Energy Conversion and Management 148 (2017) 554–568

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Non-parametric hybrid models for wind speed forecasting

Qinkai Han [a], Fanman Meng [b], Tao Hu [b], Fulei Chu [a,*]

[a] The State Key Laboratory of Tribology, Tsinghua University, Beijing 100084, China
[b] Department of Mathematics, Capital Normal University, Beijing 100048, China

ARTICLE INFO

Article history:
Received 27 March 2017
Received in revised form 26 May 2017
Accepted 7 June 2017

Keywords:
Wind speed forecasting
Hybrid algorithm
Non-parametric modeling
Autoregressive moving average

ABSTRACT

It is essential to predict the wind speed accurately in order for protecting the security of wind power integration. The aim of this study is to develop non-parametric hybrid models for probabilistic wind speed forecasting. By adopting the non-parametric models, two hybrid models, namely the hybrid autoregressive moving average/non-parametric and hybrid non-parametric/autoregressive moving average models, are proposed and their performance is compared. In the hybrid autoregressive moving average/non-parametric models, the residuals obtained after fitting with the autoregressive moving average model are studied by the non-parametric model to extract the nonlinear part of the data. In the hybrid non-parametric/autoregressive moving average models, the residuals obtained from the non-parametric model are fitted by the autoregressive moving average model. In order for comparisons, the artificial neural network with back propagation, support vector machine and random forest models are also introduced for hybrid modeling. Through conducting various tests on the real hourly wind speed time series, the prediction performance of both single and hybrid models is compared and evaluated in detail. The results of this study are to show that non-parametric based hybrid models generally outperform the other models and have more robust forecast performances. When the single autoregressive moving average model basically outperforms the single non-parametric models, the introduction of the autoregressive moving average model for the residuals from the non-parametric fitting is possible to obtain better prediction accuracy.

1. Introduction

Wind energy is a type of renewable energy that has been globally recognized as an effective method for mitigating climate change, improving energy security, and supporting low-carbon industrial and economic growth [1]. When connected with a power grid, wind energy can have a significant impact on power grid security, power system operation, and market economics due to its intermittent nature [2]. Especially in areas with high wind power penetration, the forecasting of wind speed and power generation plays a pivot role in attaining low-spinning reserves and optimal operational costs, and thus reducing fluctuation and instability [3]. Moreover, Athari and Ardehali [4] have shown potential applications of the wind speed forecasting to determine the optimal setting of fuzzy controller in a hybrid energy system, and more accurately model the uncertainties injected from high penetration of wind energy to the power system [5].

Numerous forecasting models have been proposed to the short-term prediction of wind speed. Autoregressive moving average (ARMA) family is one of the most robust and widely used approach in wind forecasting. Torres et al. [6] first introduced the ARMA model to forecast the hourly average wind speed series in Navarre (Spain). Shi and Erdem [7] expanded the ARMA model for predicting the wind speed and direction simultaneously. Then, they [8] conducted comprehensive evaluations of ARMA-GARCH(-M) approaches for modeling the mean and volatility of wind speed. Among the ARMA family, the more common models are the autoregressive (AR), autoregressive integrated average (ARIMA), fractional ARMA (fARMA), and seasonal ARMA (sARMA). The relations between the input and outputs could be explicitly described by these models. Although a novel nonlinear ARMA model [9] has been proposed recently, the ARMA models are generally considered in linear forms [10]. The other widely adopted models in wind speed forecasting are the artificial intelligence (AI) and machining learning (ML) models. Numerous AI/ML models have been presented in current literature, such as the artificial neural network (ANN) with back propagation (BP) and radial basis function (RBF) [11–13], support vector machine (SVM) [14,15], fuzzy logic [16,17], extreme learning machines [18,19], etc. Generally speaking, the AI/ML models can better handle nonlinear relationships and thus are more flexible. However, these models that rely on a

* Corresponding author.
E-mail address: chufl@mail.tsinghua.edu.cn (F. Chu).

http://dx.doi.org/10.1016/j.enconman.2017.06.021
0196-8904/

large number of historical data for constructing an input/output mapping function might be less effective in the case of inadequate available information. Following many studies in the wind forecasting field, it can be indicated that, to date, the targeted performance levels have not been attained with the individual models due to the fact that these models cannot give satisfactory results for all situations [20]. Therefore, the approaches that incorporate the individually superior features of various forecasting models have emerged, called as hybrid models, in order to obtain an advanced forecasting method for higher accuracy levels and wider forecast horizons.

Usually, the hybrid forecasting methodology employs a linear model for the prediction of the linear component and a nonlinear model for the nonlinear component in wind time series. In recent years, the hybrid models are receiving more and more attention. Cadenas and Rivera [21] first adopted the ARMA-ANN hybrid approach to forecast the wind speed for a fixed prediction horizon in different regions in Mexico. The improvement of prediction accuracy jumps 50% in terms of MAE and 85% in terms of RMSE compared with the individual ARMA or ANN models for all the tested sites. Shi, Guo and Zheng [10] systematically and comprehensively investigated the applicability of hybrid forecasting approaches for wind speed and power generation time series. Two hybrid models, namely ARIMA-ANN and ARIMA-SVM, were constructed to compare with the single ARIMA, ANN and SVM forecasting models. The results showed that the hybrid approaches are viable options for forecasting wind time series, but they do not always produce superior forecasting performance for all the time horizons investigated. Liu and Tian [22] proposed two new ARMA-ANN and ARMA-Kalman hybrid models and conducted detailed comparisons for their forecasting performance. Results showed that both of them have good performance and could be applied to the non-stationary wind speed prediction in wind power systems. Later, Liu and Tian [23] further introduced the wavelet and wavelet packet into the hybrid models and the results showed attractive. Wang and his collaborators [24–28] devoted lots of efforts to the investigation of hybrid forecasting modeling of wind speed series, and several hybrid models have been successfully used for wind speed prediction of various wind farms in China. A hybrid SARIMA-LSSVM model was significantly developed to predict the mean monthly wind speed in Hexi Corridor of China [24]. Compared with single models, the hybrid models shown powerful forecasting capacity for monthly wind speed prediction at wind parks. For improving the accuracy of an ARMA based hybrid model, the ordinary least squares algorithm was introduced to remove the outliers from original wind speed series [25]. Gaussian process regression [26] and extreme learning machine [27] were, respectively, used to make the ARMA based hybrid model more robust and self-adaptive. Later, based on a comprehensive review of the current hybrid models, two new hybrid models were proposed and the comparisons showed that the models can always provide desirable forecasting results [28]. Recently, Maatallah et al. [29] developed a new hybrid forecasting model for 1–24 h horizon by adapting the Hammerstein model into the AR model. Such nonlinear hybrid model successfully captured chaotic dynamics of wind speed time series. Shukur and Lee [30] carried out studies to improve the accuracy of wind speed forecasting by suggesting the ANN and Kalman filter into the ARMA model. Okumus and Dinler [31] presented an extensive review of recent advances in statistical wind forecasting. Then they combined the adaptive neuro-fuzzy inference system (ANFIS) and an ANN model for one hour ahead wind speed forecasts. In the latest studies, the Gaussian process regression [32] and support vector regression [33] were, respectively, adopted in the ARMA based hybrid models to improve not only the point forecast results, but also the prediction intervals.

As the above literature shown, most current hybrid forecast modeling of wind speed series adopts the ARMA models for the linear component prediction and the AI/ML models for the nonlinear component prediction. Besides the AI/ML models, the non-parametric models (such as the kernel density estimation, KDE) are also popular in the probabilistic modeling of wind speed series [34]. The most attractive feature of non-parametric models is that it directly makes use of sample data without a need of estimating characteristic parameters in a theoretical distribution [35]. In other words, there is no error caused by assumption of a theoretical distribution for wind speed and by mismatch between estimated parameters and actual behaviors of wind speed. Juban et al. [36] proposed a probabilistic prediction method based on the KDE. The performance of the model was evaluated using real-world data from French wind farms corresponding to different terrain complexity and climatic conditions. Bessa et al. [37] developed a novel KDE model, namely the time adaptive conditional KDE for wind power forecasting. By incorporating a time decay factor within the conditional KDE, Jeon and Taylor [38] proposed an approach to produce density forecasts for wind power at four Greek wind farms. In order to address two characteristics of wind power which have adverse impacts on the forecast accuracy, Zhang, Wang and Luo [39] presented a novel KDE for probabilistic wind power forecasting. The improvement of the proposed method over the standard KDE was demonstrated by short-term probabilistic forecasting results based on the data from an actual wind farm. Although the non-parametric models have been adopted in the probabilistic short-term wind power forecasting [36–39], its applications in wind speed prediction have been slow. It is even rare to see the non-parametric models in the hybrid forecasting.

Thus, the aim of this study is to develop non-parametric hybrid models for wind speed forecasting. A single non-parametric model based on the KDE is first adopted in the wind speed forecasting, and then hybrid methods for wind speed forecasting based upon the ARMA and non-parametric models are rigorously investigated. Three typical AI/ML models, including the ANN, SVM and random forest (RF) models, are also introduced for comparisons with non-parametric models. Two hybrid modeling strategies are proposed by integrating the advantages of the single linear and nonlinear forecast models. The hybrid models are compared with the individual ARMA, non-parametric forecasting models and AI/ML models by conducting various tests on the real hourly wind speed time series. The wind speed data is acquired from six wind farms of China. The reminder of the paper is organized as follows. In Section 2, the structures and procedures of the single and hybrid forecasting models are briefly explained. Three metrics are defined for performance evaluation. In Section 3, the wind speed time series are briefly introduced. In Section 4, the model parameters are obtained and the performances of different models are compared, and the results are analyzed and discussed. In Section 5, some conclusions are summarized.

2. Methods

The structures and procedures of the single forecasting models, including the conventional ARMA model and the non-parametric model, are explained in this section. Moreover, both the ANN, SVM and RF models are introduced briefly. Then, the framework for the hybrid models is proposed by integrating the advantages of the single models. Finally, several indicators are introduced for assessing the quality of the forecast.

2.1. Conventional ARMA model

ARMA models, first established by Box and Jenkins in 1976, have been widely used for the modeling of time series [40]. An