Multi-step-ahead Method for Wind Speed Prediction Correction Based on Numerical Weather Prediction and Historical Measurement Data

Han Wang¹, Jie Yan¹*, Yongqian Liu¹, Shuang Han¹, Li Li¹, Jing Zhao²
School of Renewable Energy, North China Electric Power University, Beijing, 102206, China ¹
Beijing Shenzhou Aerospace Software Technology Co., Ltd., Beijing, 100094, China ²
E-mail: yanjie_freda@163.com

Abstract. Increasing the accuracy of wind speed prediction lays solid foundation to the reliability of wind power forecasting. Most traditional correction methods for wind speed prediction establish the mapping relationship between wind speed of the numerical weather prediction (NWP) and the historical measurement data (HMD) at the corresponding time slot, which is free of time-dependent impacts of wind speed time series. In this paper, a multi-step-ahead wind speed prediction correction method is proposed with consideration of the passing effects from wind speed at the previous time slot. To this end, the proposed method employs both NWP and HMD as model inputs and the training labels. First, the probabilistic analysis of the NWP deviation for different wind speed bins is calculated to illustrate the inadequacy of the traditional time-independent mapping strategy. Then, support vector machine (SVM) is utilized as example to implement the proposed mapping strategy and to establish the correction model for all the wind speed bins. One Chinese wind farm in northern part of China is taken as example to validate the proposed method. Three benchmark methods of wind speed prediction are used to compare the performance. The results show that the proposed model has the best performance under different time horizons.

Key words. Wind speed prediction, multi-step-ahead, prediction correction, time-dependent effects, support vector machine.

1. Introduction
As an important renewable energy resource, wind power has been increasing dramatically in the power system worldwide. However, the intermittence of wind brings serious challenges to the security and stability of the power system operation as well as the power quality. Cubic relationship between wind speed and power output means that even tiny wind speed prediction error would trigger very large wind power forecasting deviation. To improve the wind speed prediction accuracy on both short-term and ultra-short-term is one of the effective ways to reduce the negative impact of wind power integration on the power system.

Previous literatures on wind speed prediction can be divided into two categories. The first category is based on historical wind speed measurements to make ultra-short-term prediction, primarily 1-4h. Its main idea is to mapping the internal relationship between wind speed time series.
The commonly used algorithms include ARMA [6-7], Kalman filter [8], gray theory [9], etc. However, the accuracy of such methods falls sharply with the increase of prediction time horizon since the impact of previous wind series on the following wind condition is weakened. Another category employs the Weather Research and Forecasting (WRF) mesoscale NWP wind speed [10]. The initial and the boundary field of the WRF model is released by the National Centers for Environmental Prediction at 6 o’clock every day. The WRF model is adopted to extract the time-serial results from 24 o’clock a day to 24 o’clock the next day as the input data for the power forecasting. The accuracy of NWP data suffers relatively small impacts from the prediction horizons, however, the deviation of NWP is not negligible and contributes largely to the wind power forecasting error. Therefore, the improvement and correction of NWP data in either physical or statistical manner is significant to serve well for the wind power forecasting.

Many researchers are working on the statistical correction of wind speed prediction, and traditionally to establish a mapping relationship between NWP wind speed and the measured wind speed at the same time slot. Commonly used algorithms include neural network [11], SVM [12-13], relevance vector machine [14], and deep learning [15], etc. These methods achieve the accuracy improvement by around 1-2m/s on average, and is slightly affected by the prediction time horizons. However, this mapping relationship, which can be interpreted as NWP deviation pattern, is not always very clear and consequently affects the correcting performance.

To solve above problem, this paper presents a multi-step-ahead wind speed prediction correction method considering the impacts passing from the wind speed in the latest time series. Both numerical weather prediction and historical measurement data are used as training inputs. Additional inputs encourage the NWP deviation pattern to be differentiated under various wind conditions during the training process and improve the correction performance. The results show that the proposed method outperforms the other three benchmark methods at different time horizons, and the accuracy degradation along with the time horizons is reduced to some extent.

The remainder of this paper is organized as follows. Section 2 analyses the probability density of NWP deviation. Section 3 describes the proposed wind speed prediction correction method construction. Section 4 discusses the prediction results among the proposed method and three benchmark methods. Section 5 concludes the paper.

2. Probability Density of NWP Deviation
Numerical weather prediction is the basic inputs of wind power forecasting and is also the main source of the power forecasting error. This section takes one wind farm in north China as example to analyse the probability density of NWP deviation in each wind speed bin and the dependency of previous time series to the following wind speed. The NWP deviation is calculated as eq. (1).

\[
\text{v}_{\text{deviation}}(t) = \text{v}_{\text{NWP}}(t) - \text{v}_m(t) 
\]

where, \(\text{v}_{\text{NWP}}(t)\) is the NWP wind speed at time of \(t\); \(\text{v}_m(t)\) is the measured wind speed at time of \(t\).

![Figure 1. Frequency of the overall NWP deviation](image-url)
Firstly, the overall NWP deviation in this wind farm is calculated. The maximum value of the positive deviation is 15.8 m/s, the average value of the positive deviation is 3.38 m/s. The maximum value of the negative deviation is 12.5 m/s, the average value of the negative deviation is 2.54 m/s. The frequency of the overall NWP deviation is shown in Fig. 1. Distribution of the NWP deviation can be fitted by using normal distribution. However, the range of NWP deviation is large, so further study on the deviation of NWP wind speed is needed.

Secondly, the NWP wind speed is segmented into several bins, for example, 1 m/s per bin. According to the range of NWP wind speed and the power curve, four wind speed bins (0-1 m/s, 7-8 m/s, 14-15 m/s, 21-22 m/s) are selected to show the results of NWP deviation. Fig. 2 shows the frequencies of NWP deviation in several wind speed bins. Tab. 1 shows the statistics of NWP deviation in several wind speed bins, including maximum positive deviation, average positive deviation, maximum negative deviation and average negative deviation. It can be seen from Fig. 2 and Tab. 1, the frequencies of NWP deviation in small bins show different shapes, and cannot be fitted such well by the normal distribution. Also, the range of the NWP deviation is still large even if a small NWP wind speed bin is considered. It means that the relationship between NWP wind speed and the measured wind speed at the same time slot is not very clear. That is one important difficulty for the traditional NWP correction methods to achieve good performance. Therefore, more impact factors need to be considered to be learning labels during the training of wind speed prediction correction.

| NWP bin (m/s) | positive deviation (m/s) | negative deviation (m/s) |
|---------------|--------------------------|--------------------------|
|               | maximum | average | maximum | average |
| 0-1           | /       | /       | 9.55     | 4.30    |
| 7-8           | 6.64    | 1.97    | 9.91     | 2.18    |
| 14-15         | 12.7    | 4.66    | 5.97     | 1.55    |
| 21-22         | 13.2    | 8.11    | /        | /       |

Figure 2. Frequencies of NWP deviation in several wind speed bins

Table 1. Statistics of NWP deviation in several wind speed bins
The transfer relationship within the wind speed sequence is the basics for time-series method to predict wind speed. Therefore, the range of NWP deviation can be reduced when the measured wind speed at previous time is also considered as an input label. And the corresponding relationship between inputs and outputs of the prediction correction model will be much clearer under this condition. For example, the frequency of NWP deviation is shown in Fig. 3 when NWP wind speed at time of $t$ and the measured wind speed at time of $t-1$ are both limited in 7-8m/s. Compared with Fig. 2. b), the frequency of NWP deviation is more concentrated, and the range of NWP deviation is reduced obviously, only between -1 m/s to 1 m/s.

Figure 3. Frequency of NWP deviation
(Wind speed of NWP at time of $t$ and HMD at time of $t-1$ are both limited in 7~8m/s)

3. The Proposed Wind Speed Prediction Correction Method

This section describes the proposed wind speed prediction correction method construction. The results from section 2 show that even if NWP wind speed is within very small bins, the corresponding relationship between NWP and the measured wind speed at the same time slot is still obscure. Therefore, the time-independent prediction correction methods cannot largely improve the accuracy of NWP. In order to further improve the correction performance, as mentioned before, the impacts passing from the wind speed in the latest time series should be considered. A multi-step-ahead wind speed prediction correction method is proposed in this paper. The main idea of this work is to take the passing effects from wind speed at the previous time slot into account by adding the historical measurement data as another input variable. The proposed model is a rolling prediction correction model, the model and input variables of the model are updated in real-time. In this paper, SVM is employed to validate the proposed modelling strategy, and any AI-based algorithm, such as neural network, relevance vector machine, deep learning, etc. can be easily applied to implement the proposed strategy.

The main modelling process of the proposed wind speed prediction correction method is shown in Fig. 4.
Figure 4. Modelling process of the proposed wind speed prediction correction method

1) To collect the training samples, and to separate the samples into several groups according to NWP wind speed segments. In this case, the segment is 1 m/s for each.

2) To train the model for mapping the relationship between NWP wind speed and the wind speed measurements. NWP wind speed at time of $t$ and the wind speed measurement at time of $t-1$ are all imported as training inputs. And the measured wind speed at time of $t$ is imported as training targets.

3) In this case, SVM is used to establish the wind speed prediction correction model for every wind speed segment, and the model parameters in each wind speed segment are stored to the future correction.

4) The prediction correction model is determined and updated in real-time according to the NWP wind speed. And then the corresponding model is applied to the wind speed prediction correction. Except for the first wind speed correction point, the measured wind speed is substituted by predicted wind speed for input variables during the prediction in the proposed method.

The modelling process of SVM is as follows.

SVM [16] is proposed by Vapnik. This algorithm not only has the ability to deal with nonlinear regression problems, but also has the global optimal solution. It is suitable for solving small sample, nonlinear and high-dimensional pattern recognition problems. To solve the nonlinear regression problem of wind speed prediction, a nonlinear mapping function is used to map the original sample space into a high dimensional space, and then the linear regression method is applied to analysis. The regression linear function is shown as eq. (2).

$$f(x) = \langle w\phi(x) \rangle + b$$  \hspace{1cm} (2)

where, $w$ is the weight vector; $\phi(x)$ is the nonlinear mapping function; $b$ is the threshold.

So the solution of the regression problem is converted to search for optimal $w$ and $b$. The expression of mathematical programming problem is shown as eq. (3).
\[
\min \left( \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \right)
\]
\[
\begin{align}
\quad y_i - w \varphi(x_i) - b & \leq \varepsilon + \xi_i \\
\quad w \varphi(x_i) + b - y_i & \leq \varepsilon + \xi_i^* \\
\quad \xi_i & \geq 0 \\
\quad \xi_i^* & \geq 0
\end{align}
\] (3)

where, \( C \) is the penalty factor to equilibrate the model complexity and error term; \( \varepsilon \) is the estimated accuracy; \( \xi_i, \xi_i^* \) are relaxation variables, whose purpose are to deal with data that cannot be estimated; \( x_i \) are the inputs of training samples; \( y_i \) are the outputs of training samples.

Gaussian radial basis kernel function is selected as kernel function in this paper, as shown in eq. (4).

\[
K(x_i, x) = \exp \left( \frac{\|x_i - x\|^2}{2\sigma^2} \right)
\] (4)

where, \( \sigma \) is the nuclear parameter. The minimization form is transformed into dual form by Lagrange function. Thus, the wind speed prediction model (regression function) can be expressed as eq. (5).

\[
f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(x_i, x) + b
\] (5)

where, \( \alpha_i, \alpha_i^* \) are Lagrange multipliers.

4. Case Study

In this section, the NWP and measured wind speed of the wind farm in north China is taken as example to analysis. Temporal resolution of the data is 10 minutes. Data of the first 20 days each month are taken as the train samples, and the rest data of each month are taken as the test samples. Based on two widely acknowledged criteria, three conventional benchmark methods are compared with the performance of the proposed model, and followed by the error analysis for different time horizons.

4.1. Performance Criteria

In order to evaluate the accuracy of wind speed prediction models, two error indexes are used as evaluation criteria. One is root mean square error (RMSE), formulated in eq. (6), to evaluate the overall error. Another is mean absolute percentage error (MAPE), formulated in eq. (7), to evaluate the real-time error.

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N} (v_{p}(t) - v_{m}(t))^2}{N}}
\] (6)

\[
MAPE = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{v_{p}(t) - v_{m}(t)}{v_{m}(t)} \right|
\] (7)

where, \( v_{m}(t) \) is the measured wind speed at time of \( t \); \( v_{p}(t) \) is the predicted wind speed at time of \( t \); \( N \) is the length of prediction.

4.2. Results and Analysis

In this case study, three widely used methods are selected as the benchmark, which are ARMA, SVM and SVM in all wind speed bins (SVM-bins). ARMA is a typical time-series method working well for the ultra-short-time prediction, it is mainly used to validate the performance of 1-4h prediction. SVM is one of mainstream algorithm for establishing the mapping relationship between NWP and the
measured wind speed at the same time slot. SVM-bins method is to establish typical SVM model for each wind speed bin, in this case 1 m/s per bin.

Fig. 5 compares the error of raw NWP data, and corrected NWP data by the proposed method and the other three benchmarks. The “NWP” represents original NWP. The “NWP & HMD” represents the results of the proposed wind speed prediction correction model. RMSE and MAPE of the proposed model are compared with those of the original NWP as well as the other three conventional models under different time horizons. The percentage of the accuracy improvement is calculated, and the comparison results to the original NWP and the second-ranking benchmark are shown in Tab. 2 and Tab. 3. It can be seen that, the proposed wind speed prediction correction model has the highest accuracy comparing with the three benchmarks under different time horizons for both evaluation criteria. For 1h and 4h, ARMA is in the second-ranking of this comparison. Compared with ARMA, the prediction accuracy of the proposed model is increased 8.15% and 1.76% of RMSE and MAPE under 1h time horizon, 8.66% and 0.80% of RMSE and MAPE under 4h time horizon. For 12h, SVM-bins and ARMA are the second-ranking of this comparison when using RMSE and MAPE as the evaluation criterion respectively. The improvement accuracy of the proposed model is 16.67% of RMSE comparing with SVM-bins, 15.26% of MAPE comparing with ARMA under 12h time horizon. For 24h, SVM-bins is in the second-ranking of this comparison. Compared with SVM-bins, the prediction accuracy of the proposed model is increased 13.41% and 20.87% of RMSE and MAPE under 24h time horizon.

![Figure 5. Comparison of wind speed prediction error using two evaluation criteria](image)

| Table 2. Improvement compared with original NWP |
|------------------------------------------------|
|          | 1h     | 4h     | 12h    | 24h    |
| RMSE     | 69.22% | 54.91% | 45.59% | 43.46% |
| MAPE     | 70.06% | 52.69% | 40.81% | 39.12% |

| Table 3. Improvement compared with the second-ranking method |
|-------------------------------------------------------------|
|          | 1h     | 4h     | 12h    | 24h    |
| RMSE     | 8.15%  | 8.66%  | 16.67% | 13.41% |
| MAPE     | 1.76%  | 0.80%  | 15.26% | 20.87% |

To further validate the adaptability of the proposed method, RMSE in different months are compared for the four models under ultra-short-term and short-term time scales. The results are shown in Fig. 6. It can be seen that in each month the proposed method outperforms the other benchmarks under various time horizons. And, for the ultra-short-term (1 and 4 hours ahead) the performance of the proposed method is similar to that of ARMA at around 1-2m/s and 1.5-2.5m/s; while for the short-term (12 and 24 hours ahead) the proposed method seems to have larger advantages comparing to the rest methods, and the deviation range of the proposed method for 12h and 24h is around 2-3m/s. Also, the degradation of the proposed method from 1h to 24h is not as large as that of ARMA. Fig. 7 shows the accuracy degradation of the proposed method with the increase of the prediction horizons. As

![Figure 6. Comparison of wind speed prediction error in different months](image)
shown, the RMSE and MAPE of the proposed method increase sharply before 8-hour time horizon, and stays constant when the time horizon is over 12h.

![Figure 6. RMSE of four prediction models in different months](image1)

![Figure 7. RMSE and MAPE of the proposed model under different time horizons](image2)

Deviation of the wind speed prediction can be divided into two parts, one is the statistical deviation, which might be able to eliminate by the statistical method and mapping algorithm to capture its pattern, for example, deviation grows with the wind speed increases; another is the random deviation, which might not be easy to capture its pattern. The correction of wind speed prediction aims to eliminate the statistical deviation and to leave the random deviation to the corrected NWP wind speed. Fig. 8 shows the corrected results for different wind speed bins. For the time horizon of 1h, the RMSE of each wind speed bin roughly ranges from 1m/s - 1.5m/s, which is different from the original NWP deviation pattern. For the rest time horizons, the results show similar but not such apparent features.
5. Conclusion
In this paper, a multi-step-ahead wind speed prediction correction method is proposed with consideration of the impacts passing from the wind speed in the latest time series. Both numerical weather prediction and historical measurement data are used as training inputs. One Chinese wind farm in northern part of China is taken as example to validate the proposed method. The conclusions of this paper are as follows.

1) Through the probabilistic analysis, the NWP deviation pattern is not very clear when only the NWP wind speed interval is considered. The actual wind speed varies widely even if the NWP wind speed is within a very small range. The involvement of wind speed at previous time slot enhances the capture of a clearer deviation pattern. This motivates to present a time-dependent method to correct the wind speed prediction in this paper.

2) In this case study, the proposed method can improve the accuracy of NWP wind speed to a large extent at different time horizons. Compared to the second-ranking benchmark method, the proposed method outperforms by 8.15%, 8.66%, 16.67% and 13.41% at the time horizons of 1h, 4h, 12h and 24h in terms of RMSE. Also, the accuracy degradation of the proposed method is not as large as the traditional time-series based method especially.

Acknowledgments
This paper is supported by National Natural Science Foundation of China (Project No. 51376062) and the project of “system development of the optimization and evaluation of the wind farm siting and wind turbine selection” from United Power.

Reference
[1] Du P, Wang J Z, Guo Z H and Yang W D 2017 Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting Energy Conversion and Management 150 90-107
[2] Li H B, Lu Z X, Qiao Y and Zeng P L 2015 Assessment on operational flexibility of power grid with grid-connected large-scale wind farms Power System Technology 39 1672-78
[3] Ren G R, Liu J F, Wan J, Guo Y F and Yu D R 2017 Overview of wind power intermittency: impacts, measurements, and mitigation solutions Applied Energy 204 47-65
[4] Giebel G, Cline J, Frank H, Shaw W, Pinson P, Hodge B.-M, Kariniotakis G, Madsen J and Moehrle C 2016 Wind power forecasting: IEA Wind Task 36 & future research issues Journal of Physics: Conference Series 53 032042-10
[5] Lazarevska E 2016 Comparison of Different Models for Wind Speed Prediction IEC ON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society (IEEE, Piscataway, NJ, USA: Elsevier) 5544-9
[6] Kavasseri R G and Seetharaman K 2009 Day-ahead wind speed forecasting using f-ARIMA models Renewable Energy 34 1388-93
[7] Urata K, Inoue M, Murayama D and Adachi, S 2016 Dynamic Modelling and Very Short-term Prediction of Wind Power Output Using Box-Cox Transformation Journal of Physics: Conference Series 744 012176

[8] Galanis G, Papageorgiou E and Liakatas A 2017 A hybrid Bayesian Kalman filter and applications to numerical wind speed modeling Journal of Wind Engineering and Industrial Aerodynamics 167 1-22

[9] Zhang Y, Sun H X, Guo Y J and Lu Z P 2016 Improved grey model with rolling method for wind power prediction 35th Chinese Control Conference (IEEE, Piscataway, NJ, USA: Elsevier) 9784-6

[10] Li L, Liu Y Q, Yang Y P, Han S and Wang Y M 2013 A physical approach of the short-term wind power prediction based on CFD pre-calculated flow fields Journal of Hydrodynamics 25 56-61

[11] Wang D Y, Luo H Y, Grunder O and Lin Y B 2017 Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction Renewable Energy 113 1345-58

[12] Sapronova A, Meissner C and Mana M. 2016 Short time ahead wind power production forecast Journal of Physics: Conference Series 749 012006

[13] Yan J, Gao X L, Liu Y Q, Han S, Li L, Ma X M, Gu C H, Bhakar R and Furong L 2015 Adaptabilities of three mainstream short-term wind power forecasting methods Journal of Renewable and Sustainable Energy 7 053101

[14] Zang H X, Fan L, Guo M, Wei Z N, Sun G Q and Zhang L 2016 Short-Term Wind Power Interval Forecasting Based on an EEMD-RT-RVM Model Advances in Meteorology 2016 8760780

[15] Wang H Z, Li G Q, Wang G B, Peng J C, Jiang H and Liu Y T 2017 Deep learning based ensemble approach for probabilistic wind power forecasting Applied Energy 188 56-70

[16] Vapnik V 1995 The nature of statistical learning theory New York: Springer-Verlag