Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory

He, Yaoyao; Liu, Rui; Li, Haiyan; Wang, Shuo; Lu, Xiaofen

DOI:
10.1016/j.apenergy.2016.10.079

Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory

Yaoyao He[a,b]*, Rui Liu[a,b] , Haiyan Li[a,b], Shuo Wang[c], Xiaofen Lu[c]

[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]School of Computer Science, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

Abstract

Penetration of smart grid prominently increases the complexity and uncertainty in scheduling and operation of power systems. Probability density forecasting methods can effectively quantify the uncertainty of power load forecasting. The paper proposes a short-term power load probability density forecasting method using kernel-based support vector quantile regression (KSVQR) and Copula theory. As the kernel function can influence the prediction performance, three kernel functions are compared in this work to select the best one for the learning target. The paper evaluates the accuracy of the prediction intervals considering two criteria, prediction interval coverage probability (PICP) and prediction interval normalized average width (PINAW). Considering uncertainty factors and the correlation of explanatory variables for power load prediction accuracy are of great importance. A probability density forecasting method based on Copula theory is proposed in order to achieve the relational diagram of electrical load and real-time price. The electrical load forecast accuracy of the proposed method is assessed by means of real datasets from Singapore. The simulation results show that the proposed method has great potential for power load forecasting by selecting appropriate kernel function for KSVQR model.

Key words: Short-term power load probability density forecasting; Support vector quantile regression; PI coverage probability; PI normalized average width; Copula theory; real-time price

---

* *Corresponding author . Tel: +86 551 2919150*

E-mail *addresses*: hy-342501y @163.com (Yaoyao He),lr_0329@163.com(Rui Liu).

1. Introduction

Load forecasting is a fundamental and vital task for economically efficient operation and controlling of power systems. It is used for energy management, unit commitment and load dispatch. The high accuracy of load forecasting guarantees the safe and stable operation of power systems. Therefore, it is necessary to improve the reliability and the forecasting accuracy of power systems. Reliable power load forecasting can decrease energy consumption and reduce environmental pollution. From different points of view, load forecasting can be divided into different categories. For instance, load forecasting has be classified into short-term, medium-term, and long-term forecasts depending on the forecast horizon [1]. The short term electric load forecasting (STLF) has attracted substantial attentions because of its definitive impact on the daily scheduling and operations of a power utility [2]. So, STLF is the focus of this paper.

In general, load forecasting contains two common types, which are point forecasts and interval predictions. Point forecasts only provide the values of predicted points, but convey no information about the prediction uncertainty [3]. As to interval predictions, a prediction interval (PI) is composed by lower and upper bounds that include a future unknown observation with a certain probability $(1\text{-}\alpha)\%$ named the confidence level [4]. Different from these two types of forecasts, probability density forecasting is able to offer much useful information by constructing probability density functions of forecasting results. Meanwhile, probability density forecasting is advantageous to the prediction accuracy [5]. This types of forecasts can provide estimates of the full probability distributions of the future load demand [6].

To improve the performance of the STLF methods, many studies have been carried out in recent decades. Various hybrid models have been used for load forecasting [7], including wavelet transform and grey model improved by PSO algorithm [8] and some hybrid methods for STLF [9,10]. Considering the complexity and potential nonlinearity of power load, support vector regression (SVR) [11] has been proposed to deal with this problem, which has become one of the most promising and effective techniques due to their attractive features and profound empirical performance in practical applications [12]. As a kernel-based method, SVR is capable of mapping the input data from a low dimensional space to a high-dimensional feature space, which can flexibly convert nonlinear regression into linear regression without assuming particular functional forms [13]. Due to the strong generalization ability of Gaussian kernel function, the SVR method based on Gaussian kernel [14,15] has been widely utilized in the field of power load forecasting.

In order to fully discover kernel-based SVR, Zhou et.al compared the performance of three kernel-based SVR in terms of forecasting accuracy [16]. To determine the ideal kernel, Che and Wang have proposed a multiply kernels model based on a combination selection algorithm for STLF [17]. Authors in [17] showed that the optimal combination is more effective than simple kernel-based SVR models and other multiply kernels combination models. However, these kernel SVR models can't completely measure the uncertainty of future power load, and only provide the accurate point prediction results.

Different from SVR, Quantile regression (QR) is a popular statistical method for estimating the quantiles of a conditional distribution on the values of covariates. QR is capable of explaining the relationships among random variables regardless of the type of the distribution function [18,19]. It is suitable to the problem with multi independent variables. If a probability density function is defined, the any shape of the predictive distribution can be determined by means of the estimated quantiles. Therefore, QR methods have been used in power load and electricity price forecasting in recent years [20-22]. However, the shortcoming of traditional linear QR is the difficulty in solving the complex nonlinear optimization problem. The difficulty in nonlinear QR lies in how to find the appropriate form of nonlinear function [23].

The accuracy of the power load forecasting is also influenced by some other factors, such as, economy, environment, historical data and real-time price. These factors make load forecasting become a complicated task [24]. Particularly, real-time price is one of the uncertain factors for smart grid. Thus, real-time price is considered as an important factor. Real-time price forecasting has become the core process of the power load system at the operational level [25-29]. With the emergence and development of smart grid, people can adjust the electricity expenditures mode according to the electrical load demand and real-time price. In other words, the consumers' activities are likely to be influenced by real-time price in smart grid. From the point of view of consumers, electricity cost can be decreased when the real-time price is referenced [30]. Although there are a few load forecasting methods, which considers the effect of real-time price for power load forecasting models [31-35], it may lose some valuable information for power load forecasting without considering the correlation between real-time price and power load. Renewable energy sources and distributed generation are integrated into power systems, which increases uncertainties in both generation and demand sides. These uncertainties of load forecasting are urgent to be addressed. The probability density forecasting method is considered as a powerful tool to quantify uncertainties associated with forecasts.

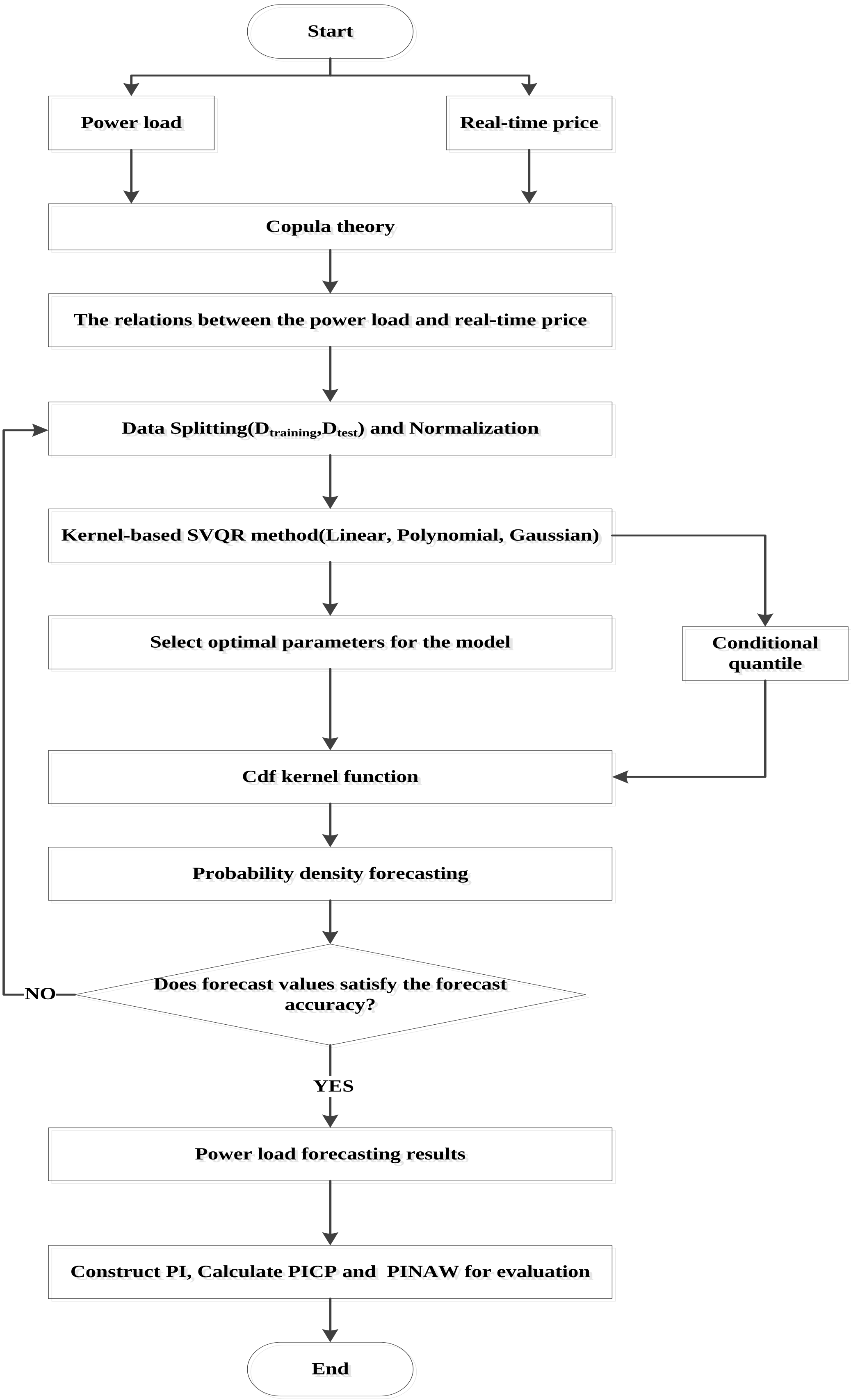
To measure the prediction uncertainty, support vector quantile regression (SVQR) models have been proposed, which incorporate SVR into the QR [36，37] to construct a nonlinear QR method. By applying SVR with a check function instead of an e-insensitive loss function into the QR model, SVQR tends to quantify more uncertain information. It is estimated by settling a Lagrangian dual problem of quadratic programming. In practical applications, SVQR provides an effective way to acquire the nonlinear QR structure by introducing a kernel function, which has shown good performance of estimating multi-period value at risk [38-40].

SVQR methods are based on kernel functions. Their better prediction performance is demonstrated based on the selection of an appropriate kernel function that fits the learning target. However, current SVQR methods only adopt the Gaussian kernel function and do not consider the performance of other kernel functions. This paper proposes a kernel-based support vector quantile regression (KSVQR) model, which chooses most appropriate kernel function from three commonly used kernels, namely linear, polynomial and Gaussian kernels. The proposed KSVQR method is applied to the probability density forecasting, which can generate complete probability distribution of the future value. The real value of power load and real-time price are employed to probability density forecasting based on Copula theory, which is used to analyze the correlation between power load and real-time price.

The contributions of this article include: 1) This paper proposes KSVQR model, which compares different kernel functions and selects the optimal kernel function for power load probability density forecasting. The proposed KSVQR model can quantify the uncertainty between power load and real-price, and provide much useful information than existing kernel SVR models. 2) Two criteria of interval prediction are adopted to evaluate the performance of probability density forecasting method considering real-time price, namely, PI coverage probability (PICP) and PI normalized average (PINAW). 3) A short-term power load probability density forecasting method based on Copula theory is presented to verify the importance of real-price in the smart grid. The *t*-Copula function is adopted to draw the relational diagram and explain the nonlinear correlation between the power load and real-time price. 4) The accuracy of power load forecasting is assessed by three cases of Singapore. Moreover, the comparison of prediction results from SVQR, SVR and Back propagation（BP) method exhibits that the proposed method can achieve better prediction performance. The main structure of this paper is shown in Fig.1, which provides the main ideas so the reader can see a roadmap before moving on to the rest of the paper.

The organization of this paper is as follows. Section 2 introduces the mathematical formulation of SVR, QR and KSVQR models. This section also introduces probability density prediction based on Copula theory. Section 3 introduces two point forecasting metrics and two PI assessment metrics to measure the errors and uncertainty of the power load forecasting. Practical cases of Singapore are used to evaluate the performance of the proposed KSVQR model in Section 4. Finally, the conclusions and future work are summarized in Section 5.

1



2 Fig.1 The flowchart of the paper structure

2. A kernel-based support vector quantile regression

2.1. SVR method

SVM is proposed by Vapnik which is based on statistical learning theory and structural risk minimization principle [41]. It is called SVR when SVM is applied to regression problems. Given a data set $T=\{x_i, y_i\}_{i=1}^{n}$, where $x_i \in R^d$ and $y_i \in R$, the main aim of SVR is to obtain a regression model which has good forecasting performance on future cases. When the data set T is nonlinearly dependent, $m(x)$ can be regarded as a nonlinear function of the input vector x. In order to solve this problem, a feasible way for $m(x)$ estimation is to perform the locally polynomial regression in parametric form. To implement the nonlinear mean regression, the paper projects the input vector x into a higher dimensional feature space using a nonlinear mapping function $\varphi(\cdot)$, which is defined by a kernel function. $m(x)$ of SVR can be obtained by the following linear functional form.

$$f(x)=m(x)=w\cdot\phi(x)+b \tag{1}$$

where w is the weight vector, b represents the threshold and $\phi(\cdot)$ denotes a nonlinear mapping function. The optimal parameter (w,b) of the model can be solved by the following formula.

$$\min_{\omega,b} \tfrac{1}{2}\|\omega\|^2 + C\sum_{t=1}^{k}|y_t - f(x_t)| \tag{2}$$

where C denotes penalty parameter, k is sample size.

2.2. QR method

QR is introduced to replace the classical mean regression [19]. It provides a comprehensive strategy for the entire conditional distribution of a response variable y when x is a explanatory variable instead of the conditional mean only. The idea behind QR can be ascended to the loss functions in advance. The check function (the asymmetric loss function) was proposed by Koenker [18]. It can obtain the optimal parameters through the check function minimization. The check function $\rho_\tau(\mu)$ is defined as follows.

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

The indicator function $I(\mu)$ is shown as following.

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

where $\tau \in (0,1)$ is able to generate quantiles.

2.3.KSVQR model

The regression variables of QR are capable of being adopted to provide information for estimating conditional quantile of response variables. It has been found that explanatory variables have influence on the response variable under different quantiles [18]. However, QR is based on the linear regression, which is difficult to solve complex nonlinear problems. Takeuchi and Furuhashi [36] first utilized SVR to study quantile regression problems, which are better in solving the nonlinear structure of the economic system and the heterogeneity of economic behavior. Li et al [37] proposed a SVQR method and deduced a formula for the effective dimension of the proposed model, which allows suitable selection of the hyperparameters. Shim introduced a semiparametric method which combines SVR with QR to construct SVQR model [39]. Without loss of generality, $u_t = x_t = (h_t, p_t)^T$ are used as input variables, in which h_t represents historical load, and p_t stands for real-time price. Therefore, SVQR model can be obtained by applying a check function of QR in formula (3) instead of penalty function in formula (2) as follows.

$$\min_{\omega_\tau, b_\tau} \tfrac{1}{2}\left\|\omega_\tau\right\|^2 + C\sum_{t=1}^{k} \rho_\tau (y_t - b_\tau - \beta_\tau^T u_t - \omega_\tau^T \phi(x_t)) \tag{5}$$

where C denotes penalty parameters and $\phi(\cdot)$ denotes the nonlinear mapping function. We can rewrite (5) to the quadratic programming by formulation as follows.

$$\min \frac{1}{2} w^T w + C\sum_{t=1}^{k} \left(\tau\xi_t + (1-\tau)\xi_t^*\right) \tag{6}$$

$$\text{s.t.} \begin{cases} y_t - b - \beta^T z_t - w^T \phi(x_t) \le \xi_t \\ -y_t + b + \beta^T z_t + w^T \phi(x_t) \le \xi_t^* \\ \xi, \xi_t^* \ge 0 \end{cases} \tag{7}$$

To solve the optimization problem, the slack variables are introduced to construct the Lagrange function, and the estimators of SVQR are calculated by the following equation:

$$
\begin{cases}
\omega_\tau = \sum_{t=1}^{k}(\alpha_t - \alpha_t^*)\phi(x_t) \\
(b_\tau, \beta_\tau)^T = (U^T U)^{-1} U^T (y - K(\alpha - \alpha^*)) \\
Q_{y_t}(\tau | u_t, x_t, \eta) = b_\tau + \beta_\tau{}^T u_t + K_t(\alpha - \alpha^*)
\end{cases}
\tag{8}
$$

The meaning of the above parameters are expressed respectively as: k is sample size, α, α^* denote the optimal Lagrange multipliers. Following principles in SVM, the index set of support vectors are acquired in the SVQR model. $I_{SV} = \{t = 1,2,\cdots k | 0 < \alpha_t < \tau C, 0 < \alpha_t^* < (1-\tau)C\}$ is obtained by exploiting Karush-Kuhn-Tucker conditions [42], U is the matrix consisting of $U = (1, u_t^T)$, $y = \{y_t | t \in I_{SV}\}$. K_t is a kernel function in the input space ,which is equal to the inner product of vector x_s and x_t in the feature space, that is $K(x_s, x_t) = \phi(x_s)^T \phi(x_t), t \in I_{sv}$ （$s = 1,2,\cdots T$）.

How to select regularization parameter C and the parameters of kernel function play a vital role in the performance of the KSVQR approach. The researchers need to select in advance the type of kernel function and the associated parameters for KSVQR. In this study, we choose three types of kernel functions, namely, linear, polynomial, and Gaussian kernels, which are commonly employed in the related area [17]. The linear kernel is

$$K(x,z) = x^T z, \tag{9}$$

the expression of polynomial kernel is

$$K(x,z) = \left(x^T z + c\right)^d \tag{10}$$

and the formula of Gaussian kernel is

$$K(x,z) = \exp\left(\frac{-(x-z)^2}{2 \times \delta^2}\right) \tag{11}$$

where c is the offset of polynomial, d is the degree of the polynomial kernel, δ is the width of Gauss kernel. The selection of the parameters of the kernel function presents a considerable impact on performance of the KSVQR approach.

2.4.Probability density prediction based on Copula theory

Considering uncertainty factors and the correlation of input variables are of great importance for accurate power load prediction. The renewable energy power generation such as wind power has strong randomness and volatility, which brings more uncertainty for the planning and operation of power systems. The randomness and volatility are brought by output power of wind and photovoltaic units in micro grids (MGs) [30]. MGs have the indispensable infrastructure in smart grid, which consist of distributed energy resources, customers and energy storage units [43]. In the environment of smart grid, renewable energy sources and distributed power are applied to MGs, which become an indispensable important segment in the development of smart grid. However, the volatility and intermittency of renewable energy have significant impact on electricity market under the real-time price guidance, which increases uncertainties in both generation and demand sides [44,45]. In addition, there exists correlation between many random variables in the power system. If the correlation of diverse factors is ignored, it may cause calculation error, which can have a direct effect on the safety of power system and economic operation. Hence, in order to gain accurate prediction results, power load forecasting should consider correlation factors. Copula theory is introduced to build correlation of input variables probability density prediction. The theory not only describes the correlation between input variables in detail, but also has a certain influence for power load forecasting results.

Copula theory is proposed firstly by Sklar [46], which describes accurately the correlation of input variables for nonlinear and asymmetric variable analysis. It is flexible and important to analyze the tail correlation between input variables which is based on Copula theory. In this article, the t - Copula function is used to describe the correlation between the input variables. Power load and real-time price are the input of the Copula function, which provides the correlation diagram of input variables. Gaussian kernel function is adopted to conduct probability density prediction. The kernel density estimation is defined as follows:

$X_1, X_2, \cdots X_n$ are taken from one-dimensional continuous total samples, and the kernel density estimation of the overall density function $f(x)$ at any point x is defined as:

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - X_i}{h}) \tag{12}$$

where $K(\ )$ denotes kernel function, h is bandwidth, Gaussian kernel function is adopted as the kernel density estimation function.

N-dimensional t - copula density function is defined as follows:

$$c\left(u_1,u_2,\cdots u_N;\rho,k\right)=\left|\rho\right|^{-\frac{1}{2}}\frac{\Gamma\left(\frac{k+N}{2}\right)\left[\Gamma\left(\frac{k}{2}\right)\right]^{N-1}}{\left[\Gamma\left(\frac{k+1}{2}\right)\right]^{N}}\frac{\left(1+\frac{1}{k}\zeta'\rho^{-1}\zeta\right)^{-\frac{k+N}{2}}}{\prod_{i=1}^{N}\left(1+\frac{\zeta_i^2}{k}\right)^{\frac{k+1}{2}}} \tag{13}$$

where the ρ is N order symmetric positive definite matrix for all the elements of 1 on the diagonal, $\left|\rho\right|$ denotes determinant of square matrix ρ . k indicates the degrees of freedom. $\zeta'=\left[t_k^{-1}(u_1),t_k^{-1}(u_2),\cdots t_k^{-1}(u_N)\right]$, in which t_k^{-1} denotes inverse function of one-dimensional t distribution when the degree of freedom is k. $\Gamma(\cdot)$ indicates a Gamma function. u_i $\left(i=1,2\cdots,N\right)$ is input variable.

In addition, Pearson correlation coefficient is considered as measurement index for linear correlation of random variable. Fig.2 represents clearly the structure of the KSVQR model.

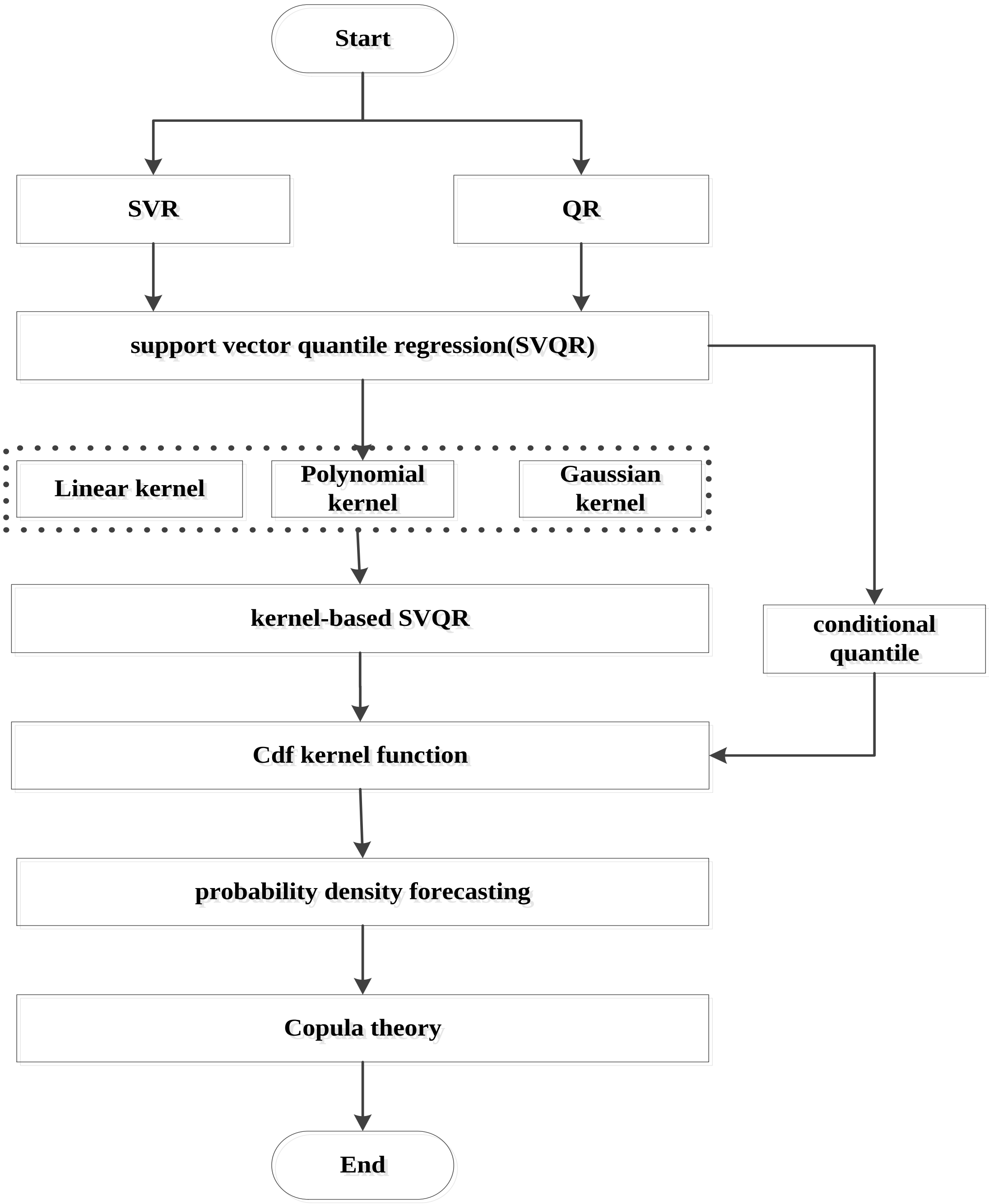


Fig. 2 The flowchart of KSVQR probability density forecasting model

3. Evaluation metrics

3.1.Evaluating the prediction error

Many measures have been proposed to evaluate the errors of the power load forecasting for point prediction, including MAPE (mean absolute percentage error) and the MAE (mean absolute error). MAPE and MAE are defined as follows:

$$MAPE = \frac{1}{n}\left(\sum_{i=1}^{n}\left|\frac{P_i - L_i}{L_i}\right|\right)\times 100\%, i = 1, 2, ..., n \tag{14}$$

$$MAE = \frac{1}{n}\left(\sum_{i=1}^{n}\left|P_i - L_i\right|\right) \tag{15}$$

where i denotes the hour and n indicates the total number of the hour over forecasting period. P_i and L_i represent the *i*-th predicted value and actual value, respectively.

3.2.PICP (PI coverage probability) criterion

PICP (PI coverage probability) and PINAW (PI normalized average width) are usually considered as the criterion for assessing the accuracy of the prediction interval. PICP is defined as the cardinal feature of the PIs (prediction intervals), which demonstrates the percentage of targets that will be covered by the upper and lower bounds. A larger PICP means more targets are located in the constructed PIs. PICP is defined as follows [47]:

$$PICP = \frac{1}{N}\sum_{i=1}^{N} c_i \tag{16}$$

in which N is the total number of predictions and c_i is a Boolean variable, which demonstrates the coverage of PIs. From the perspective of mathematics, c_i is defined as follows:

$$c_i = \begin{cases} 1, & if \ \ y_i \in [L_i, U_i]; \\ 0, & if \ \ y_i \notin [L_i, U_i]. \end{cases} \tag{17}$$

where L_i and U_i are the lower and upper bounds of target y_i, respectively. To obtain effective PIs, PICP needs to be more than the confidence level of PIs. Otherwise, PIs are invalid and unreliable . The ideal value of PICP is equal to100%, which indicates all the target values are covered completely (namely, 100% coverage).

3.3.PINAW (PI normalized average width) criterion

To evaluate the quality of PIs, researchers are more interested in PICP rather than the width of PIs [48]. On one hand, if the width of the interval is large enough, the request for high PICP can be easily met. However, on the other hand, too wide intervals transmit little information about the target values, which are useless for decision making. Width of PIs should be as small as possible, and determines the informativeness of PIs. In the literature, PINAW has been introduced, which is an important quantitative measure. PINAW is defined as follows [3]:

$$PINAW = \frac{1}{NR}\sum_{i=1}^{N}\left(U_i - L_i\right) \quad (18)$$

in which R denotes the maximum minus minimum of the target values. The aim of using R is the normalization of the PI average width in percentage. Thus, PINAW can be adopted for performance comparisons.

4. Case studies

In this section, a comprehensive experimental analysis is given. Three KSVQR models are compared with BP and SVR. The software MATLAB 7.14 is used for all the models. All the programs were run on a 3.20-GHz-based Intel dual-core processor (i5-3470) with 4 GB of random access memory.

The real-world datasets are used from Singapore network [49] to demonstrate the effectiveness of the proposed KSVQR model. It has a tropical rainforest climate with no distinctive seasons, and uniform temperature and pressure. Throughout the year, the climate is hot and humid, with temperatures in the range of 23 to 32 ℃. So, the power load mode in Singapore is fixed on account of the climate and regional reasons. In addition, Singapore lacks of land and resources. Hydropower and wind power generation are infeasible, and solar energy can not be utilized in the large area widely. At present, 80% power load of Singapore comes from power generation of the natural gas. There are also a small amount of photovoltaic power stations which incorporate into power grid. With the development of opening up policy in Singapore electricity market, most of customers demand more competitive price. Under the condition of the environment and social background of Singapore, the power load forecasting is affected by electricity price, which is the main factor. So, real-time price is considered as an important factor for Singapore power load forecasting. The power load and real-time price data of Singapore are chosen as the input variables for the proposed model.

In case studies, section 4.1 carries out a correlation analysis between power load and real-time price based on the Copula theory. Real-time price is proved to influence the consumption of electricity. In section 4.2, a small size dataset and a medium size dataset in 2014 are selected to compare with existing point prediction methods. To further evaluate the generalization capability of SVQR, a smaller training sample and longer test sample example in 2016 is chosen in section 4.3. Section 4.4 summarizes the observations for the experiments.

4.1.The correlation analysis between power load and real-time price based on the Copula theory

Many uncertainty factors exist in power systems, such as equipment failure, power load fluctuations and so on. Especially, the rapid development of the smart grid and wide application of renewable energy increase uncertainty, which bring a certain impact on the operation and control of power systems. These uncertainty factors are considered as input variables to influence the power load forecasting. However, the correlation of these input variables has usually been ignored, which may cause calculation error and even have a direct effect on safe and economic operation of the power system. The section mainly analyzes the correlation between power load and real-time price based on the copula theory. This theory can deal with the correlation between random variables of normal and non-normal distributions.

This paper adopts Jarque-Bera (J-B) test, Kolmogorov-Smirnov test (K-S) and Lilliefors (L) test for real-time price and power load normality test, respectively. The level of significance is set as 0.01. The real-time price and power load datasets are chosen from November 2014 to December 2014 in Singapore, with 48 points in each day. The results of the case show that the h value of three kinds of the test is 1.The p-value results are summarized in Table 1, which are all smaller than 0.01. Therefore, real-time price and power load do not follow the normal distribution. By the calculation, the value of Pearson linear correlation coefficient is 0.4112, which shows that the linear correlation of real-time price and power load is not significant. Copula theory is applicable to any distribution. The paper adopts t - copula function to explain the correlation between real-time price and power load.

Correlation analysis has drawn more and more attentions in many fields. Fig.3 gives Binary t - Copula density function of correlation diagram between real-time price and power load. U is real-time price, and V represents power load. It clearly demonstrates the distributed situation and heavy tail of real-time price and power load. In addition, it shows strong correlation between the point (0, 0) and (1, 1) in the tail. In other words, power load consumption has a great influence on the customers when real-time price suddenly becomes high or low. The fluctuation of real-time price is able to cause the change of the power load. Similarly, the change of the power load tends to cause real-time price fluctuations.

Table 1

The results of J-B, K-S and L test

	p-value of real-time price			*p*-value of power load		
	J-B test	K-S test	L test	J-B test	K-S test	L test
For the case of Singapore	1.0000e-03	1.5810e-170	1.0000e-03	1.0000e-03	4.8325e-18	1.0000e-03

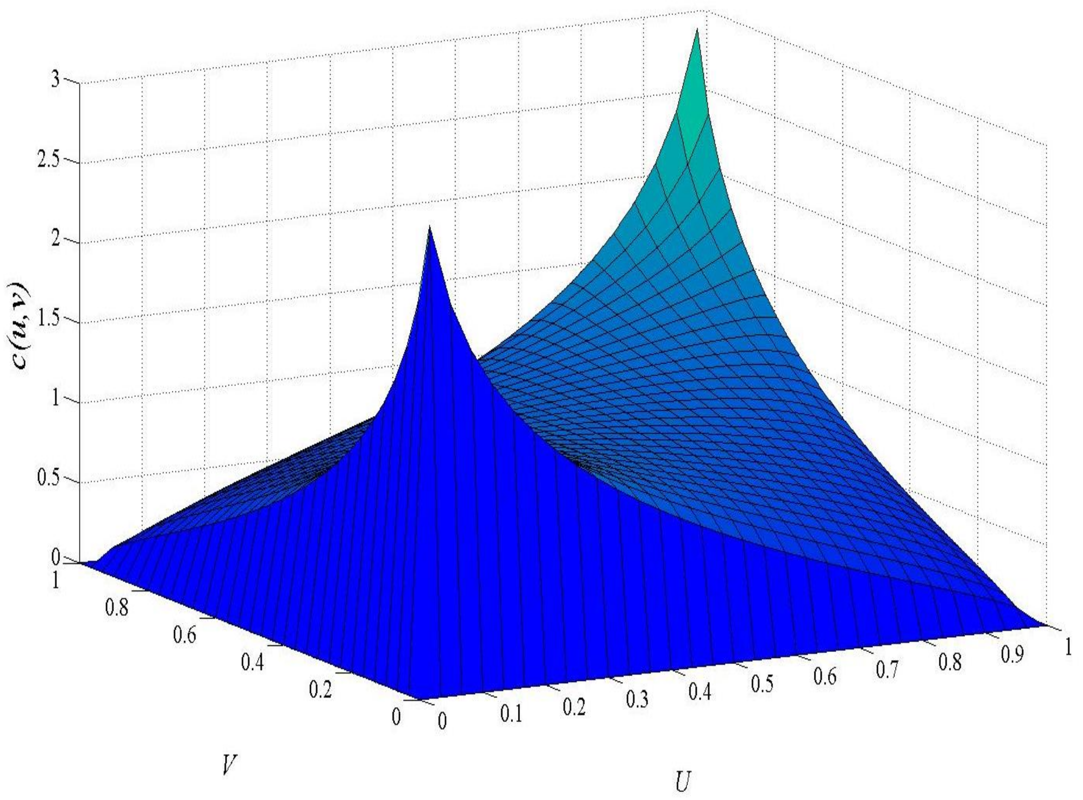


Fig. 3 Binary *t* - Copula density function of correlation diagram between real-time price and power load for the case of Singapore

4.2.Empirical results and analysis of Singapore in 2014

In this subsection, the practical historical load and real-time price data of Singapore include a small size dataset and a medium size dataset in the winter of 2014. Namely, a small size dataset from November, 2014 is selected to predict the future load of 1 day, and a medium size dataset from December, 2014 is selected to predict the future load of 4 days. Without considering real-time price, the power load dataset from November 8, 2014 to November 19, 2014 are selected as training datasets, and the power load dataset on November 20, 2014 is chosen as testing data subset. Under the condition of considering real-time price, the power load and real-time price datasets from November 8, 2014 to November 19, 2014 are chosen as training data subset, and the power load and real-time price datasets on November 20, 2014 are chosen as testing data subset. The case forecasts the power load of 1 day with 48 load points.

Moreover, in order to further demonstrate the satisfactory performance of the proposed KSVQR model, the another medium size dataset of Singapore is selected. Without considering real-time price, the power load datasets from December 1, 2014 to December 26, 2014 are selected as training datasets. The power load from December 27, 2014 to January 30, 2014 are chosen as testing data set. Under the condition of considering real-time price, the power load and real-time price datasets from December 1, 2014 to December 26, 2014 are chosen as training data set, the power load from December 27, 2014 to December 30, 2014 is chosen as testing dataset. This case predicts power load for 4 days, and there are 48 points for a day in both conditions. All samples are normalized in advance

The empirical results, comparison and discussion among these cases are demonstrated to verify our research. We choose three different kernel-based SVQR models in our comparative studies. The two cases select 20 quantiles with the interval of 0.05, and the quantile is from 0.01 to 0.96. All the parameters for the three different kernel functions of two cases are shown in Table 2, where C is penalty parameter; c, d and δ^2 are the parameters of kernel functions. The selection of penalty parameters and the parameters of the kernel functions have a considerable impact on prediction ability of the KSVQR method [40]. The forecasting errors and time of mode (the highest probability point) in the probability density curve from the datasets in 2014 are summarized in Table 3. Two forecasting error measurements, MAPE and MAE, are employed to verify the forecasting accuracy of the proposed model. For the case of Singapore in November, 2014, Gaussian kernel SVQR obtains the optimal result with considering real-time price, the MAPE and MAE values are 0.81% and 47.20, respectively. For the case of Singapore in December, 2014, Gaussian kernel SVQR also obtains the optimal result with considering real-time price, the MAPE and MAE values are 1.91% and 98.71, respectively. Furthermore, it is easy to find that the calculation time of three different KSVQR methods is similar, and it spends much time to obtain accurate results when the forecasting period is extended to 4 days. The results are reasonable because the principle of SVQR probability density forecasting method is to construct probability density function by means of the kernel density function and point prediction results under different quantiles. The complexity will continually increase with the expansion of forecasting period and sample scale.

In order to further explain the superiority of the proposed method, the PICP and PINAW are considered as the evaluation metrics and the values of the two measures are shown in Table 4. We can discover that the constructed PIs cover the real values in a great percentage. For the case of Singapore in November, 2014, the PICP of the three different kernels SVQR is 100%, regardless of whether real-time price is considered or not, which means that all the targets are covered by PIs. It shows that the perfect results are obtained by KSVQR model when the forecasted period is only one day. On the other hand, the width of PIs for Singapore load is narrow. For the case of Singapore in December, 2014, it has the high PICP values with 99.48% for Gaussian kernel function with considering real-time price. The other two KSVQR models also have high coverage. This also means that the constructed PIs cover the target values with a high probability. However, PINAW value of the Gaussian-based SVQR models is wider than the other two models.

In order to better illustrate the advantages of KSVQR method, this paper compares the forecasting errors of KSVQR, SVR and BP models in Table 5. It shows the prediction errors of three different methods under the condition of considering real-time price and the condition of not considering real-time price. The penalty parameter of SVR is 8000 and the insensitive loss function is 0.001 for 2014.11 datasets of Singapore. For 2014.12 datasets of Singapore, the penalty parameter of SVR is 1000 and the insensitive loss function is 0.1. The iteration number of BP neural network is 1000, and the neural network structure for the two cases are 11-3-1 and 7-1-1, respectively. KSVQR method is superior to the other methods based on the results in both cases of considering or not considering the real-time power price. The proposed KSVQR method shows strong generalization capability by comparing the forecasting results of different methods. Also, Forecasting results of KSVQR method with considering real-time price are better than that of KSVQR method without considering real-time power price. Therefore, real-time price factors should be considered as an important factor for STLF. No matter what kind of method is adopted.

Fig.4 demonstrates that the prediction results of Singapore on November 20, 2014 and prediction intervals based on real-time price and Gaussian kernel SVQR, which show that the actual value always falls in the prediction interval and the mode of predicted results is close to the true value curve. It comes to the conclusion that the proposed method can accurately depict power load fluctuations. Fig.5 shows the prediction results and prediction intervals for the case of Singapore from December 27, 2014 to December 30, 2014 based on real-time price and Gaussian kernel SVQR. It can be seen from the diagram that the actual value almost falls in the prediction interval. This also illustrates the Gaussian kernel SVQR method can better describe power load fluctuations. Fig.6 and Fig.7 give the diagram of probability density curve based on real-time price and Gaussian kernel SVQR on November 20, 2014 and December 30, 2014, respectively. It gives completely probability distribution of future power load and the real value also appears in the density function with high probability, which can explain the advantages of probability density forecasting method in quantifying the uncertainty and improving prediction accuracy. It can be seen from Fig.6 that the rest of the actual values are mostly appear in the probability density curve with the highest probability, in addition to the actual value on 12:00 that appears in the tail of probability density curve. Similarly, it can be seen from Fig.7 that the rest of the actual values arise in the middle of the probability density curve, except for the actual value on 23:30. However, the density curve drawn in Fig 7 is less smooth than the results of Fig. 6.

Table 2

The parameters used in the KSVQR model.

Kernel function type	Parameters value		
For 2014.11 datasets of Singapore			
Linear	C=2		
Polynomial	C=1	c=1	d=1
Gaussian	C=8000	δ^2 =10000	
For 2014.12 datasets of Singapore			
Linear	C=1		

Polynomial	C=1000	c=1	d=1
Gaussian	C=1000	δ^2 =4500	

Table 3

Forecasting errors and time of mode from the datasets in 2014

	Without considering real-time power price			Considering real-time power price		
	MAPE(%)	MAE(MW)	Time(s)	MAPE(%)	MAE(MW)	Time(s)
For the case of Singapore in 2014.11						
Linear SVQR	1.29	75.94	0.97	1.16	67.94	1.42
Polynomial SVQR	1.13	66.73	1.03	0.87	50.63	1.44
Gaussian SVQR	1.10	64.43	1.13	0.81	47.20	1.52
For the case of Singapore in 2014.12						
Linear SVQR	2.28	119.80	141.58	2.06	106.61	150.43
Polynomial SVQR	2.29	120.77	173.95	2.03	105.44	238.91
Gaussian SVQR	2.20	115.51	167.29	1.91	98.71	169.64

Table 4

PI evaluation indices from the datasets in 2014

	Without considering real-time	Considering real-time power

	power price		price	
	PICP(%)	PINAW(%)	PICP(%)	PINAW(%)
For the case of Singapore in 2014.11				
Linear SVQR	100	16.48	100	11.62
Polynomial SVQR	100	15.95	100	11.83
Gaussian SVQR	100	15.76	100	11.75
For the case of Singapore in 2014.12				
Linear SVQR	91.15	23.93	93.23	22.20
Polynomial SVQR	82.81	23.68	93.23	21.11
Gaussian SVQR	96.35	30.65	99.48	29.60

Table 5

Forecasting errors of SVR , SVQR and BP from the datasets in 2014

algorithm	Without considering real-time power price		Considering real-time power price	
	MAPE(%)	MAE(MW)	MAPE(%)	MAE(MW)
For the case of Singapore in 2014.11				
SVR	3.41	200.00	2.97	154.36
SVQR(Gaussian kenel)	1.10	64.43	0.81	47.20
BP	3.65	215.96	3.30	174.80

For the case of Singapore in 2014.12				
SVR	4.22	203.16	3.23	155.79
SVQR(Gaussian kenel)	2.20	115.51	1.91	98.71
BP	4.33	223.68	3.32	167.43

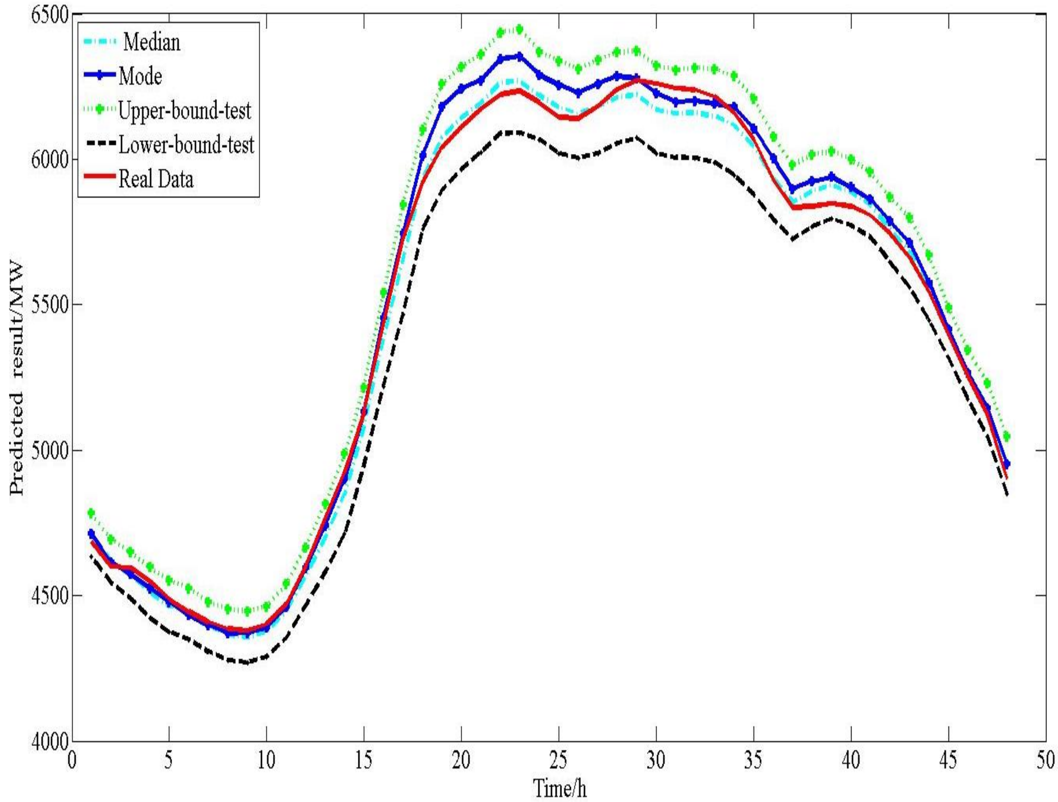


Fig. 4 Prediction results and prediction intervals based on Gaussian kernel SVQR for the case of Singapore on November 20, 2014.

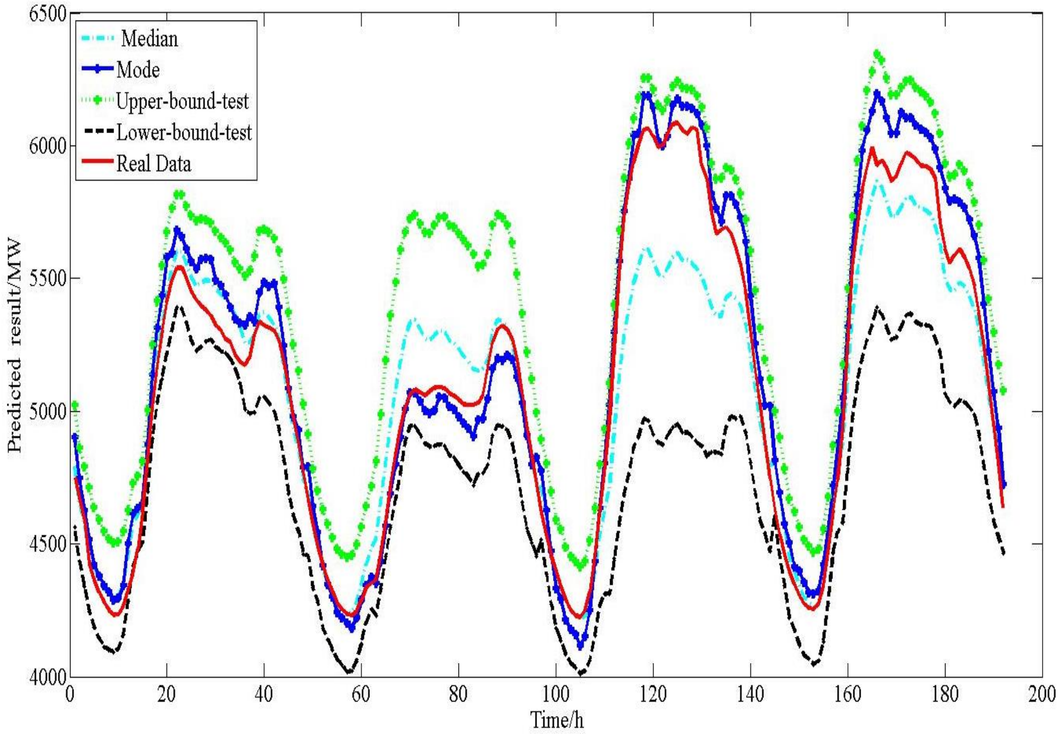


Fig. 5 Prediction results and prediction intervals based on Gaussian kernel SVQR for the case of Singapore from December 27, 2014 to December 30, 2014.

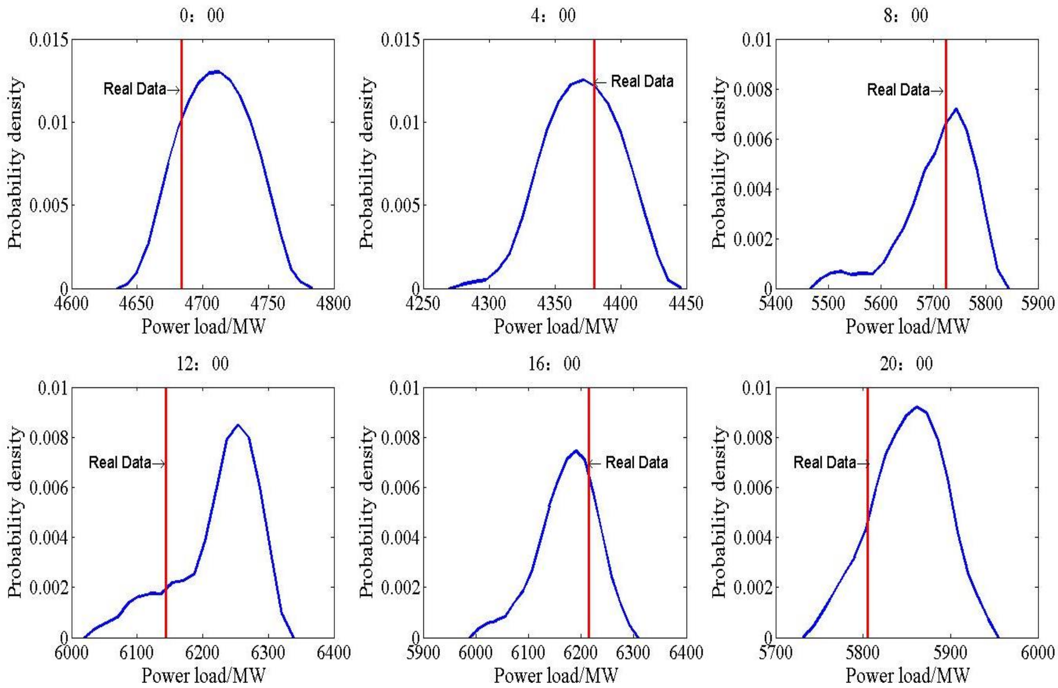


Fig. 6 Diagram of probability density curve based on real-time price and Gaussian kernel SVQR on November 20, 2014.

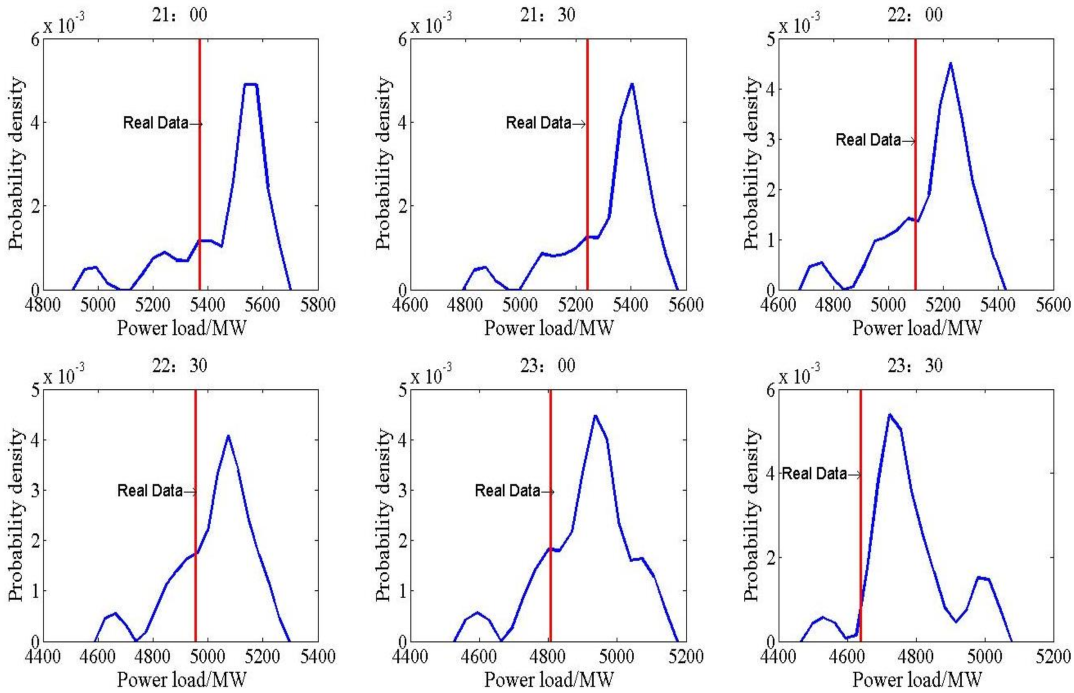


Fig. 7 Diagram of probability density curve based on real-time price and Gaussian kernel SVQR on December 30, 2014 at the last 6 half-hours.

4.3.Empirical results and analysis of Singapore in July, 2016

In order to further evaluate the generalization capability of the proposed method, an real-time dataset of Singapore in the summer of 2016 is adopted to predict the future load of 10 days. The case selects the practical power load and real-time price datasets of Singapore from July 1, 2016 to July 30, 2016 as a smaller training sample and longer test sample example. As compared with the case in December, 2014, the period of training sample is decreased to 20 days, and the period of test sample is increased to 10 days. Under the conditions of considering and not considering real-time price, this dataset is divided into training datasets and testing datasets. Namely, the datasets from July 1, 2016 to July 20, 2016 are selected as training datasets, and the datasets from July 21, 2016 to July 30, 2016 are chosen as testing datasets. Our task is to predict power load for 10 days, and there are 48 points for a day in both conditions. It comes to the conclusion that the total datasets are same as the other cases, but the sample size of training datasets is smaller. Because we have shown that the Gaussian kernel function can lead to the optimal results in our previous two case studies, we also choose Gaussian kernel SVQR for our analysis here. Under the both conditions of considering and not considering real-time price, all parameter values of the KSVQR model are kept the same. The penalty parameter of the KSVQR model is

1($C=1$), the width of Gauss kernel is 20($\delta^2=20$).The experimental results are summarized in Table 6. It can be concluded that the KSVQR model also has stable forecasting results. At the same time, with the decrease of training data and the increase of forecasting period, it can be seen that real-time price has a certain influence on the results of power load forecasting. However, the effect is not obvious. On the other hand, it has similar PICP and PINAW values though the result is a bit better when the real-time power price is considered. Furthermore, the PINAW value of this case is wider than the other two cases.

In order to clearly show the advantages of KSVQR method, Fig. 8 provides prediction results and prediction intervals for the case of Singapore from July 21, 2016 to July 30, 2016 based on Gaussian kernel SVQR, which shows the actual value mostly falls in the prediction interval. Furthermore, the mode of predicted results is closer to the true value curve, which illustrates the KSVQR method accurately depicts power load fluctuations. Fig. 9 shows the probability density curve based on real-time price and Gaussian kernel SVQR on July 30, 2016. It is clear to see that the curve is less smooth than those in Figs. 6 and 7, though the real value also appears in the probability density curve.

Table 6

Experimental results of KSVQR (Gaussian kernel) from the datasets in 2016

	For the case of Singapore in 2016.7			
	MAPE(%)	MAE(MW)	PICP(%)	PINAW(%)
Without considering real-time power price	1.84	102.40	94.79	37.04
Considering real-time power price	1.80	99.83	95.21	36.98

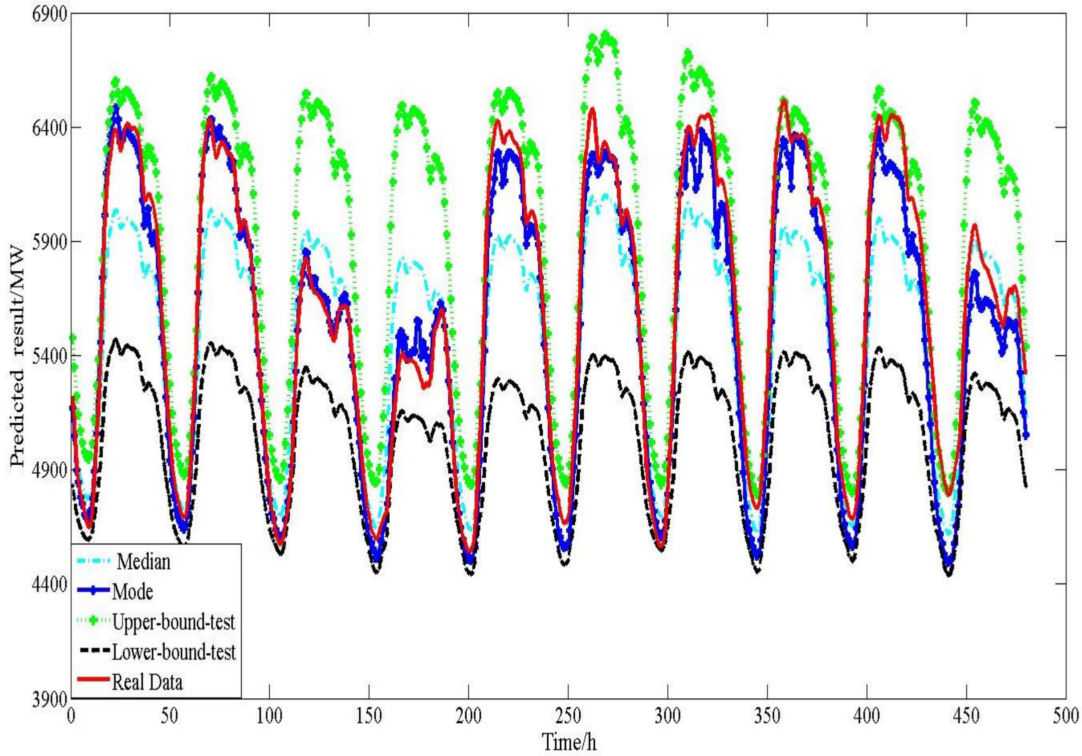


Fig. 8 Prediction results and prediction intervals based on Gaussian kernel SVQR for the case of Singapore from July 21, 2016 to July 30, 2016.

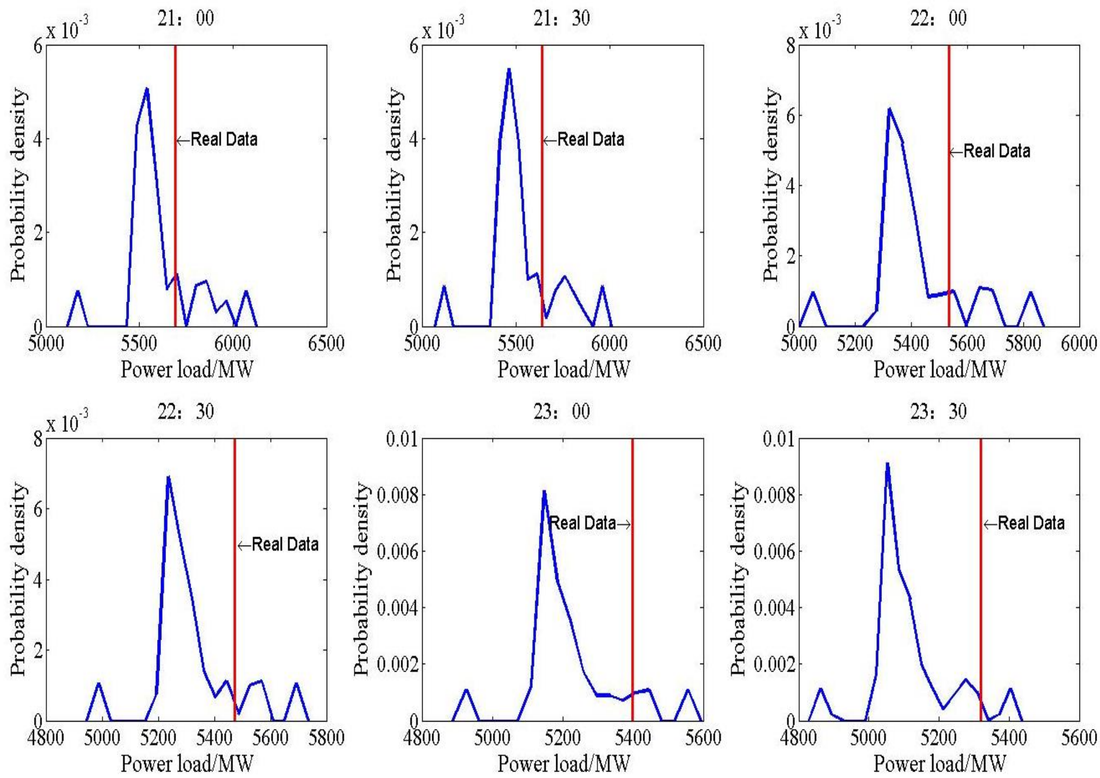


Fig. 9 Diagram of probability density curve based on real-time price and Gaussian kernel SVQR on July 30, 2016 at the last 6 half-hours.

4.4. Observations for the experiments

Base on the above results in the experiment, several observations can be made: (1) The KSVQR probability density forecasting method not only shows better performance than traditional point prediction methods, but also provides much useful information, which is the probability distribution of forecasting results. (2) Gaussian kernel SVQR is more effective than Polynomial SVQR and linear SVQR. It can obtain nearly 100% PICP when the real-time price is considered and the forecasting period is less than 4 days. (3) The real-time price can improve PICP, PINAW, and the forecasting accuracy of all models adopted. However, the improvement may be reduced when the training data are decreased and the forecasting period is increased. (4) The smoothness of obtained probability density curves will reduce when the test data are increased

5 Conclusion and future work

STLF is a fundamental and vital task for operation and controlling of smart grids. The random factors and the penetration of the renewable energies greatly increase the uncertainty of power systems and calculation errors. In order to achieve better prediction results, the paper proposes a short-term power load probability density forecasting method using KSVQR and Copula theory. Since kernel function can effectively approximate any function, the selection of appropriate kernel function is important to the learning target of forecasting model. Therefore, this paper adopts KSVQR method to compares three different kinds of kernel functions and selects the optimal kernel function for the learning target. Copula theory is introduced to provide the correlation analysis between real-time price and power load. As to the probability density forecasting method, the forecasting results under the different quantiles are input into the kernel density estimation function [5]. Hence, we design a multi-stage scheme to construct probability density functions for measuring the uncertainty of electricity load. The advantage of our probability density prediction method is the effectiveness on quantifying the uncertainty, which contributes to the improvement of the forecasting accuracy of the power load.

Based on the analysis of three datasets in Singapore, the three KSVQR methods have excellent forecasting performance, and Gaussian kernel is shown to be the optimal kernel function. Couple function can measure the nonlinear relationship between power load and real-time price. The fluctuation of power load and real-time price affects each other. The prediction accuracy can be improved, the high PICP and narrow PINAW are acquired when the real-time is considered. However, the quality of PICP and PINAW is reduced when forecasting sample scale is expanded. The smoothness of probability density function may also decrease with the increase of forecasting period.

In the future, we will carry out the following research: (1) More information on probability density function will be considered to improve the forecasting potential of SVQR. (2) It is difficult to obtain high PICP and narrow PINAW because a higher PICP may result in a wider PINAW. To improve the reliability of probability density forecasting methods, the appropriate objective function need to be constructed to balance the result between PICP and PINAW. (3) The quality of probability density curve will reduce when the forecasting period and sample scale are expanded. Better bandwidth estimation methods will be investigated (4) The parameters of SVQR are very important for probability density forecasting. Several intelligence optimization algorithms may be applied to tune the parameter of SVQR.

Acknowledgement

This paper is funded by the National Natural Science Foundation (No.71401049)，the Anhui Provincial Natural Science Foundation (No.1408085QG137, 1408085MG136), 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), and the Specialized Research Fund for the Doctoral Program of Higher Education (No. 20130111120015).

References

[1] Fan S, Hyndman R. Short-term load forecasting based on a semi-parametric additive model. IEEE Transa Power Syst 2012;27(1):134-41.

[2] Masa-Bote D, Castillo-Cagigal M, Matallanas E, et al. Improving photovoltaics grid integration through short time forecasting and self-consumption. Appl Energy 2014;125(15):103–13.

[3] Khosravi A, Nahavandi S, Creighton D, Atiya A.F, Comprehensive review of neural network-based prediction intervals and new advances, IEEE Trans. Neural Netw. 2011:22 (9) 1341–356.

[4] Chatfield.C, Calculating interval forecasts, J. Bus. Econ. Stat.1993: 11 (2) 121–135.

[5] He YY, Xu QF, Wan JH, Yang SL, Short-term power load probability density forecasting based on quantile regression neural network and triangle kernel function. Energy 2016;114:498-512

[6] Rob J. Hyndman and Shu Fan, Density Forecasting for Long-Term Peak Electricity Demand , IEEE T Power Syst 2009:1-12.

[7] Xiao L, Wang J, Hou R, Wu J A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting. Energy 2015:524-49.

[8] Bahrami S, Hooshmand R, Parastegari M Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm. Energy 2014:434-42.

[9] Ghofrani M, Ghayekhloo M, Arabali A, Ghayekhloo A A hybrid short-term load forecasting with a new input selection framework. Energy 2015:777-86.

[10] Ghayekhloo M , Menhaj M.B, Ghofrani M, A hybrid short-term load forecasting with a new data preprocessing framework. Electr Pow Sys Ters 2015:138-48.

[11] Jain RK, Smith KM, Culligan PJ, Taylor JE. Forecasting energy consumption of multi-family residential buildings using support vector regression: investigating the impact of temporal and spatial monitoring granularity on performance accuracy. Appl Energy 2014;123:168–78.

[12] Chen K, Yu J. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. Appl Energy 2014;113:690–705.

[13] Wu, S., & Akbarov, A. Support vector regression for warranty claim forecasting. Eur J Oper Res, 2011; 213:196–204.

[14] Che J, Wang J, Wang G. An adaptive fuzzy combination model based on selforganizing map and support vector regression for electric load forecasting. Energy 2012;37(1):657-64.

[15] Wang J, Li L, Niu D, Tan Z. An annual load forecasting model based on support vector regression with differential evolution algorithm. Appl Energy 2012,94:65–70.

[16] Zhou J, Shi J, Li G. Fine tuning support vector machines for short-term wind speed forecasting. Energy Convers Manage 2011;52:1990–9.

[17] Che J, Wang J. Short-term load forecasting using a kernel-based support vector regression combination model, Appl Energy 2014:602-9.

[18] Koenker R, Bassett, G. W. Regression quantiles. Econometrica, 1978: 33–50.

[19] Koenker, R. Quantile regression. New York: Cambridge University Press.2005.

[20] Hagfors LI, Bunn D, Kristoffersen E, Staver TT, Westgaard S, Modeling the UK electricity price distributions using quantile regression. Energy. 2016; 102: 231-43.

[21] Bunn D, Andresen A, Chen D, Westgaard S. Analysis and forecasting of electricity price risks with quantile factor models. Energy J 2016;37(1):101-22.

[22] Maciejowska K , Nowotarski J, Weron R, Probabilistic forecasting of electricity spot prices using Factor Quantile Regression Averaging. Int J Forecasting. 2016;32: 957-65.

[23] Wang J Bayesian quantile regression for parametric nonlinear mixed effects models. Stat Methods Appl 2012:279–95.

[24] Elattar EE, Goulermas J, Wu QH. Generalized locally weighted GMDH for short term load forecasting. IEEE Trans Syst Man Cybern 2012;42(3):345-56.

[25] Lin WM, Gow HJ, Tsai MT. An enhanced radial basis function network for short-term electricity price forecasting. Appl Energy 2010; 87(10):3226–34.

[26] Carolina GM, Julio R, Maria JS. Modelling and forecasting fossil fuels, CO2 and electricity prices and their volatilities. Appl Energy 2013;101:363–75.

[27] David P. C , Daniel R, Aggregate modeling of fast-acting demand response and control under real-time pricing .Appl Energy 2016;181:288-98.

[28] Tan Z, Zhang J, Wang J, Xu J. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. Appl Energy 2010;87:3606–10

[29] Khosravi A， Nahavandi S, Creighton D，Quantifying uncertainties of neural network-based electricity price forecasts. Appl Energy 2013;112：120-9

[30] Niknam T, Azizipanah-Abarghooee R, Narimani M R, An efficient scenario-based stochastic programming framework for multi-objective optimal micro-grid operation. Appl Energy. 2012; 99:455-70.

[31] Yun Z, Quan Z, Caixin S, Shaolan L, Yuming L, Yang S. RBF neural network and NFIS-based short-term load forecasting approach in real-time price environment. IEEE Trans Power Syst 2008; 23(3):853–8.

[32] Khotanzad A, Zhou E, Elragal H. A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment. IEEE Trans Power Syst 2002;17(4):1273–82.

[33] Amir M, Hamidreza Z, William DR. Electricity price and demand forecasting in smart grids. IEEE Trans Smart Grid 2012;3(2):664–74.

[34] Shayeghi H, Ghasemi A, Moradzadeh M, Nooshyar M. Simultaneous day-ahead forecasting of electricity price and load in smart grids. Energy Convers Manage 2015:371-84

[35] Che J, Wang J. Short-term electricity prices forecasting based on support vector regression and auto-regressive integrated moving average modeling. Energy Convers Manage 2010;51(10):1911–7.

[36] Takeuchi I, Le QV, Sears TD, Smola AJ Nonparametric quantile estimation. J Mach Learn Res 2006:1231–64.

[37] Li Y, Liu Y, Zhu J Quantile regression in reproducing kernel Hilbert spaces. J Am Stat Assoc 2007:255–68.

[38] Shim J, Hwang C Support vector censored quantile regression under random censoring. Comput Stat Data Anal 2009;53:912–9.

[39] Shim, J, Kim, Y, Lee, J, & Hwang, C. Estimating value at risk with semiparametric support vector quantile regression. Comput Stat, 2012; 27: 685–700.

[40] Xu QF, Zhang JX, Jiang CX, Huang X, He YY. Weighted quantile regression via support vector machine. Expert Syst Appl, 2015:5441-51.

[41] Cristianini Nello, Shawe-Taylor John, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Publishing House of Electronics Industy, 2004.

[42] HW Kuhn, Tucker AW Nonlinear programming. In: Proceedings of 2nd Berkeley symposium. University of California Press, Berkeley, 1951: 481–92.

[43] Liu N , Tang QF, Zhang JH, Fan W, Liu J, A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids Appl Energy. 2014;129: 336-45.

[44] Wang J， Botterud A， Bessa R, Keko H, Carvalho L, Issicaba D, Sumaili J. Miranda V， Wind power forecasting uncertainty and unit commitment. Appl Energy.2011; 88: 4014-23.

[45] Ketan P, Optimal autonomous microgrid operation: A holistic view. Appl Energy.2016; 173: 320-30.

[46] Sklar A. Fonctions de répartition n dimensions et leurs marges. Publication de l Institut de Statistique de L Université de Paris 1959:229-31.

[47] Khosravi A, Nahavandi S, Creighton D, Construction of optimal prediction intervals for load forecasting problems IEEE Trans Power Syst., 2010; 25(3) 1496–503

[48] Khosravi A, Nahavandi S, Creighton D, Prediction interval construction and optimization for adaptive neurofuzzy inference systems, IEEE Trans Fuzzy Syst. 2011; 19 (5) 983–8.

[49] Price information; 2014 [Online; accessed 27-March-2014], https://www.emcsg.com/MarketData/PriceInformation#priceDataView.