

# Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network

Yaoyao He <sup>a,b,\*</sup>, Yang Qin <sup>a,b</sup>, Shuo Wang <sup>c</sup>, Xu Wang <sup>d</sup>, Chao Wang <sup>d</sup>

<sup>a</sup> School of Management, Hefei University of Technology, Hefei 230009, China

<sup>b</sup> Key Laboratory of Process Optimization and Intelligent Decision-Making (Hefei University of Technology), Ministry of Education, Hefei 230009, China

<sup>c</sup> School of Computing and Digital Technology Birmingham City University Millennium Point, Curzon Street, Birmingham, B4 7XG, England, UK

<sup>d</sup> China Institute of Water Resources and Hydropower Research, Beijing 100048, China

---

## Abstract

The electricity consumption forecasting is a challenging task, because the predictive accuracy is easily affected by multiple external factors, such as society, economics, environment, as well as the renewable energy, including hydro power, wind power and solar power. Particularly, in the smart grid with large amount of data, how to extract valuable information of those external factors timely is the key to the success of electricity consumption forecasting. A method of probability density forecasting based on Least Absolute Shrinkage and Selection Operator-Quantile Regression Neural Network (LASSO-QRNN) is proposed in this paper. First, important features are extracted from external factors affecting the electricity consumption forecasting by LASSO regression. Then, the LASSO-QRNN model is constructed to predict annual electricity consumption. The results of electricity consumption forecasting under different quantiles in the next several years are evaluated. Besides, we introduce kernel density estimation into our LASSO-QRNN model, which can give a probability distribution instead of a single-valued prediction. The prediction accuracy is evaluated through the empirical analyses from the Guangdong province dataset in China and the California dataset in the United States. The simulation results demonstrate that the proposed method provides better performance for electricity consumption forecasting, in comparison with existing quantile regression neural network (QRNN), back-propagation of errors neural network (BP), radial basis function neural network (RBF) methods, quantile regression (QR) and nonlinear quantile regression (NLQR). LASSO-QRNN can not only better learn the high-dimensional data in electricity consumption forecasting, but also provide more precise results.

*Keywords:* LASSO Quantile Regression Neural Network, probability density forecasting, electricity consumption forecasting, uncertainty analysis, high dimensional data

---

## 1. Introduction

Electricity consumption around the world is rising rapidly, due to the ever-increasing population, value pursuit for better living standards and attention to large-scale industrialization, which leads to the positive economic growth rates [1]. Medium and long term electricity consumption forecasting is essential to energy investment planning [2]. For example, in response to the excessive consumption of electricity, huge quantities of fossil fuel were used to generate electricity at the end of the 20th century, which caused the depletion of fossil fuel resources. Since the beginning of the 21st century, renewable energy sources (such as hydro, solar, wind) have been increasingly utilized to generate electricity. Compared with traditional fossil fuels, the wind, solar and hydro power generations have the advantages of large development potential, non-pollution and recyclability. They established a better and cleaner energy structure. However, on account of the uncertainty and complementarity of hydro, wind and solar power, electricity consumption forecasting becomes utterly important, but challenging. An accurate forecasting can promote the effective use of renewable energy. Moreover, precise electricity consumption forecasting can guide government strategies for future energy usage and development.

Most commonly used methods for electricity consumption forecasting include regression models [3], time series models [4], the fuzzy theory [5], neural networks [6], Bayesian networks [7], hybrid method [8] and so on. Regression analysis and time series models are the most acclaimed modeling techniques in electricity consumption forecasting [9]. With the rapid development of modern artificial intelligence methods, artificial neural network and population evolutionary algorithm are introduced into electricity forecasting [10]. Graditi G et al. [11] demonstrated that statistical machine learning approaches provided more accurate power predictions. In terms of prediction output, existing methods of electricity consumption forecasting can be divided into deterministic point prediction and probabilistic forecasting based on uncertainty analysis [12]. An exact point prediction cannot reflect the fluctuation of electricity consumption. In reality, real electricity consumption and the load growth are influenced by various factors, including economic development, industrial structure, income level of residents, climate, geographical environment, national policy (electricity price) etc. All of these factors interact and influence each other. Furthermore, the integration of informatization and industrialization in power industry promotes the rapid growth of power data, resulting in a large number of data with extensive sources, diverse features, huge volume and fast growth. It makes electricity consumption forecasting increasingly complicated [13].

To deal with the uncertainty in electricity consumption, a new Kullback-Liebler (K-L) divergence based similarity measure strategy was designed by Shao et al. [14] to forecast the electricity demand in China considering the significant impact factors of the electricity consumption, showing better accuracy. Amina

---

\*Corresponding author. School of Management, Hefei University of Technology, Hefei 230009, China  
Email addresses: hy-342501y@163.com (Yaoyao He <sup>a,b,\*</sup>), 1206474485@qq.com (Yang Qin <sup>a,b</sup>), shuo.wang@bcu.ac.uk (Shuo Wang <sup>c</sup>), wangxu04@126.com (Xu Wang <sup>d</sup>), wangchao@iwhr.com (Chao Wang <sup>d</sup>)

1  
2  
3 et al. [15] implemented a novel fuzzy wavelet neural network model for electricity consumption forecasting  
4 in the power system of the Greek Island of Crete, which provided significantly better results. In [16], a  
5 small-sample adaptive hybrid model (AHM) based on trend extrapolation method was proposed to forecast  
6 electricity consumption in China from 1991 to 2014, which showed robustness to stochastic changes and  
7 obtained more precise forecasting results. Al-Ghandoor et al. [17] presented an empirical model based on  
8 multivariate linear regression of time series for the Jordanian industrial sector, which identified the main  
9 influential factors of electricity consumption. Van der Meer et.al reviewed the probabilistic forecasting of  
10 photovoltaic power production and electricity consumption [18] and proposed a probabilistic forecasting  
11 method using Gaussian processes [19]. Vincenzo Bianco et al. [20] developed a long-term forecasting model  
12 considering the influence of economic and demographic variables on the annual electricity consumption in  
13 Italy from 1970 to 2007. It utilized the linear regression to build different models to reduce the uncertainty  
14 of forecasting, and showed consistency with the official projections.  
15  
16

17  
18 As a kind of probabilistic forecasting method [21], probability density forecasting method not only can  
19 get the future electricity consumption variation interval under the corresponding confidence interval, but  
20 also depict the full probability distributions of the future electricity consumption. It can provide more  
21 insights into the distribution of electricity consumption, which can help power system decision-makers to  
22 make effective decisions for the power system to avoid the large error and economic losses. However, owing to  
23 the complexity and uncertainty of external influence factors, existing probability density forecasting methods  
24 are difficult to characterize the uncertainty of electricity consumption [22].  
25  
26

27  
28 In order to achieve accurate prediction of electricity consumption in large data environments, extracting  
29 the feature from the external factors has become a key problem. The dimension of explanatory variables will  
30 increase remarkably, if all influential factors are considered without selection. Identifying the most related  
31 factors for training the model can improve predictive performance and reduce computational cost [23].  
32 Traditional variable selection methods, such as forward stepwise, backward stepwise, ridge regression, may  
33 suffer from the co-linearity problem, if there are too many tedious independent variables in high-dimensional  
34 situations. So, it is likely to lose some important factors in the process of feature selection, greatly reducing  
35 the explanatory power of the model. To tackle this issue, Tibshirani [24] proposed the Least Absolute  
36 Shrinkage and Selection Operator (LASSO) algorithm, which was further extended by Hastie, Tibshirani  
37 and Friedman [25]. By constructing a penalty function to acquire a refined model, the algorithm can compress  
38 the coefficients of variables and make some regression coefficients set exactly to zero, so as to achieve the aim  
39 of variable selection [26]. The LASSO is computationally practicable for high-dimensional data, reducing  
40 amounts of computation greatly. For example, Florian et al. [27] proposed an accurate method of wind power  
41 forecasting combined with LASSO, greatly reducing the computation time under the condition of the vast  
42 parameter space. Huebner et al. [28] utilized LASSO regression for variable selection from different types  
43 of factors influencing the energy consumption in residential buildings. Zhang et al. [29] built a classifier to  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58

1  
2  
3 recognize the patients at high risk of disease progression using the LASSO method from the high-dimensional  
4 microarray data. In the paper, we first present a novel method, namely LASSO-quantile regression neural  
5 network (LASSO-QRNN), to combine the LASSO regression with the quantile regression neural network  
6 (QRNN).  
7  
8

9 QRNN model was put forward by Taylor [30], which is a nonparametric nonlinear method combing the  
10 neural network with the quantile regression. Dadabada et al. [31] used the particle swarm algorithm to opti-  
11 mize the QRNN for predicting the volatility of financial time series. Cannon [32] applied the QRNN method  
12 to the precipitation downscaling task, considering the characteristics of prediction with mixed discrete-  
13 continuous variables. Xu et al. [33] proposed a composite quantile regression neural network (CQRNN)  
14 model to solve underlying nonlinear problems among variables. He et al. [34] utilized the quantile regression  
15 neural network method to implement the probability density forecasting of wind power, which validly quan-  
16 tified the indeterminacy of wind power generation. These empirical studies show the superiority of QRNN  
17 method. Not only can it reveal the entire conditional distribution of response variables, but also deal with  
18 complex nonlinear problems. In this paper, we will make use of the superior performance of QRNN for  
19 electricity consumption forecasting. It can produce the detailed conditional quantiles of the consumption  
20 values, and more accurate predictions.  
21  
22

23 Motivated by the existing research, we propose a new learning method to predict electricity consumption,  
24 which combines LASSO and QRNN. The details will be given in the next section. LASSO can produce high-  
25 quality features and reduce data dimensions effectively. QRNN is used to build the model, which estimates  
26 the probability density function of the consumption value. To the best our knowledge, this is the first work  
27 that uses and combines LASSO and QRNN for electricity consumption prediction. We verify our method on  
28 two real-world datasets from Guangdong province in China and California in U.S. The probability density  
29 forecasting results from the LASSO-QRNN method are compared to the results without considering the  
30 external factors and the results without variable selection considering the external factors. Comparisons  
31 with the state of the art further exhibit that the LASSO-QRNN method is able to improve the forecasting  
32 accuracy of the electricity consumption significantly and minimize the uncertainty in electricity consumption.  
33  
34

35 The main contributions of this article include: 1) LASSO dimension reduction technique is applied  
36 to reduce the uncertainty of annual electricity consumption associated with social and economic factors.  
37 2) At different quantiles, LASSO-QRNN produces the probability distribution of electricity consumption  
38 in details. More smooth probability density curves can be obtained, which provides more insights into  
39 electricity consumption. 3) Two criteria are utilized to evaluate the performance of LASSO-QRNN, namely,  
40 mean absolute percentage errors (MAPE) and relative mean squared error (RMSE). 4) The superiority of  
41 the method is proven through the experiment on two real-world data sets, Guangdong province in China  
42 and California in U.S. We study and compare three different settings: without considering external factors,  
43 considering external factors without variable selection and considering external factors with LASSO-QRNN  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58

1  
2  
3 method. Our results indicate that the proposed LASSO-QRNN method provides informative prediction  
4 than QRNN. Meanwhile, it achieves better accuracy than radial basis function neural network (RBF) [35],  
5 back-propagation of errors neural network (BP) [36], quantile regression (QR) [37] and nonlinear quantile  
6 regression (NLQR) [38]. In conclusion, the main novelty of the proposed method consists in effectively  
7 reducing data dimensions and nondeterminacy in the forecasting process to improve prediction precision  
8 considering the external and internal factors.  
9

10  
11 The subsequent sections are organized as follows. Section 2 introduces the correlation theories including  
12 the mechanism of LASSO regression and the mathematical formulation of QRNN. Section 3 proposes the  
13 LASSO-QRNN method for electricity consumption. The simulation studies are conducted to evaluate the  
14 performance of LASSO-QRNN in Section 4. Finally, conclusions are given in Section 5.  
15  
16  
17  
18

## 19 20 21 **2. Correlation theory**

### 22 23 *2.1. LASSO regression*

24 With the advent of the big data era, how to extract useful information from data has been paid close  
25 attention by scholar. One of the most effective approaches is statistical modeling technology. In order to  
26 reduce model deviation at the beginning of modeling, researchers usually consider adding as many indepen-  
27 dent variables as possible. However, in the actual modeling process, the subset of independent variables are  
28 chosen in order to improve the prediction accuracy. It is called the variable selection. Variable selection is  
29 an all-important step in the process of statistical modeling [39].  
30  
31  
32  
33

34 The LASSO method is a contraction estimation method, which can effectively deal with complex high-  
35 dimensional data problems [40].  
36

37 A multivariate linear regression model that satisfies the classical hypothesis is shown below.  
38  
39

$$40 \quad Y = X\beta + \varepsilon \quad (1)$$

41 where  $\varepsilon_i \sim \text{i.i.d } N(0, \sigma^2)$  means that each residual term is independent and has the same distribution with  
42 normal distribution. The response variables satisfy the condition  $Y \sim N(X\beta, \sigma^2 I)$ , in which  $I$  is the  
43 identity matrix. By optimizing the objective function Eq(2), the ordinary least squares estimation result  
44 obtained is  $\hat{\beta}_{ols} = (X'X)^{-1}X'Y$ .  
45  
46  
47  
48

$$49 \quad \min_{\beta} \frac{1}{2} \|Y - X\beta\|_2^2 \quad (2)$$

50 Suppose the row and the column from the design matrix are respectively recorded as  $s$  and  $k$ . If  $s < k$ ,  
51 it indicates the sample size is less than the number of explanatory variables. At this time,  $\text{rank}(X'X) \leq$   
52  $\text{rank}(X) \leq s$  where  $(X'X)$  is a matrix with  $k$  row and  $k$  column. Hence, ordinary least squares estimation  
53 fails because  $(X'X)$  is non-existent.  
54  
55  
56  
57  
58

Tibshirani [24] added L1 penalty to the end of the Eq(2) as follows.

$$\min_{\beta} \frac{1}{2} \|Y - X\beta\|_2^2 + \frac{\lambda}{2} \|\beta\|_1 \quad (3)$$

where  $\lambda$  is penalty parameter that is positive correlated with punishment intensity,  $\|\beta\|_1$  is 1-norm satisfying  $\|\beta\|_1 = |\beta_1| + |\beta_2| + \dots + |\beta_k|$ . In the formula, the first item is the loss function, which measures the fitting effect of the regression model on the data. And the second item is the penalty function, which condenses some nonsignificant coefficients into 0. By optimizing the Eq(3), the regression coefficient vector estimation  $\hat{\beta}_{lasso}$  can be obtained, which is called LASSO regression. There is an equivalent constraint form that corresponds to the Eq(3) as follows.

$$\begin{aligned} \min_{\beta} \frac{1}{2} \|Y - X\beta\|_2^2 \\ s.t. \|\beta\|_1 \leq C \end{aligned} \quad (4)$$

where  $C$  is the constraint parameter corresponding to  $\lambda$ , which is negative correlation with punishment intensity. The optimal solution can be obtained by Eq(3) and Eq(4). Because the penalty function is a first order function with absolute value, it cannot be solved by the calculus method. It requires highly complicated numerical calculation to solve the regression equation. Efron et al. [41] proposed least angle regression (LARS) algorithm to solve the planning problem. By using LARS algorithm to solve LASSO regression efficiently, LASSO algorithm is highly regarded by the academic community.

## 2.2. Quantile regression Neural network (QRNN)

Koenker and Bassett presented a method of quantile regression (QR). Compared with the traditional mean regression, QR can more accurately describe the influence of the conditional distribution shape and the variation range of the dependent variables for independent variables. However, this method only learns the linear relationship between the response variables and the input variables under different quantiles, which is not appropriate for the nonlinear relationship in most practical data [42]. Therefore, Taylor put forward a nonlinear QR method based on neural networks. The expression is shown as follows.

$$Q_{y_i}(\tau|x_i) = f(x_i, T_i(\tau), U_i(\tau)), i = 1, 2, \dots, n \quad (5)$$

$$f(x_i, T_i(\tau), U_i(\tau)) = g_2\left\{\sum_{k=1}^K u_{i,k}(\tau)g_1\left[\sum_{j=1}^J t_{i,j,k}(\tau)x_i\right]\right\} \quad (6)$$

where  $\tau \in (0, 1)$  is able to generate different quantiles,  $T_i = \{t_{i,j,k}\}_{j=1,2,\dots,J;k=1,2,\dots,K}$  represents the estimated weight matrix between the input layer and the hidden layer,  $U_i = \{u_{i,k}\}_{k=1,2,\dots,K}$  represents the connection weight vector between the hidden layer and the output layer.  $g_1(\cdot)$  is expressed as an activation

function of the hidden layer using the hyperbolic tangent sigmoid function  $g_1(v) = \frac{1}{1+e^{-v}}$ .  $g_2(\cdot)$  is the output layer function, which is represented by a general linear model.  $f(x_i, T_i(\tau), U_i(\tau))$  is a nonlinear function that is composition of weight vectors  $T_i$  and  $U_i$ .

The parameter estimation values of  $T(\tau) = \{T_1, T_2, \dots, T_n\}$ ,  $U(\tau) = \{U_1, U_2, \dots, U_n\}$  in the QRNN model can be acquired by a programming problem as follows.

$$H(T(\tau), U(\tau)) = \min \sum_{i=1}^n \rho_{\tau}[(y_i(\tau) - f(x_i, T_i(\tau), U_i(\tau)))] + s_1 \sum_{i,j,k} t_{i,j,k}^2 + s_2 \sum_{i,k} u_{i,k}^2 \quad (7)$$

where  $n$  is the number of the sample,  $\rho_{\tau}$  is loss function defined as following Eq(8).  $s_1$  and  $s_2$  are model penalty parameters, which effectively prevent the model from overfitting to improve the accuracy of prediction. The optimal estimate value of  $T(\tau)$  and  $U(\tau)$  can be obtained by optimizing Eq(7), which are recorded as  $\hat{T}(\tau)$  and  $\hat{U}(\tau)$ . With bringing the two values into Eq(5), the conditional quantiles of the response variables are obtained.

$$\rho_{\tau}(u) = \begin{cases} \tau u & , u \geq 0 \\ (\tau - 1)u & , u < 0 \end{cases} \quad (8)$$

### 3. Electricity consumption probability density forecasting method based on LASSO-QRNN

#### 3.1. LASSO-QRNN model

It is known that electricity consumption forecasting is disturbed by diverse external factors and these factors influence the accuracy of forecasting greatly. How to identify significant factors that influence the electricity consumption becomes important. Based on LASSO-QRNN method, the paper mainly explores the influence of historical load and external factors on the electricity consumption forecasting.

The estimate value of the parameters can be obtained by solving the programming problem in Eq(7). However, when the influence factors in the model are voluminous considering the influence of historical load and external factors on electricity consumption forecasting, the variable selection method ought to be used to improve the prediction accuracy. The LASSO regression method is used in this paper, which is based on the original objective function of QRNN and adds the  $L_1$  penalty. The model is shown as follows.

$$\min \sum_{i=1}^n \rho_{\tau}[(y_i(\tau) - f(x_i, T_i(\tau), U_i(\tau)))] + s_1 \sum_{i,j,k} t_{i,j,k}^2 + s_2 \sum_{i,k} u_{i,k}^2 + \lambda(\|T(\tau)\|_1 + \|U(\tau)\|_1) \quad (9)$$

where  $\lambda$  is penalty parameter that is positive correlation with punishment intensity,  $\|T(\tau)\|_1$  and  $\|U(\tau)\|_1$  are 1-norm. In the formula, the first item is the loss function, which measures the fitting effect of the QRNN model on the data. The second item is the penalty function, which can compress nonsignificant coefficients into zero. Then, the process of variable selection is achieved. In order to solve the LASSO-QRNN model, an equivalent constraint form is obtained according to Eq(9) as follows.

$$\begin{aligned} & \min \sum_{i=1}^n \rho_{\tau}[(y_i(\tau) - f(x_i, T_i(\tau), U_i(\tau)))] + s_1 \sum_{i,j,k} t_{i,j,k}^2 + s_2 \sum_{i,k} u_{i,k}^2 \\ & s.t. \|T(\tau)\|_1 + \|U(\tau)\|_1 \leq C \end{aligned} \quad (10)$$

in which  $C$  is the constraint parameter corresponding to  $\lambda$ . The smaller the value of  $C$  is, the stronger the effect of compressing the coefficients into zero is. Furthermore, the programming problem can be solved by using LARS algorithm proposed by Efron, so as to obtain the parameter estimates of the model and the conditional quantile of the electricity consumption. The LASSO-QRNN model not only produces the distribution of electricity consumption values in different quantiles, but also selects more explanatory factors to reduce the data dimension. In conclusion, the LASSO-QRNN method optimizes the prediction model, which lays the foundation for the follow-up prediction work.

### 3.2. Probability density forecasting method based on LASSO-QRNN with kernel density estimation

Kernel density estimation (KDE) method is a nonparametric method for estimating probability density functions [43], fitting observations to simulate a true probability distribution by a smooth spike function. According to the above research, the conditional quantile of electricity consumption in different quantiles can be obtained, which is represented as  $Z_i = Q_{y_i}(\tau | x_i)$ . The conditional quantile  $Z_i$  is utilized as the input values of the kernel function. In addition, by choosing appropriate bandwidth for probability density prediction, the probability density functions of electricity consumption are obtained eventually. Supposing that  $Z_1, Z_2, \dots, Z_n$  are the independent identically distributed random samples, the probability density function at any one point is denoted as  $f$ . The formula is shown as follows.

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - Z_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - Z_i}{h}\right) \quad (11)$$

where  $h$  is the bandwidth that is a smoothing parameter requiring self-actuated setting,  $n$  is the sample size, and  $K(\cdot)$  is regarded as a non-negative kernel function. Among different kernel functions, such as gaussian, rectangular, triangular, epanechnikov, biweight, cosine and optcosine, this paper chooses epanechnikov and gaussian, because the predicted results are more precise compared with the other kernel functions through contrast experiments. The function expressions are respectively shown as follows.

$$K(u) = \frac{3}{4}(1 - u^2) \quad (12)$$

$$s.t. |u| \leq 1$$

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (13)$$

The flow chart of electricity consumption probability density forecasting method based on LASSO-QRNN in this paper is shown in Fig.1.



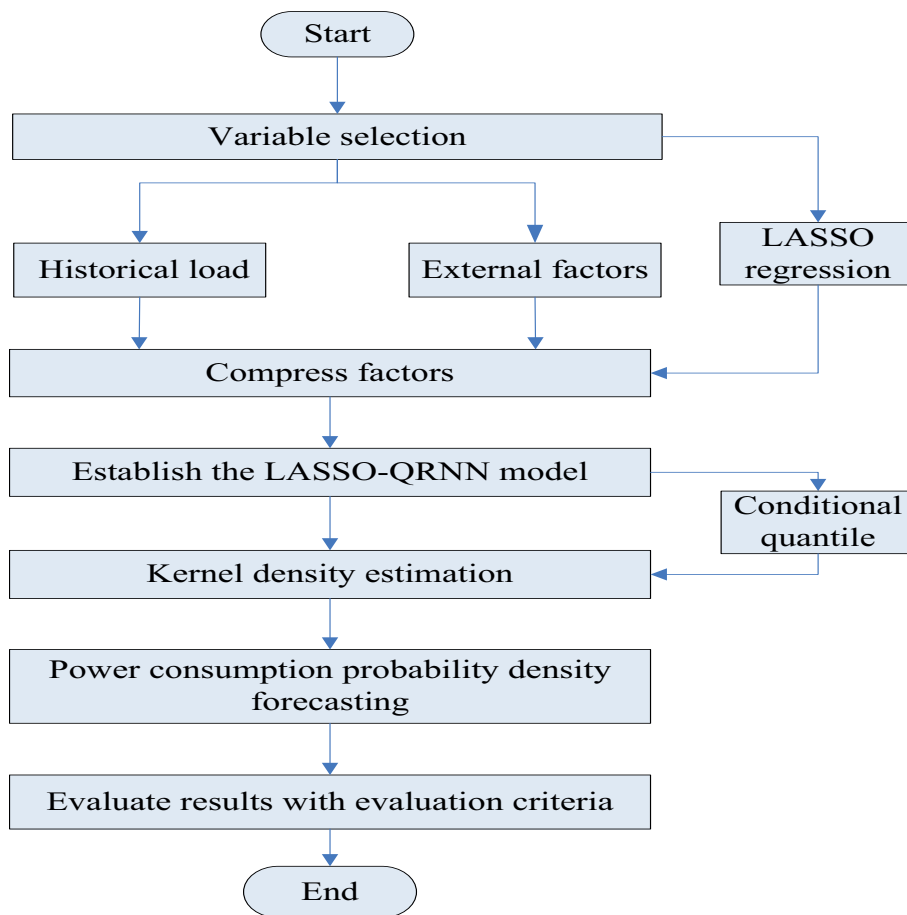


Figure 1: The flow chart of the algorithm

## 4. Experimental Analysis

### 4.1. Evaluation criteria

In this section, we compare LASSO-QRNN with two other situations to show the benefit of introducing LASSO in our method. 1) the QRNN without considering external factors; 2) the QRNN considering external factors without variable selection. Then, LASSO-QRNN is also compared with other state-of-the-art approaches in electricity consumption forecasting, including RBF, BP, QR and NLQR. To verify the effectiveness of our method, the paper employs the following evaluation criteria:

The mean absolute percentage errors (MAPE) is used for error analysis to point prediction results. MAPE is defined as follows.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (14)$$

in which  $n$  indicates the total number of years for electricity consumption to be predicted,  $y_i$  and  $\hat{y}_i$  respectively represent the actual and predicted values of electricity consumption at time  $i$ .

The relative mean square errors (RMSE) is used to reflect the deviation degree of the predicted value to the actual value. RMSE is defined as follows.

$$RMSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n y_i^2} \quad (15)$$

### 4.2. Case study 1: Guangdong province in China

The first case is from Guangdong province of China. The experimental data are available from Guangdong Statistical Yearbook published in 2017 [44]. As a dominating electricity consumption province in China, the electricity consumption in Guangdong province can reach the sum of the surrounding provinces and cities [45]. In terms of electricity consumption structure and energy efficiency, the proportion of coal in fuel sources has declined, and the proportion of high-quality energy (including hydro, solar, wind, etc) in consumption has increased. In recent years, accompanied by the economic development and the adjustment of industrial structure, the electricity consumption in Guangdong province has achieved sustainable growth, which is shown in Fig. 2. Hence, it is important to predict electricity consumption accurately.

In this case study, the historical electricity consumption and external factors from the past four years are used as input data to forecast the annual electricity consumption. Suppose  $Y_t = (y_{t-1}, y_{t-2}, \dots, y_{t-c})$  indicates the historical load of the lag  $c$ , in which  $c$  represents the maximum lag phase of the electricity consumption sequence  $y_t$  to be predicted. The sequence of external influence factors is expressed as  $X_t = (X_t^1, X_t^2, \dots, X_t^m)$ , in which  $m$  is denoted as the number of the external influence factors. The  $d$ -th influence

Table 1: The statement of independent variables for Guangdong province data

Name	independent variables			
	Lag 1	Lag 2	Lag 3	Lag 4
Electricity consumption $Y_t$	$\mathbf{y_{t-1}}$	$y_{t-2}$	$\mathbf{y_{t-3}}$	$y_{t-4}$
GDP $X_t^1$	$x_{t-1}^1$	$x_{t-2}^1$	$x_{t-3}^1$	$x_{t-4}^1$
Second industry proportion $X_t^2$	$x_{t-1}^2$	$x_{t-2}^2$	$\mathbf{x_{t-3}^2}$	$\mathbf{x_{t-4}^2}$
Third industry proportion $X_t^3$	$x_{t-1}^3$	$x_{t-2}^3$	$x_{t-3}^3$	$x_{t-4}^3$
Permanent residents $X_t^4$	$x_{t-1}^4$	$x_{t-2}^4$	$x_{t-3}^4$	$\mathbf{x_{t-4}^4}$
CPI $X_t^5$	$x_{t-1}^5$	$x_{t-2}^5$	$x_{t-3}^5$	$x_{t-4}^5$
Electricity consumption per unit of GDP $X_t^6$	$x_{t-1}^6$	$x_{t-2}^6$	$x_{t-3}^6$	$x_{t-4}^6$
Elasticity ratio of electricity consumption $X_t^7$	$x_{t-1}^7$	$x_{t-2}^7$	$x_{t-3}^7$	$x_{t-4}^7$
PCDI $X_t^8$	$x_{t-1}^8$	$x_{t-2}^8$	$x_{t-3}^8$	$x_{t-4}^8$
Per capita consumption expenditure $X_t^9$	$x_{t-1}^9$	$x_{t-2}^9$	$x_{t-3}^9$	$x_{t-4}^9$
Urbanization rate $X_t^{10}$	$x_{t-1}^{10}$	$x_{t-2}^{10}$	$x_{t-3}^{10}$	$x_{t-4}^{10}$
Total volume of retail sales $X_t^{11}$	$x_{t-1}^{11}$	$x_{t-2}^{11}$	$x_{t-3}^{11}$	$x_{t-4}^{11}$
Fixed investment $X_t^{12}$	$x_{t-1}^{12}$	$x_{t-2}^{12}$	$x_{t-3}^{12}$	$x_{t-4}^{12}$
Mean annual precipitation $X_t^{13}$	$x_{t-1}^{13}$	$x_{t-2}^{13}$	$x_{t-3}^{13}$	$x_{t-4}^{13}$
Mean annual temperature $X_t^{14}$	$x_{t-1}^{14}$	$x_{t-2}^{14}$	$x_{t-3}^{14}$	$x_{t-4}^{14}$

factor  $X_t^d$  satisfies the condition  $X_t^d = (x_{t-1}^d, x_{t-2}^d, \dots, x_{t-q}^d)$ . In the case,  $m$  is set to 14. In addition, both  $c$  and  $q$  are equal to 4. There are 60 independent variables in total, as shown in Table.1. LASSO regression is used to select the most influential factors of electricity consumption. The selected factors are historical load with a lag of 1 and 3 periods, second industry proportion with a lag of 3 and 4 periods and permanent residents with a lag of 4 periods, as shown in boldface text of Table.1. The procedure chart of compressing factors is shown in Fig. 3. The dotted line indicates the trajectory of the factor selection process. The abscissa indicates the constraint coefficient  $C$ , and the ordinate indicates the regression coefficient  $\beta$ . The two coefficients correspond to the parameters in Eq(4). From the output of LASSO, we can see that the electricity consumption in Guangdong province closely correlates with the historical electricity consumption, the secondary industry proportion and the permanent residents. The description of the pertinent literature further reflects this result [46]. During 1995-2016, with the development of the city, the number of permanent residents in Guangdong province is growing and the second industry has achieved great development, which leads to the growth of electricity consumption.

The annual total electricity consumption and extracted factors from 1995 to 2016 are selected as experimental data, which are already normalized. The training sample is the electricity consumption and extracted

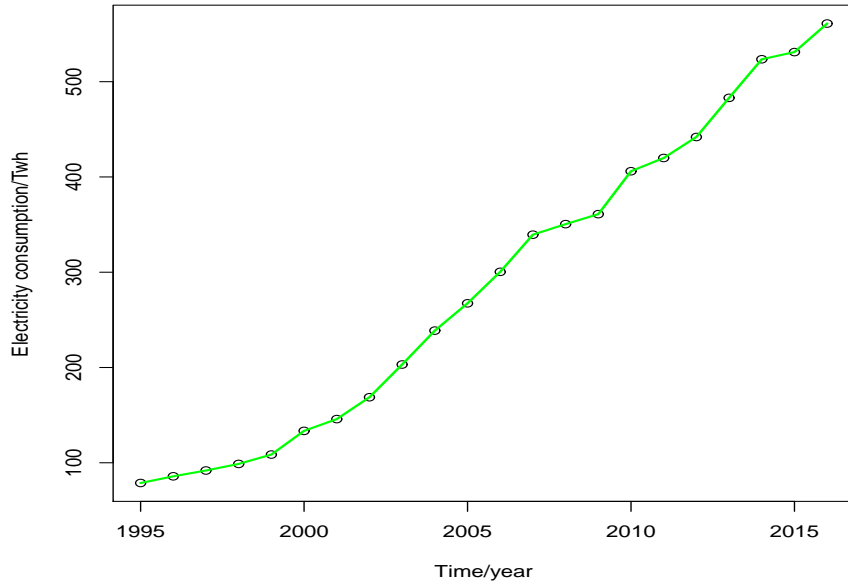


Figure 2: The electricity consumption in Guangdong Province from 1995 to 2016

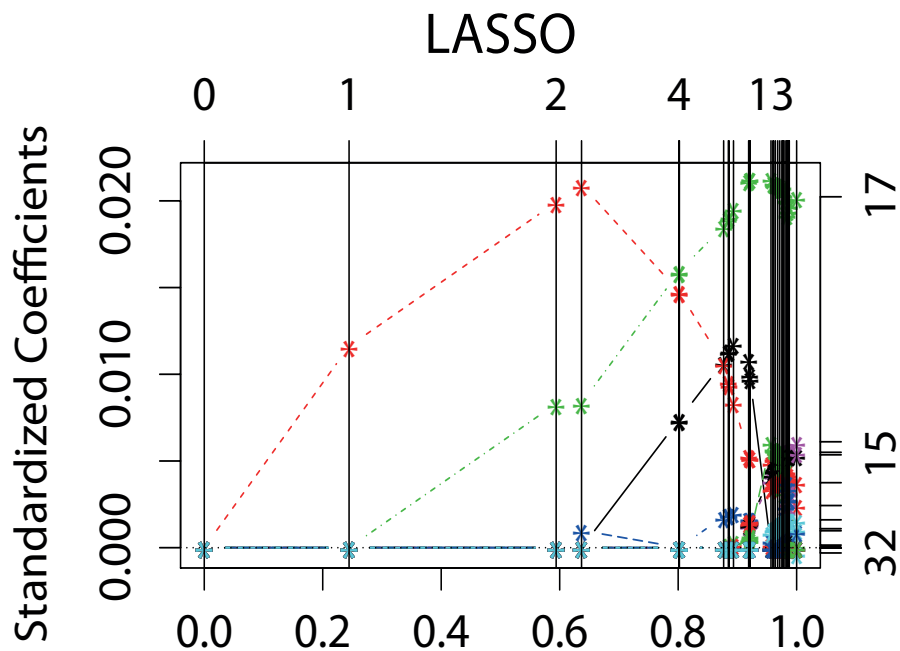


Figure 3: The procedure chart of compressing factors (Guangdong)

1  
2  
3 factors from 1995 to 2010. Through rolling forecasting method, the historical electricity consumption and  
4 influence factors in previous four years are considered as input variables to forecast the annual electricity  
5 consumption. The electricity consumption values for a total of 12 years from 1999 to 2010 are the expected  
6 output values. In other words, a total of 12 groups of samples are selected to train the LASSO-QRNN  
7 model, then we can further determine the optimal structure of the neural network. According to the opti-  
8 mal structure, the electricity consumption values from 2011 to 2016 are predicted. Without considering the  
9 external factors, the input layer of the neural network structure is 4 nodes, the hidden layer is 1 node, and  
10 the output layer is 1 node. Considering external influence factors without variable selection, the nodes of  
11 input layer, hidden layer and output layer are 60, 1, 1, respectively. When the LASSO regression is used for  
12 variable selection considering external influence factors, the input structure of the neural network structure  
13 is 5 nodes, the hidden layer is 1 node, and the output layer is 1 node. The neural network training number  
14 is set to 100. And the penalty parameter is set to 2, which denotes that  $s_1$  and  $s_2$  are equal to 2 in Eq(7).  
15 Through the trained neural network structure, consecutive quantiles under different quantiles from 2011 to  
16 2016 are obtained, which have been restored from normalization. This paper selects the quantile range from  
17 0.01 to 0.99, and the interval is 0.05. Then the predicted electricity consumption quantile is used as the  
18 input variable of epanechnikov kernel function. Combined with the kernel density estimation method, the  
19 annual probability density curve from 2011 to 2016 is obtained.

20  
21  
22 Fig. 4 shows the probability density curve from 2011 to 2016 without considering the external factors. We  
23 compare it with the one with factor selection, as shown in Fig. 5. In Figs. 4 and 5, the value of the abscissa  
24 corresponding to the red vertical line represents the actual value of the annual electricity consumption in  
25 Guangdong province of China. The following conclusions can be drawn: Without considering the external  
26 factors, the true values almost deviate from the highest probability point of the probability density curve.  
27 Most of the true values appear at the end of the probability density curve, and the predicted values are far  
28 from the true values. However, in consideration of external factors with variable selection, the true values  
29 are almost in the vicinity of the highest probability point in the probability density curve. There is almost  
30 no case in the tail of the probability density curve. The probability density forecasting method based on  
31 LASSO-QRNN can produce better probability density curve of predicted values closer to the true values.

32  
33  
34 To show the effect of variable selection in our method, we compare three situations as we introduced  
35 earlier: QRNN without considering external factors, QRNN considering external factors without variable  
36 selection and LASSO-QRNN considering external factors, as shown in the Table. 2. According to the Table. 2,  
37 the MRPE and RMSE of the predicted values considering the external factors are smaller than those without  
38 considering the external factors in the mode and median. Moreover, considering the external factors, the  
39 MRPE and RMSE of the predicted values with variable selection are smaller than those without variable  
40 selection in the mode and median. It can be concluded that the predicted results using the LASSO-QRNN  
41 method for variable selection are better than the predicted results without variable selection and without  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

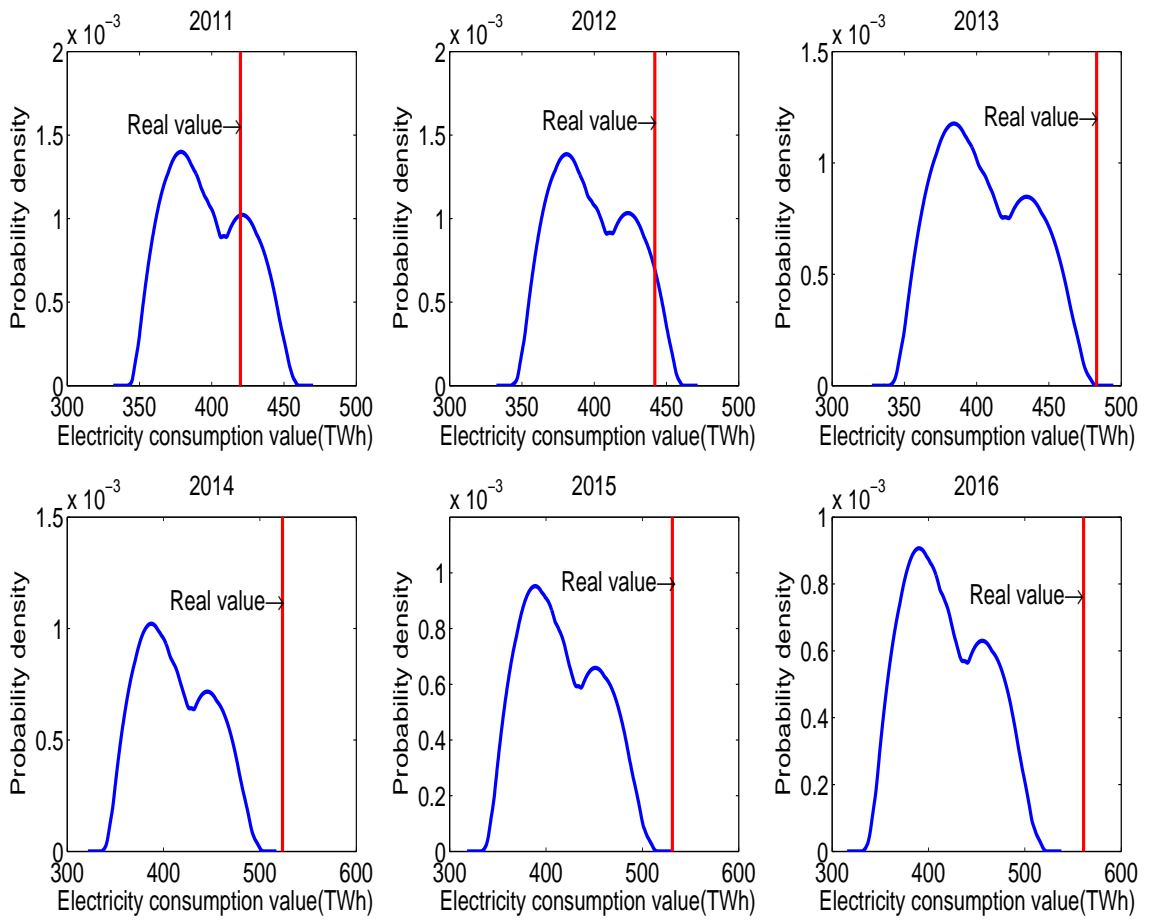


Figure 4: Probability density curve without considering the external factors in Guangdong province of China

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

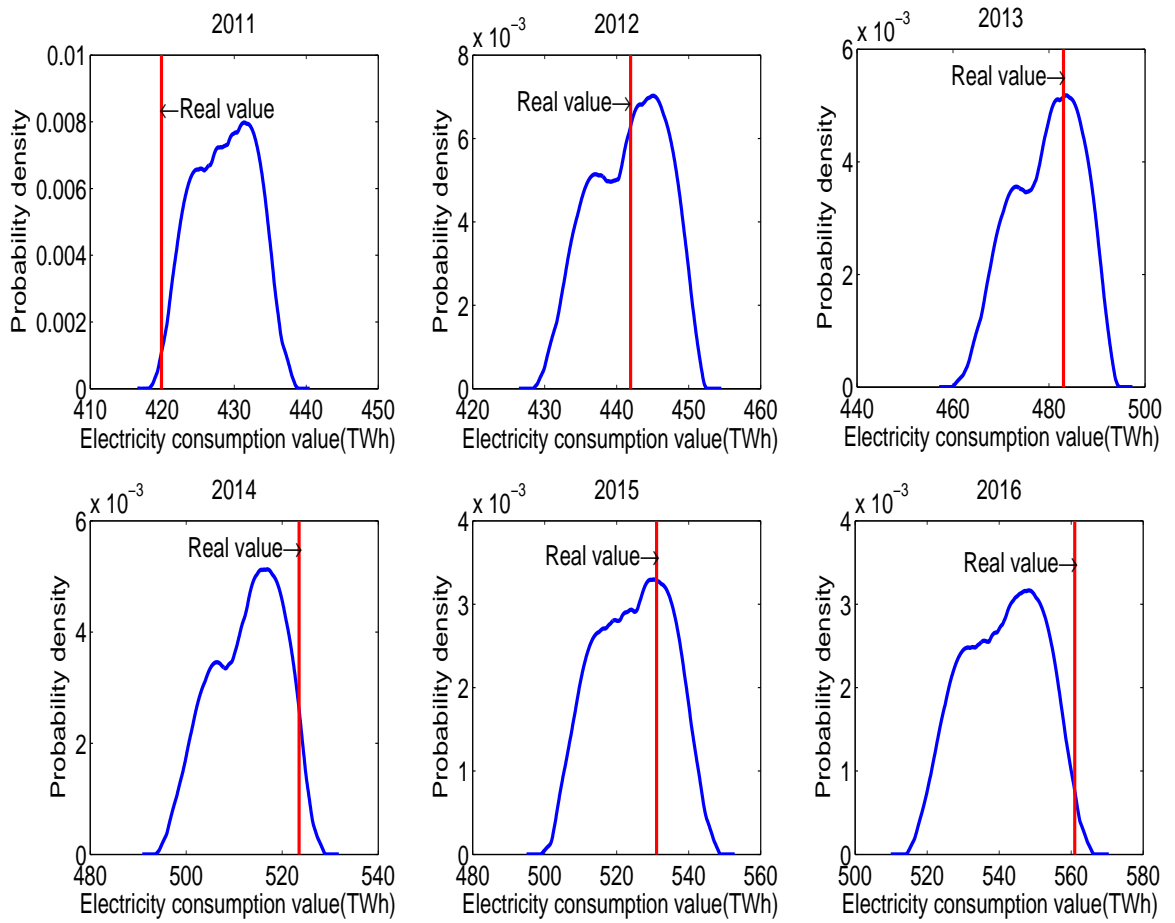


Figure 5: Probability density curve considering the external factors (LASSO) in Guangdong province of China

Table 2: Prediction errors in three situations of Guangdong province data

Evaluation criteria		Without considering external factors	Considering external factors	
		(QRNN)	QRNN	LASSO-QRNN
MRPE(%)	Median	15.41	8.60	1.83
	Mode	21.21	5.25	1.21
RMSE(%)	Median	3.22	1.10	0.05
	Mode	5.63	0.50	0.02

Table 3: Forecasting results of Guangdong province data based on LASSO-QRNN

Year	Actual value/TWh	Optimal forecast value/TWh	APFE/%
2011	419.90	431.41	2.74
2012	441.94	445.06	0.71
2013	483.01	483.60	0.12
2014	523.52	516.89	1.27
2015	531.07	530.25	0.15
2016	561.01	548.32	2.26

considering the external factors. The predicted results of the mode is slightly better than the median and the MRPE are 1.21 % and 1.83 %, respectively. In the big data environment with smart grid, this shows that probability density forecasting of electricity consumption ought to fully consider the external influential factors to achieve the purpose of optimal prediction model from the perspective of high dimensional data analysis.

Based on the probability density forecasting method of LASSO-QRNN, we present the optimal point predictions and absolute percentage forecast errors (APFE) of annual electricity consumption from 2011 to 2016, as shown in Table. 3. In consideration of external factors, the MAPE of the probability density forecasting based on LASSO-QRNN is 1.21%. The minimum APFE of the probability density prediction is 0.12 %. It can be concluded from Table. 3 that the probability density forecasting method based on LASSO-QRNN can obtain more accurate prediction results, which are close to the actual value.

More details on the state of the art are provided to better illustrate the advantages of LASSO-QRNN method. The paper compares the prediction error obtained from LASSO-QRNN with the error obtained by RBF [35], BP [36], QR [37] and NLQR [38] in Table. 4. LASSO-QRNN method shows significantly better performance, reducing the indeterminacy of electricity consumption forecasting.



Table 4: Prediction errors of state-of-the-art methods for Guangdong province data

Evaluation criteria	RBF	BP	QR (Median)	QR (Mode)	NLQR (Median)	NLQR (Mode)
MRPE(%)	14.00	30.06	36.51	8.91	19.89	23.99
RMSE(%)	2.83	9.44	13.41	0.99	5.24	7.09

#### 4.3. Case study 2: California in U.S.

The second case is from California in U.S. The experimental data are available from U.S. Energy Information Administration website [47], which contains annual historical retail sales (consumption) of electricity and external factors including total energy, average electricity price, number of electricity consumers and so on. California’s electricity generation, transmission and distribution network are a management system with a wide range and complexity [48]. Its fuel sources include higher percentages of natural gas, hydro, and non-hydro renewables, and lower percentages of coal, petroleum, and nuclear [49]. Until 2003, the level of electricity consumption in California ranked the second among the 50 states. The electricity consumption in California from 1990 to 2015 is presented in Fig. 6. It shows a declining period after the peak value of electricity consumption in 2008, which increases the difficulty and challenge tremendously for forecasting the electricity consumption. In [50], the statewide annual electricity consumption prediction was presented during the period 2008-2018, assuming average temperatures. It is worth noting that the electricity consumption in California used in this paper is different from that in this literature.

According to the above observations, the uniform rolling forecast approach is employed to forecast the annual electricity consumption in California, and the number of external factors  $m$  is set to 8. In addition, both  $c$  and  $q$  are equal to 4. As can be seen in Table.4, there are 36 independent variables in total. In this case, the LASSO regression is also adopted to select the most appropriate variables for electricity consumption forecasting. The procedure chart of compressing factors is shown in Fig. 7. The selected factors are total energy with a lag of 1 ,2 and 4 periods, average electricity price with a lag of 1 and 4 periods, and number of electricity customers with a lag of 2 and 3 periods, as shown in bold section of Table. 5. As similar as [51], it appears that economic growth and the fluctuation of electricity price are the main factors correlated with electricity consumption. Similarly to the previous case study, LASSO regression alleviates high-dimensional data problems through the selection process of key factors, which is of great significance in reality.

The electricity consumption and influence factors from 1990 to 2009 are used as the training sample. Similar to the Guangdong province case, the rolling forecasting method is also adopted to process data. The electricity consumption values for a total of 16 years from 1994 to 2009 are the expected output values. As a result, 16-year samples are selected to train the LASSO-QRNN, which identifies the parameters of the

Table 5: The statement of independent variables for California data

Name	independent variables			
	Lag 1	Lag 2	Lag 3	Lag 4
Electricity consumption $Y_t$	$y_{t-1}$	$y_{t-2}$	$y_{t-3}$	$y_{t-4}$
Personal income $X_t^1$	$x_{t-1}^1$	$x_{t-2}^1$	$x_{t-3}^1$	$x_{t-4}^1$
Population $X_t^2$	$x_{t-1}^2$	$x_{t-2}^2$	$x_{t-3}^2$	$x_{t-4}^2$
Per capita personal income $X_t^3$	$x_{t-1}^3$	$x_{t-2}^3$	$x_{t-3}^3$	$x_{t-4}^3$
Personal consumption expenditures $X_t^4$	$x_{t-1}^4$	$x_{t-2}^4$	$x_{t-3}^4$	$x_{t-4}^4$
GDP $X_t^5$	$x_{t-1}^5$	$x_{t-2}^5$	$x_{t-3}^5$	$x_{t-4}^5$
Total energy $X_t^6$	$x_{t-1}^6$	$x_{t-2}^6$	$x_{t-3}^6$	$x_{t-4}^6$
Average electricity price $X_t^7$	$x_{t-1}^7$	$x_{t-2}^7$	$x_{t-3}^7$	$x_{t-4}^7$
Number of electricity customers $X_t^8$	$x_{t-1}^8$	$x_{t-2}^8$	$x_{t-3}^8$	$x_{t-4}^8$

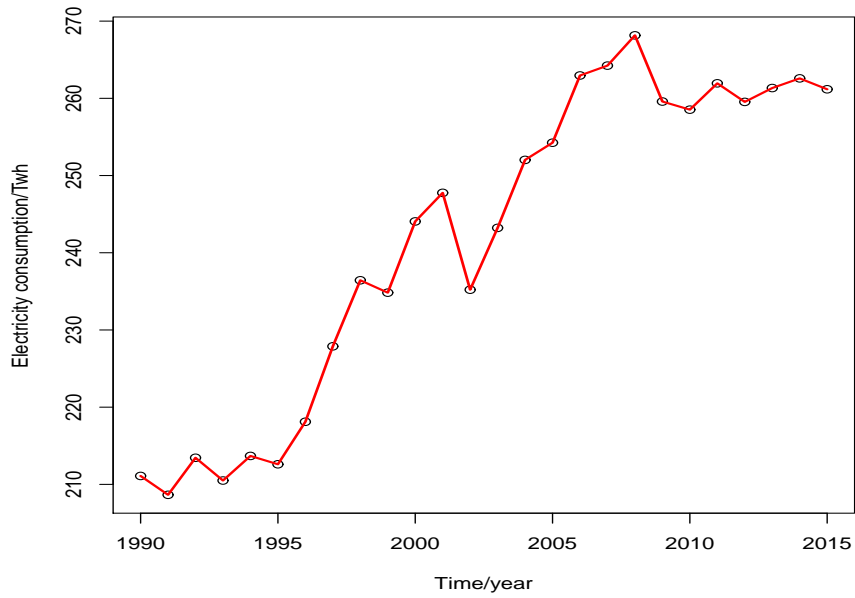


Figure 6: The electricity consumption of California from 1990 to 2015

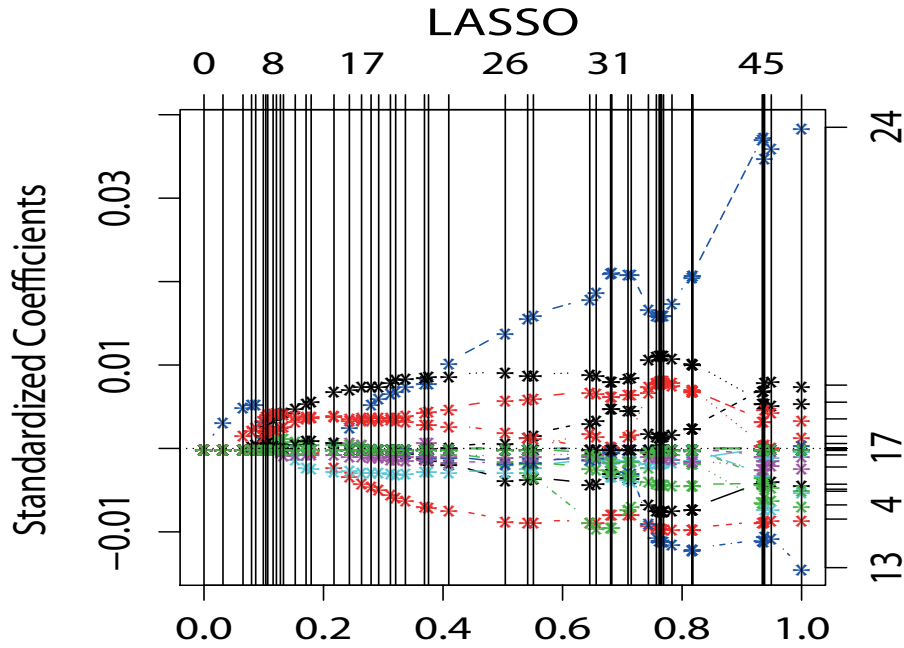


Figure 7: The procedure chart of compressing factors (California)

neural network for obtaining the optimal structure. From 2010 to 2015, a total of six years of electricity consumption forecasting is implemented. Without considering the external factors, the input layer of the neural network is 4 nodes. Considering external influence factors without variable selection, the node number of input layer is 36. When LASSO regression is used for variable selection considering external influence factors, the input structure of the neural network is 7 nodes. In all three settings, the hidden layer is 2 nodes, and the output layer is 1 node. The neural network training number is set to 10, and the penalty parameter is set to 5. Then the predicted electricity consumption quantile is considered as the input variable of Gaussian kernel function, which have been inversely normalized. Through the trained neural network structure and kernel density estimate algorithm, the results of probability density forecasting under different quantiles for any year from 2010 to 2015 are obtained.

From the above, the probability density curve of consecutive quantiles can be obtained from 2010 to 2015. Fig. 8 shows the probability density curve from 2010 to 2015 without considering the external factors. Similar to the previous case study, it is compare with the results considering external factors with variable selection, as shown in Fig. 9. In Figs. 8 and 9, the value of the abscissa corresponding to the red vertical line indicates the actual value of the annual electricity consumption in California, USA. We obtain the following findings: Without considering the external factors, the true values almost deviate from the highest probability point of the probability density curve. In addition, the obtained probability density curve is inconsecutive and rough.

Nevertheless, the true values are almost all in the vicinity of the highest probability point in the probability density curve considering the external factors. Therefore it is easy to find that the LASSO-QRNN method can acquire consecutive and slippy the probability density curves. It means that the predicted values are more likely to get close to the true values. The probability density curve can not only provide more detailed information for electricity consumption forecasting, but also open up new ideas and methods for electricity consumption forecasting.

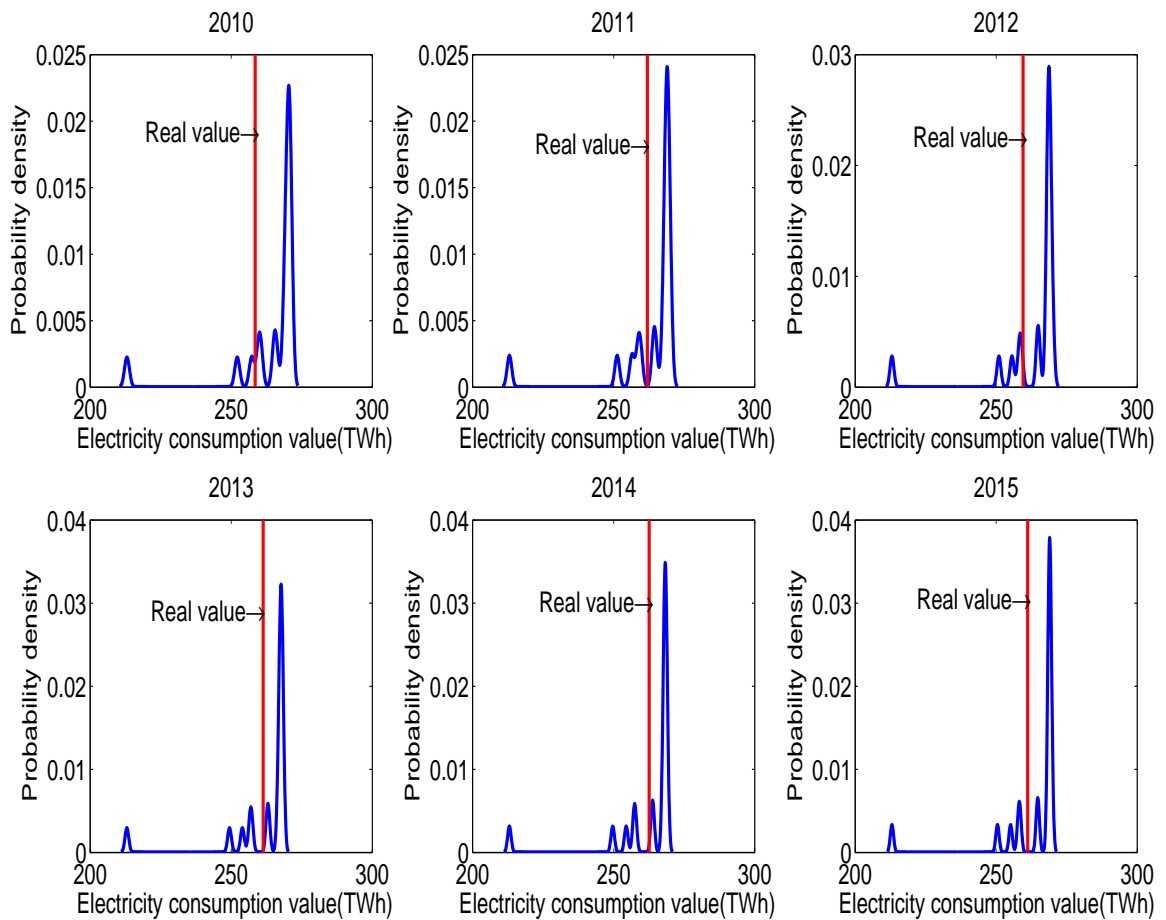


Figure 8: Probability density curve without considering the external factors in California of U.S

In order to illustrate the effectiveness of this method, as similar as the case above-mentioned, the prediction error of the three situations are also compared in Table. 6. According to the experimental results in Table. 6, the MRPE and RMSE of the predicted results considering the external factors are less than the predicted results without considering the external factors in the median. It can be concluded that the predicted results using the LASSO-QRNN method for variable selection are better than the predicted results

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

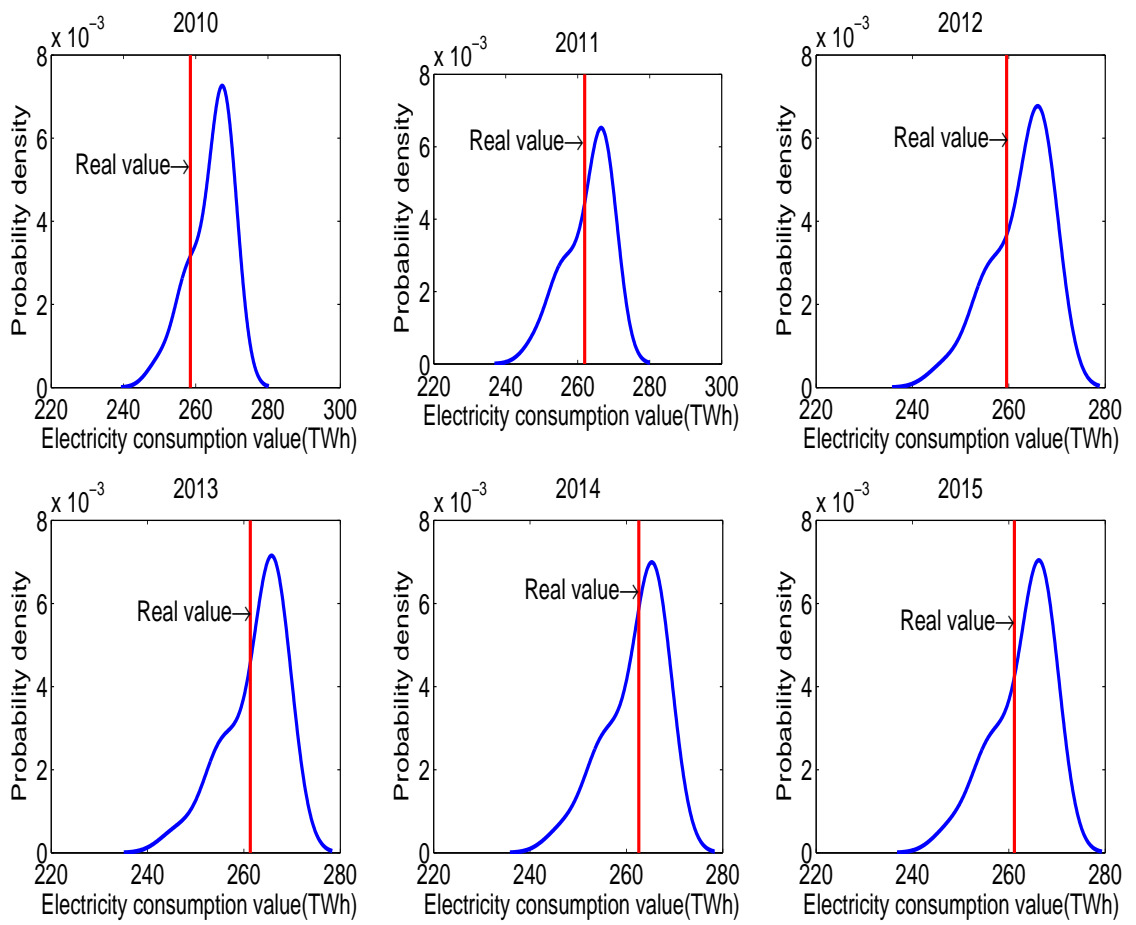


Figure 9: Probability density curve considering the external factors (LASSO) in California of U.S

Table 6: Prediction errors in three situations of California data

Evaluation criteria		Without considering external factors	Considering external factors	
		(QRNN)	QRNN	LASSO-QRNN
MRPE(%)	Median	7.44	2.09	1.29
	Mode	3.12	3.89	2.04
RMSE(%)	Median	0.56	0.05	0.02
	Mode	0.10	0.16	0.05

Table 7: Forecasting results of California data based on LASSO-QRNN

Year	Actual value/TWh	Optimal forecast value/TWh	APFE/%
2010	258.33	259.76	0.48
2011	261.94	258.45	1.33
2012	259.54	257.31	0.86
2013	261.33	256.72	1.77
2014	262.59	257.07	2.10
2015	261.17	258.07	1.19

without variable selection and without considering the external factors. The predicted results of the median is slightly better than the mode and the MRPE are 1.29% and 2.04%, respectively. This demonstrates that the electricity consumption probability density forecasting method based on LASSO-QRNN is able to lessen the nondeterminacy caused by external factors for improving the prediction precision of electricity consumption, thus avoiding larger prediction errors and economic losses.

The forecasting results considering external factors based on LASSO-QRNN are shown in Table. 7, which includes actual values, optimal point forecasts and the APFE of electricity consumption in California during the period of 2010-2015. From Table. 7, it can be seen that the MAPE of the probability density forecasting based on LASSO-QRNN is 1.29%. In addition, the minimum APFE of the probability density prediction is 0.48%. Consequently, it can be concluded that the probability density forecasting method based on LASSO-QRNN can obtain high-precision prediction results.

Similar to the case of Guangdong province data, the prediction error obtained by LASSO-QRNN is also compared with that of BP, RBF, QR and NLQR as shown in Table. 8. The prediction accuracy of BP, RBF, QR and NLQR is lower than that of LASSO-QRNN for California data, which justifies the superiority of the proposed method.

Table 8: Prediction errors of state-of-the-art methods for California data

Evaluation criteria	RBF	BP	QR	QR	NLQR	NLQR
			(Median)	(Mode)	(Median)	(Mode)
MRPE(%)	6.28	7.05	7.06	3.87	2.02	2.08
RMSE(%)	0.40	0.50	0.50	0.16	0.04	0.05

## 5. Conclusions

Aiming at the electricity consumption forecasting, this paper proposes a probability density forecasting method – LASSO-QRNN. Firstly, the method selects the most informative features from external factors by using LASSO. It reduces data dimensions effectively without hurting predictive performance. Secondly, the QRNN method produces overall probability distribution of the electricity consumption at different quantiles in details, which is valuable information for managing electricity consumption. The performance of LASSO-QRNN is analyzed on two real-world datasets from Guangdong province of China and California of U.S. Our main findings are:

By taking into account the external factors, the MRPE and RMSE of the probability density forecasting results using LASSO-QRNN method are smaller than those without considering the external factors in terms of the median and mode. With respect to the probability density prediction curve, the true values are almost all in the vicinity of the highest probability point in the probability density curve considering the external factors. However, the probability of the true value appearing in the tail is significantly increased if the external factors are not considered. Derived from the probability density curve, the MRPE considering the external factors is limited within 2% by using LASSO-QRNN method. Furthermore, the comparative analysis with QRNN, RBF, BP, QR and NLQR displays that our method provides a more accurate prediction and reduces the uncertainty of electricity consumption forecasting in power systems. From the simulation results, the most relevant factors affecting electricity consumption in different regions are significantly discrepant. The historical electricity consumption, second industry proportion and permanent residents are selected as the feature to train the model in the data of Guangdong province, while total energy, average electricity price and number of electricity customers are selected to train the model for the U.S. data. The sequence diagram of Guangdong province data (Fig.2) is monotonically increasing other than the California data, which verifies that the electricity consumption of California (Fig.6) is more intractable to achieve accurate prediction than Guangdong data for the aggravation of instability. In summary, LASSO-QRNN is capable of important features and learning from high-dimensional data effectively. It produces more informative and accurate results than other existing methods.

## ACKNOWLEDGEMENT

This paper is funded by the National Natural Science Foundation (No.71771073,71401049,U1765201), the CRSRI Open Research Program (Program SN CKWV2017525/KY) and the Open Research Fund of State Key Laboratory of simulation and Regulation of Water Cycle in River Basin (China Institute of Water Resources and Hydropower Research) (Grant NO IWHR-SKL-201605).

## References

- [1] Pao HT. Forecasting energy consumption in taiwan using hybrid nonlinear models. *Energy* 2009;34(10):1438–144611.
- [2] Kaytez F, Taplamacioglu MC, Cam E, Hardalac F. Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power & Energy Systems* 2015;67(67):431–8.
- [3] Ming M, Niu D. Annual electricity consumption analysis and forecasting of china based on few observations methods. *Energy Conversion & Management* 2011;52(2):953–7.
- [4] Sadownik R, Barbosa EP. Short-term forecasting of industrial electricity consumption in brazil. *Journal of Forecasting* 1999;18(3):215–24.
- [5] Torrini FC, Souza RC, Oliveira FLC, Pessanha JFM. Long term electricity consumption forecast in brazil: A fuzzy logic approach. *Socio-Economic Planning Sciences* 2016;54:18–27.
- [6] Azadeh A, Ghaderi SF, Sohrabkhani S. Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. *Energy Conversion & Management* 2008;49(8):2272–8.
- [7] Bassamzadeh N, Ghanem R. Multiscale stochastic prediction of electricity demand in smart grids using bayesian networks. *Applied energy* 2017;193:369–80.
- [8] Fan S, Chen L. Short-term load forecasting based on an adaptive hybrid method. *IEEE Transactions on Power Systems* 2006;21(1):392–401.
- [9] Abdel-Aal RE, Al-Garni AZ. Forecasting monthly electric energy consumption in eastern saudi arabia using univariate time-series analysis. *Energy* 1997;22(22):1059–69.
- [10] Taylor JW. Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research* 2010;204(1):139–52.
- [11] Graditi G, Ferlito S, Adinolfi G. Comparison of photovoltaic plant power production prediction methods using a large measured dataset. *Renewable Energy* 2016;90:513–9.
- [12] Hong T, Fan S. Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting* 2016;32(3):914–38.
- [13] Wu X, Zhu X, Wu GQ, Ding W. Data mining with big data. *IEEE Transactions on Knowledge & Data Engineering* 2013;26(1):97–107.
- [14] Shao Z, Gao F, Zhang Q, Yang SL. Multivariate statistical and similarity measure based semiparametric modeling of the probability distribution: A novel approach to the case study of mid-long term electricity consumption forecasting in china. *Applied Energy* 2015;156:502–18.
- [15] Amina M, Kodogiannis VS, Petrounias I, Tomtsis D. A hybrid intelligent approach for the prediction of electricity consumption. *International Journal of Electrical Power & Energy Systems* 2012;43(1):99–108.
- [16] Meng M, Fu Y, Shi H, Wang X. A small-sample adaptive hybrid model for annual electricity consumption forecasting. *Mathematical Problems in Engineering* 2017;:4–12.



- 1  
2  
3  
4 [17] Al-Ghandoor A, Al-Hinti I, Jaber JO, Sawalha SA. Electricity consumption and associated ghg emissions of the jordanian  
5 industrial sector: Empirical analysis and future projection. *Energy Policy* 2008;36(1):258–67.
- 6 [18] van der Meer D, Widén J, Munkhammar J, et al. Review on probabilistic forecasting of photovoltaic power production  
7 and electricity consumption. *Renewable and Sustainable Energy Reviews* 2018;81(P1):1484–512.
- 8 [19] van der Meer D, Shepero M, Svensson A, Widén J, Munkhammar J. Probabilistic forecasting of electricity consump-  
9 tion, photovoltaic power generation and net demand of an individual building using gaussian processes. *Applied Energy*  
10 2018;213:195–207.
- 11 [20] Bianco V, Manca O, Nardini S. Electricity consumption forecasting in italy using linear regression models. *Energy*  
12 2009;34(9):1413–21.
- 13 [21] He Y, Xu Q, Wan J, Yang S. Short-term power load probability density forecasting based on quantile regression neural  
14 network and triangle kernel function. *Energy* 2016;114:498–512.
- 15 [22] He Y, Liu R, Li H, Wang S, Lu X. Short-term power load probability density forecasting method using kernel-based  
16 support vector quantile regression and copula theory. *Applied Energy* 2017;185:254–66.
- 17 [23] Fan J, Lv J. A selective overview of variable selection in high dimensional feature space. *Statistica Sinica* 2009;20(1):101–48.
- 18 [24] Tibshirani RJ. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society* 1996;58:267–88.
- 19 [25] Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. World Publishing Corporation; 2015.
- 20 [26] Tibshirani RJ. The lasso problem and uniqueness. *Electronic Journal of Statistics* 2013;7(1):1456–90.
- 21 [27] Ziel F, Croonenbroeck C, Ambach D. Forecasting wind power - modeling periodic and non-linear effects under conditional  
22 heteroscedasticity. *Applied Energy* 2016;177:285–97.
- 23 [28] Huebner GM, Hamilton I, Chalabi Z, Shipworth D, Oreszczyn T. Explaining domestic energy consumption - the compar-  
24 ative contribution of building factors, socio-demographics, behaviours and attitudes. *Applied Energy* 2015;159:589–600.
- 25 [29] Zhang JX, Song W, Chen ZH, Wei JH, Liao YJ, Lei J, et al. Prognostic and predictive value of a microrna signature in  
26 stage ii colon cancer: a microrna expression analysis. *Lancet Oncology* 2013;14(13):1295–306.
- 27 [30] Taylor JW. A quantile regression neural network approach to estimating the conditional density of multiperiod returns.  
28 *Journal of Forecasting* 2000;19(4):299–311.
- 29 [31] Pradeepkumar D, Ravi V. Forecasting financial time series volatility using particle swarm optimization trained quantile  
30 regression neural network. *Applied Soft Computing* 2017;58:35–52.
- 31 [32] Cannon AJ. Quantile regression neural networks: Implementation in r and application to precipitation downscaling.  
32 *Computers & Geosciences* 2012;46(9):9.
- 33 [33] Xu Q, Deng K, Jiang C, Sun F, Huang X. Composite quantile regression neural network with applications. *Expert Systems*  
34 *with Applications* 2017;76:129–39.
- 35 [34] He Y, Li H. Probability density forecasting of wind power using quantile regression neural network and kernel density  
36 estimation. *Energy Conversion and Management* 2018;164:374–84.
- 37 [35] He YY, Xu QF, Yang SL, Yu BG. A power load probability density forecasting method based on rbf neural network  
38 quantile regression. *Proceedings of the CSEE* 2013;33(1):93–8.
- 39 [36] Cui K, Qin X. Virtual reality research of the dynamic characteristics of soft soil under metro vibration loads based on bp  
40 neural networks. *Neural Computing & Applications* 2017;:1–10.
- 41 [37] Koenker R, Bassett G. Regression quantiles. *Econometrica* 1978;46(1):33–50.
- 42 [38] Andr s JD, Landajo M, Lorca P. Using nonlinear quantile regression for the estimation of software cost. In: *International*  
43 *Conference on Hybrid Artificial Intelligence Systems*. 2018, p. 422–32.
- 44 [39] Wang D, Zhang Z. Variable selection for linear regression models: a survey. *Journal of Applied Statistics And Management*  
45 2010;(04):615–27.
- 46 [40] Geer SAVD. High-dimensional generalized linear models and the lasso. *Annals of Statistics* 2008;36(2):614–45.

- 1  
2  
3  
4 [41] Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. *Annals of Statistics* 2004;32(2):407–51.  
5 [42] Xu Q, Jiang C. Quantile partial adjustment model and its application. *The Journal of Quantitative & Technical Economics*  
6 2011;(8):115–33.  
7 [43] Sheather SJ, Jones MC. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the*  
8 *Royal Statistical Society* 1991;53(3):683–90.  
9 [44] province statistic bureau G. *GUANGDONG STATISTICAL YEARBOOK*. China Statistics Press; 2017.  
10 [45] Wang R. Study on the relationship between economic growth and electricity consumption in guangdong province. *Prices*  
11 *Monthly* 2010;(7):64–7.  
12 [46] LianhongLV , Hong , Baoliu , ZHANG . Relationship between electricity consumption and economic growth of guangdong  
13 province in china. *Frontiers in Energy* 2012;6(4):351–5.  
14 [47] Electric sales, revenue, and average price. Website; 2018. [https://www.eia.gov/electricity/sales\\_revenue\\_price/](https://www.eia.gov/electricity/sales_revenue_price/).  
15 [48] Franco G, Sanstad AH. Climate change and electricity demand in california. *Climatic Change* 2008;87(1):139–51.  
16 [49] USEIA . State electricity profiles 2003. US Department of Energy 2006;0629.  
17 [50] California energy commission. Website; 2007. [https://www.energy.ca.gov/2007publications/CEC-100-2007-008/](https://www.energy.ca.gov/2007publications/CEC-100-2007-008/CEC-100-2007-008-CMF.PDF)  
18 [CEC-100-2007-008-CMF.PDF](https://www.energy.ca.gov/2007publications/CEC-100-2007-008/CEC-100-2007-008-CMF.PDF).  
19 [51] Brown RE, Koomey JG. Electricity use in california: past trends and present usage patterns. *Energy Policy* 2003;31(9):849–  
20 64.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65