

Matevosyan J, and Soder L, 2006. Minimization of imbalance cost trading wind power on the short-term power market. *IEEE Transactions on Power Systems* 21(3):1396–404.

Mathew S, 2006. Wind energy – fundamentals, resource analysis and economics. Berlin: Springer-Verlag.

McLachlan GJ, and Krishnan T, 1997 The EM algorithm and extensions. New York: John Wiley & Sons.

Mohandes M, 2004. Support vector machines for wind speed prediction. *Renewable Energy* 29(6):939–47.

Monfared M, Rastegar H, and Kojabadi H,2009. A new strategy for wind speed forecasting using artificial intelligent methods. *Renewable Energy* 34(3):845–8.

Morales JM, Conejo AJ, and Perez-Ruiz J, 2010. Short-Term trading for a wind power producer. IEEE Transactions on Power Systems 25(1):554–64.

Morales JM, Conejo AJ. and Pérez-Ruiz J, 2011. Simulating the impact of wind production on locational marginal prices. IEEE Transactions on Power Systems 26(2):820–8.

More A, and Deo MC, 2003. Forecasting wind with neural networks. *Mar. struct.* 16(1):35–49.

Mideksa TK, and Kallbekken S, 2010. The impact of climate change on the electricity market: a review. *Energy Policy* 38(7):3579–85.

Minoia A, Ernst D, and Marija I, 2004. Market dynamics driven by the decision-making of both power producers and transmission owners. In: *Proc. Of IEEE Power Engineering Society General Meeting* 255–60.

Naghibi-Sistani M, Akbarzadehtootoonchi M, Javididashtebayaz M, Rajabimashhadi H, 2006. Application of q-learning with temperature variation for bidding strategies in market based power systems. *Energy Conversion and Management* 47(11–12):1529–38.

Newton MA, and Raftery AE, 1994. Approximate Bayesian Inference by the Weighted Likelihood Bootstrap (with discussion). *J. Roy. Statistical Society (Ser. B)* 56:3–48.

Ni E, Luh PB, and Rourke S, 2004. Optimal integrated generation bidding and scheduling with risk management under a deregulated power market. *IEEE Trans. Power Syst.* 19(1):600–9.

Niu H, 2005. Models for electricity market efficiency and bidding strategy analysis. *Ph.D. Dissertation,* Dept. Elec. Eng., Univ. Texas, Austin.

Nogales FJ, Contreras J, Conejo AJ, and Espinola R, 2002. Forecasting next-day electricity prices by time series models. *IEEE Trans. Power Syst.* 17(2):342–8.

Orponen P, 1994. Computational complexity of neural networks: a survey. *Nordic J. of Computing* 1(1):94–110.

Ostergaard P, 2008. Geographic aggregation and wind power output variance in Denmark. *Energy* 33(9):1453–60.

Ott AL, 2003. Experience with PJM market operation, system design, and implementation. *IEEE Transactions on Power Systems* 18(2):528–34.

Park JB, Kim BH, Kim JH, Jung MH, and Park JK, 2001. A continuous strategy game for power transactions analysis in competitive electricity markets. *IEEE Transactions on Power Systems* 16(4):847–55.

Petrov V, and Sheblé GB, 2000. Power auctions bid generation with adaptive agents using genetic programming. In: *Proc. of the 2000 North American Power Symposium*, Institute of Electrical and Electronic Engineers, Waterloo-Ontario, Canada.

Pinson P, Chevalier C, and Kariniotakis G, 2004. Optimizing benefits from wind power participation in electricity markets using advanced tools for wind power forecasting and uncertainty assessment. In: *Proc. Of EWEC 2004*, London.

Pinson P, Chevallier C, and Kariniotakis GN, 2007. Trading wind generation from short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems* 22(3):1148–56.

Press WH, Flannery BP, Teukolsky SA, and Vetterling WT, 1992. Numerical recipes in C: The art of scientific computing, 2nd Edition, Cambridge University Press.

Raftery AE, Gneiting T, Balabdaoui F, and Polakowski M, 2005. Using bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review* 133(5):1155–74.

Rahimiyan M, and Rajabimashhadi H, 2008. Supplier's optimal bidding strategy in electricity pay-as-bid auction: comparison of the Q-learning and a model-based approach. *Electric Power Systems Research* 78(1):165–75.

Rahimiyan M, and Rajabimashhadi H, 2007. Risk analysis of bidding strategies in an electricity pay as bid auction: A new theorem. *Energy Conversion and Management* 48(1):131–7.

Renewableenergyworld.com, 2009. BTM forecasts 340-GW of wind energy by 2013. URL: http://www.renewableenergyworld.com/rea/news/article/2009/03/btm-forecasts-340-gw-of-wind-by-2013?src=rss. Accessed on January 10, 2012.

Richter CJ, and Sheble GB, 1998. Genetic algorithm evolution of utility bidding strategies for the competitive marketplace. *IEEE Transactions on Power Systems* 13(1):256–61.

Richter CW, Sheble, GB, and Ashlock D, 1999. Comprehensive bidding strategies with genetic programming/finite state automata, *IEEE Trans. Power Syst.* 14(4):1207–12.

Rodriguez CP, and Anders GJ, 2004. Bidding strategy design for different types of electric power market participants. *IEEE Transactions on Power Systems* 19(2):964–71.

Rogersa AL, Rogersb JW, and Manwell JF, 2005. Comparison of the performance of four measure–correlate–predict algorithms. *Journal of Wind Engineering and Industrial Aerodynamics* 93:243–64.

Sahraei-Ardakani M, and Rahimi-Kian A, 2009. A dynamic replicator model of the players' bids in an oligopolistic electricity market. *Electric Power Systems Research* 79(5):781–8.

Saidur R, Islam MR, Rahim NA, and Solangi KH, 2010. A review on global wind energy policy. *Renewable and Sustainable Energy Reviews* 14(7):1744–62.

Saini LM, 2008. Peak load forecasting using Bayesian regularization, resilient and adaptive back-propagation learning based artificial neural networks, *Int. J. Electr. Power Energy Syst.* 78(7):1302–10.

Sánchez, I., 2005. Short-term prediction of wind energy production. *Int. J. Forecasting* 22(1):43–56.

Sanchez I, 2008. Adaptive combination of forecasts with application to wind energy. *Int. J. of Forecasting* 24(4):679–93.

Schweppe FC, 1988. Management of a spot price based energy marketplace. *Energy Policy* 16(4):359–68.

Sen S, Yu L, and Genc T, 2006. A stochastic programming approach to power portfolio optimization. *Operation Research* 54(1):55–72.

Shahidehpour M, Yamin H, and Li Z, 2002. Market operation in electric power systems. Wiley, New York.

Sfetsos A, 2000. A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renewable Energy* 21(1):23–35.

Sheble GB, 2001. Economically destabilizing electric power markets for profit. IEEE Power Engineering Society Winter Meeting 1:50–4.

Sioshansi R, and Oren S, 2007. How good are supply function equilibrium models: An empirical analysis of the ERCOT balancing market. *Journal of Regulatory Economics* 31(1):1–35.

Sloughter JM, Raftery AE, Gneiting T, and Fraley C, 2007. Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Monthly Weather Review* 135(9):3209–20.

Sloughter JM, Gneiting T, and Raftery AE, 2010. Probabilistic wind speed dorecasting using ensembles and Bayesian model averaging, Journal of the American Statistical Association, 105(489):25–35

Soleymani S, Ranjbar A, and Shirani A, 2007. New approach for strategic bidding of GenCos in energy and spinning reserve markets. *Energy Conversion and Management* 48(7):2044–52.

Soleymani S, Ranjbar A, and Shirani A, 2008. Strategic bidding of generating units in competitive electricity market with considering their reliability. *International Journal of Electrical Power & Energy Systems* 30(3):193–201.

Soman SS, Zareipour H, Malik O, and Mandal P, 2010. A review of wind power and wind speed forecasting methods with different time horizons. In: *North American Power Symposium (NAPS), 2010. IEEE* pp. 1–8.

Song KB, Ha SK, Park JW, Kweon DJ, and Kim KH, 2006. Hybrid load forecasting method with analysis of temperature sensitivities. *IEEE Trans. Power Syst.* 21(2):869–76.

Spall JC, 2003. Introduction to Stochastic Search and Optimization. Wiley. http://www.jhuapl.edu/ISSO.

Sueyoshi T, and Tadiparthi GR, 2005. A wholesale power trading simulator with learning capabilities. *IEEE Trans. Power Syst.* 20(3):1–11.

Sueyoshi T, and Tadiparthi GR, 2007. Agent-based approach to handle business complexity in U.S. wholesale power trading. *IEEE Transactions on Power Systems* 22(2):532–43.

Sueyoshi T, and Tadiparthi GR, 2008. An agent-based decision support system for wholesale electricity market. *Decision Support Systems* 44(2):425–46.

Sueyoshi T, 2009. An agent-based approach equipped with game theory: strategic collaboration among learning agents during a dynamic market change in the California electricity crisis. *Energy Economics* 32(5):1009–24.

Tamaschke R, Docwra G, and Stillman R, 2005. Measuring market power in electricity generation: A long-term perspective using a programming model. *Energy Economics* 27(2):317–35.

Tesfatsion L, 2002. Agent-based computational economics: growing economies from the bottom up. *Artificial Life* 8(1):55–82.

Tesfatsion L, 2006. Agent-based computational economics: a constructive approach to economic theory. *Handbook of computational economics*, North-Holland 2:831–80.

The Federal Energy Regulatory Commission (FERC). Order no. 890. Fact Sheet: FERC Docket Nos RM5-25-000 and RM05-17-000. URL:http://www.ferc.gov/industries/electric/indus-act/oatt-reform/order-890/fact-sheet.pdf. Accessed on January 18, 2012.

Trudnowski DJ, McReynolds WL, and Johnson JM, 2001. Real-time very short-term load prediction for power-system automatic generation control *IEEE Trans. Power Syst.*

9(2):254–60.

U.S. Department of Energy, 2008. 20% Wind Energy by 2030: Increasing wind energy's contribution to U.S. electricity supply. URL:http://www.20percentwind.org/20percent_wind_energy_report_05-11-08_wk.pdf. Accessed on January 18, 2012.

Usaola J, and Angarita J, 2006. Benefits of short term wind power prediction programs for the integration of wind energy in electricity markets. *EWEC 2006*.

Usaola J, and Angarita J, 2007. Bidding wind energy under uncertainty. In: *Proc. of IEEE 2007 International Conference on Clean Electrical Power* 754–9.

Usaola J, Ravelo O, Gonzalez G, Soto F, Davila M, and Diaz-Guerra B, 2004. Benefits for wind energy in electricity markets from using short term wind power prediction tools: a simulation study. *Wind Engineering* 28(1):119–27.

Vahidinasab V, and Jadid S, 2009. Multiobjective environmental/techno-economic approach for strategic bidding in energy markets. *Applied Energy* 86(40):496–504.

Vahidinasab V, and Jadid S, 2010. Normal boundary intersection method for suppliers' strategic bidding in electricity markets: An environmental/economic approach. *Energy Conversion and Management* 51(6):1111–19.

Valenzuela J, and Mazumdar M, 2003. Commitment of electric power generators under stochastic market prices. *Oper. Res.* 51(6):880–93.

Veit DJ, Weidlich A, and Krafft JA, 2009. An agent-based analysis of the German electricity market with transmission capacity constraints. *Energy Policy* 37(10):4132–44.

Von der Fehr NHM, and Harbord D, 2002. Competition in electricity spot markets: economic theory and international experience. *Industrial Organization 0203006, EconWPA*. http://ideas.repec.org/p/wpa/wuwpio/0203006.html.

Von der Fehr NHM, and Harbord D, 1993. Spot market competition in the UK electricity industry. *The Economic Journal* 103(418):531–46.

Vytelingum P, and Cliff D, 2008. Jennings N. Strategic bidding in continuous double auctions. *Artificial Intelligence* 172(14):1700–29.

Wallace SW, and Fleten SE, 2003. Stochastic programming models in energy. *Handbooks in Operations Research and Management Science,* Elsevier 10:637–77.

Viallefont V, Raftery AE, and Richardson S, 2001. Variable selection and Bayesian model averaging in epidemiological case-control studies. *Statistics in Medicine* 20(21):3215–30.

Walter I, and Gomide F, 2003. Evolving fuzzy bidding strategies in competitive electricity markets. *IEEE international conference on systems* 4:5–8.

Walter I, and Gomide F, 2008. Coevolutionary fuzzy multiagent bidding strategies in competitive electricity markets. In: *Proc. of the 3rd International Workshop on Genetic and Evolving Systems (GEFS 2008)* 53–8.

Wanas N, Auda G, Kamel MS, and Karray F. 2002.On the optimal number of hidden nodes in a neural network. In: *Proc. 1998 IEEE Canadian Conference on Electrical and Computer Engineering* (2):918–21

Wang X, Sideratos G, Hatziargyriou N, and Tsoukalas LH, 2004. Wind speed forecasting for power system operational planning. In: *Proc. of 2004 International Conference on Probabilistic Methods Applied to Power Systems*. pp. 470–74.

Wasserman L, 2000. Bayesian model selection and model averaging. *J Math Psychol* 44(1):92–107.

Weidlich A, and Veit D, 2008. A critical survey of agent-based wholesale electricity market models. *Energy Economics* 30(4):1728–59.

Wen F, and David AK, 2001a. Optimal bidding strategies and modeling of imperfect information among competitive generators. *IEEE Transactions on Power Systems* 16(1):15–21.

Wen F, and David AK, 2001b. Strategic bidding for electricity supply in a day-ahead energy market. *Electric Power Systems Research* 59(3):197–206.

Wen F, and David AK, 2001c. A genetic algorithm based method for bidding strategy coordination in energy and spinning reserve markets. *Artificial Intelligence in Engineering* 1(1):71–9.

Widrow B, and Lehr MA, 2002. 30 years of adaptive neural networks: perceptron, Madaline, and backpropagation. *Proc. the IEEE* 78(9):1415–42.

Willems B, 2002. Modeling Cournot competition in an electricity market with transmission constraints. *The Energy Journal* 23(3):95–126.

Wind energy the facts, 2011. Wind energy penetration and integration. http://www.wind-energy-the-facts.org/en/part-2-grid-integration/chapter-1-setting-the-scene/. Accessed on August 31, 2011.

Wu YK, and Hong JS, 2007. A literature review of wind forecasting technology in the world. In: Power Tech, 2007 IEEE Lausanne. IEEE, pp. 504–9.

Xia X, and Elaiw AM, 2010. Optimal dynamic economic dispatch of generation: a review. *Electric Power Systems Research* 80(8):975–86.

Xiong G, Hashiyama T, and Okuma S, 2002. An evolutionary computation for supplier bidding strategy in electricity auction market. *IEEE Power Eng. Soc. Winter Meeting* 1:83–8.

Xiong G, Okuma S, and Fujita H, 2004. Multi-agent based experiments on uniform price and pay-as-bid electricity auction markets. *IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies* 1:72–6.

Yamin HY, 2004. Review on methods of generation scheduling in electric power systems. *Electric Power Systems Research* 69(2–3):227–48.

Yamin H, and Shahidehpour SM, 2003. Unit commitment using a hybrid model between Lagrangian relaxation and genetic algorithm in competitive electricity markets. *Electric Power Systems Research* 68(2):83–92.

Yamin H, and Shahidehpour SM, 2004. Self-scheduling and energy bidding in competitive electricity markets. *Electric Power Systems Research* 71(3):203–9.

Younes Z, and Ilic M, 1999. Generation strategies for gaming transmission constraints: will the deregulated electric power market be an oligopoly? *Decision Support Systems* 24(3–4):207–22.

Yuan Z, Liu D, and Jiang C, 2007. Analysis of equilibrium about bidding strategy of suppliers with future contracts. *Energy Conversion and Management* 48(3):1016–20.

Yucekaya A, Valenzuela J, and Dozier G, 2009. Strategic bidding in electricity markets using particle swarm optimization. *Electric Power Systems Research* 79(2):335-45.

Zareipour H, Canizares CA, Bhattacharya K, and Thomson J, 2006. Application of publicdomain market information to forecast Ontario's wholesale electricity prices. *IEEE Trans. Power Syst.* 21(4):1707–17.

Zhang D, Wang Y, and Luh PB, 2000. Optimization based bidding strategies in the deregulated market. *IEEE Transactions on Power Systems* 15(3):981–6.

Zhang L, and Luh PB, 2005. Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method, *IEEE Trans. Power Syst.* 20(1):59–66.

Zhang N, 2009. Generators' bidding behavior in the NYISO day-ahead wholesale electricity market. *Energy Economics* 31(6):897–913.

Zhao B, Guo CX, and Cao YJ, 2005. A multiagent-based particle swarm optimization approach for optimal reactive power dispatch. *IEEE Trans Power Syst Delivery* 20(2):1070–78.

Zhou J, Erdem E, Li G, and Shi J, 2010. Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites. *Energy Conversion and Management* 51(7):1449–58.

Zimmerman RD., Murillo-Sánchez CE., and Thomas RJ, 2011. MATPOWER steady-state operations, planning and analysis tools for power systems research and education, *Power Systems, IEEE Transactions on* 26(1):12–19.