Available online at www.sciencedirect.com
ScienceDirect
Energy Reports 8 (2022) 1386–1397
www.elsevier.com/locate/egyr

2021 The 2nd International Conference on Power Engineering (ICPE 2021), December 09–11, 2021, Nanning, Guangxi, China

Short-term microgrid load probability density forecasting method based on *k*-means-deep learning quantile regression

Zilong Zhao, Jinrui Tang*, Jianchao Liu, Ganheng Ge, Binyu Xiong, Yang Li

School of Automation, Wuhan University of Technology, Wuhan, Hubei 430070, China

Received 27 February 2022; accepted 12 March 2022
Available online 25 March 2022

Abstract

Traditional short-term load forecasting (STLF) methods for large utility grid systems usually provide the forecasted load with deterministic points. However, deterministic load forecasting cannot reveal the load pattern and uncertainty of controllable load in a microgrid, where the prediction errors may exceed the expected range due to the high volatility and strong randomness. In order to deal with this matter, a probability density forecasting method is proposed to predict the microgrid load with uncertainty for robust power scheduling in this paper. The proposed probability forecasting method effectively combines several data-driven and statistical algorithms, including the *k*-means algorithm, quantile regression long short-term memory neural network (QRLSTM), and kernel density estimation (KDE). Firstly, similar days related to the prediction day are selected through the *k*-means algorithm, and the historical load data of these selected days are divided into two subsets including the training dataset and the testing dataset. Secondly, a QRLSTM-based model is established and used to predict the microgrid load for different quantiles. Finally, the probability density function of the predicted points is obtained by KDE on the target day. The prediction accuracy is evaluated roundly and the results demonstrate that the proposed method can effectively reproduce the probability density distribution of the load and provide noticeably better performance than some benchmark methods.

Peer-review under responsibility of the scientific committee of the 2021 The 2nd International Conference on Power Engineering, ICPE, 2021.

Keywords: Short-term load forecasting; Probability density; Quantile regression; Long short-term memory neural network; Kernel density estimation; Microgrid

1. Introduction

Short-term load forecasting (STLF) plays a vital role in power system operation, and the accuracy of STLF results will affect the security, stability, and economy of the power systems [1,2]. As an essential part of electricity generation scheduling, the STLF can be used for balancing the power supply and load demand, serving as a basis for energy dispatch and management [3]. Accurate STLF results can also minimize electricity waste and prevent energy crises [4]. The advanced control strategies and economic dispatch algorithms always rely on accurate and

* Corresponding author.
E-mail address: tangjinrui@whut.edu.cn (J. Tang).

https://doi.org/10.1016/j.egyr.2022.03.117

Peer-review under responsibility of the scientific committee of the 2021 The 2nd International Conference on Power Engineering, ICPE, 2021.

reliable STLF methods, but many factors like weather, economic, and social affect the STLF prediction results in microgrid [5,6]. Therefore, it is challenging to achieve accurate STLF results, especially in a microgrid environment, which motivates our investigation in this paper.

Traditional STLF methods can be generally divided into two groups, including the deterministic point forecasts [7] and the probability forecasts [8]. In terms of deterministic point forecasts, substantial progress has been made, and various intelligent methods have emerged to pursue high accuracy. These including multiple linear regression [9], support vector machine (SVM) [10], autoregressive integrated moving average (ARIMA) [11], artificial neural network [12], clustering methods [13,14], and hybrid methods [15,16]. These essential methods have been widely applied in system-level load forecasting applications and achieved accurate prediction results. Nevertheless, the microgrid load is more difficult to forecast than a regional system due to the high randomness and lower similarities in its historical load curves [17]. In addition, the drastic load fluctuation leads to low-precision prediction due to the limited load capacities in a microgrid. A detailed description of commonly-used indicators to evaluate the accuracy of the prediction in microgrids is given in [18], including the mean absolute percentage errors (MAPE) and the relative mean square errors (RMSE). In [19], several forecasting steps including similar days loads selection, load reconfiguration through empirical mode decomposition, and building prediction model by long–short memory (LSTM), with which the MAPE is only 1.59% for a regional system. In [20], the backpropagation (BP) neural network is used to predict microgrid load several days ahead. However, the maximum prediction error of MAPE can reach high up to 24.4%, leading to the requirement of further improvement in the forecasting accuracy.

Based on the performance of prediction results, probabilistic forecasting consists of quantile forecasting, prediction interval forecasting, and probability density function forecasting [21]. Probability density forecasting can effectively illustrate the load uncertainties by using conditional probability density curves [22]. Various advanced probability load forecasting methods were proposed and compared with each other in 2014 [23]. Furthermore, a hybrid model is proposed and kernel density estimation (KDE) and quantile regression (QR) is used in the method [24]. The QR can obtain the complete distribution relationship between the explained and the explanatory variables by estimating the different conditional quantiles, while the KDE can bridge the prediction results based on the quantile and the probability density curve. Using the quantile, He et al. [25] developed a diverse quantile regression neural network, which replace the single-core with multi-core in the master–slave model, trained in different cores to be fully utilized in multi-core CPU. Wang et al. [26] replaced the regular mean square error loss function with a pinball loss function in LSTM neural network, the results show that this improved method performs better. However, related research of short-term load probability density forecasting is scarce in a microgrid. The prediction results accuracy varies substantially in microgrids with diverse capacities. In [27], a sparse heteroscedastic model is proposed to achieve the day-ahead probabilistic system-level load forecasting results. By contrast, a stochastic model for microgrid load forecasting is proposed in [28], but the load features are not taken into account in the constructed model. Therefore, due to its smaller capacity, higher volatility, and higher randomness, the microgrid load is more challenging to forecast than in a large power grid. A combination of the clustering method and probability load forecast method can potentially be used to reduce the load forecasting error in a microgrid and for analyzing the relationship between forecasting accuracy with load characteristics.

With these in mind, a novel method named *k*-means-deep learning quantile regression LSTM (K-QRLSTM) is proposed to provide the day-ahead load forecast in a microgrid. First, the multivariate time load series are built using the datasets of the Smart Grid Smart City (SGSC) project in Australia. Second, the *k*-means algorithm is adopted to find similar days related to the prediction day. The selected days are divided into the training set and the testing set. Third, the different conditional quantiles are acquired by training the QRLSTM only once. Finally, the KDE algorithm is used to obtain the load probability density function. The primary contributions of this paper can be summarized as follows:

(1) The proposed K-QRLSTM produces the probability distribution of electrical loads instead of deterministic prediction results, in which *k*-means and LSTM can reduce the forecast error in a microgrid.

(2) The proposed method K-QRLSTM can save training time as running once to obtain different quantiles.

(3) A real-world data set of Australia is utilized to prove the superiority of K-QRLSTM.

The remainder of this paper is organized as follows. First, the proposed K-QRLSTM is elaborated in Section 2. Next, Section 3 conducts a case study to evaluate the performance of K-QRLSTM. And finally, the conclusion is given in Section 4.

2. K-QRLSTM algorithm-based STLF method

2.1. Overall network

The framework of our proposed K-QRLSTM method for electrical load probability density forecasting is presented in Fig. 1, mainly consisting of the following five steps.

The program of K-QRLSTM algorithm

Start:

Step 1: The data are divided into several categories according to the K-means algorithm; the clustering results of the load are labeled as input features.

Step 2: Select the input data for the neural network and generate time series load from the selected data, the obtained load data will be divided into two subsets including training dataset and testing dataset after normalized;

Step 3: Train the QRLSTM model by using the selected training dataset, and the rest testing dataset is applied to verify the validity of the compound prediction model, which employs the quantile loss to replace the regular mean square error loss;

Step 4: After obtaining the prediction results, the inverse normalization is required correspondingly, three types of prediction results, including deterministic prediction (including median and mode), prediction intervals, and probability density function, are produced by KDE;

Step 5: Calculate the error metrics of the forecast results by relative indicators.

End

2.2. K-means clustering analysis

As a typical and efficient algorithm, the k-means algorithm has been widely used in various applications. Large datasets can be divided into predefined clusters by the distance from the core point to the samples [29]. The k-means algorithm implementation process in STLF is detailed as follows.

Step 1: Select the number N of historical load days; the sampling interval for the whole historical load sequence is dx minutes; the ith load sequence is the farthest from the forecast day that is denoted as x_i, and the number of dimensions is $P = 24 \times 60/dx$. k samples of clustering centers $\{\mu_1, \mu_2, \ldots, \mu_k\}$ are selected as the initial clustering centers. The first clustering center represents the load sequence for a particular day. The load represented by the jth clustering center is denoted as u_j.

Step 2: Calculate the Euclidean distance from the remaining samples, including the ith day to each initial clustering center. The Euclidean distance between the load sequence x_i and the load u_j represented by the jth clustering center is calculated by

$$d_{ij} = \sqrt{\sum_{m=1}^{P} \left(x_i(m) - \mu_j(m)\right)^2} \tag{1}$$

Step 3: We denote each column as a category after sorting each row of the distance d_{ij} matrix from the smallest to the largest, and k clusters $\{C_1, C_2, \ldots, C_k\}$ are chosen The kth cluster set contains a total of N_k samples, and we update the clustering center set by calculating the mean vector of each cluster as the new cluster centers, i.e.,

$$\mu_i(m) = \frac{1}{N_k} \sum_{i=1}^{N_k} sample^k(m) \tag{2}$$

where $\mu_i(m)$ and $sample^k(m)$ denote the mth dimensional element of the new clustering center μ_i and kth sample in the kth cluster respectively.

Step 4: The constant iteration procedure is repeated between Step 2 and Step 3 until the cluster centers do not change or the number of maximum iterations exceeds the predefined threshold.

The final load clustering results are divided into k classes according to the pre-set parameters, and the different categories will be brought into the model as a label in QRLSTM.

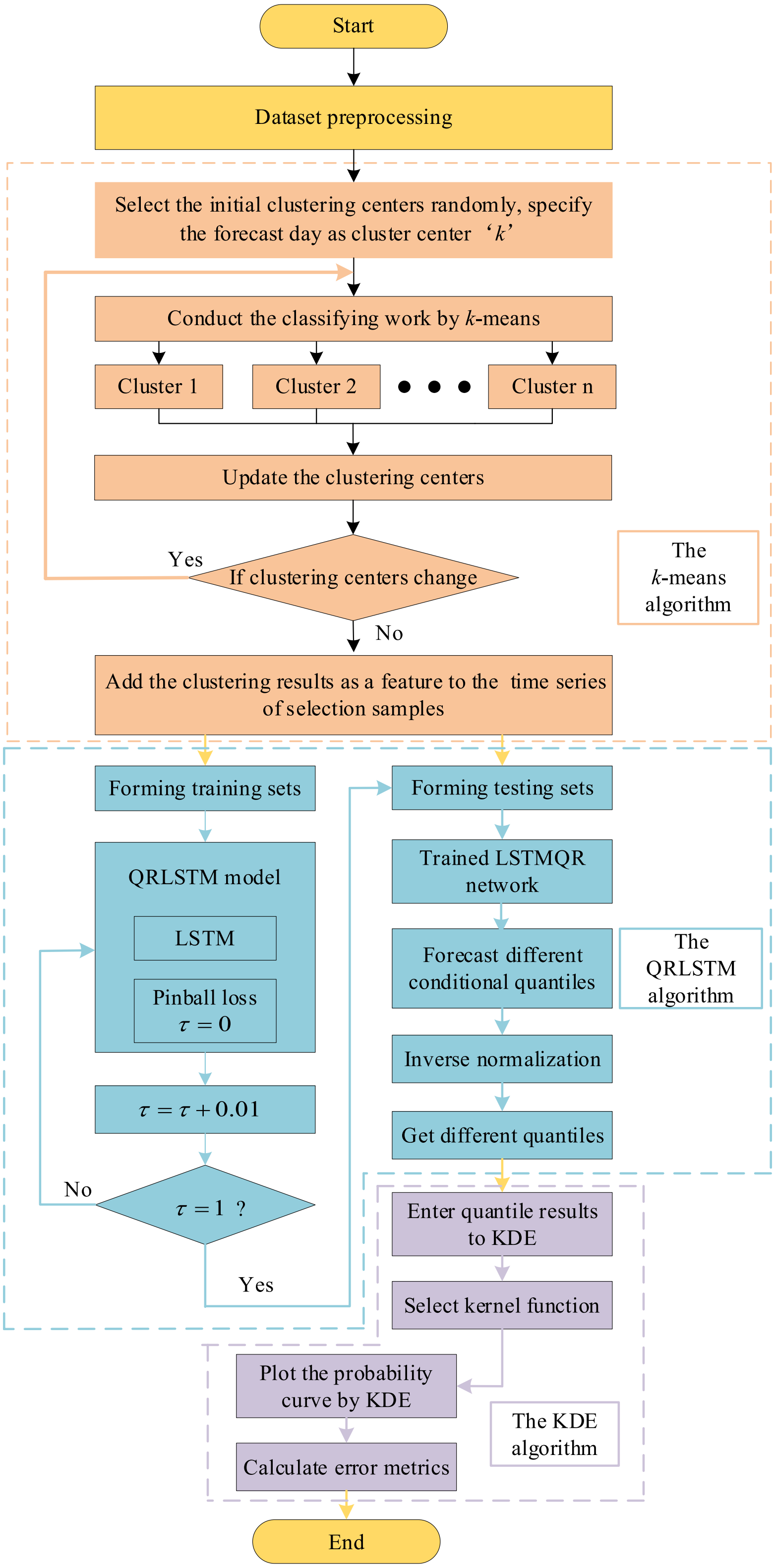


Fig. 1. Flowchart of our proposed load forecasting method based on the K-QRLSTM algorithm.

2.3. Quantile regression long short-term memory neural network

According to the above clustering results, we consider this result part of the input features while marking $H_i \in [0, k]$. H_i represents the clustered index of the ith day. LSTM is an excellent choice for the prediction model due to its excellent properties in dealing with time series. LSTM has achieved outstanding prediction works because of the memory structure, it has been widely used in various applications. The whole LSTM cell can be displayed in Fig. 2.

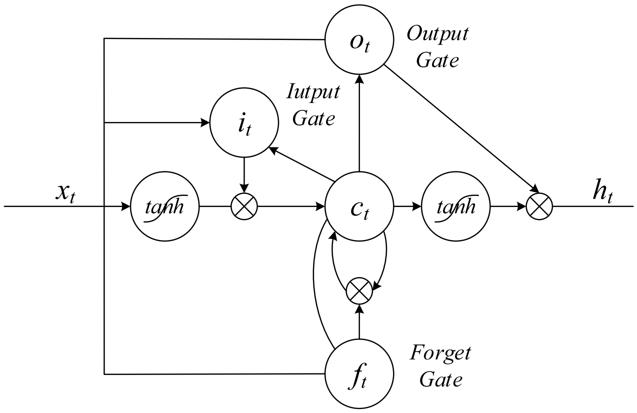


Fig. 2. A standard LSTM cell expansion diagram.

The QRLSTM takes advantage of the LSTM and makes the further adjustment, where the quantile loss replaces the mean square error completely. The other calculation procedure is the same as the LSTM. We arrange the electric loads in chronological order as $\{e_1, e_2, \ldots, e_n\}$ after selecting the historical load data. The general time series prediction uses the sequence $x_c = \{H_{c+1}, e_1, e_2, \ldots, e_c\}$ to predict the load in the next time instant e_{c+1}. Here, c represents the number of lag periods, and H denotes the other characteristics. The output e_t of LSTM at time t is presented by the input sequence $x_{t-1} = \{H, e_1, e_2, \ldots e_{t-1}\}$, the hidden state h_{t-1}, and the cell state C_{t-1}. The calculation process is described in [30,31] but not elaborated for the sake of brevity.

In this paper, we integrate the LSTM with QR, and the resulting QRLSTM algorithm is more suitable for solving the present nonlinear STLF problems. The training model can offer better prediction results by continuously optimizing the quantile loss function, i.e.,

$$L = \min_{W,b} \sum_{i=1}^{n} \rho_{\tau_j} \left[e_i - f\left(x_{i-1}, W, b\right)\right] \tag{3}$$

$$\rho_{\tau_j}(X) = \begin{cases} \tau_j X & e_i \geq f\left(x_{i-1}, W, b\right) \\ \left(1-\tau_j\right) X & e_i < f\left(x_{i-1}, W, b\right) \end{cases} \tag{4}$$

where e_i is forecast load at each time instant, and $f(x_{i-1}, W, b)$ represents the output of LSTM block for load sequence x_{i-1}. $W = \{W_f, W_i, W_c, W_o\}$ denotes the neural network weight vector, which includes the forget gate weight W_f, input gate weight W_i, cell state weight W_c, and output gate weight W_o. $b = \{b_f, b_i, b_c, b_o\}$ is the neural network bias vector, the forget gate bias b_f, input gate bias b_i, cell state bias $b_{c,}$ and output gate bias b_o are the element of it. Furthermore, $\tau_j \in [0, 1]$ represents the jth quantile. The optimal weights $\hat{W}$ and bias $\hat{b}$ can be estimated by Eqs. (3) and (4). The conditional quantiles for a load of each forecast moment can thus be expressed as.

$$e_t\left(\tau_j \middle| x_{i-1}\right) = f\left(x_{i-1}, \hat{W}_{\tau_j}, \hat{b}_{\tau_j}\right) \tag{5}$$

where $e_t\left(\tau_j \middle| x_{i-1}\right)$ denotes the jth quantile in the ith load, and $\hat{W}_{\tau_j}$ and $\hat{b}_{\tau_j}$ represent the optimal weights and bias respectively. The detailed training process of the QRLSTM model is presented Fig. 3. The quantile results are the final predicted output, which gives the complete load distribution.

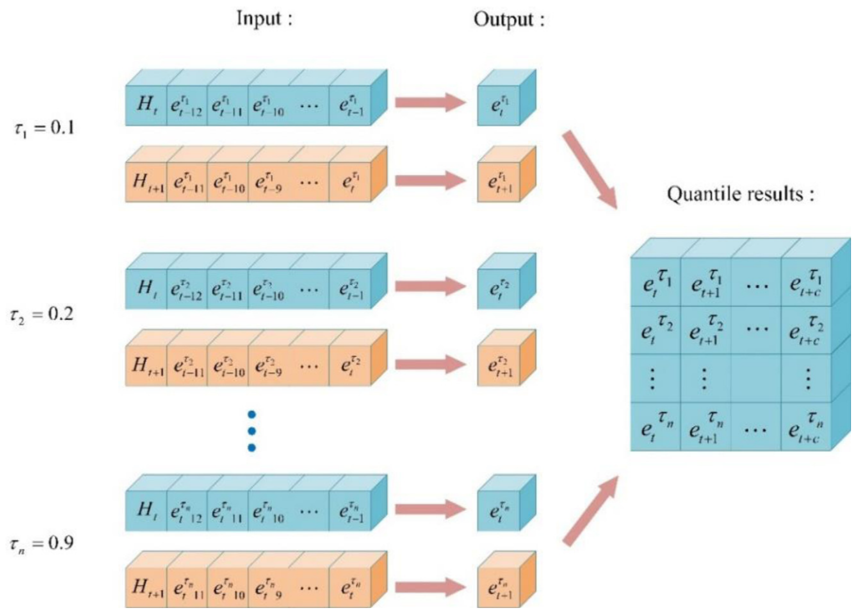


Fig. 3. Detailed training process of the QRLSTM model.

2.4. Kernel density estimation

In our proposed method, KDE is selected to obtain the probability density of the microgrid load. This method does not require any assumption about the prior distribution of the predictor variables to obtain the probability density distribution of the random variables.

Based on the proposed method, Suppose that $Z^i_{\tau_1}, Z^i_{\tau_2}, \ldots, Z^i_{\tau_n}$ represent the different conditional quantiles of the ith load after training the QRLSTM, $Z^i_{\tau_j} = e_t\left(\tau_j \middle| x_{i-1}\right)$ is the jth quantile in ith load. The conditional quantile is utilized as the input values of kernel function The kernel function can be described as

$$\hat{f}_h(x) = \frac{1}{h}\sum_{j=1}^{n} K_h\left(x - Z^i_{\tau_j}\right) = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{x - Z^i_{\tau_j}}{h}\right) \tag{6}$$

where h is the smoothing parameter, that affects the slope of the probability density curve. Furthermore, $K(\cdot)$ is a kernel function, which can be one of the following: the uniform function, the Gaussian function, or the Epanechnikov function.

The probability density of the load at each time instant can be obtained by (5), and each probability density value corresponds to a microgrid load value. A complete load probability curve can be fitted using the probability density and the load data.

2.5. Algorithms integration

The QRLSTM model integrates k-means, LSTM, QR, and KDE algorithms. The K-QRLSTM model is designed to save training time in this paper. It can run once to obtain different quantiles and plot the load probability density curves. Firstly, the clustering results will be produced by the k-means algorithm, and the different categories will be marked to a vector H_N. Secondly, we construct all samples into time series $X_{N,c}$ by taking c lag periods, then the vector of H is spliced with this time series as input sequence of the QRLSTM model. Thirdly, the outputs of different quantiles can be obtained after training the QRLSTM model. Finally, the completed probability distribution can be obtained by entering the different conditional quantile into the KDE algorithm. Error calculation can also be achieved in this process.

2.6. Analysis of model parameters

The prediction results of the deep learning algorithm are always variable and uncontrollable because of the random parameters and weights, the output can be affected by changing the parameters of neural network. In this paper, this impact on the results will further increase due to the several integrated models. We made a reasonable selection of the parameters in each model in order to pursue higher prediction precision.

(1) Structures of our proposed LSTM model

The LSTM model cannot map the connection between the input and the output with a few nodes in the hidden layer. However, the LSTM model will be overfitted if massive nodes are used in the hidden layer. We chose one hidden layer in this paper because it performs better than other parameters, and the number m of nodes in the hidden layer is set as 20 by

$$m = \sqrt{n + l} + \alpha \tag{7}$$

where n and l represents the number of nodes in the input layer and output layer respectively, $\alpha \in [1, 10]$.

(2) Look-back time steps in the LSTM network

The look-back step refers to the first n moments of the time series is taken as input to predict the next time instant of load. The MAPE of prediction results is affected by different look-back steps, because of the different load prediction training processes. In this paper, the number of the look-back time steps will be set as 12 to obtain satisfactory prediction results.

(3) Kernel functions and bandwidth in KDE

We choose the Epanechnikov function (8) as the kernel function because the predicted load results are found to be more accurate than other ones.

$$K(x) = \begin{cases} \frac{3}{4}\left(1 - x^2\right), & |x| \le 1 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

The bandwidth plays a vital role to improve the slope of the probability density curve, The larger bandwidth is, the more smooth the probability density curve can be obtained. The bandwidth is set as 0.5 in this work.

3. Simulation verification and results analysis

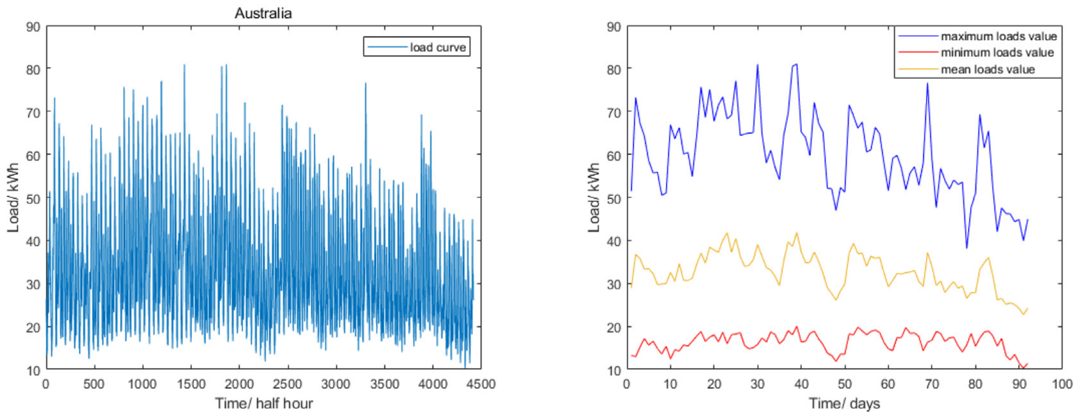
3.1. Dataset analysis

The SGSC project is the first commercial-scale smart grid project in Australia. This project had collected smart meter datasets from New South Wales since 2012. Due to this paper focus on short-term microgrid load forecasting, and thus a subset of the SGSC dataset is adapted for validating the proposed method in this paper. We selected a reasonably sized testing subset according to the criterion of possessing a hot water system, including 69 customers. We use 4416 samples from 92 days, i.e., from 0:30 on 1 June 2013 to 24:00 on 31 August 2013. Fig. 4 shows the load data in the winter, and each daily profile consists of 48 half-hourly samples. As shown in Fig. 4(a), the entire load fluctuates between 10 and 80 kWh. In Fig. 4(b), the blue curve represents the maximum daily load, the red curve is the minimum daily load, and the yellow curve represents the mean daily load. In this dataset, the maximum daily load is high as 80.96, but the minimum daily data is 10.28, which leads to a maximum daily loading rate of 0.696 and a minimum daily loading rate of 0.4792. All of the above indicators exhibit the high volatility of the selection data in a small microgrid.

3.2. Experiment setting and results analysis

In this subsection, a real-world case study is conducted to validate our proposed load forecasting method. All forecasting results with deterministic, interval, and probabilistic prediction types are given and compared with each other. To show the improved performance of the K-QRLSTM model, the MAPE and RMSE are used for error analysis.

Firstly, we apply the k-means algorithm to the aggregated load over three months. This can reveal which factors affect the prediction accuracy. The clustering results in Fig. 5 show no outliers, where the blue curves represent



a. Selected load profile

b. Statistical analysis result of load data

Fig. 4. Load profiles from the selected dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Cluster 1, and the red curves mean Cluster 2. We tested different k and found that $k = 2$ is the most reasonable, leading to 44 days in Cluster 1 and 48 days in Cluster 2. The two clusters are used to label the training sets and the testing tests for prediction, respectively.

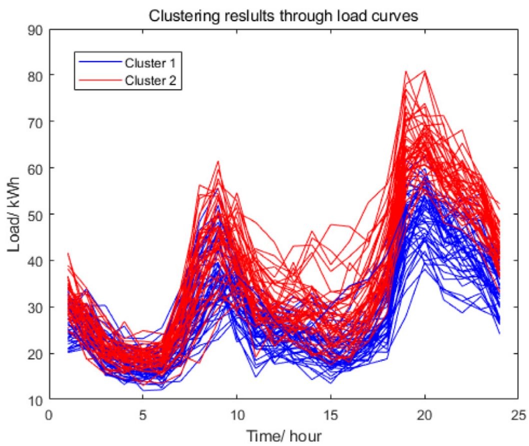


Fig. 5. Aggregated daily profiles with two clusters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Secondly, the input data is normalized into [0, 1] due to the rule function is sensitive to the input data scale. The input features consist of the sequence of electricity consumption for the past 12 steps $\{e_{t-12}, \ldots, e_{t-2}, e_{t-1}\}$. The corresponding day is labeled with a mark H, either 1 or 2.

Thirdly, we divide each dataset into two subsets: the training set (from 01-June-2013 to 22-August-2013) and the testing set (from 23-August-2013 to 31-August-2013). All parameters of the K-QRLSTM model are configured as follows. The quantile range from 0.01 to 0.99 is adopted for obtaining predicted quantile, and the interval is 0.01. The number of nodes in the input layer and the output layer is predefined as 13 and 1. The rest parameters in the neural network including the maximum epoch and learning rate are set as 200 and 0.001 respectively.

Finally, the KDE method is used to obtain the probabilistic density curves of load after different conditional quantiles have been calculated. The median and the mode of the probability density are selected for the point load prediction results of the probabilistic forecasting methods. The median represents the median value in prediction results, as the name suggests, while the mode represents the maximum value in the probabilistic density curve. The corresponding interval is adopted for different confidence levels. Here, a 100% confidence level means all quantiles are used, while if the confidence level is 80%, we adopt the quantiles in [0.1, 0.9].

The deterministic forecast error of each method is shown in Table 1. The evaluation criteria MAPE and RMSE have the same change trend in different prediction methods. The classic approach has the unsatisfactory performance in prediction results, while the proposed K-QRLSTM method outperforms the other algorithms.

Table 1. Forecast errors with different algorithms.

Evaluation criteria	BP	LSTM	QRLSTM (Median)	QRLSTM (Mode)	QR (Median)	QR (Mode)	K-QRLSTM (Median)	K-QRLSTM (Mode)
MAPE (%)	12.80	10.10	9.44	9.86	10.01	10.17	9.20	9.19
RMSE	3.65	3.22	3.04	3.12	3.33	3.50	2.97	2.96

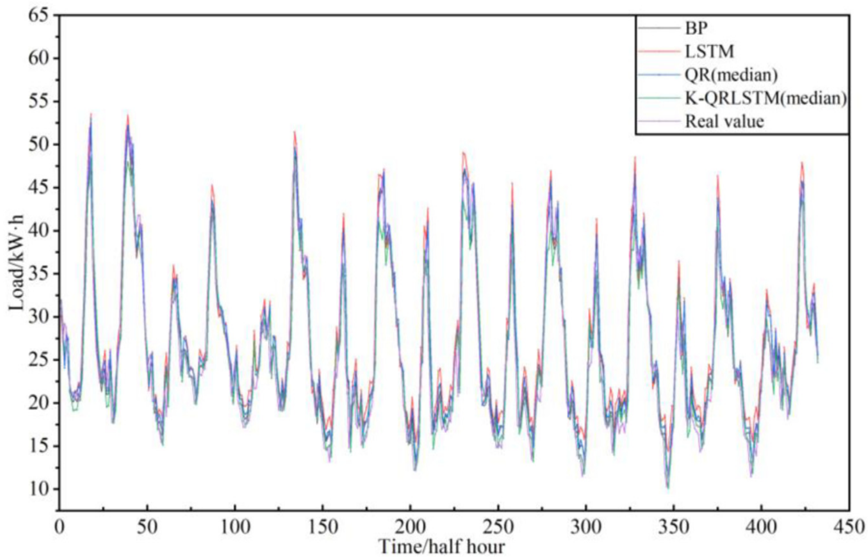


Fig. 6. The load forecasting results within nine days for diverse methods. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The forecasting results of the microgrid are illustrated in Fig. 6. To highlight the superiority of the proposed K-QRLSTM, the other forecasting models are also presented for comparison. The blue curve represents the prediction results of the BP neural network, the red curve reflects the prediction results of LSTM, the yellow curve conducts the prediction results of the QR method, the purple curve indicates the prediction results of the K-QRLSTM method, and the green curve stands for true load profile. To better exhibit the changes in the predicted results, we plot the detail of the first day's prediction results in Fig. 7, where it can be clearly shown that the proposed K-QRLSTM can achieve higher prediction accuracy than the other methods.

Fig. 8 reflects the interval forecast results under different confidence levels. The blue lines describe the prediction interval under a 100% confidence level, the red dotted lines represent the prediction interval under an 80% confidence level, and the yellow lines indicate the true loads. Most of the true loads fall within the prediction interval. The prediction interval under the 80% confidence level is closer to the actual values than that under the 100% confidence level.

Due to the high randomness and lower similarity of history load curves in the microgrid, the probability density curves in the 6th, 13th, 17th, and 23rd hours from the test dataset on 23-August-2013 are randomly selected and shown in Fig. 9. In each subfigure, the blue curve is the probability density, and the red straight line indicates the actual load. It can be observed that the actual values are close to the maximum probability in the probability density curve. The results validate that the proposed method can provide accurate prediction as well as highlight the uncertainty in the microgrid.

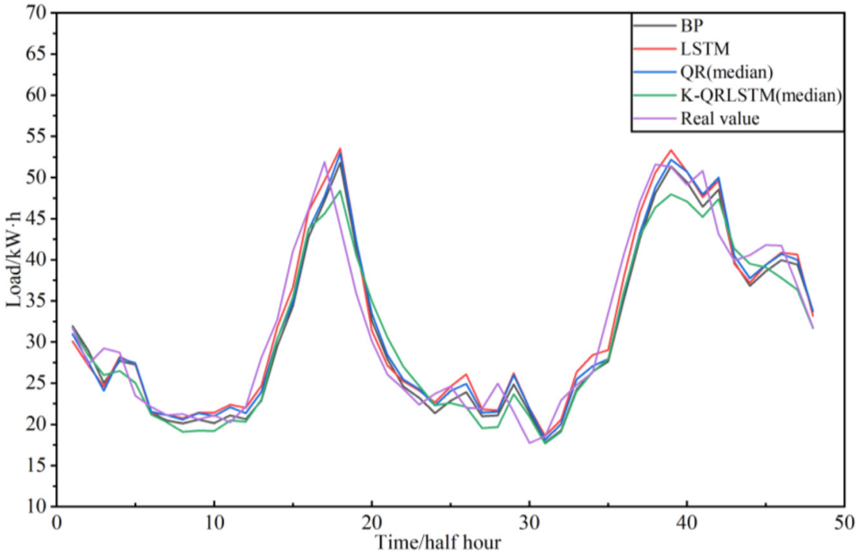


Fig. 7. Detailed forecasting results on the first day of the selected dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

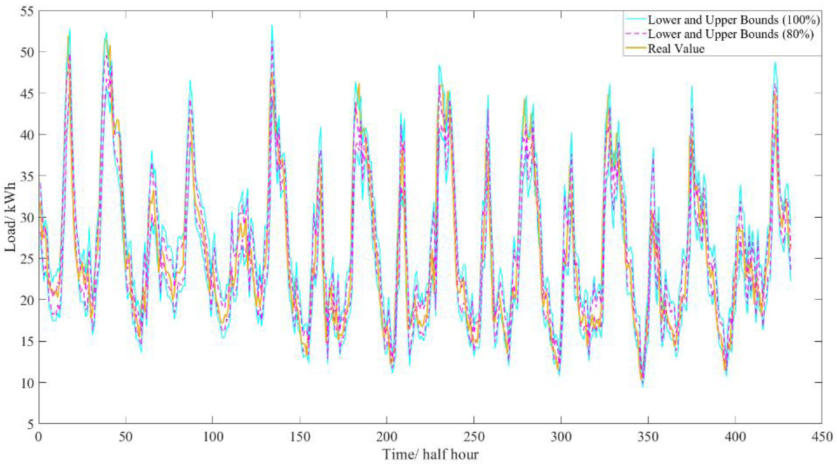
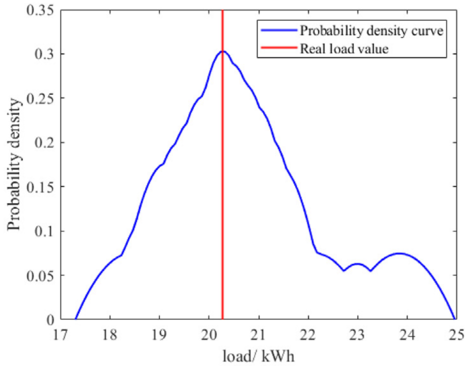


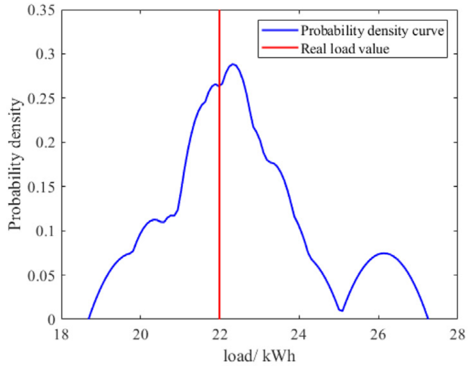
Fig. 8. Load forecasting interval results under confidence levels of 80% and 100%. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Conclusions

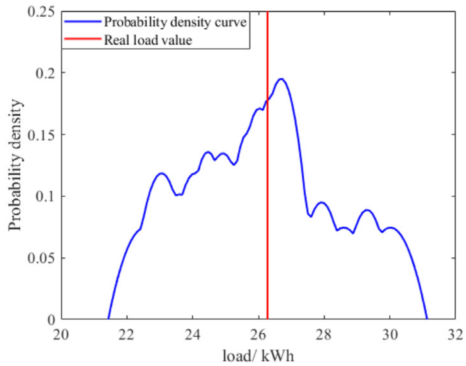
To improve the accuracy and practicability of STLF in the microgrid optimal dispatch, a short-term microgrid load probability density forecasting method based on *k*-means and deep learning regression is presented. The model integrates *k*-means, LSTM, QR, and KDE with the Epanechnikov kernel. Real-world load data are utilized to evaluate the performance of the K-QRLSTM model. From the simulation results, In terms of training speed, K-QRLSTM only needs to train the model once to get different quantiles. The error of QRLSTM is smaller than many existing methods such as BP, LSTM, and QR. The comprehensive evaluation shows that our proposed method can produce good informative and acceptable precise deterministic, interval, and probability density load prediction results.



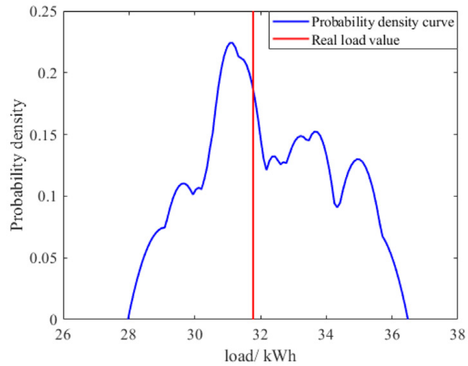
a. Load probability density curve in the 6th hour.



b. Load probability density curve the 13th hour.



c. Load probability density curve for the 17th hour.



d. Load probability density curve for 23rd hour.

Fig. 9. Probability density functions at different time instants. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant 51707139.

References

[1] Zhang S, Wang Y, Zhang Y, et al. Load probability density forecasting by transforming and combining quantile forecasts. Appl Energy 2020;277:115600. http://dx.doi.org/10.1016/j.apenergy.2020.115600.
[2] Li F, Yu X, Xin T, et al. Short-term load forecasting for an industrial park using LSTM-RNN considering energy storage. In: 2021 3rd Asia energy and electrical engineering symposium (AEEES). 2021, p. 684–9.
[3] He Y, Xu Q, Wan J, et al. Short-term power load probability density forecasting based on quantile regression neural network and triangle kernel function. Energy 2016;114:498–512.
[4] Li T, Wang Y, Zhang N. Combining probability density forecasts for power electrical loads. IEEE Trans Smart Grid 2020;11:1679–90.
[5] Kong Z, Zhang H, Lv F, et al. Multimodal feature extraction and fusion deep neural networks for short-term load forecasting. IEEE Access 2020;8:185373-185383.
[6] Imani M, Ghassemian H. Residential load forecasting using wavelet and collaborative representation transforms. Appl Energy 2019;253:113505.
[7] Pei S, Qin L, Yao L, et al. Multi-step ahead short-term load forecasting using hybrid feature selection and improved long short-term memory network. Energies 2020;13:4121.
[8] Hong T, Fan S. Probabilistic electric load forecasting: A tutorial review. Int J Forecast 2016;32:914–38.
[9] Christiaanse W. Short-term load forecasting using general exponential smoothing. IEEE Trans Power Appar Syst 1971;2:900–11.
[10] Chen Y, Xu P, Chu Y, et al. Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings. Appl Energy 2017;195:659–70.

[11] Lee C, Ko C. Short-term load forecasting using lifting scheme and ARIMA models. Expert Syst Appl 2011;38:5902–11.
[12] Bisht B, Holmukhe R. Electricity load forecasting by artificial neural network model using weather data. IJEET Trans Power Syst 2013;4:91–9.
[13] Lei J, Jin T, Hao J, et al. Short-term load forecasting with clustering–regression model in distributed cluster. Cluster Comput 2019;22:10163–73.
[14] Huang N, Wang W, Wang S, et al. Incorporating load fluctuation in feature importance profile clustering for day-ahead aggregated residential load forecasting. IEEE Access 2020;8:25198–209.
[15] Aurangzeb K, Alhussein M, Javaid K, et al. A pyramid-CNN based deep learning model for power load forecasting of similar-profile energy customers based on clustering. IEEE Access 2021;9:14992–5003.
[16] Sun Z, Zhao S, Zhang J. Short-term wind power forecasting on multiple scales using VMD decomposition, K-means clustering and LSTM principal computing. IEEE Access 2019;7:166917-166929.
[17] Liu N, Tang Q, Zhang J, et al. A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids. Appl Energy 2014;129:336–45.
[18] He Y, Liu R, Li S, et al. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory. Appl Energy 2017;185:254–66.
[19] Zheng H, Yuan J, Chen L. Short-term load forecasting using EMD-LSTM neural networks with a xgboost algorithm for feature importance evaluation. Energies 2017;253:1168.
[20] Sun Z, Chen T, Lian K, et al. Study on load forecasting of an actual micro-grid system in Taiwan. IEEE; 2013, p. 1–7.
[21] Li T, Wang Y, Zhang N. Combining probability density forecasts for power electrical loads. IEEE Trans Smart Grid 2020;11:1679–90.
[22] He Y, Zheng Y. Short-term power load probability density forecasting based on Yeo-Johnson transformation quantile regression and Gaussian kernel function. Energy 2018;154:143–56.
[23] Hong T, Pinson P, Fan S, et al. Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. Int J Forecast 2014;32:896–913.
[24] Haben S, Giasemidis G. A hybrid model of kernel density estimation and quantile regression for GEFCom2014 probabilistic load forecasting. Int J Forecast 2016;32:1017–22.
[25] He Y, Zhang W. Probability density forecasting of wind power based on multi-core parallel quantile regression neural network. Knowl-Based Syst 2020;209:106431.
[26] Wang Y, Gan D, Sun N, et al. Probabilistic individual load forecasting using pinball loss guided LSTM. Appl Energy 2014;235:10–20, (2019).
[27] Kou P, Gao F. A sparse heteroscedastic model for the probabilistic load forecasting in energy-intensive enterprises. Int J Electr Power Energy Syst 2014;55:144–54.
[28] Popov S, Stanev R, Baeva S, et al. Stochastic model for microgrid load forecasting. AIP Conf Proc 2021;2333.
[29] Sun Z, Zhao S, Zhang J. Short-term wind power forecasting on multiple scales using VMD decomposition, K-means clustering and LSTM principal computing. IEEE Access 2019;7:166917-166929.
[30] Stratigakos A, Bachoumis A, Vita V, et al. Short-term net load forecasting with singular spectrum analysis and LSTM neural networks. Energies 2021;14:4107.
[31] Kong W, Dong Z, Jia Y, et al. Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Trans Smart Grid 2019;10:841–51.