Wind speed spatial estimation using geostatistical kriging

Yu Wang 1,2, Deji Wang3*, Jianghai Zhao2, Changan Zhu 1*

1 School of Engineering Science, University of Science and Technology of China, Hefei 230026, China
2 Hefei Institutes of Physical Science, Chinese Academy of Sciences, Changzhou 213164, China
3 Staff development institute of CNTC, Zhengzhou 450000, China

Corresponding authors: 1) Changan Zhu (postlily@mail.ustc.edu.cn);

Abstract. Accurate wind speed prediction plays an important role in the planning and development of wind power plant. Wind speed prediction can reduce the impact of wind power grid connection and improve the control of wind power systems. In this study, wind speed is considered as a geographical variable. We firstly used exploratory methods to examine the spatio-temporal dependence on wind field. The variogram was taken into account in analysing the spatial structure in the rectangular grid. Then the exponential model was adopted to fit the anisotropic feature of variogram. The de-trended data have been computed by using a geostatistical kriging estimator. The results were compared with the leave-one-out cross-validation (LOOCV) method. The accuracy of the proposed method was verified to obtain the best estimated results.

1. Introduction
Wind power is a kind of renewable energy with a great potential for reducing greenhouse gas emissions and coping with the climate change around the world. In recent years, the construction scale of wind farms has become larger and larger in China and all over the world [1–3]. With the rapid development of big data science in [6] and intelligent prediction methods [7] over the past decades. Traditional numerical methods for wind speed prediction have been proposed, including the Numerical Weather Prediction (NWP) method, artificial neural network (ANN) method, the statistical method and hybrid method. The statistical method includes auto-regressive moving-average (ARMA) model [15], hybrid ARMA model [16]. Compared with the statistical method, the hybrid method exhibits higher accuracy for wind speed prediction in wind farms[17–21].The geostatistical kriging has also been used in the prediction of wind resource. For examples, the geographical statistics kriging combining with Bayesian model was applied in the hierarchical modeling of wind farms [26]; and the Taylor kriging model was used to forecast the wind speed of time series in North Dakota, USA [27]. It was found that the kriging methods could produce more accurate results and outperform the traditional methods in wind speed prediction. However, these kriging methods previously were mainly used in the wind speed prediction for different timescales. Moreover, the prediction on spatial and time serious wind speed was carried out separately. In addition, the variation of wind speed shows complex atmospheric mechanisms relating to both the temporal and spatial behaviors.

In response to the challenges, the geostatistical kriging was applied for wind speed prediction for comprehensive understanding of the spatial and temporal characteristics of wind speed. The wind data
were sampled at a set of regularly distributed sites and treated as geostatistical variables. The geostatistical variables were modeled through the spatial semivariogram. The correlations between distances were analyzed. To validate the proposed kriging method, the predicted results are compared. This study demonstrates the validation to extend the application of geostatistical king method in wind speed prediction.

2. Methodology

2.1 The geostatistical kriging technique

The prediction at the unobserved point $Z(s_e)$ is based on the observed points $Z(s_i)$ and the best linear unbiased predictor of $Z(s_e)$ is:

$$
\hat{Z}(s_e) = f(s_o)\hat{\beta} + \nu V^{-1}(z(s) - F\hat{\beta})
$$

Where $\hat{\beta} = (F'V^{-1}F)^{-1}F'V^{-1}z(s)$ is generalized least squares (GLS) estimation of the trend coefficients and $F'$ is the transpose of the design matrix $F$. The estimation error, $\nu = (\hat{\beta} - \beta) = (F'V^{-1}F)^{-1}F'V^{-1}z(s) - F\beta$. It is zero if $s_o$ is an observed location and increases if $F(s_e)$ is far more distant from $F$. The predictor consists of an estimated mean value at location $s_o$ and a weighted mean value of the residuals from the mean function, where the weights is $\nu V^{-1}$. The predictor in (5) has prediction error variance:

$$
\text{Var}(\hat{Z}(s_e)) = \sigma^2_0 - \nu V^{-1}V + \delta (F'V^{-1}F)^{-1}\delta
$$

Where $\sigma^2_0 = \text{Var}(Z(s_o))$, $\delta = f(s_e) - \nu V^{-1}F$, the term $\nu V^{-1}V$ is zero if all observations are uncorrelated with $Z(s_o)$.

There are basic theoretically restricted variograms to describe the spatial dependencies: the spherical, exponential, Gaussian and cubic models [28]. The exponential theoretical model is given as follow.

$$
\gamma(h) = \begin{cases} 
c_0 + c \left[1 - \exp \left(\frac{-h}{a}\right)\right], & 0 < h \\
0, & h \geq 0.
\end{cases}
$$

The parameters $c$ and $a$ denote the distance lag. The function asymptotically approaches its sill and so has no finite range. Its effective range is usually taken as $r' = 3a$.

2.2 Accuracy assessment

The model’s predictive accuracy can be assessed through cross-validation. The quality of all methods was assessed using statistical measures. For example, Mean error (ME), Mean squared error (MSE) and Mean squared deviation ratio (MSDR) between the predictions and measurements at known locations was calculated using leave-one-out cross-validation (LOOCV) method. LOOCV is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice.

$$
\text{ME} = \frac{1}{N} \sum_{i=1}^{N} (Z(x_i) - \hat{Z}(x_i))
$$

$$
\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (Z(x_i) - \hat{Z}(x_i))^2
$$
\[
MSDR = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{Z(x_i) - \hat{Z}(x_i)}{\hat{\sigma}^2(x_i)} \right)^2
\]  

3. Results and discussion

3.1. Wind speed data

The wind speed profile in Cartesian coordinates was generated by using the TurbSim [30]. TurbSim is a stochastic, full field, turbulent-wind simulator that uses a statistical model to generate time series of three-component wind vectors at points in a two-dimensional vertical rectangular grid. Three wind fields were considered: one with step changes and the other two with time changes in the wind speed. This study selected wind data. Fig. 1 (a) shows the wind speed curves at different instants of time. Fig. 1(b) shows the residuals after the first order difference. One can see the de-trended data have a normal characteristic with mean null.

![Figure 1](a) Wind speed curves at a rectangular grid; (b) the residuals

3.2. Model fitting

3.2.1 Experimental semivariogram

The theoretical variogram model in Section2 has been fully discussed for structural analysis and modeling of the wind field. We are now trying to construct an effective space model through the minimum optimization technique by fitting theoretical semivariogram function with the observed data. Wind speed exhibits non-trivial spatio correlation structure, when restricted to a time lag, the space correlation is quite strong. Figure 2 shows the simultaneous spatial correlation between the wind point grids of the portfolio. As can be observed, spatial dependence is quite strong even at distances of about 150 meters. This can be partly explained by the wind map landscape. As expected, the number of pairs firstly increases and then decreases with the increasing distance, which implies that the semivariogram values for large distances are not reliable. As shown Fig.4, the fluctuations in the empirical semivariogram decrease with the increasing tolerance angle, albeit at the expense of increasing the discrimination error committed. The anisotropies in different directions almost do not affect the empirical semivariogram. However, the semi-variance shows different trends with the lag distance. For the lag distance less than about 80m, the semi-variance in four directions increases almost linearly with the increase of the lag distance; then the semi-variance increases slowly; after the lag distance is more than about 220m, the semi-variance rapidly increases.
3.3 Wind speed spatial estimation

The exponential model has been proved adequate to fit the empirical semi-variogram in Section 2. The UK method is used for spatial estimation of wind speed. And then, in every point of the same grid used in the sample the wind speed mean has been estimated by using OK (Ordinary Kriging) and UK (Universal Kriging). Figure 4 (a) shows the corresponding prediction results. From such figure it can be deduced that the prediction results match the observed values well and are consistent. The prediction error values are in a similar stable range. The results indicate that the OK method can be used to successfully predict the spatial wind speed. The LOOCV validation used in the companion has been carried out here.

When the number of random splits goes to infinity, the repeated random sub-sampling validation becomes arbitrary close to the leave-one-out cross-validation (LOOCV). Figure 5 shows the LOOCV validation results of the UK. The validation yields very good results. It must be observed that most of the residual is composed of noise, which is a signal uncorrelated with the observations and thus not estimated by kriging. On the other hand, this lack of generalization could also be due to a modelling error for the fact that an isotropic model of wind speed has been assumed.

Figure 2  Experimental semivariogram (splattering points) and theoretical model.

Figure 3. Effects of the empirical semivariogram with different tolerance angles.

Figure 4. (a) Plot of the predicted and the observed values; (b) Frequency histogram of the residuals.

Figure 5 (a) LOOCV for UK applied to the residuals. (b) Plot of the prediction errors of the UK.
3.4 Model validation
In order to verify the model’s performance, a subset of data has been considered as a finite realization of structural analysis and parameter estimation. We now check the four semivariogram models by weighted least square (WLS) method. Cross validation for the four models based on the validation set, and then validation measurements can be compared to their predictions. As shown in Table 1, the validated error measurements ME, MSE and MSDE are the mean of the predicted errors divided by the corresponding kriging variances. As described in [29], if the semivariogram model has been successfully applied, the ME should be approximately zero; the MSE should be very small; the MSDE should be approximately 1; and the prediction errors are compatible with the kriging prediction variance. Thus, the use of the model and the prediction can be confident based on that such a semivariogram is approximately unbiased and that the mean-square prediction errors are approximate. As expected, the ME values are close to zero for these four models. The MSE values are small and approximately the same for the Exponential, Spherical and Cubic models, but clearly larger for the Gaussian model. The MSDE shows little differences. According to the judged criterion, the Exponential model is the best one which also indicates that the exponential model prediction is used for prediction. Besides, the correlation coefficients between the predicted and observed values are 0.8581, 0.8313, 0.8336 and 0.7670, respectively.

Table 1. Errors comparison of the four models

| Model      | ME    | MSE  | MSDE |
|------------|-------|------|------|
| Exponential| -0.00091 | 0.001 | 1.86 |
| Spherical  | -0.00087 | 0.001 | 2.33 |
| Cubic      | -0.00018 | 0.001 | 0.77 |
| Gaussian   | -0.00011 | 0.002 | 1.34 |

4. Conclusion
In this paper, wind speed data are chosen to investigate the application of geostatistical kriging technology. The wind speed was treated as geostatistical variables that were modelled through the spatial variogram. The correlations of wind speed between lag distances were analyzed with different fitting models. Besides, the UK and OK were employed to carry out short-term wind speed prediction and validate the proposed method’s superiority. The results indicated that the UK method could be successfully used for short-term wind speed prediction.

Acknowledgment
This work was supported in part by the Key R&D Program of Jiangsu Province under Grant BE2017007-1, in part by Project of National Natural Science Foundation of China under Grant 61703390, and in part by Anhui Natural Science Foundation under Grant 1808085QF193.

References
[1] deCastro M, Salvador S, Gómez-Gesteira M, Costoya X, Carvalho D, Sanz-Larruga FJ, et al. Europe, China and the United States: Three different approaches to the development of offshore wind energy. Renew Sustain Energy Rev 2019;109:55–70. https://doi.org/10.1016/j.rser.2019.04.025.
[2] Study on China’s wind power development path—Based on the target for 2030. Renew Sustain Energy Rev n.d.;51:197–208.
[3] Giri NC. Accelerating India’s Energy Sector to Sustainable Sources, Potentials and Prospects 2020.
[4] Simão HP, Powell WB, Archer CL, Kempton W. The challenge of integrating offshore wind power in the U.S. electric grid. Part II: Simulation of electricity market operations. Renew Energy 2017;103:418–31. https://doi.org/10.1016/j.renene.2016.11.049.
[5] Martin J, Ramsey D. The Economics of Wind Energy. J Appl Corp Financ 2009;21:100–9. https://doi.org/10.1111/j.1745-6622.2009.00231.x.
[6] Zhu K, Joshi S, Wang QG, Hsi JFY. Guest Editorial Special Section on Big Data Analytics in Intelligent Manufacturing. IEEE Trans Ind Informatics 2019; 15: 2382–5. https://doi.org/10.1109/TII.2019.2900726.
[7] Huang R, Member S, Li J, Li W, Member S, Cui L. Deep Ensemble Capsule Network for Intelligent Compound Fault Diagnosis Using Multisensory Data 2019;9456:1–11. https://doi.org/10.1109/TIM.2019.2958010.

[8] Liu X, Pu K, Bo L, Yang K, Zhao Y, Liu Y, et al. Quantitative estimation of nonlinearity parameter of noised lamb waves using a chaotic oscillator. J Sound Vib 2019;438:476–89. https://doi.org/10.1016/j.jsv.2018.09.043.

[9] Chen N, Qian Z, Nabney IT, Meng X. Wind power forecasts using gaussian processes and numerical weather prediction. IEEE Trans Power Syst 2014. https://doi.org/10.1109/TPWRS.2013.2282366.

[10] Gauterin E, Kammerer P, Kühn M, Schulte H. Effective wind speed estimation: Comparison between Kalman Filter and Takagi-Sugeno observer techniques. ISA Trans 2016;62:60–72. https://doi.org/10.1016/j.isatra.2015.11.016.

[11] Zhao X, Jiang N, Liu J, Yu D, Chang J. Short-term average wind speed and turbulent standard deviation forecasts based on one-dimensional convolutional neural network and the integrate method for probabilistic framework. Energy Conver Manag 2020;203:112239. https://doi.org/10.1016/j.enconman.2019.112239.

[12] Peng Z, Peng S, Fu L, Lu B, Tang J, Wang K, et al. A novel deep learning ensemble model with data denoising for short-term wind speed forecasting. Energy Conver Manag 2020;207:112524. https://doi.org/10.1016/j.enconman.2020.112524.

[13] Zhang Y, Chen B, Pan G, Zhao Y. A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting. Energy Conver Manag 2019;195:180–97. https://doi.org/10.1016/j.enconman.2019.05.005.

[14] Harbola S, Coors V. One dimensional convolutional neural network architectures for wind prediction. Energy Conver Manag 2019;195:70–5. https://doi.org/10.1016/j.enconman.2019.05.007.

[15] Zhang Y, Zhao Y, Kong C, Chen B. A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic. Energy Conver Manag 2019;203:112254. https://doi.org/10.1016/j.enconman.2019.112254.

[16] Wang J, Li Y. An innovative hybrid approach for multi-step ahead wind speed prediction. Appl Soft Comput J 2019;78:296–309. https://doi.org/10.1016/j.asoc.2019.02.034.

[17] Jiang Y, Zhao N, Peng L, Liu S. A new hybrid framework for probabilistic wind speed prediction using deep feature selection and multi-error modification. Energy Conver Manag 2019;199:111981. https://doi.org/10.1016/j.enconman.2019.111981.

[18] Jiang Y, Liu S, Peng L, Zhao N. A novel wind speed prediction method based on robust local mean decomposition, group method of data handling and conditional kernel density estimation. Energy Conver Manag 2019;200:112099. https://doi.org/10.1016/j.enconman.2019.112099.

[19] Wang Y. Prediction of Spatial Wind Speed at Different Timescales Using Geostatistical Universal Kriging 2017.

[20] Jiang Y, Huang G, Yang Q, Yan Z, Zhang C. A novel probabilistic wind speed prediction approach using real time refined variational model decomposition and conditional kernel density estimation. Energy Conver Manag 2019;185:758–73. https://doi.org/10.1016/j.enconman.2019.02.028.

[21] Hoolohan V, Tomlin AS, Cockerill T. Improved near surface wind speed predictions using Gaussian process regression combined with numerical weather predictions and observed meteorological data. Renew Energy 2018;126:1043–54. https://doi.org/10.1016/j.renene.2018.04.019.

[22] Yang Y, Wu J, Christakos G. Prediction of soil heavy metal distribution using Spatiotemporal Kriging with trend model. Ecol Indic 2015;56:125–33. https://doi.org/10.1016/j.ecolind.2015.03.034.

[23] Geostatistics A. Spatio-temporal analysis and interpolation of PM 10 and PM 1 in Eindhoven 2015:1–13.

[24] Zhang B, Yang Y. Spatiotemporal modeling and prediction of soil heavy metals based on spatiotemporal cokriging. Sci Rep 2017;7:1–10. https://doi.org/10.1038/s41598-017-17018-5.

[25] Cellura M, Cirrincione G, Marvuglia A, Miraoui A. Wind speed spatial estimation for energy planning in Sicily: A neural kriging application. Renew Energy 2008;33:1251–66. https://doi.org/10.1016/j.renene.2007.08.013.

[26] Qian H, Hongkun C, Yubo T, Ruixi Y, Ling W, Pan H. Wind speed spatio-temporal forecasting of wind farms based on universal kriging and Bayesian dynamic model. POWERCON 2014 - 2014 Int Conf Power Syst Technol Towar Green, Effic Smart Power Syst Proc 2014:2897–902. https://doi.org/10.1109/POWERCON.2014.6993915.

[27] Liu H, Shi J, Erdem E. Prediction of wind speed time series using modified Taylor Kriging method. Energy 2010;35:4870–9. https://doi.org/10.1016/j.energy.2010.09.001.
[28] Oliver MA, Webster R. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. Catena 2014;113:56–69. https://doi.org/10.1016/j.catena.2013.09.006.

[29] Montero JM, Fernández-Avilés G, Mateu J. Spatial and Spatio-Temporal Geostatistical Modeling and Kriging. Spat Spat Geostatistical Model Kriging 2012:1–357. https://doi.org/10.1002/9781118762387.

[30] Jonkman BJ, Buhl ML. TurbSim User’s Guide TurbSim User’s Guide. Renew Energy 2006.