**University of Wisconsin Milwaukee**
**UWM Digital Commons**

Theses and Dissertations

August 2014

# Short-Term Forecasting of Wind Speed and Wind Power Based on BP and Adaboost_BP

Sibei Mi
*University of Wisconsin-Milwaukee*

Follow this and additional works at: https://dc.uwm.edu/etd

Part of the Electrical and Electronics Commons

# SHORT-TERM FORECASTING OF WIND SPEED AND

# WIND POWER BASED ON BP AND ADABOOST_BP

by

Sibei Mi

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Engineering

at

The University of Wisconsin-Milwaukee

August 2014

# ABSTRACT

## SHORT-TERM FORECASTING OF WIND SPEED AND WIND POWER BASED ON BP AND ADABOOST_BP

by
Sibei Mi

The University of Wisconsin-Milwaukee, 2014
Under the Supervision of Professor David C. Yu

Due to wind is intermittent and less dispatchable, wind power fluctuates as the wind fluctuating and is uncontrollable. Therefore, when wind power accounts for a higher proportion of total electricity generation of the system, power generation plan needs to arrange according to the variation of wind power output. The way to solve the problem is forecasting the wind power.

In this paper, we focus on the wind speed and wind power forecasting in the time scales of 10min, 1h and 3h in the future. BP Neural Network and Adaboost_BP Neural Network are selected as the forecasting model for wind speed forecasting. And for wind power forecasting model, we use BP Neural Network as the method.

The wind farm which the data used in this paper comes from has 31 wind turbines in the same type. As the geographical distribution of this wind farm is unknown, we pick one No.17 wind turbine as the optimal one by analyzing the data of 31 wind turbines and making the curve fitting of wind speed and wind power. Then we analyze the influencing factor of wind power, and find out wind speed as the most influential factor. For the wind speed, we deal with the raw data which selected from SCADA (Supervisory Control and Data Acquisition) system before using it.

For the wind speed forecasting model, we find the optimal number of the training data for each training sets for the BP Neural Network in each time scale. Then we make a contrast of the accuracy of the single-step forecasting accuracy between BP Neural Network model and Adaboost_BP model in 10min, 1h and 3h at their respectively optimal number of the training data. And there is a comparison between the accuracy of the single-step and iterative multi-step wind speed forecasting model in 1h and 3h time scales at the number of the training data of 10 for the two models.

For the wind power forecasting model, we use the forecasting wind speed and

its corresponding wind power to build the input matrix for the network training.

And we find that not only the number of the training data for each training sets

but also the range of wind power affects the forecasting accuracy. Then make a

contrast of the wind power forecasting result with forecasting wind speed and

actual wind speed as the input.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGEMENTS

This thesis wouldn't have been possible without the help and support of several people.

I would like to express my deep and sincere gratitude to my supervisor Prof. David C. Yu. With his invaluable support, guidance and patience, I make this thesis possible. He built up the communicational bridge between the universities in China and the University of Wisconsin-Milwaukee, and this gave me the opportunity to come to United States for further study.

I am grateful to Xiaoxuan Lou and Zhenze Long. Their insightful comments helped me to improve this thesis.

Last but not least, I would like to thank my family, my mother Yanxi Zhou, my father Ge Mi and my aunt Xiaoling Mi, for their unconditional love and support they have shown me throughout my life.

# I. INTRODUCTION

## 1.1. Introduction

Electricity can be generated by a variety of ways. Apart from solar photovoltaic power generation, other forms of electricity generation are in the same way that primary energy pushes prime mover, and then the prime mover drives generators to generate electricity. Wind power has many characteristics which other fossil energy does not have, such as clean, intermittent and randomness. This is because the wind is a natural phenomenon. Wind energy converts into mechanical energy in the way that wind blow through fans to drive rotor rotation. And then the energy converts into electricity without generating pollution and radiation which will be generated in the electricity conversion process of conventional energy. The reason why the demand for wind power around the world grows involves many aspects, including the shortage of energy, change in climate, the progress of economy and technology, etc.

In 2006, International Energy Agency pointed out the world's energy demand would be 60% higher than in that year, and the supply of fossil fuels would be reduced [1]. The energy in some countries would be relied on imports, and this would affect the stability of geopolitics in these countries. On the contrary, as a new energy without energy costs, wind energy exists in various countries and regions all over the world. Therefore the use of wind power as a new energy will become an important strategy for the world to solve the problem of energy crisis.

Besides, another impetus for the rapid development of wind power is the increasingly serious problem of global climate change which is known as the most seriously worldwide recognized environmental threat. In the 1997 Kyoto protocol, developed countries reached an agreement to decrease the totality of the emissions of six kinds of greenhouse gas at least 5% from 1990 levels in 2008-2010 (8% for Europe, 7% for United States and 6% for Canada). In addition, the development of nuclear energy is constrained by concerning about the radioactive waste. Especially the leakage of nuclear materials of Japan Fukushima nuclear power plant in March 2011 made countries around the world to reconsider their nuclear policies. Nuclear power development in China has therefore been put off, while Germany will shut down all nuclear power plants by 2022. Thus, the declination of the status of nuclear power in clean energy makes the development of other clean energy like wind power more urgent.

In the 1880's, with the development of the global wind power market, sharp decline in the price of wind power equipment made contribution to the development of wind power industry. In the area which is rich in wind resources, the cost of electricity generation of wind power is not expensive which can be competitive with that of coal and natural gas. And with the increase in the price of fossil fuels in recent years, the competitiveness of wind power gets greatly improved. Besides, considering the extra charges and the impact on people's health as parts of generating cost which caused by the pollution from fossil fuels and nuclear radiation, the price of wind power is more competitive [2].

In recent years, the rapid development of wind power industry is mainly driven by national policy, environment and economy. And the development of wind turbine technology plays an essential role. Since the first commercial wind turbine launched in the 1880's, the capacity, efficiency and the visual design of wind turbines have been greatly improved. The modern wind turbines are modularity and easy to install. The capacity of wind farm Increased from a few MW to several hundred MW.

According to the report from global wind energy council, the growth rate of the wind power are faster than that of other renewable energy sources in the past ten years. Starting in 2000, the annual average growth rate of installed wind energy capacity reaches 28%. And it is estimated that the global installed wind energy capacity will be 233-486GW by 2015 and reach 352-1081GW by 2020.

Compared with traditional fossil fuel power plants, the electricity production in wind power plants is quite different. The most different one is the production of wind power depends on the existence of the wind which is an uncontrollable natural phenomenon [3]. However, the generating capacity of both traditional fossil fuel power plants and nuclear power plants can be controlled according to the load demand. But, due to wind is intermittent and less dispatchable, wind power also fluctuates as the wind fluctuating and is uncontrollable. When the power system operates, the output of electricity and the load must be equal at any time. Therefore, when wind power accounts for a higher proportion of total electricity generation of the system, power generation plan needs to arrange according to the variation of wind power output.

Wind power forecasting is an important way to solve this challenge. Under the conventional power generation way, power system operation needs to forecast the load and develops electricity generation schedule for traditional power plants. However, after adding wind power to the system, the wind speed and power forecasting is needed in order to make wind power in the electricity system scheduling plan. Unlike the traditional electricity generation schedule, the accuracy of the wind speed and power forecasting is far less than that of load forecasting for its uncontrollability. Hence, it is important to reserve more spinning reserve or add another energy storage device while making the electricity system scheduling plan.

## 1.2. Wind Power Forecasting Methods

*A.    Classification of Wind Power Forecasting Methods*

Wind speed and wind power forecasting have been developed over 30 years. A number of wind speed and wind power forecasting models for commercial and research use are made in recent years. In countries such as Denmark, Germany and Spain, Wind power occupies a large proportion in the whole power system, and wind power forecasting system has become an important part of power and control system.

Wind power forecasting model is divided into statistical model, physical model and the combination of the two models according to the input data [4]. In general, different forecasting models require different inputs. Considering the most important influence factor of wind farm output power is the local wind speed, wind speed is the

essential input for all forecasting model. The data types of inputs which can be used for predictive models include Numerical Whether Predictions, online data of wind farm, physical environment data (including topography, surface roughness and the distribution of wind farm, etc.) and user information [5]. For the statistical model, it only need the online measurement data of wind farm, such as the wind farm output power, the power output of a single wind turbine and meteorological data (wind speed and direction, etc.). And it also can use the meteorological information provided by the Numerical Weather Prediction (NWP) [4]. As for the physical model, it mainly depends on the numerical weather prediction (NWP), and use some physical conditions of geography, such as surface features (topography and roughness and obstacles), and the technical information of the wind turbine (for example, the center height, power curve and penetrating power) [4,6,7]. Combination of the two forecasting models can have the high accuracy in a relatively short time which comes from the statistical model as well as the longer time scale forecasting of the physical model.

Wind power forecasting model also can be classified according to the forecast time length, and the forecast time length depends on the application purpose of forecasting model and the technology it uses. For instance, the forecast result which the forecast time length is a few seconds to a few minutes can be used for the dynamic control of the wind turbine system [5,8], and the forecast result which the forecast time length is one hour or a few hours can be used for the safe operation and the economic dispatch of the power system [4, 9]. Classifying by the forecasting technique is another way for

the classification of wind power forecasting model. The forecasting model can be divided in physical model, time sequence model, artificial intelligence model and spatial correlation model, etc.

*B.    Physical Models of Wind Power Forecasting Methods*

Currently, the forecast time length of physical model can be up to 48 hours. These models mainly depend on the numerical weather prediction (NWP), and need to take some other physical factors into consideration, such as the    local surface roughness, the influence of terrain and obstacles, the layout of wind farm and the wind turbine power curve [4,6,7]. The physical model has higher forecast accuracy for a long forecast time length. However, it is very complex and expensive to make forecasting by using physical models. Besides, for the short-term forecasting (a few hours forecasting), the forecasting results are often not satisfactory as the delay of data updating [7]. Generally physical models are used as the first step of the wind speed forecasting, and its forecasting results will be the input of other statistical models.

Nowadays, there are multiple physical models which can be used for wind power forecasting. One of them is high-resolution limited area model (HIRLAM) which established in 1999 by Landberg [10]. The model uses the Numerical weather forecasting data which is provided by HIRLAM as the input. The research of numerical weather forecasting data is in a long-term scale forecasting in power plants. The collecting data of wind farm must be related to a rough model which considers the influences of local territory, and use WASP mode for local wind speed forecasting

by taking the local impact into account. In order to solve the effect of the wake flow of wind turbines, PARK scheme is adopted. The fan chart is presented in PARK which declares the reduce of wind farm output power due to the wake flow of wind turbines. Then obtain the wind farm output power which calculates from the model output data. The model is successfully applied to multiple wind farms in Denmark. The model runs twice a day and provides the forecasting for 36 hours in advance. Compared with continuous model, the forecasting results of this model is better than that of the continuous model when forecasts for more than 4 hours.

In the reference [11], Focken studied the influence of the decrease of the wind power forecasting error due to the spatial smoothness. The model used the NWP data provided by Deutscher Wetterdienst, including the forecasting data for 6, 12, 18, 24, 36 and 48 hours in advance. The forecasting model provides the wind farm output power forecasting for 48 hours in advance. In this model, the wind speed of the center height of wind turbines is calculated by the formula of wind profile. In addition, the model also considers the differential thermal of the boundary layer, the surface roughness and terrain of wind farms and the wake effect of wind turbines. Using the statistical model as the forecasting method for wind farm power output, the final output of the model is an output power of a wind farm. Compared the average error of wind power forecasting in 30 wind farms in Germany with that of a single wind farm, the forecasting precision of the spatial correlation model is higher than that of a single address model. The reason is the influence of spatial smoothing effect, and the improvement of the forecasting precision depends on the size of the region and the

number of wind farms in the region.

*C.*    *Time Sequence Models of Wind Power Forecasting Methods*

At present, several time sequence models have been put forward which can be used for wind speed and wind power forecasting, such as AR model, MA model, ARMA model, ARIMA model, time sequence model based on neural network and gray model etc. The characteristics of time sequence models are simple and easy to modeling, and the input data of the model only need to be wind speed and wind power in the past. However, the accuracy of time sequence forecasting model as will have fallen sharply as the increase of the length of the forecasting time. Therefore, these time sequence models usually used to forecast short-term wind speed and wind power.

Sfetsos (2002) used the auto-regressive integrated moving average (ARIMA) model and the feed forward neural network as the forecasting model, and put forward the wind speed forecasting model an hour in advance based on the two different time sequence models in the reference [12]. One forecasting model used the average wind speed in an hour to make one step forecasting in advance. The other one used the average wind speed in 10 min to make six steps forecasting to get wind speed forecasting data an hour in advance. Comparing the two forecasting methods with the persistence approach, the prediction accuracy of Artificial Neural Network (ANN) model is higher than that of the persistence approach. However, the prediction accuracy of ARIMA model is higher than that of ANN model.

Huang and Chalabi (1998) presented an auto-regression (AR) model which could use to forecast wind speed an hour or several hours in advance in the reference [6]. This

algorithm considered the unstable characteristics of wind speed and used Kalman filter to estimate the time variation parameter of AR model.

Erasmo and Wilfrido (2007) used ARIMA model and ANN model to make a monthly time-step size wind speed forecasting with the wind speed data for past 7 years in advance in the reference [13]. And comparing this method with the persistence approach, ARIMA model has a higher prediction accuracy for wind speed.

EI-Fouly (2006) presented the gray model (GM(1,1)) to forecast wind speed for an hour in advance in the reference [14]. The wind speed forecasting data which is provided by GM(1,1) model is used to forecast wind power as the input of the 600kW wind turbine power formula. Comparing the forecasting result of GM(1,1) model with that of the persistence approach, the forecasting accuracy of wind speed was improved about 11.2% than the persistence approach method and the forecasting accuracy of wind power is increased by about 12.2%.

EI-Fouly (2006) put forward a time sequence model for wind speed forecasting which could carry out the wind speed 24 hours in advance [5]. In this model, the data came from three times of three years data set of a meteorological station. And compared this method with the persistence approach, this model has a better forecasting result while the forecasting time length is 3, 6, 12 and 24 hours.


*D.    Artificial Intelligence Models and Other Models of Wind Power Forecasting*

In recent years, the artificial intelligence method，such as artificial neural network, fuzzy logic and hybrid model, has been widely used. As the artificial intelligence

method can simulate nonlinear problems very well, many researchers found that the wind speed and wind power forecasting can get good result by using these methods.

Shuang (2007) used Tabu Search Algorithm optimized the link weight of BP neural network, and forecasted the wind speed 1 hour in advance by using this model [15]. The number of input neurons was decided by correlation analysis. Thus, ANN model could learn better in this way. The result showed that ANN model which used the new training method has a better accuracy in wind speed forecasting than traditional BP neural network model.

Leigh and Ran (2008) decomposed the wind speed time series which is highly nonlinear into several relatively stable time series. Every time series is decomposed from the original wind speed sequence according to the different resolution. And ARMA model forecasting models are established for every time series to forecast the next value of this series, then use the combination method of the forecasting data to synthetize the forecasting wind speed 1 hour in advance. The actual mean hourly wind speed used in this article to establish the forecasting model came from meteorological station. Compared this method with the continuous approach and the original ARMA method, this model improved the forecasting accuracy.

Senjyu (2006) presented a feedback neural network (RNN) to forecast the wind speed 3 hours to 3days in advance by time series method [16]. Compared this method with feed forward neural network (FNN), RNN method can reduce the forecasting error and improve the forecasting effect of wind speed and wind power.

# II.  DATA PROCESSING

At present, the common method used for wind power forecasting is using neural network to establish the corresponding forecasting model of the forecast data and wind power, and the forecasting data is the meteorological forecasting data of the wind tower which is near the wind farm. In order to use the neural network to establish the model, the most essential step is selecting the appropriate network input variables by analyzing the influence factors of wind turbine power output. Then the data used in the forecast is the processed one of the raw data.

## 2.1.  Wind Turbine Selection

The wind farm which the data used in this paper comes from has 31 wind turbines in the same type. While the geographical distribution of this wind farm is unknown, one wind turbine is picked out from the wind farm in this paper. By analyzing the data of 31 wind turbines and making the curve fitting of wind speeds and wind powers, No.17 wind turbine has been found out as the optimal one. The steps are:

*A.  Data Normalization*

Due to the data ranges of variables are different and the values may vary widely, large numerical variables have bigger influence on the results when establish mathematical model to solve problems. If one small numerical variable had a great influence on the result, it might play a small role in building the mathematical model. In order to avoid such a situation, the data which is used in modeling need to be normalized first.

Normalization can transform the value of variables into a specified range, usually [-1, 1] or [0, 1].

Considering the range of wind speed is 0-30m/s and the range of wind power is 0-2000kW, the maximums of wind speed and wind power set as 50m/s and 5000kW, and the minimum are 0m/s and 0kW.

For wind speed,

$$v*(t) = \frac{v(t) - 0}{50 - 0}$$

where $v*(t)$ is normalization value of wind speed and $v(t)$ refers to the original value of wind speed.

For wind power,

$$P*(t) = \frac{P(t) - 0}{5000 - 0}$$

where $P*(t)$ is normalization value of wind power and $P(t)$ refers to the original value of wind power.


*B.  Eliminate the Zero Point of Wind Power*

By analyzing the wind speed and wind power which will be used to do the carving fitting, we found that when wind speed was very low and changed suddenly to a high speed, and then quickly went back to low speed, the wind power of the high speed point was 0W likes that of very low wind speed, such as

Table II-1. These points affected the accuracy of the carving fitting Figure II-1, so we eliminate these points before the carving fitting.

Table II-1 Low Wind Speed and Its Corresponding Wind Power

| speed | power |
|---|---|
| 1.383682 | 0 |
| 0.780178 | 0 |
| 2.47893 | 0 |
| 0.377842 | 0 |
| 4.356498 | 0 |
| 5.898786 | 0 |
| 1.965843 | 0 |
| 1.004706 | 0 |
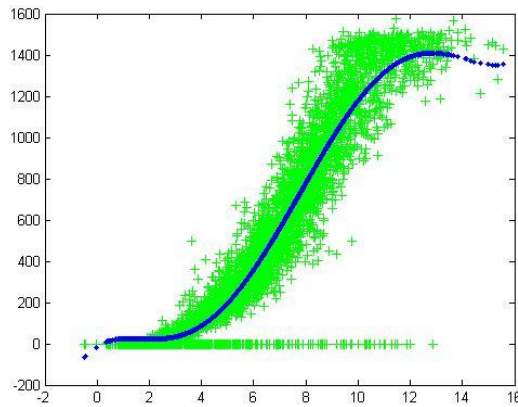


Figure II-1 5th-order Carving Fitting without Eliminating the Zero Points

*C.  Order of Carving Fitting*

In the progress of the carving fitting of wind speed and wind power, the order of carving fitting should be taken into consideration as it greatly affects the results of convergence. To decide the order of the carving fitting, this paper makes contrast of 7 different orders to find out the most suitable one for its convenience and accuracy. In Figure II-3 the green spot stands for the actual value and the blue one is the fitted data.
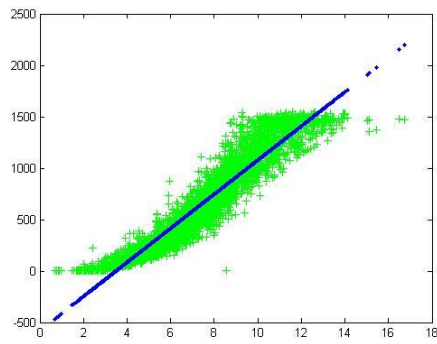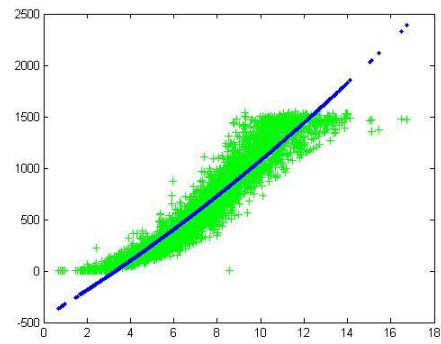
Figure 1(a)                    Figure 1(b)
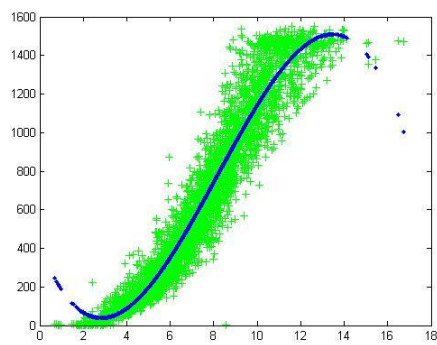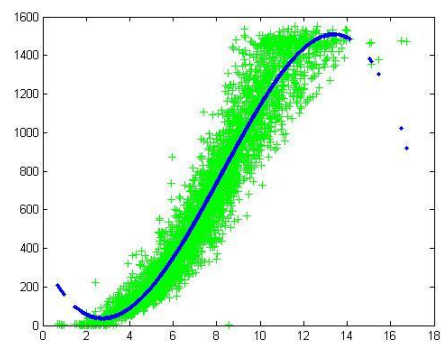
Figure 1(c)                    Figure 1(d)
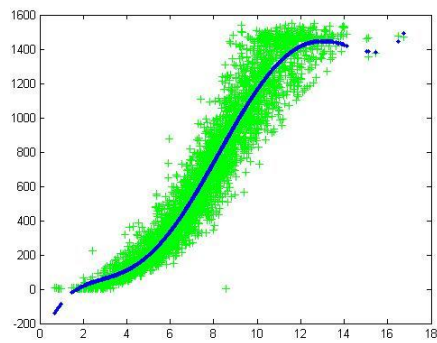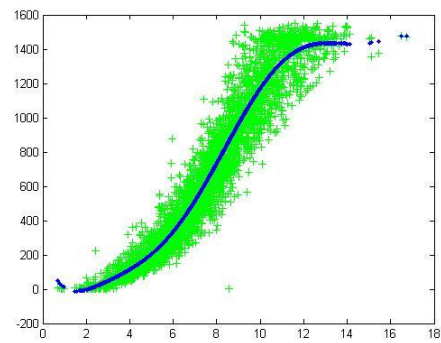
Figure 1(e)                    Figure 1(f)

Figure II-2. Carving Fitting

(a) 1st-order fitting. (b). 2nd-order fitting. (c) 3rd-order fitting. (d). 4th-order fitting.

(e). 5th-order fitting. (f). 7th-order fitting. (g). 10th-order fitting
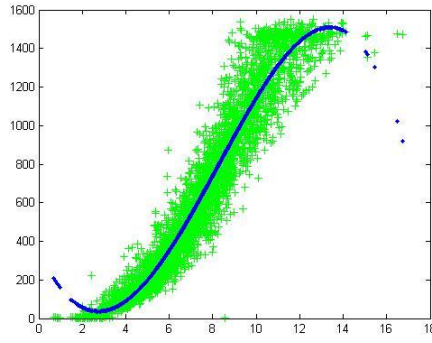
Figure 2(g)

Figure II-3. Carving Fitting

(a) 1st-order fitting. (b). 2nd-order fitting. (c) 3rd-order fitting. (d). 4th-order fitting.

(e). 5th-order fitting. (f). 7th-order fitting. (g). 10th-order fitting

In this paper, we select 7 sets of wind speed and wind power data from different wind turbines and different months randomly. And the results are almost unanimous. Thus these figures are the carving fitting results of one of wind turbines in the same month. From Figure II-3, we can tell that the fifth-order fitting and the seventh-order fitting are very good. And considering the absolute value of the error and the simplicity of carving fitting, we choose the fifth-order fitting as the order of carving fitting in this paper.

*D.  Select the Wind Turbine*

The data used in this paper starts from March 2011 to January 2012. Use fifth-order for carving fitting to find out one wind turbine with the most optimal data. Firstly carve fitting the wind speed and wind power of 31 wind turbines in March 2011. And calculate the RMS (Root Mean Square) of the actual value and fitted value for each

turbine. Then make a list for the wind turbines as the ascending order of RMS, and select the first five wind turbines to carve fitting their data in April 2011. And by comparing the RMS of the five wind turbines in the Table II-2, we choose No.17 as the objective wind turbine. Figure II-4 is the crave fitting of No.17 in March 2011, and the green spot stands for the actual value and the blue one is the fitted data.

Table II-2 RMS of Error of Wind Turbine

| Year-Month-No. | RMS | Year-Month-No. | RMS |
|---|---|---|---|
| 2011-3-25# | 4.32E+03 | 2011-4-25# | 4.32E+03 |
| 2011-3-17# | 4.98E+03 | 2011-4-17# | 4.93E+03 |
| 2011-3-23# | 4.98E+03 | 2011-4-20# | 5.03E+03 |
| 2011-3-18# | 5.21E+03 | 2011-4-23# | 5.09E+03 |
| 2011-3-20# | 5.22E+03 | 2011-4-18# | 5.44E+03 |
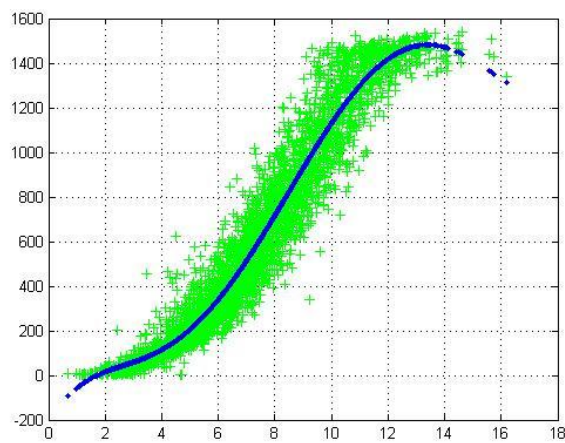


Figure II-4. Crave Fitting of No.17 in March 2011

## 2.2. Influencing Factors of Wind Power

In order to analyze the influencing factors of wind power, we start from analyzing the wind turbine capture capacity of wind power. The equation of the wind turbine capture capacity shows below:

$$P = C_P A \rho v^3 / 2$$

where, $P$ stands for actual available power(W), $C_P$ refers to the wind power coefficient, $A$ is the perpendicular area that wind passing through ($m^2$), $\rho$ is the density of air ($kg/m^3$) and $v$ stands for the wind speed ($m/s$).

Form the equation, wind power varies with the cube of its speed, so wind speed is the most important influencing factor of the wind turbine output. Therefore, wind speed is one of the input variables of neural network.

The wind power coefficient is one of the important parameters stand for the wind turbine efficiency. Every wind turbine has its fixed wind power coefficient, and it indicates the proportion of useful energy obtained by the wind turbine from the wind. According to Bates theory, the maximum of wind power coefficient is 0.593, and the wind power coefficient varies from the type of wind turbines.

For the area that wind passing through perpendicular to the wind,

$$A = \pi D^2 / 4$$

where $D$ is the diameter of rotor.

Thus, the air density is the rest of the factors which influence the wind turbine power output, and it is changing over time.

### 2.2.1 Air Density

Air density is determined by the pressure, temperature and humidity of the air. The equation of the air density shows below:

$$\rho = 3.48 \frac{P}{T}(1 - 0.378 \frac{\varphi P_b}{P})$$

where $P$ stands for standard atmospheric pressure ($kP_a$), $T$ is thermodynamic temperature ($K$), $P_b$ refers to the saturated vapor pressure ($kP_a$) and $\varphi$ is relative air humidity.

At present, there are different views on whether use the pressure, temperature and humidity as the neural network input in the wind power forecasting research which is based on the meteorological forecast data. The reference [17,18] point out that the correlation of these three variables (pressure, temperature and humidity) and wind power is relatively small. Using them as the neural network inputs has little contribution to improve the effect of the forecasting model, and adding the three variables as the inputs into the neural network will slow down the training speed. Therefore, we do not use the three variables as the inputs in this paper.

### 2.2.2 Wind Direction

Apart from the influence factor talked about before, wind direction is another important wind power influence factor. As the distribution of wind turbines is scattered, wake flow will be forming after wind passing through the fan blades, and the wind speed will slow down. When the wind speed is low, the wind farm efficiency will reduce in some wind direction as the influence of the wake and roughness. With

the increase of wind speed, the effect of wake flow on the output power will be reduced, and the efficiency coefficient of wind farm will increase. When the wind speed exceeds a certain value, the wind farm efficiency coefficient is 100% in any wind direction. Normally, because the wind changes randomly, the wind farm efficiency coefficient cannot stay 100% all the time. However, because of the relationship between wind power and wind direction is difficult to carve fitting, it is hard to determine the quality of the wind direction. In this paper, the wind direction is not considered as one of the inputs of the neural network. And in this way, the training speed can be faster and the difficulty of network training caused by the wind direction normalization can be prevented. In the further study, the wind direction can be taken as the inputs for more accurate prediction.

### 2.2.3    Wind Speed

Wind speed has different varying patterns in different length of time, different height and different area. And the change of wind speed is also related to temperature. Meanwhile, the distribution of wind speed is not the same as well. General Weibull distribution is used to describe the statistical distribution of wind speed.

The time scale has an important significance for the study of wind power, and wind power researches under different time scales have different effects on the power system decision-making. Long term wind power forecasting involves in national policies, regulations and energy resources long-term planning, so its uncertainty is quite strong. Use grey modeling to deal with the impact of its uncertainty and predict

the installed capacity and the output of the wind power system [19]. Most wind power forecasting researches are focused on the medium and short term forecasting [20]. The accurate and effective wind power forecasting of wind farm is not only the important reference which helps the power system dispatching operation personnel make the most effective decisions, but also for the reference for the amount of wind power integration under the condition of electricity market. In this paper, we focus on the wind speed and wind power forecasting under time scales of 10min, 1h and 3h.

*A. Data Processing*

The data we use in this paper is selected from SCADA (Supervisory Control and Data Acquisition) system. And time sampling interval of the data is 10min. Generally, the variation of wind speed doesn't exceed 6m/s within an hour. And by exploring the data, we find that some wind speed is negative when the wind speed is quite slow. However, it is not correct that the wind speed is negative. Therefore, we should deal with the raw data before considering the sampling frequency is not high.

Firstly, we deal with the negative wind speed. By analyzing the data trending, the negative wind speed is assigned the value zero.

Then, we find out the wind speed which does not meet the condition that the variation of wind speed doesn't exceed 6m/s within an hour, and calculate the average value of the successively four wind speed. And the bad data $v_i$ is replaced by the average value of $v_{i-2}$, $v_{i-1}$, $v_{i+1}$ and $v_{i+2}$.
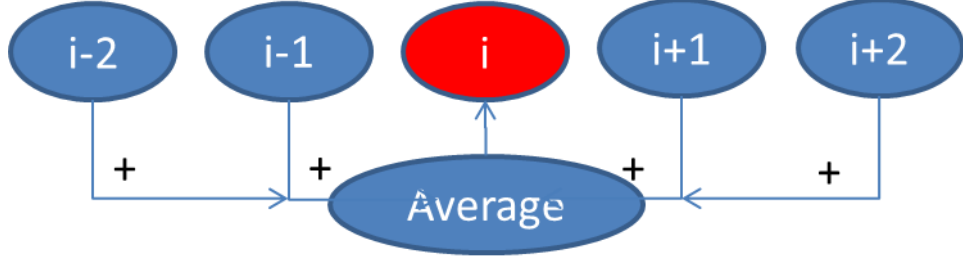
Figure II-5. Bad Data Replacement

*B. Training Matrix*

The matrix for neural network training is composed of the wind speed in the chronological order. The number of columns of the training matrix is the number of training data set, and the number of rows of the training matrix is the number of training data in each training set.

For the training matrix, there is a relationship between the adjacent columns. The first data of a column is the second data of previous column, and the rest data arranges in the chronological order in each column.

For example, data $a_{i,2}$ is the second data of column i, then, according to the rule of the matrix generating, $a_{i+1,1}$ the first data of column i+1 is equal to $a_{i,2}$. So $a_{i+1,2}$ the second data of column i+1 is equal to $a_{i,3}$ which is the third data of column i.

Table II-3 Example of Training Matrix

| column i | column i+1 | column i+2 |
|---|---|---|
| $a_{i,1}$ | $a_{i+1,1}=a_{i,2}$ | $a_{i+2,1}=a_{i+1,2}$ |
| $a_{i,2}$ | $a_{i+1,2}=a_{i,3}$ | ... |
| $a_{i,3}$ | ... | ... |
| ... | ... | ... |
| $a_{i,n}$ | $a_{i,n+1}$ | $a_{i,n+2}$ |

In this way, we get the training matrix. The numbers of columns and rows vary from

different time scales, and we will discuss later.

*C.  Testing Matrix*

For the testing matrix, the number of rows of testing matrix is the same as the training matrix. However, the number of columns of testing matrix is up to the test data and there is no relationship between any two columns. And the rest data arranges in the chronological order in each column.

# III. WIND SPEED FORECASTING

## 3.1. Introduction

Due to Neural Network can approximate any nonlinear mapping through its learning ability, it can be applied to the forecasting of nonlinear system in the time series prediction [21].

First, this section introduces the basic concept and principle of Neural Network. Then, we build the BP neural network forecasting model and AdaBoost_BP improved forecasting mode on Matlab software. And make the short-term wind speed forecasting based on the actual historical data collected from wind farms. Finally, analyze the results of each forecasting model and make contrast.

## 3.2. The Basic Principle of Artificial Neural Network

Neural Network (NN), also called as Artificial Neural Network (ANN) or Neural Computing (NC), is to abstract and model the human brain or biological neural network. It has the ability that can learn from the circumstance and adapt to the environment by the interaction similar to that between organisms to achieve the information handle ability which is similar to brain, like learning, recognition, memory, etc [22].

Neural Network not only has the good non-linear ability and the adaptive ability, and it also has the associative learning ability and strong fault tolerance. Therefore, in the

aspect of forecasting there is a large development space and good practical application for Neural Network.

Specifically, Neural Network has the following advantages: [23]

1) Because the information is scattered stored in the neuron of Neural Network, Neural Network has strong robustness and fault tolerance.

2) By using parallel processing method, the structure of Neural Network is parallel and each unit of Neural Network can run the similar progress at the same time. And improve the computation speed in this way.

3) Neural Network can deal with the uncertain or unknown system for its self-learning, self-organizing, adaptivity, various connections between neurons and certain plasticity of the structural junctions between neurons. Neural Network can be trained by using the actual statistics data, and the trained network has the extensive adaptability, which means Neural Network can work even the inputs are not provided in the training. And Neural Network can be trained online.

4) Neural Network can approximate any complex non-linear relationship. For the modeling and forecasting of complex nonlinear systems, neural network is more practical and economical than other methods.

5) Neural network has strong information comprehensive ability. It can handle quantitative and qualitative information at the same time, coordinate the relationship of a variety of input information and deal with complex non-linear and uncertain objects.

Wind speed and wind power forecasting is a highly uncertain and complicated

nonlinear process which involves various factors. The No.4 and No.5 advantages of Neural Network are a suitable for this part. In addition, the wind power farm is flexible, in other words it is general stage construction. The No. 3 advantage of Neural Network can works with the changing object of forecasting. And No.2 advantage of Neural Network marches the high computation speed requirement of the online wind power forecasting. Therefore, it will largely simplify the modeling work and improve the forecasting precision by selecting Neural Network method for the nonlinear forecasting research on the wind speed and power.

### 3.2.1 Neuron Model

Neuron is the basic data processing unit of Neural Network, and it is also a nonlinear component with multiple inputs and single output. The inputs of neurons are the external inputs or the outputs of other neurons in the Neural Network, and they determine the response of the neuron. There are three important parts of the neuron:

A. Connection Weight

Input $x_i$ connects to Neuron $k$ through Path $i$. It needs to multiply with the associated weight when it passes the Path.

B. Sum Function

Add the inputs of one neuron (the external inputs or the outputs of other neurons in the Neural Network) together to get the sum of the weighted value.

C. Transfer Function

Use the sum of the weighted value through the transfer function to get the only

output of the neuron. The transfer function can limit the amplitude the neuron output. Generally, the range of the amplitude is [0,1] or [-1,1].

Transfer function, also called as excitation function, describes the transfer characteristic of the neurons. Its basic function is:

a) controlling the activation of inputs to output,

b) the function transformation of inputs and output,

c) transforming the inputs which may be the infinite range to the limited range output.

The relationship of the inputs and output can be described in the formula below [24]:

$$I = \sum_{i=1}^{n} w_i x_i - \theta$$

$$y = f(I)$$

where, $x_i$ stands for the inputs (the external inputs or the outputs of other neurons), $\theta$ refers to the threshold value, $w_i$ is the connection weight stands for the strength of the connection , and $f(x)$ is the excitation function or transfer function.

For simplicity, take $-\theta$ as the weight of $x_0$ which identically equal to 1. And the formula will be:

$$I = \sum_{i=0}^{n} w_i x_i$$

Where $w_0 = -\theta$ and $x_0 = 1$.

### 3.2.2 Commonly Used Transfer Function

The commonly used transfer functions [24,25] show in Figure III-1. The different

transfer function determines the different characteristics of the neuron output.
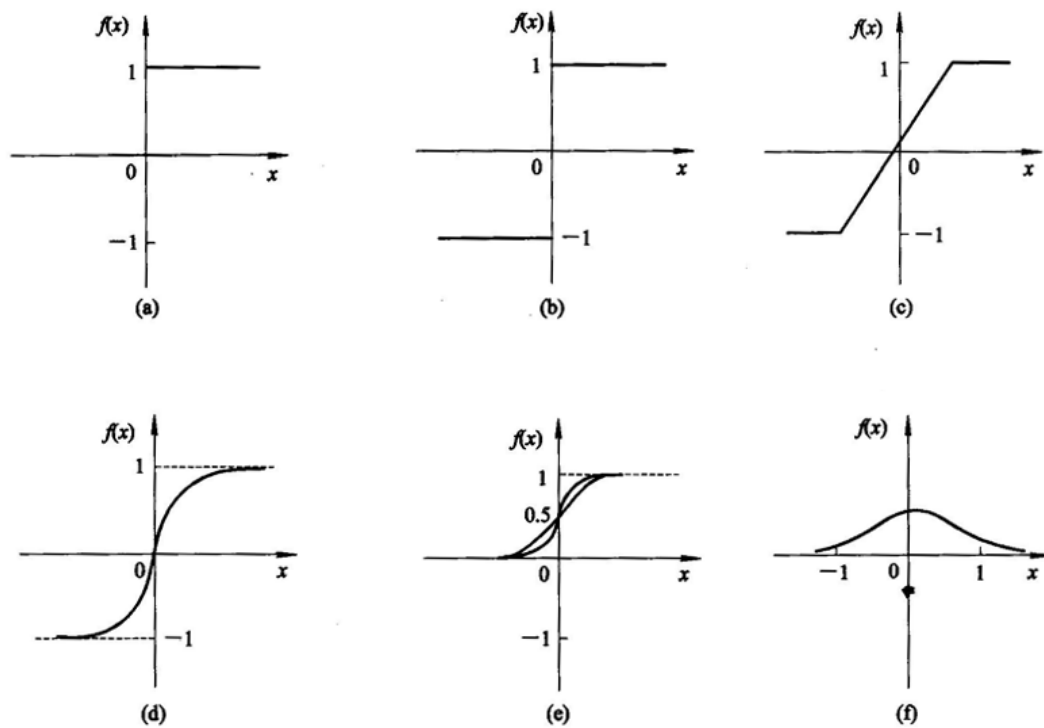


Figure III-1. Commonly Used Transfer Functions

## A. *Threshold Function*

In Figure III-1, (a) and (b) are the threshold functions. This function turns any input to two kinds of output ±1 or (0,1).

When $y_i$ equal to 0 or 1, $f(x)$ is the step function shows in (a).

$$f(x) = \begin{cases} 1 & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

When $y_i$ equal to -1 or 1, $f(x)$ is the sign function (sgn) shows in (b).

$$\text{sgn}(x) = f(x) = \begin{cases} 1 & (x \geq 0) \\ -1 & (x < 0) \end{cases}$$

Its main feature is non-differentiable and step type. And it often uses in Perceptron Model, M-P Model and Hopfield Model.

*B. Saturated Function*

In Figure III-1, (c) is the saturated function. Its formula shows:

$$f(x) = \begin{cases} 1 & \left(x \geq \dfrac{1}{k}\right) \\ kx & \left(-\dfrac{1}{k} \leq x < \dfrac{1}{k}\right) \\ -1 & \left(x < -\dfrac{1}{k}\right) \end{cases}$$

Saturated function is essentially the combination of the threshold function and the linear function. Its main feature is non-differentiable and step type. And it often uses in Cellular Neural Network, such as Pattern Recognition, Character Recognition or Noise Control, etc.

*C. Hyperbolic Function*

In Figure III-1, (d) is the hyperbolic function, also called as sigmoid symmetric function. Its formula shows:

$$f(x) = \tanh(x) = \frac{1 - e^{-I}}{1 + e^{-I}}$$

Its main feature is differentiable and step type. And it often uses in BP Model or Fukushina Neocognition Model.

*D. Sigmoid Function*

In Figure III-1, (e) is the sigmoid function. Its formula shows:

$$f(x) = \frac{1}{1 + e^{-\beta I}}$$

The relationship between the input and the neuron state is a monotone function which can value continuously in (0, 1). When $\beta \to \infty$, sigmoid function trends to step function. Generally, $\beta$ equals 1. Similar to the hyperbolic function, it often uses in

BP Model or Fukushina Neocognition Model.

*E.  Gaussian Function*

In Figure III-1, (f) is the Gaussian function:

$$f(x) = \exp\left[\frac{(I-c)^2}{b^2}\right]$$

where, $c$ stands for the central value of Gaussian function, and the function is longitudinal axisymmetric when $c$ equals to 0 threshold value. $b$ is the scale factor of Gaussian function， and it determines the width of the Gaussian function. Hyperbolic Function, Sigmoid Function and Gaussian function are continuous transfer function.

### 3.2.3    Neural Network Structure

The structure of Neural Network largely determines the effect of neural network, and the learning algorithm of Neural Network is selected according to the structure of Neural Network. Therefore, choosing a suitable Neural Network structure, according to the specific problem, for the system identification, modeling or forecasting is an important step in establishing the neural network model. Neural Network can be classified in forward and feedback neural network according to the network structure. They are also known as can be static neural network and dynamic neural network. In this paper, there are the brief introductions of single layer forward neural network, multilayer forward neural network and Elman neural network.

*A.  Single layer forward neural network*

Single layer forward neural network connects the input variables and output

directly by one layer of neurons. It complete the training of network by adjusting

the deviation $b$ and weight $w_{ij}$

B. *Multilayer forward neural network*

The neurons of Multilayer forward neural network connect with the neurons of

one or two adjacent layer. The first layer is input layer, the middle is the hidden

layer and the last layer is output layer. Neural Network has one or multiple hidden

layers. In the network, adjacent layers connect each other in interconnect mode.

There is no link between the neurons in the same layer and there is no feedback

loop in the whole network. And there is no direct connection between the input

layer and output layer. Hidden layer is used as the intermediary of the external

inputs and the network output.

C. *Elman Neural Network*

Elman feedback Neural Network is a kind of two layers neural network with

feedback. Feedback connects to the inputs from the output of first layer by

delaying and storing. The number of feedback neurons is the same as that of the

neurons of hidden layer, and its input is one step delay of the output of hidden

layer neurons.

### 3.2.4    Neural Network Training

After determined the construction of Neural Network model, Neural Network needs to

establish the model to solving a specific problem by learning and training. The

training process of Neural Network is a adjusting the weights of network. The weight of network is the only element for the performance of Neural Network processing unit. Therefore, adjusting the connection weights can make processing unit reflect the model of the given data correctly.

The Neural Network learning aims to reduce the error caused by equations by adjusting the network parameters, such as weights and thresholds of the network which can storage the knowledge from learning stage, according to previous experience. Generally, learning algorithms can be divided into supervised learning method and unsupervised learning method.

In supervised learning method, inputs and the actual output sometimes are requested. Based on the current output and the required target output, the network will adjust the network parameters until the network complete the study of the desired response. A supervised learning process is a closed loop adjustment process from the output to the inputs. On the contrary, the learning progress of the unsupervised learning method just needs a certain number of input learning samples, and the network can make the response to the input patterns without the output learning samples.

In this paper, all the trainings of the neural network forecasting model use the supervised learning algorithm.

## 3.3. BP Neural Network

BP Neural Network is a kind of multilayer feed forward neural network. It names

from the adjustment of the network weight is backward propagation learning algorithm. BP learning algorithm was put forward by Rumelhart in 1986. Since then, the BP Neural Network was obtained in a wide range of practical applications. According to statistics, 80% - 90% of Neural Network models adopt BP network or its transformation form.
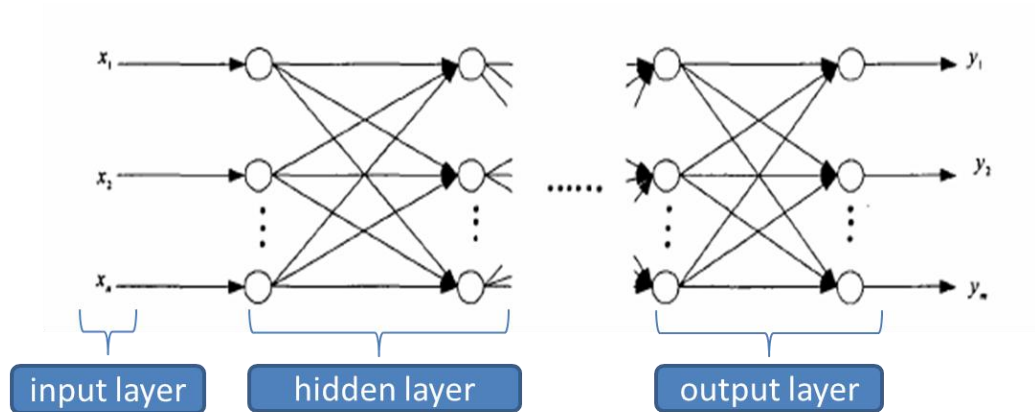


Figure III-2. BP Neural Network

Figure III-2 [26] is the BP Neural Network diagram. The signal propagation of the network consists of two parts, the forward and backward propagation. In the forward propagation stage, learning samples feed into the input layer, and send to output layer after the step by step operation of hidden layer. The neuron state of each layer only affects the neuron state of next layer. If the output layer did not get the desired output, that is there is an error between the network actual output and the desired output, calculate the error of the output layer, and then go to the error back propagation stage. At this time, the error signal goes back to the input layer from the output layer along the original connection. Then adjust the connection weight layer by layer to make the error minimize.

BP neural network has been used in a variety of forecasting process, and achieved good effects. Therefore, the classical BP Neural Network is one of the choices in this paper.

### 3.3.1    Steps of BP Neural Network Learning Algorithm

（1）    Initialization: In this progress, all weighting coefficient are set to be the smallest random number.

（2）    The input vectors $x_1$, $x_2, \dots$, $x_n$ and the desired output vectors $t_1$, $t_2, \dots$, $t_n$ are the training sets for the network.

（3）    Calculate the actual output of each neuron of the hidden layer and output layer.

（4）    Get the error of the desired output and the actual output based on Step 3.

（5）    Adjust the weighting coefficients of output layer and hidden layer based on Step 4.

（6）    If the error did not meet the requirements, continue the progress by returning to Step 3.[27]
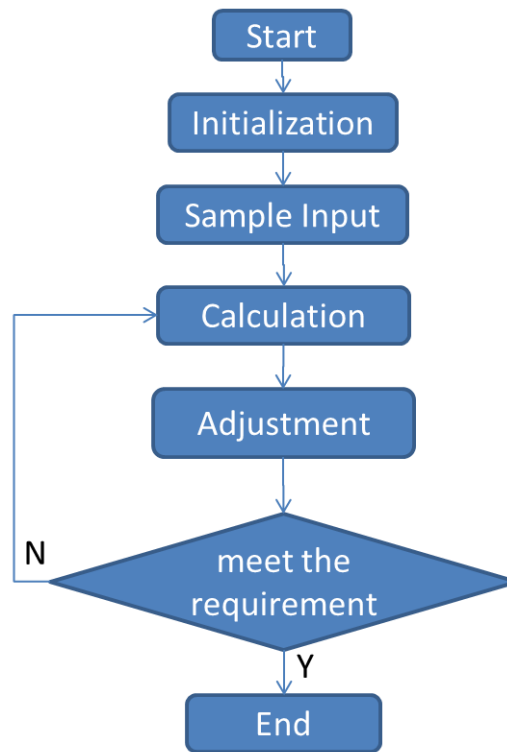
Figure III-3. BP Neural Network Learning Steps

### 3.3.2   The Short-term Wind Speed Forecasting Based on BP Neural Network Learning Algorithm

In this paper, by wind turbine selection we decide to use the data of No.17 wind turbine in 2011 as the sample data. And the forecasting time scales are 10min, 1h and 3h.

*A.  Matlab Used in Wind Speed Forecasting Based on BP Neural Network*

In Matlab software, there is a Matlab neural network toolbox which based on artificial neural network theory [28]. The BP Neural Network in this sample can be established by calling three functions *newff* , *sim*  and *train* , in Matlab neural network toolbox. Figure III-4 is the BP Neural Network model built in Matlab software.
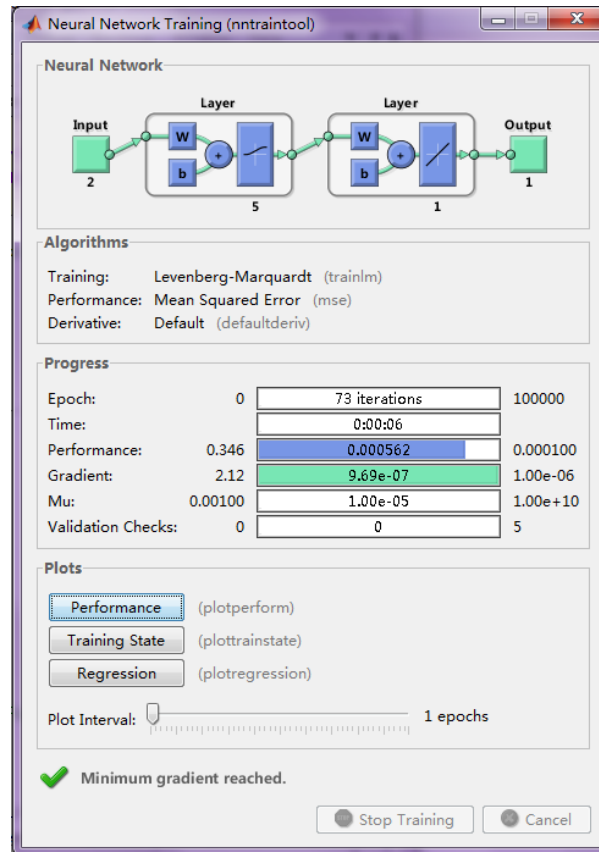
Figure III-4. BP Neural Network Model in Matlab

The function of *newff* is mainly setting parameters of the BP neural network. The functional form shows in two ways.

a) $net = newff(P,T,[S1\ S2\ ...S(N-1)],\{TF1\ TF2...TFN1\},BTF,BLF,PF,IPF,OPF,DDF)$

Where:

P: R x Q1 matrix of Q1 sample R-element input vectors

T: SN x Q2 matrix of Q2 sample SN-element target vectors

Si: Size of $i_{th}$ layer, for N-1 layers, (Default = [Output layer size SN is determined from T].)

TFi: Transfer function of $i_{th}$ layer. (Default = 'tansig' for hidden layers and 'purelin' for output layer.)

BTF:     Back propagation network training function (Default = 'trainlm')

BLF:     Back propagation weight/bias learning function (Default = 'learngdm')

IPF:Row cell array of input processing functions.

    (Default={'fixunknowns','removeconstantrows','mapminmax'})

OPF:     Row cell array of output processing functions.

    (Default={'removeconstantrows','mapminmax'})

DDF:     Data divison function (Default = 'dividerand')

Usually, the first six parameters are mainly set. P is the input data matrix, and in this sample, P is the matrix form of the wind speed history data arranged for network training. T is the output data matrix which is used to store the result of the short-term wind speed forecasting, and its function is similar with input data matrix.

b)    $net = newff\,(PR,[S1\,S2\,...SN],\{TF1\,TF2...TFN\},BTF,BLF,PF)$

PR：     Rx2 matrix of min and max values for R input elements.

Si：     Size of $i_{th}$ layer, for Nl layers.

TFi：     Transfer function of $i_{th}$ layer, Default = 'tansig'.

BTF：     Back propagation network training function, Default = 'trainlm'.

BLF：     Back propagation weight/bias learning function, Default = 'learngdm'.

PF：     Performance function, Default = 'mse'.

The first expression has less training times than the second one, but the training accuracy of the first one is less than that of the second one. In this study, we take the second expression.

The number of input layer node is the number of wind speed which is prior to the forecasting moment in each time scales. In this paper, we will discuss three time scales. We will discuss the most suitable number of the training data for each training sets for each time scale, so the input layer node numbers of each time scales are not the same. In this sample, the hidden layer node number is set to be 5. Output layer node number is 1, that is, the predicted value export through the Neural Network.

In this sample, we adopt the linear transfer function ( *purelin* function) from the node transfer functions and the BP gradient descent training algorithm function (*trainglm*) from BTF training function. The last parameters which needs to set is network learning function BLF, and we choose *learngdm* which has the momentum as the network learning function.

For the whole network, we also need to set its training function and forecasting function.

The functional form of the BP neural network training function is:

$$[net, tr] = train(NET, X, T, Pi, Ai)$$

As the *newff* function has already been set, we can call the input data matrix, output data matrix directly to set the parameters of the training function. The last two parameters are used to initialize the conditions of the input and output layer. Generally, they are set as the system default parameters, and there is no do need to decide by ourselves.

In the training function, we need to set the learning rate, and it is set to be 0.05 in this paper (general range is 0.01-0.8). In order to ensure the stability of the system, generally choose the small learning rate. Besides, we also need to select an appropriate expected objective error, and the expected objective error in this sample is set to be 0.0001. And setting an appropriate iteration number can prevent that neural network go into an iteration for not reaching the expected objective. The iteration number in this sample is 100000.

The last function for the wind speed BP Neural Network forecasting modeling in Matlab neural network toolbox is the forecasting function. The functional form of the forecasting function is:

$$y = sim(net, x)$$

Where $net$ is the BP Neural Network model which is established in the $newff$ function and well trained by the training function $train$. The $x$ is the history wind speed input used to forecast. As the network is already trained, the functional relationship between the input and the forecasted value can be established by the functions, and get the objective forecasted wind speed at last.

*B. Wind Speed Forecasting by BP Neural Network*

According to the mathematical analysis above and the optimization algorithm referred, we can make short-term wind speed forecasting with the sampling points we get from SCADA system. The number of the input and output layer in the forecasting process can be set according to the number of independent variables and dependent variables

of the function.

In this paper, we will talk about the wind speed forecasting in the time scales of 10min, 1h and 3h in the future. As the sampling interval is 10min, the data used for 10min forecasting is the sampling data and there are two ways of building the training data for 1h and 3h forecasting that we select. For example, we can use the average hourly value as the element of the input matrix of the training function in 1h forecasting, so the time scale of the element of input matrix is the same as that of the forecasting. We name this model as the single-step wind speed forecasting model. And the second way for 1h forecasting is different in two places. One place is the element of the input matrix of the training function which is the sampling value of 10min in the second way, and the other place is the result of the forecasting which is the average hourly value of six 10min forecasting speeds in the hour. For the forecasting characteristic of the six speeds which we will talk about later, we call this model the iterative multi-step wind speed forecasting method.

a) The single-step and iterative multi-step wind speed forecasting model

The wind speed forecasting models are for the wind speed of 10min, 1h and 3h in the future. As the sampling interval is 10min, the 10min time scale forecasting has only single-step wind speed forecasting model, while 1h and 3h time scale have both single-step and iterative multi-step wind speed forecasting model.

i. The single-step wind speed forecasting model

For 10min time scale forecasting model, the element of the neural network training input matrix is the wind speed of every 10min sampling interval, and

the element of training output matrix is the wind speed of 10min in the future. The testing result is the wind speed of 10min in the future.

For 1h time scale forecasting model, the element of the neural network training input matrix is the hourly wind speed which is the average of six adjacent 10min sampling intervals, and the element of training output matrix is the wind speed of 1h in the future. And the testing result is the wind speed of 1h in the future.

For 3h time scale forecasting model, the element of the neural network training input matrix is the wind speed which is the average of eighteen adjacent 10min sampling interval, and the element of training output matrix is the wind speed of 3h in the future. And the testing result is the wind speed of 3h in the future.

ii. The iterative multi-step wind speed forecasting model

For 1h time scale forecasting model, the element of the neural network training input matrix is the wind speed of every 10min sampling interval, and the element of training output matrix is the wind speed of 10min in the future. The element of the testing result is the hourly wind speed which is the average of six adjacent 10min forecasting wind speeds. The first 10min forecasting wind speed is $Y_1$, and the testing input matrix of $Y_1$ is composed by $z$ 10min sampling intervals. The second forecasting wind speed is $Y_2$, and the testing input matrix of $Y_2$ is composed by $z-1$ 10min sampling intervals and $Y_1$ as the last one element of the matrix. Similarly, $Y_3$ is

composed by $z-2$ 10min sampling intervals and $Y_1, Y_2$ as the last two

elements of the matrix. $Y_4$ is composed by $z-3$ 10min sampling intervals

and $Y_1, Y_2, Y_3$ as the last three elements of the matrix. $Y_5$ is composed by

$z-4$ 10min sampling intervals and $Y_1, Y_2, Y_3, Y_4$ as the last four elements of

the matrix. $Y_6$ is composed by $z-5$ 10min sampling intervals and

$Y_1, Y_2, Y_3, Y_4, Y_5$ as the last five elements of the matrix. And the testing result is

the average value of $Y_1, Y_2, Y_3, Y_4, Y_5, Y_6$ which stands wind for speed of 1h in

the future.

Table III-1　Element of Multi-step Wind Speed Forecasting Matrix

| | column i | column i+1 | column i+2 | column i+3 | column i+4 | column i+5 |
|---|---|---|---|---|---|---|
| Training Input Matrix | $a_{i,1}$ | $a_{i+1,1}=a_{i,2}$ | $a_{i+2,1}=a_{i+1,2}$ | $a_{i+3,1}$ | $a_{i+4,1}$ | $a_{i+5,1}$ |
| | $a_{i,2}$ | $a_{i+1,2}=a_{i,3}$ | ... | ... | ... | ... |
| | $a_{i,3}$ | ... | ... | ... | ... | $Y_1$ |
| | ... | ... | ... | ... | $Y_1$ | $Y_2$ |
| | ... | ... | ... | $Y_1$ | $Y_2$ | $Y_3$ |
| | ... | ... | $Y_1$ | $Y_2$ | $Y_3$ | $Y_4$ |
| | $a_{i,n}$ | $Y_1$ | $Y_2$ | $Y_3$ | $Y_4$ | $Y_5$ |
| 10min Forecasting Speed | $Y_1$ | $Y_2$ | $Y_3$ | $Y_4$ | $Y_5$ | $Y_6$ |
| 1h Forecasting Speed = | $\dfrac{Y_1+Y_2+Y_3+Y_4+Y_5+Y_6}{6}$ | | | | | |

For 3h time scale forecasting model, the element of the neural network

training input matrix is the hourly wind speed which is the average of six

adjacent 10min sampling intervals, and the element of training output matrix

is the wind speed of 1h in the future. The element of the testing result is the

wind speed of 3h in the future which is the average of 3 adjacent 1h

forecasting wind speeds. The first 1h forecasting wind speed is $Y_1$, and the testing input matrix of $Y_1$ is composed by $z$ hourly wind speeds which is the average of six adjacent 10min sampling intervals. The second forecasting wind speed is $Y_2$, and the testing input matrix of $Y_2$ is composed by $z-1$ hourly wind speeds and $Y_1$ as the last one element of the matrix. $Y_3$ is composed by $z-2$ hourly wind speeds and $Y_1, Y_2$ as the last two elements of the matrix. And the testing result is the average value of $Y_1, Y_2, Y_3$ which stands wind for speed of 3h in the future.

Here is Table III-2 of the single-step and iterative multi-step wind speed forecasting model.

Table III-2. Single-step and Iterative Multi-step Wind Speed Forecasting Model

| Model | | training input wind speed | training output wind speed | testing result wind speed |
|---|---|---|---|---|
| 10min | | 10min sampling interval | 10min in the future | 10min in the future |
| 1h | single-step | average of 1h sampling wind speed | 1h in the future | 1h in the future |
| | multi-step | 10min sampling interval | 10min in the future | 1h in the future $\frac{Y_1,Y_2,Y_3,Y_4,Y_5,Y_6}{6}$ |
| 3h | single-step | average of 3h sampling wind speed | 3h in the future | 3h in the future |
| | multi-step | average of 1h sampling wind speed | 1h in the future | 3h in the future $\frac{Y_1,Y_2,Y_3}{3}$ |

b) The number of the training data for each training sets

When we established the BP Neural Network of wind speed forecasting model, we found that the number of the training data for each training matrix, that is to say the length of the column of the training matrix, affected the final result of the accuracy of forecasting speed.

Therefore, we want to find the optimal number of the training data for each

training sets for the BP Neural Network, and the optimal number for each time scale should be discussed separately. For each time scale forecasting model, we use three different training matrixes for the BP Neural Network training, and test the trained network by the same testing data. Thus, we can make contrasts of the accuracy based on the same testing data. The percentage is the relative error. Its formula shows below:

$$Relative\ Error = \frac{\left|Real\ Value - Predicted\ Value\right|}{Real\ Value} \times 100\%$$

For 10min time scale forecasting model, we test the number 2, 4, 6, 8, 10, 15, 20 of the column of the training matrix. And the accuracy of number 15 and 20 are not high, so we only list the number 2, 4, 6, 8, 10 below. For each number, we test seven times and the result is the average value.

Table III-3. Number of Training Data for Each Training Sets for 10min

| Number | matrix 1 | matrix 2 | matrix 3 |
|--------|----------|----------|----------|
| 10 | 15.19% | 17.14% | 17.54% |
| 8 | 16.55% | 14.61% | 17.44% |
| 6 | 15.72% | 16.41% | 19.75% |
| 4 | 18.48% | 17.91% | 17.89% |
| 2 | 15.39% | 17.85% | 16.51% |

From Table III-3, we can see the number 2 can achieve the relatively stable accuracy for 10min time scale forecasting.

For 1h time scale forecasting model, we test the number 2, 4, 6, 8, 10 of the column of the training matrix. For each number, we test seven times and the result is the average value.

Table III-4. Number of Training Data for Each Training Sets for 1h

| Number | matrix 1 | matrix 2 | matrix 3 |
|--------|----------|----------|----------|
| 10 | 13.43% | 13.85% | 21.10% |
| 8 | 14.70% | 14.46%• | 23.62%• |
| 6 | 14.96% | 15.27%• | 21.98% |
| 4 | 15.35% | 14.18%• | 18.43% |
| 2 | 15.44% | 14.57% | 20.27% |

From Table III-4, we can see the number 10 can achieve the relatively stable accuracy for 1h time scale forecasting.

For 3h time scale forecasting model, we test the number 2, 4, 6, 8, 10 of the column of the training matrix. For each number, we test seven times and the result is the average value.

Table III-5. Number of Training Data for Each Training Sets for 3h

| Number | matrix 1 | matrix 2 | matrix 3 |
|--------|----------|----------|----------|
| 10 | 20.86% | 23.82% | 24.06% |
| 8 | 20.08% | 18.21% | 21.44% |
| 6 | 19.86% | 22.26% | 20.62% |
| 4 | 19.42% | 22.90% | 20.99% |
| 2 | 21.42% | 24.54% | 22.15% |

From Table III-5, we can see the number 6 can achieve the relatively stable accuracy for 3h time scale forecasting.

After analyzing the effect of different number of the column of the training matrix on the final result of the accuracy of forecasting speed, we can make a conclusion: when establish the forecasting model, choose a different number of the column of the training matrix can get a different accuracy according to the different samples and the forecasting time scales. This is a way to improve the

accuracy of the forecasting result.

## 3.4. Adaboost

In light of the accuracy of present BP Neural Network wind speed forecasting model is not high, we decide to use Adaboost [29] algorithm to improve the BP Neural Network algorithm. Adaboost is an iterative algorithm. Its core concept is to get a powerful learner by training the weak learners in the same training set and bringing them together, and the generalization ability and forecasting accuracy of the algorithm can be improved in this way.

### 3.4.1   Adaboost Algorithm Overview

Adaboost algorithm is an important machine learning algorithm of feature classification, and mainly used to solve the problems of classification and regression. At present, this algorithm has been applied to the forecasting of power system load [30] and traffic volume[31], and get a good forecasting result.

Adaboost algorithm improves the generalization ability of the powerful learner through the combination to the complementary of weak learners. Its main work focuses on the samples whose prediction error is big and the weak learners which have good performance. It works in enhancing the weight of the samples who has poor training effect and the weak learners whose learning capacity is good in the training set, and reducing the weight of the samples who has good training effect and

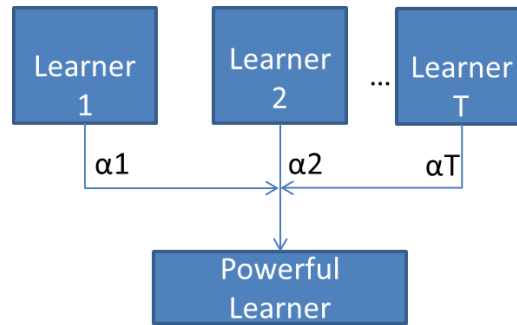the weak learners whose learning capacity is weak.



Figure III-5 Adaboost Algorithm

In this paper, for the limitations of BP neural network itself and the subjective factor of the training sample, we build the Adaboost_BP Neural Network forecasting model which is the combination of Adaboost and BP Neural Network to enhance the prediction accuracy. This model uses BP Neural Network as the weak predictor (weak learner), and reduce or increase its corresponding weight according to the pros and cons of each training sample. And then the model will retrain the weak predictor with the samples of the changed weight. At last, integrate the results of the weak predictor training and get the final result.

The weight of each sample is the same at the initial state, and we can train a basic weak predictor $h_1(x)$ under this sample distribution. Then determine the weight $w_1$ of this weak predictor $h_1(x)$ and the weights of the samples in this training set, and increase the corresponding weights of the samples with relative large forecast error while decreasing the corresponding weights of the samples with small forecast error in $h_1(x)$. Next, get the weak predictor $h_2(x)$ and its weight $w_2$ under the new sample distribution. After T times of circuit trainings, we can get T times of weak

predictors and the corresponding weights. Then get a powerful predictor which consists of T times of weak predictors according to their corresponding weights.

The procedure of Adaboost algorithm shows below:

1) First we set the weak learner algorithm and the training set.

$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, $x_i \in X$, $y_i \in Y$. $X$ and $Y$ represent a domain or Instance space.

2) Initialize the weights of N samples. For example, if the sample distribution is uniform, its weight is $D_t(i) = 1/N$ where $D_t(i)$ stands for the sample weight at t-th iteration and N is the number of samples in the training set.

3) Train the weak learner $h_t(x)$ under the probability distribution of its samples.

4) Calculate the error of each sample $\varepsilon_i$ and the average error $\varepsilon_t = \dfrac{1}{N}\sum_{i=1}^{N}\varepsilon_i$ of this weak learner.

5) Update the sample weight $D_{t+1}(i) = \dfrac{D_t(i)\beta_t^{-\varepsilon_i}}{Z_t}$. The weight of this weak learner is $W_t = \dfrac{1}{2}\ln(1/\beta_t)$, where $\beta_t = \dfrac{\varepsilon_t}{1-\varepsilon_t}$ and $Z_t$ is the normalization factor for

$$\sum_{i=1}^{N} D_t(i) = 1.$$

6) Then go back to step 3 for next iteration until T times of iterations.

Get the powerful learner composed by the trained weak learners $H(x) = \sum_{t=1}^{T} W_t h_t(x)$.

## 3.4.2 BP Neural Network Algorithm Based on Adaboost

In Adaboost_BP wind speed forecasting model, we use BP Neural Network wind

speed forecasting model as the weak learner, and then train the BP Neural Network

with the samples according to their weights. After 10 times of circuit trainings, we can

get 10 BP Neural Network predictors and their corresponding weights. Then we get

the powerful predictor Adaboost_BP wind speed forecasting model with 10 BP

weighted predictors.
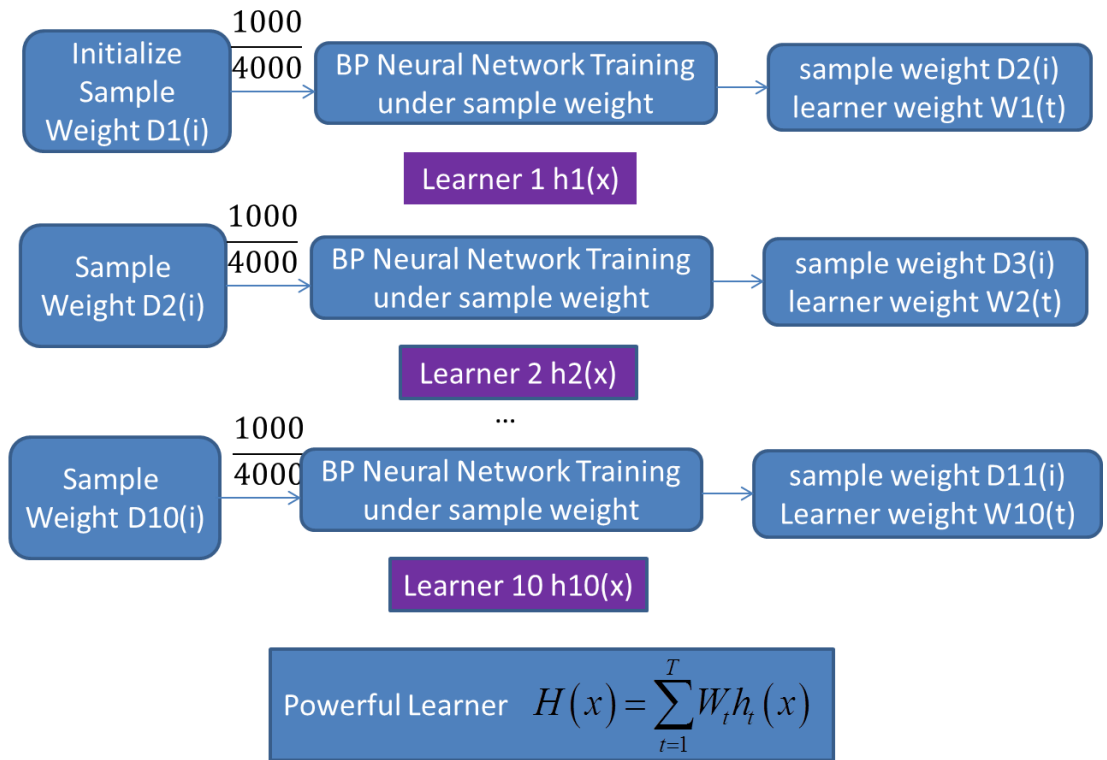


Figure III-6. Adaboost_BP Neural Network Learning Steps

For the weak predictors of Adaboost_BP, the functions and parameters of these

functions are the same as these of the BP model. In the BP model, we set the hidden

layer node number to be 5 and the output layer node number to be 1. *logsig* is the

transfer function of the first layer and the output layer transfer function is *purelin*.

And the training function is *trainlm*. In the training function, the learning rate is 0.05.

The expected objective error is set to be 0.0001. And the iteration number in is 100000. The minimum gradient requirement is $10^{-6}$. For the parameters of Adaboost, the training set has 4000 samples, and every time there are 1000 samples selected as the input according to its weight. The samples whose weight is heavy will have a greater possibility to be chosen as the inputs of the next training than whose weight is light, so that the next network focuses on the samples whose prediction error.

After we set up all these parameters, we can make forecast of the wind speed. Here is the comparison of the single-step wind speed forecasting accuracy between BP Neural Network model and Adaboost_BP model in 10min, 1h and 3h at their respectively optimal number of the training data for each training sets. The comparison of each time scale is made under the same random data.

Table III-6  Wind Dpeed Forecasting Result of BP and AdaBoost_BP
at Their Respectively Optimal Number of Training Data

| Time Scale | BP | AdaBoost_BP | Number |
|---|---|---|---|
| 10min | 15.39% | 13.31% | 2 |
| 1h | 13.44% | 11.83% | 10 |
| 3h | 19.86% | 16.72% | 6 |

The accuracy of Adaboost_BP model is higher than that of BP Neural Network model in the three time scales at their respectively optimal number of the training data for each training sets.

Then we can make the contrast between the accuracy of the single-step and iterative

multi-step wind speed forecasting model. This contrast is made in two time scales, 1h

and 3h, and the number of the training data for each training sets is set to be 10. And

the comparison of each time scale is made under the same random data.

Table III-7 Wind Speed Forecasting of Single-step and Iterative Multi-step Model

| Time Scale | BP | AdaBoost_BP |
|---|---|---|
| 1h_ multi-step | 11.91% | 10.88% |
| 1h_ single-step | 13.44% | 12.72% |
| 3h_ multi-step | 16.56% | 15.32% |
| 3h_ single-step | 20.86% | 17.11% |

The accuracy of Adaboost_BP model is higher than that of BP Neural Network model

in the same model. Compared with the accuracy of single-step, the multi-step has a

higher accuracy in both BP Neural Network and Adaboost_BP model. However,

arithmetic speed of the BP Neural Network is faster than that of Adaboost_BP model

for Adaboost_BP model calculates T times BP Neural Network. Thus, selecting the

algorithm for wind speed forecasting should consider the requirement both arithmetic

speed and accuracy.

# IV. WIND POWER FORECASTING

## 4.1. Introduction

From the previous work in this thesis, we discussed about the influencing factors of the wind power, and find out that wind speed is the most important one among these factors. Therefore, we decide to use the wind speed as the one of the input for the wind power forecasting model. In the part of **BP Neural Network Algorithm Based on Adaboost**, we already make a contrast of the forecasting accuracy between the BP Neural Network forecasting model and Adaboost_BP Neural Network forecasting model for wind speed. When we build the forecasting model for wind power, we should consider the arithmetic speed and the accuracy requirement and choose an optimal one for all the requirements.

## 4.2. Wind Power Forecasting Model

For the wind power forecasting model, we choose BP Neural Network as the method in this paper.

### 4.2.1 Input matrix

For the limited set of data we have in the research, the parameter we chose from the SCADA system as one of the inputs for the wind power forecasting model is wind speed.

Considering we can get the forecasting value of the wind speed from previous work, we will use the forecasting wind speed and its corresponding value of wind power as the input of the network training for the wind power forecasting model. As we want to forecast the wind power of 10min, 1h and 3h in the future, its corresponding wind speed is the wind speed of 10min, 1h and 3h in the future. Thus, when we put the wind speed into the training matrix as a part of the input, we prefer to use the forecasting wind speed instead of the original one. In this way, we can reduce the error which is caused by the training data. And if we put the corresponding wind power of the forecasting wind speed in the network training, we can get the relationship between forecasting wind speed and the real time wind power. Therefore, the input matrix is made up by two parts. One part is the forecasting wind speed, and the other part is its corresponding wind power.

Similar to the wind speed forecasting model, the number of the training data for each training sets affected the final result of the accuracy of forecasting power. If we choose $N$ moments prior to the forecasting point, the first $N$ elements of the input training matrix are the forecasting wind speed at the previous $N$ moments and they are in the time order of sequence arrange. And the $(N+1)th$ data is the forecasting wind speed of the forecasting point of the forecasting model. Then the last $N$ elements of the input training matrix are the measured values of the wind power which are corresponding to the first $N$ the forecasting wind speed on the timeline. The total number of the column of the input matrix is $2N+1$. Like the wind speed, the number of the training data for each training matrix, that is to say the length of the

column of the training matrix, affected the final result of the accuracy of wind power

forecasting result. For each time scale forecasting model, we test the number 10, 15,

20, 25 and 30 of the column of the training matrix. And we test seven times and the

result is the average value.

Table IV-1 Number of Training Data for Each Training Sets

|  | 10min | 1h | 3h |
|---|---|---|---|
| N | 15 | 20 | 15 |

The percentage is the relative error. And for each number, we test seven times and the

result is the average value. From this table, we can see 15, the number of N, can

achieve the optimal forecasting accuracy for10min time scale forecasting model.

By contrasting the forecasting accuracy of different number, we can get the suitable

number for each time scale.

### 4.2.2 The Range of Wind Power

When we study the error of the forecasting wind power in the previous work, we find

that the error varies from range of the wind power. After analyzing the differences of

the error, we find that the size of difference has a certain rule. The range of wind

power is the key point. Thus, we divide the wind power into four value ranges, they

are less than 100 kW, 100-500 kW, 500-1000 kW, more than 1000 kW.

We discuss the problem in the time scale of 10min. And in order to make the result

more convincing, we use different number of the training data for each training sets to

make a contrast. Here is the result.

Table IV-2     Wind Power Forecasting in Different Wind Power Range

| Length | <100kW | 100-500kW | 500-1000kW | >1000kW |
|--------|--------|-----------|------------|---------|
| 10 | 25.24% | 13.83% | 15.38% | 13.51% |
| 15 | 23.00% | 14.88% | 15.31% | 13.17% |
| 20 | 23.29% | 15.22% | 15.66% | 13.20% |
| 25 | 29.10% | 17.25% | 19.29% | 13.76% |
| 30 | 25.73% | 17.10% | 16.74% | 14.46% |

From the table, we can see the error varies from the size of the wind power. Although the errors in different number of the training data vary from each other, they have the same tend that the wind power range which is less than 100 kW has the maximal error and it is far from the other values. Therefore, we can make a conclusion that not only the number of the training data for each training sets but also the range of wind power affects the forecasting accuracy.

While we get the conclusion, there is another problem need to solve that which the wind power forecasting range is the one that we should focus on in our study.

The wind farm which is selected in this study has 31 wind turbines of the same type. The one we focus on is the No.17 wind turbine. When we want to find out the most suitable wind power forecasting range, we should consider the universality of the range. Therefore, the SCADA monitoring data of No.3, No.10, No.23 and No.31 were randomly selected.

In statistically analyzing wind speed data, the Bin method [32] was used to analyze these data. As we have made the curve fitting in previous work, we know that the wind speed and wind power are corresponding to each other. And the wind speed is the variable in the nature. So we focus on the study of the range of the wind speed.

First, we analyze the cumulative distribution probability of the wind speed of the No.17 wind turbine. The result is shown in Figure IV-1. The cumulative distribution probability of one wind speed is the sum of the probability which is equal or lesser than that wind speed. From this figure, we can see when the wind speed is 13.5 m/s, its cumulative probability is 99.22%. This means the wind speed which is equal or lesser than 13.5 m/s accounts for 99.22% of all the wind speed. That is to say the monitoring data is less than 0.8% when wind speed was greater than 13.5 m/s.

Then we analyze the wind speed cumulative distribution probability of the No.3, No.10, No.23 and No.31 wind turbines. As illustrated in Figure IV-2, the wind speed probability distributions of No.3, No.10, No.23 and No.31 wind turbines are similar to that of the No.17 wind turbine.
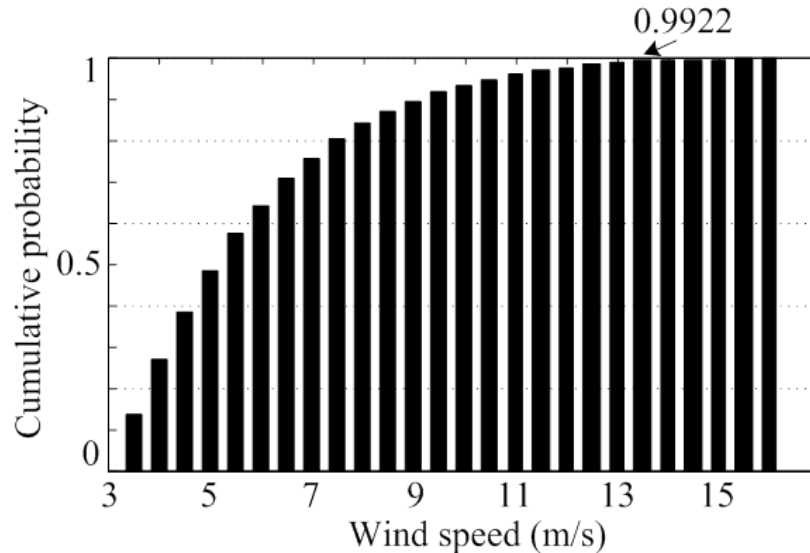


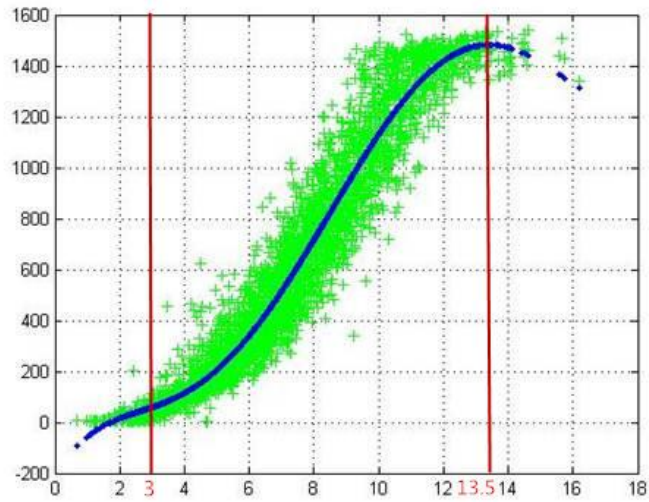Figure IV-1 Wind Speed Distribution of Wind Turbine No.17.

Figure IV-2 Forecasting Range of Wind Speed and Wind Power

Then we look for the range in the carving fitting of wind speed and wind power in

Figure IV-2, and we can see the actual value (the green spot stands for the actual value

and the blue one is the fitted data) of 13.5m/s can be 1500kW which is the rated

power of the wind turbine we used in the study. Considering the cut-in wind speed is

3m/s and the rated wind speed is 12m/s, the monitoring wind speed within the range

of [3 m/s to 13.5 m/s] is considered as the research data in this study. And the range of

the wind speed and wind power is shown in Figure IV-2. The wind power value is set

to be 1500kW when the wind speed is more than 13.5 m/s.

### 4.2.3    The Wind Power Forecasting

With the previous work, we get the wind speed range and the composition of the input

matrix. The wind power forecasting model is BP Neural Network method. The

functions are the same as that in the BP wind speed Neural Network method. The

learning rate is set to be 0.05, and the expected objective error is set to be 0.0001. The

iteration number in this sample is 100000.

First, we should get the forecasting wind speed data which will be used in the wind power forecasting model, and the data is obtained in the optimal forecasting result considering both arithmetic speed and accuracy in its time scale. That is to say, for 10min time scale wind power forecasting, the model is BP Neural Network single-step and the number of the training data is 2. For 1h time scale wind power forecasting, the model is BP Neural Network multi-step and the number of the training data is 2. For 3h time scale wind power forecasting, the model is Adaboost_BP multi-step and the number of the training data is 10.

Then we use the actual wind speed as the wind speed input, so that we can make a contrast of the wind power forecasting result with different wind speed input.

Here is the result of the wind power forecasting result.

Table IV-3 Wind Power Forecasting

| Time Scale | wind speed | wind power (forecasting wind speed) | wind power (actual wind speed) |
|---|---|---|---|
| 10min | 15.39% | 17.23% | 19.98% |
| 1h_ multi-step | 11.91% | 14.19% | 15.31% |
| 3h_ multi-step | 15.32% | 16.62% | 17.21% |

For wind power forecasting result, the wind speed used in the input training matrix affect the final result. The accuracy of the forecasting power with the forecasting wind speed is higher than that with actual wind speed. This is because using the forecasting wind speed instead of the original one can reduce the error which is caused by the training data and the noise which is produced in sampling process or data

transmission.

For the wind power forecasting model, considering the inputs are only wind speed and wind power, the accuracy can be improved by adding other influencing factors in the further study such as wind direction.

# V.   CONCLUSION

Chapter II selects the appropriate network input variables by analyzing the influence

factors of wind power. Then process the raw data for forecasting model. And find that

when establish the forecasting model, choose a different length of the column of the

training matrix can get a different accuracy according to the different samples and the

forecasting time scales. This is a way to improve the accuracy of the forecasting

result.

In Chapter III, we make a contrast of two wind speed forecasting models, BP Neural

Network and Adaboost_BP. Then find that the accuracy of Adaboost_BP model is

higher than that of BP Neural Network model in the same model for wind speed

forecasting. However, the improvement of forecasting accuracy from BP Neural

Network model to Adaboost_BP model has a certain limit, and the calculating time of

Adaboost_BP is 10 times as much as that of BP Neural Network. Thus, select the

algorithm should consider the requirement both arithmetic speed and accuracy.

Chapter IV we find that the accuracy of the forecasting power with the forecasting

wind speed is higher than that with actual wind speed. This is because using the

forecasting wind speed instead of the original one can reduce the error which is

caused by the training data and the noise which is produced in sampling process or

data transmission.

For the wind power forecasting model, considering the inputs are only wind speed and

wind power, the accuracy can be improved by adding other influencing factors in the

further study such as wind direction.

# VI. REFERENCES

[1] The Global Wind Energy Council. Global Wind Energy Outlook 2006[M]. The Global Wind Energy Council (GWEC), Belgium, 2006：5-10.

[2] Millais C. Wind Force 12 - A Blueprint to Achieve 12% of the World's Electricity from Wind Power by 2020[M]. GWEC, 2005：1-17.

[3] Liu H. Wind Speeed Forecasting for Wind Energy Applications[D]. Toronto：York University, 2009：1-10.

[4] Giebel G, Brownsword R, Kariniotakis G. The State-of-the-art in Short-term Prediction of Wind Power: A Literature Overview[R]. ANEMOS Project, 2003：1-36.

[5] Landberg L, Giebel G, Nielsen H A, et al. Short-term Prediction - an Overview[J]. Wind Energy, 2003, 6(3)：273-280.

[6] Landberg L. A Mathematical Look at a Physical Power Prediction Model[J]. Wind Energy, 1998, 1(1)：23-28.

[7] Wang X, Sideratos G, Hatziargyriou N, et al. Wind Speed Forecasting for Power System Operational Planning[C]. IEEE Int. Conf. on Probabilistic Methods Applied to Power Systems, 2004：470-474.

[8] Riahy G H, Abedi M. Short Term Wind Speed Forecasting for Wind Turbine Applications Using Linear Prediction Method[J]. Renewable Energy, Elsevier Science Ltd, 2008, 33(1)：35-41.

[9] Miranda M S, Dunn R W. One-hour-ahead Wind Speed Prediction Using a Bayesian Methodology[C]. IEEE Power Eng. Society General Meeting, 2006：1-6.

[10] Landberg L. Short-term Prediction of the Power Production from Wind Farms[J]. Wind Eng. & Ind. Aerodyn, Elsevier Science Ltd, 1999, 80(1-2)：207-220.

[11] Focken U, Lange M, Monnich K, et al. 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[J]. Wind Eng. Ind. Aerodyn, Elsevier Science Ltd, 2002, 90(3)：231-246.

[12] Sfetsos A. A Novel Approach for the Forecasting of Mean Hourly Wind Speed Time Series[J]. Renewable Energy, Elsevier Science Ltd, 2002, 27(2)：163-174.

[13] Cadenas E, Rivera W. Wind Speed Forecasting in the South Coast of Oaxaca,

Mexico[J]. Renewable Energy, Elsevier Science Ltd, 2007, 32(12)：2116-2128.

[14] El-Fouly T H M, El-Saadany E F, Salama M M A. Grey Predictor for Wind Energy Conversion Systems Output Power Prediction[J]. IEEE Trans on Power System, 2006, 21(3)：1450-1452.

[15] Shuang H, Yongqian L, Yongping Y. Taboo Search Algorithm Based ANN Model for Wind Speed Prediction[C]. IEEE Conf. on Industrial Electronics and Apply, 2007：2599-2602.

[16] Senjyu T, Yona A, Urasaki N. Application of Recurrent Neural Network to Long-term-ahead Generating Power Forecasting for Wind Power Generator[C]. IEEE Power Systems Conf. and Exposition, 2006：1260-1265.

[17] Wang Chengxu. Practical Technology of Wind Power[M]. Beijing：Shield Press, 1999：40-41.

[18] Thanasis G B, John B. Theocharis. Long-TermWind Speed and Power Forecasting Using Local Recurrent Neural Network Models[J]. IEEE Trans on Energy Conversion, 2006, 21(1)：273-284.

[19] Xie Jianhua, Cui Xinwei, et al. Prediction of Wind Powe Installed Capacity in China by Grey ForecastTheory[J].Xinjiang Agricultural Mechanization,2005(3):28-29.

[20] Wang Lijie, Liao Xiaozhong, Gao Yang, et al. Summarization of Modeling and Prediction of Wind Power Generation[J]. Power System Protection and Control,2009, 37(13):118-121.

[21] Wang Kejun, Wang Kecheng. Neural Network Modeling, Prediction and Control[M]. Harbin：Harbin Institute ofTechnology Press, 1996：104-108.

[22] Simon Haykin.Fundamentals of Neural Network [M].Beijing: China Machine Press, 2004: 109-121.

[23] Zhong Luo, Rao Wenbi, Zou Chengming. Artificial Neural Network and Its Application Technology[M]. Beijing: Science Press,2007

[24] Hou Yuanbin, Du Jingyi, Wang Mei. Neural Network[M]. Xi'an :Xian University of Electronic Science and Technology Press,2007

[25] Chen Xiangguang, Pei Xudong. Technology and Application of Artificial Neural Network[M]. Beijing:China Electric Power Press ,2003

[26] Han Wenlei.Nonlinear Analysis and Forecasting in China Stock Market:[D].Northwestern Polytechnical University,2005

[27] Martin Hagan T, Howard demuth B, Mark Beale H, etal.Neural network

design[M]. China Machine Press, 2002: 197-221.

[28] Shi Feng, Wang Xiaochuan, Yu Lei, et al. 30 Case Studies of Matlab Neural Network[M], Beijing: Beijing University of Aeronautcs and Astronautics Press, 2010.

[29] Morra J H，Zhuowen T U，Apostolova L G．Comparison of adaboost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation[J]．IEEE Trans on Medical Imaging，2010，29(1)：30-43．

[30] Gao Lin，Gao Feng，Guan Xiaohong，et al．Neural networks ensemble model based on boosting algorithm for short-term load forecasting[J]．Journal of Xi'an Jiaotong University，2004，38(10)： 1026-1030(in Chinese)．

[31] Chen Shuyan，Wang Wei，Qu Gaofeng，et al．Traffic flow forecast based on neural network ensemble[J]．Journal of Highway and Transportation Research and Development，2004，21(12)：80-83(in Chinese)．

[32] IEC International Standard. IEC61400-12-1 Wind turbines − Part 12-1: Power performance measurements of electricity producing wind turbines[S]. 2005.