

Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm

Yu JIANG¹ , Xingying CHEN¹, Kun YU¹, Yingchen LIAO¹



Abstract Day-ahead wind power forecasting plays an essential role in the safe and economic use of wind energy, the comprehending- intrinsic complexity of the behavior of wind is considered as the main challenge faced in improving forecasting accuracy. To improve forecasting accuracy, this paper focuses on two aspects: ①proposing a novel hybrid method using Boosting algorithm and a multi-step forecast approach to improve the forecasting capacity of traditional ARMA model; ②calculating the existing error bounds of the proposed method. To validate the effectiveness of the novel hybrid method, one-year period of real data are used for test, which were collected from three operating wind farms in the east coast of Jiangsu Province, China. Meanwhile conventional ARMA model and persistence model are both used as benchmarks with which the proposed method is compared. Test results show that the proposed method achieves a more accurate forecast.

Keywords Hybrid method, Multi-step-ahead prediction, Wind power forecast, Boosting algorithm, Time series model

1 Introduction

In the context of increasing power demand as well as the energy and environmental issues, wind energy has been growing rapidly around the world [1–4]. However, the intermittent behavior of wind poses a great challenge to increase the wind energy penetration. Thus wind power forecasting technique has been extensively studied in order to reduce the impact of wind intermittency on the power system [3, 4].

Generally, wind power forecasting techniques can be classified into three groups, numeric weather prediction (NWP) methods, statistical methods, and hybrid approaches [3, 5]. NWP models are based on complex mathematical models, thus have advantages in prediction accuracy, especially for longer horizons. However, it is difficult to develop an accurate mathematical model without the knowledge of aerodynamics and atmosphere physics, and the model calculation is time-consuming and requires super computers to get solutions. The statistical methods such as autoregressive moving average (ARMA) and artificial neural network (ANN) provide a way to predict the wind power using only the data, thus they are of special interests for a large number of engineering applications [6]. However, the prediction capability of statistical methods drops as the forecast horizon grows. In addition, support vector regression-based methods [7, 8] and generalized locally weighted group method of data handling (GMDH) [9] were also proposed in recent years. However, the existing forecasting techniques still cannot adequately meet the engineering requirements.

CrossCheck date: 23 June 2015

Received: 21 July 2014 / Accepted: 23 June 2015 / Published online: 30 November 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

✉ Xingying CHEN
xychen@hhu.edu.cn

Yu JIANG
jiangyu_nanjing@126.com

Kun YU
kun.yu@vip.sina.com

Yingchen LIAO
yingchenliao2@163.com

¹ College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China

This paper focuses on improving the problem of forecast accuracy drops as time horizon grows, when time series model is used for forecasting. It embraces two aspects: first, a horse racing principle is introduced, which assumed a gambler named A, who is frustrated by persistent horse-racing losses and envious of his friends' winnings decides to make some positive changes [10]. An easy and efficient way to help him is to allocate his wager in accordance with the luckiest fellows to bet, based on the previous performances of his friends. Boosting theory exactly focuses on the problem how a set of weak capacity forecasts can create a single accurate forecasting result by a weighted vote process. It was first introduced by Michael Kearns and Leslie Valiant [11–13]. Boosting method also has many engineering advantages [14, 15], which have been tested empirically by many researchers, for example, it is fast, simple and easy to program. Supposing those gamblers can be replaced by forecasting models, the boosting model can be used to improve forecasting accuracy of existing models. On the other side, the multi-step-ahead (MS) technique was indicated to improve the issue of forecasting accuracy fades rapidly, and it has a wide application in economic time series forecasts [16]. Therefore, at the first step, the MS technique is applied to build a base model using an ARMA model for forecasts. At the second step, furthermore, the base model is used to construct the proposed method by boosting algorithm in Section 3. In this study it is called ARMA based MS base forecasting model (ARMA-MS), and more details for MS approach and its parameter identification can be found in [16].

Based on the above derivation, this paper proposes a methodology of applying boosting method and MS technique to significantly improve the accuracy of time series forecasting model for day-ahead wind power forecasts. In the first part, the advantages of boosting method for wind power forecast are generalized. Then, a novel method is proposed which combining boosting algorithm with ARMA-MS model. The procedure of the proposed modeling method can be summarized as follows: ① combining multi-step approach and ARMA model to build basic forecasting model (weak forecasting capacity); ② using boosting algorithm to combine these weak forecasting models into an accurate assembled model H by a weighted vote process; ③ applying the model H to forecast the wind power output of forecast date τ .

In the second part, based on the results of [10], this paper theoretically proves the existence of error bounds of the proposed forecast method. For instance, one of the results shows that the forecasting error of the proposed algorithm can be bounded by

$$\left(\frac{2}{T}\right) \frac{\sqrt{\varepsilon_1(1-\varepsilon_1)}}{1-q_1} + \frac{\varepsilon_1}{T} \quad (1)$$

where T is the number of iteration times; ε_1 is forecast error of the first round; q_1 is an constant defined in Section 3.

Finally, in order to test the proposed method, hourly wind power output data of one year from three operating wind farms in the east coast of Jiangsu Province of China is used for analysis. The results indicate that the proposed algorithm can improve the wind power forecast accuracy with respect to traditional ARMA model and persistence model (PM). Econometric views (EViews) is a statistical package for Windows operating systems. It is used for time-series oriented econometric analysis, general statistical analysis, and wind power forecast. In this paper, EViews (Version 8) is used as the analysis tool to get simulation results.

Structure of this paper is organized as follows: Section 2 reviews the main aspects of boosting techniques. Section 3 proposes a novel forecasting method and analyzes the calculating error bounds. Section 4 shows simulation results to evaluate the performances of the proposed method, and presents discussions. Section 5 highlights final remarks for concluding the paper.

2 Related algorithms

According to [17], boosting algorithm has its roots in a theoretical framework for studying machine learning called probably approximately correct (PAC) learning model [18], regarding learning as a phenomenon of knowledge acquisition in the absence of explicit programming. After that, whether a “weak” learning algorithm which performs just slightly better than random guessing using the PAC learning algorithm can be “boosted” into an arbitrarily accurate “strong” learning algorithm was analyzed in [11]. An efficient boosting algorithm named adaptive boost (AdaBoost) was proposed to solve many of the practical difficulties of earlier boosting algorithms [10]. Genetic programming boost (GPBoost) algorithm was proposed based on genetic programming (GP) [19]. After that, based on the GPBoost method, boosting algorithm using a correlation coefficients (BCC) was proposed to update the weights of calculation and improve accuracy of prediction [14, 20].

2.1 Conventional boosting algorithm

Boosting is traditionally considered as a general technique of combining rules of thumb, or weak classifiers, to form highly accurate combined classifiers. References [10, 13] presented that a class C is existing learnable concepts, and there exists a class of estimated models h such that for all $n \geq 1$, an actual concept $c \in C$, for all distributions D on

X , and $0 < \gamma, \delta \leq 1$, given parameters δ, n, γ , and the size s of the target concept c . Using running time polynomial in $1/\delta, n, 1/\gamma$ and s , boosting algorithm outputs an accurate forecast model H that with probability at least $1 - \delta$ is γ -close to c under D . As defined the class \mathbf{h} required having a prediction error slightly better than $1/2$ with respect to the distribution D on which it is trained. The boosting procedure is as follows: ① the learner receives M examples $\{(x_1, y_1), \dots, (x_M, y_M)\}$ chosen according to the distribution D on $X \times Y$, where Y is an actual data set of the forecasting target associated with training-example set X ; ② on each round $t = 1, \dots, T$, the booster devises a distribution D_t over the set of examples, and requests a rule-of-thumb $h_t \in \mathbf{h}$ with low error ε_t with respect to D_t ; ③ after T rounds, the booster combines the original weak forecast algorithm into a single strongly accurate forecast algorithm H .

2.2 ARMA model

ARMA model is generally one of the most widely used approaches for forecast. ARMA models can effectively be used to predict the behavior of a time series from past values alone. In [21], ARMA model improves significantly the wind speed forecasts as compared to those obtained with persistence models, for forecasting one hour in advance or even ten hours in advance.

The notation ARMA (p, q) refers to the model with p autoregressive terms and q moving-average terms. On the other hand, by seasonal variations the parameters of the ARMA model cannot simply be accepted as constant. Generally, ARMA models incorporated into prediction requiring three main steps, identification, estimation and diagnostic check.

In the step of model estimation, it is indicated that AR(1) and AR(2) are both appropriate models for wind power forecast [22]. Reference [23] indicated that ARMA(2, 3) is a suitable model. In this paper, tools of the sample autocorrelation function (SACF) and the sample partial autocorrelation function (SPACF) are both used to identify the parameters (p, q) of the ARMA model [6]. In this paper, based on the analysis of the one-year period wind power data, the ARMA(1, 1) model is more favorable than others using Akaike information criterion (AIC), and this simplest model was also used in [24].

2.3 Persistence model

The persistence model is considered as the simplest time series models. However, it surpasses many other models in very short-term prediction. Therefore, the PM is used as a benchmark. The persistence forecast can be written as [21]:

$$\hat{P}(t+k|t) = \frac{1}{\hat{T}} \sum_{i=0}^{n-1} P(t-i\Delta t) \quad (2)$$

where $\hat{P}(t+k|t)$ is the wind power forecast for time $t+k$ made at time origin t ; k is the prediction horizon; \hat{T} is the prediction interval length ($\hat{T} = k$); $P(t-i\Delta t)$ is the measured wind power for time t and the previous i time steps within \hat{T} ; n is the number of time steps within \hat{T} ; Δt is the time step length of the measured time series ($\hat{T} = n\Delta t$).

3 Boosting based hybrid method and its analysis

3.1 Structure of hybrid method

For more than half a century, as a well-known time series technique ARMA models have been widely applied in the construction of accurate hybrid models and the engineering forecasts of wind power. Therefore, based on the advantages of boosting algorithm, a hybrid forecasting method that combines boosting algorithm with the ARMA-MS model is proposed to improve the forecast accuracy.

The generation process of the final model H which is the output model for wind power forecast via T boosting iterations. It shows that forecast model H is a weighted vote of T weak models $h \in \mathbf{h}$ where α is the weight assigned to each. Intuitively, α_t measures the importance that is assigned to h_t , and that α_t gets larger as ε_t gets smaller as shown in (6).

The proposed method includes four steps as shown in Fig. 1.

Step 1: Define the forecast date τ and let parameter “training target date” equal $\tau - 1$ (measured wind power Y_0), considering the existence of wind speed persistence. Define the parameter T the number of iteration rounds and the training set for each iterative forecast.

Step 2: Apply ARMA-MS model h_t to forecast wind power of date $\tau - 1$ based on the training date $\tau - (t + 1)$. Calculate the forecasting error ε_t of model h_t using (4), and the weight value α_t of model h_t using (6). Due to h_t is a week capacity forecasting model, its forecast error ε_t should be less than $1/2$ defined by Freund and Schapire [10]. In other words, each generated forecasting model h with error larger than $1/2$ is ignored.

Step 3: After T round forecasts, boosting algorithm combine these weak forecasting models into an accurate assembled model H by a weighted vote process as shown in (8).

Step 4: Finally, the model $H(X_0)$ is used to forecast the wind power output of the final forecast date τ .

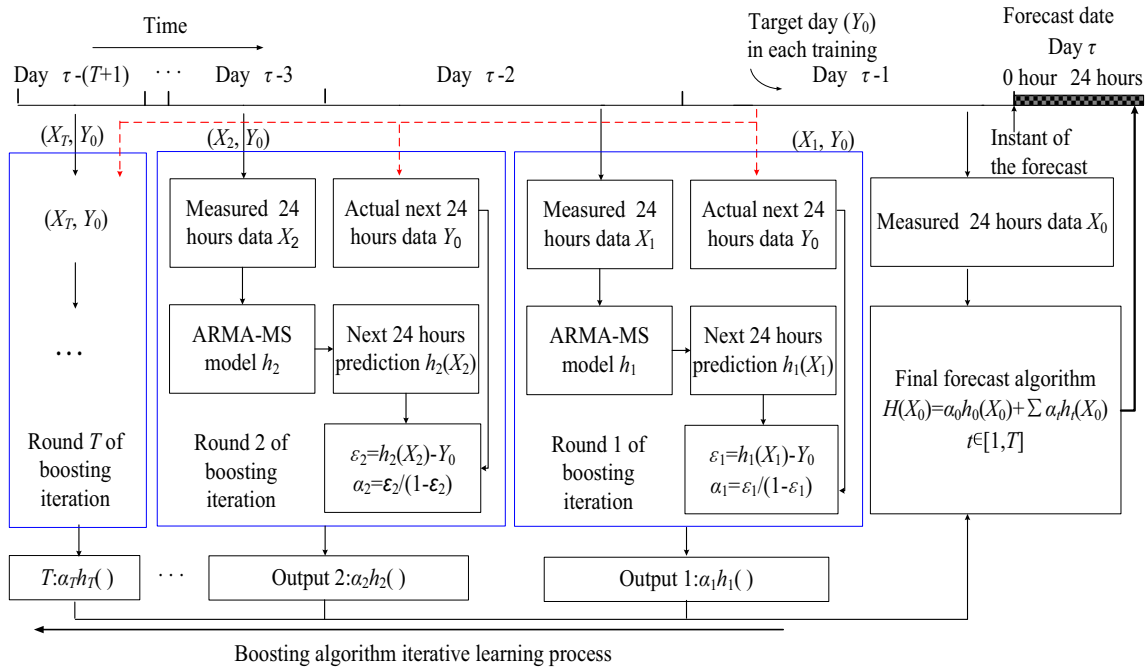


Fig. 1 Logic structure of proposed method for wind power forecast

3.2 Forecasting process

The calculation process and pseudo-code of the new method for wind power prediction is shown below: Given $(X_1, Y_0), \dots, (X_T, Y_0)$, where Y_0 is the actually measured wind power, $(X_t, Y_0) = \{(x_t^1, y_1^0), \dots, (x_t^M, y_M^0)\}$ where $t \in [1, T], i \in [1, M]$, and M is the number of wind farms.

Initialize: $w^1 = \{w_1^1, \dots, w_M^1\} = \{1, \dots, 1\}$.

Do for $t = 1, 2, \dots, T$

Set ARMA-MS model as

$$D^t = \frac{w^t}{\sum_{i=1}^M w_i^t} \tag{3}$$

Choosing forecasting hypothesis of wind power $h_t : X \rightarrow R$ with the calculated error.

$$\varepsilon_t = D^t \ell^t = \frac{\sum_{i=1}^M D^t(i) |h_t(x_i) - y_i|}{P_{Cap}} \tag{4}$$

where $h_t(x_i^t) \neq y_i^0$, $\varepsilon_t \in (0, 1/2)$ guaranteed by forecast-accuracy selection, and P_{Cap} is the installed capacity.

$$\text{Set } \beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t} \quad \beta_t \in (0, 1) \tag{5}$$

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right) \quad \alpha_t \in [0, +\infty) \tag{6}$$

Set the new weights vector to be

$$w_i^{t+1} = w_i^t \beta_t^{1 - \ell_i^t} = w_i^t \beta_t^{1 - \left[\frac{|h_t(x_i^t) - y_i^0|}{P_{Cap}} \right]} \tag{7}$$

Output the final forecasting algorithm as

$$H(X_0) = \alpha_0 h_0(X_0) + \frac{\sum_{t=1}^T \alpha_t h_t(X_0)}{\sum_{t=1}^T \alpha_t} \tag{8}$$

In the first step, it is assumed that the sequence of M training examples $\{(x_1^1, y_1^0), \dots, (x_M^1, y_M^0)\}$ is drawn from $X \times Y$ according to distribution D . The value of $h(x_i)$ is the forecasted result on x_i . $Y_0 = \{y_1^0, \dots, y_M^0\}$ is the actually measured wind power in the target date $\tau - 1$ from M wind farms, which is used for model check in the whole training process. Then a boosting process is started to find out the hypothesis H which is consistent with most of the forecasting sub-models.

After each iteration t , h_t suffers forecasting accuracy loss $\ell_i^t \in [0, 1]$ named loss function in (4), which can be expressed as

$$\ell_i^t = \frac{|h_t(x_i) - y_0|}{P_{Cap}} \tag{9}$$

The loss suffered from D can be written in the form $\sum_{i=1}^M D_i \ell_i = D \ell$, and the average loss of the forecasting algorithm h_t with respect to D^t is shown below.

$$D^t \ell^t \tag{10}$$

As described in Fig. 1, the objective of the novel method is to find out a final forecasting model H for the next day, which has a higher close-to the actual and most accurate model c defined in Section 2, which cannot be known in advance.



3.3 Forecasting error bounds analysis

According to [25], a number of bounds on forecasting errors are theoretically proven in this section. For instance, one of the results shows that the forecasting error of the new algorithm can be bounded by (11).

Theorem 1: Suppose the weak learning algorithm $h \in \mathbf{h}$, when called by boosting algorithm, generates hypothetical forecast models with error $\varepsilon_1, \dots, \varepsilon_T$. Then the error $\varepsilon = Pr_{i \sim D}[h(x_i) \neq y_i]$ of the output of the final hypothesis H is bounded by

$$\varepsilon \leq \left(\frac{2}{T}\right) \frac{\sqrt{\varepsilon_1(1-\varepsilon_1)}}{1-q_1} + \frac{\varepsilon_1}{T} \tag{11}$$

where $q_1 = \max_{t \in [1, T]} 2\sqrt{\varepsilon_t(1-\varepsilon_t)}$, $0 < q_1 < 1$.

Proof: Use w_i^t defined in (7), and for simplicity, the initial distribution D is set to be uniform so that $D(i) = 1/M$. Using a result of the error $\varepsilon = Pr_{i \sim D}[h(x_i) \neq y_i]$ for binary forecast problem, which has been proved in [10], we get that

$$\sum_{i=1}^M w_i^{T+1} \leq \prod_{t=1}^T [1 - (1 - \varepsilon_t)(1 - \beta_t)] \tag{12}$$

and

$$\begin{aligned} \sum_{i=1}^M w_i^{T+1} &\geq \sum_{i: h_{T+1}(x_i) \neq y_i} w_i^{T+1} \\ &\geq \left(\sum_{i: h_{T+1}(x_i) \neq y_i} D(i)\right) \left(\prod_{t=1}^T \beta_t\right)^{\frac{1}{2}} \end{aligned} \tag{13}$$

where β_t is defined in (5).

Clearly, in this paper

$$\varepsilon_{T+1} = \sum_{i: h_{T+1}(x_i) \neq y_i} \frac{D(i)|h_{T+1}(x_i) - y_i|}{P_{\text{Cap}}} \leq \sum_{i: h_{T+1}(x_i) \neq y_i} D(i) \tag{14}$$

since $\frac{|h_{T+1}(x_i) - y_i|}{P_{\text{Cap}}} \in (0, 1]$, combining (13) and (14) we get

$$\sum_{i=1}^M w_i^{T+1} \geq \varepsilon_{T+1} \left(\prod_{t=1}^T \beta_t\right)^{\frac{1}{2}} \tag{15}$$

where ε_{T+1} is the error of h_{T+1} .

Combining (12) and (15) we get

$$\varepsilon_{T+1} \leq \prod_{t=1}^T \frac{1 - (1 - \varepsilon_t)(1 - \beta_t)}{\sqrt{\beta_t}} \tag{16}$$

According to (5), we get

$$\varepsilon_{T+1} \leq \prod_{t=1}^T 2\sqrt{\varepsilon_t(1-\varepsilon_t)} = 2^T \prod_{t=1}^T \sqrt{\varepsilon_t(1-\varepsilon_t)} \tag{17}$$

The form of error ε is written as

$$\varepsilon = \frac{\sum_{t=1}^T \alpha_t \varepsilon_t}{\sum_{t=1}^T \alpha_t} \leq \frac{1}{T} \sum_{t=1}^T \varepsilon_t \tag{18}$$

where α_t is a convex function on ε_t as defined in (6).

According to the definition in (8) and combining (17) and (18), the bound on the error of H in (8) can be written as

$$\begin{aligned} \varepsilon &\leq \left(\frac{1}{T}\right) \sum_{i=2}^T \left(2^{i-1} \prod_{t=1}^{i-1} \sqrt{\varepsilon_t(1-\varepsilon_t)}\right) + \frac{\varepsilon_1}{T} \\ &= \left(\frac{1}{T}\right) \sum_{i=2}^T \left(\prod_{t=1}^{i-1} 2\sqrt{\varepsilon_t(1-\varepsilon_t)}\right) + \frac{\varepsilon_1}{T} \\ &\leq \left(\frac{1}{T}\right) \sum_{i=2}^T \left(2\sqrt{\varepsilon_1(1-\varepsilon_1)} q_1^{i-2}\right) + \frac{\varepsilon_1}{T} \end{aligned} \tag{19}$$

where $q_1 = \max_{t \in [1, T]} \{2\sqrt{\varepsilon_t(1-\varepsilon_t)}\}$, $0 < q_1 < 1$.

The bound can also be written as

$$\begin{aligned} \varepsilon &\leq \left(\frac{2}{T}\right) \frac{\sqrt{\varepsilon_1(1-\varepsilon_1)}(1-q_1^{T-1})}{1-q_1} \\ &\quad + \frac{\varepsilon_1}{T} \leq \left(\frac{2}{T}\right) \frac{\sqrt{\varepsilon_1(1-\varepsilon_1)}}{1-q_1} + \frac{\varepsilon_1}{T} \end{aligned} \tag{20}$$

Then, the proof is completed.

Through iterative selection the accurate forecast model h is selected by (4), who's forecast error is less than 1/2. In other words, the estimated forecast models, whose forecast errors are lower than 1/2, are ignored. Only the high accurate models established in iterative learning process are used to the final forecast according to (8). According to the above derivation, the final forecast accuracy is theoretically guaranteed by the bound-control capability of the boosting algorithm.

4 Simulation results

4.1 Simulation data and settings

This study focuses on the forecast accuracy of aggregated wind power in the Nantong-Yancheng regional power grid. It is a region that features well-developed wind power energy and heavily industrial load. Therefore, this study uses hourly real wind power data from this power grid to test proposed method. The three wind farms have a total installed power capacity of 701.3 MW. They are located in the east coast of Jiangsu Province, China: Dong Yuan (100.5 MW), Long Yuan (400.5 MW), Da Feng (200.3 MW). The study data covers a period ranging from May 2012 to August 2013. Data were continuously acquired over this period with the only unavailability occurred for 85 days (from November 9, 2012 to January 20, 2013, and from April 26, 2013 to May 7, 2013) due to continuous faults of data acquisition system. The availability of wind power output data is 80%.

Two important issues are considered in designing and conducting the experiment to test the proposed method: ① how well does a model retain accuracy in its time horizons; ② how robust is the algorithm to the choice of test set. To address these problems two cases with different time horizons are collected for test: the short-term data of one-month period (July 2012) is mainly used for principle description, and the long-term data of fourteen-month period is used for the capability validation and an economic analysis.

Several literatures have proposed various indices to evaluate the efficiency of predictive methods [5]. Specially, mean absolute error (MAE), normalized average absolute energy production error (NMAE) of (21) and root mean-square error (RMSE) were widely accepted metrics used by the statistical community [14]. In this paper, for the purpose of comparing the prediction accuracy of the ARMA model, the persistence model, and the proposed boosting method, the indices of MAE, NMAE and RMSE are all used to study the absolute errors between the measured values and the forecasts.

$$100 \times \frac{|P_{\text{real}} - P_{\text{forecast}}|}{P_{\text{Cap}}} \quad (21)$$

where P_{real} is the measured wind power data; P_{forecast} is the forecast result.

4.2 Case 1

A preliminary study has been carried out using EViews (Version 8) software as an analysis tool. According to the

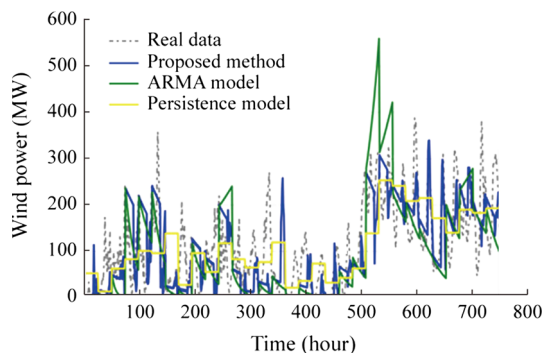


Fig. 2 Comparison of forecasting results

Table 1 Forecast results analysis

Method	MAE (MW)	NMAE (%)	Improvement (%)	RMSE (MW)	Improvement (%)
Proposed method	57.09	8.14		51.89	
ARMA model	67.69	9.66	15.73	61.41	15.50
Persistence model	62.70	8.94	8.95	53.61	3.21

validation set measured in July 2012 and the training set of June, Fig. 2 presents the comparison of forecasting results of the ARMA model and the proposed novel method. Parameter T equals 30 (the length of training set of June), and M equals 3 (Dong Yuan, Long Yuan, and Da Feng wind farms).

From Fig. 2 it can be clearly observed that the proposed method is advantageous for wind power forecasts due to its capability of quick convergence of forecasting errors as described in the Theorem 1 of Section 3. It explicitly shows that only in very few cases the accuracies of the proposed method are lower than the ones obtained by the normal ARMA model. The test results show that the index NMAE of the proposed method is 8%, the normal ARMA is 9.94%, and the persistence model is 9.09%. The accuracy improvement of the proposed method is 19.52% compared with the normal ARMA model, and 11.99% compared with the PM model, 24 hours in advance. It can be intuitively seen that the proposed forecasting method can effectively improve the forecasting accuracy than that uses only ARMA and PM. Especially, due to the application of MS approach, the proposed method (blue line) tracks the real data better than the benchmark models (green & yellow lines) in the tail of each forecasting series, as shown in Fig. 2.

4.3 Case 2

Table 1 shows that compared with traditional ARMA and PM models, the indices of NMAE, MAE and RMSE of the proposed method have much better performance, based on one-year period of real data collected from three operating wind farms in the east coast of Jiangsu Province, China. The indices have an improvement ranging from 3.21% to 15.73%.

5 Conclusions

This paper proposes a novel hybrid method using the boosting algorithm to boost the forecasting capability of ARMA models. Comparing with traditional time series forecasting models' poor accuracy as forecasting time



horizon grows, the proposed method improves this limitation through combining the boosting model and the MS technique. Especially, the forecasting accuracy of the proposed method is theoretically guaranteed by the error bound deduced out in this paper. To validate the accuracy of the proposed method, a real data collected from operating wind farms and covering fourteen months is used in experiments. Simulation results show that from point view of the indices of MAE, NMAE and RMS, the proposed hybrid method is more accurate and more efficient than those of traditional ARMA model and persistence model.

In the future work, the proposed novel method will focus on performing accuracy validation compared with more existing approaches, such as wavelet transform and fuzzy ARTMAP networks combined forecast approach, wavelet-ARIMA forecast, hybrid Kalman filters forecast method and other hybrid algorithms.

Acknowledgment This work was supported by the National High Technology Research and Development of China (863 Program) (No. 2012AA050214), the National Natural Science Foundation of China (No. 51077043) and the State Grid Corporation of China (Impact research of source-grid-load interaction on operation and control of future power system). We would like to thank the Jiangsu Electric Power Dispatching and Control Center, the State Grid Corporation of China, for providing the past fourteen-month historical wind power output data.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- [1] Greenwood C, Hohler A, Hunt G et al (2007) Global trends in sustainable energy investment 2007: analysis of trends and issues in the financing of renewable energy and energy efficiency in OECD and developing countries. DTI/0985/PA, United Nations Environment Programme and New Energy Finance Ltd, Nairobi, Kenya
- [2] Ridge R (2012) 2011 wind technologies market report. DOE/GO-102012-347, Energy Efficiency and Renewable Energy, US Department of Energy (DOE), Washington, DC, USA
- [3] Wu YK, Hong JS (2007) A literature review of wind forecasting technology in the world. In: Proceedings of the 2007 IEEE Lausanne power tech conference, Lausanne, Switzerland, 1–5 Jul 2007, pp 504–509
- [4] Kehler J, Hu M, McMullen M et al (2010) ISO perspective and experience with integrating wind power forecasts into operations. In: Proceedings of the 2010 IEEE power and energy society general meeting, Minneapolis, MN, USA, 25–29 Jul 2010, 5 pp
- [5] Bilgili M, Sahin B (2010) Comparative analysis of regression and artificial neural network models for wind speed prediction. *Meteorol Atmos Phys* 109(1/2):61–72
- [6] Lojowska A, Kurowicka D, Papaefthymiou G et al (2010) Advantages of ARMA-GARCH wind speed time series modeling. In: Proceedings of the IEEE 11th international conference on probabilistic methods applied to power systems (PMAPS'11), Singapore, 14–17 Jun 2010, pp 83–88
- [7] Zhu L, Wu QH, Li MS et al (2013) Support vector regression-based short-term wind power prediction with false neighbours filtered. In: Proceedings of the 2013 international conference on renewable energy research and applications (ICRERA'13), Madrid, Spain, 20–23 Oct 2013, pp 740–744
- [8] Elattar EE, Goulermas J, Wu QH (2010) Electric load forecasting based on locally weighted support vector regression. *IEEE Trans Syst Man Cybern C* 40(4):438–447
- [9] Elattar EE, Goulermas J, Wu QH (2012) Generalized locally weighted GMDH for short term load forecasting. *IEEE Trans Syst Man Cybern* 42(3):345–356
- [10] Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci* 55(1):119–139
- [11] Kearns M (1988) Thoughts on hypothesis boosting. Unpublished manuscript 45:105
- [12] Kearns M, Valiant L (1994) Cryptographic limitations on learning Boolean formulae and finite automata. *J ACM* 41(1):67–95
- [13] Schapire RE (1990) The strength of weak learnability. *Mach Learn* 5:197–227
- [14] de Souza LV, Pozo ATR, da Rosa JMC et al (2007) The boosting technique using correlation coefficient to improve time series forecasting accuracy. In: Proceedings of the 2007 IEEE congress on evolutionary computation (CEC'07), Singapore, 25–28 Sept 2007, pp 1288–1295
- [15] de Souza LV, Pozo ATR, Neto AC (2006) Using correlation to improve boosting technique: an application for time series forecasting. In: Proceedings of the IEEE 18th international conference on tools with artificial intelligence (ICTAI'06), Arlington, VA, USA, 13–15 Nov 2006, pp 26–32
- [16] Chevillon G, Hendry DF (2005) Non-parametric direct multi-step estimation for forecasting economic processes. *Int J Forecast* 21(2):201–218
- [17] Freund Y, Schapire RE (1999) A short introduction to boosting. *J Jpn Soc Artif Intell* 14(5):771–780
- [18] Valiant LG (1984) A theory of the learnable. *Commun ACM* 27(11):1134–1142
- [19] Paris G, Robilliard D, Fonlupt C (2001) Applying boosting techniques to genetic programming. In: Artificial evolution: proceedings of the 5th international conference on evolution artificielle (EA'01), Le Creusot, France, 29–31 Oct 2001. LNCS, vol. 2310, pp 267–278
- [20] de Souza LV, Pozo A, da Rosa JMC et al (2010) Applying correlation to enhance boosting technique using genetic programming as base learner. *Appl Intell* 33(3):291–301
- [21] Torres JL, García A, de Blas M et al (2005) Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Sol Energy* 79(1):65–77
- [22] Brown BG, Katz RW, Murphy AH (1984) Times series models to simulate and forecast wind speed and wind power. *J Clim Appl Meteorol* 23(8):1184–1195
- [23] Kamal L, Jafri YZ (1997) Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. *Sol Energy* 61(1):23–32
- [24] Gnana Sheela K (2011) Computing models for wind speed prediction in renewable energy systems. *IJCA Special Issue on*

Computational Science—New Dimensions & Perspectives, vol 3, pp 108–111

[25] Freund Y (1995) Boosting a weak learning algorithm by majority. *Inf Comput* 121(2):256–285

Yu JIANG was born in Sichuan, China. He received B.Eng. and M.Eng. degrees from Nanjing University of Science & Technology, China, in 2003 and 2009, respectively. Currently he is pursuing his Ph.D degree in energy and electrical engineering in Hohai University, Nanjing, China. His major fields of study are operation analysis and control of power systems, time series analysis and forecasts.

Xingying CHEN was born in Wuxi, China. She received B.Eng. and Ph.D degrees from Southeast University, China, in 1984 and 2002, respectively, and an M.Eng. degree from Hohai University, China, in 1995, all in electrical engineering. She joined the faculty of Hohai University, China, in 1984, and has become a professor there since 2002. Her current research interests include operation analysis and

control of power systems, distribution automation and advanced application analysis, and power marketing and economics.

Kun YU was born in Chongqing, China. He received B.Eng. and M.Eng. degrees from Hohai University, China, in 2000 and 2003, respectively, and Ph.D degree from Zhejiang University, China, in 2011, all in electrical engineering. Since 2003 he has become a lecturer in Hohai University. His current interests include operation analysis and control of power systems, distribution automation and advanced application analysis.

Yingchen LIAO received her B.Eng degree from Hohai University, Nanjing, China, in 2000. Currently she is a Ph.D candidate in electrical engineering in Zhejiang University, Hangzhou, China, and a lecture in Hohai University. Her major fields of study are Electricity Market and Demand Side Management.

