Renewable Energy 132 (2019) 43–60

Contents lists available at ScienceDirect

Renewable Energy

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

Robust functional regression for wind speed forecasting based on Sparse Bayesian learning

Yun Wang [a], Haibo Wang [b,*], Dipti Srinivasan [c], Qinghua Hu [a]

[a] School of Computer Science and Technology, Tianjin University, Tianjin, China
[b] School of Economics and Management, Hubei University of Technology, Wuhan, Hubei, China
[c] Department of Electrical and Computer Engineering, National University of Singapore, Singapore

ARTICLE INFO

Article history:
Received 13 June 2017
Received in revised form
13 June 2018
Accepted 18 July 2018
Available online 30 July 2018

Keywords:
Robustness
Sparsity
Functional regression
Variational Bayesian
Wind speed forecasting

ABSTRACT

Accurate wind speed forecasting is helpful for reducing the instantaneous fluctuation of voltage, and also has great practical significance on power dispatching and plan. There are two problems in wind speed forecasting: loss of details when transforming the high-resolution data into the low-resolution data and outliers existing in our data. So, a sparse Bayesian-based robust functional regression model is proposed in this paper. First, both the low-resolution and high-resolution data are considered as inputs to forecast future wind speed. Specifically, besides the historical 10-min mean wind speed, the corresponding functional variables, constructed by wind speed data recorded every 5 s in each 10-min interval, are also taken as inputs to make 10-min-ahead wind speed forecasting. But, not all functional variables contribute to the accurate wind speed forecasts. So, a multi-Laplace prior is given to the corresponding coefficient functions to get sparse solutions, which can reduce the adverse effects of the redundant functional variables on the final forecasting results. Second, a multi-mixture of Gaussians prior is assumed for the forecasting error to enhance the robustness of the forecasting model. Results of spatial-temporal and multi-step-ahead wind speed forecasting show that the proposed model provides more accurate forecasts than the compared models.

1. Introduction

Currently, researchers around the world spare no efforts to develop renewable energy due to the increasing depletion of fossil fuel and the serious environmental pollution problems. Wind energy has become one of the fastest growing renewable energy today due to its clean, inexhaustible, inexpensive and widely distributed characteristics. However, the main factor, which restricts the large-scale integration of wind power, is the intermittence and randomness of wind speed. Large wind speed perturbation will make the system voltage and system frequency change greatly [1].

Recently, researchers have developed various forecasting models to get accurate wind speed and wind power forecasts. Based on the modelling theories, they can be divided into five categories: physical models, conventional statistical models, artificial intelligence-based models, spatial correlation models and hybrid models [2,3].

Physical models suffer from large-scale computation and difficult physical data acquisition [2]. Statistical models (e.g. factional autoregressive integrated moving average (f-ARIMA) model [4], Hammerstein autoregressive model [5]) are widely employed to capture the linear relationship in the wind speed data, while the nonlinear relationship can be captured by the artificial intelligence-based models (e.g. artificial neural network (ANN) and support vector machine (SVM)) [6]. Sometimes, hybrid models are designed by combining the statistical and artificial intelligence-based forecasting models [7]. Shi et al. [7] proposed ARIMA-ANN and ARIMA-SVM, which outperformed the single forecasting models. The similar results were also obtained in Refs. [8] and [9].

As to spatial correlation models, they take the spatial relationship of wind speed in different sites into account. Velázquez et al. [10] showed that forecasting error decreases when the wind speed and wind direction recorded at neighbouring sites were taken as inputs for ANN model. Owing to the similar fluctuation of wind speed between the observed site and its neighbouring sites, there is no surprise to see that the spatial correlation models outperform the forecasting models which just use wind speed data at the target site. Similar conclusions can also be drawn from Refs. [11] and [12].

The purpose of spatial correlation models is to seek more

* Corresponding author.
E-mail addresses: wangyun15@tju.edu.cn, e0266173@u.nus.edu (H. Wang).

https://doi.org/10.1016/j.renene.2018.07.083
0960-1481/

explanatory variables from the neighbouring sites. If there is no additional information in the neighbouring sites, other researchers try their best to find the other explanatory variables from the target sites. Gallego et al. [13] showed that taking wind direction with wind speed can reduce forecasting error, while De Giorgi et al. [14] demonstrated that temperature and pressure were also helpful to improve the forecasting accuracy.

However, the rich information contained in the data has not been fully utilized. Certain detailed information will be lost when transforming the observed high-resolution data into the required low-resolution data. The original collected wind speed, recorded every 1 s, 2 s or 5 s, is the high-resolution data [13]. In reality, 10-min mean wind speed time series (low-resolution data) are usually employed to make wind speed forecasting [11,13]. Using the mean to represent all wind speed recorded every 1 s, 2 s or 5 s in 10-min interval cannot depict the wind speed fluctuation in 10 min comprehensively.

In addition, the collected wind data often contain some outliers, which will prevent us from obtaining accurate wind speed forecasts. Zheng et al. [15] divided the collected wind data into six categories, in which irrational data and unnatural data are considered as outliers. These outliers are mainly caused by the data entry error and the defective data collection instruments [16].

In literature, as shown in Fig. 1, two solutions are employed to reduce the adverse effect of outliers on our models. The first solution is to pre-process the collected data before training our models. Recent studies show that a pre-processing stage for outliers helps enhance the forecasting accuracy. The pre-processing methods can be divided into two categories: outlier detection methods [17,18] and signal processing methods [19,20]. In Refs. [17] and [18], the reweighed least squares-based and the support vector regression-based outlier detection methods were employed to pre-process the raw wind speed data, respectively. Wang et al. [18] showed nearly 34% improvement was obtained regarding the mean squared error by pre-processing. The wavelet transform was employed to detect and remove outliers, then the performance of SVM was enhanced when compared with SVM without data pre-processing [19]. A similar conclusion can also be drawn from Refs. [20–22].

However, the main drawback of the forecasting methods with a pre-processing stage is that the final forecasting accuracy depends largely on the quality of the processed data [6,20]. For outlier detection methods, we don't know whether all outliers are detected completely. As to signal processing methods, no matter normal wind speed points or outliers are processed with no difference. We cannot guarantee that all outliers are removed during processing. From the above discussion, the processed data handled by some pre-processing methods may still contain some outliers. So, the forecasting models also need the ability to handle outliers.

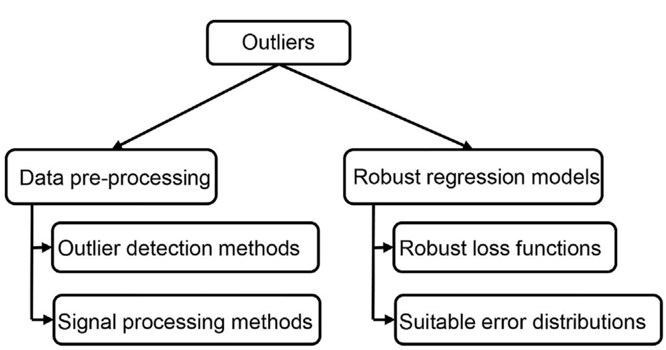

Fig. 1. Methods to deal with outliers.

The second solution for outliers is to develop robust regression models, which can not only be used to deal with outliers directly, but also be employed to analysis the data processed by some pre-processing methods. From Fig. 1, there are two ways for constructing robust regression models. The first one is to construct robust regression functions with the principle that the values of loss functions don't change much as the error increases [23]. The second one is to give the regression error a suitable error assumption. In fact, commonly used wind speed forecasting models (e.g. LSSVM, SVM) lack robustness due to the short tail error distribution assumption [24–26]. From Bayesian inference, the loss function is equal to the negative logarithm of the error distribution [26,27]. According to the fact that wind speed forecasting error obeys a Bate distribution, the corresponding loss function is obtained from the Bayesian perspective. Then, they were employed to replace the original loss functions in LSSVM and SVM. The proposed BN-KRR and BN-SVM show better wind speed forecasts [26,27]. However, the challenge to develop a robust regression model for wind speed forecasting is that the wind speed forecasting error distribution is unknown and complex, and single distributions cannot fit it properly.

Summarily, there are two main shortcomings for the current wind speed forecasting methods: (1) the high-resolution wind data, which contains more information that reflect wind speed fluctuation, are often neglected; (2) the robustness of the current wind speed forecasting models is not sufficient. In this paper, we develop a sparse Bayesian-based robust functional regression model to overcome the above two shortcomings. Functional variables, constructed by high-resolution data recorded every 5 s in 10-min interval, are taken to represent the wind speed fluctuation during 10 min to avoid the loss of information during data transformation. So, besides considering the 10-min mean wind speed, historical functional variables are also taken as inputs to forecast future wind speed.

When there are tens of wind turbines in a wind farm, hundreds of functional variables from different wind turbines can be used to forecast the wind speed of the selected wind turbine. Obviously, most of these variables are redundant or irrelevant, which may lead to model over-fitting. Sparse Bayesian learning can produce sparse coefficients on all variables by specific prior, those sparse coefficients can reduce the impact of redundant variables on final forecasting results [28]. In this paper, a multi-Laplace prior [29,30] is assumed for the functional coefficients to obtain a sparse representation among functional variables.

Mixture of Gaussians (MoG), which have superior fitting ability to any continuous distributions [31], is assumed to model the unknown and complex error distribution [32] to construct a general robust forecasting model. When encountering with the forecasting tasks of multiple targets under the same inputs (e.g. multi-step-ahead wind speed forecasting), we use multiple mixture of Gaussians (MMoG) prior to the error term of multiple forecasting tasks, which indicates that the forecasting error of each task is modeled by MoG. The main contributions of the paper can be summarized as.

- The proposed model uses the multi-resolution information to forecast future wind speed. Except for the historical 10-min mean wind speed data, the corresponding functional variables constructed by high-resolution wind speed data (sampled every 5 s) in each 10 min interval are also taken as inputs to forecast future 10-min mean wind speed.
- The sparsity of functional coefficients is realized by multi-Laplace prior within the variational Bayesian framework. Sparse functional coefficients will reduce the adverse effects of