Applied Energy 143 (2015) 472–488

Contents lists available at ScienceDirect

Applied Energy

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

The study and application of a novel hybrid forecasting model – A case study of wind speed forecasting in China

Jian-Zhou Wang [a], Yun Wang [b,*], Ping Jiang [a]

[a] School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China
[b] School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, Gansu Province, China

HIGHLIGHTS

- The Wavelet Packet Transform (WPT) methodology is exploited to process the original wind speed data.
- Phase Space Reconstruction is performed to select the input form.
- The LSSVM model whose parameters are tuned by an artificial intelligence (PSOSA) model is built to make forecast.
- Grey Relational Analysis and hypothesis test also indicate that the proposed model has the better forecast performance.

ARTICLE INFO

Article history:
Received 22 May 2014
Received in revised form 8 January 2015
Accepted 9 January 2015

Keywords:
Wavelet Packet Transform
Least Square Support Vector Machine (LSSVM)
PSOSA algorithm
Grey Relational Analysis
Hypothesis test

ABSTRACT

Given the current increasingly serious energy crisis, the development and utilization of new energy resources are attracting increasing attention, and wind power is widely used among these renewable energy resources. However, the randomness of wind power can cause a series of problems in the power system. Furthermore, the integration of large-scale wind farms into the whole power grid can place a great burden on stability and security. Accurate wind speed forecasting would reduce the randomness of wind power, which could effectively alleviate the adverse effects on the power system. In this paper, a hybrid wind speed forecasting model is proposed with the hope of achieving better forecasting performance. Wavelet Packet Transform (WPT) was employed to decompose the wind speed series into several series with different frequencies. A Least Square Support Vector Machine (LSSVM), the parameters of which were tuned by a particle swarm optimization based on simulated annealing (PSOSA), was built to model those series. The optimal input form of the model was determined by Phase Space Reconstruction (PSR). To verify the effectiveness of the proposed model, the daily average wind speed series from four wind farms in Gansu Province, Northwest China, were used as a case study. The results of the simulation and Grey Relational Analysis indicate that the proposed model outperforms the comparison models, and the null hypothesis of the predicted series having the same mean of the real series was accepted.

1. Introduction

Along with the continuous increase in world energy consumption and vigorous development in traditional energy resources, the storage of fossil fuel is on the decrease, and the worldwide energy crisis is gradually becoming significant. Therefore, alleviating the energy crisis, developing renewable energy sources and achieving the sustainable development of energy has become a major initiative of the world's energy development strategy. Wind energy, an important category of renewable energy, is abundant, renewable, widely distributed and clean, resulting in wind power becoming an important renewable energy development direction. Presently, wind power is widely used not only in the developed countries but also in many developing countries, and even in some developed countries, wind power has partly replaced the traditional power generation modes and is providing the basic driving force of economic development.

China is the world's largest developing country and has abundant wind energy resources. With increasingly mature wind power technology and the government's strong support, wind power has become the fastest growing renewable energy in the country. According to the plan of the Chinese government, by 2020, the installed capacity of wind power will reach 30 GW [1]. In today's rapid development of wind power generation, the proportion of wind power in the whole power system is becoming larger.

* Corresponding author. Tel.: +86 18293191191; fax: +86 931 8912481.
E-mail address: wangyun12@lzu.edu.cn (Y. Wang).

http://dx.doi.org/10.1016/j.apenergy.2015.01.038
0306-2619/

However, due to the randomness and intermittent of wind energy, random wind speed and wind direction lead to evident fluctuations of wind turbines' output power, such fluctuations adversely affect the grid frequency and voltage stability. When the proportion of wind power reaches a certain amount, it poses serious challenges to the security and stable operation of the power system and to the quality of the generated electric energy. In addition, to cope with the intermittent and random nature of wind power, sufficient backup power is necessary to protect the normal power supply to the users, which results in increases of the reserve capacity of the power system, which undoubtedly increases the operation cost of the power system. Thus, accurate wind speed forecasting can enhance the foreseeability of otherwise random wind power, reducing the reserve power requirements and accordingly increasing the reliability of the grid. A reduction of the operation cost and the spinning reserve would make it possible to increase the proportion of wind power in the grid.

To avoid the challenges of wind power integration, relatively accurate wind speed and wind power forecasting are extremely important. Currently, wind speed forecasting errors are approximately in the 25–40% range in Chinese wind farms [2,3]. These results are not satisfactory, and they are related not only to the forecasting methods but also to the forecasting period. Based on the requirements of wind power operation, forecasts can be grouped into four horizons [4,5]: ultra-short term, short term, mid and long term and long term. Ultra-short-term forecasts and short-term forecasts are mainly used for load tracking and preload sharing. The power system management and maintenance scheduling of the wind turbines are conducted using mid and long-term and long-term [6,7] forecasts, respectively.

Recently, many researchers have thoroughly studied wind speed and wind power forecasts, and many methods have been proposed and applied to wind farms. These methods can be classified into four categories [8]: physical models, statistical models, spatial correlation models, and artificial intelligence models.

Physical models not only make use of historical data but also consider the weather and geographical condition to aid wind speed forecasting with the expectation of achieving better forecasting accuracy [9,10]. On the contrary, statistical models, which are known as stochastic time series models, only employ historical wind speed. This model method is easy to apply and simple to implement. Thus, several types of time series models commonly appear in wind speed forecasting, including the autoregressive model (AR), moving average model (MA), autoregressive moving average model (ARMA) [11], and autoregressive integrated moving average model (ARIMA) [12]. A survey of the literature leads to the conclusion that [4] the statistical models can perform well when applied to short-term, medium-term, and long-term wind speed forecasting in the vast majority of case studies, whereas the physical models present satisfactory results in the ultra-short-term and short-term horizons. Typically, the spatial correlation models [13,14] are mostly used when the available study information for wind farms is not sufficient, but the essential information in several adjacent wind farms is available. Unlike other models, to build the spatial correlation wind speed forecasting models, the wind speed and other necessary information, which contain delay times, have to be measured from multiple spatial correlated sites. Thus, the measurements and their delay times add complexity and cost to the implementation of spatial correlation forecasting. Recently, along with the rapid development and popularization of the artificial intelligence technology, it is common to apply different intelligence algorithms, including artificial neural networks (ANNs) [15–19], Support Vector Machine (SVM) [20–22] and fuzzy logic methods [23,24], to wind speed forecasting.

Hybrid wind speed forecasting models [25–28] are commonly and extensively adopted with satisfactory forecasting results. The reason for this that such models can mine the hidden information of different wind speed time series to a large extent. According to many studies, there are at least three methods that can enhance the wind speed forecasting accuracy. First, before being entered into the forecasting model, the original wind speed series is processed to achieve a relatively higher forecasting accuracy. The Wavelet Transform (WT) [29–31] and Empirical Mode Decomposition (EMD) [19,32,33] are the most common technology for processing wind speed series. The former can remove the irregular fluctuation of the original series, whereas the latter can decompose the original series into several intrinsic model functions (IMFs) for modeling. Second, as is well-known, the input form has a significant effect on the forecasting results. On occasion, the input form is decided by trial and error and the researchers' experience. In many papers, the Partial Auto Correlation function (PACF) [32,33] is applied to choose the best input form. In this paper, we use Phase Space Reconstruction [34] to determine the input form. Third, the parameters of models play a significant role in the forecasting accuracy of the modeling process. With the development of science and technology, common intelligent optimization algorithms, such as Genetic Algorithms (GAs) [29,35], simulated annealing (SA) [36,37], Practical Swarm Algorithms (PSOs) [5,35], and Ant Colony Optimization (ACO) [38] are widely used in parameter optimization. In addition, many papers have proved that the models that are optimized by intelligent optimization algorithms have a better forecasting performance.

In this paper, a hybrid model is proposed with the expectation of more precise forecasting of wind speed. In comparison with the WT methodology, the WPT methodology, which decomposes not only the low-frequency series but also the high-frequency series, can mine the traits of the original series more meticulously. First, the WPT methodology is adopted to process the original wind speed series to enhance the forecasting capacity, and several series with different frequencies that contain different characters of wind speed can be achieved. Then, Phase Space Reconstruction is performed for all the sub-series to select the input form except for the sub-series with the highest frequency, and the embedding dimension and delay time is calculated by the C–C method because of its easy operation and minimal calculation requirements. Finally, Least Square Support Vector Machine (LSSVM) models are built to forecast the selected sub-series. As all parameters in the LSSVM model haves a significant effect on the forecasting accuracy, the particle swarm optimization based on simulated annealing (PSOSA), which has the ability to avoid falling into local extreme points compared with the basic PSO algorithm, is employed to optimized the two parameters in the LSSVM model. The wind speed forecasting values can be obtained by adding up the forecasting values of all selected sub-series except for the one with the highest frequency. To evaluate the effectiveness of the hybrid method, case studies from four wind farms located in Gansu Province, Northwest China were conducted. In addition, Grey Relational Analysis and statistical hypothesis tests were used to evaluate the rationality of the forecasting series produced by the proposed model.

The advantages of the proposed model, which result in the better forecasting performance, are represented in the following several aspects. To begin with, many single methods make wind speed forecasting by using raw wind speed series directly, but the forecasting accuracy is not very satisfactory due to the influence of random noise in original series. In this paper, WPT is employed to preprocess the original wind speed series and reduce the effect of random noise. Then, the input form determination in the proposed model is smarter and more novel. By reviewing many papers, we found that determination of input form is based either on ones' experience or on the figure of PACF [29]. However, the fluctuation of wind speed affected by many factors such as temperature, pressure, and humidity is a