J. Wind Eng. Ind. Aerodyn. 153 (2016) 26–33

Contents lists available at ScienceDirect

Journal of Wind Engineering and Industrial Aerodynamics

journal homepage: www.elsevier.com/locate/jweia

Wind power prediction method based on regime of switching kernel functions

Tinghui Ouyang*, Xiaoming Zha, Liang Qin, Yi Xiong, Tian Xia

School of Electrical Engineering Wuhan University, Wuhan 430072, China

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form
13 March 2016
Accepted 13 March 2016
Available online 28 March 2016

Keywords:
Wind power prediction
Time series
Kernel functions
Support vector machine
Switching regime

ABSTRACT

The fluctuation of wind cause threat to power grid, this paper proposed a wind power prediction method to improving this situation. The proposed method is based on time series and regime of switching kernel functions. First, the mutual information method and the false nearest neighbor method were used to calculate parameters to reconstruct the original data. The recurrence figure and the Lyapunov exponent were applied to verify that the time series data was from a chaotic system. Then, this paper proposed a prediction method based on the kernel function and also a switching regime based on the support vectors machine. The new prediction method combining these two parts was proposed to predict wind power. The comparison of wind power prediction by the proposed method and traditional methods were present, the results validated that the proposed method is feasible to predict wind power, and that the precision of prediction is improved, which will be useful for the future analysis of wind power.

1. Introduction

As the energy crisis becomes more serious currently, renewable energy has been developed vigorously around the world (Council, 2013). Among these renewable sources, wind energy has been widely utilized for power generation in power grid due to its richness and exploitability. Because of the randomness, fluctuation and uncertainty of wind, great challenges are brought to the power grid, especially in the large-scale wind farms (Yang et al., 2012). For instance, a large down-ramp event was reported to cause destructive harm to the electric power system in Texas of America, 2008 (Francis, 2008). If these wind events or variance of wind power are predicted in advance, valid control strategy would be taken by power system operators and power loss would be avoided. In addition, more analysis about wind power are studied based on prediction to ensure the safety and stability of power system, e.g. ramp analysis. Therefore, it is important to predict wind power accurately.

Currently, the methods on wind power prediction are mainly divided into two models: physical models and statistical models (Monteiro et al., 2009). Generally, physical models utilize the results of numerical weather prediction (NWP) system to predict wind speed which was computed into wind power base on wind power curve (Chen et al., 2014). Since NWP systems are based on physical formulations, it is effective to capture the developing trend of wind in the long term. The drawback of these models is the low precision in local prediction. Statistical models obtain high precise prediction performance in the short term. They train the prediction models based on the history inputs and outputs, and find the optimal model parameters (Zheng and Kusiak, 2009), such as the time series models (Croonenbroeck and Ambach, 2015), neural network (NN) models (Saavedra-Moreno et al., 2013), Kalman model (Stathopoulos et al., 2013), support vector machine (SVM) model (Liu et al., 2012) and so on. It is common among these models that history wind data is used as the research basis. Time series data is extracted from original systems. It not only behaves some feature in exterior, e.g. the variation trend, but also contains some invisible information of original systems. Therefore, it validly explains the reason why statistical models based on time series data perform well in short-term prediction. In other word, the essence of statistical models is to study the nature from the phenomenon. Time series is acquired from industrial systems including stochastic system or deterministic system. The atmospheric dynamics system is a deterministic system following physical principles. Wind is a basic element of atmospheric system, it is reasonable to be predicted. Some random factors also affect the formation of wind (Lei et al., 2008), which is the main limitation for improving wind power prediction performance. In general, the average error of statistical prediction models located at between 25% and 40% (Mohandes et al., 2014).

* Corresponding author. Tel.: +86 15171508249.
E-mail address: thouyang@whu.edu.cn (T. Ouyang).

http://dx.doi.org/10.1016/j.jweia.2016.03.005
0167-6105/

To improve the performance of wind power prediction, a model based on prediction of time series is proposed in this paper. Professor Lorenz expounded that the atmospheric system belongs to chaotic systems in Chatterjee and Biswas (2015). Wind power time series was validated from a chaotic system and used in short-term prediction (Lei et al., 2008). The models based on time series of chaotic systems are helpful for prediction. Some physical features are mined to analyze the original systems, and direct to select suitable statistical algorithms to improve prediction performance.

Compared with those models predicting time series of chaotic systems, a new method based on switching regime of kernel functions was proposed in this paper. Usually, the local linear model, the maximum Lyapunov exponent method and the Volterra method are applied to predict time series of chaotic systems. These methods have some advantages and disadvantages in different applications. For example, the former two models predict trend of time series based on the historical similar tracks. However, chaotic systems are sensitive to the selected initial values, and similarity is only useful to improve the performance in short term. The Volterra method is a non-linear algorithm, it was useful to improve the precision of prediction. However, nonlinear models have troubles to train parameters in high-order models. To overcome these shortages of nonlinear models, kernel functions were introduced to map a nonlinear model of low dimension space into a linear model of high dimension space in Lu et al. (2013). A weighting method was proposed in Zhong-Da et al. (2014) to average the performance of different kernel functions. Considering the different performance of prediction models, a regime of Markov chain is applied to select models according to transition probability in Sorensen et al. (2009). Summarizing the advantages of these models, a method based on deterministic regime of switching kernel functions instead of Markov chain was proposed to avoid the uncertainty of probability. In case study, industrial data were taken to verify the validity of the proposed method.

To realize the wind power prediction method based on the regime of switching kernel functions, the rest parts of the paper is organized as follows. Section 2 presents the algorithm of phase space reconstruction, which is the basis of research on time series of chaotic systems. Section 3 tests the certainty and predictability of the system which wind power data belongs to before building the prediction model. Based on the results of Sections 2 and 3, a prediction model based on regime of switching kernel functions was established in Section 4. Section 5 took the industrial data as study case to realize wind power prediction and validated the performance of the proposed method.

2. Reconstruction of phase space

Before building the model predicting time series of chaotic systems, the first task is to reconstruct the phase space of data. In Packard et al. (1980), Pakard thought that any variable time series of a given system contain the dynamics information of the original system. Therefore, when embedding a time series into a new space, the system in new space still reserves the features of the original system. This was called these two systems were diffeomorphism, and dynamic behaviors of these two systems are equal (Takens, 1981). Based on the Takens embedding theory (Takens, 1981), wind power time series is reconstructed into a new phase space according to (1).

$$\mathbf{x}_n = (x_n, x_{n+\tau}, \cdots x_{n+(m-1)\tau}) \in R^m,$$
$$n = 1, 2\cdots, N_0 = N-(m-1)\tau \tag{1}$$

where: $\{x_n\}$ is the original wind power time series; $\mathbf{x}_n$ is the reconstructed vector, belonging to a $\boldsymbol{R}^m$ space; N is the length of original time series; N_0 is the number of reconstructed vectors. τ and m are the parameters representing the delay time and embedding dimension, respectively. If the time series is infinite ($N \to \infty$) and with no noise, there are abundant choice for the delay time τ according to the condition $m \geq 2d+1$. Noise is an unavoidable element in industrial time series, so de-noising process is needed before selecting parameters of delay time and embedding dimension.

In this paper, the industrial wind power data from Bonneville Power Administration (BPA) website (http://transmission.bpa.gov/Business/Operations/Wind/WindGenTotalLoadYTD_2013.xls) is taken as the study cases. The data set spanning from 01/01/13 00:00 to 06/30/13 23:55 totally has 52,116 data points with a sampling interval of 5 min. The data of former four months is taken as training set, the data of rest two months is taken as validation set.

2.1. Delay time

The purpose of selecting delay time τ is to reconstruct the original system effectively. There are some algorithms to select the delay time, e.g. autocorrelation function and mutual information method (Fraser and Swinney, 1986). Since autocorrelation function is not suitable to analyze non-linear time series or that with small sampling intervals, the mutual information method is applied in this paper.

Assuming a wind power time series $\{x_n\}$, the data point having a delay time τ with data point x_n is defined as $x_{n+\tau}$, the mutual information value of these two data is computed by (2).

$$\begin{aligned} I(\tau) &= \sum_{n=1}^{N} P(x_n, x_{n+\tau}) \log \left[\frac{P(x_n, x_{n+\tau})}{P(x_n)P(x_{n+\tau})}\right] \\ &= -\sum_n P(x_n) \log\ P(x_n) - \sum_n P(x_{n+\tau}) \log\ P(x_{n+\tau}) \\ &\quad + \sum_n P(x_n, x_{n+\tau}) \log\ P(x_n, x_{n+\tau}) \end{aligned} \tag{2}$$

where: $I(\tau)$ represents the mutual information function; $P(*)$ represents probability density function; $\log(*)$ represents the logarithmic function. Assuming a variable s, $H(s)$ is defined as the information entropy to express $-\sum_i P(s_i) \log\ P(s_i)$. Then the mutual information in (2) is rewritten through forms of information entropy in (3).

$$I(q,s) = H(q) + H(s) - H(q,s) = H(q) - H\left(q \middle| s\right) \tag{3}$$

where: s and q represent variable x_n is defined as $x_{n+\tau}$, respectively. The mutual information $I(q,s)$ describes the dependence between s and q. $H(q|s)$ represents the uncertainty of q when s is given. If its value is small, the mutual information value $I(q,s)$ will be large, and vice versa.

According to (2) and (3), it is necessary to calculate the probability of $P(x_n)$, $P(x_{n+\tau})$ and $P(x_n,\ x_{n+\tau})$, which are obtained from the statistical analysis of history wind power data. Fraser and Swinney (1986) proposed a meshing method to speed up the process of calculating $P(x_n, x_{n+\tau})$. A suitable value τ is important. If the value of τ is too small, the reconstructed track is close to the diagonal line in phase space, which means the reconstruction is not effective. Conversely, if the value of τ is too large, the track may be mapped into two irrelevant dimensions, which would not reflect the evolution of original system effectively, too.

Fig. 1 depicts the variance of mutual information entropy when delay time increasing, and the value $I(x_n,\ x_{n+\tau})$ of training data set has a decreasing trend. Generally, the delay time τ is decided when $I(x_n,\ x_{n+\tau})$ reaches the first local minimum value. Therefore, the delay time of the given time series for phase space reconstruction is $\tau = 9$, and the selected two variables will have the largest independence at this time.