Theoretical and Applied Climatology (2020) 141:419–432
https://doi.org/10.1007/s00704-020-03201-8

ORIGINAL PAPER

A statistical method to construct wind speed at turbine height for study of wind power in China

X. Zhou[1] · J. Qin[1,2] · H. D. Li[3] · W. Tang[1,2] · X. Pan[1,2] · B. Huang[4] · X. Li[1,2]

Received: 11 July 2019 / Accepted: 29 March 2020 / Published online: 29 April 2020

Abstract

The accurate estimation of wind power is crucial not only to conduct wind energy research but also to improve policy and planning decisions for the national government. However, reliable wind speed datasets at turbine height that can be used for wind power estimation in China are very limited. In this study, the simulated wind profile (expressed by a power-law approach) is optimized using radiosonde wind speed measurements. The wind speed at a typical turbine height (80 m) is then constructed using the optimized wind profile and 10-m wind speed measured at each conventional meteorological station throughout the country. Compared with measurements, the constructed wind speed has reasonable errors (0.08 m s^{-1} in mean bias, 0.89 m s^{-1} in root mean square error, and 0.90 in correlation coefficient), which are much smaller than the model errors. Further comparison shows the constructing method in the current study outperforms the other similar statistical methods. Therefore, this method and the constructed wind speed dataset are suggested for wind energy research in China.

1 Introduction

The demand for energy and resources (e.g., electric energy, fossil oil, and coal) is increasing with the industrialization of society. Simultaneously, the greenhouse gases' emission is increasing because of the increasing consumption of conventional fuels, which has led to intensifying global warming in recent decades. Wind is an important clean energy source and has intrinsic superiority: It is a sustainable energy and easy to exploit. It can meet the demand for carbon dioxide emissions reductions over the globe. To meet the increasing demand for clean energy to keep up with fast society's industrialization and alleviate global warming, the investigation and estimation of wind energy are drawing increasing attentions from the scientists, industrialists, and governmental policymakers.

This work focuses on wind power study in China. The evaluation of China's wind resources plays a central role in policy and planning decisions in the country, and therefore, wind power requires accurate estimations. However, China has a vast territory with multiple climatic and geological conditions, which indicates both great potentials, and also difficulties, for wind power exploitation. One of the main difficulties may originate from basic statistical uncertainties in wind speed, which is the a key variable for calculating the wind power. Evaluations and projections of China's wind resources differ in the choices of wind speed datasets [Davidson et al. 2016]. Therefore, a reliable wind speed dataset is necessary for wind energy studies, which will further support the energy dispatching and planning by the national government.

In situ measurements and numerical simulations are widely used methods to quantitatively estimate the wind power. Zhang and Li [2012] estimated the wind power distribution

Key points
- This work proposes a new statistical method to construct wind speed at turbine height.
- The constructed wind speed has obviously reduced the errors than direct model output and previous similar work.
- This work has potential to provide a more accurate statistical basis for further application of wind speed on wind energy study in China.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00704-020-03201-8) contains supplementary material, which is available to authorized users.

✉ X. Zhou
xuzhou@itpcas.ac.cn

[1] National Tibetan Plateau Data Centre, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China

[2] University of Chinese Academy of Sciences, Beijing 100049, China

[3] CAS Key Laboratory of Forest Ecology and Management, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China

[4] Department of Energy and Process Engineering, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway

in China using the wind speed measured at conventional meteorological stations. However, the wind turbines are generally higher than 60 m, and thus, the applicability of their studies may be limited. Besides, there are very few measurements at turbine height that is free and can be used for model evaluation and the investigation of wind power in China. Wind speed in numerical simulation may be an alternative for estimating the potential wind power. For example, the wind power distribution map for China from the Energy Research Institute (http://www.cnrec.org.cn/cbw/fn/2014-12-29-459.html) was created based on numerical simulation. However, the simulated wind speed may have systematic biases [Carvalho et al. 2014; Deppe et al. 2013; Li et al. 2018b] owing to the intrinsic disadvantages of numerical models (e.g., errors owing to the physical approximation during numerical discretization, and errors caused by the model resolution and sub-grid parameterizations). McElroy et al. [2009] and Davidson et al. [2016] used the wind speed from the Goddard Earth Observing System (GEOS) and atmospheric data assimilation system (ADAS) to estimate the density of wind power in China. However, this system did not assimilate the observed 10-m wind speed at conventional meteorology stations [Gelaro et al. 2017], which is similar in other reanalysis datasets (such as ERA-Interim [Dee et al. 2011]. Therefore, the near-surface wind speed in these datasets may also have systematic errors [Li et al. 2018b].

Considering the above shortcomings in estimating wind power, developing more reliable wind speed datasets at turbine height is necessary. In this paper, we propose a statistical method (Section 3) to construct wind speed at 80 m (a typical wind turbine height) by combining the 10-m measurements, radiosonde data, and numerical simulation results. The goal of this work is to develop a daily wind speed dataset at a given wind turbine height that can be used directly for model evaluation and studies of wind power in China. The constructed wind speed would provide an important basis for creating a high-resolution-gridded wind speed dataset at turbine height.

2 Data and model

2.1 Data

Two datasets of measured wind speed are used in the current study. One is a set of hourly 10-m wind data at 729 weather stations belonging to the China Meteorology Administration (CMA). An anemometer was used for wind speed measurement at these conventional stations. Although the 10-m station data cannot be used directly to estimate the wind energy in China, it can provide some basic information that can be used to calculate the wind speed at turbine height. Meanwhile, the 10-m data can also be used in the model evaluation. Another dataset is wind speed, air temperature, and air pressure measured at 89 radiosonde stations, which can be downloaded from the National Ocean and Atmospheric Administration (NOAA) and Integrated Global Radiosonde Archive (IGRA) website [Durre et al. 2006]. The IGRA dataset was measured twice a day at approximately 11:45 and 23:45 universal time at each station. The wind speed at different heights was detected using the Global Positioning System (GPS) displacement of the balloon. The geographical locations of CMA stations (black dots) and radiosonde stations (black triangles) are shown in Fig. 1.

The 10- and 100-m wind speed data in the ERA5 global reanalysis dataset from ECMWF (European Centre for Medium-Range Weather Forecasts) based on the IFS model [ECMWF 2017] has also been used in the current study. It has a spatial resolution of approximately 30 km (0.25°) and temporal resolution of 1 h. The 10- and 100-m wind speed at ERA5 is forecasting variables and can be treated as simulation results. The wind speed datasets have been used to construct the wind speed at 80-m turbine height and compared with that using the wind speed in the WRF model.

2.2 Model simulation

2.2.1 Model introduction and configuration

The numerical model used in the current study is the Weather Research and Forecasting (WRF) model, which is developed by the National Center for Atmospheric Research (NCAR). The model is widely used not only in meteorological [Gao et al. 2011, 2015] and climatic studies [Li et al. 2018a, 2019; Wang et al. 2016] but also in wind power research from many aspects, e.g., making decisions on the choices of land surface models [Wharton et al. 2015], planetary boundary layer (PBL) schemes [Deppe et al. 2013; Krogsaeter and Reuder 2015; Wang and Jin 2014], sensitivity to initial/lateral conditions [Carvalho et al. 2014; Deppe et al. 2013], and wind speed or wind power estimations based on further corrections by fuzzy systems [Zhao et al. 2016, 2017].

In the current study, we conduct a 1-year simulation in 2014 for China with the WRF model version 3.7 (Fig. 1). The horizontal resolution is 0.1° (approximately 10-km), at which resolution the model can resolve large-scale topographic variation. A turbulent orographic form drag scheme [Beljaars et al. 2004] is switched on to represent the obstacle effect by the sub-grid complex terrain. Model configurations with the consideration of sub-grid orography are demonstrated to have more reasonable simulation results for wind speed [Jimenez and Dudhia 2012; Lee et al. 2015; Zhou et al. 2017]. The implementation of this scheme is particularly effective in reducing positive bias in modeled wind speed over mountainous regions [Zhou et al. 2018, 2019]. The total number of vertical levels is 37, from the surface to the top of the atmosphere (set to 50 hPa). The model is initialized and driven

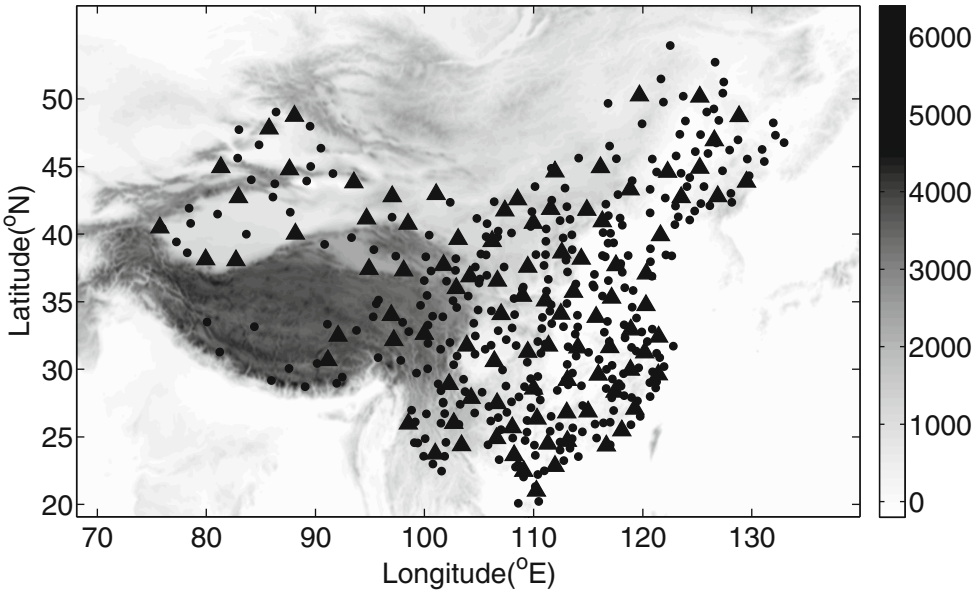

Fig. 1 Topography (m; gray shaded) of study region and geographic locations of CMA stations (dots) and IGRA radiosondes (triangles)

(for upper and lateral boundaries) by the ERA-Interim reanalysis [Dee et al. 2011] at 6-h intervals. The sea surface temperature is also from ERA-Interim and updated at the same intervals. All the simulations are configured with the Noah land surface scheme [Mahrt and Pan 1984; Pan and Mahrt 1987], the Yonsei University PBL scheme [Hong et al. 2006], and the rapid radiation transfer model (RRTM) scheme [Mlawer et al. 1997] for longwave and solar radiative transfer. The model simulation results have been bi-linearly interpolated to the station locations during the data processing.

2.2.2 Model performance

Wind speed is the basic key parameter used in the current study. The simulated 10-m wind speed ($V10$) was therefore evaluated using CMA station wind speed measurements.

Figure 2 shows the scatter plots of seasonal mean wind speed at each station. The letters on each graph represent the months of the year (spring, MAM; summer, JJA; autumn, SON; winter, DJF). The error is largest in spring with a larger bias (0.77 m s^{-1}) and root mean square error (RMSE; 1.36 m s^{-1}) and a lower correlation coefficient (CORR; 0.46) than in the other seasons. The biases in the other seasons are similar, in the range of 0.59–0.66 m s^{-1}. The RMSE is smallest in summer (1.15 m s^{-1}) and the CORR is highest in winter (0.61).

Figure 3 gives the statistical metrics at each station (points in Fig. 1) derived from daily wind speed: mean bias, RMSE CORR, and mean relative bias. Systematic overestimation of the 10-m wind speed occurs over the study domain. The errors are larger for Southeast China with larger positive bias, RMSE, and relative bias and lower correlation coefficients than for other regions. The wind speed at stations over the Middle Eastern Tibetan Plateau has the smallest biases and RMSE values.

In general, systematic errors exist in the simulated 10-m wind speed simulated by the WRF model, which is consistent with the findings of other studies [Carvalho et al. 2014; Deppe et al. 2013; Li et al. 2018b] although those studies covered different regions. Nevertheless, their studies also demonstrated that the representation of large-scale features of wind speed by the model is generally reasonable by resolving large-scale climatic and orographic processes, which will be beneficial for the current work.

The purpose of engaging the WRF model is to gain the large-scale features of the wind profile, which could be associated with atmospheric circulation. This profile is then optimized using nearby measurements. The optimized wind profile is expected to be closer to the reality than that in the WRF model and is used to construct the wind speed at turbine height together with 10-m wind speed measured at conventional meteorological stations.

3 Method

Now, we have observed 10-m wind speed and radiosonde wind speed. How to combine the two to derive the wind speed at turbine height? The idea is to employ a numerical simulation. In this section, we give the details of the statistical method in the following sub-sections.

3.1 Wind profile expressed by power-law approach

The power-law approach is widely used to express the relationship between the wind speed at a certain height and that at

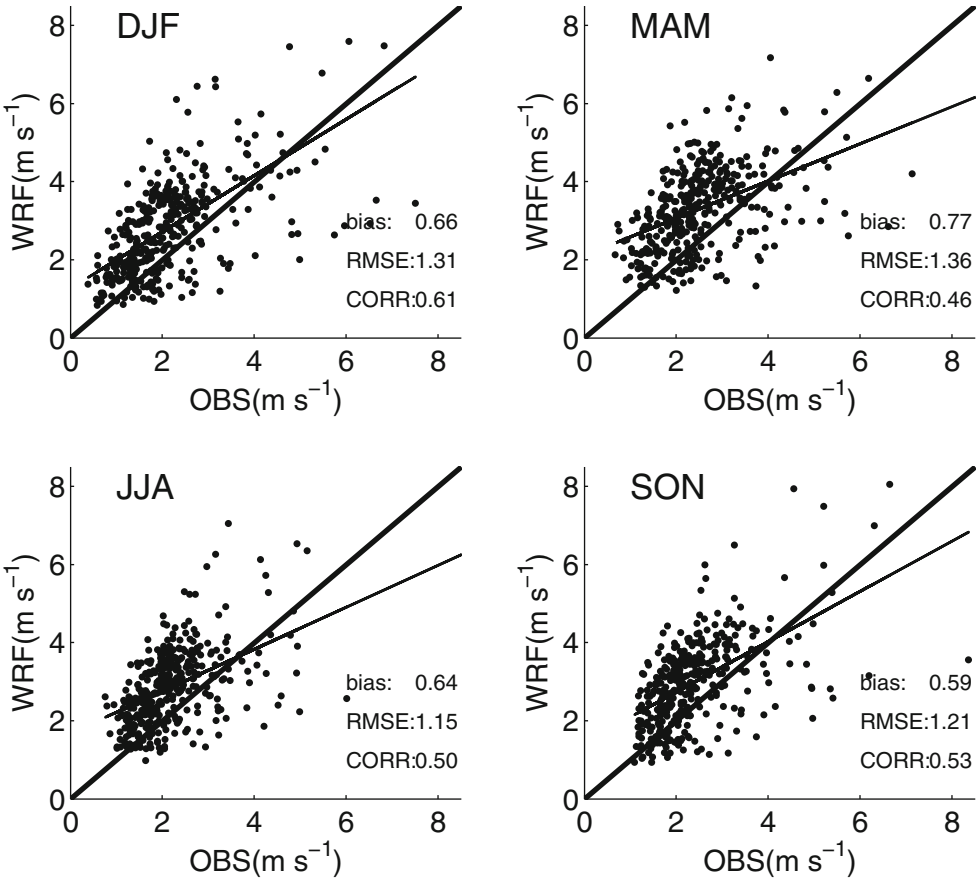

Fig. 2 Scatter plots of seasonal mean wind speed (m s^{-1}) at station location derived from WRF simulation and 10-m measurements, and the statistical metrics of mean bias (m s^{-1}), root mean square error (RMSE; m s^{-1}), and correlation coefficient (CORR)

a reference height [Archer and Jacobson 2003, 2005; Li et al. 2010], as shown by the following equation:

$$V_z = V_r\left({}^{z}/_{z_r}\right)^{a} \quad (1)$$

where V_z is the wind speed at a certain height; z and V_r is the wind speed at reference height z_r. This equation is applicable for neutral flow [Tobin et al. 2015]. The index a is a key parameter that generally represents the wind profile at a given

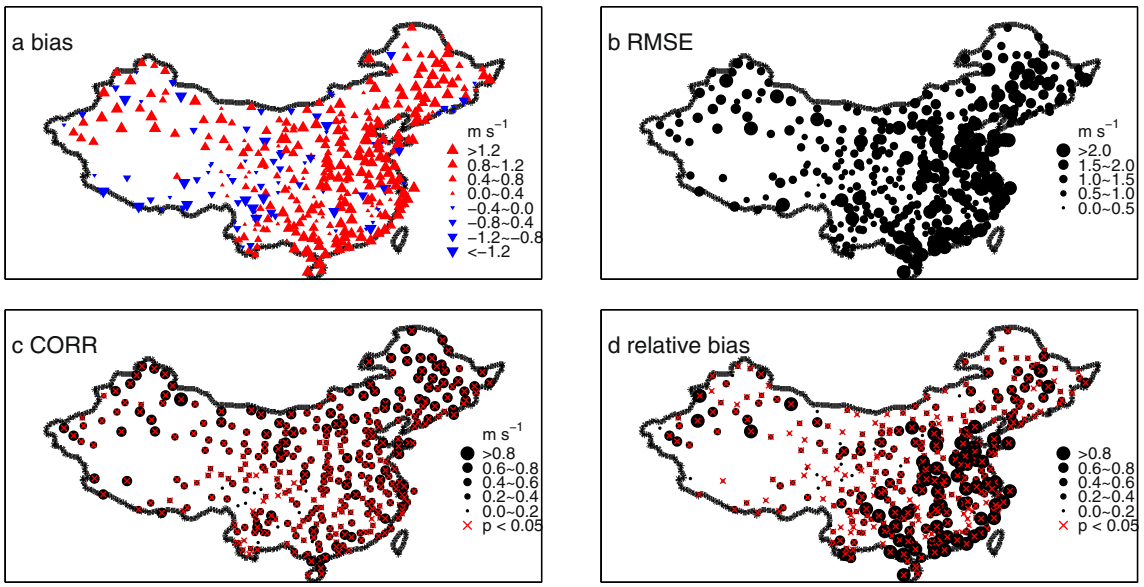

Fig. 3 Statistical metrics in simulated 10-m wind speed. **a** Mean bias (m s^{-1}), **b** RMSE (m s^{-1}), **c** correlation coefficient (CORR), and **d** mean absolute relative bias (m s^{-1}), derived from WRF versus measurements at each station based on daily mean 10-m wind speed; Red cross indicates significant test at 95% level ($p < 0.05$)

location. It is time-varying and can be easily computed from the model by least square method using the simulated wind speed. Statistically, the value of a depends on the geographical location and the local surface condition.

We assume that the WRF model can generally and reasonably represent large-scale features of the wind profile by resolving large-scale climatic and orographic processes. The index a derived from the WRF model (a_{WRF}) may have local errors and should be optimized before being used for the construction of wind speed at 80-m height at station locations where 10-m wind speed has been measured. The wind speed at 80 m can then be derived as follows:

$$V_{80} = V_{10\_OBS}\left({}^{80}/_{10}\right)^{a_{OPT}} \tag{2}$$

where V_{80} represents the constructed wind speed at 80 m that we need, V_{10_OBS} is the measured 10-m wind speed at CMA station, and a_{OPT} is the optimized a_{WRF}, which is developed using station and radiosonde measurements. The optimization procedure is described in Section 3.3.

3.2 Processing the radiosonde data

IGRA radiosonde data (triangles in Fig. 1) is used to optimize a_{WRF}. Radiosonde data are obtained from a detector carried by a balloon. The data are then transferred back to the receiver in the form of radio waves. It does not measure precisely at 10 and 80 m above the surface. Instead, it measures near the surface (level 1) and at upper pressure levels (here we take the second level as level 2). Besides, the IGRA radiosonde has not recorded the height at each level. Instead, it has records for air temperature, pressure and humidity. Therefore, the measurement heights at the two lowest levels, the measured wind speed at 80-m and the index a based on radiosonde data (a_{OBS}) can be estimated using the following three steps.

First, the height difference ($\Delta h = h_2 - h_1$) between the two levels is calculated based on the measured air temperature and pressure by combining the ideal gas law (Eq. (3)) and hydrostatic equilibrium equation (Eq. (4)):

$$P = \rho RT \tag{3}$$

$$\left.{}^{(P_2-P_1)}\right/_{(h_2-h_1)} = -\rho g \tag{4}$$

where P_2 and P_1 are the air pressures at levels 2 and level 1, respectively; h_2 and h_1 are the heights above surface at levels 2 and level 1, respectively; $\rho$is the air density; g is the gravity constant; $R$is the gas constant; T_1 and T_2 are the air temperatures (K) at level 1 and level 2. The air densities at the two levels are assumed to be the same. $P = (P_2 + P_1)/2$and $T = (T_1 + T_2)/2$ are the averaged pressure and temperature, which are used in Eq. (3). Note that Eq. (4) is widely applicable for the atmosphere except for strong convective conditions.

Second, there is 10-m wind speed measurement at each CMA station that corresponds to a radiosonde station. The following relationship can be derived based on Eq. (1):

$$V_1 = V_{10}\left({}^{h_1}/_{10}\right)^{a_{OBS}} \tag{5}$$

$$V_2 = V_1\left({}^{h_2}/_{h_1}\right)^{a_{OBS}} \tag{6}$$

where V_2 and V_1 are wind speed at level 2 and level 1 from radiosonde. Combining Eqs. (5), (6), and $h_2 = h_1 + \Delta h$, a_{OBS}, h_2, and h_1 can be numerically derived based on an iteration method. Two aspects will cause errors during the calculation and cannot be eliminated by statistical methods: (1) the CMA measurement and the corresponding radiosonde measurement have a 10–20-min time difference, and (2) the measurement methods and instruments are different between CMA station and radiosonde station (anemometer versus GPS-detected displacement of balloon).

Finally, reasonable measured data samples are selected that could be used to optimize and verify index$a$. IGRA measurements at level 1 lower than 15 m were excluded because the balloon may take a moment to stabilize after release and the measured wind speed is more reliable at higher elevations than at lower ones. A threshold of 800 m was used at level 2, considering two aspects: (1) a sufficiently large set of measured samples is needed to statistically verify the robustness of the optimization, and (2) level 2 should not be too high, such that Eq. (1) remains applicable (the reference height is 80 m, which is eight times higher than surface observations and the value 800 m is ten times higher than 80 m). Sensitivity tests show that the statistical biases and RMSEs in a_{OPT} and the constructed wind speed are not sensitive to the selected height thresholds (e.g., when selecting 10 m or 15 m at level 1 and 600 m, 800 m, or 1000 m at level 2; not shown). The wind profiles from the near surface to this threshold of 800 m are most relevant to the wind power study.

3.3 Optimization of index *a*

Wind profiles represented by index a have particular geographical characteristics. It is generally accepted that the nearer the two stations the more similar wind characteristic they have [Achberger et al. 2006], which provides a scientific basis for optimization of the simulated wind profile using the nearest radiosondes, and can be further extended to each surface station and each hour. Additionally, assimilation of observed wind speed can effectively improve the near-surface wind speed forecasting [Zhang et al. 2015], which also supports our work that the construction of wind speed could be improved by statistically optimizing the simulated wind profile based on observations. In this section, a least square method has been used to optimize the index a at each CMA station using the measurements from the nearest five radiosondes (the

number of five is selected based on the significant test during the linear regression to make sure that the p value at each station is smaller than 0.05). The available pairs of observations and simulations at these 5 radiosondes are used for optimization.

On the assumption that the wind speed profile follows the rule of Eq. (1), the two terms ($a_{WRF} - a_{OBS}$ and $\log(V_{10\_WRF}/V_{10\_OBS})$) should have a linear relationship (see Appendix A for a detailed deduction):

$$a_{WRF} - a_{OBS} = A \times \log(V_{10\_WRF}/V_{10\_OBS}) + B \quad (7)$$

Assuming that A and B do not vary with time, they can be derived from the above equation using the least square regression for each CMA station using the available observation samples at its nearest five radiosonde stations. The a_{WRF} can then be optimized accordingly at each hour:

$$a_{OPT} = a_{WRF\_h} - A \times \log(V_{10\_WRF\_h}/V_{10\_OBS\_h}) + B \quad (8)$$

where a_{WRF_h} is the wind profile index (at hourly) derived from the WRF model; $V_{10\_WRF\_h}$ and $V_{10\_OBS\_h}$ are the hourly wind speed from the WRF model and CMA station measurements, respectively.

The correlation coefficients between the two terms are calculated for each CMA station based on the measurements from its nearest five radiosonde stations. The coefficients are generally high (0.5–0.7) with statistical significance greater than 95% (not shown). Therefore, we think that the relationship between the two terms is physically reasonable and statistically reliable, such that it can be used for the optimization of a_{WRF} for the CMA stations by the least square method. Note that a_{OBS} in Eq. (7) is calculated based on all selected samples at each radiosonde station in Section 3.2. It is used to empirically calculate the time-invariant parameters A and B. The a_{OPT} is then calculated at each hour for each conventional CMA stations based on Eq. (8). Consequently, 80-m wind speed at each hour at these CMA stations can be derived from Eq. (2).

4 Results

4.1 Validation with radiosonde data

The a_{OPT} is validated at each radiosonde station before being further used to construct 80-m wind speed. Figure 4 shows the scatter plots of a_{WRF} and a_{OPT} against a_{OBS}. For a fair comparison, the cross-validation is adjusted to a_{OPT}: each a_{OPT} is derived considering the wind speed at the five nearest radiosonde stations without considering the measurement at itself. The statistical metrics in Fig. 4 show clear improvements by the above optimization with reduced mean bias (from − 0.087 to 0.0006) and RMSE (from 0.18 to 0.11). a_{OPT} is then used for constructing 80-m wind speed using measured 10-m wind speed with Eq. (2).

The simulated and constructed 80-m wind speed is also compared with radiosonde observations. Figures 5a–c show the scatter plots of the WRF-simulated 80-m wind speed, the 80-m wind speed constructed with a_{WRF}, and the 80-m wind speed constructed with a_{OPT} against the IGRA measurements. Figure 5a shows that the wind speed in WRF has large errors. Thus, its applicability on the wind power study is quite limited. Figure 5b is 80-m wind speed constructed based on a 10-m observation and a_{WRF}. The errors are reduced but still considerable, which indicates that a_{WRF} needs to be optimized. Statistical metrics show that the wind speed constructed with a_{WRF} is obviously better than the simulated one, with a smaller mean bias (− 0.29 m s^{-1}) and RMSE (1.24 m s^{-1}), and a higher CORR (0.85). The optimization further reduces the bias (0.08 m s^{-1}) and RMSE (0.89 m s^{-1}), and increases the CORR (0.90) (Fig. 5c). These comparisons indicate that the error in the optimized wind profile and the constructed wind speed at 80-m turbine height can be much reduced at the stations without radiosonde measurements.

The wind speeds at daytime and nighttime show different characteristics in magnitude and direction, which is proved by observation and simulations [Hu et al. 2013; Klein et al. 2016]. Thus, the proposed statistical method is further tested at two typical times when the wind speed is observed at each radiosonde: 06:00 and 18:00 local time. Evaluations show that the optimization can also effectively reduce the biases and RMSEs as well as increase the CORRs in constructed wind speed (not shown). It is also worth noting that the power-law approach as shown in Eq. (1) is only an empirical relationship. The accuracy of our results may therefore be limited by its precision. For example, computational fluid dynamics (CFD) modeling can provide more accurate wind speed simulation results and is widely used for wind speed and power forecasting. However, CFD modeling would be too expensive and its application to the entire region of China may not be practically feasible. Nevertheless, the proposed construction procedure in this paper is easy to follow and can be applied to a large area (e.g., countrywide). The above validation shows that the errors in the constructed wind speed at 80-m turbine height are reduced significantly. Therefore, this work is expected to be beneficial for studies such as those of McElroy et al. [2009] and Davidson et al. [2016].

4.2 The constructed wind speed at 80 m

The 80-m wind speed constructed at each CMA station is shown in Fig. 6a. The wind speed is higher along north Xinjiang-Mongolia and along the coastline in Southeast of China than in the rest of the country, which is consistent with other studies [Feng et al. 2015; Liu and Kokko 2010; Sun et al. 2015].

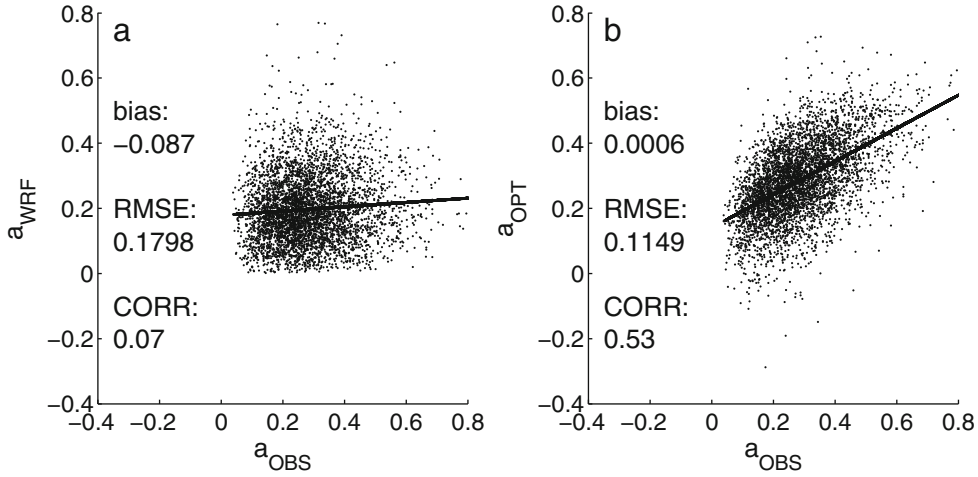

Fig. 4 Scatter plots of (**a**) index *a*in WRF model (a_{WRF}) and (**b**) optimized index (a_{OPT}) versus measurement

Furthermore, Fig. 6b shows the difference between the WRF-simulated 80-m wind speed and the constructed one. The difference is consistent with the 10-m wind speed bias (Fig. 3a) that the wind speed from the WRF model is obviously higher over the regions where the simulated 10-m wind is also obviously higher (e.g., over the south, east, and northeast regions of China). The logic is that the construction is expected to be smaller than the simulated 80-m wind speed at stations where the 10-m wind speed is overestimated in the model (i.e., at stations where the 10-m wind speed is overestimated, the 80-m wind speed is also expected to be overestimated by the WRF model). This consistency indicates that the constructed wind speed data using this statistical procedure is reasonable.

4.3 Application to wind power study

In this section, we provide a simple analysis by applying the constructed 80-m wind speed to a wind power study. To save computational resources, this analysis was carried out based on daily mean of the 80-m-constructed wind speed. The wind speed can be classified based on its magnitude (Table 1). According to the previous studies, wind speed in class 3 and higher is considered as suitable for wind power generation (and considered marginal in class 2) [Archer and Jacobson 2003, 2005; Li et al. 2010].

Figure 7 shows the number of marginal days and suitable days that can be used to produce wind energy by wind turbine. They are derived from both the measured 10-m wind speed and the constructed 80-m wind speed. The spatial distributions of the number of marginal days (Fig. 7a versus Fig. 7c) and suitable days (Fig. 7b versus Fig. 7d) generally agree between the two datasets. However, the numbers are higher when derived from the constructed 80-m wind speed. The differences between the two could be associated with the classification methods: the classification thresholds for 80-m are simply derived from the 10-m classification thresholds based on Eq. (1) with$a = 1/7$. However, statistical calculation based on radiosonde observation shows that the index a is much higher (0.23–0.26 for the global average) [Archer and Jacobson 2005]. Our results are consistent with those findings: the mean $a_{WRF}$is 0.21 and the mean a_{opt} is 0.25 in the current study. Thus, it is natural to speculate that the construction results are within reasonable scales. Therefore, the optimized wind

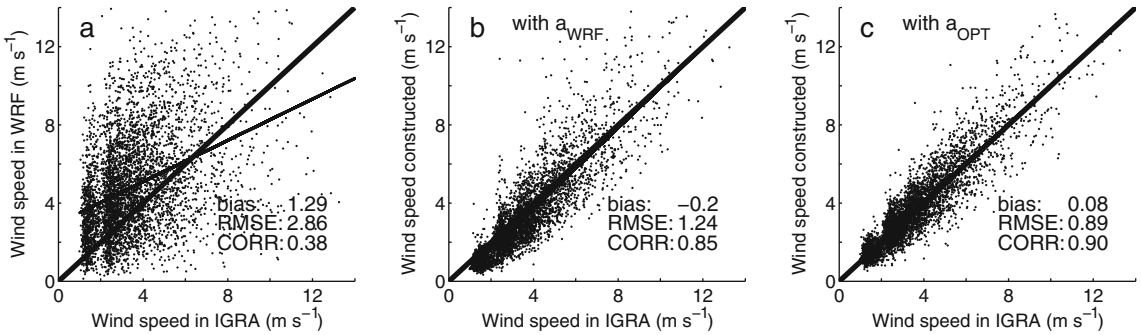

Fig. 5 Scatter plots of (**a**) WRF simulated 80-m wind speed (m s^{-1}), (**b**) 80-m wind speed (m s^{-1}) constructed using a_{WRF}, and (**c**) 80-m wind speed (m s^{-1}) constructed using a_{OPT} versus radiosonde measurement

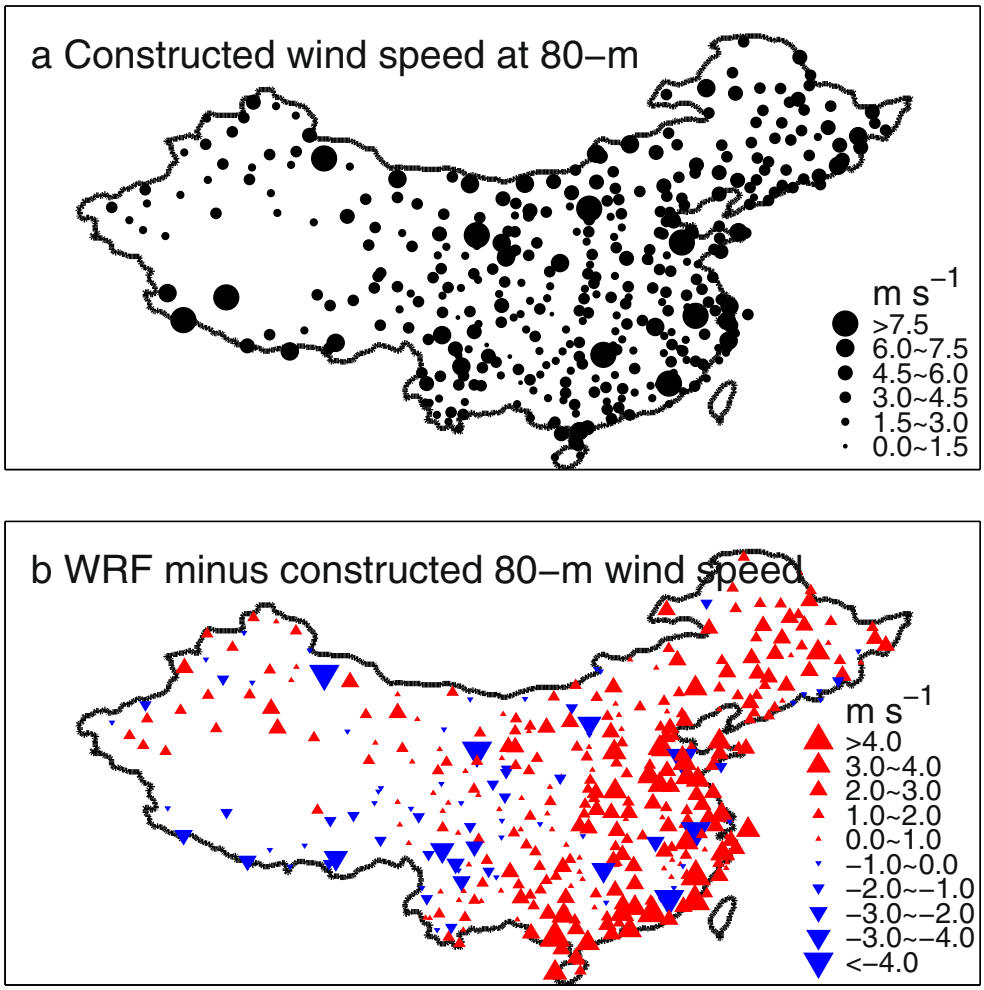

Fig. 6 **a** Mean constructed wind speed (m s^{-1}) at 80 m using a_{OPT} and **b** difference between the WRF and constructed wind speed

profiles and the constructed wind speed are expected to be reliable and are recommended for use in model evaluations and wind power investigations.

The wind speed at both 10 and 80 m is widely used for the classification of wind speed by the empirical method. However, wind power statistics are more relevant for energy estimation. Because the turbine blades are generally located at a specific height above the surface, e.g., 80-m in the current study, therefore, only the wind speed at turbine height can be used to calculate the wind power, which is further used for wind energy estimation. Hence, the potential wind power can be calculated based on daily mean constructed 80-m wind speed with the following empirical equation [Archer and Jacobson 2003],

$$P = \frac{1}{2}\frac{6}{\pi}\rho V^3 \tag{9}$$

where P is the power in W m^{-2} day^{-1}, and V is the scalar daily mean wind speed. The power of one turbine needs to be multiplied by the rotor area for further calculations. The applicability of this equation has been demonstrated particularly when the wind speed is high (the mean wind speed that is suitable for wind power production) with the assumption that the wind speed follows a Rayleigh distribution [Archer and Jacobson 2003].

Figure 8 shows the exploitable potential wind energy and highly exploitable potential wind energy based on the Eq. (9). The former is calculated by adding wind power derived from both the marginal days and suitable days. The latter is derived from suitable days only. Figure 8a–b show that the number of stations that are suitable for wind power development and have the potential of producing energy more than 2000 kWh m^{-2} in the study year is limited. Most of these

Table 1 Wind speed (m s^{-1}) classifications (following Archer and Jacobson 2003)

Wind class	1	2	3	4	5	6	7
10-m wind	< 4.4	4.4–5.1	5.1–5.6	5.6–6.0	6.0–6.4	6.4–7.0	> 7.0
80-m wind	< 5.9	5.9–6.9	6.9–7.5	7.5–8.1	8.1–8.6	8.6–9.4	> 9.4

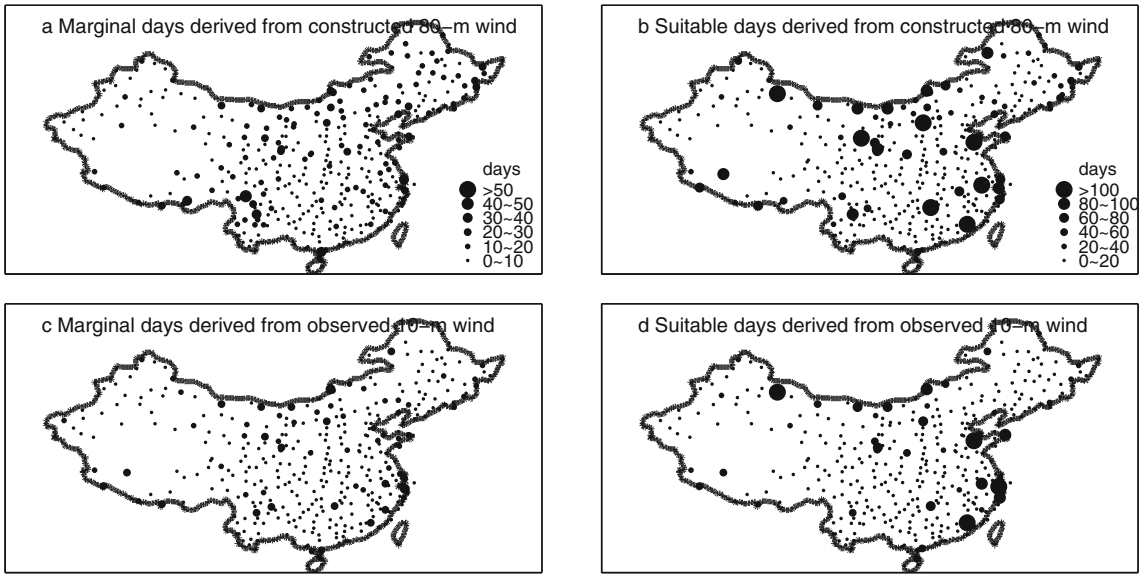

Fig. 7 Number of marginal days and suitable days accumulated in the study year, derived from daily mean of constructed 80-m wind speed (**a**–**b**) and measured 10-m wind speed (**c**–**d**)

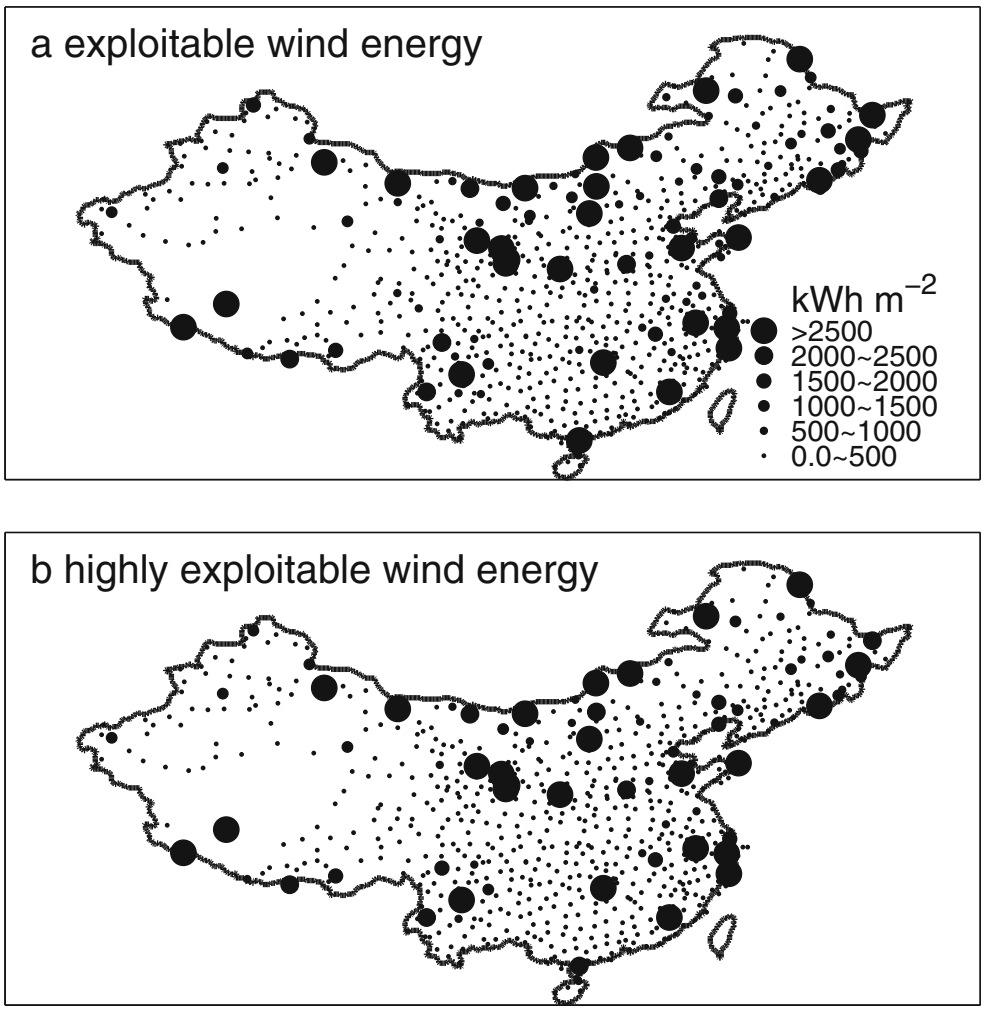

Fig. 8 **a** Exploitable potential wind energy (kWh m^{-2}) and **b** highly exploitable potential wind energy (kWh m^{-2}) derived from constructed 80-m wind speed

stations are located in Xinjiang-Mongolia and along the coastlines of Southeast China and Northeast China. Locations at coastlines suffer extreme weather conditions (e.g., tropical cyclones), which may have some negative impacts on turbine construction. The region of Xinjiang-Mongolia appears to be a good choice. A few stations are located at the Himalaya Mountains, where harsh geological conditions and some special treatments are required for the wind turbine installation over this region [Duan 2008]. Besides, the wind power should be low at these stations due to small air density. However, the wind power derived from the constructed wind speed are high for specific stations in the Tibetan Plateau (Fig. 8), which shows disagreement with Jiang et al. [2010]. The reasons could be as follows: (1) these stations are located at the boundary of the Plateau, where the thermal wind is very strong at day-time; and (2) the radiosonde stations are very sparse and located far away in this region, which will cause large errors in optimizing the wind profile as well as the wind speed construction.

The differences between highly exploitable potential wind energy derived from the WRF model and that derived from the constructed wind speed are shown in Fig. 9. Obviously, the wind power can differ significantly while the difference in wind speed (Fig. 6b) is moderate, which indicates that the wind power estimation is very sensitive to the accuracy of the selected wind speed. The differences in highly exploitable potential wind energy are generally consistent with those in the wind speed except for the areas along north Xinjiang-Mongolia. The reason could be that the former is the sum of wind energy at suitable days based on daily mean wind speed, while the latter is only an average of all days. This result emphasizes the need for highly accurate wind speed data for wind power assessment and highlights the importance of the current study.

This section only gives a very preliminary example on the application of the constructed wind speed. Statistically, the calculation of these parameters (e.g., marginal days, suitable days, and the energy) could have considerable uncertainties due to the length of the data (only limited to 1 year). Thus, developing a longer time dataset and more thorough analysis are planned for further work. Additionally, two disadvantages during the data processing described in this section should be noted. First, the marginal days and suitable days shown in Fig. 7 are calculated based on daily mean wind speed, such that some days that have suitable hours for wind energy production may be excluded. Second, the wind speed is highly sensitive to the very nearby small-scale topography. This work is carried out from the statistical aspect concerning the entire region of China. Thus, there might be obvious spatial heterogeneity exist. For example, individual points with very good resources can be surrounded by multiple points with marginal resources although the topography looks similar. Nevertheless, this work has generally provided a more accurate estimation of wind speed data at turbine height over the country without spending too expensive computing resources compared with CFD modeling.

5 Discussion

One origination of errors in this method is the spatial heterogeneity, i.e., the mismatch between observation and model. The representative of observed 10-m wind speed is different from that in a model: single point versus mean in a grid box. Another origination of errors is the application of the power-law approach. As described in Section 3.1, the power-law approach is only applicable to neutral flow without considering the atmospheric stability, which will cause errors in the

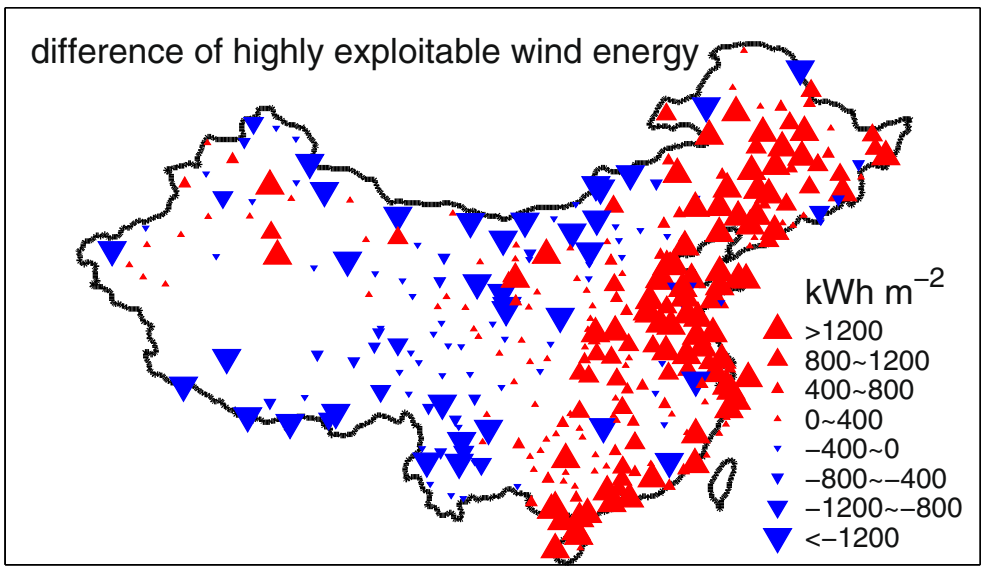

Fig. 9 Differences of highly exploitable wind energy (kWh m^{-2}) derived from the WRF minus that calculated from constructed 80-m wind speed

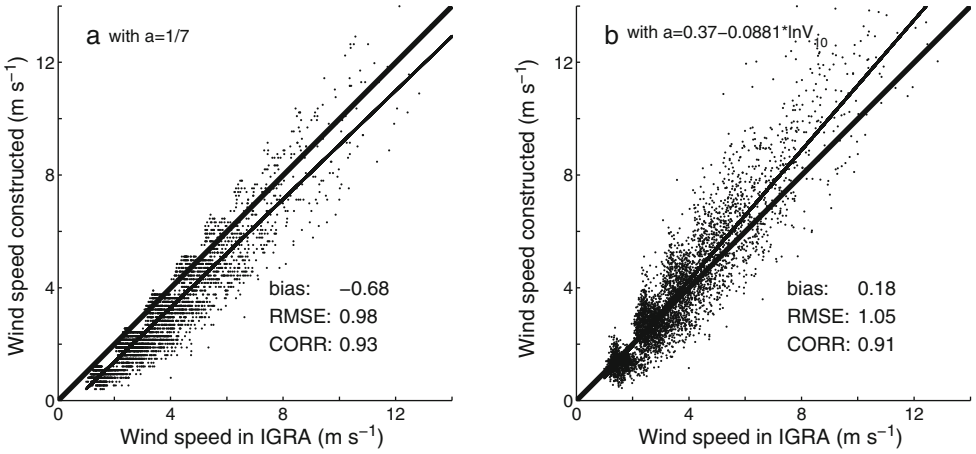

Fig. 10 Similar as Fig. 5b–c but for the 80-m wind speed constructed with (**a**) $a = 1/7$and (**b**) $a = 0.37 - 0.0881 \bullet \ln V_{10}$, respectively

current work. However, this approach is still a good choice in most current studies focusing on the large area [Liu et al. 2018; Tobin et al. 2015], such as China region. There are many other statistical methods to derive the wind speed at turbine height. In the current study, only the work based on power-law approach will be discussed.

Based on the power-law approach, another two empirical methods have been used to interpolate the wind speed from 10 m up to turbine height in previous studies. One is commonly used with $a = 1/7$ (e.g., used by Archer and Jacobson 2003, Bichet et al. 2012, Tobin et al. 2015). The other one is established by Justus and Mikhail [1976] with $a = 0.37 - 0.0881 \times \ln V_{10}$ and has been used in the most recent studies from Tobin et al. [2015] and Liu et al. [2018].

Therefore, we also validate the 80-m wind speed, constructed using the above two methods, against radiosonde measurements to make a comparison with the statistical method proposed in the current study. Figure 10 shows the scatter plots of the constructed 80-m wind speed (a) using the above two methods. Compared with radiosonde measurements, the constructed 80-m wind speed using above two methods has systematic biases (Fig. 10a–b): -0.68 m s^{-1} and 0.19 m s^{-1}, while the constructed wind speed (in the current work) using a_{OPT} has much smaller mean bias (0.08 m s^{-1}; Fig. 5c). Furthermore, the constructed wind speed using a_{OPT} has smaller RMSE (0.89 m s^{-1}) than that using $a = 1/7$ (0.98 m s^{-1}) and $a = 0.37 - 0.0881 \times \ln V_{10}$ (1.05 m s^{-1}), although the correlation coefficients are slightly smaller (0.90 versus 0.93 and 0.91). These statistical metrics show that the procedure to construct the wind speed in the current work generally outperforms the two used in the previous two studies, which demonstrates the improvements and novelty of the current study.

Noting that the above two methods are independent of numerical simulation. It is worth to know the dependence of the current method on the model. Consequently, we select ERA5 hourly wind speed data to construct the 80-m wind speed. Figure 11 shows the scatter plots of a_{ERA} (wind speed profile index in ERA5) and a_{OPT}(based on ERA5) against a_{OBS}. In comparison with Fig. 4, the wind profiles in the two simulated show difference performances when evaluated with observation (RMSE 0.1526 versus 0.1798; Fig. 11a versus Fig. 4a), while performances of the optimized wind profile are quite close (RMSE 0.1125 versus 0.1149; Fig. 11b versus Fig. 4b). Furthermore, the performances of simulated 80-m wind speed also show considerable differences (RMSE 2.06 versus 2.86; Fig. 12a versus

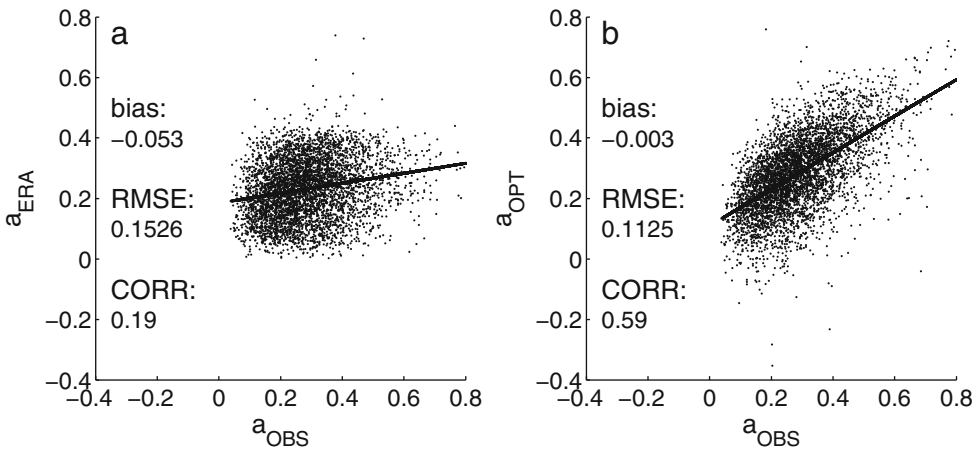

Fig. 11 Similar as Fig. 4a–b but based on ERA5

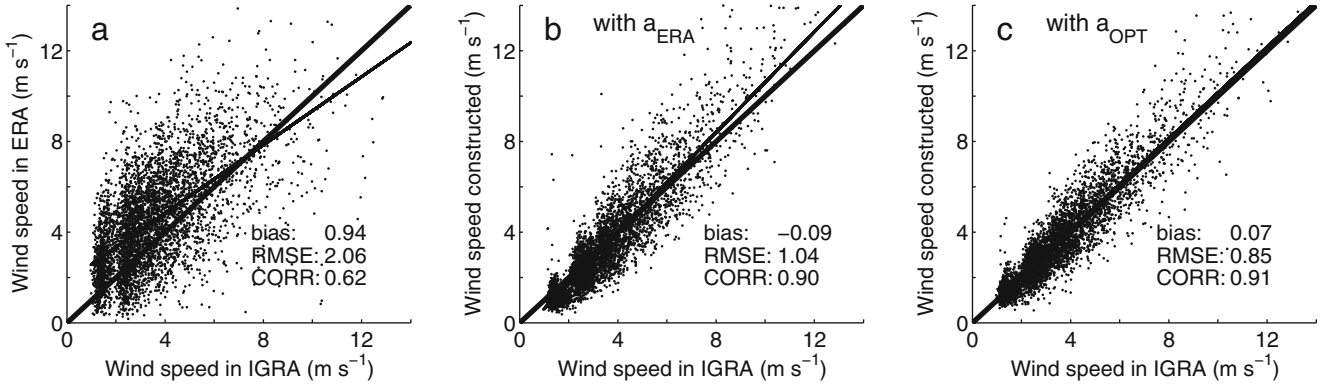

Fig. 12 Similar as Fig. 5a–c but based on ERA5

Fig. 5a), while the performances of the constructed ones are quite close to each other (RMSE 0.85 versus 0.89; Fig. 12c versus Fig. 5c). Those comparisons show that the optimized wind profile and the constructed 80-m wind speed are very close when using the current WRF simulation and ERA5 data. At the same time, Shin et al. [2012] stated that the simulated near-surface wind speed can be effectively improved by increasing the horizontal and vertical resolutions of the model. Considering both our simulation and ERA5 data are relatively within a coarse resolution, we further design a more higher resolution simulation (approximately 3 km horizontally and up to 76 vertical layers) over the selected sub-region of China (Appendix B, Fig. B1) to quantify the dependence of the current statistical method to a much higher resolution simulation. Comparisons show that the optimized wind profile and the constructed wind speed are also quite close when derived from the coarse-resolution simulation and the high-resolution simulation (Appendix B). Therefore, we conclude that the dependence of the optimized wind profile and the constructed wind speed on the model resolution is very limited in our case. The reason could be that improving the (horizontal and vertical) resolution did not improve the simulation results in this sensitivity test. Nevertheless, the construction of wind speed may benefit from better simulation results.

6 Summary and outlook

In this work, a new statistical method is proposed to construct the wind speed at a typical turbine height: We firstly use the least square method to optimize the wind profiles derived from the WRF simulation. The 80-m wind speed is then constructed using the optimized wind profiles and 10-m wind speed measurements at conventional weather stations throughout China. Validation with radiosonde data shows that the optimized wind profiles and constructed 80-m wind speed show considerable reductions in the bias and RMSE, as well as increases in the correlation coefficient when compared with the simulation results from the WRF model. Comparing the wind classifications with that at 10 m, the constructed wind speed generally shows agreement with respect to marginal days and suitable days for wind power production. Furthermore, comparison with previous studies shows that the constructed 80-m wind speed has considerable smaller errors than their work. Additional sensitivity tests show that the constructed wind speed has similar results based on ERA5 and a much higher simulation, which could be due to the similar performance of these datasets, and the construction may benefit from better simulation results.

The proposed statistical method to construct wind speed at turbine height as described in Section 3 is physically reasonable and easy to follow. It is demonstrated to be effective and outperforms the other methods based on the same approach in previous work. Therefore, this procedure and the constructed 80-m wind speed dataset have the potential to be widely used in wind energy studies.

Further work is planned to develop gridded wind speed data based on the current work and to extend the constructed wind speed dataset to a longer time period considering from two aspects: (1) application to wind energy study with different types of turbine (e.g., turbine height of 60 m, 100 m, and 120 m), and (2) evaluation of numerical models and investigation of wind energy variability at different time scales (e.g., decadal variability and its linear trends). Much detailed investigations on the temporal-spatial characteristics of wind speed and wind energy are then expected.

Funding information This work was funded by the National Natural Science Foundation of China (Grant No. 41705084, 91537210), National Key Development Program of China (Grant No. 2018YFA0605400), and by the 13th Five-Year Information Plan of Chinese Academy of Sciences (Grant No. XXH13505-06).

Data availability The station data were provided by the National Meteorological Information Center of the China Meteorological Administration. The ERA-Interim reanalysis data were downloaded from https://www.ecmwf.int/en/forecasts/datasets. The radiosonde data can be downloaded from the IGRA website provided. We wish to thank the China Meteorological Administration and the European Centre for Medium-Range Weather Forecasts (ECMWF) for offering the necessary data. The daily constructed 80-m wind speed data can be obtained by contacting the authors.

Compliance with ethical standards

Conflict of interest The authors declare that there are no conflict of interest.

References

Achberger C, Chen DL, Alexandersson H (2006) The surface winds of Sweden during 1999-2000. Int J Climatol 26(2):159–178. https://doi.org/10.1002/joc.1254

Archer CL, Jacobson MZ (2003) Spatial and temporal distributions of US winds and wind power at 80 m derived from measurements. J Geophys Res-Atmos 108(D9). https://doi.org/10.1029/2002jd002076

Archer CL, Jacobson MZ (2005) Evaluation of global wind power. J Geophys Res-Atmos 110(D12). https://doi.org/10.1029/2004jd005462

Beljaars ACM, Brown AR, Wood N (2004) A new parametrization of turbulent orographic form drag. Q J Roy Meteor Soc 130(599):1327–1347. https://doi.org/10.1256/qj.03.073

Bichet A, Wild M, Folini D, Schar C (2012) Causes for decadal variations of wind speed over land: sensitivity studies with a global climate model. Geophys Res Lett 39. https://doi.org/10.1029/2012gl051685

Carvalho D, Rocha A, Gomez-Gesteira M, Santos CS (2014) WRF wind simulation and wind energy production estimates forced by different reanalyses: comparison with observed data for Portugal. Appl Energy 117:116–126. https://doi.org/10.1016/j.apenergy.2013.12.001

Davidson MR, Zhang D, Xiong WM, Zhang XL, Karplus VJ (2016) Modelling the potential for wind energy integration on China's coal-heavy electricity grid. Nat Energy 1:1–7. https://doi.org/10.1038/Nenergy.2016.86

Dee DP et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteor Soc 137(656):553–597. https://doi.org/10.1002/qj.828

Deppe AJ, Gallus WA, Takle ES (2013) A WRF ensemble for improved wind speed forecasts at turbine height. Weather Forecast 28(1):212–228. https://doi.org/10.1175/Waf-D-11-00112.1

Duan Y (2008) Speciality of wind farm construction in Qinghai-Tibetan plateau. Renew Energy Resour 3:37 in Chinses

Durre I, Vose RS, Wuertz DB (2006) Overview of the integrated global radiosonde archive. J Clim 19(1):53–68. https://doi.org/10.1175/Jcli3594.1

ECMWF (2017), ERA5 reanalysis. National Center for Atmospheric Research Computational and Information Systems Laboratory. Accessed 1 June 2018, https://doi.org/10.5065/D6X34W69., doi: https://doi.org/10.5065/BH6N-5N20

Feng Y, Lin HY, Ho SL, Yanc JH, Dong JN, Fang SH, Huang YK (2015) Overview of wind power generation in China: status and development. Renew Sust Energ Rev 50:847–858. https://doi.org/10.1016/j.rser.2015.05.005

Gao YH, Xue YK, Peng W, Kang HS, Waliser D (2011) Assessment of dynamic downscaling of the extreme rainfall over East Asia using a regional climate model. Adv Atmos Sci 28(5):1077–1098. https://doi.org/10.1007/s00376-010-0039-7

Gao YH, Xu JW, Chen DL (2015) Evaluation of WRF mesoscale climate simulations over the Tibetan Plateau during 1979-2011. J Clim 28(7):2823–2841. https://doi.org/10.1175/Jcli-D-14-00300.1

Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, Randles C, Darmenov A, Bosilovich MG, Reichle R, Wargan K, Coy L, Cullather R, Draper C, Akella S, Buchard V, Conaty A, da Silva A, Gu W, Kim GK, Koster R, Lucchesi R, Merkova D, Nielsen JE, Partyka G, Pawson S, Putman W, Rienecker M, Schubert SD, Sienkiewicz M, Zhao B (2017) The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J Clim 30(14):5419–5454. https://doi.org/10.1175/Jcli-D-16-0758.1

Hong SY, Noh Y, Dudhia J (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. Mon Weather Rev 134(9):2318–2341. https://doi.org/10.1175/Mwr3199.1

Hu XM, Klein PM, Xue M (2013) Evaluation of the updated YSU planetary boundary layer scheme within WRF for wind resource and air quality assessments. J Geophys Res-Atmos 118(18):10490–10505. https://doi.org/10.1002/jgrd.50823

Jiang Y, Luo Y, Zhao Z, Tao S (2010) Changes in wind speed over China during 1956–2004. Theor Appl Climatol 99:421–430. https://doi.org/10.1007/s00704-009-0152-7

Jimenez PA, Dudhia J (2012) Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model. J Appl Meteorol Climatol 51(2):300–316. https://doi.org/10.1175/Jamc-D-11-084.1

Justus CG, Mikhail A (1976) Height variation of wind speed and wind distributions statistics. Geophys Res Lett 3(5):261–264. https://doi.org/10.1029/GL003i005p00261

Klein PM, Hu XM, Shapiro A, Xue M (2016) Linkages between boundary-layer structure and the development of nocturnal low-level jets in Central Oklahoma. Bound-Layer Meteorol 158(3):383–408. https://doi.org/10.1007/s10546-015-0097-6

Krogsaeter O, Reuder J (2015) Validation of boundary layer parameterization schemes in the weather research and forecasting (WRF) model under the aspect of offshore wind energy applicationspart II: boundary layer height and atmospheric stability. Wind Energy 18(7):1291–1302. https://doi.org/10.1002/we.1765

Lee J, Shin HH, Hong SY, Jimenez PA, Dudhia J, Hong J (2015) Impacts of subgrid-scale orography parameterization on simulated surface layer wind and monsoonal precipitation in the high-resolution WRF model. J Geophys Res-Atmos 120(2):644–653. https://doi.org/10.1002/2014jd022747

Li X, Zhong S, Bian X, Heilman WE (2010) Climate and climate variability of the wind power resources in the Great Lakes region of the United States. J Geophys Res-Atmos 115. https://doi.org/10.1029/2009jd013415

Li HD, Wolter M, Wang X, Sodoudi S (2018a) Impact of land cover data on the simulation of urban heat island for Berlin using WRF coupled with bulk approach of Noah-LSM. Theor Appl Climatol 134(1–2):67–81. https://doi.org/10.1007/s00704-017-2253-z

Li X, Gao Y, Pan Y, Xu Y (2018b) Evaluation of near-surface wind speed simulations over the Tibetan Plateau from three dynamical downscalings based on WRF model. Theor Appl Climatol 2017(1–2):1–13. https://doi.org/10.1007/s00704-017-2353-9

Li HD, Zhou YY, Wang X, Zhou X, Zhang HW, Sodoudi S (2019) Quantifying urban heat island intensity and its physical mechanism using WRF/UCM. Sci Total Environ 650:3110–3119. https://doi.org/10.1016/j.scitotenv.2018.10.025

Liu YQ, Kokko A (2010) Wind power in China: policy and development challenges. Energ Policy 38(10):5520–5529. https://doi.org/10.1016/j.enpol.2010.04.050

Liu F, Sun F, Liu W, Wang T, Wang H, Wang X, Lim W (2018) On wind speed pattern and energy potential in China. Appl Energy

Mahrt L, Pan H (1984) A 2-layer model of soil hydrology. Bound-Layer Meteorol 29(1):1–20. https://doi.org/10.1007/Bf00119116

McElroy MB, Lu X, Nielsen CP, Wang YX (2009) Potential for wind-generated electricity in China. Science 325(5946):1378–1380. https://doi.org/10.1126/science.1175706

Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J Geophys Res-Atmos 102(D14):16663–16682. https://doi.org/10.1029/97jd00237

Pan HL, Mahrt L (1987) Interaction between soil hydrology and boundary-layer development. Bound-Layer Meteorol 38(1–2):185–202. https://doi.org/10.1007/Bf00121563

Shin HH, Hong SY, Dudhia J (2012) Impacts of the lowest model level height on the performance of planetary boundary layer parameterizations. Mon Weather Rev 140(2):664–682. https://doi.org/10.1175/Mwr-D-11-00027.1

Sun SP, Liu FL, Xue S, Zeng M, Zeng FX (2015) Review on wind power development in China: current situation and improvement strategies to realize future development. Renew Sust Energ Rev 45:589–599. https://doi.org/10.1016/j.rser.2015.02.018

Tobin I, Vautard R, Balog I, Breon FM, Jerez S, Ruti PM, Thais F, Vrac M, Yiou P (2015) Assessing climate change impacts on European wind energy from ENSEMBLES high-resolution climate projections. Clim Chang 128(1–2):99–112. https://doi.org/10.1007/s10584-014-1291-0

Wang CH, Jin SL (2014) Error features and their possible causes in simulated low-level winds by WRF at a wind farm. Wind Energy 17(9):1315–1325. https://doi.org/10.1002/we.1635

Wang ZQ, Duan AM, Li MS, He B (2016) Influences of thermal forcing over the slope/platform of the Tibetan Plateau on Asian summer monsoon: numerical studies with the WRF model. Chin J Geochem 59(9):3175–3187. https://doi.org/10.6038/cjg20160904

Wharton S, Simpson M, Osuna JL, Newman JF, Biraud SC (2015) Role of surface energy exchange for simulating wind turbine inflow: a case study in the Southern Great Plains, USA. Atmosphere-Basel 6(1):21–49. https://doi.org/10.3390/atmos6010021

Zhang SF, Li XM (2012) Large scale wind power integration in China: analysis from a policy perspective. Renew Sust Energ Rev 16(2):1110–1115. https://doi.org/10.1016/j.rser.2011.11.007

Zhang F, Yang Y, Wang C (2015) The effects of assimilating conventional and ATOVS data on forecasted near-surface wind with WRF-3DVAR. Mon Weather Rev 143(1):153–164. https://doi.org/10.1175/MWR-D-14-00038.1

Zhao J, Guo ZH, Su ZY, Zhao ZY, Xiao X, Liu F (2016) An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. Appl Energy 162:808–826. https://doi.org/10.1016/j.apenergy.2015.10.145

Zhao J, Guo YL, Xiao X, Wang JZ, Chi DZ, Guo ZH (2017) Multi-step wind speed and power forecasts based on a WRF simulation and an optimized association method. Appl Energy 197:183–202. https://doi.org/10.1016/j.apenergy.2017.04.017

Zhou X, Beljaars ACM, Wang Y, Huang B, Lin C, Chen Y, Wu H (2017) Evaluation of WRF simulations with different selections of sub-grid orographic drag over the Tibetan Plateau. J Geophys Res-Atmos 122:9759–9772. https://doi.org/10.1002/2017jd027212

Zhou X, Yang K, Wang Y (2018) Implementation of a turbulent orographic form drag scheme in WRF and its application to the Tibetan plateau. Clim Dyn 50(7–8):2243–2255. https://doi.org/10.1007/s00382-017-3677-y

Zhou X, Yang K, Beljaars A, Li HD, Lin CG, Huang B, Wang Y (2019) Dynamical impact of parameterized turbulent orographic form drag on the simulation of winter precipitation over the western Tibetan Plateau. Clim Dyn 53(1–2):707–720. https://doi.org/10.1007/s00382-019-04628-0