Energy Conversion and Management 164 (2018) 374–384

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Probability density forecasting of wind power using quantile regression neural network and kernel density estimation

Yaoyao He*, Haiyan Li

School of Management, Hefei University of Technology, Hefei 230009, China
Key Laboratory of Process Optimization and Intelligent Decision-Making, Ministry of Education, Hefei 230009, China

ARTICLE INFO

Keywords:
Short-term wind power
Probability density forecasting
Quantile regression neural network (QRNN)
Epanechnikov kernel function
Unbiased cross-validation
Evaluation metrics

ABSTRACT

Owing to the increasingly serious social energy crisis nowadays, wind power and other renewable energy are paid more attention. However, penetration of wind power prominently enhances the degree of complexity and difficulty in planning and dispatching of electric power systems. High-precision and more-information short term wind power forecasting (STWPF) results can effectively alleviate the uncertainly of wind power and balance the electrical power. Kernel function and bandwidth selection method have significant impact on the results of STWPF. A hybrid wind power probability density prediction method based on quantile regression neural network and Epanechnikov kernel function using Unbiased cross-validation (QRNNE-UCV) is presented. The wind power predicting results at different conditional quantiles are used as the input of kernel density estimation (KDE), which is capable of estimating the comprehensive wind power probability density forecasting information at any time in the future. In order to evaluate the wind power prediction results, the paper constructs two evaluation criteria, including evaluation metrics of point prediction results and evaluation metrics of prediction interval (PI). As a point prediction result, the probability mean is first constructed in the paper. Two real datasets of wind power from Ontario, Canada, are used to verify the QRNNE-UCV method. Moreover, by comparing with the probability density results at various confidence levels, the influence of confidence level on STWPF is investigated in this article. Experiment results show that the QRNNE-UCV method can construct more accurate PI and probability density curves, and the calculated probability mean is superior to the other point predictions. Meanwhile, the quality of PICP and PINAW improves with the increase of confidence level. The above prediction results have the ability to validly quantify the indeterminacy of wind power generation in contrast to existing support vector quantile regression (SVQR) and quantile regression neural network and triangle kernel function (QRNNT) probability density forecasting methods.

1. Introduction

As substitute for the traditional fossil fuels, the clean and renewable energy has been widely used in modern industry for the recent decades. Low-carbon energy with large-scale form is integrated to electrical power grid, which is a critical step of constructing a strong, efficient and interactive smart grid [1,2].However, the uncertainty in intelligent dispatch of smart grid has been prominently increasing because of the highly non-schedulable and randomness of the renewable energy [3,4]. In order to dispatch and utilize the renewable energy in a massive scale smart grid, Rifkin put forward the concept of energy internet based on new renewable energy and information technologies. Energy internet is a strong smart grid with global and interconnection in essence [5]. The decision departments of power systems for the comprehensiveness and precision of the forecasting results is escalating in the pattern of energy internet. In different renewable energy technologies, wind power has been widely considered in the various countries due to its clean, recyclability, exploitability and technology mature [6]. Comprehensive and precise short-term wind power forecasting (STWPF) is a necessary tool to reduce the uncertainty of wind power, which can optimize the unit commitment and load dispatch, and improve the stability and economic efficiency of power systems [1,7]. Large space-time wind generation forecasting is needed for smart grid economic operation from the global perspective. And it is essential to guarantee the stability of power systems and improve the competitiveness of wind power in the electricity market [8].

To construct accurate prediction model of wind power generation, many scholars have carried on a series of investigates. In recent years, point forecasts and interval predictions are the two most common types of wind power prediction methods [9,10]. The researchers acquired an

* Corresponding author.
E-mail addresses: hy-342501y@163.com (Y. He), 2465460550@qq.com (H. Li).

https://doi.org/10.1016/j.enconman.2018.03.010
Received 21 January 2018; Received in revised form 2 March 2018; Accepted 4 March 2018

estimate of wind farm power output by point prediction, which is hard to measure the nondeterminacy of wind power [9]. Different from the point prediction results, prediction intervals (PIs) are composed of the maximum and minimum of prediction results under certain confidence level [11]. Interval prediction can offer accurate waving interval of wind power for electricity utilities and decision makers. Different from the above-mentioned two kinds of wind power forecasting methods, probability density forecasting is able to construct complete probability density curves of prediction results to qualify the uncertainty of wind power [12,13]. And it can provide fully and high-accuracy predictive information for deciders in the power system.

According to different realization approaches in previous studies, wind power prediction methods mainly include statistical methods, physical and artificial methods [14,15]. Compared with physical methods, statistical and artificial methods do not use relevant numerical weather prediction (NWP) information and can obtain forecasting results with higher prediction accuracy [3]. Furthermore, some hybrid methods have been proposed to improve the quality of renewable energy prediction results, combining the advantages of multiple predicting methods [16–21]. Therefore, the predicting of wind power is discussed based on statistical and artificial approach in this article.

On the basis of the analysis of various types of generalized autoregressive conditional heteroscedasticity (GARCH) wind power forecasting models, Ref. [17] put forward a Markov regime switching (MRS-GARCH) method to measure the uncertainty of wind power. In [18], an evolutionary-adaptive method, combining the merits of wavelet transform, adaptive neuro-fuzzy inference system, and evolutionary particle swarm optimization model, was presented to acquire high-precision prediction results by a relatively simple calculation method. Anneela et al. [19] certificated the validity of the ensemble predicting method based on regressor and artificial neural networks (ANN) by the tested data of Europe wind power. As an artificial intelligence method, ANN has been extensively applied in the field of wind power forecasting because of its obvious advantages of strong generalization, self-adaptability and non-linear mapping [22–24]. By specifying the mean and variance of a Gaussian distribution about tested dataset, ANN is able to obtain the probabilistic prediction results [25]. MDMW-based (multi-dimensional Morlet wavelets) wavelet neural network (WNN) was utilized in [23] for predicting wind power generation, and Maximum Correntopy Criterion was used to train the constructed model. In [24], a new wind power forecasting method based on the least square support vector machine (LSSVM) and radial basis function neural network (RBFNN) was adopted, and the associated arguments were obtained by variant wind velocity scope. However, the drawback of these methods mentioned above is that it only provides the point prediction and PIs, which is hard to completely handle the volatility and uncertainty of the wind farm output power.

Quantile regression (QR) is commonly used for regression analysis in the statistics and econometrics fields. QR has the ability to estimate the conditional distribution of the explained variables [26]. The integrated distribution characteristic of response variables can be obtained by a prescribed suitable regression function, without considering the distribution type of random variables [27]. QR is more robust to deal with the outliers in the explained measurements. Therefore, QR has been widely utilized in wind power prediction. However, simple linear QR is hardly to handle the complex non-linear problems [28]. It is necessary to explore suitable nonlinear function for QR. In order to thoroughly reveal the nonlinear relationships of the dependent and independent variables, Taylor [29] proposed a quantile regression neural network (QRNN) model, which incorporated the advantage of ANN and QR model. Ref. [30] proposed a robust neural network method to depict the condition density at any future time based on QR and ANN model. However, it is unable to directly acquire the probability density function of future wind power only with a single QRNN model. The current studies of STWPF are restricted to gain the PIs information by the forecasting model [6].

Different from parametric estimation methods, KDE is a popular approach for estimating the data distribution without prior assumption of datasets [31]. Meanwhile, bandwidth also has a significant impact on density function of random variables [32]. In this article, in order to acquire the complete wind power probability density curve, Epanechnikov kernel function and Unbiased cross-validation (UCV) bandwidth selection method are combined with the QRNN model, which is named QRNNE-UCV. The proposed QRNNE-UCV method is used to STWPF, which helps to acquire the overall conditional probability density at any moment in the future by choosing the appropriate kernel function and bandwidth.

Finally, two real wind power time series of summer and winter from Ontario, Canada are used to verify the proposed QRNNE-UCV method. For studying the effects of confidence interval in the QRNNE-UCV model, the paper evaluates the wind power prediction results obtained with 90% and 80% confidence level. Meanwhile, the proposed QRNNE-UCV method is compared with the support vector quantile regression method (SVQR) in [4] and quantile regression neural network and triangle kernel function (QRNNT) in [8] in order to demonstrate its advantages and forecasting performance.

The differences between this article and the previous researches can be briefly summarized as follows. (1) The previous four papers [4,8,12,21] focus mainly on power load probability density forecasting. However, this paper is the research on wind power probability density forecasting. It is easy to find that wind power forecasting is different from power load forecasting because the wind energy time series are more volatile. Hence, it varies from time to time without obvious daily and weekly pattern. (2) Previous work only adopted the mode and median of probability density curve as the metrics of point forecasting. In this paper, we firstly construct the model to get the probability mean of probability density curve for wind power forecasting. It has more accurate forecasting results for wind power. (3) Compared with the SVQR, kernel-based support vector quantile regression (KSVQR) methods proposed in [12,4] and the RBF neural network quantile regression (RBFQR) proposed in [21], this paper uses the quantile regression neural network model. The above methods employed the method of quantile regression, but the essential method used in this paper for forecasting is different with Ref. [4,12,21]. In addition, the foresaid two methods are difficult to obtain the appropriate parameters and the best prediction results. (4) Kernel density estimation is the key part of probability density forecasting. Simple quantile regression cannot implement probability density forecasting. Different from the triangle kernel function and the direct plug-in bandwidth selection method in [8], the kernel density estimation and bandwidth choice method adopted in this paper are the Epanechnikov kernel function and Unbiased cross-validation. Epanechnikov kernel function is optimal in a mean square error sense [33]. Compared with the direct plug-in bandwidth selection method, UCV can well describe the unknown distribution data and construct narrower PIs. (5) In order to clearly illustrate the role of confidence level in probability density forecasting, the wind power probability density forecasting is implemented separately with different confidence levels. However, previous work do not consider the influence of confidence level.

There are four mainly contributions in this article as follows: (1) A QRNNE-UCV method is constructed in this paper, combining Epanechnikov kernel function and UCV method with QRNN model. (2) Different from existing probability density forecasting method, the probability mean of the probability density curve is regarded as point forecasting results for the first time. By comparative analysis with the traditional forecasting methods, it demonstrates higher accuracy. (3) In order to clearly illustrate the role of confidence level in probability density forecasting, the wind power forecasting are implemented separately with different confidence levels. (4) Seasonal factors of wind power are considered in this paper, the wind power data in different seasons are tested and compared.

The rest organization of this paper is as follows. Section 2 introduces