J Electr Eng Technol.2018; 13(3): 1089-1098
http://doi.org/10.5370/JEET.2018.13.3.1089

ISSN(Print) 1975-0102
ISSN(Online) 2093-7423

Assessment of Wind Power Prediction Using Hybrid Method and Comparison with Different Models

Mohammed Eissa†, Yu Jilai*, Wang Songyan* and Peng Liu*

Abstract – This study aims at developing and applying a hybrid model to the wind power prediction (WPP). The hybrid model for a very-short-term WPP (VSTWPP) is achieved through analytical data, multiple linear regressions and least square methods (MLR&LS). The data used in our hybrid model are based on the historical records of wind power from an offshore region. In this model, the WPP is achieved in four steps: 1) transforming historical data into ratios; 2) predicting the wind power using the ratios; 3) predicting rectification ratios by the total wind power; 4) predicting the wind power using the proposed rectification method. The proposed method includes one-step and multi-step predictions. The WPP is tested by applying different models, such as the autoregressive moving average (ARMA), support vector machine (SVM), and artificial neural network (ANN). The results of all these models confirmed the validity of the proposed hybrid model in terms of error as well as its effectiveness. Furthermore, forecasting errors are compared to depict a highly variable WPP, and the correlations between the actual and predicted wind powers are shown. Simulations are carried out to definitely prove the feasibility and excellent performance of the proposed method for the VSTWPP versus that of the SVM, ANN and ARMA models.

Keywords: Wind power prediction, Hybrid, ARMA, SVM, ANN, Error.

1. Introduction

In the last ten years, the renewable-energy industry, especially wind-energy, has been rapidly developed in China, which has significantly contributed to the rapid economic growth and sustainable development in the country. Generally, wind energy has several advantages, for instance, lower cost, relatively mature technology, high reliability, etc. Thus, wind energy now plays an important role in the energy strategy all around the world. WPP is one of the most important topics of the wind power grid integration. WPP should be accurate enough to facilitate the practical control of the power grid.

It is imperative for the WPP to deal with the power variability and should be implemented before wind generation is integrated into the power grid. The predicting techniques can be classified into several types depending on the different predicting horizons. The horizon of very-short-term predicting can be 5 minutes up to 1 hour. For the VSTWPP, many statistical approaches, such as the ARMA, and ARIMA models are found to be useful to develop time-series models. An accurate prediction is crucial for ensuring operational security and enhancing the economic feasibility of the power system [1, 2].

† Corresponding Author: School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China. (moheissa183@gmail.com)

* School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China. (yupwrs@hit.edu.cn, {wangsongyan, liupeng_hit}@163.com)

Received: February 14, 2017; Accepted: January 26, 2018

An accurate WPP of up to 1 hour ahead could significantly contribute to reliable wind power integration. It can improve the stability of the wind power generation to the power grid [3]. As known, a good prediction ensures grid stability and has a positive trading performance of the power market, and can reduce the potential risks of the high penetration of wind power into the power grid [4, 5]. The accuracy of the short-term WPP was achieved by several studies such as physical numerical weather prediction models, statistical models based on historical data and statistical models with numerical weather prediction data as additional exogenous inputs [6].

In recent years, utilizing wind energy has shown rapid growth in many regions around the world [2, 3, 7]. China has a growing wind power that revealed a growth rate of 34% in 2020, and the generating capacity is assumed to reach 610 MW at that time [3-6]. Most of the wind power projects are achieved based on national goals without causing environmental pollution, and producing a high-output and reducing electricity cost [8].

The estimation of the wind power output with an accurate prediction would significantly reduce the uncertainty to the system operation. This can be achieved by using meteorological data. VSTWPP is required for the monitoring and scheduling of the power grid. The probabilistic prediction is definitely representing the best prediction solution, since it provides a wealth of uncertainty information on wind power [5]. The variability and high penetration of wind power in to the power grid cause many challenges to power system operators and instability [9, 10], if not properly managed.

In this study, we developed a modern hybrid model to predict the wind power with a high accuracy. This is achieved by applying a time-series analysis for the data obtained from the wind power plant, which relies on the predicting methods to use a time-series with different widths and period lengths of the prediction sample. The developed approach is the MLR&LS based hybrid model, which utilizes a new principle. The developed approach by using the MLR&LS for a hybrid model paves the way to utilize a new principle for the predication methods. The proposed hybrid model involves two main steps to predict the wind power. The first step is data transformation, which converts the historical wind power data into ratios and using these ratios as values for the prediction. The second step is inverse-transformation, which converts the predicted ratios into actual values and then predicts the total wind-power. In many research cases, there are sets of alternative predicting methods such as the SVM and ANN models. The WPP with a single model (PSM) may not always show the appropriate values. On the other hand, predicting with the hybrid model (PHM) is more suitable because the hybrid strategy can use the advantages of each PSM and is more likely to provide better results.

We propose a MLR&LS-based hybrid model for the VSTWPP. Simulation work is achieved via the proposed model, the ARMA model, the SVM model and ANN model. This is carried out by using the historical wind power obtained from the Northeast China electricity demand. The simulation results showed that the proposed hybrid model outperforms the other prediction models, i.e., the ARMA, SVM and the ANN.

2. Related Work and Model

2.1 Time-series models in prediction

Time-series modeling has attracted the attention of research community over the last few years. Its main objective is to carefully collect and rigorously analyze the past observations of a time-series to develop an improved model which describes the inherent structure of the series. The main idea of the time-series methods is that the evolution data in the past will continue into the future by stationary, trend-based and ext. From the applied time-series data analysis method, VSTWPP is important and contributes to the stability of a power system.

There are many different modern and traditional techniques to time-series modeling such as artificial intelligence and statistical method. The time-series analysis is now performed by using the ARMA model and MLR&LS model. There are three main objectives for the time-series analysis. The first one is to identify the nature of this phenomenon represented by the sequence of observations. The second one is to predict the future values of the time-series variable, and the third one is the ability of the time-series to predict the future based on the past history data. These objectives require identifying and describing the pattern of the observed time-series data. The time-series data of the WPP processing is developed by understanding the behavior of the wind power signals on different time scales. Researchers from different study fields have applied machine learning, data mining techniques, and statistical approaches in an attempt to solve problems associated with the time-series predictions [10-13].

2.2 ARMA model

The ARMA model is a type of time-series analysis with a statistical model proposed by the American G. Box and British G. Jenkins statisticians in 1970 [14]. This model is one of the most popular and frequently used models that utilize stochastic time-series models to predict the wind power with different times. Additionally, it can also be used to predict the behavior of the time-series from only historical values. This prediction serves as a baseline to evaluate the possible importance of other variables to the system [15]. The objective of the ARMA model for the WPP is to investigate the wind variability and its effect on the prediction accuracy [16].

ARMA models are mathematical models of the persistence, and can be described by a series of equations in the time-series. The general form of the ARMA model is stated as in (1):

$$W_t = \sum_{i=1}^{p} \phi_i w_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{1}$$

Where t is the time to be predicted; W_t represents the prediction value at time t; w_{t-i} represents the historical sample data at time $t-i$; ϕ_i is the autoregressive part; θ_j is the moving average part; ε_{t-j} represents the error at time $t-j$, is a random variable; p is the order of the autoregressive terms; and q is the order of the moving-average process.

The algorithm steps of the ARMA model for the WPP in Matlab code are as follows:

Step 1: Data processing

Assume that the known data (actual) $\{w_1, w_2, ..., w_n\}$ as the research object. Process the data according to the following equation to stabilize at a two order difference will be carried out

$$\begin{aligned} y_i &= w_i - w_{(i-1)} \qquad & i &= m, m+1, ..., n \\ l_i &= y_i - y_{(i-1)} \qquad & i &= m+1, m+2, ..., n \end{aligned}$$

Step 2: Correlation coefficient computing

Use the auto correlation coefficient (ACF) to find the stationary random sequence data and partial auto correlation coefficient (PACF) as the formulas

$$ACF_{(c)} = r(t,t+c)\Big/\sqrt{Dw_{(t)}Dw_{(t+c)}}$$
$$PACF_{(c)} = E[w_{(t)} - Ew_{(t)}][w_{(t-c)} - E_{(t-c)}]\Big/ E\{[w_{(t-c)} - Ew_{(t-c)}]^2\}$$

where $r(t,t+c)$ is the covariance.

Step 3: Order determining

Determine the order number of model (1) by ACF and PACF. If PACF is truncated at p, then it's an AR(p) sequence. If ACF is truncated at q order, it's an MA(q) sequence.

Step 4: Parameter estimating

Estimate the parameters of the model by the appropriate methods. There are three common methods that can be used for this task: moment estimation method, maximum likelihood estimation method, and least square's estimation method.

Step 5: Performance testing and application

Test the performance of the estimated model. If the error satisfies the requirements of the engineering application, the estimated data can be applied to the VSTWPP.

2.3 SVM model

The SVM model technique is closely related to the ANN model, in which the nonlinear functions are searched to model the complex composition of the variables. In the SVM nonlinear functions, it is converted by the *kernel* into a more dimensional space, where it is possible to express them in linear equations. The basic idea of the SVM is to choose the linear transformation function in order to transform the multi-dimensional training samples from the original space in to a high-dimensional space, and then construct the optimum regression function in the high-dimensional space structure. The SVM has been successfully used in terms of dealing with the nonlinear regression and time-series problems [17]. The SVM is a new powerful method for predicting the load consumption. The SVM based on the structural risk minimization principle has a better performance for the nonlinear short-term WPP and small sample problems [18], and has been developed for solving pattern recognition and nonlinear regression estimation problems.

The aim of the SVM is to minimize the upper bound of the generalization error, and to deal with nonlinear regression and time-series problems [17]. The SVM model has the ability to increase the prediction accuracy via the suitable selecting parameter [19]. The main idea of the SVM for regression is to map the data into a high-dimensional feature space by using a nonlinear mapping and to apply a linear regression in the feature space. Its general regression function is formulated as follows:

$$\hat{W}_i = mw_i + c \tag{2}$$

Where w_i and $\hat{W}_i$ are the actual and predicted wind powers at the i^{th} time interval, respectively, m is the scaling factor, and c the prediction axis.

2.4 ANN model

ANN has been widely utilized in various fields, including time-series prediction and pattern recognition. The ANN model is considered as an advanced technology of very-short-term prediction, as well as becomes an alternative technique for the time-series prediction, and gained immense popularity in the past. The ANNs have a comfortable and flexible structure to model a wide variety of nonlinear problems. The main advantage of the ANN model compared to the other classes of the nonlinear model is that it has a great potential to approximate a large class of functions with a high degree of accuracy [20]. Beside the other goal of use, ANN for the WPP is to find the neural network weights that reduce the model errors. Recently, ANNs have attracted increasing attention in the domain of time-series prediction [21]. The wonderful feature of ANNs is their inherent capability of non-linear modeling without any presumption around the statistical distribution followed by the observations when applied to time-series prediction problems. The ANN has general nonlinear mapping capabilities, and has increasingly attracted attention in the field of time-series prediction [17].

The ANN consists of many interconnections of identical basic processing units called neurons. Each contact of a *neuron* has the ability to edit the weight factor associated with the input and output. Every *neuron* in the network sums its weighted inputs to produce an activity level. In this paper, the ANN is composed of a three-layer feed-forward BP neural network, and all of the processes are performed on the input layer, hidden layer and output layer. A brief description of the sample set is as follows: training, validation and testing samples account for 70%, 15% and 15%, respectively. For this study, the data are used in the three phases to model the relationship between the input and the output target variable as shown in Fig. 1.

In general, the input-output relationship of the radial basis function (RBF) as ANN, widely used as a time-series prediction, can be written as follows:

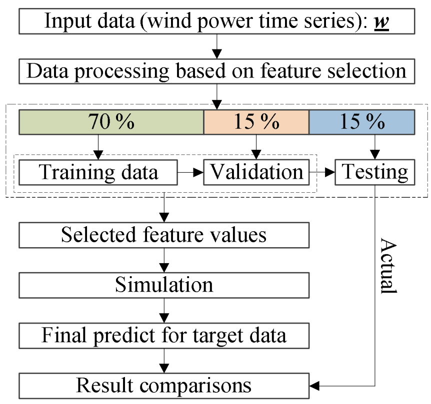


Fig. 1. ANN process method of prediction

$$W^{out} = \sum_{i=1}^{m} w_i \varphi_i \left(w^{in}, c_i, B_i \right) + b \quad (3)$$

Where W^{out} represents the dependent output variables (predicted wind power values); w^{in} represents a set of input variables (historical wind power data); m is the number of RBF units in the hidden layer; w_i and b are the weight and bias between the i^{th} RBF unit and the output; $\varphi_i(.)$ is the activation function; c_i and B_i are the center and width, respectively.

A number of tests has been performed in order to choose the suitable number of inputs and number of neurons in the hidden layers.

2.5 Multiple linear regressions

We first applied the multiple linear regression analysis to the WPP technique using the weighted least square estimation method based on a statistical analysis. This method is simple and commonly used to calculate the relationship between dependent and independent variables. The least square method is widely used to minimize the sum of the squared differences between the parameter estimating predicted values by MLR. Based on this method, it can lead to the least mean square error between the actual and predicted values.

2.6 One-step and multi-step predictions

Assume the time-series sequence of the actual wind power is denoted as $W = (w_1, w_2, ..., w_t)$, and the actual wind power w_t corresponds to the time t with a particular time period every one-minute. The one-step prediction is to get the predicted value of $\hat{w}_{t+1}$ with one step ahead based on the historical observations of $W_{t-p+1:t}$. The multi-step prediction is to get the predicted values of $\hat{w}_{t+1}$, $\hat{w}_{t+2}$, …, $\hat{w}_{t+n}$ with n steps ahead based on the historical observations of $W_{t-p+1:t}$. Typically, n can be 5 to10 steps ahead.

The prediction error (e) for the lead time $t+n$ is defined as the difference between the predicted and actual wind power values:

$$e_{t+n} = \hat{w}_{t+n} - w_{t+n} \quad (4)$$

3. Methods for VSTWPP

In this part, we will describe the detailed procedure of the VSTWPP with a maximum width of the prediction time being one hour. The output power of a wind farm (WF) in the power grid will be determined using historical wind power data by the hybrid time-series model.

3.1 Procedure of VSTWPP

Wind has random and intermittent features, and the output power of a wind farm represents this characteristic. Therefore, it is vital to reduce these influences for improving the accuracy of the WPP. This issue will be analyzed from the following aspects:

For different widths of the prediction (T), what is the most appropriate length of prediction in this study?

In this paper, the T will be assigned to 30 and 60 min; the length of prediction (Δt) will be assigned to 5 and 10 min.

- What's the performance difference between the predictions of PSM and PHM?

In this study, PSM is implemented by ARMA, SVM and ANN respectively, while PHM is implemented by the MLR&LS method.

- What's the performance difference between the one-step prediction and multi-step prediction?

In this paper, the one-step prediction is to predict the wind power of the next one length period, and the multi-step prediction is to predict the wind power of the next few lengths within the T.

- For multiple wind farms (WF), what's the performance difference between the direct predictions for the total wind power and indirect prediction for single wind power?

For the indirect prediction, first predict the wind power ratio of each WF; second, sum them to obtain the total wind power ratio of the multiple WF, and then predict the rectification ratios to calculate the WPP. The wind power of each WF can be from the direct WPP.

- What is the importance of the WPP?

WPP is essential for the higher penetration of the wind power into the power systems. Since no wind forecasting system is perfect, a thorough understanding of the errors that may occur is a critical factor for system operation functions, such as the setting of the operating reserve levels. However, the strong randomness and high uncertainty of wind power generation pose challenges to power system plans and operation. Thus, to improve the accuracy of the WPP and reduce the prediction errors are some of the fundamental and key issues of enhancing the wind power management and controlling the power system operation.

- Why the prediction values by the proposed method have a higher accuracy than those of other predictions used in this paper?

The main reason lies in the fact that the prediction with the proposed method uses the wind power ratio and its corrected value for the WPP. To a certain extent, the relationship between the single wind farm (WF) and the total wind power is better determined through the wind power ratio. In the process of the hybrid method, the historical data of the wind power time-series is first transformed into the form of the wind power ratio coefficients. The ratios for different WFs are then predicted using the MLR&LS model. The MLR allows finding the best linear prediction equation between the dependent and other predictor variables. The basic idea of the hybrid model in this study is to combine the MLR and LS (MLR

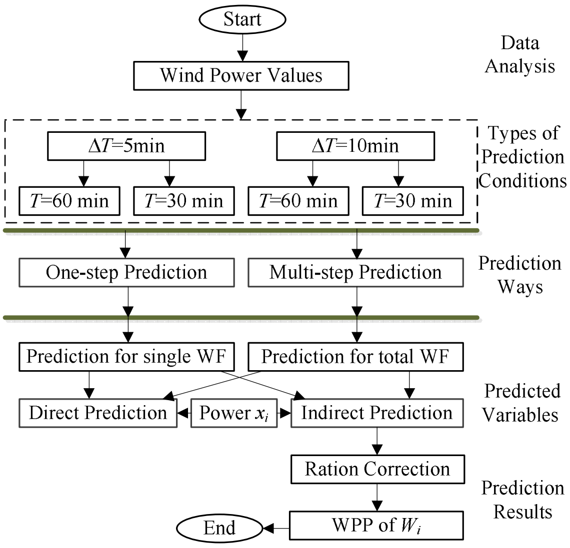


Fig. 2. Detailed procedure of VSTWPP

& LS), which retains the advantages of each approach and accesses to the largest combination advantages. The main goal of this is to improve the forecast accuracy. The LS method is widely used to minimize the sum of the squared differences between the estimated parameter values and the predicted parameter values by MLR. Finally, the WPP values of the different wind farms can be calculated based on the corrected wind power ratios and the predicted total wind power of the system.

Fig. 2 shows the detailed procedure of the VSTWPP designed in this study. In Fig. 2, the VSTWPP procedure collects data from the SCADA systems of the WF, and the power grid dispatches the center, and predicts the WPP of a single WF, total WPP of multiple WF, and the power ratios of different WFs at different Δt & T and with different prediction method and prediction types.

3.2 Processing framework of MLR&LS by the hybrid model

A hybrid model, which is based on a combination of the MLR and LS, is a new analysis tool that contains the statistical learning mechanism to explain the behavior of one of the dependent variables as a function of many other predictor variables. The MLR allows a flexible and easy way to find the best linear prediction equation between the dependent and other predictor variables. The object of this model is to benefit from the advantages of each model and obtain the optimal prediction performance. The performance of the WPP and the predict accuracy depends on the historical time-series data.

Next, we will review the hybrid model for the prediction method. The general form of the MLR model with response W and term's $w_1,........,w_p$ will have the form:

$$E(W/w)=\beta_0+\beta_1 w_1+........+\beta_p w_p \tag{5}$$

That conditioning of all the terms on the right side of the equation, and conditioning for specific values for the prediction as follows:

$$E(W_i/w_j)=\beta_0+\sum_{j=1}^{p}\beta_j w_j+\varepsilon_i \tag{6}$$

Where W_i is the predicted values; w_j is the historical values; p is the number of predictor variables; ε is the error terms; and β is the regression coefficient.

Using the ordinary least square to minimize a quantity called the residual sum of the squares. The fitted values for this case are given by:

$$\hat{W}=\hat{\beta}_0+\hat{\beta}_1 w_1+........+\hat{\beta}_p w_p \tag{7}$$

The ordinary least squares estimator's values β_0 and β_i that minimize the function are:

$$\sum_{i=1}^{n}(W_i-\beta_0-\sum_{j=1}^{p}\beta_j w_j)^2 \tag{8}$$

The optimal solution of the regression coefficients is:

$$\hat{\beta}=(w^T w)^{-1}w^T W \tag{9}$$

Transform the historical wind power values into the ratio of the total wind power by:

$$x_i=w_i\bigg/\sum_{i=1}^{n}w_i \tag{10}$$

Where $\sum w$ is the summing of the wind power from the total WF. Predict the wind power by the ratios and the total wind power as in (11):

$$W_i=x_i\times\sum_{i=1}^{n}w_i \tag{11}$$

The summation of the predicted power ratios X_i must be equal to 1 as in (12):

$$\sum_{i=1}^{n}X_i=1 \tag{12}$$

The predicted $[X_1,X_2,.....,X_n]$ is a vector. The summation of X_i might not be equal to 1. Thus, a rectification is needed as the following:

$$X_i^{\text{repred}}=X_i\bigg/\sum_{i=1}^{n}X_i \tag{13}$$

After predicting the total wind power $\sum W_i$ and power ratio by the rectification equation X_i^{repred}, the proposed method to calculate WPP by the rectification is as follows:

$$W_i^{\text{repred}}=X_i^{\text{repred}}\times\sum_{i=1}^{n}W_i \tag{14}$$

Where W_i^{repred} is the WPP value of the i^{th} WF by using MLR&LS with the rectification ratio.

4. Simulation and Results

In this section, simulations were carried out for VSTWPP using the time-series analytical method. The test data in this study are divided into two sets: 1) the comparing set; and 2) the confirmed set.

4.1 Comparison between one-step & multi-step predictions

The prediction accuracy is a function to describe the difference between the predicted value and actual values. In this section, the comparison of a numerical result of the one-step and multi-step prediction way is displayed in Table 1. In this table, the data are taken from the real wind power data, the T is assigned to 60 and 30 min, and the Δt is assigned to 5 and 10 min. To evaluate the accuracy of the approach in the WPP, different criterions are used, such as the root mean square error (RMSE), the mean absolute percentage error (MAPE), and the mean square error (MSE).

Table 1 shows that the one-step prediction method gives the better results, and is slightly more prone to the MSE values. This way could outperform the others if fewer T sets were used, and a lower Δt was predicted.

Fig. 3 gives the total WPP comparison of the one-step and multi-step prediction methods. In Fig. 3, T = 60 min, and Δt =5 min.

Table 1. RMSE, MAPE and MSE comparisons of the two methods

Error	One-step	Multi-step	One-step	Multi-step
	T = 60 & Δt = 5 min		T=60 & Δt =10 min	
RMSE	2.18	2.31	2.21	2.45
MAPE	2.61	2.78	2.64	3.31
MSE	16.17	16.68	16.26	18.19
	T =30 & Δt = 5 min		T=30 & Δt =10 min	
RMSE	1.48	1.50	2.02	3.27
MAPE	1.12	1.27	2.19	5.32
MSE	10.57	11.27	14.81	23.08

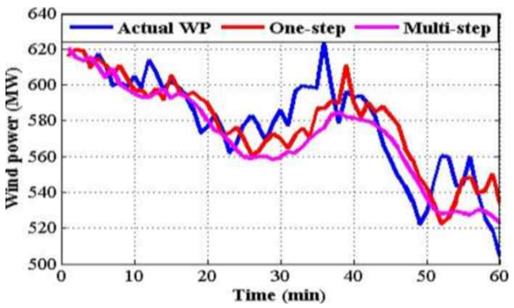


Fig. 3. Total WPP with one-step and multi-step predictions

From Fig. 3, the WPP curve with the one-step prediction method is closer to the actual wind power curve, and the correlation coefficient between the two curves is about 0.85; while the WPP curve in the multi-step prediction method is less accurate with a correlation coefficient of about 0.82.

4.2 Comparisons of single WF at WPP by direct and indirect prediction types

This part shows the comparisons between the single WPP at WF by the direct and indirect prediction types. During the process of comparison, the values of the single WPP at the two wind farms are used, and T = 60 min, and Δt =5 min. Figs. 4 and 5 give the actual and predicted wind power curves of the two wind farms, respectively, with direct and indirect predictions.

From Figs. 4 and 5, it was observed that the predicted curves were closer to the actual curves for the indirect prediction, thus the results of the indirect prediction are more satisfactory than those of the direct prediction for the two wind farms.

As part of the comparison between the direct and indirect WPP, the correlation of the predicted and actual variables under the conditions for the latter is higher than that for the former. Therefore, the indirect WPP using the rectification values causes it to be better than the direct WPP.

4.3 Comparisons between ARMA, SVM, ANN and proposed method

This paper studied the VSTWPP using four methods; i.e., the ARMA, SVM, ANN and proposed method. The simulation results for the VSTWPP using the statistical

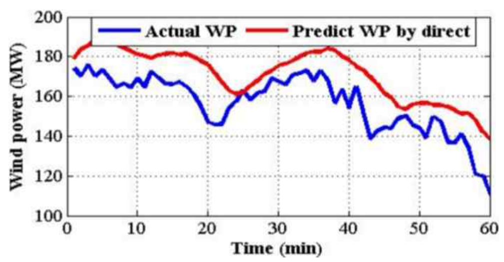


Fig. 4. Single WPP by direct prediction

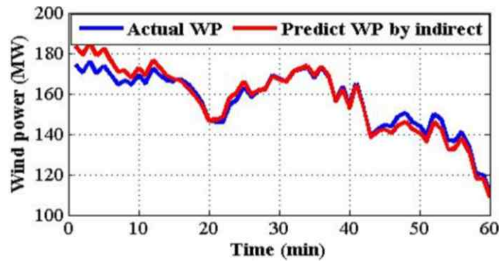


Fig. 5. Single WPP by indirect prediction

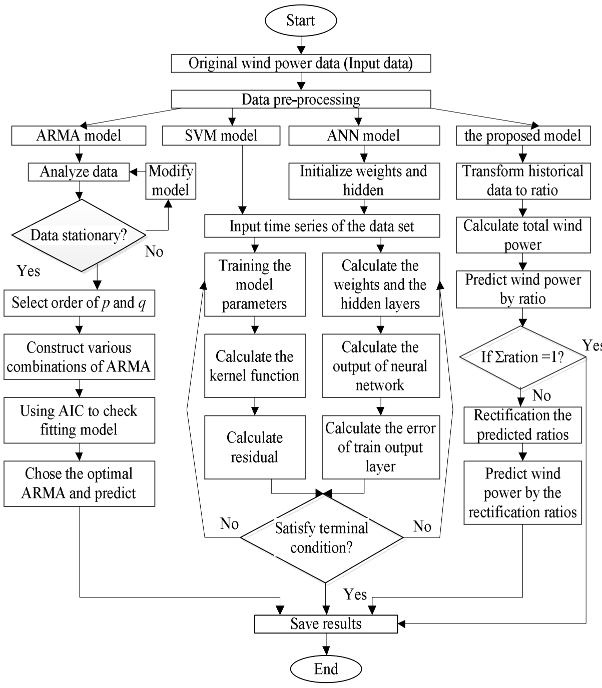


Fig. 6. Comparison procedures of the flowcharts among the 4 methods

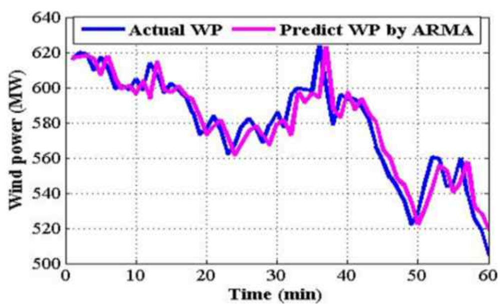


Fig. 7. Actual & predicted total wind powers by ARMA model

methods are presented. Fig. 6 gives the comparison procedures of the four methods. To assess the performance of the proposed method WPP scheme, the prediction method was built as follows.

The ARMA model is known to provide accurate results. Therefore, it is used as the benchmark in this study. The results indicated in Fig. 7 using the ARMA model show that the predicted values are very close to the actual values in the previous stage of the width of the prediction time window, and the ARMA model can get relatively more accurate results when dealing with the smallest length of the prediction. We can see that the predication values are almost similar to the actual values, and the error is very small.

As shown in Fig. 8, the WPP using the SVM model appeared to be close to the actual values. However, the maximum errors are less than 3.20%, as given in previous studies. The predictions by the SVM model are more

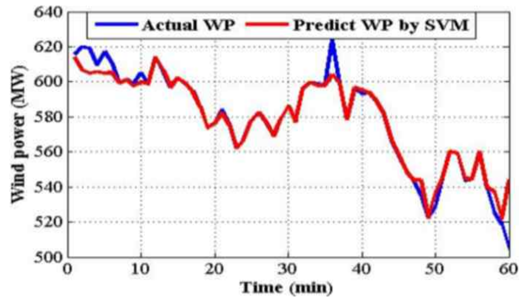


Fig. 8. Actual & predicted total wind powers by SVM model

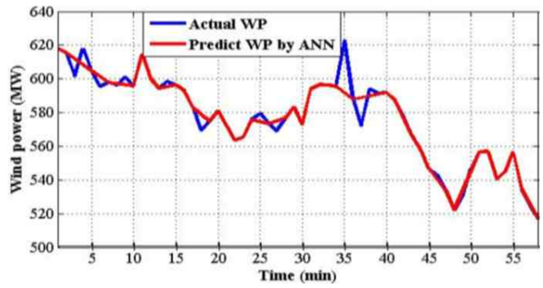


Fig. 9. Actual & predicted total wind powers by ANN model

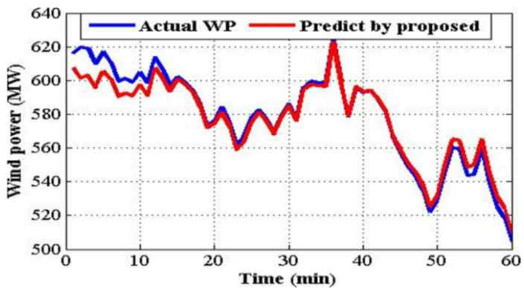


Fig. 10. Actual & predicted total wind powers by proposed method

accurate and suitable for the VSTWPP of the time-series models.

The prediction results of the ANN model are demonstrated after several experiments with various network architectures based on the BP algorithm as shown in Fig. 9 and Fig. 10. The high predicted values and the solid correlation between the actual and predicted values indicate high prediction accuracy. The ANNs have been successfully used to predict the integrated wind power.

The proposed model for the VSTWPP has been successfully implemented in this paper. Fig. 10 shows a time-series of the WPP that is closest to the time-series of the actual wind power. In this study, the final predicted value by the proposed model is more accurate than that by the other models. It is observed that the minimum errors are obtained by the proposed model. Fig. 10 shows the skills of the proposed MLR&LS model as a function of the WPP. This indicates that the proposed model significantly outperforms the other models used in this paper.

4.4 Comparison of the predicted results

The numerical results of the ARMA, SVM, ANN and the proposed models are depicted in Table 2. From Table 2 and the Fig. (7-10), a highly accurate WPP with the SVM is observed compared to the ANN and ARMA models. It can be said that each model has achieved a good predicted response and their error values are low. This result is consistent with the findings of previous studies. Moreover, we observed that the WPP accuracy level with SVM compared to ANN is not quite significant. However, it is observed that the proposed method has a high-accuracy level for the total WPP, when compared to the other methods in this paper. The numerical results in Table 2 showed that the proposed model is better than the other methods. We also noted that ARMA demonstrated a highly accurate prediction when using suitable parameters according to the algorithm steps and smallest Δt. These findings agree with several studies. Finally, all the comparisons, analysis and simulation results prove that the proposed method provides a better performance.

4.5 Prediction accuracy comparisons of the four models

In this section, we describe the framework for the evaluation of the error index about the prediction accuracy of the WPP methods. ARMA, SVM, ANN and the proposed models were adopted as time-series to the WPP in the total WFs of Northeast China. With the view to inspect the accuracy of the predicted results in general, the mean absolute percentage error (MAPE), mean square error (MSE) and correlation coefficient (ρ) were all investigated. The computation formula is shown as:

$$MAPE = \left[\frac{1}{N} \sum_{i=1}^{N} \frac{\left[w_i - \hat{W}_i \right]}{|w_i|} \right] \times 100 \tag{15}$$

$$MSE = 1/N \sum_{i=1}^{N} \left(w_i - \hat{W}_i \right)^2 \tag{16}$$

$$\rho = \frac{\sum_{1}^{N} \left[\left(w_i - \overline{w} \right) \left(\hat{W}_i - \overline{\hat{W}} \right) \right]}{\sqrt{\left(\sum_{i=1}^{N} \left(w_i - \overline{w} \right)^2 \right) \left(\sum_{i=1}^{N} \left(\hat{W}_i - \overline{\hat{W}} \right)^2 \right)}} \tag{17}$$

Where N is the number of data points; w_i and $\overline{w}$ are the actual and means wind power values; and $\hat{W}_i$ and $\overline{\hat{W}}$ are the predicted and means wind power values, respectively.

The prediction accuracy is a function of the difference between the predicted and actual values. Table 2 gives the errors for the four models. It can be inferred that the predicted values are close to the actual values when the proposed model is applied. In these calculations, it is also shown that there is a high correlation between the actual curve and predicted curve in all the models, but the proposed model provides the better result.

Table 2. Error and correlation coefficient of the 4 methods for VSTWPP

Model	MAPE	MSE	ρ
ARMA	1.42	1.03	0.93
SVM	0.48	0.49	0.97
ANN	1.31	1.01	0.94
proposed	0.69	0.32	0.98

From Table 2, it is observed that WPP by the proposed model outperformed all the other models. The proposed model has both the lowest error and the maximum correlation coefficient. Compared to SVM, ANN and ARMA, the proposed model has the superiority of a high VSTWPP accuracy. Based on Table 2 and Figs. 7-10, we evaluated the predictive performances of the four methods in detail. In the total WPP of the multiple WF, it is shown that the proposed model is better than SVM, ANN and ARMA.

Based on the simulation training for the historical data of the eight WFs, the VSTWPP is reasonable to avoid volatility due to the use of the proposed model. From Table 2, the proposed model clearly performs much better than the SVM, ANN and ARMA models as the calculation of the total WPP, MAPE and MSE of the proposed model are reduced, and ρ of the proposed model is increased.

5. Conclusion

In this study, the VSTWPP was evaluated by a combination of four methods, including the proposed method. The WPP obtained from SVM and ANN is suggested to be more accurate compared to the ARMA model. Overall, the proposed method is better than the SVM and ANN. In the WPP, wind power follows the peaks and valleys of the actual power component patterns. The shape of the WPP obtained from the proposed method is more accurate than that obtained from the other methods in this study, and the VSTWPP by the proposed method is close to the actual power. Therefore, it is concluded that the forecasting accuracy with the combination method performs better than the individual ones in general. As our simulation results indicated, the proposed method has three improvements and best performance versus the same method for the individual forecasting process and contributes to a reasonable improvement in the final WPP. After several experiments with different widths as the prediction time window for the proposed method, the smallest MSE was noted to give the best prediction accuracy for the test and validity data. To assess the results of different prediction methods, the errors are reports in Table 2, including the results of the prediction MAPE, MSE and ρ for the different predictor models.

Acknowledgements

This study was supported by the National Natural Science Foundation of China under Grant 51377035

References

[1] Gargeya M Anuj and Sai Praneeth Pabba, "Economic load dispatch using genetic algorithm and pattern search methods," *IJAREEIE*, vol. 2, no. 4, pp. 1203-1212, Apr. 2013.

[2] Milligan M, Schwartz M, and Wan Y, "Statistical wind power forecasting models: Results for US wind farms," *National Renewable Energy Laboratory, Golden*, CO, 1-1-14, May, 2003.

[3] Tao Yubo and Hongkun Chen, "A hybrid wind power prediction method," *PESGM IEEE*, pp. 1-5, Nov. 2016.

[4] Zeng Jianwu and Wei Qiao, "Support vector machine-based short-term wind power forecasting," *PSCE IEEE/PES*, pp. 1-8, 2011.

[5] Wu Wenzu, Chen K, Qiao Y, and Lu Z, "Probabilistic short-term wind power forecasting based on deep neural networks," *PMAPS, 2016 International Conference on. IEEE*, pp. 1-8, Dec.2016.

[6] Chen Niya, Qian Z, Nabney I T and Meng X, "Wind power forecasts using Gaussian processes and numerical weather prediction," *IEEE Transactions on Power Systems*, vol. 29, no 2, 656-665, Mar. 2014.

[7] Chen Niya, Qian Z, Nabney I T and Meng X, "Short-Term Wind Power Forecasting Using Gaussian Processes," *In IJCAI*, 2790-2796, Aug. 2013.

[8] Mao Yang and Jian Ma, "Data completing of missing wind power data based on adaptive BP neural network," *PMAPS, 2016 International Conference on. IEEE*, pp. 1-6, Dec.2016.

[9] Chang Wen-Yeau, "Short-term wind power forecasting using the enhanced particle swarm optimization based hybrid method," *Energies*, vol. 6, no. 9, pp. 4879-4896, Sep. 2013.

[10] Reliability; Vargas Luis, Gonzalo Paredes and Gonzalo Bustos, "Data mining techniques for very short term prediction of wind power," *In Bulk Power System Dynamics and Control (iREP)-VIII* (*iREP*) *IEEE*, pp. 1-7, Aug. 2010.

[11] Pai Ping-Feng, "System reliability forecasting by support vector machines with genetic algorithms" *Mathematical and Computer Modelling*, vol. 43, no. 3, pp. 262-274, 2006.

[12] Chang Wen-Yeau, "A literature review of wind forecasting methods," *Journal of Power and Energy Engineering*, pp. 161-168, Feb.2014.

[13] Aggarwal S K and Gupta M, "Wind power forecasting: A review of statistical models-wind power forecasting," *International Journal of Energy Science*, vol. 3, no. 1, Feb. 2013.

[14] Zhang Jianyong and Wang Cong, "Application of arma model in ultra-short term prediction of wind power," *CSA, 2013 International Conference on. IEEE*, pp. 361-364, Dec. 2013.

[15] Chan Ngai Hang, "Time series: applications to finance with R and S-Plus," *John Wiley & Sons, Second Edition*, 23-37, (2010).

[16] Rajagopalan S and Santoso S, "Wind power forecasting and error analysis using the autoregressive moving average modeling," *Power & Energy Society General Meeting, 2009. PES'09. IEEE*, pp. 1-6, Jul. 2009.

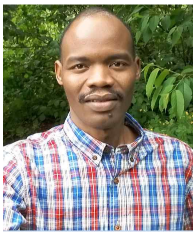
[17] Pai Ping-Feng, "System reliability forecasting by support vector machines with genetic algorithms," *Mathematical and Computer Modelling*, vol. 43, no. 3, pp. 262-274, 2006.

[18] Wang Jidong, Sun Jiawen and Zhang Huiying, "Short-term wind power forecasting based on support vector machine," *PESA, 2013 5th International Conference on. IEEE*, pp. 1-5, Dec.2013.

[19] Pai Ping-Feng, Wei-Chiang Hong and Yu-Shen Lee, "Determining Parameters of Support Vector Machines by Genetic Algorithms-Applications to Reliability Prediction," *International Journal of Operations Research*, vol. 2, no. 1, pp. 1-7, 2005.

[20] Areekul Ph, Senjyu T, Toyama H and Yona A, "A Hybrid Neural Network Model for Short-Term Price Forecasting in Deregulated Market," *IEEE Trans. Power Syst*, vol. 25, no. 1, pp. 524-530, Feb. 2010.

[21] Zhang G Peter, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing 50, 159-175, 2003.



Mohammed Eissa He received B.S and M.S degree in Dept. of Electrical Engineering at Kordofan and Sudan University in 2010 and 2012 respectively. And he is currently pursuing his Ph.D. degree in power engineering and its automation at HIT, Harbin China. He is interested in renewable energy resources, power system stability and security analysis, power grid planning and risk assessment.

Jilai Yu joined the School of Electrical Engineering and Automation in HIT in 1992. From 1994 to 1998, he was an Associate Professor with the School of Electrical Engineering and Automation, where he is currently a Professor and the Director of Electric Power Research Institute. His current research interests include power system analysis and control, optimal dispatch of power system, green power and smart grid.

Songyan Wang received the B.S., M.S., and Ph. D. degrees from the School of Electrical Engineering and Automation, HIT, in 2007, 2009, and 2012, respectively. He was a Visiting Scholar at VA, United States, in 2010. He is currently an Assistant Professor with HIT. His research interests include power system operation and transient stability analysis and control.

Peng LIU He received the B.S. and M.S. degrees in electrical power engineering and its automation from HIT, Harbin, China, in 2010 and 2012, respectively. He is currently pursuing his Ph.D. degree in power engineering and its automation at HIT. His current research interest lies in the dispatch of PEVs and renewable energies, integration of smart energy systems, etc.