Applied Energy xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Applied Energy

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

Deterministic and probabilistic wind power forecasting using a variational Bayesian-based adaptive robust multi-kernel regression model

Yun Wang[a], Qinghua Hu[a,*], Deyu Meng[b], Pengfei Zhu[a]

[a] School of Computer Science and Technology, Tianjin University, Tianjin, China
[b] School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

HIGHLIGHTS

- Multi-resolution data are used to forecast wind power by multi-kernel regression.
- Mixture of Gaussians is applied to model the complex wind power forecasting errors.
- The number of Gaussian components in a mixture of Gaussians is determined by stick breaking construction.
- The derived model can output the deterministic and probabilistic forecasts simultaneously.

ARTICLE INFO

Keywords:
Wind power
Probabilistic forecast
Deterministic forecast
Outlier
Robust multi-kernel learning
Variation Bayesian

ABSTRACT

Accurate wind power forecasting has great practical significance for the safe and economical operation of power systems. In reality, wind power data are recorded at high time resolution (5 s, etc.). The original high-resolution data are averaged to produce the low-resolution time series (10 min, etc.) used in wind power forecasts. Therefore, the current wind power forecasting models neglect certain information in the high-resolution data. Moreover, the common Gaussian assumption used for the error term in the current wind power forecasting model is not consistent with the real, complex wind power forecasting error distribution. In this paper, an adaptive robust multi-kernel regression model is proposed to deal with the two disadvantages mentioned above. First, a multi-kernel regression model is constructed to process the multi-resolution wind power data. Second, a Gaussian mixture model is employed to model the complex wind power forecasting error. Finally, a variational Bayesian method is introduced to optimize the proposed model and to cause the simultaneous output of both the deterministic and probabilistic forecasts. Two case studies have been conducted on real wind power data from Chinese wind farms. The results show that the proposed model provides more accurate deterministic forecasts and more useful probabilistic forecasts, and has great potential for practical application in power systems.

1. Introduction

The demand for energy increases rapidly with economic development. Considering the limited reserves of conventional resources (e.g., coal, oil, and gas) and the environmental pollution caused by their consumption, a growing number of countries have begun to explore renewable alternatives. Among the various renewable resources (e.g., solar and geothermal energy), wind energy is receiving extensive attention in the world due to its clean, inexhaustible, inexpensive, and widely distributed nature [1].

However, as wind power penetration increases, the high volatility and intermittency of wind power will challenge the stability of power system operations [2]. Wind volatility is primarily manifested as large-scale wind power ramping events (WPREs), while strong intermittency is manifested as difficulty in accurate wind power forecasting. Therefore, accurate wind power and wind power ramping event forecasting are essential, as they can enable adjustments to wind power advance scheduling, improve power quality, and reduce both the operational costs and the reserve capacity of the power system, etc. [3,4].

WPREs are sudden changes in wind power within a short period of time [4]. Currently, WPRE forecasting approaches are divided into two categories: direct models and indirect models [5]. In [5], a hybrid model based on orthogonal tests and a SVM (support vector machine) employed historical wind power ramp series and meteorological information to make direct WPRE forecasts. The authors of [4] used a reservoir computing-based model to realize binary (ramp/non-ramp)

* Corresponding author.
E-mail address: huqinghua@tju.edu.cn (Q. Hu).

http://dx.doi.org/10.1016/j.apenergy.2017.09.043
Received 22 May 2017; Received in revised form 24 August 2017; Accepted 10 September 2017

Abbreviations

ARMA	autoregressive and moving average model
ARIMA	autoregressive integrated moving average model
ANNs	artificial neural networks
ARMKR	adaptive robust multi-kernel regression
ARSKR	adaptive robust single-kernel regression
ACE	average coverage error
CWC	coverage width-based criterion
GP	Gaussian process
GMM	Gaussian mixture model
LUBE	lower upper bound estimation
LSSVM	least squares support vector machine
MMD	maximum mean discrepancy
MMAPE	modified version of the mean absolute percentage error
NMAE	normalized mean absolute error
NRMSE	normalized root mean square error
PIs	prediction intervals
PDF	probability density function
PACF	partial autocorrelation function
PICP	PI coverage probability
PINC	PI nominal confidence level
PINAW	PI normalized average width
QR	quantile regression
RVM	relevance vector machine
SVM	support vector machine
WPREs	wind power ramping events
WS	Winkler score

prediction. Cui et al. proposed an indirect WPRE forecasting model that used a neural network to forecast wind power for different scenarios, and then extracted information about WPREs from the wind power forecasting results mentioned above [6]. Similarly, the Wind Forecast Improvement Project, which was formed to enhance short-term wind power forecasting accuracy, extracted WPREs from actual and forecasted wind power time series using an optimized swinging door algorithm [7]. Accurate wind power forecasting is important for indirect WPRE forecasting models.

A number of models have been proposed recently to enhance wind power forecasting accuracy. These models can be categorized in terms of modeling theory into four types: physical models, statistical models, learning models, and hybrid models [8]. Physical models use physical quantities to calculate final forecasts of wind speed and wind power. The physical quantities used in these models include location information and meteorological factors provided by numerical weather prediction, such as wind speed, wind direction, and barometric pressure [9,10]. For statistical models, only historical wind power information is used. ARMA (autoregressive and moving average model) and ARIMA (autoregressive integrated moving average model), as well as fractional-ARIMA [11] and Hammerstein autoregressive model [12], are commonly applied to generate wind speed/wind power forecasts. With the rapid development of machine learning technologies, various learning methods such as ANNs (artificial neural networks), SVM, GP (Gaussian process), and fuzzy logic methods are widely employed to forecast wind power [13]. Recently, deep neural networks have also been used in wind energy forecasting [14,15]. However, due to the limited performance of the single models listed above, hybrid models, which combine different techniques, have become increasingly popular for forecasting wind power. Model combination is conducted in the preprocessing or post-processing stages; the forecasting performance of both types of hybrid methods is generally enhanced [13]. For instance, in the preprocessing stage, input data is sometimes handled by signal processing methods such as wavelet transform and empirical mode decomposition, while, in the post-processing phase, the results of several predictors are taken into consideration [13]. Unfortunately, the existing deterministic forecasting models do not always generate accurate forecasts due to the following issues.

First, information will be lost when transforming the collected wind power data into the required timescale. The collected wind power data are recorded at a high time resolution, such as 5-s [16]. 10-min, 30-min, 1-h, or even 1-day forecasts are often required for power system operation (including power system management and maintenance scheduling, etc.) [17]. The representative point in the required timescale is calculated by averaging, after which a 10-min or 30-min resolution time series can be obtained [16,18]. Therefore, fluctuations in the high-resolution data will be lost during data averaging to the required timescale.

Second, there is the inconsistency between the real wind power forecasting error and the error distribution assumption used in forecasting models. Recent studies have reported that the wind power forecasting error obeys not a Gaussian distribution, but a Beta distribution [19], a mixed distribution based on Laplace and normal distributions [20], a Levy alpha-stable distribution [21], etc. The findings above are inconsistent with the underlying Gaussian error distribution assumed in many current forecasting models, such as the LSSVM (least squares support vector machine) and the GP. Therefore, a proper and general error distribution assumption should be used in wind power forecasting models.

With the integration of large-scale wind power into the power grid, the estimation of power generation uncertainty becomes increasingly important and plays significant roles in risk assessment and risk decision-making in power system operation [22]. Further, large-scale wind power integration requires distribution information and PIs (prediction intervals) for the forecasted wind power [23]. Providing various aspects of power forecasts is equally important when integrating large-scale wind energy into the power grid [24,25]. Current uncertainty forecasting technologies require further research when compared with the relatively mature deterministic forecasting models [24].

QR (quantile regression) is the most commonly used model for wind power uncertainty analysis [26–28]. The advantages of QR include no additional assumptions about the distribution shape and the flexible inclusion of predictive information [23]. However, QR also has disadvantages such as discontinuity in the resulting PDF (probability density function) for each forecast, which may be time-consuming for the requirement of overall quantile distribution [29].

Recently, LUBE (lower upper bound estimation) ANN models have been used successfully for PI construction [30,31]; these models feature theoretical approximation for any continuous nonlinear functions and possess fine generalization ability. LUBE models make no assumptions about data distribution and avoid the calculation of matrices such as Jacobin and Hessian matrices [31]. Moreover, LUBE models are constructed using machine-learning methods [32]. However, the accuracy of the PIs constructed by LUBE models depends largely on the objective function; also, the PDF resulting from each forecast is discontinuous.

Bayesian models offer another way to obtain PIs [25,29,33]. When compared with the interval forecasting models discussed above, Bayesian models have the distinct advantage of producing continuous PDFs, which can produce interval forecasts at any confidence level. The primary challenge in Bayesian models lies in the selection of optimal prior variable distributions. However, the error term is assumed have a Gaussian distribution in the current Bayesian models, which is inconsistent with the complex wind power forecasting error observed in reality.

In this paper, a novel wind power forecasting model using ARMKR (adaptive robust multi-kernel regression) is proposed to overcome the disadvantages in the current deterministic and probabilistic forecasting models. First, high-resolution wind power data recorded every 5-s