Very Short-term Wind Speed Prediction Using Geostatistical Kriging with External Trend

Yu Wang1,2, Kunpeng Zhu2*, Changan Zhu1*

1 School of Engineering Science, University of Science and Technology of China, Hefei 230026, China
2 Institute of Advanced Manufacturing Technology, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Changzhou 213164, China

Corresponding authors: Changan Zhu (e-mail: changan@ustc.edu.cn)

Abstract. The prediction of wind speed plays a key role in the determination of the optimal reserve capacity for different time instants and the reduction of the standby cost of wind power grid. Short-term wind speed prediction can reduce the impact of wind power grid and improve the control of wind turbines. The geostatistics kriging as an unbiased estimation method applied to predict the wind speed in a two-dimensional vertical rectangular grid. The wind data were as a geographical variable considered. We firstly used exploratory methods to examine the spatial variability of wind speed and its dependence on geographical variables. The covariance and variogram models took into account in analysing the spatial structure of wind speed at three different instants of time. Then the product model adopted to fit wind data due to the variogram. The accuracy of the method was assessed through a cross validation. The prediction results were comparable with the observed values. The prediction error was about ±0.15 and the correlation coefficient was about 0.85. This study demonstrated that the kriging method could be successfully used in the short-time wind speed prediction in space and time direction.

1. Introduction

Wind power is a kind of renewable energy with 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[1,2]. Short-term wind speed prediction is the most important thing to determine the optimal reserve capacity for different time instance and reduce the standby cost of wind power grid. Regarding this, there research results have done on wind speed and wind power prediction.

Traditional wind predictions of wind speed are mainly included the statistical method and physical method. Statistical method uses mathematical models, such as time-series models and data mining techniques[3–5], to compute for the forecasting results based on the a large number of historical wind speed data. Physical method, which is focused on the physical characteristics of wind, applies the terrain analysis, wake effect analysis, spatial correlation analysis and other methods to calculate the relevant meteorological information of the wind locations [6–11]. Physical method does not require a large amount of historical data, but it has a high demand on the model accuracy.

Compared with traditional statistical predicting methods, geographical theory can give a temporal and spatial correlation based on the historical wind data information, which makes the demands of wind speed data greatly reduce under the same forecasting accuracy. The forecasting results in the form of
the optimal linear unbiased estimation are given and the parameters estimated by means of partial prior knowledge. Meanwhile, the method also takes into account the spatial-temporal feature of wind data, comprehensively analyzes the temporal and spatial correlation for the wind data in both space and time directions. Finally, a simulation of wind speed in a rectangular space given to verify the validity of the method.

2. Methods

2.1. Geostatistics modeling and prediction

Let the velocity of wind field be denoted as \( z(s, t) \) at each point grid, where \( s \) denotes a spatial point grid. Typically, \( z(s, t) \) may exhibit spatial dependencies, which can interpolate via kriging by a function of known coverable over the geostatistics domain. The wind velocity \( z(s, t) \) can express as fellow:

\[
Z(s, t) = U(s, t) + V(s, t)
\]

(1)

Where \( U(s, t) \) is the deterministic part, or spatial trend and \( V(s, t) \) is a zero-mean second-order stationary spatial random field with covariance function residual and considered an intrinsically stationary random function with a zero expectation and the residual covariance function defined as:

\[
C(h) = \text{Cov}(Z(s), Z(s+h))
\]

(2)

Where \( h \) represents the spatial lag distance between two observed points \((s, t)\). The covariance function (2) that characterizes the wind speed variation can be calculated in [12].

\[
C(h) = \sigma^2 \text{R}(s, t; \Phi)
\]

(3)

Where \( \sigma^2 \) is the process variance for the 1, component of the response; and \( \text{R}(s, t; \Phi) \) is the spatial correlation of stochastic process between the different locations \( s \) and \( t \) with smoothness and decay parameter \( \Phi \).

For simplicity of exposition, we assume in this section that the trend has removed up to a constant and that a stationary model is suitable for the residual wind data. The commonly covariance model is built by the modifying or combining generic covariance function. Commonly known generic models are covariance functions of exponential, spherical, Gaussian and cubic type, amongst many others.

2.2 The variogram models

The experimental variogram is one that we estimate from data. It is usually computed by the method of moments and attributed in [14]:

\[
\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{j=1}^{N(h)} \left( z(x_j) - z(x_j + h) \right)^2
\]

(4)

Where \( N(h) \) is the number of paired comparisons at distance lag \( h \). By incrementing \( h \) in steps, we obtain an ordered set of values, as shown by the points plotted in each of the graphs in Fig. 1. This is the experimental variogram, also known as the sample variogram. It estimates points on the regional variogram.

In this paper, the generalized product model proposed for modelling the spatio-temporal correlation structures under study. The generalized product model [15] in the variogram form is defined as follows:

\[
\gamma_{st}(h, u) = \sigma^2 \exp\left(-\|h\|/a_s - |u|/a_t\right)
\]

(5)

Where \( a_s \) and \( a_t \) are two positive space and time scaling parameters, respectively and \( \sigma^2 \) is the a priori variance. This kind of choice can be justified by fitting of the generalized product model in the variogram form; it is flexible because it regards the temporal and spatial marginal structures, and just
one parameter. Several authors have used the available spatio-temporal models for prediction purposes in their applications[16–19].

3. Results and discussion

3.1. Dataset
The dataset used in this paper are publicly available by the department of Energy’s national renewable energy laboratory in US. The page at http://wind.nrel.gov/designcodes/. The original wind values of 817600 scores are modified based on the hub height of 50-m wind towers of a large wind farm during January 2001-April 2005, in San Gorgonio Pass in southern California, US. Figure 1 shows the original wind speed curves. It shows a strong trend effect of the original wind data. The first task is to eliminate the trend, which is usually of low frequency and followed a deterministic pattern due to the dominant physics. Here, the first order difference computed to remove the trend effect from the original data. Figure2 showed the residual data that the de-trended data had a mean null and the variance exists and is finite. Therefore, the operation performed very well, and the resulting quantity referred to as the wind velocity. Subsequent statistical analysis, modeling, and prediction performed on the de-trended data.

3.2. Computing and modeling variogram
Let us consider a sample consisting of 9600 values at 48 consecutive instants of time in a space of rectangle grid size of 800 *400. First, the empirical variogram function (4) used to estimate the space–time variability. As shown in Figure 4, the theoretical models follow the patterns of the experimental variogram rather closely, indicating that the fitted models adequately express the existing space-time correlations between the residual data. The plotting of the sample marginal variogram showed in Figure 3, the sample spatial variogram for each data time point and then averaging over time. It shows a great correlation when the distance between two positions is smaller than 4 km. As shown in Figure 3 (b), sample temporal variogram for each data spatial location and then averaging over space. It estimates points on the regional variogram.

The generalized product model (5) depends on the experimental structures from wind speed data. the spatial-time correlation and exponential function have been successfully applied for wind energy planning in [10].The result depends on the precise way we apply the program and the decisions we make, which we discuss after modeling. In this study, we also use this model for our wind data modeling. More specifically, we will optimize parameters according to the above model with initial values and some of the parameters fixed and form different models to estimate.

3.3. Short-term prediction in space and time
The exponential model proved adequate to fit the empirical variogram in section 3.2. Now the kriging used to predict short-term prediction in space and time. The prediction results show in Figure 5, Figure 6, Figure 7 and Figure 8, respectively. From such figures, it can deduce that the prediction results match the observed values very well and are consistent at different instants of time. Although the predicted plots in the Figures 5 (a) and Figure 6 (a) appear to be the same, the legends indicate that the predicted values tend to the mean of the process when the time horizon increases. However, the predictions RMSE plots in Figures 5 (b) and Figure 6 (b) : it appears to be the same, but, as expected, the longer the prediction time horizon, the greater the space prediction variance, its maximum being sill in the covariance function used in the kriging equations. For short-term time prediction, hourly and daily predicted at 24-hour and 2-day ahead, respectively.

3.4. Accuracy assessment
A subset of validation points reserved among the samples obtained during the period 2001- 2005 in San Gorgonio Pass used for the prediction accuracy assessment. More specifically, the numerical CV validation results obtained based on the criteria of the ME, MSE and MSDE listed in Table 1 for the various techniques and models used. As described in[20], if the variogram models were successfully applied, the ME should be approximately zero; the MSE should be very small; the MSDE should be approximately 1, the prediction errors are compatible with the prediction variance. Thus, the use of the
model and the prediction can be confident based on that such a variogram is approximately unbiased and the mean-squared prediction errors are approximate. According to the judged criterion, the exponential model shows a better forecasting accuracy.

**Figure 1** The original wind speed curves.

**Figure 2** Plotting of the residuals by the first order difference.

**Figure 3** Sample marginal variogram in space (a) and time (b) and their models.

**Figure 4**(a) Sample space–time variogram surface for the wind speed data and (b) its model.
Figure 5 (a) Space predictions at each point of the estimation grid. (b) Plotting of the RMSE difference of the observations and SK estimates at each point of the estimation grid.

Figure 6 (a) Space predictions at each point of the estimation grid; (b) the RMSE difference of the observations and SK estimates at each point of the estimation grid.

Figure 7 24-hour ahead prediction at p=11.

Figure 8 Two-day ahead prediction at p=21.

4. Conclusion
In this study, the wind speed data chosen to investigate the application of space–time model in the prediction of wind speed. The wind data treated as geostatistics variables modeled through the spatial and time variogram. The cross validation analyzed with the different models. The wind speed values predicted by the proposed model matched the observed values well. The results indicated that the kriging method be successfully used for short-time wind speed prediction in space and time.

Table 1 Quantitative criteria for the comparison of the four models.

| Model  | ME        | MSE       | MSDE      |
|--------|-----------|-----------|-----------|
| Exponential | -0.0009132623 | 0.001052534 | 1.856445  |
| spherical | -0.0008782672 | 0.001023111 | 2.328915  |
| Gaussian | -0.0001813599 | 0.001420726 | 0.7666526 |
| cubic   | -0.0001101884 | 0.002057135 | 1.337914  |
5. References

[1] Zhang Z, Qin H, Liu Y, Wang Y, Yao L, Li Q, Li J and Pei S 2019 Long Short-Term Memory Network based on Neighborhood Gates for processing complex causality in wind speed prediction Energy Convers. Manag. 192 37–51

[2] Anon Study on China’s wind power development path—Based on the target for 2030 Renew. Sustain. Energy Rev. 51 197–208

[3] Liu H, Tian H Q, Liang X F and Li Y F 2015 Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks Appl. Energy 157 183–94

[4] Kavasseri R G and Seetharaman K 2009 Day-ahead wind speed forecasting using f-ARIMA models Renew. Energy 34 1388–93

[5] Santhosh M, Venkaiah C and Vinod Kumar D M 2018 Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction Energy Convers. Manag. 168 482–93

[6] Feng S L, Wang W S, Liu C and Dai H Z 2010 Study on the physical approach to wind power prediction Proc. Csee 30 1–6

[7] Flay R G J, King A B, Revell M, Carpenter P, Turner R, Cenek P and Safaei Pirooz A A 2019 Wind speed measurements and predictions over Belmont Hill, Wellington, New Zealand J. Wind Eng. Ind. Aerodyn. 195 104018

[8] Simonetti R, Moretti L, Molinaroli L and Manzolini G 2020 Energetic and economic optimization of the yearly performance of three different solar assisted heat pump systems using a mixed integer linear programming algorithm Energy Convers. Manag. 206 112446

[9] Zhao J, Wang J, Guo Z, Guo Y, Lin W and Lin Y 2019 Multi-step wind speed forecasting based on numerical simulations and an optimized stochastic ensemble method Appl. Energy 255 113833

[10] Cellura M, Cirrincione G, Marvuglia A and Miraoui A 2008 Wind speed spatial estimation for energy planning in Sicily: A neural kriging application Renew. Energy 33 1251–66

[11] Zheng C W, Li C Y, Pan J, Liu M Y and Xia L L 2016 An overview of global ocean wind energy resource evaluations Renew. Sustain. Energy Rev. 53 1240–51

[12] Lophaven S N, Nielsen H B and Sondergaard J 2002 DACE - A Matlab Kriging Toolbox

[13] Oliver M A and Webster R 2014 A tutorial guide to geostatistics: Computing and modelling variograms and kriging Catena 113 56–69

[14] Wu S-J, Chen P-H and Yang J-C 2011 Application of Weighted Semivariogram Model (WSVM) based on fitness to experimental semivariogram on estimation of rainfall amount Hydrol. Earth Syst. Sci. Discuss. 8 4229–59

[15] Yang Y, Wu J and Christakos G 2015 Prediction of soil heavy metal distribution using

[16] Baxevani A and Lenzi A 2018 Very short-term spatio-temporal wind power prediction using a censored Gaussian field Stoch. Environ. Res. Risk Assess. 32 931–48

[17] Liu X L, Fu X Q, Li Y, Shen J L, Wang Y, Zou G H, Wu Y Z, Ma Q M, Chen D, Wang C, Xiao R L and Wu J S 2016 Spatio-temporal variability in N2O emissions from a tea-planted soil in subtropical central China Geosci. Model Dev. Discuss. 1–45

[18] De Iaco S and Posa D 2012 Predicting spatio-temporal random fields: Some computational aspects Comput. Geosci. 41 12–24

[19] Xie L, Gu Y, Zhu X and Genton M G 2014 Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch IEEE Trans. Smart Grid 5 511–20

[20] Montero J M, Fernández-Avilés G and Mateu J 2012 Spatial and Spatio-Temporal Geostatistical Modeling and Kriging Spat. Spat. Geostatistical Model. Kriging 1–357

Acknowledgments
This work is 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.