Energy 61 (2013) 673–686

Contents lists available at ScienceDirect

Energy

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

A Gaussian mixture copula model based localized Gaussian process regression approach for long-term wind speed prediction

Jie Yu*, Kuilin Chen, Junichi Mori, Mudassir M. Rashid

Department of Chemical Engineering, McMaster University, Hamilton, Ontario L8S 4L7, Canada

ARTICLE INFO

Article history:
Received 28 June 2013
Received in revised form
6 September 2013
Accepted 8 September 2013
Available online 3 October 2013

Keywords:
Renewable wind power
Wind speed forecasting
Gaussian mixture copula model
Gaussian process regression
Random uncertainty
Non-stationary seasonality

ABSTRACT

Optimizing wind power generation and controlling the operation of wind turbines to efficiently harness the renewable wind energy is a challenging task due to the intermittency and unpredictable nature of wind speed, which has significant influence on wind power production. A new approach for long-term wind speed forecasting is developed in this study by integrating GMCM (Gaussian mixture copula model) and localized GPR (Gaussian process regression). The time series of wind speed is first classified into multiple non-Gaussian components through the Gaussian mixture copula model and then Bayesian inference strategy is employed to incorporate the various non-Gaussian components using the posterior probabilities. Further, the localized Gaussian process regression models corresponding to different non-Gaussian components are built to characterize the stochastic uncertainty and non-stationary seasonality of the wind speed data. The various localized GPR models are integrated through the posterior probabilities as the weightings so that a global predictive model is developed for the prediction of wind speed. The proposed GMCM–GPR approach is demonstrated using wind speed data from various wind farm locations and compared against the GMCM-based ARIMA (auto-regressive integrated moving average) and SVR (support vector regression) methods. In contrast to GMCM–ARIMA and GMCM–SVR methods, the proposed GMCM–GPR model is able to well characterize the multi-seasonality and uncertainty of wind speed series for accurate long-term prediction.

1. Introduction

Renewable energy has attracted considerable attention recently because of the deterioration of the environment and the depletion of conventional energy resources. As a non-polluting and renewable energy source for power generation, wind energy is increasingly appealing and has been well recognized as environmentally friendly, socially beneficial and economically competitive [1,2]. The efficiency of wind turbines can be significantly improved if its routine operation is controlled through intelligent and predictive decisions on load demand and production scheduling based on short-term wind speed forecasting [3,4]. Meanwhile, accurate estimation of the trend and persistency of long-term wind speed is critical towards optimizing the selection of wind farm cites and planning of wind power generation [5–7]. However, the intermittency, non-stationarity, stochastic uncertainty and unpredictable nature of wind speed pose great challenges as to the reliable and efficient production of wind power [8]. Therefore, the development of efficient modeling tools becomes necessary to overcome the above technical challenges and ensure accurate long-term and short-term forecasting of wind speed, which is vital to optimize the efficiency of wind power generation.

Improving wind power generation has significant technical, economic and environmental advantages. The electric power output of a wind turbine depends on wind speed, which exhibits variability on a wide range of time scales from seconds to hours, days and months. Local and regional weather patterns, seasonal variations, terrain and the proximity of other nearby wind turbines are a few of the factors that may have impact on the electrical output of wind power generator. The forecasting horizons concerning short-term wind speed prediction are usually in the order of days with time-scales ranging from minutes to hours. In contrast, long-term forecasting is referred to prediction horizon of a few weeks to months with sampling period typically in the order of days. Short-term wind speed prediction is focused on the electricity market pricing and regulation of load variations whereas long-term prediction is targeted at the operational and maintenance scheduling of wind turbines [9,10]. In literature reports, two groups of

* Corresponding author.
E-mail address: jieyu@mcmaster.ca (J. Yu).

0360-5442/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.energy.2013.09.013

methods have been developed for wind speed prediction including physical and statistical models. Physical model based approaches take geological and atmospheric conditions like terrain, obstacle, pressure and temperature into consideration for wind prediction while statistical models aim to find the relationships within the time series of wind speed measurements [11,12]. The former type of approaches are derived from meteorological model of wind dynamics in order to predict wind power. Typically physical models are advantageous in short-term prediction because the influence of atmospheric dynamics is becoming more important while statistical models provide good predictions in the short-term horizon. In practice, however, physical and statistical models can be utilized simultaneously so as to combine the merits of both kinds of methods.

NWP (numeric weather prediction) method is a type of physical model based approaches to wind forecasting where complex mechanistic models are built to utilize weather and geographical measurements such as temperature, atmospheric pressure, surface roughness and obstacles [13,14]. NWP models have certain limitations in handling medium or long-term wind speed forecasting because of the difficulties of estimating meteorological dynamics accurately. Moreover, NWP methods are more appropriate for the situations where the weather conditions are relatively stable but may not be efficient to deal with random uncertainty of weather data. Kalman filter has been explored for wind speed prediction in order to remove the systematic forecast errors [15]. Nevertheless, the use of analytical models for prediction may not well characterize the complex random variations in wind speed. As a kind of statistical approach, ARMA (auto-regressive moving average) model is popular for characterizing time series of wind speed data and has been widely used for wind power forecasting [16,17]. Various ARMA models used for wind speed prediction include ARIMA (auto-regressive integrated moving average), seasonal- and fractional-ARIMA, and ARMAX (ARMA with exogenous input) [18,19].

In addition to the time-series models, artificial intelligence techniques such as ANNs (artificial neural networks) and fuzzy logic are also extensively studied for wind speed prediction. Fuzzy logic is proposed for generating fuzzy rules from numerical input–output data and used for time-series prediction of wind speed. A clustering algorithm based adaptive neuro-fuzzy inference system is developed to estimate wind speed at the higher heights by utilizing the knowledge of wind speed at the lower heights [20]. However, the main drawback of fuzzy logic approach lies in the fact that a large number of historical wind speed measurements are needed for accurate modeling and reliable prediction. As a result, the large amount of data can cause high dimension of fuzzy rule base that increases computational complexity. ANN is a powerful technique in wind speed prediction and has been widely utilized in various subjects including transient detection, pattern recognition and time-series prediction [21–25]. In ANN models, significant effort is needed to determine the network structure, number of neurons, preprocessing and activation functions in order to obtain accurate wind speed forecasting [26,27]. Consequently, it may lead to complex network structure and substantial learning time [28,29]. Evolutionary computing algorithms such as particle swarm optimization are used to identify parameter settings for ANN based predictive models [30]. In addition, AIM (abductory induction mechanism) is employed to determine the optimal network size, element types, connectivity and coefficients of ANN in forecasting wind speed at a location depending on a few other dispersed sites [31]. Meanwhile, ANN and MC (Markov chain) models have been integrated for wind speed forecasting in short-term time scale, though the model has restrictions to long-term predictions because the uncertainty in weather conditions and patterns may result in deteriorated model performance [32]. Support vector regression models are also applied to wind speed forecasting where a kernel function based nonlinear predictive model is constructed in a high-dimensional feature space [4,33–35]. Despite the strong nonlinear modeling capacity, the SVR (support vector regression) approach may not well characterize the random variations and seasonality in wind-speed series [36]. In parallel, LSSVM (least squares support vector machine) is employed to characterize the prediction residuals of the seasonal ARIMA model so that the prediction errors of wind speed can be corrected and mitigated [37]. Meanwhile, wavelet analysis has been adopted in wind speed forecasting to decompose the original time series into different sub-sequences for aggregated forecasting models [38]. However, this strategy cannot exactly characterize the feature of multi-seasonality in long-term forecasting of wind speed.

More recently, neural networks with adaptive Bayesian learning and Gaussian process approximation are proposed for short-term wind power prediction [39]. Moreover, BMA (Bayesian model averaging) technique is applied to predict the probability density functions of wind speed for reliable forecasting through probabilistic ensembles. This kind of approach estimates the predictive probability density function as a weighted average of distributions centered on the individual bias-corrected forecasts. The weights reflect the relative contributions of the individual wind speed forecasting models based on their predictive power over certain training period [29,40]. However, the appropriate number of local models needs to be pre-identified in order to build multiple models and combine them into a global model. Although the BMA algorithm is incorporated with various types of neutral networks for short-term wind speed forecasting, the issues of neural network technique including local minima and poor generalization capability may affect the effectiveness of the integrated models.

To address the existing challenges in wind speed prediction, a GMCM (Gaussian mixture copula model) and Bayesian inference strategy based global GPR (Gaussian process regression) approach is developed, where various non-Gaussian components of the mixture model are identified to characterize the non-stationary multi-seasonality of wind speed for long-term forecasting. The estimated means, covariances and prior probabilities across the different Gaussian copula components can accurately characterize the multi-seasonality of wind speed data [41]. Further, a Bayesian inference strategy is used to estimate the posterior probabilities of different Gaussian copula components so that the localized Gaussian process regression models corresponding to the multiple Gaussian copula components are integrated through the posterior probabilities into a global model for wind speed prediction [42,43]. The new aggregated GPR model within Bayesian framework is able to account for the random uncertainty and non-stationarity in the time series of wind speed and thus the predictability of the proposed approach can be substantially enhanced.

The remainder of the article is organized as follows. Section 2 briefly reviews the ARIMA and SVR based wind speed forecasting approaches. Then the Gaussian mixture copula model based Gaussian process regression framework for long-term wind speed prediction is developed in Section 3. Section 4 applies the GMCM based global GPR approach to wind speed prediction of three different wind farm locations and compares the results of GMCM–GPR method with those of the GMCM–ARIMA and GMCM–SVR based methods. The conclusions of this paper are summarized in Section 5.

2. Preliminaries

2.1. Auto-regressive integrated moving average model

Various types of auto-regressive models including ARMA and ARIMA models are well suited to capture the dynamic correlations