Very-short-term wind power prediction by a hybrid model with single- and multi-step approaches

E Mohammed¹, S Wang¹ and J Yu¹
¹School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China
²E-mail: moheissa183@gmail.com

Abstract. Very-short-term wind power prediction (VSTWPP) has played an essential role for the operation of electric power systems. This paper aims at improving and applying a hybrid method of VSTWPP based on historical data. The hybrid method is combined by multiple linear regressions and least square (MLR&LS), which is intended for reducing prediction errors. The predicted values are obtained through two sub-processes: 1) transform the time-series data of actual wind power into the power ratio, and then predict the power ratio; 2) use the predicted power ratio to predict the wind power. Besides, the proposed method can include two prediction approaches: single-step prediction (SSP) and multi-step prediction (MSP). WPP is tested comparatively by auto-regressive moving average (ARMA) model from the predicted values and errors. The validity of the proposed hybrid method is confirmed in terms of error analysis by using probability density function (PDF), mean absolute percent error (MAPE) and means square error (MSE). Meanwhile, comparison of the correlation coefficients between the actual values and the predicted values for different prediction times and window has confirmed that MSP approach by using the hybrid model is the most accurate while comparing to SSP approach and ARMA. The MLR&LS is accurate and promising for solving problems in WPP.

1. Introduction
The electric power systems worldwide are undergoing rapid transformation in structure and functionality in search of a quest to increase efficiency and responsiveness. VSTWPP is considered as an important issue in the daily operations and control of the power system. Over the past few years, China has become the country with the largest wind power. By the end of 2015, total wind power installed capacity in China has reached 129 GW. It should be noted that wind-power load share in the power system is increasing. At the same time, there is lacked of high accurate prediction method. In addition, intermittent wind power penetration is increasing in the power grid, which poses new operational and regulatory challenges to power system. Wind energy is the most important renewable energy to share in the power system. Fluctuation from wind farm output power is a fundamental issue to power system operation and control.

Accurate VSTWPP with prediction time window being from minutes to 1h is critical to improve the schedule of maintenance and reserve generation for power system with high penetration wind power. Short-term WPP can help to reduce scheduling error, which was suggested to have a great impact on the power grid [1], and which is one work on this paper. However, fluctuations of wind power impact the area control system, when the system is interconnected [2-4].
The short-term power prediction is necessary in some applications, particularly in the plan, economic dispatch and security to the power system [5]. However, wind power is subjected to high penetration in the power grid where this penetration causes many challenges to power system operators [6].

During the last few years, more attentions to short and VSTWPP have concentrated on developing new prediction methods and strategies, such as using combined models [7]. Moreover, other methods, such as ANN, have been widely used for the field of WPP. ANN can compute nonlinear relationships between inputs and outputs, which conform to the non-linear characteristics to the power system [8-9]. ANNs have attracted increasing attention in the domain of time series prediction [10-11]. It is also applied to power load forecast, which is one of the necessary parts for smart grid [12-13].

Short-term WPP, up to 24 hours, using MLR combined with polynomial and steps composes, is suggested to provide good results [14]. The accuracy of short-term WPP is of key importance not only for reliability of power system operation but also for continuity of consumer service. WPP is a significant tool of the power market. In this regard, WPP method has been continuously improved over the last decade. WPP can be used in market analysis and operation [15].

Time series method is widely used in WPP. Generally, it has been suggested to use combined methods to obtain more accurate results [16]. Similarly, the hybrid strategy, such as the combination of seasonal ARIMA method and BP neural network model, can increase prediction accuracy [7]. ARMA model is a statistical model about persistence or autocorrelation. It is a powerful statistical tool for depicting the dynamics of time series by estimating the forthcoming value [17].

MLR method is one group of prediction techniques most commonly for short term. The mean feature of it is the flexibility, which explains its wide applications [18]. The common use of MLR method in power load prediction is for modeling the relationship of load consumption and other factors. The best fit in LS method is to minimize the sum of squared residuals. A residual is the difference between the actual value, and the fitted value provided by regression.

The purposes of this paper were to investigate prediction accuracy of wind power by a proposed hybrid method and to compare the hybrid method with other prediction techniques. In the current study, wind power data sets, including time series results, were obtained from the historical records of wind power. These data sets were applied to the WPP with time window being up to one hour. Contributions of this paper could be listed as following:

A-The proposed algorithm and its comparison to ARMA model. The algorithm can reduce errors based in the correction value.
B-Algorithm charts. Multi stages of the algorithm can reduce error.
C-Comparison from the PDF. The proposed method showed higher accuracy.
D-Accuracy evaluation. MSP has better prediction performance calculating numerical results of error.

2. Structure for MLR & LS based hybrid model
In this paper, development and application of a hybrid model for WPP is presented, which involves the combination of MLR & LS by using SSP & MSP approaches. This hybrid method provides a preliminary WPP based on the power ratio which is defined as the ration of wind power of a farm to total wind power within the system. The proposed method is illustrated in figure 1.
Figure 1. Outline of the proposed method (flowcharts of MSP and SSP are shown in the left and the right, respectively).

We describe the underlying structure of the hybrid method for WPP. This method is flexible, dynamic and accurate, easily understood and applied. It has been used for this study. As an example of the basic scheme of WPP by the proposed method, the steps are discussed in figure 2 showing the main processes of the MSP.

Figure 2. Simulation algorithm chart of proposal method (MSP).

The steps of the proposed method are provided as below:

Actual input data: historical data of wind power from actual wind farms based on the minute scale of time. These data are used during the research on WPP.

Transformation data: all data from this section are converted to ratios according to equation (1). These ratios are used to get the predicted ratio values.

\[ X_i(t) = \frac{P_i(t)}{P_{total}(t)} \]  

(1)

The \( X_i \) is fraction at any time \( t \) for the \( i^{th} \) wind farm; \( P_i \) is the actual wind power; \( P_{total} \) is the total wind power. The summation of \( X_i \) must be equal to 1 as that in equation (2). Since the summation may be not equal to 1, so it is needed to calculate a new ratio as equation (3).

\[ \sum_{i=1}^{n} X_i = 1 \]  

(2)

\[ X_i' = \frac{P_i}{\sum_{i=1}^{n} P_i} \]  

(3)

Where, \( X_i' \) is the relative asset value.

Next, use prediction method to predict a new ratio by using the modified data calculated as equation(3), and transform this predicted ratio to power finally as equation (4).

\[ \hat{P}_i = X_i' \times P_{total} \]  

(4)
3. Method used for VSTWPP

WPP methods are continuously updated to improve accuracy. The current research focuses on how to obtain predict accuracy probability distribution. In generation dispatching of power system, decisions must be made on real information with certainty and uncertainty, which must be predicted. There are many techniques for very-short-term WPP to improve accuracy and efficiency. They are broadly classified as two types of procedures, traditional and advanced model.

Traditional statistical method for WPP is generally multi parameter regression that uses input data depending on parameters such as the wind power for the current time, temperature and other factors [18]. These methods are commonly techniques to be used in short-term load forecasting of the power system. However, from numerous simulations, MLR is more flexible and quicker for WPP.

In this part, we provide some necessary details on the methods. First, give information from ARMA model; second, discuss how to reduce prediction error of the proposed method. Figure 3 shows the detailed procedure of WPP.

![Figure 3. Structure of the proposed method and comparison to ARMA.](image)

3.1. ARMA time series

In this work, we used ARMA model for the process of prediction, and the obtained results are compared with those of the proposed method. In this case; the proposed method showed better results that may be attributed to a fact that ARMA model depends on the order selection. (The comparison shape in (8) and table 1 and 2 validate such a statement through the prediction process). ARMA model can be described by time series equations in this paper ARMA using as tested. These equations can be simplified using an order only. The computation in this model depends on coefficients of the linear value and parameters within the model. The prediction model of ARMA is expressed as following:

\[
\hat{P}_t = \sum_{i=1}^{p} \varphi_i P_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}
\]  

(5)

Where, \( \hat{P}_t \) represents the predicted value at time \( t \), \( P_{t-i} \) represents the actual wind power at time \( t-i \), \( \varphi_i \) is the autoregressive parameter, \( \theta_j \) is the moving average parameter, and \( \varepsilon \) is random error.

3.2. ARMA modeling

Modeling with ARMA involves inclusion of wind power data, identification of shape, choice of parameter and estimation. Model validation is illustrated in figure 4.
3.3. The proposed method

Currently, there are several statistical techniques for prediction of time-series data. However, each one has some drawbacks. Therefore, we develop an approach to provide superior prediction results in the power system. Figure 4, table 1 and 2 give the prediction result of the hybrid method with MSP approach, which revealed more accurate results that is closer to actual values than ARMA. The general linear regression equation model can be written as follows:

$$\hat{P} = \beta P + \epsilon$$

(6)

Where, $\hat{P}$ is the dependent variable (output power, predicted values), $P$ is a vector of the variables (input data of wind powers); $\beta$ is the vector of the regression coefficients and $\epsilon$ is the error term.

In this study, the mathematical approach of MLR can be written as:

$$\hat{P} = \beta_0 + \beta_1 P_{i1} + \beta_2 P_{i2} + \ldots + \beta_{p-1} P_{i,p-1} + \epsilon_i$$

(7)

Minimize the error of MLR with LS methods as following:

$$\min \sum (\hat{P}_i - \beta_0 - \beta_1 P_{i1} - \ldots - \beta_{p-1} P_{i,p-1})^2$$

(8)

The optimal solution of $\beta$ is:

$$\hat{\beta} = (P^T P)^{-1} P^T \hat{P}$$

(9)

4. Results for VSTWPP

In this section, we present the WPP framework and some benchmarks: the proposed method and ARMA model used the current value to predict, and the waveforms of the predicted power curve confirm the effectiveness of the method. First, we analyze the prediction result of ARMA and the proposed hybrid models respectively with SSP & MSP approaches from figures 5-7.

4.1. Comparisons between ARMA model and the proposed hybrid model with SSP and MSP approaches

The simulation results of wind power value are shown in figure 5-6. The WPP values are not very similar to the actual values when ARMA is used to prediction with time window $\leq$ 1h. In addition, the
average relative error between actual and expected values is about 2.468% note (ARMA model used in this work as a comparison with the proposed method).

![Figure 5. Actual & WPP values by ARMA model.](image1)

![Figure 6. Actual & WPP value by hybrid model SSP.](image2)

From figure 6, the prediction accuracy of the proposed hybrid model with SSP approach has been improved while compared against the ARMA model. The WPP values are close to the actual values, and the average relative error is about 2.451%.

From figure 7, the WPP by the hybrid model with MSP approach has greatly improved comparing with ARMA model and the hybrid model with SSP approach. The WPP values are close to the actual values. The average relative error is 0.693%, which is the smallest one.

![Figure 7. Actual & WPP value by hybrid model MSP.](image3)

![Figure 8. WPP comparisons of 3 methods with time=1h.](image4)

Figure 8 gives the comparisons between the WPP values and the actual values about three methods above. Table 1 shows the average and maximum predicted errors. In this test, we evaluate the prediction performances of three methods. Considering eight wind farms, the results show that the hybrid model with MSP approach is the best. Furthermore, the MAPE and MSE of the hybrid model with MSP approach have the smallest prediction error.

When comparing the numerical results within the VSTWPP time frames, the MAPE of the hybrid with MSP approach is less than 1% and more than 2.4% for SSP approach and ARMA model. Although this value represents a strong positive correlation, it is noticeably smaller than the correlation coefficient calculated within 1 h.

4.2. VSTWPP error comparisons between ARMA model and the proposed hybrid model respectively with SSP and MSP approaches

Table 1 provides a comparative study between the proposed hybrid model and ARMA model based on the average and maximum predicted error criterion. The average predicted error of the proposed hybrid model with MSP approach is only 0.693%, which is the lowest among the three models. The predicted errors of the proposed hybrid model with SSP approach and ARMA model are more than 2.4%. Moreover, table 1 compares the proposed hybrid model with the ARMA model for maximum predicted error criterion, and the error of the proposed hybrid model with MSP approach is less than
that of other two models. This result confirms the accuracy of the proposed hybrid model to reduce prediction error.

Table 1. 1 h WPP and error values of the ARMA model & the MLR&LS with SSP&MSP approaches.

| Time(min) | Actual (MW) | ARMA Model | Error (%) | MLR&LS (SSP) | Error (%) | MLR&LS (MSP) | Error (%) |
|-----------|-------------|------------|-----------|--------------|-----------|--------------|-----------|
| 1         | 614.02      | 591.92     | 3.59      | 601.01       | 2.12      | 607.52       | 1.05      |
| 2         | 606.65      | 589.70     | 2.79      | 601.53       | 0.84      | 601.38       | 0.87      |
| 3         | 596.87      | 596.57     | 0.05      | 599.63       | 0.46      | 593.25       | 0.60      |
| 4         | 602.11      | 590.18     | 1.99      | 597.86       | 0.70      | 600.82       | 0.21      |
| 5         | 598.93      | 607.45     | 1.42      | 598.22       | 0.12      | 597.85       | 0.18      |
| ...       | ...         | ...        | ...       | ...          | ...       | ...          | ...       |
| 58        | 525.23      | 540.95     | 2.99      | 535.15       | 1.89      | 531.26       | 1.15      |
| 59        | 519.01      | 538.28     | 3.71      | 531.46       | 2.40      | 524.72       | 1.10      |
| 60        | 504.31      | 549.48     | 8.95      | 528.35       | 4.77      | 509.49       | 1.02      |

From table 1, average error (ARMA) = 2.468%, average error (hybrid model with SSP) = 2.451%, and average error (hybrid model with MSP) = 0.693%; maximum error (ARMA) = 8.96%, maximum error (hybrid model with SSP) = 8.57%, and maximum error (hybrid model with MSP) = 2.97%.

Figure 9 & 10 show the errors with time window ≤ 1h and normal probability distribution respectively. MSP provides lower errors than SSP, and ARMA does. The MAPE of them is 0.693%, 2.451% and 2.468% respectively. Besides, MSP has given the MSE of 0.32%, which is the smallest one while comparing with SSP and ARMA. The errors listed in table 2 confirm the credibility of the proposed hybrid model.

Using normal PDF to analyze WPP error, the function depends on two parameters, mean and standard deviation [19-21]:

\[ f(E) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(E-\mu)^2}{2\sigma^2}} \]  

(10)

Where, \( E \) is the error for predict values, \( \mu \) is the mean of the error and \( \sigma \) is the standard deviation. We use these parameters to describe the error out of this curve.

Figure 10 shows the distributions of WPP error using different prediction method with time window ≤ 1h. The WPP errors are displayed by the normal PDF. The figure shows that MSP performs the best, because it has the narrowest error distribution curves while comparing to SSP and ARMA. The largest and the smallest percentage errors are concentrated between -17% and 17% in the time to 1 h by MSP. At that time, errors mainly lie in the ranted [-46% 46%] and [-50% 50%] respectively by SSP and ARMA. This is an indicator of the accuracy of the proposed hybrid method.

Figure 9. WPP error comparisons of 3 methods with time=1h.  
Figure 10. Comparisons PDF error with time=1h.
The actual and predicted values of wind power are in high correlation. Error is significant to evaluate the prediction performance. Accuracy evaluation is used for WPP from MAPE, MSE and $\rho$:

$$\text{Error} : E_t = P_t - \hat{P}_t$$

$$\text{MAPE} = \left[ \frac{1}{n} \sum_{t=1}^{n} \left| \frac{P_t - \hat{P}_t}{P_t} \right| \right] \times 100$$

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^{n} (\hat{P}_t - P_t)^2$$

$$\rho = \frac{\sum_{t=1}^{n} (\hat{P}_t - \mu_{\hat{P}})(P_t - \mu_P)}{\sqrt{\sum_{t=1}^{n} (\hat{P}_t - \mu_{\hat{P}})^2 \sum_{t=1}^{n} (P_t - \mu_P)^2}}$$

Where, $P_t$ is the actual wind power at time $t$; $\hat{P}_t$ is the WPP value at time $t$; $n$ number of data points; $\mu_{\hat{P}} \& \mu_P$ represents mean values of $\hat{P}$ and $P$. $\rho (|\rho| \leq 1)$ is a measurement of linear correlation between two variables, a larger positive $\rho$ indicates that the prediction values are more correlated with actual ones.

![Wind Power Plant](Northeast China)

![Wind power simulation](One-minute wind power data)

![Single step prediction](Actual power - Predict power)

![Multi steps prediction](Actual power - Predict power)

![ARMA models](Actual - MSP - SSP - ARMA models)

![Comparison values](Comparison error value - Error PDF)

Figure 11. Simulation chart for WPP and error of wind farm with time window = 1h.

Table 2. WPP error and correlation coefficient comparisons of three methods with time window = 1h.

| Model | MAPE  | Improvement | MSE   | Improvement | $\rho$ |
|-------|-------|-------------|-------|-------------|-------|
| MSP   | 0.693 |             | 0.32  |             | 0.98  |
| SSP   | 2.451 | 71.73%      | 3.27  | 90.21%      | 0.80  |
| ARMA  | 2.468 | 71.92%      | 3.29  | 90.27%      | 0.79  |

Table 2 represents the effectiveness of the proposed hybrid method in WPP. The MAPE, MSE and $\rho$ are criteria of efficiency and the best prediction value is occurred while using the proposed hybrid method.
method. Both errors of MSP and SSP are improved slightly in the proposed method which confirms the accuracy of the proposed method. Moreover, figure 11 shows WPP and its error by using the hybrid model with MSP & SSP approaches and ARMA model. From error PDF graphics, it can be seen that the proposed hybrid method with MSP approach is better than other methods.

6. Conclusion
In this paper, we proposed a hybrid method to reduce error of VSTWPP based on recorded data. Simulation studies by using three methods are conducted. The results show that in terms of prediction accuracy (reduction of prediction error), the proposed hybrid model with MSP approach significantly outperformed of SSP approach and ARMA model. The results in table 1, 2 and figure 11 shows that the proposed hybrid model can reduce error from views of MAPE, MSE, $\rho$ and normal PDF. The WPP error was uniformly distributed with positive and negative error values.

For time-series data, the hybrid method is a technique with the high response. In addition, errors are less than 0.6%; the shape of the predicted curve is smooth and less fluctuating, and the normal probability distribution shown in the shapes above is the best-fit shape.

It has been proved that the proposed hybrid method with MSP approach reduces error comparing to SSP approach and ARMA model. Moreover, the proposed hybrid model with MSP approach shows best-fit curves symmetrical around at smallest error value. Generally, the statistical error analysis shows that WPP error PDF by MSP approach is better than other methods used throughout this paper. This finding reveals the advantage of hybrid method with MSP approach from a probabilistic approach.

Acknowledgment
This work is supported by the National Natural Science Foundation of China under Grant 51377035.

References
[1] Hu J, Wang Jand Zeng G 2013 J. Renewable Energy. 60 185-194
[2] Sorensen P, Cutululis N A, Vigueras-Rodriguez A, Jensen L E, Hjerrild J, Donovan M Hand Madsen H 2007. IEEE Trans. Power Syst. 22 958-965
[3] Han C, Huang A Q, Baran M E, Bhattacharya S, Litzenberger W, Anderson L. ... and Edris A A 2008. IEEE Trans. Energy conversion. 23 226-233
[4] Spahic Eand Balzer G 2005. IEEE International Conference. F P S. pp. 6-pp
[5] JIANG Y, Xingying C, Kun Y and Yingchen L 2015 J. Mod. Power Syst. Clean Energy. 1-8
[6] Chang W Y 2013 J. Energies. 6 4879-4896
[7] Yang Y, Wu J, Chen Yand Li C 2013 J. ABSTR APPL ANAL. 20139 pages
[8] Skolthanarat S, Lewlomphaisarl Uand Tungpimolrut K 2014. IEEE Conference on SusTech. 165-171
[9] He D and Liu R 2012. IEEE 2012 7th International in IPEMC. 32108-2112
[10] Jyothi M N and Rao P R 2016. IEEE Conference on PESTSE. 1-6
[11] Sheikh S K and Unde M G 2012 J. IJETT.1 97-107
[12] Adepoju G, Ogunjuyigbe S and Alawode K 2007 J. PJST. 8 68-72
[13] Quaiyum S, Khan Y I, Rahman S and Barman P 2011 J. IJCA. 30 1-7
[14] Amral N, Ozveren C Sand King D 2007. IEEE UPEC 2007.1192-1198
[15] Botterud A, Wang J, Miranda Vand Bessa R J 2010 J. Electricity Journal. 23 71-82
[16] Wang C, Lu Z and Qiao Y. 2009. IEEE Conference on SUPERGEN09. 1-6
[17] Gomes Pand Castro R. 2012 J. IJSED.1 36-45
[18] Rothe M, Wadhwani D Aand Wadhwani D 2009 J. IJCSIS.6 303-306
[19] Jiang X, Dong B, Xie Land Sweeney L 2010 J. In ECAI. 661-666
[20] Daneshi Hand Daneshi A2008. IEEEThird International Conference on DRPT. 689-695
[21] Hodge B Mand Milligan M2011. IEEE PES,general meeting. 1-8