

# A Hybrid Model for Wind Speed Prediction based on Spatial Correlation

Dong Wang, Jianping Fei and Haidong Qian  
Zhoushan Electric Power Company,  
Zhoushan, China

Yuhong Zhao  
School of Mathematics and Statistics,  
Xidian University,  
Xi'an, China

**Abstract**—The accurate prediction of wind speed is greatly significant for evaluating the security and economy of the wind farm which is connected to the grid. This paper proposed a hybrid model for wind speed prediction based on spatial correlation and nonlinear state space model. Firstly, a rank correlation coefficient is calculated according a Coupla function and the coefficient is used to choose the input data of the model. A support vector regression machine is introduced to build a nonlinear state space model, and then a hybrid model is obtained. Finally, experiment results based on some public data are given, the results demonstrate that the proposed hybrid model has higher prediction accuracy and stable performance while comparing with some classical algorithm.

**Keywords**—wind speed forecasting; spatial correlation; coupla functions; state space

## I. INTRODUCTION

The intermittent and randomness of wind energy determines the volatility of the wind power, along with the increasing number of wind farms and installed capacity increasing, once the wind power integrate the power grid, the fluctuation will bring great challenge to the safe and economic operation of power grid. Accurate projections for wind speed and wind power in advance can alleviate the pressure of the power system in peak load and frequency regulation, and improve acceptable capacity of the power grid to the wind power<sup>[1]</sup>.

With wind power prediction at home and abroad research more and more abroad and deeper, in all kinds of forecasting method what the most simple and direct method is using the nearest wind speed or power point to predict under observation named continuous method<sup>[2]</sup>. Model of this method is simple but the increase of time scale result in rapid decline in accuracy. Other advanced methods like Kalman filter<sup>[3]</sup>, time series<sup>[4-5]</sup>, the artificial neural network<sup>[6-7]</sup> and support vector machine (SVM)<sup>[8]</sup>, which are parts of intelligent method of statistics. The applicability and limitations of these methods have their own shortage, but they all have a common denominator that forecasting need only the wind field in the process of historical data and real-time data and it is easy to implement. But these methods ignore the relation between space position of wind farm around and the result of prediction, so prediction method system and model prediction accuracy are subject to further perfect and improve.

Different from the above methods, spatial correlation method take into account the wind farm and several near sites of wind farm in multiple sets of wind speed data and use spatial correlation between several sites of wind speed for wind speed forecasting. This paper use Coupla function to analysis the sequence similarity of wind farms around with the target wind farm, with the target of using the correlation coefficient of choice to predict the size of the wind speed of wind farm around. Then, support vector machine (SVM) regression is used to establish the nonlinear state space model for more hour forecast wind speed. Eventually the predictive results of the hybrid model compared with the measured data, the accuracy of the verifiable hybrid model can be proved.

## II. WIND SPEED FORECASTING MODEL

### A. Spatial correlation

The experimental data of this method is public data which source from the speed data of 22 wind farm wind (on January 1, 2006 to December 31, 2009, Record every 1 hours) of Wisconsin state, choose 9 wind field for wind farms to predict target, the location of the 22 wind field distribution is shown in figure 1.

Because of the influence of the wind itself inherent physical properties and the terrain, the wind speed change curve in the same direction between different locations could not have exactly the same, But always have a strong similarity. The wind speed change of three wind field in a certain period of time is shown in figure 2. Figure 2 illustrates that the change curve of wind speed of the three wind field has obvious similarities, there is a big difference only between the individual points, so it is feasible to predict the wind speed by using the spatial relationships.

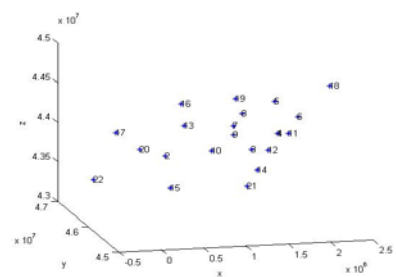


FIGURE 1. DISTRIBUTION OF WIND FIELD POSITION

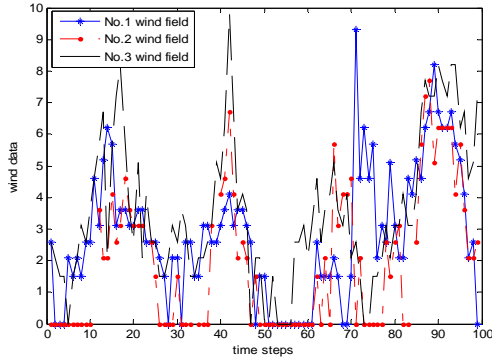


FIGURE II. NO 1,2,3 WIND FIELD SPEED CHANGE CURVE

### B. Based on Coupla of the Correlation Coefficient between the Wind Field

Correlation coefficient is used to evaluate the similarity of two groups of data sequence. In statistical analysis for quantitative analysis of the degree of close relationship between two variables. Commonly used methods for the Pearson correlation coefficient method, Calculating formula for:

$$\rho(X,Y) = \frac{\text{cov}(X,Y)}{\sigma(X)\sigma(Y)} \quad (1)$$

Where  $\text{cov}(X,Y)$  as covariance of random  $X$  and  $Y$ .  $\sigma(X)$  and  $\sigma(Y)$  respectively as standard deviation of random  $X$  and  $Y$ .  $\rho(X,Y)$  is the linear correlation coefficient.

The linear correlation coefficient is used to measure linear correlation between the variables, but the linear correlation coefficient, the correlation among the variables under linear transformation can be the same, under the nonlinear gain transformation will not be able to guarantee the invariance of the correlation coefficient, and to obey the normal distribution between variables, using the linear correlation coefficient, the correlation between variables, there will be a wrong conclusion. In order to overcome this defect, it introduces a Copula continuous function, it is the joint distribution of multidimensional random variables connected with the edge of the one dimensional distribution function, used to describe the multiple joint distribution of random variables and relation.

Copula function<sup>[9]</sup> containing many types can be divided into elliptical and Archimedes. Elliptical class mainly consist of normal function and t copula function, otherwise Gumbel, Frank and Clayton copula function belonging to Archimedes. Ellipse function is applied in this paper. Parameter estimation on copula function can be divided into parametric and nonparametric method. Parameter method need to know in advance model of marginal distribution, which the wind speed is volatile and instable and then we use the nonparametric method to estimate parameters of copula function, namely nuclear smoothing method based on empirical distribution. On the basis of historical samples of wind speed observation

data. We use kernel density to determine the overall probability distribution of wind speed sequence.

To measure the correlation strength between two variables, commonly used indicators are Kendall rank correlation coefficient and Spearman rank correlation coefficient.

Before calculating the rank correlation coefficient of each wind field and the 9th wind field, we exclude the 1st, 8th, 12th, 13th, 15th wind field for the wind speed of them are serious lost. The calculation results are shown in table 1.

TABLE I. THE RANK OF THE PHASE RELATIONSHIP

Wind field	2	3	4	5	6
Kendall_t	0.553	0.597	0.489	0.549	0.573
Spearman_t	0.748	0.79	0.677	0.744	0.769

Wind field	7	10	11	14	16
Kendall_t	0.587	0.591	0.471	0.570	0.373
Spearman_t	0.781	0.787	0.657	0.765	0.535

Wind field	17	18	19	20	21	22
Kendall_t	0.604	0.321	0.618	0.234	0.598	0.599
Spearman_t	0.799	0.423	0.812	0.346	0.768	0.795

On the basis of two rank correlation coefficient, we can know the quantitative relationship between the rest of the wind field and the 9th. In order to better demonstrate the validity of this relationship, we draw the scatter plot of rank correlation coefficient and distance. As shown in figure 3, it is fitted using linear mode.

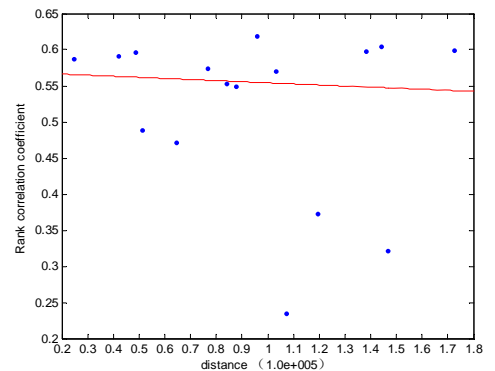


FIGURE III. COEFFICIENT AND DISTANCE RELATIONSHIP

As the scatter plot distribution show, with the increase of distance, rank correlation coefficient decreases markedly, and it is accordance with the actual. We use Kendall coefficient for this figure, in fact, when using the spearman coefficient, the conclusion is the same.

By selecting of the above methods, we finally choose 3th , 4th , 5th , 6th , 7th , 11th , 14th and 19th wind fields as input.

### C. Nonlinear State Space Model

In this paper, a discrete-time nonlinear stochastic system is considered, where  $x_k \in \mathbb{R}^{n_x}$  and  $z_k \in \mathbb{R}^{n_y}$  represent the state and the measurement at time instant  $k$ , respectively,  $f_k: \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$ ,  $h_k: \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_y}$  are known nonlinear functions, and  $w_k \in \mathbb{R}^{n_x}$  and  $v_k \in \mathbb{R}^{n_y}$  are zero-mean state and measurement white noises with covariance matrices  $\Sigma_k^w$  and  $\Sigma_k^v$ .

$$\begin{aligned} x_{k+1} &= f_k(x_k) + w_k, \quad k = 0, 1, 2, \dots \\ z_k &= h_k(x_k) + v_k, \quad k = 0, 1, 2, \dots \end{aligned} \quad (2)$$

For this model, the most difficult is how to get the two nonlinear functions. In the literature[10], based on support vector machine regression to get these two functions. Here, we use this method.

### III. PREDICTION RESULTS AND ERROR ANALYSIS

Here, we choose the speed data of December 18, 2009 to December 24, 2009 as the training data to predict the speed data of December 25, 2009. Using the above modeling method, we can get the prediction results in figure 4.

By calculating, the root-mean-square of our method is 0.4635. In order to verify the prediction accuracy of the methods, we give the predict residual histogram of Kalman filter and the wavelet neural network.

It can be drawn from figure 4, the trend of predicted wind speed and actual wind speed is basically the same, only in a few points, the error is a little bigger. What more, on the basis on figure 5, 6, and 7, it can be seen that the residual of the hybrid model is more concentrated around zero.

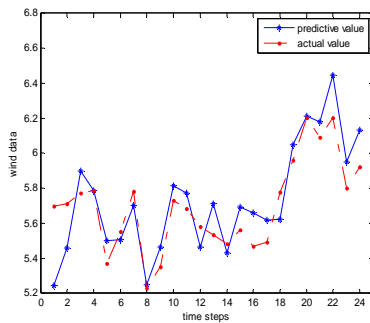


FIGURE IV. PREDICTED RESULTS FIGURE

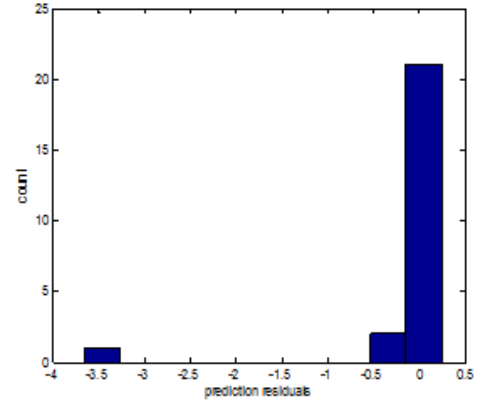


FIGURE V. HYBRID MODELS TO PREDICT RESIDUAL

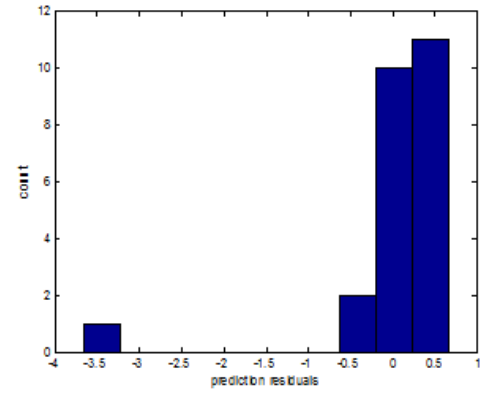


FIGURE VI. KALMAN FILTER PREDICT RESIDUAL

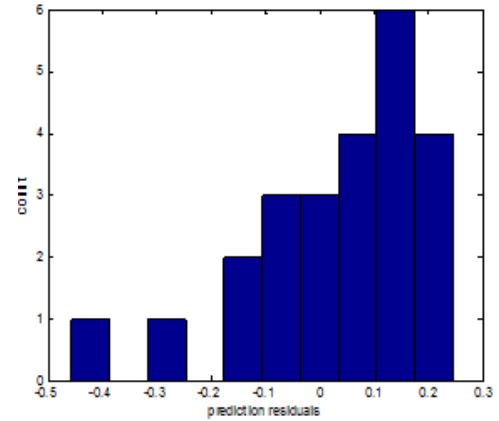


FIGURE VII. WAVELET NETWORK PREDICT RESIDUAL

### IV. CONCLUSION

Based on the spatial correlation between each wind field, this paper uses the copula function to calculate the rank correlation coefficient. Then, using the strength of the correlation to choose the input data. Finally, the nonlinear state space model is used for the wind speed prediction. This paper shows the hybrid model can improve the prediction accuracy

effectively. But, when using in nonlinear state space model, this method has a bigger shortage, latter we will study in this field

#### ACKNOWLEDGMENT

This work is supported by the state grid corporation of science and technology projects: Marine electric (group) power prediction algorithm research; The national natural science fund project (No:71271165).

#### REFERENCES

- [1] CHEN Ying, ZHOU Hai, WANG Wenpeng, Improvement of ultra-term forecast of wind power [J].Automation of Electric Power Systems,2011,35(15):30-33.
- [2] ALEXIADIS M C,DOKOPILOS P S,SAHSAMANOGLUOH S.Short-term forecasting of wind speed and related electrical power[J].Solar Energy,1998,63(1):61-68.
- [3] Bossanyi E A. Short-term wind prediction using Kalman filters[J]. Wind Engineering, 1985,9(1):1-8.
- [4] Kamal L, Jafri Y Z. Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan[J]. Solar Energy, 1997, volume 61(1):23-32(10).
- [5] Jun L I, Zhao F. Time Series Prediction Using Support Vector Regression Neural Networks[J]. Journal of System Simulation, 2008, 20(15):4025-4030.
- [6] LIU Ruiye, HUANG Lei. Wind power forecasting based on dynamic neural networks[J].Automation of Electric Power Systems,2012,36(11):19-22.
- [7] SHI Hongtao, YANG Jingling,DING Maosheng, A short-term wind power prediction metod basd on wavelet decomposition and BP neural network [J]. Automation of Electric Power Systems,2011,35(16):44-48.
- [8] YE Lin,LIU Peng. Combined model based on EMD-SVM for short-term wind power prediction[J].Proceedings of the CSEE,2011,31(31):102-108.
- [9] KANG Yali,ZHANG Zongyi,GUO Xinglei. Portfolio risk analysis in electricity market based on coupla approach[J].Power System Protection and Control,2012,40(6):50-56
- [10] Chen K, Yu J. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach[J]. Applied enEnergy, 2014, 113(6):690-705.