Renewable Energy 76 (2015) 91–101

Contents lists available at ScienceDirect

Renewable Energy

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

Medium-term wind speeds forecasting utilizing hybrid models for three different sites in Xinjiang, China

Jianzhou Wang [a], Shanshan Qin [b, c, *], Qingping Zhou [c], Haiyan Jiang [c]

[a] School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China
[b] MOE Key Laboratory of Western China's Environmental Systems, Research School of Arid Environment and Climate Change, Lanzhou University, Lanzhou 730000, China
[c] School of Mathematics & Statistics, Lanzhou University, Lanzhou 730000, China

ARTICLE INFO

Article history:
Received 7 March 2014
Accepted 5 November 2014
Available online

Keywords:
Wind speed forecasting
Outlier detection
Support vector regression
Elman recurrent neural network
Hybrid model

ABSTRACT

Interest in renewable and clean energy sources is becoming significant due to both the global energy dependency and detrimental environmental effects of utilizing fossil fuels. Therefore, increased attention has been paid to wind energy, one of the most promising sources of green energy in the world. Wind speed forecasting is of increasing importance because wind speeds affect power grid operation scheduling, wind power generation and wind farm planning. Many studies have been conducted to improve wind speed prediction performance. However, less work has been performed to preprocess the outliers existing in the raw wind speed data to achieve accurate forecasting. In this paper, Support Vector Regression (SVR), a learning machine technique for detecting outliers, has been successfully combined with seasonal index adjustment (SIA) and Elman recurrent neural network (ERNN) methods to construct the hybrid models named PMERNN and PAERNN. Then, this paper presents a medium-term wind speed forecasting performance analysis for three different sites in the Xinjiang region of China, utilizing daily wind speed data collected over a period of eight years. The experimental results suggest that the hybrid models forecast the daily wind velocities with a higher degree of accuracy over the prediction horizon compared to the other models.

1. Introduction

Over the past few decades of rapid economic development and industrialization, China has become a large energy consumption power. Fossil fuels, primary energy resources that play a crucial role in meeting energy demands, are non-renewable resources that generate large amounts of greenhouse gases, which have caused an emerging crisis in global climate warming [1]. However, the rapid depletion of fossil fuel reserves and increasing environmental problems encouraged policy makers and planners to search for substitutes from renewable energy sources [2]. Wind energy, as a promising source of renewable and green energy, has received increasing attention due to its inexhaustibility, sustainability, ecological awareness and contribution to energy security. With the global installed wind turbine capacity increasing exponentially, the wind power industry has gradually expanded into a large-scale business [3].

Large wind power penetration, however, inevitably induces numbers of problems to be solved, such as management of wind power production, power quality, power scheduling, turbine maintenance, market integration and the stability and reliability of power grid system, etc. And the cubic relationship between wind power and wind speed necessitates accurate wind speed forecasting. Otherwise, any deviation in the wind speed forecasts will cause a large deviation in wind power generation [4]. For instance, in an electricity market, a 10% deviation in wind speed forecasts leads to a 30% deviation in wind power generation [5]. Thereby, precise wind speed forecasting is crucial for the various respects of the wind industry, including such as wind turbine predictive control, wind power grid integration and economic dispatch, wind farm siting and so on [3].

* Corresponding author. MOE Key Laboratory of Western China's Environmental Systems, Research School of Arid Environment and Climate Change, Lanzhou University,Lanzhou 730000, China.
E-mail address: qinsslzu@gmail.com (S. Qin).

http://dx.doi.org/10.1016/j.renene.2014.11.011
0960-1481/

Nevertheless, wind speed, with its intrinsically highly irregular and volatile series, is considered one of the most difficult weather parameters to model and forecast due to its random and intermittent nature [6,7]. And numbers of researches have devoted large efforts underway to develop high quality models for wind speed prediction.

According to the prediction time horizon and the purpose of wind forecasting, wind speed forecasting can be roughly classified into three categories [8]: short-term, medium-term and long-term forecasting. Short-term forecasting (over minutes, hours or days) is generally utilized in the control of wind turbines, real-time grid operations, regulation, economic load dispatch planning and load reasonable decisions and operational security in electricity market; Medium-term forecasting (over weeks to months) is primarily utilized in maintenance planning, operation management and optimizing operating costs; Long-term forecasting (over years) is usually utilized in wind farm design and calculating the annual generating capacity of wind farms. Up to now, much work has been conducted on wind speed forecasting of different time-scales. These efforts have focused on short-term forecasting [3,4,7–29], and despite growing awareness of the significance of medium-term and long-term wind prediction, relatively fewer studies have been performed [6,30–34]. Forecasting tools specific to the time-scale are not abundant because long-term prediction is difficult [35].

According to the references, the models proposed and launched on wind farms in world-wide locations are based on physical, statistical and spatial correlation approaches [36]. Physical methods are based on the lower atmosphere or numerical weather prediction (NWP) utilizing weather prediction data such as pressure, temperature, obstacles and surface roughness [9–13,37]. Statistical methods draw on vast historical data without considering meteorological conditions, which usually involves time series analysis [14,15,38,39] and artificial intelligence, i.e., artificial neural networks (ANNs) and support vector machine (SVM) [16,17,25,30,31,35,40,41]. Spatial correlation approaches utilize multi-dimensional data from several study stations to forecast wind speeds [36]. ANNs are data-driven, non-parametric models, which can be useful for wind speed forecasting because they can capture subtle functional relationships existing in the empirical data even though the underlying relationships are unknown or difficult to describe [18]. Special recurrent neural networks, Elman neural recurrent networks (ERNN), have been studied extensively by researchers in time series forecasting and nonlinear identification [42]. The recurrent Elman architecture provides considerable nonlinear mapping ability, which motivates the utilization of ERNN for medium-term wind speed prediction in this study.

Recently, hybrid or combination model, the basic idea of which is to combine different approaches, retaining strengths of each method, to improve model's forecasting performance [43], has been widely utilized in various fields, especially for wind energy/power forecasting [19–25,34,36,44,45]. In practice, as compared to single technique models hybrid models tend to yield more accurate forecasts by taking advantage of different techniques while mitigating the restrictions of each model [43].

Although the above hybrid models have demonstrated admirable predictive performance, no models has been determined to be superior because wind patterns can vary significantly not only on different wind farms but also at different times or heights on the same farm [6]. In other words, no model structure outperforms the others for all farms, periods and heights, which suggests that potential models should be constructed and evaluated for different wind speed datasets to detect the most appropriate model.

Wind speed time series are highly volatile and non-stationary, which complicates its forecasting. It is well known that the input samples play a significant role in modeling; however, forecasting wind speeds with abnormal data will yield inaccurate estimates of the wind speed fluctuation trend, which is usually subject to deviation and errors in prediction. Therefore, removing abnormal data is a meaningful and significant procedure prior to modeling. The dominant methods for detecting and removing abnormal data are Wavelet De-noising (WD) [20,23,24,29] and Empirical Mode Decomposition (EMD) [35,44,45]. However, the support vector regression (SVR) technique, a tool for detecting outliers, is increasing in popularity. For this reason, this study applies SVR to preprocess abnormal data in the original wind speed series.

In this paper, a novel hybrid model to forecast medium-term wind speeds is developed by incorporating SVR and SIA into the ERNN method. In an application of the proposed hybrid model, daily wind speed data are collected from three observation sites in the Xinjiang region during the period 2005–2012. Daily wind speeds for autumn 2012 are forecast using daily wind speed values obtained in the corresponding quarter from the period 2005–2011. The empirical results clearly demonstrate that the proposed hybrid model outperforms methods that utilize the raw wind speed series, that is, without abnormal data preprocessing and seasonal index adjustment.

The main contributions of this study are as follows:

(1) This study presents an effective approach based on the SVR technique to address the general phenomenon of outliers in wind speed series.
(2) The additive and multiplicative SIA techniques are proposed to separate the seasonal and trend components.
(3) K–W test is utilized to identify any significant differences among the input datasets.
(4) The proposed approach can improve forecasting performance significantly in three different study sites.

The rest of this paper is organized as follows: Section 2 presents a brief analysis of the study sites and wind speeds. Section 3 describes the related methods, including SVR, K-W test, SIA and ERNN. A detailed case study to illustrate the process of proposed models is given in Section 4. Section 5 provides the predictive results comparison. Finally, conclusions are drawn in Section 6.

2. Brief analysis of study sites and wind speeds

The Xinjiang Uygur Autonomous Region, which accounts for more than one-sixth of China's total territory, possesses abundant wind resources. The total area devoted to the development and utilization of wind power is approximately 150,000 km^2, and the total installed capacity is more than 80 million kw. By the end of September 2012, the installed wind capacity of the Xinjiang region was 1.907 million kw, which had grown seven times compared to capacity in 2006 [46]. According to preliminary planning, by 2020, total wind power installed capacity will exceed ten million kw, and wind power delivery goals will be achieved [47]. In this paper, three different sites in the Xinjiang region have been selected as case our studies, with the latitude and longitude of 45°N and 87.5°E, 42.5°N and 87.5°E, 40°N and 90°E.

The wind speed data utilized in this paper are sampled from three observation sites in the Xinjiang region during the period 2005–2012. The dataset sources are from the NCEP/NCAR 40-Year Reanalysis Project [48] with a standard height 10 m. Fig. 1 displays the basic statistical measures (maximum/minimum, average, standard deviation) of wind speed in three different study sites. As