ORIGINAL RESEARCH

Open Access

System bias correction of short-term hub-height wind forecasts using the Kalman filter

Jingjing Xu[1] , Ziniu Xiao[2*], Zhaohui Lin[1] and Ming Li[3]

Abstract

Wind energy is a fluctuating source for power systems, which poses challenges to grid planning for the wind power industry. To improve the short-term wind forecasts at turbine height, the bias correction approach Kalman filter (KF) is applied to 72-h wind speed forecasts from the WRF model in Zhangbei wind farm for a period over two years. The KF approach shows a remarkable ability in improving the raw forecasts by decreasing the root-mean-square error by 16% from 3.58 to 3.01 m s^{-1}, the mean absolute error by 14% from 2.71 to 2.34 m s^{-1}, the bias from 0.22 to -0.19 m s^{-1}, and improving the correlation from 0.58 to 0.66. The KF significantly reduces random errors of the model, showing the capability to deal with the forecast errors associated with physical processes which cannot be accurately handled by the numerical model. In addition, the improvement of the bias correction is larger for wind speeds sensitive to wind power generation. So the KF approach is suitable for short-term wind power prediction.

Keywords: Wind forecasts, Wind energy, Numerical model, Bias correction, Kalman filter, Atmospheric boundary layer

1 Introduction

Wind energy is one of the most important sources of renewable energy, and over the last decade there has been rapid growth in wind energy production in China. Because of its stochastic nature, wind is inherently varying and wind energy is a fluctuating source for power systems [1]. Accurate wind power prediction can reduce the risk of uncertainty and lead to better grid planning and integration of wind energy. Thus, short-term wind power prediction, which has been discussed widely [2–5], is useful in power system planning. According to the demands of the wind energy industry, 15-min wind speed forecast at turbine height with a forecast period of 24–72 h should be provided for short-term wind power prediction.

Numerical weather prediction (NWP) models can exhibit systematic errors in wind speed and ramp event forecasts [6–8], especially near the surface. This can be caused by many factors, such as shortcomings in the physical parameterization, inability to successfully handle sub-grid phenomena, inaccurate initial and lateral boundary conditions, interpolations to areas not close to grid points or model levels, and smoothed orographic and landscape characteristics associated with the model horizontal resolution. To reduce the influence of these drawbacks in the final output of a NWP model, a variety of approaches based on statistical methods have been used [9–11], such as linear regression, moving average correction, partial least squares method, etc. These approaches are called model output statistics (MOS) methods, and employ postprocessing bias correction to improve the direct model outputs.

One of the most convenient postprocessing approaches is Kalman filter (KF) [12], which is a statistically optimal sequential estimation procedure for dynamic systems. Observations are recursively combined with recent forecasts using weights that minimize the corresponding biases. The KF approach carries minor computational

*Correspondence: xiaozn@lasg.iap.ac.cn
[2] State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
Full list of author information is available at the end of the article

costs and can easily adapt to any alterations in observations. In addition, this recursive and adaptive method does not need an extensive database for training.

The KF method has been widely used for improving weather forecasts of continuous variables such as 2 m temperature, 10 m wind speed and ozone concentration etc. [13–17]. As a postprocessing bias correction method, the KF uses recent past observations and forecasts to estimate the model bias in future forecasts to correct the raw forecasts. The nonlinear formation of Kalman filter has been applied to improve short-term wind speed forecasts at turbine height [18] and to forecast wind speed and wind power at a wind farm site [19]. Reference [20] modifies the nonlinear Kalman Filter proposed in [18] and applies the method to improve the wind ramp event forecasts for wind power industry. As well as the MOS methods and the KF approach, many studies have employed various other approaches to improve short-term wind speed and power forecasts, such as analog bias correction [21, 22], Gaussian process regression [23], machine learning techniques such as neural network [24, 25] and random forests [21], analog ensemble [26] and time-dependent bias correction [27] for probability forecasting. Other related studies include data preprocessing technique [28] and circular-circular regression method for wind direction forecasts [29].

Reference [30] shows that in the lower atmosphere, forecast errors can be correlated for several days and contain strong diurnal variability. So in this paper, diurnal behavior of the forecast error is taken into consideration and the KF is applied to Zhangbei wind farm in northeastern China to perform a detailed statistical analysis. The state-of-the-art MOS method is also employed as a comparison to evaluate the forecast skill of the KF approach. The discussion is focused on the capability of the KF to improve the raw forecasts, especially on systematic and random errors.

The rest of this paper is organized as follows. In Sect. 2, information on the wind farm, the WRF model and the experimental design is provided. Since the turbine height (60–100 m) is in the lower boundary layer, a special Kalman filter for each hour of the day is presented in Sect. 3, so that the forecast error can be estimated from data with lags of multiple of 24-h. In Sect. 4 a case study on wind speed forecasts in Zhangbei wind farm over two years is illustrated. Finally, Sect. 5 draws conclusions.

2 Wind speed datasets and numerical simulations
2.1 The wind farm and observations
Zhangbebi wind farm is located northwest of Hebei Province in China. The farm consists of 66 wind turbines with a total installed capacity of around 100 MW. The wind turbine height is 70 m, and the wind farm covers complex mountain terrain with the altitude ranging from 1600 to 1800 m (see Fig. 1).

Observational data comes from one mast situated at the wind farm, at latitude 41° 03′ N, longitude 114° 29′ E and altitude 1660 m. Wind speed data was recorded by the anemometer with a sampling rate of 5 min. The dataset used in the present study spans over 2 years, from 07 November 2015 to 22 February 2018 (839 days in total), and the measurements are averaged over 15 min in order to meet the demands of the wind power industry.

2.2 The WRF model and experimental design
The WRF model is a nonhydrostatic compressible-equation, mesoscale NWP model [31], and the proposed bias correction methods have been applied to improve the hub-height wind speed forecasts provided by the WRF model.

In the present study, the WRF model is run over northeastern China with three nested domains centered on Zhangbei wind farm with 27-km, 9-km, and 3-km horizontal grid increments. The forecast periods are 24-h, 48-h and 72-h, and the wind speeds are issued in 15-min increments. Vertical levels in the model have fine resolution around wind turbine height with 48 levels in total and 6 levels in the lowest 100 m. The 70-m wind speed forecasts within the finest domain are bilinearly interpolated to the observation location directly. The parameterizations chosen for these experiments include the Purdue Lin microphysics scheme, the Mellor–Yamada–Janjic planetary boundary layer (PBL) scheme, the Eta scheme for the surface layer, the Kain–Fritsch scheme for the convective processes (only in the two coarser domains), and the Noah land surface model for the land surface scheme.
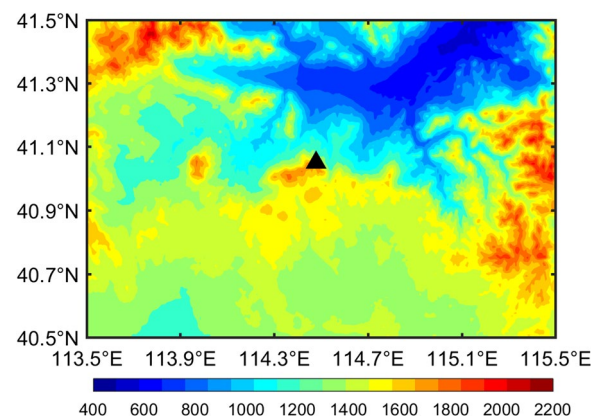
Fig. 1 Location of Zhangbei wind farm (the black solid triangle denotes the observation station)

3 Bias correction procedure

3.1 The Kalman filter (KF) approach

In numerical weather forecasts, the specific KF approach predicts the future bias as equal to the old bias plus uncertainty, which is corrected by the linear function of the difference between the previous prediction and verifying bias.

In our study, the KF method models the true (unknown) forecast bias $\mathbf{x}_t$ at time t by the previous true bias plus a white noise $\mathbf{w}_t$ term as:

$$\mathbf{x}_t = \mathbf{x}_{t-\Delta t} + \mathbf{w}_t \tag{1}$$

where Δt is the time delay. Assuming the forecast error $\mathbf{y}_t$ is corrupted from truth by a random error term $\mathbf{v}_t$ as:

$$\mathbf{y}_t = \mathbf{x}_t + \mathbf{v}_t \tag{2}$$

thus, $\mathbf{y}_t$ includes both the systematic bias and random error. Therefore, the system matrix and the observation matrix are designated as the identity matrix for the proposed algorithm. The random vectors $\mathbf{w}_t$ and $\mathbf{v}_t$ must follow the normal distribution with zero mean and be independent.

Kalman [12] showed that the optimal recursive predictor of $\mathbf{x}_t$ can be written as a combination of the previous bias estimation and the previous forecast error. The optimal estimates given for the state vector $\mathbf{x}_t$ and covariance matrix $\mathbf{P}_t$ at time t are:

$$\mathbf{x}_{t/t-\Delta t} = \mathbf{x}_{t-\Delta t}, \ \ \mathbf{P}_{t/t-\Delta t} = \mathbf{P}_{t-\Delta t} + \mathbf{W}_t \tag{3}$$

As soon as the new value $\mathbf{y}_t$ becomes known, the new value of state vector $\mathbf{x}_t$ is calculated as:

$$\mathbf{x}_t = \mathbf{x}_{t/t-\Delta t} + \mathbf{K}_t(\mathbf{y}_t - \mathbf{x}_{t/t-\Delta t}) \tag{4}$$

where $\mathbf{K}_t$ is the Kalman gain matrix and is the most crucial parameter of the filter. $\mathbf{K}_t$ determines how easily the filter will adjust to any possible new conditions, and is given as:

$$\mathbf{K}_t = \mathbf{P}_{t/t-\Delta t}(\mathbf{P}_{t/t-\Delta t} + \mathbf{V}_t)^{-1} \tag{5}$$

Finally, the new value of the covariance matrix $\mathbf{P}_t$ of the unknown state is given by:

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t)\mathbf{P}_{t/t-\Delta t} \tag{6}$$

Equations (3)–(6) are known as updating equations, which update the Kalman algorithm from time $t - \Delta t$ to t. Note that -1 denotes the inverse matrix while $\mathbf{I}$ stands for the identity matrix. This procedure is recursive because values of the KF coefficients at any one-time step depend on the values at the previous time step, and it is optimal in a least-squares sense.

Since $\mathbf{x}_t$ and $\mathbf{P}_t$ converge to their true values rapidly, the initial values $\mathbf{x}_0$, $\mathbf{P}_0$ do not seriously affect the results of the algorithm. However, the way in which the covariant matrices $\mathbf{W}_t$ of the random vector $\mathbf{w}_t$ and $\mathbf{V}_t$ of $\mathbf{v}_t$ are calculated during the process can significantly affect the final outcome. In our case, the system covariant matrix $\mathbf{W}_t$ and observation covariant matrix $\mathbf{V}_t$ are estimated based on the samples of the last 7 values of $\mathbf{w}_t = \mathbf{x}_{t-}\mathbf{x}_{t-1}$ and $\mathbf{v}_t = \mathbf{y}_{t-}\mathbf{x}_t$ respectively, as:

$$\mathbf{W}_t = \frac{1}{n-1}\sum_{i=1}^{n}\left((\mathbf{x}_{t-i} - \mathbf{x}_{t-i-1}) - \frac{\sum_{j=1}^{n}(\mathbf{x}_{t-j} - \mathbf{x}_{t-j-1})}{n}\right)^2 \tag{7}$$

$$\mathbf{V}_t = \frac{1}{n-1}\sum_{i=1}^{n}\left((\mathbf{y}_{t-i} - \mathbf{x}_{t-i}) - \frac{\sum_{j=1}^{n}(\mathbf{y}_{t-j} - \mathbf{x}_{t-j})}{n}\right)^2 \tag{8}$$

where $n=7$ and $n-1$ represents unbiased estimation. The period of 7 values has proved to be the optimal choice in the study for successful correction and fast adaptability. Here the initial values of $\mathbf{x}_t$, $\mathbf{P}_t$, $\mathbf{W}_t$ and $\mathbf{V}_t$ are set to 0, 4, 1 and 6, respectively.

For a time delay of $\Delta t=24$ h, it means today's forecast bias is estimated using yesterday's bias, which in turn was estimated using the bias from the day-before-yesterday. The forecast bias may have time-varying behavior at different times of the day (e.g., different wind speed activities during daytime versus nighttime), so the influence of diurnal variation in atmospheric boundary layer is considered. Different from the nonlinear version of Kalman filter in [18, 20], the KF algorithm used in this study is run on data for each hour of the day to eliminate the error caused by the diurnal behavior, using only values from previous days at the same hour of the day. In this way, a given hour is corrected using only the past forecasts and observations at the same hour.

Because the original data is in 15-min increments, the observations and forecasts are transformed to one-hour averages as pre-processing of the bias correction. The results for the first 39 days are removed to eliminate the effects of the KF algorithm's spin-up, and the results for the remaining 800 days are then analyzed. Finally, the corrected one-hour wind speeds are transformed back to 15-min wind speeds.

3.2 Model output statistics (MOS) method

Since the MOS method is the most widely used in post-processing techniques for numerical models, the correction results by the MOS method are also provided to compare with the KF correction results.

MOS uses a multiple linear regression to correct systematic errors in NWP models using NWP outputs of certain weather variables as predictors and fitting these predictors

Table 1 The overall performance of the KF bias correction for 24-h, 48-h, 72-h forecasts and MOS bias correction for 72-h forecasts based on the five metrics (bias, MAE, RMSE, CRMSE, and correlation coefficient)

Statistics	Bias (m s^{-1})	MAE (m s^{-1})	RMSE (m s^{-1})	CRMSE (m s^{-1})	Correlation coefficient
Raw forecasts	0.22	2.71	3.58	3.57	0.58
KF bias correction (24 h)	− 0.16	2.28	2.92	2.91	0.67
KF bias correction (48 h)	− 0.18	2.32	2.99	2.98	0.66
KF bias correction (72 h)	− 0.19	2.34	3.01	3.00	0.66
MOS bias correction (72 h)	0.19	2.49	3.22	3.21	0.59

to observations. To find the optimal predictors, the MOS multivariate regression equations are used.

MOS aims at correcting forecasts by means of a regression function r between the variable Y to be predicted and some explanatory variable(s) (or predictors) X which may be NWP model output(s) or any other source of information. This regression function is estimated based on past forecasts and associated observations, and it is then applied to future forecasts to increase their accuracy. In the present study, a classical regression method is applied [32], and the regression function is a second-order polynomial relationship of the explanatory variables as:

$$\hat{r}(X) = \beta_0 + \boldsymbol{\beta} \cdot \mathbf{X}^{1,2} \tag{9}$$

where β_0 is a real number, $\boldsymbol{\beta}$ is a vector of real numbers, and $\mathbf{X}^{1,2}$ is the vector containing every possible combination of product of explanatory variables of order 1 and 2 (called interactions). The parameters β_0 and $\boldsymbol{\beta}$ are fitted onto the training dataset with an ascending selection of predictors based on the Bayesian information criterion. WRF-modeled U, V, and W wind components, wind speed (*wspd*), temperature T, pressure P and relative humidity (*rh*) are used as potential predictors here.

4 Results and discussion

The bias correction methods presented in Sect. 3 are applied to improve the wind speed forecasts in this section. The root-mean-square error (RMSE) can be decomposed as:

$$R^2 = R_C^2 + B^2 \tag{10}$$

where

$$R_C = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} [(F_i - \overline{F}) - (O_i - \overline{O})]^2} \tag{11}$$

$$B = \overline{F} - \overline{O} \tag{12}$$

In Eqs. (10)–(12), the variable R represents RMSE, R_C represents the centered root-mean-square error

(CRMSE), and B denotes bias. O_i is the observation value at time i and F_i is the corresponding forecast value. N_p is the size of sample which means the number of pairs (F_i, O_i). $\overline{F}$ and $\overline{O}$ are the mean forecast value and observation value, respectively. CRMSE is considered to be the random component of RMSE, while the systematic component is represented by bias. CRMSE can be associated with the intrinsic predictive skill of the forecast because of the limitation of physical parameterization, and bias estimates the systematic errors which may be associated with biased initial conditions, smoothed topography, offset parameter values, etc.

4.1 Global performance

To evaluate the performance of the bias correction, the global metrics of the results are investigated first. Table 1 characterizes the ability of improving forecast skill for the bias correction quantitatively, based on metrics of bias, mean absolute error (MAE), RMSE, CRMSE and the correlation coefficient. Because the study mainly analyzes the bias correction results for 72-h forecasts, only 72-h MOS bias correction results are presented as comparison. From Table 1, it can be seen that the KF approach for 72-h forecasts has decreased RMSE by 16% from 3.58 to 3.01 m s^{-1}, MAE by 14% from 2.71 to 2.34 m s^{-1}, bias from 0.22 to − 0.19 m s^{-1}, and improved the correlation from 0.58 to 0.66. The result for CRMSE is similar to that for RMSE. Table 1 also shows the results of the KF bias correction for 24-h and 48-h forecasts. It is not surprising that the KF for 24-h forecast shows the best performance, as when the forecast period gets longer, the correction effect is weakened. For results of 24-h forecasts, the KF bias correction has decreased RMSE by 18%, MAE by 16%, bias by 27%, and improved the correlation from 0.58 to 0.67, with all the metrics showing remarkable improvement. Although 0.6 m s^{-1} improvement in RMSE in some circumstances is not seen to be high, this is a substantial improvement in wind energy. In a 2008 Department of Energy publication, it was stated that, "a 1% error in wind speed estimates for a 100-MW wind generation facility can lead to losses approaching \$12,000,000 over the lifetime of that plant" [33]. The MOS bias correction also
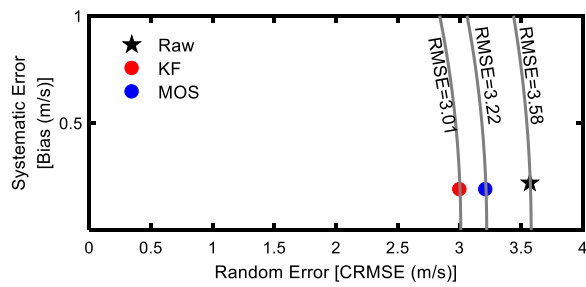
Fig. 2 Bias (m s^{-1}) as a function of CRMSE (m s^{-1}), showing the raw forecasts and the KF and MOS bias corrections

shows moderate improvement, decreasing RMSE by 10%, MAE by 8% and bias by 14%. From the above results, it is obvious that the KF is suitable for short-term wind power prediction with the forecast period of 24–72 h. In the rest of this paper, only the results for 72-h forecast are analyzed as representative forecasts.

From the relationship shown by (10), if CRMSE is plotted against the bias, the distance of the point from the origin is equal to RMSE. Figure 2 illustrates such a diagram for the raw forecasts and the KF and MOS bias

corrections, where the radial distances from the origin to the gray circular lines equal RMSE. Obviously, wind speed errors in this forecast system are dominated by the random component. The KF not only reduces the bias component of RMSE, but also the remaining part. The latter can be seen as the ability of the KF to add predictive skill to the raw forecasts by reducing random errors.

### 4.2 Seasonal and diurnal variability for bias correction

To investigate the seasonal and diurnal characteristics of the forecast errors, the two metrics with all the available data at a given time of the 24 forecast hours are computed. Figures 3 and 4 show the temporal evolution of bias and CRMSE, respectively for the raw forecasts, and the KF and MOS bias correction results.

The biases in summer and autumn are much larger than in winter and spring, and exhibit violent diurnal variation. The maximum bias is almost 3 m s^{-1} which appears in summer. The growth of the planetary boundary layer (PBL) is often a challenging process for prediction, which is reflected by the jump in the bias values before sunrise (05 LST-07 LST) in summer and autumn (Fig. 3). Throughout the daytime and early evening, the
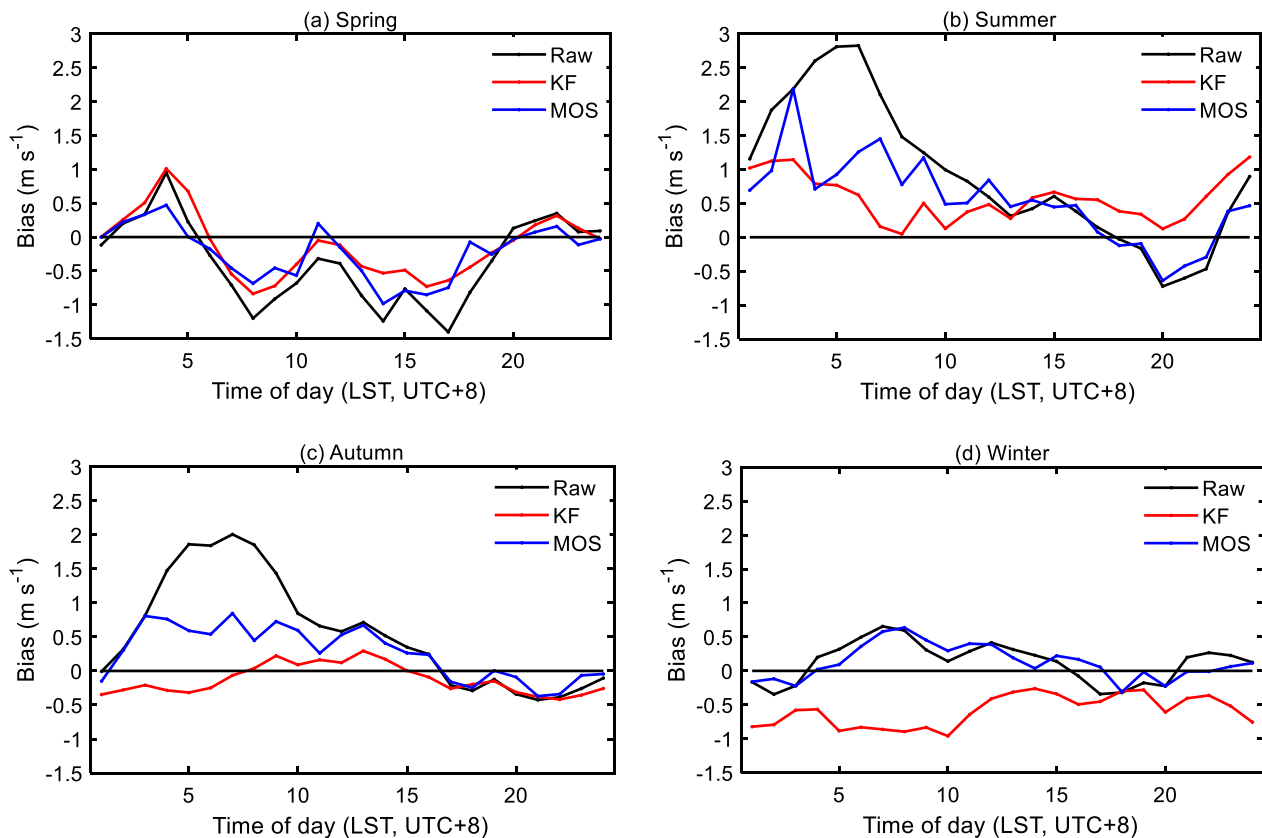
Fig. 3 Temporal evolution of bias in **a** spring, **b** summer, **c** autumn, **d** winter for the raw forecasts, and the KF and MOS bias corrections across the 24 h of forecast

Xu *et al. Protection and Control of Modern Power Systems*   (2021) 6:37

Page 6 of 9
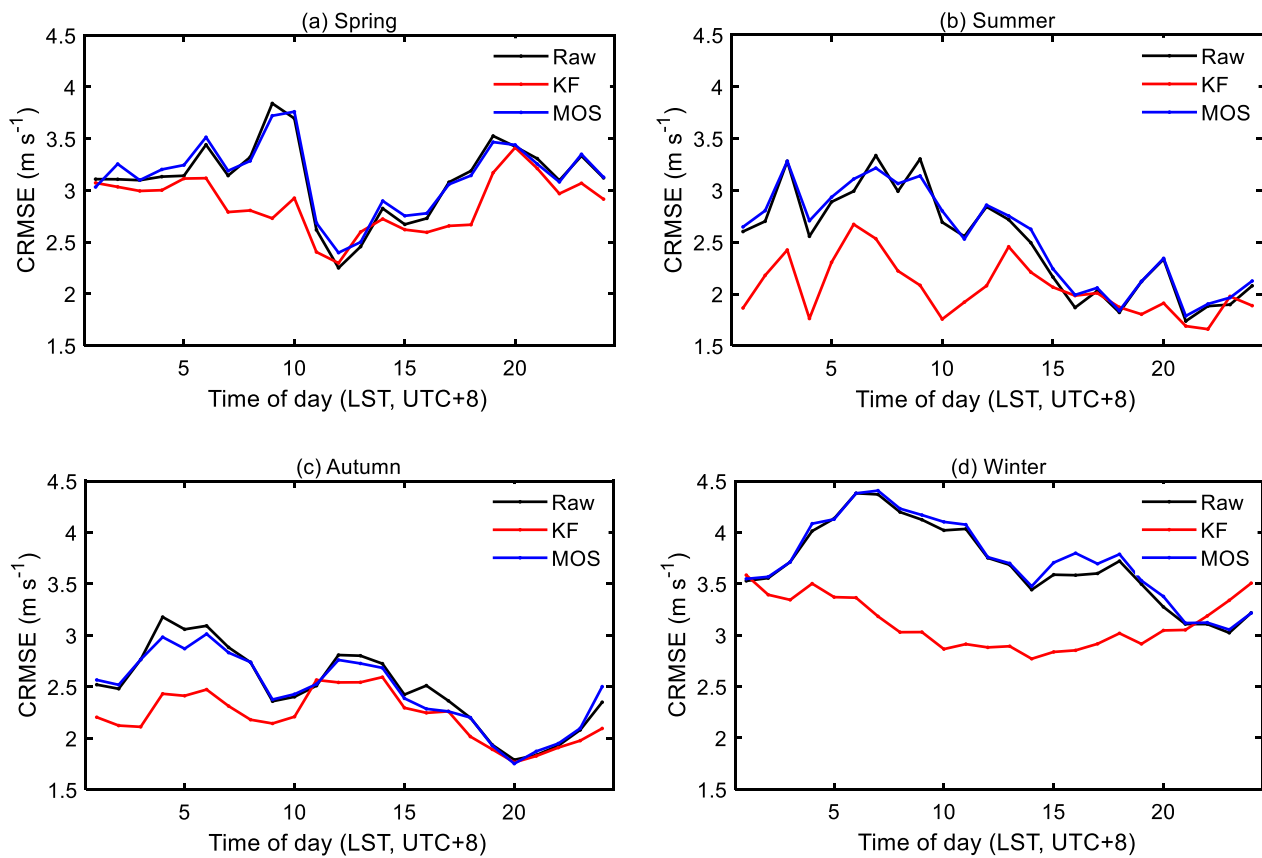
Fig. 4 Temporal evolution of CRMSE in **a** spring, **b** summer, **c** autumn, **d** winter for the raw forecasts, and the KF and MOS bias corrections across the 24 h of forecast

bias values stay constant around 0.5 m s^{-1}, and before sunrise, the raw forecasts show an increase in bias, given the uncertainty associated with the PBL growth. The KF shows improvement across all the forecast hours in the range of 0.5–2.5 m s^{-1} with respect to the raw forecasts, while decreasing bias around 1.5–2.5 m s^{-1} in 05 LST-07 LST. This is the indication of the KF method's ability to improve the predictive skill. In winter and spring, the diurnal cycles are not so obvious compared to those in summer and autumn, and the KF correction effect is not so remarkable. The MOS bias correction also shows moderate improvement in summer and autumn, especially in PBL growth time, though the KF approach performs better than the MOS method nearly all the time.

For CRMSE (Fig. 4), the results are contrary to those for bias. The CRMSE in winter and spring is larger than in summer and autumn, and the maximum CRMSE appearing in winter is almost 4.5 m s^{-1}. The KF drastically reduces CRMSE of the raw forecasts by 1 m s^{-1} in winter. In spring, there is a high peak around 09 LST-10 LST, and then it drops abruptly by almost 1.5 m s^{-1}. The KF approach reduces CRMSE by around 1 m s^{-1} in this

time interval, showing the capability to handle the forecast error associated with PBL growth. Figure 4b shows similar results in summer. The KF reduces CRMSE of 0.7–1.0 m s^{-1} from midnight to noon. The MOS method shows little improvement for CRMSE in all four seasons, which indicates that the MOS method mainly reduces bias of the forecast error rather than the CRMSE component.

4.3 Evaluation of wind speed forecasts sensitive to wind power prediction

It is well known that wind speed forecasts for wind power prediction are different from routine wind speed forecasts according to the wind turbine power curve. When wind speed is equal to or less than 5 m s^{-1} a turbine will not generate power output, while once the wind speed exceeds 12 m s^{-1}, the wind turbine reaches its maximum capability. So forecast accuracy for wind speed between 5 and 12 m s^{-1} is essential to wind power prediction. Hence the wind speed forecasts sensitive to power prediction are evaluated in this subsection.
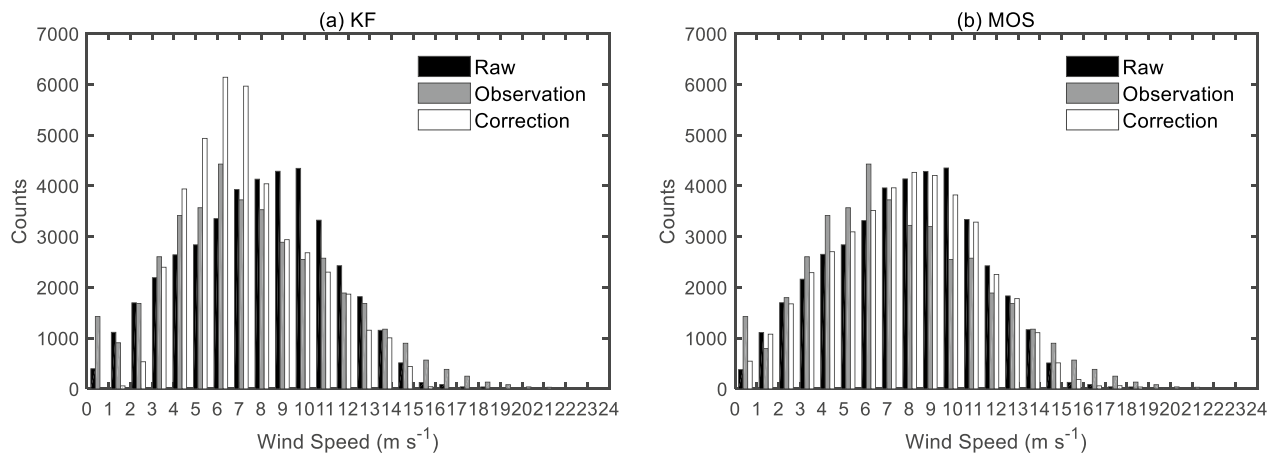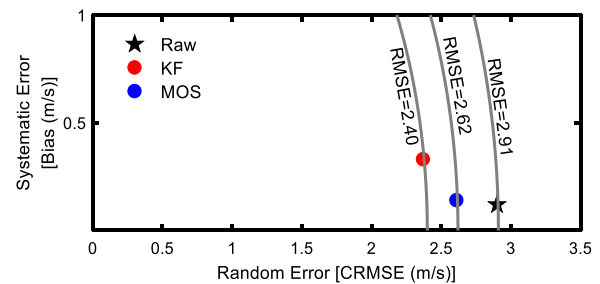
Fig. 5 The wind speed distribution of the observation, the raw forecasts and the **a** KF, **b** MOS

Table 2 The overall performance of the KF and MOS bias corrections for wind speeds sensitive to wind power prediction based on four metrics (bias, MAE, RMSE, and CRMSE)

Statistics	Bias (m s^{-1})	MAE (m s^{-1})	RMSE (m s^{-1})	CRMSE (m s^{-1})
Raw forecasts	0.12	2.28	2.91	2.90
KF	-0.35	1.88	2.40	2.36
MOS	0.14	2.03	2.62	2.61

Fig. 6 Bias (m s^{-1}) as a function of CRMSE (m s^{-1}) for wind speeds sensitive to wind power prediction, showing the raw forecasts, and the KF and MOS bias corrections

Figure 5 gives the wind speed distribution of the observation, the raw forecasts, and the correction. For the KF approach (Fig. 5a), the bias correction mainly improves the raw forecasts in 4–12 m s^{-1}, except for wind speeds in the 5–7 m s^{-1} interval. In the wind speed interval of 9–12 m s^{-1}, which are included in the area sensitive to wind power prediction (5–12 m s^{-1}), the correction shows significant efficiency. While for the MOS method (Fig. 5b), the bias correction only provides small improvements on the raw forecasts in most wind speed intervals.

To quantitatively evaluate the performance of the two bias correction methods for wind speeds sensitive to wind power prediction, the results are presented in Table 2. The raw forecasts for wind speed sensitive to wind power show better forecast skill than in Table 1 based on all four metrics. This is rational because the model exhibits large errors when wind speed is low. In complex terrain, low wind speed is always associated with local circulations which cannot be accurately handled by the numerical model. The KF has decreased RMSE by 18% from 2.91 to 2.40 m s^{-1}, MAE by 17% from 2.28 to 1.88 m s^{-1}. The larger reduction of CRMSE by 19% from 2.90 to 2.36 m s^{-1} compared to RMSE, is due to the promoted bias. Thus, for wind speeds sensitive to wind power, the improvement of the bias correction is

larger. This result proves that the KF approach is especially suitable for wind power prediction. The same fact can be seen in Fig. 6, while the MOS method also shows moderate improvement for wind speeds sensitive to wind power prediction.

The overall performance of the two bias correction methods has been evaluated. In contrast to MOS and other statistical methods which require a long training period and behave in a static manner, the KF approach adapts its coefficients during each time step. It requires a much shorter training period, and can adapt to changing synoptic conditions, changing seasons, and even changing weather forecast models. However, the disadvantage of the method is that it is less likely to predict extreme bias events, e.g., it is unable to anticipate a large bias when all biases for the past few days have been small.

5 Conclusions

To improve the short-term hub-height wind speed forecasts for wind power prediction, a postprocessing bias correction Kalman filter approach is applied to 72-h wind speed forecasts at 70-m height from the WRF

model in Zhangbei wind farm for a period of more than two years. The main conclusions can be summarized as follows:

The KF approach for 72-h forecasts has decreased RMSE by 16% from 3.58 to 3.01 m s^{-1}, MAE by 14% from 2.71 to 2.34 m s^{-1}, bias from 0.22 to -0.19 m s^{-1}, and improved the correlation from 0.58 to 0.66. Wind speed errors in this forecast system are dominated by the random component, while the KF not only reduces systematic errors of the model, but also significantly reduces random errors of it.

In all seasons, the forecast errors (including systematic errors and random errors) of the raw forecasts exhibit a strong diurnal cycle with a peak before sunrise in PBL growth time. The KF approach significantly reduces random errors of the model when the PBL growth occurs, showing the capability to deal with the forecast errors associated with physical processes which cannot be accurately handled by the numerical model. In contrast, the MOS method mainly reduces systematic component of the error rather than the random component.

The improvement of the KF is larger within the range of the power curve where power generation is most sensitive to wind speed. This fact shows that the KF approach is especially suitable for wind power prediction. And with the advantages of shorter training periods and faster adaption time than traditional statistical methods, the KF approach is suitable for short-term wind power prediction with the forecast period of 24–72 h.

Abbreviations
NWP: Numerical weather prediction; MOS: Model output statistics; KF: Kalman filter; WRF: Weather Research and Forecasting; PBL: Planetary boundary layer; RMSE: Root-mean-square error; CRMSE: Centered root-mean-square error; MAE: Mean absolute error; LST: Local standard time; UTC: Coordinated universal time.

Acknowledgements
Not applicable

Authors' contributions
JJX performed the study of the experiments and conducted the draft of the manuscript, ZNX the corresponding author, engaged in modifying the paper and submitted it to the PCMP. ZHL and ML checked the grammar and writing of the paper. All authors read and approved the final manuscript.

Availability of data and materials
The datasets generated and analyzed during the current study are not publicly available because the dataset of the target wind farm is confidential.

Declarations

Competing interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author details
[1]International Center for Climate and Environment Science (ICCES), Institute of Atmospheric Physics, Chinese Academy of Science, Beijing, China. [2]State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China. [3]Data61, Commonwealth Scientific and Industrial Research Organization (CSIRO), Perth, Australia.

References
1. Yan, C., Tang, Y., Dai, J., Wang, C., & Wu, S. (2021). Uncertainty modeling of wind power frequency regulation potential considering distributed characteristics of forecast errors. *Protection and Control of Modern Power Systems, 6*, 22.
2. Lange, M., & Focken, U. (2006). *Physical approach to short-term wind power prediction* (p. 208). Berlin: Springer.
3. Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H., & Feitosa, E. (2008). A review on the young history of the wind power short-term prediction. *Renewable & Sustainable Energy Reviews, 12*, 1725–1744.
4. Li, Z., Ye, L., Zhao, Y., Song, X., Teng, J., & Jin, J. (2016). Short-term wind power prediction based on extreme learning machine with error correction. *Protection and Control of Modern Power Systems, 1*, 1.
5. Cheng, X. L., Li, J., Hu, F., Xu, J., & Zhu, R. (2015). Refined numerical simulation in wind resource assessment. *Wind and Structures, 20*(1), 59–74. https://doi.org/10.12989/was.2015.20.1.059
6. Yang, Q., Berg, L. K., Pekour, M., Fast, J. D., Newsom, R. K., Stoelinga, M., & Finley, C. (2013). Evaluation of WRF-predicted near-hub-height winds and ramp events over a Pacific Northwest site with complex terrain. *Journal of Applied Meteorology and Climatology, 52*, 1753–1763.
7. Yim, S. H. L., Fung, J. C. H., Lau, A. K. H., & Kot, S. C. (2007). Developing a high-resolution wind map for a complex terrain with a coupled MM5 CALMET system. *Journal of Geophysical Research, 112*, D05106. https://doi.org/10.1029/2006JD007752
8. Siuta, D., West, G., & Stull, R. (2017). WRF hub-height wind forecast sensitivity to PBL scheme, grid length, and initial condition choice in complex terrain. *Weather and Forecasting, 32*, 493–503.
9. Mitchell, M. J., & Ancell, B. (2020). Configuration of statistical postprocessing techniques for improved low-Level wind speed forecasts in west Texas. *Weather and Forecasting, 35*, 129–147.
10. Hacker, J., & Rife, D. (2007). A practical approach to sequential estimation of systematic error on near-surface mesoscale grids. *Weather and Forecasting, 22*, 1257–1273.
11. Zamo, M., Bel, L., Mestre, O., et al. (2016). Improved gridded wind speed forecasts by statistical postprocessing of numerical models with block regression. *Weather and Forecasting, 31*(6), 1929–1945.
12. Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, 82*, 35–45.
13. Muller, M. D. (2011). Effects of model resolution and statistical postprocessing on shelter temperature and wind forecasts. *Journal of Applied Meteorology, 50*, 1627–1636.
14. Delle Monache, L., Nipen, T., Liu, Y., Roux, G., & Stull, R. (2011). Kalman filter and analog schemes to post-process numerical weather predictions. *Monthly Weather Review, 139*, 3554–3570.
15. Delle Monache, L., Nipen, T., Deng, X., Zhou, Y., & Stull, R. B. (2006). Ozone ensemble forecasts: 2—A Kalman-filter predictor bias correction. *Journal of Geophysical Research, 111*, D05308. https://doi.org/10.1029/2005JD006311

16. Galanis, G., Louka, P., Katsafados, P., Kallos, G., & Pytharoulis, I. (2006). Applications of Kalman filters based on non-linear functions to numerical weather predictions. *Annales Geophysicae, 24*, 2451–2460.
17. McCollor, D., & Stull, R. (2008). Hydrometeorological accuracy enhancement via post-processing of numerical weather forecasts in complex terrain. *Weather and Forecasting, 23*, 131–144.
18. Louka, P., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I., & Kallos, G. (2008). Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics, 96*, 2348–2362.
19. Cassola, F., & Burlando, M. (2012). Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output. *Applied Energy, 99*, 154–166.
20. Xu, J. J., Xiao, Z. N., & Lin, Z. H. (2020). Nonlinear Kalman filter bias correction for wind ramp event forecasts at wind turbine height. *Wind and Structures, 30*(4), 393–403.
21. Carreno-Madinabeitia, S., Ibarra-Berastegi, G., Saenz, J., et al. (2020). Sensitivity studies for a hybrid numerical-statistical short-term wind and gust forecast at three locations in the Basque country (Spain). *Atmosphere, 11*, 45.
22. Xu, J. J., Hu, F., Xiao, Z. N., & Li, J. (2013). Analog bias correction of numerical model on wind power prediction. *Journal of Applied Meteorological Science (in Chinese)., 24*(6), 731–740.
23. Cai, H. S., Jia, X. D., Feng, J. S., et al. (2020). Gaussian process regression for numerical wind speed prediction enhancement. *Renewable Energy, 146*, 2112–2123.
24. Chang, G. W., Lu, H. J., Chang, Y. R., et al. (2017). An improved neural network-based approach for short-term wind speed and power forecast. *Renewable Energy, 105*, 301–311.
25. Madhiarasan, M. (2020). Accurate prediction of different forecast horizons wind speed using a recursive radial basis function neural network. *Protection and Control of Modern Power Systems, 5*, 22.
26. Alessandrini, S., Delle Monache, L., Sperati, S., & Nissen, J. N. (2015). A novel application of an analog ensemble for short-term wind power forecasting. *Renewable Energy, 76*, 768–781. https://doi.org/10.1016/j.renene.2014.11.061
27. Chen, S. H., Yang, S. C., Chen, C. Y., et al. (2019). Application of bias corrections to improve hub-height ensemble wind forecasts over the Tehachapi wind resource area. *Renewable Energy, 140*, 281–291.
28. Song, J. J., Wang, J. Z., & Lu, H. Y. (2018). A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting. *Applied Energy, 215*, 643–658.
29. Xu, J. J., Hu, F., Xiao, Z. N., & Cheng, X. L. (2014). Bias correction to wind direction forecast by the circular–circular regression method. *Atmospheric and Oceanic Science Letters, 7*, 87–91. https://doi.org/10.3878/j.issn.1674-2834.13.0057
30. Dee, D. P. (2005). Bias and data assimilation. *Quarterly Journal of the Royal Meteorological Society, 131*, 3323–3343.
31. Skamarock, W. C., Klemp, J. B., Dudhia, J., et al. (2008). *A description of the advanced research WRF version 3*. NCAR technical note, 113pp.
32. Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*, 3rd edn. International Geophysics Series (Vol. 100, p. 704). Academic Press.
33. Schreck, S., Lundquist, J., & Shaw, W. (2008). *US Department of energy workshop report: Research needs for wind resource characterization*. NREL report TP-500-43521, 107 pp.