Energy Conversion and Management 151 (2017) 737–752

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Composite quantile regression extreme learning machine with feature selection for short-term wind speed forecasting: A new approach

Weiqin Zheng[a], Xiangang Peng[a,*], Di Lu[a], Dan Zhang[a], Yi Liu[a], Zhehao Lin[a], Lixiang Lin[b]

[a] School of Automation, Guangdong University of Technology, Guangzhou 510006, Guangdong, China
[b] Guangzhou Power Supply Bureau Co., Ltd, Guangzhou 510620, Guangdong, China

A R T I C L E I N F O

Keywords:
Wind speed forecasting
Composite quantile regression outlier-robust extreme learning machine
Feature selection
Hybrid particle swarm algorithm
Time adaptive filter based empirical mode decomposition

A B S T R A C T

With the growing wind penetration of wind resources into power generation worldwide, accurate and comprehensive wind speed forecasting (WSF) is becoming increasingly significant to ensure continuously economical and reliable power system operations. In this paper, a novel WSF framework based on composite quantile regression outlier-robust extreme learning machine (CQR-ORELM) with feature selection and parameter optimization using a hybrid population-based algorithm is developed. The CQR-ORELM model offers flexibility and efficiency to explore potential nonlinear characteristic among wind speed variables and improve model robustness and predictive capability. A hybrid algorithm with the combination of particle swarm optimization and gravitational search algorithm (PSOGSA) is utilized to fine-tune the optimal value of weights and bias in the ORELM network structure, while the binary version of PSOGSA (BPSOGSA) is exploited as a feature selection method to construct the most relevant input feature matrix for the model. Aimed to alleviate the influence of uncertainty in wind speed time series, a time adaptive filter based empirical mode decomposition (TVF-EMD) approach is used to decompose the original wind speed series into several intrinsic mode functions (IMFs). Each IMF is well-developed using the proposed method and the final forecasting results are obtained through aggregate calculation. This proposed wind forecasting methodology is compared to several benchmarks. The effectiveness of the proposed forecast strategy to predict wind speed is evaluated by the trials of 10 min ahead wind speed forecasting at two locations of the National Renewable Energy Laboratory (NREL). Comparative results confirm that the developed CQR-ORELM with feature selection procedure can describe the complete conditional distribution information hidden in variables and provide satisfactory wind speed information.

1. Introduction

Forecasting future wind speed is an important component in the management of modern energy systems. Accurate short-term WSF in the electricity market helps wind power suppliers optimize their bidding strategies to maximize profits [1,2]. System operators need short-term WSF to ensure adequate supply, thus operating the system securely, in terms of system stability, ancillary service and power quality [3].

Wind speed time series usually has complex and uncertain behavior. The complexities of wind speed series, such as its nonlinear and non-stationary behaviors, have been discussed in [2,4]. As Kani and Ardehali stated [5], the intermittent and stochastic characteristic of wind speed series requires more complex functions to capture the nonlinear relations, thus increasing the difficulty of wind speed prediction.

In the literature available, various forecasting approaches have been studied and proposed, each utilizing a different technique and performing well with a one and multi prediction horizon. Recent studies in the area of wind speed prediction predominantly focus on the short-term WSF, and the term of which ranges from minutes to a few days depends on the importance of these data to the power systems [1]. The state-of-the-art methods can generally be classified into different categories: (i) physical modes that usually based on numerical weather prediction (NWP); (ii) statistical models; (iii) intelligent approaches and (iv) hybrid (or combined) methods. NWP models simulate the physics of the atmosphere by utilizing physical laws and terrain conditions [6]. However, directly adopting the NWP model for WSF still suffers from some challenges, including accuracy, spatial and temporal resolutions, domain and hierarchical importance of the physical processes [1,7].Statistical models are identical to the direct random time-series models, including autoregressive (AR), and auto regressive integrated moving average (ARIMA) models [8,9]. The artificial intelligent (AI)

* Corresponding author at: Room 315, No. 2, Laboratory Building, Guangdong University of Technology, No. 100, West Waihuan Road, Higher Education Mega Center, Panyu District, Guangzhou City, Guangdong Province 510006, China

E-mail address: epxg001@163.com (X. Peng).

http://dx.doi.org/10.1016/j.enconman.2017.09.029
Received 12 July 2017; Received in revised form 7 September 2017; Accepted 9 September 2017

based WSF methods, such as artificial neural networks (ANN) [5,10,11] and support vector machine (SVM) [9,12,13], normally attempt to minimize training errors using the historical wind speed data. Due to the excellent ability of non-linear mapping, generalization and self-learning, AI-based methods prove to be of widespread utility in the WSF field. In the hybrid framework, the signal decomposition, individual approaches mentioned above and optimization technique are combined to forecast wind speed. A review of hybrid approaches can be found in [1]. Recently, another neural networks (NN) structure based regression model, known as extreme learning machine (ELM), has been selected to operate wind speed forecasting due to its fast learning speed and simple network structure [14–17]. In consideration of the good performance of ELM, an improved ELM in [18], termed as outlier-robust ELM (ORELM) network structure, is selected as a base forecasting engine in our proposed WSF approach.

Quantile regression (QR) analysis allows estimating various quantile functions of a conditional distribution. Each quantile regression characterizes a particular point such as the center point or tail point of a conditional distribution. This analysis is particularly useful when the conditional distribution is heterogeneous and has a nonstandard shape, such as an asymmetric, fat-tailed, or truncated distribution [19]. The outstanding merits of composite quantile regression neural network (CQRNN) have been proved in [20]. As demonstrated, the CQRNN model takes advantages of both CQR and ANN. It is not only flexible to explore the nonlinear relationship among variables through the property of ANN, but also efficient to improve the estimation and prediction ability of CQR. QR technique has been extensively applied in the field of short-term power load and electric price forecasting [21–24], yet is under-explored in the field of wind speed forecasting. Inspired by Li and Liu et al. [25] who proposed a support vector quantile regression (SVQR) model, we develop a novel ORELM based regression strategy by considering a composite quantile regression via ORELM, denoted by CQR-ORELM hereafter. It is the first time to propose and adopt the CQR-ORELM model to implement WSF.

Wind speed time series are complex signal and directly forecasting the wind speed data is usually subjected to large errors. In order to improve the performance of the single soft computing approach, hybrid approaches have been applied to the field of WSF. One popular attempt is to utilize data preprocessing technique to realize a preliminary process on data sets by decomposing the nonlinear wind speed time series into more stationary and regular subseries which are generally easier to analyze, thus improving the forecasting accuracy [26]. In the literatures, signal decomposition technique, such as wavelet transform (WT) [9,27], empirical mode decomposition (EMD) [28–30], ensemble empirical mode decomposition (EEMD) [16,31] and variational mode decomposition (VMD) [14,17,32], are widely used to extract the trends and harmonics and filter irrelevant and redundant features of the wind speed data set. However, as pointed out by Zhang et al. [17], the performance of WT depends highly on the selection of mother wavelets and optimal numbers of decomposition levels. EMD is widely used to recursively decompose complex, multi-component signals into several amplitude-modulated-frequency-modulated modes, known as intrinsic mode functions (IMFs). However, EMD has two disadvantages: (1) EMD fails to distinguish components whose frequencies lie within an octave and (2) EMD is vulnerable to intermittence such as noise. Wang et al. have shown EEMD having better performance than EMD [31]. The main drawback of EEMD is that it is difficult to select parameters such as amplitude of noise and ensemble number. VMD is a filtering based method which has been proposed recently to overcome the drawback of EMD [33] and has been exploited in wind speed preprocessing. Although VMD has been proved to be robust to noise, it requires an appropriate choice of mode number. Most recently, a modified version of EMD, referred as time varying filtering based EMD (TVF-EMD), has been presented by Li et al. to solve the mode mixing problem in EMD (referred to separation problem and intermittence problem) [34]. Unlike the original EMD, which is known for limitation like sensitivity to noise and sampling, TVF-EMD is able to achieve robust performance under low sampling rates and is suitable for non-stationary signals. Therefore, TVF-EMD is applied to the wind speed time series decomposition in this paper.

Most of the existing works in WSF field focus on prediction methodologies. However, the study of feature selection for WSF has not received enough attention, which is another contribution in this paper. Feature (variable) selection (FS) is a key component in building reliable prediction strategies when a large number of candidate inputs exist during forecast processes [35]. By removing ineffective candidate features and reducing the size of input set through feature selection, the forecasting engines can better learn the relevant input/output mapping function of the process, thus improving forecast accuracy and decreasing computation time. In the literature, feature selection technique can be divided into three categories: correlation and principle component analysis based [35,9,36], optimization algorithm based [17,37,38] and information theory based [39,40] feature selection. In terms of WSF, Liu et al. [9] analyzed the relationship between the historical speeds and environment variables on different lags using the autocorrelation function (ACF) and partial correlation function (PCF), along with Granger causality test. Kong et al. utilized the principle component analysis (PCA) to determine necessary factors (including wind speed, temperature, pressure and wind direction) for forecasting the wind speed. Li et al. [40] developed a feature selection technique based on the conditional mutual information to choose a compact set of input features for the forecasting model. In the [17], the binary hybrid backtracking search algorithm (BHBSA) is exploited for feature selection applying on the candidate inputs predefined by partial autocorrelation function (PACF) values. However, most of the existing feature selection methods consider only linear relations between time series and nonlinear methods always take a continuous feature subset as a factor but do not analyze the individual lags. Recently in [41], Feng et al. discussed a deep feature selection process for wind forecasting to track these two issues above. Furthermore, feature selection for the WSF process is a difficult task, demanding more effective methods to analyze the non-linear relation between wind speed time series.

In this paper, a new prediction strategy is proposed for short-term WSF. The main contribution of the paper can be summarized as follows:

(1) model for short-term wind speed forecasting for the first time. We reconstruct the empirical loss function of ORELM model [18] and provide a solution scheme to the loss Quantile regression is integrated into ORELM function. We further provide a model selection scheme to select optimal number of nodes and regularization parameter using the generalized approximate cross-validation (G-ACV) [42]. A hybrid population-based algorithm (PSOGSA), with the combination of particle swarm optimization (PSO) and the gravitational search algorithm (GSA) [43], is used to fine-tune the parameters of ORELM network structure.
(2) It is the first time to adopt TVF-EMD technique to decompose wind speed time series, which is a newly developed time varying filter based method to decompose the signal into series of local narrow-band components. To further verify the performance of TVF-EMD, comparison with EMD is made.
(3) Inspired by the strategy of feature selection in [17], this paper introduces an optimization algorithm based feature selection method to select a minimum subset of the most informative features for the WSF process. A binary version of PSOGSA (BPSOGSA) [44] is exploited to carry out this task.
(4) To demonstrate the effectiveness of the proposed hybrid WSF model, a number of comparable experimental studies are carried out. The support vector quantile regression (SVQR), linear quantile regression (LQR) and persistence model (PM) with and without signal decomposition technique (i.e. TVF-EMD and EMD), with and without using the hybrid PSOGSA (HPSOGSA) for parameter optimization and feature selection, are compared with the proposed hybrid WSF model.