Energy 114 (2016) 498–512

Contents lists available at ScienceDirect

Energy

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

# Short-term power load probability density forecasting based on quantile regression neural network and triangle kernel function

Yaoyao He [a, b, c, *], Qifa Xu [a, b], Jinhong Wan [c], Shanlin Yang [a, b]

[a] School of Management, Hefei University of Technology, Hefei, 230009, China
[b] Key Laboratory of Process Optimization and Intelligent Decision-Making, Ministry of Education, Hefei, 230009, China
[c] State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing, 100048, China

ARTICLE INFO

*Article history:*
Received 17 March 2016
Received in revised form
19 July 2016
Accepted 7 August 2016

*Keywords:*
Load forecasting
Quantile regression neural network
Probability density forecasting
Triangle kernel function
Bandwidth selection method

ABSTRACT

Highly accurate short-term power load forecasting (STLF) is fundamental to the success of reducing the risk when making power system planning and operational decisions. For quantifying uncertainty associated with power load and obtaining more information of future load, a probability density forecasting method based on quantile regression neural network using triangle kernel function (QRNNT) is proposed. The nonlinear structure of neural network is applied to transform the quantile regression model for constructing probabilistic forecasting method. Moreover, the triangle kernel function and direct plug-in bandwidth selection method are employed to perform kernel density estimation. To verify the efficiency, the proposed method is used for Canada's and China's load forecasting. The complete probability density curves are obtained to indicate the QRNNT method is capable of forecasting high quality prediction interval (PIs) with higher coverage probability. Numerical results also confirm favorable performance of proposed method in comparison with the several existing forecasting methods.



## 1. Introduction

In recent years, extensive studies have been devoted in power load forecasting as power generation costs increase and market competition exacerbates. The accurate electric load forecasting plays an important role in power system planning and management. The purpose of the short-term load forecasting (STLF) is to forecast in advance the system load for the day-to-day operation, scheduling and load-shedding plans of power utilities [1]. However, it is difficult to improve the prediction accuracy of STLF due to the nonlinear and random-like behaviors of system load, weather conditions, variations of social and economic environments, and so on [2]. In order to improve the prediction accuracy of load forecasting model, various STLF methods have been introduced in the past years, including artificial neural network (ANN) [3], grey Bernoulli model [4], wavelet transform combined with neuro-evolutionary algorithm [5], radial basis function (RBF) [6], particle swarm optimization [7], support vector regression (SVR) [8], combining sister forecasts [9], ensemble Kalman filter [10], combined model [11], and hybrid methods [12,13].

In terms of forecasting outputs, load forecasts can be categorized into point forecasts, interval prediction [4,14,15] and probability density forecasting [16]. Interval prediction method focuses mainly on constructing high quality prediction interval (PI) with higher coverage probability [14]. As the most complete prediction method, probability density forecasting have the capability to represent uncertainty as the probability distributions around a prediction interval. However, many assumptions and mathematical derivations are usually made in advance and hinder the application of probability density forecasting. The distribution type and its parameters need to be estimated for the error distributions [17]. Hence, most applications of STLF are based on point forecasts generated using different methods.

Due to excellent ability of non-linear mapping, generalization and self-learning, ANN has been proved to be of widespread utility in engineering optimization fields [3]. The ANN is capable of solving

* Corresponding author. School of Management, Hefei University of Technology, Hefei, 230009, China.
*E-mail addresses:* heyaoyaohfut@hfut.edu.cn, hy-342501y@163.com (Y. He), xuqifa1975@163.com (Q. Xu), wanjh@iwhr.com (J. Wan), hgdysl@gmail.com (S. Yang).

http://dx.doi.org/10.1016/j.energy.2016.08.023
0360-5442/

the probabilistic prediction problems by means of specifying the mean and variance of a Gaussian distribution when sufficient observed data are supplied [18]. The greatest advantage of ANN method is that it is very simple. However, the calculation results are likely to overflow or fluctuate between the optima because the simple ANN model is sensitive to selection of weight values, threshold value and topology. Different hidden nodes and transfer function tend to lead to distinct results when network is trained and tested on the same database. At the same time, ANN is easy to be trapped in some static points or oscillate in these points in the process of learning. In this case, the forecasting system exists large error as well, no matter how many times for the iteration [19]. In addition, traditional ANN have difficult in realizing the probability density forecasting.

Quantile regression model is able to estimate point values of individual quantiles directly. If a probability density function with cumulative distribution function has been defined, the shape of the predictive distribution can be determined by making estimates throughout the range of quantiles [20,21]. Thus, assumptions about the parametric form of the distribution function are not required. Furthermore, regression quantiles are equivariant to monotonic transformations, which means that transformed quantiles of a predictand is able to be estimated from quantiles of a transformed predictand [22]. This property leads to the censored quantile regression model that has the capacity to solve a great many complex prediction problems, including precipitation amounts, wind speeds, and power load [23].

To improve the prediction task of the conditional probability distribution of financial returns, a more flexible type of model, which is named quantile regression neural network (QRNN), was introduced by Taylor in 2000 [24]. A neural network was used to estimate the potential nonlinear quantile regression models for estimating the conditional density [22,25]. However, for probability density forecasting problem, a single QRNN tend to be insufficient to describe the regularity and probability of future power load satisfactorily. Previous works on probabilistic power load forecasting focused mainly on prediction interval construction by coverage probability of target values [14]. For a forecasting task, the ideal prediction interval is difficult in covering all target values.

To overcome this problem, we take advantage of the quantile regression theory and neural network technology, and develop the kernel density estimation and bandwidth selection methods. By adopting fused penalization, we penalize the weights and thresholds of neural network and their successive differences at neighboring quantile levels, so that the QRNN model of coefficients can be identified simultaneously. Triangle kernel function is employed to construct probability density function that have important impacts on the average and given quantile of the response distribution [26]. The proposed method differs from these existing works, as triangle kernel function is used to replace Gauss kernel density estimator and direct plug-in implements a reliable data-based bandwidth selection method using pilot estimation of derivatives [27]. Therefore, the proposed method not only implements the conditional probability density forecasting, but also improves the estimation efficiency by adjusting bandwidth to estimate the scope of prediction interval.

The main contributions of this paper are listed below: 1) A reliable and accurate probability density forecasting method is implemented. The complete conditional probability density curve of future load is drawn, and most actual values (target values) are covered in the forecasted probability density curve. 2) Triangle kernel function is firstly integrated into QRNN model for short-term power load probability density forecasting. The smoothness and continuity of probability density curve are further enhanced through integration of obtained quantile function. 3) The problem of estimating the optimal bandwidth is transformed to estimate three quadratic functions by the direct plug-in approach. 4) Two important interval prediction assessment indices are employed to evaluate the performance of presented probability density forecasting method. 5) The median and the highest probability point (mode) of probability density curve are introduced as important assessment indices of probability density forecasting. Compared with existing probability density forecasting method and traditional point forecasting algorithms, and hybrid Artificial intelligence methods, obtained results from three case studies indicate that the quality of prediction has been significantly improved.

The remainder of the paper is organized as follows. Section 2 provides the formulation of quantile regression. The quantile regression neural network using triangle kernel function (QRNNT) is described in section 3. Section 4 shows three practical numerical examples to show the performance of probability density forecasting method, and section 5 outlines the conclusions and future research.

## 2. Quantile regression

Quantile regression is an important expansion of traditional mean regression model. It is closely related to models for the conditional median, and makes regression model fitting and analysis of the relevant data more robust and accurate [20,21]. Suppose there is a random variable $Y$, which is characterized by its (right continuous) distribution function.

$$F(y) = P(Y \le y) \tag{1}$$

whereas for any quantile $0<\tau<1$,

$$Q_Y(\tau) = F^{-1}(\tau) = \inf(y : F(y) \ge \tau) \tag{2}$$

is called the $\tau$-th quantile of $Y$.

Consider a simple decision theoretic problem: a point estimate is required for a random variable with distribution function $F$. If loss is described by the piecewise linear function illustrated in equation (3).

$$\rho_\tau(\mu) = \mu(\tau - I(\mu)) \tag{3}$$

where$I(\mu)$ is indicative function, and satisfies

$$I(\mu) = \begin{cases} 0, & \mu \ge 0 \\ 1, & \mu < 0 \end{cases} \tag{4}$$

Asymmetric loss function $\rho_\tau(\mu)$ can be described as follows:

$$\rho_\tau(\mu) = \begin{cases} \tau\mu, & \mu \ge 0 \\ (\tau - 1)\mu, & \mu < 0 \end{cases} \tag{5}$$

A linear quantile regression is based on the assumption that there is linear relationship between dependent variables $Y$ and explanatory variables $X = [X_1, X_2, \ldots, X_K]'$. Simplified expression of linear quantile regression model is as follows.

$$Q_Y(\tau|X) = \mathbf{X}'\boldsymbol{\beta}(\tau) \tag{6}$$

where $Q_Y(\tau|\boldsymbol{X})$ is the conditional $\tau$ quantile of response variables $Y$ under explanatory variables $X$, $\beta(\tau) = [\beta_0(\tau),\beta_1(\tau),\beta_2(\tau), \ldots,\beta_k(\tau)]'$ is the regression coefficient vector.

According to equation (6), the estimates of parameter vector $\beta(\tau)$ can be transformed to solve the following optimization problem: