Membrane Technology
ISSN (online): 1873-4049

# An improved prediction model for residential user response potential based on DTW-K-medoids double-layer clustering and Bi-LSTM

Wenhui Zhao[1], Mutian Guo[1,*], Jinlong Yu[2], Zilin Wu[3]
[1] College of Economics and Management, Shanghai University of Electric Power, Shanghai 200090, China
[2] Huangshan Power Supply Company, State Grid Anhui Electric Power Company, Huangshan 245000, China
[3] Beijing Branch, China Tower Co., Ltd, Beijing 100142, China
* Corresponding Author.

**Abstract:**
As demand response (DR) participants gradually loosen up, the number of new users participating in DR will continue to increase. Due to the small electricity consumption and wide distribution of residential users, they cannot participate in DR individually and need to be represented by load aggregators. However, for users who haven't participated in DR before, the lack of historical response data makes it more difficult for load aggregators to predict their response potential. In this context, a response potential prediction model was constructed for such users. Firstly, based on the user's electricity consumption load and energy-saving awareness as the clustering criteria for the upper and lower layers, the DTW-K-medoids model is used to perform double-layer clustering on users, grouping users with similar electricity consumption behaviors and concepts into one group. Then, based on the clustering results, the electricity consumption data of residential users who have participated in DR among various types of users are used as the training set, and the Bi-LSTM algorithm is used to sequentially construct a response potential prediction model for each type of user. Finally, case analysis shows that the proposed model can effectively solve the problem of low prediction accuracy of response potential for users who haven't participated in DR, and further reduce the assessment cost of load aggregators.

**Keywords:** Residential users, Response potential prediction, DTW-K-medoids dual-layer clustering; Bi-LSTM algorithm.

## INTRODUCTION

The current Demand Response (DR) mechanism divides participating entities into categories such as industrial users, commercial users, and residential users. Due to the high electricity consumption and high proportion of electricity costs for industrial and commercial users (Deng et al., 2023), they have great potential for response. Therefore, many references have conducted in-depth research on the application of industrial and commercial users in DR (Lu et al., 2021; Xu et al., 2023). As the proportion of renewable energy in the power system gradually increases, demand side resources become increasingly valuable. Government departments expect more types of entities to participate in DR to further tap into the response potential of end users. However, the electricity consumption of low-power users such as residential users is relatively low and widely distributed, making it impossible for them to participate in DR as independent users. Therefore, government departments have introduced third-party entities - load aggregators (LA), and low-power users can seek LA to act as agents for their participation in DR (Li et al., 2024). As the agent, the LA needs to declare the response volume on the trading platform. Therefore, aggregators need to accurately predict the response potential of their proxy users to avoid corresponding penalties. Due to the irregular electricity consumption behavior of the residential user group (Lazzari et al., 2022), it increases the difficulty of LA in predicting the potential response volume. Therefore, in the context of user side liberalization, more and more scholars are studying how to improve the LA's prediction of the response potential of its proxy users.

Reference (Zhou et al., 2024) uses K-means algorithm to cluster flexible loads such as air conditioners and dishwashers to determine the DR response potential of residential users. Case analysis shows that this method can improve the management ability of LA; Reference (J. Wang et al., 2023) proposes an evaluation method for the potential response of air conditioning for residential users. Obtain all user devices with similar adjustable features through aggregation algorithms, and establish an air conditioning load model to effectively solve the problem of LA delivering fines due to insufficient response ability; Reference (Duman et al., 2023) conducted a survey on the electricity usage behavior and sensitivity of residential users to electricity bills. The user's household energy management system will optimize their electricity consumption behavior based on the survey results. Residential users responding to electricity usage based on the instructions provided by the management system will effectively enhance their response potential; Reference (Sridhar et al., 2023) investigated the motivation of residential users to participate in DR through a questionnaire survey, and quantified the impact of various motivations on response potential to further improve prediction accuracy.

The above references have made important contributions in predicting the response potential of low-power users such as residential users. However, there are problems in the data collection process, such as a wide range of data sources and leakage of customer privacy, which make it difficult to operate in practical applications. It is worth noting that in November 2021, the National Development issued the "Management Measures for Power Selling Companies" document (National Development and Reform Commission, 2021), which stated that LA can access user historical electricity consumption data after obtaining authorization from their proxy customers. Therefore, studying how to use easily available data such as electricity consumption to predict the response potential of residential users is of great significance for improving operability and data privacy. Reference (Shi & Jiao, 2023) suggests that factors such as electricity consumption, peak electricity consumption during response periods, interruptible loads, electricity prices, and range of electricity consumption fluctuations are closely related to response behavior. They are used as characteristic variables to evaluate user response potential using machine learning algorithms; Reference (Shirsat & Tang, 2021) selected electricity consumption, response time, and temperature sensitivity factors as characteristic variables, and used a mixed density recurrent neural network to evaluate the DR potential of residential users; Reference (Kong et al., 2023) uses response time, incentive price, load baseline, actual load and other factors as characteristic variables, and uses a combination of long short-term memory network and mixed density network to obtain the probability distribution of user response quantity; Reference (Zhang et al., 2020) proposed a distributed modeling method based on the fully distributed alternating direction multiplier method to evaluate user DR potential, using electricity price, electricity consumption, historical response quantity, and response time as characteristic variables to construct the model; Reference (Kong et al., 2020) suggests a close relationship between electricity prices and user DR potential, and uses price elasticity as a characteristic variable to predict user DR potential using neural networks.

The research approach of the above references is to select factors related to user response potential as feature variables to construct a prediction model. The variables taken, such as subsidy prices, temperature, time, etc., do indeed have a correlation with the user's response, but most references have demonstrated and believed that historical response is one of the key indicators for predicting response potential (Afzalan & Jazizadeh, 2019; Zhang et al., 2020). As more and more users learn about DR, more new users will participate in the future. Under this trend, the difficulty of LA predicting the response potential of users who haven't participated in DR increases due to the lack of historical response data. Therefore, this article aims to fill the gap in this research field and explore how to accurately predict the response potential of users who have not participated in DR. This article proposes a response potential prediction model for residential users who have not participated in DR, which combines a DTW improved double-layer clustering model and a Bi-LSTM model. Firstly, divide the 24-hour day into response time periods and non-response time periods, and use a clustering algorithm improved by DTW to perform upper-level clustering on residential users, grouping users with similar electricity consumption behaviors during non-response time into one category. Then, based on the energy-saving awareness of each user, lower-level clustering is carried out for each type of user, further grouping users with similar electricity consumption behaviors and energy-saving concepts into one category. Finally, the electricity consumption data of residential users who have participated in DR in various types are used as the training set, and a Bi-LSTM response potential prediction model is constructed for each type of user in sequence. By using predictive models, the response potential values of residential users who did not participate in DR in each type can be obtained.

## CLUSTER ANALYSIS OF RESIDENTIAL USERS

Reference (F. Wang et al., n.d.) indicates that if the load curve trends of users are similar, then these users have similar response behaviors during the response period. Therefore, this article takes this as a starting point, using load data as clustering basis, categorizing users with similar electricity consumption behaviors, and predicting the response behavior of new users based on the response behavior of users in the same category. Due to differences in education levels and environmental concerns among resident users, even if their load curve trends are similar, their response behaviors may differ during the DR period (Ballenger et al., 2017; Shekari et al., 2021; B. Wang et al., 2020). Therefore, this article uses the user's load curve and energy-saving awareness as the clustering basis to conduct a double-layer clustering of users, as shown in Figure 1.

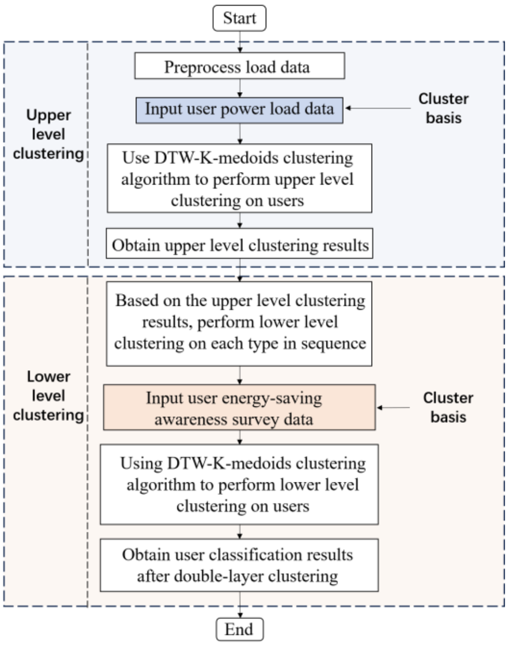



Figure 1. Double layer clustering flowchart

In the upper-level clustering model, users with similar electricity consumption behaviors are classified into the same category based on load data clustering. On the basis of the upper-level clustering results, based on the user's energy-saving awareness, lower-level clustering is carried out for each type of user in sequence. Through double-layer clustering, users with similar electricity consumption behaviors and energy-saving awareness are grouped together, and users in the same group have similar response behaviors during the DR period.

**Data preprocessing**

Due to the significant differences in electricity consumption among different residential users, this will result in the model being more sensitive to larger values. Therefore, this article normalizes the user electricity load data to ensure that the contribution of each value to the model is relatively balanced. Assuming the number of measurement points for the user's load data within a day is $T$, it $P_{i,j,t}$ indicates the power consumption of user $i$ at time $t$on day $j$, where $t \in T$. $\bar{P}_{i,j,t}$ is the normalized value of $P_{i,j,t}$, subsequent models are constructed based on the normalized load data. The expression is as follows:

$$\bar{P}_{i,j,t} = \frac{P_{i,j,t}}{max(P_{i,j,t})} \tag{1}$$

**Quantify the level of user energy-saving awareness**

This article refers to relevant references and combines practical situations (Ballenger et al., 2017; Shekari et al., 2021; B. Wang et al., 2020), selecting economic awareness, environmental change awareness, social responsibility awareness, and electricity conservation awareness to comprehensively reflect the user's electricity conservation awareness. The specific content of each awareness is shown in Table 1.

Table 1. Specific Content of Energy Conservation Awareness

| Factors affecting energy-saving awareness | Concrete content |
|---|---|
| Economic awareness | Do you value the proportion of electricity expenses to household income and pay attention to the monthly changes in household electricity bills? |
| Environmental change awareness | Are you paying attention to the current environmental changes such as climate and ecosystem? |

| | |
|---|---|
| Social responsibility awareness | Do you pay attention to personal lifestyle in daily life, such as choosing green transportation and reducing excessive electricity consumption, to fulfill corresponding social responsibilities? |
| Electricity conservation awareness | Have you developed the habit of turning off lights casually in daily life, and have you gained a general understanding of the power consumption of various electrical appliances? |

The current research references mainly use questionnaire surveys to understand the level of user awareness (B. Wang et al., 2020; Z. Wang et al., 2020). However, this method has some shortcomings, such as respondents being easily influenced by social expectations and filling in answers that are not true to maintain their personal image. In addition, the questionnaire survey method has problems such as high labor costs and low efficiency. To solve the previously listed problems, this study proposes incorporating a survey on user energy-saving awareness into power service application software. After successfully registering in the power service application software, the user will enter the energy-saving awareness survey to comprehensively understand their attitude towards energy-saving awareness. The energy-saving awareness survey page is shown in Figure2.

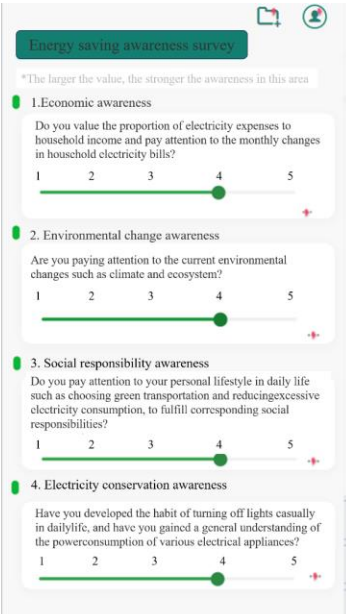



Figure 2. Energy saving awareness information collection page

**Principle of DTW algorithm**

Compared to the classic time series similarity calculation method of Euclidean distance, the outstanding feature of Dynamic Time Warping (DTW) algorithm is that it does not require the length of the time series to be consistent (Lee & Leung, 2023). Due to various real-world factors that constrain the data collection process, datasets often struggle to ensure one-to-one correspondence at time points, thereby limiting the applicability of Euclidean distance(Melnykov & Michael, 2020). The basic principle of the DTW algorithm can be summarized as follows(Yan et al., 2022):

(1) For time series $a = (a_1, a_2, \dots, a_n)$ and $b = (b_1, b_2, \dots, b_n)$, the Euclidean distance between data points $a_i$ and $b_j$ can be expressed as $p(a_i, b_j) = |a_i - b_j|$. Therefore, a $m \times n$ distance matrix of dimensions can be formed. The expression is as follows:

$$p_{m\times n} = \begin{pmatrix} p(a_1, b_1) & \dots & p(a_i, b_1) \\ \vdots & \ddots & \vdots \\ p(a_1, b_j) & \cdots & p(a_n, b_m) \end{pmatrix} \tag{2}$$

(2) The combination of elements in the distance matrix $p_{m\times n}$ forms multiple curved paths $Q$ starting from $(a_1, b_1)$ and ending at $(a_n, b_m)$, as shown in the following expression:

$$Q = \{q_1, q_2, \ldots, q_K\} \quad (3)$$

In the formula, $q_K$ represents the $K$-th element in the curved path $Q$, which needs to meet the requirements of boundary constraint, continuity, monotonic increasing, and number limitation (El Amouri et al., 2023).

(3) The optimal bending path is obtained with the goal of minimizing the cumulative distance, as shown in the following expression:

$$DTW(a, b) = min\sqrt{\sum_{i=1}^{k} p_i} \quad (4)$$

**DTW-K-medoids clustering algorithm**

Due to the frequent problem of getting stuck in local optima in K-Means clustering algorithm (Sieranoja & Fränti, 2022), relevant scholars have proposed K-Medoids clustering algorithm to improve the accuracy of clustering. The K-Medoids clustering algorithm effectively solves the shortcomings of K-Means clustering by using the actual samples with the highest similarity to other sample points as the representative clustering centers. However, due to the fact that the K-medoids clustering algorithm uses Euclidean distance as a measure of similarity between sample points, the model is prone to problems such as abnormal sensitivity to noisy data, poor clustering performance for non spherical clusters, and lack of time elasticity in practical applications (Sobrinho Campolina Martins et al., 2024). This article combines the DTW algorithm with the K-medoids clustering algorithm, and uses the DTW algorithm to calculate the distance between two sample data in the K-medoids algorithm. The DTW algorithm can find the optimal bending path between two time series data, effectively solving problems such as local scaling and drift of the sequence. The flowchart is shown in Figure 3. The specific steps of the DTW-K-medoids algorithm are as follows:

For dataset $A = \{a_1, a_2, \ldots, a_N\}$, the sample point $a_i$is a $M$vector of dimensions, $a_{ij}$ represents the $j$-th data of sample point $i$. Assuming the initial cluster center point is $a_x{}^*, x = 1,2, \ldots K$.

1. Use the DTW algorithm to calculate the distance between each sample point and the initial cluster center point $d(a_i, a_x^*)$.

2. Divide the dataset into $K$ classes to form a cluster set $G = \{g_1, g_2, \ldots, g_K\}, K \leq N$, where $g_i$ represents the $i$-th cluster. Calculate the cumulative distance between the sample point in the $i$-th cluster and the initial cluster center point, expressed as follows:

$$E = \sum_{x=1}^{K} \sum_{p \in g_x} d(p, a_x^*) \quad (5)$$

In the formula, $p$ represents all sample points in the $x$-th cluster, and $E$ represents the sum of distances between sample points in each cluster and the cluster center.

3. Calculate the change in the sum of squared cumulative errors during iteration, expressed as follows:

$$S = E_i - E_{i-1} \quad (6)$$

In the formula, $E_i$ represents the sum of squared cumulative errors obtained from the $i$-th iteration, and $E_{i-1}$ represents the sum of squared cumulative errors obtained from the $(i-1)$-th iteration.

4. If $S < 0$, the latest center point is used as the cluster center point, and the remaining samples are divided into clusters closest to the latest cluster center point; If $S > 0$, no changes will be made.

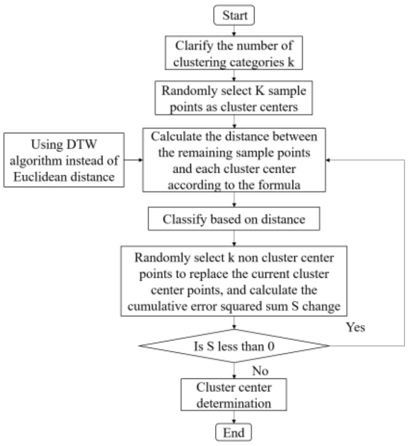



Figure 3. DTW-K-medoids clustering algorithm flowchart

**Cluster Effect Evaluation Index System**

Evaluating clustering effectiveness requires comprehensive consideration of multiple aspects, such as intra cluster compactness and inter cluster dispersion. Therefore, this article constructs an evaluation index system to comprehensively reflect the clustering effect.

Davies-Bouldin Index (DBI) can be used to evaluate the tightness within clusters, focusing on the tightness within clusters and the separation between clusters (Guo et al., 2022). DBI calculates the relative distance between each cluster and its nearest cluster, with smaller values indicating tighter samples within the cluster. The expression is as follows:

$$DBI = \frac{1}{k}\sum_{i=1}^{k} \max_{i \neq j}\left(\frac{S_i + S_j}{C_{ij}}\right) \tag{7}$$

In the formula, the distances between the $i$-th and $j$-th sample sets and that class's cluster center are denoted by $S_i$ and $S_j$, respectively; $C_{ij}$ represents the distance between the $i$cluster center of the class and the $j$cluster center of the class.

The Calinski - Harabasz (CH) index is calculated based on intra cluster dispersion and inter cluster compactness, using the intra class dispersion matrix to reflect intra cluster dispersion and the inter class dispersion matrix to reflect inter cluster compactness(Yu et al., 2023). The larger the CH value, the smaller the intra cluster variance and the larger the inter cluster variance. The specific formula is as follows:

$$CH = \frac{\frac{tr\left(S_B(K)\right)}{K-1}}{\frac{tr\left(S_W(K)\right)}{N-1}} \tag{8}$$

$$S_B(K) = \sum_{k=1}^{k} P_k(m_k - m)\,(m_k - m)^T \tag{9}$$

$$DBI = \frac{1}{k}\sum_{i=1}^{k} max\left(\frac{S_i + S_j}{C_{ij}}\right) \tag{10}$$

In the formula, $K$ represents the number of clusters; $N$ represents the number of samples; $N_K$ represents the number of samples in class $K$. $S_B(K)$ and $S_W(K)$ represent the inter class and intra class dispersion matrices, respectively. $tr(S_B(K))$ and $tr(S_W(K))$ represent their traces, respectively; $x^{(k)}{}_i$ represents the input vector of user $i$ in class $K$; $P_k$ represents the prior probability of the $K$-th type of user; $m$ represents the vector mean of all sample data; $m_k$ represents the vector mean of the $K$-th type of user.

## A PREDICTION MODEL FOR THE RESPONSE POTENTIAL OF RESIDENTIAL USERS

The input variables for resident users used in this article are time series. Due to the characteristics of multiple time scales and nonlinearity in time series data, machine learning has poor information mining performance on time series data (K. Wang et al., 2019). Deep learning models can effectively mine complex time dependencies in time series data and are specifically designed to handle long-term dependencies in data (Chen et al., 2023). Therefore, this article adopts deep learning algorithms to construct a resident user DR potential prediction model.

### Model Input Variables Selection

This article constructs a prediction model based on the clustering results. Since the execution period $T_{dr}$ of each DR is not fixed, The training set of the model consists of 75% of the load data $P_{k,i,j,t}^{dr}, t \in T_{dr}$ and response amount $d_{k,i,j,t}, t \in T_{dr}$ of users who have participated in DR, and the remaining 25% of the samples are used as the testing set of the model. $P_{k,i,j,t}^{dr}$ and $d_{k,i,j,t}$ represent the power and response of user $i$ in category $K$ who have participated in DR at time t on the $j$-th day, respectively. The specific framework diagram is shown in Figure 4.

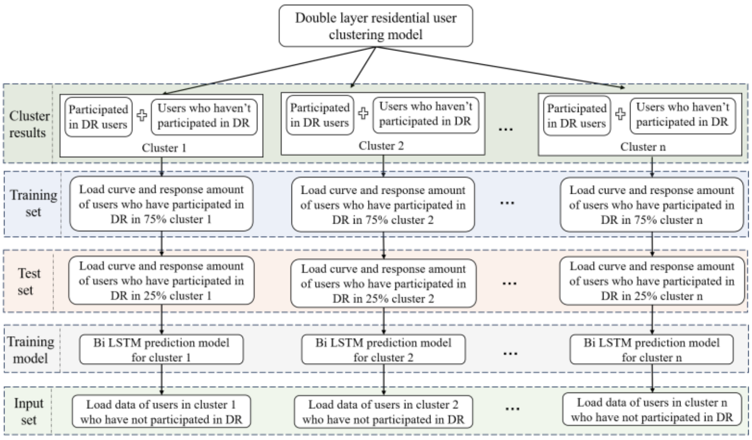



Figure 4. Flow Chart of Input Variables Selection Process for Prediction Model

### Principle of Bi-LSTM model

In recent years, LSTM has been widely used in fields such as speech recognition and computer vision. The structure of LSTM is shown in Figure 5.

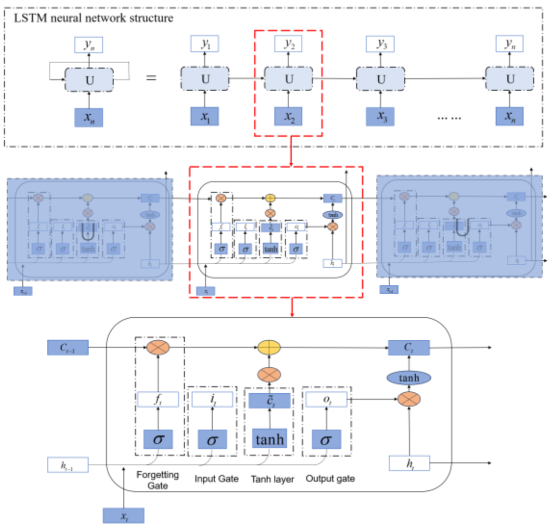



Figure 5. Working Steps of LSTM Model

LSTM introduces gating mechanisms such as input gates, forget gates, and output gates to manage the information from the previous moment, resulting in dynamic changes in the weight of the network during each time step's cyclic process (Fang et al., 2024). The sigmoid activation function will determine which information is transmitted to the next neural unit. The specific work steps of LSTM are as follows (Dao et al., 2024):

1. The forgetting gate calculates the output vector $f_t$ using the following formula:

$$f_t = \sigma\left(w_i * [h_{t-1}, x_t] + b_f\right) \tag{11}$$

In the formula, $\sigma$ represents the excitation function; $w_f$ represents the weight of input features; $x_t$ represents the input value of the sample at time $t$; $h_{t-1}$ represents the output value of the previous unit; $b_f$ represents the offset amount.

2. The input gate calculates the value of $i_t$, which determines the information to be retained in the current stage. The expression is shown in formula 12. The Tanh layer calculates the candidate memory vector $\tilde{c}_t$, which is obtained from $x_t$ and the output $h_{t-1}$, as shown in formula 13:

$$i_t = \sigma(w_i * [h_{t-1}, x_t] + b_i) \tag{12}$$

$$\tilde{c}_t = tanh(w_c * [h_{t-1}, x_t] + b_c) \tag{13}$$

In the formula, $w_i$ and $w_c$ represent coefficients in their respective functions; $b_i$ and $b_c$ represent the deviations in their respective functions.

3. Calculate the new state $C_t$, which is obtained by combining various outputs. The formula is as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{c}_t \tag{14}$$

4. Calculate the judgment condition $O_t$ for the output gate, and obtain the model output value from $O_t$ and $C_t$. The formula is as follows:

$$O_t = \sigma(w_0 * [h_{t-1}, x_t] + b_0) \tag{15}$$

$$h_t = O_t * tanh(C_t) \tag{16}$$

In the formula, $w_0$ represents the coefficients in the function; $b_0$ represents the deviation in the function.

From the above analysis, it can be seen that LSTM networks can only train past information and cannot mine future information. To make up for this deficiency, this article adds a reverse LSTM layer to the original LSTM infrastructure, forming a Bi-directional Long Short Time Memory (Bi-LSTM), whose structure is shown in Figure 6.

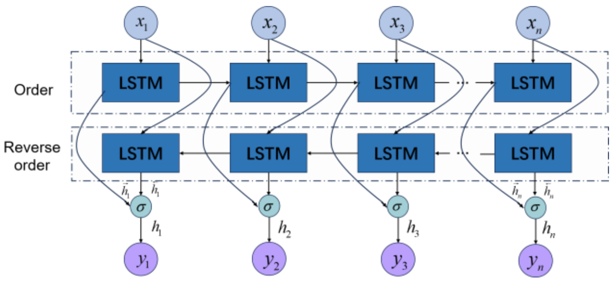



Figure 6. Structure diagram of Bi-LSTM network model

In Figure 6, $\vec{h}$ represents the hidden output sequence of the forward layer, and $\overleftarrow{h}$ represents the hidden output sequence of the reverse layer. The overall output sequence $h$ of the Bi-LSTM network model is composed of $\vec{h}$ and $\overleftarrow{h}$ together, and the following are the expressions:

$$\vec{h}_t = \overrightarrow{LSTM}\left(\vec{h}_{t-1}, x_t, \vec{C}_{t-1}\right), t \in [1, T] \tag{17}$$

$$\overleftarrow{h}_t = \overleftarrow{LSTM}\left(\overleftarrow{h}_{t+1}, x_t, \overleftarrow{C}_{t+1}\right), t \in [1, T] \tag{18}$$

$$h_t = \sigma\left(w_h\left[\vec{h}_t, \overleftarrow{h}_t\right] + b_h\right) \tag{19}$$

In the formula, $\vec{h}_{t-1}$ and $\overleftarrow{h}_{t+1}$ respectively represent the $(t-1)$-th hidden state in the forward LSTM and the $(t+1)$-th hidden state in the reverse LSTM; $\vec{C}_{t-1}$ and $\overleftarrow{C}_{t+1}$ represent the $(t-1)$-th cell state in forward LSTM and the $(t+1)$-th cell state in reverse LSTM, respectively.

The prediction accuracy of the Bi-LSTM model can be reflected by the MSE index. The expression is as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - f(x_i)\right)^2 \tag{20}$$

In the formula, $y_i$ represents the actual value; $f(x_i)$ represents the predicted value of the model; $N$ represents the number of model samples.

## EXAMPLE ANALYSIS

### Data Description

In order to verify the effectiveness of the prediction model constructed in this article, 986 residential users were selected as case studies from actual horizontal projects. In order to fully analyze user electricity consumption behavior, this article selects weekend electricity load data as the model dataset. Each user dataset includes user electricity load data and user electricity awareness data. Among them, user load data is recorded every 30 minutes, with 48 data points per day for each user, and there is a significant difference in the magnitude of user load data. Therefore, in the clustering stage, the normalized user load dataset is taken as the research object. The energy-saving awareness data of each user is a vector data of 1 row and 5 columns.

### Residential user clustering results

Before clustering users, it is necessary to determine the number of aggregated cluster classes. This article uses the elbow method to determine the appropriate number of clusters, and the results are shown in Figure 7. From the graph, it can be seen that the curve shows a sharp downward trend when the number of aggregated clusters is between 2 and 4. When the number of clusters exceeds 4, the curve shows a slow downward trend. Therefore, the optimal number of set clusters determined by the elbow method is 4.

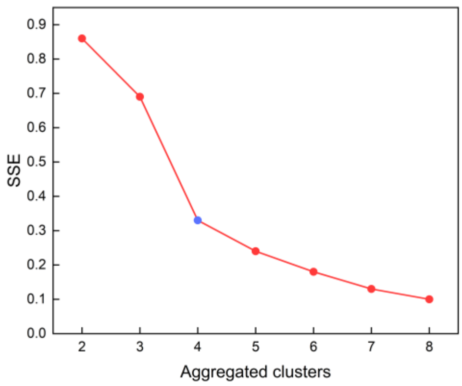



Figure 7. Elbow Method Results

The clustering results of each user group are shown in Figure 8. From Figure 8, it can be seen that users are divided into four distinct categories: double peak users, flat peak users, inverted peak users, and evening peak users. Among them, the number of inverted peak users is the lowest. The load curve of dual peak users shows two significant peak periods, occurring between 8:00-10:00 and 17:30-20:30, respectively. The electricity load power of peak users is relatively stable, and the electricity consumption period is mainly concentrated from 8:30 to 19:30. The characteristic of this type of users is that they do not reduce their electricity consumption during the lunch break period. The peak electricity consumption period of inverted users is different from other types of users. The number of this type of users is smaller than that of other types of users. They are in a low electricity consumption period during the day and their electricity consumption increases from midnight to early morning. The electricity consumption behavior of evening peak users is similar to that of double peak users. These users also have two peak electricity consumption periods, but the electricity consumption of evening peak users is more concentrated in the evening.

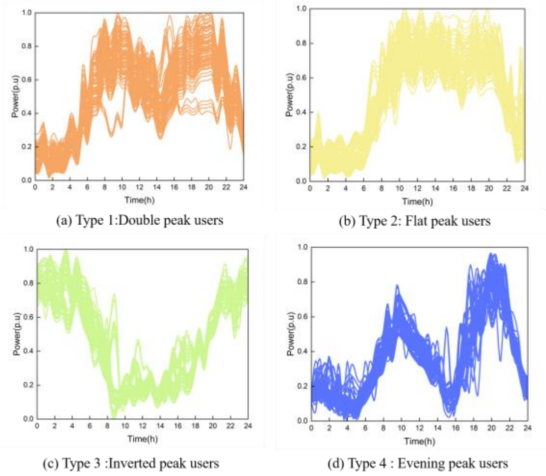



Figure 8. Upper-level clustering results graph

The upper-level clustering divides users into four categories with different electricity consumption behaviors. Based on the results of the upper-level clustering, the lower-level clustering is carried out, and the clustering results are shown in Figure 9. From Figure 9, it can be seen that each user category is further divided into two subcategories in the lower-level clustering: strong energy-saving conscious users and weak energy-saving conscious users. Although there are users with similar electricity consumption behaviors, their energy-saving awareness shows certain differences. Overall, the resident user group shows a strong awareness of energy conservation. After double-layer clustering, residential users are divided into 8 categories.

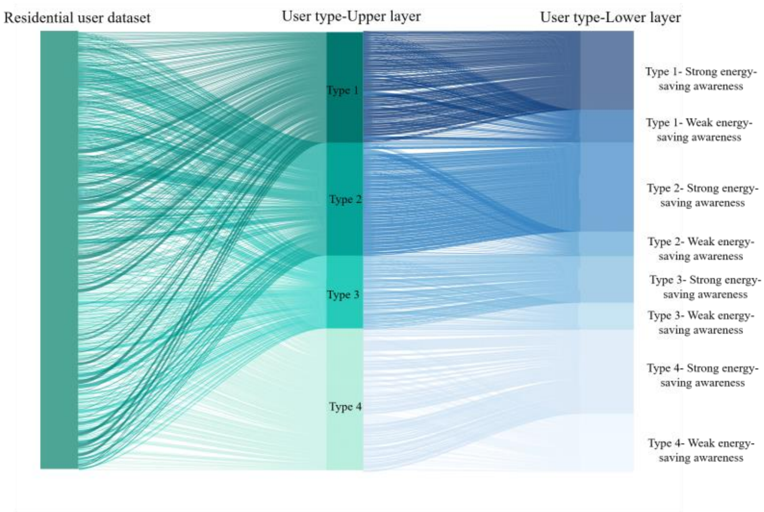



Figure 9. Importance ranking of feature variables

In order to better reflect the relationship between energy-saving awareness and user response, this article defines a response rate indicator. Response rate refers to the ratio of the response completed by users at the response time on the response day to the electricity load at the same time on the day before the response day. The comparison chart of response rates for different types of users is shown in Figure 10. The expression is as follows:

$$c_{a,b,t_{dr}} = \frac{\sum_{t\in T_{dr}} d_{i,j,t}^{a,b}}{\sum_{t\in T_{dr}} p_{i,j-1,t}^{a,b}}, a = 1,2,\dots,k \tag{21}$$

In the formula, $b$ represents two types of users with strong energy-saving awareness and weak energy-saving awareness; $c_{a,b,t_{dr}}$ represents the response rate of users with energy-saving awareness type $b$ among type $a$ users at response time $t_{dr}$; $d_{i,j,t}^{a,b}$ represents the response amount of user with energy-saving awareness type $b$ and ID $i$ in type $a$ at time $t$ on the $j$-th day; $p_{i,j-1,t}^{a,b}$ represents the response amount of user with energy-saving awareness type $b$ and ID $i$ in type $a$ at time $t$ on the $(j-1)$-th day.

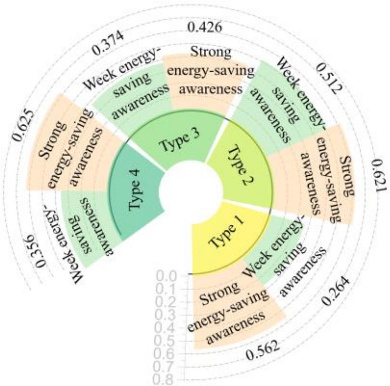



Figure 10. Relationship between energy-saving awareness and response rate

From the graph, it can be seen that among various user types, users with strong energy-saving awareness exhibit higher response rates than users with weaker energy-saving awareness. This indicates that although users belonging to the same type of electricity consumption behavior have varied degrees of energy-saving awareness, their response behavior shows significant differences. This indicates the need for more detailed classification of users of the same electricity type to improve the accuracy of model predictions.

**Comparison of clustering algorithm effectiveness**

On the basis of the original K-Medoids clustering algorithm, this article adds the DTW algorithm and proposes the DTW-K-Medoids clustering algorithm to improve the shortcomings of the original K-Medoids clustering algorithm. This article compares the clustering performance of the DTW-K-Medoids clustering algorithm with the classic clustering algorithms K-Means and K-Medoids clustering algorithms, in order to analyze the superiority of the improved algorithm.

In order to better display the differences in clustering performance between clustering algorithms in the form of images, Multidimensional Scaling (MDS) is used to process the dataset, mapping the 48-dimensional user load data to a 2D space (H. Wu et al., 2022). By processing the data, a point in the two-dimensional graph can be used to represent a residential user, which includes the user's 48-dimensional electricity load curve data. The comparison between DTW-K-Medoids clustering and K-Means clustering is shown in Figure 11.

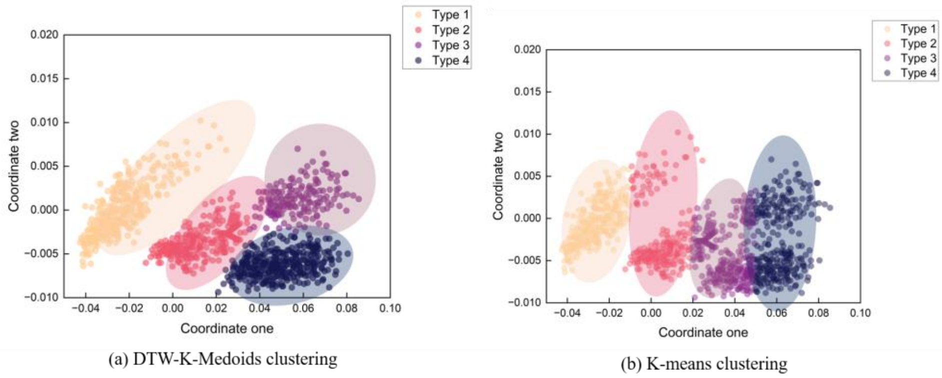



Figure 11. Comparison of DTW-K-Medoids clustering and K-Means clustering effects

From Figure 11, it can be observed that the DTW-K-Medoids clustering algorithm and K-means clustering algorithm maintain consistency in the number of clusters, but there is a significant difference in their classification results. Through comparative analysis, it was found that the clustering results of the DTW-K-Medoids clustering algorithm are more reasonable, and data points with similar distances are classified as the same type. In contrast, the K-means clustering algorithm is more susceptible to the influence of noise points. In Figure 11(b), we can observe that the K-means clustering algorithm is affected by noise points (0.012, 0.011), (0.016, 0.0091), etc., resulting in distorted clustering results for user type 2. In addition, based on the K-means clustering principle, the contour shape of its clustering results tends to be circular, making the clustering results of certain data points appear unreasonable and unable to effectively explore the potential information between data points.

Section 2.5 of this article constructs an evaluation system that includes multiple clustering algorithm evaluation indicators, aiming to comprehensively evaluate the clustering effect of the model. The performance of various clustering algorithms under each indicator is shown in Table 2. From the perspective of DBI indicators, the DTW-K-Medoids clustering algorithm proposed in this article has the smallest DBI value, indicating that the algorithm clusters the closest intra cluster samples after clustering; From the perspective of CH index, the DTW-K-Medoids clustering algorithm has the highest CH value, indicating that the variance between clusters is relatively large, and the distance between clusters is far, which can effectively distinguish between clusters; In terms of model computation time, due to the complex principle of the DTW-K-Medoids clustering algorithm, it requires the longest computation time. On the contrary, the K-means clustering algorithm requires the shortest computation time. Taking into account various factors, it can be concluded that the DTW-K-Medoids clustering algorithm proposed in this article performs the best in terms of overall performance.

Table 2. Performance of Various Indicators of Different Clustering Algorithms

| Types of clustering algorithms | DBI | CH | Calculation time/s |
|---|---|---|---|
| DTW-K-Medoids | 0.587 | 61.325 | 24.69 |
| K-Medoids | 0.646 | 54.326 | 20.37 |
| K-means | 0.712 | 36.785 | 1.56 |

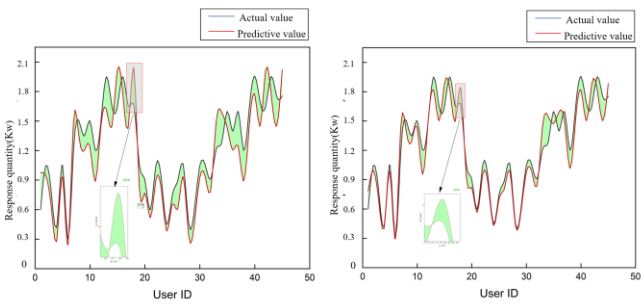
**Prediction model for residential user response volume**

This article uses the Python 3.9 platform and its sklearn function package to construct a prediction model. During the construction of the Bi-LSTM model, some parameters need to be set, and the specific parameter settings are shown in Table 3.

Table 3. Bi-LSTM Model Parameter Setting Values

| Parameter | Numerical value |
|---|---|
| Number of hidden layers | 2 |
| Output power savings | 1 |
| Loss rate | 0.2 |
| loss function | MSE |
| Activation function | Sigmoid |
| Learning rate | 0.1 |
| The number of neurons in the hidden layer | 128 |
| Number of hidden layer two neurons | 64 |

According to the clustering results in section 4.2, users are subdivided into 8 categories, and predictive models are constructed for each category of users in sequence. The training set of each model consists of the load curve and response dataset of 75% of users who have participated in the response in that category, while the test set consists of the electricity load curve data and response dataset of 25% of users who have participated in the response. In order to highlight the superiority of the Bi-LSTM model, this article also constructed an LSTM model and compared the prediction results of the two models. This article randomly selected 246 users from 986 users as the test set to evaluate the effectiveness of the model. Due to the large amount of data in the test set, 45 users were randomly selected and the prediction accuracy of different prediction models was compared and analyzed. The comparison chart of the prediction effect of the prediction model is shown in Figure 12.



(a) Comparison between LSTM predicted values and actual values
(b) Comparison between Bi-LSTM predicted values and actual values

Figure 12. Comparison of Different Prediction Models

By comparing and analyzing Figure 12(a) and Figure 12(b), it can be observed that the error between the predicted and true values of the Bi-LSTM model is generally smaller than that of the LSTM model, and the overall error rate has decreased by 9.5 percentage points. This improvement will enhance the market competitiveness of LA. Based on the Bi-LSTM response prediction model, LA can more accurately provide personalized energy management services and preferential schemes to users. This can enhance the competitiveness of LA in the market, attracting more users to join LA services.

After further zooming in on the image, we found that the LSTM model had a significant error in predicting the response of the 18th resident user, while the Bi-LSTM model predicted a value closer to the actual value for that user. Therefore, an in-depth analysis was conducted on the 18th user, with a focus on the potential relationship between their electricity load curve during the response period and the model's predicted values, as shown in Figure 13.

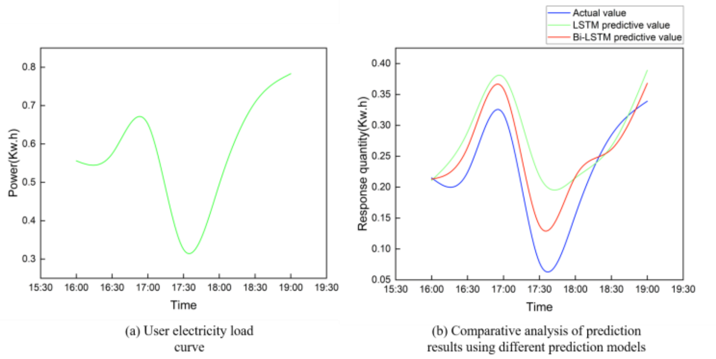



Figure 13. Analysis of Electricity Consumption Behavior of User No.18

Figure 13(a) shows that the user's electricity consumption showed a slight increase from 16:00 to 17:00, but suddenly decreased between 17:00 and 17:30, followed by a rapid increase in electricity consumption. The trend of the user's actual response curve is similar to that of the electricity load curve. From the analysis of Figure 13(b), it can be found that during the time period of 17:30, the LSTM model has the most significant error between the predicted values and the true values, with an error value of 0.14Kw.h. However, the Bi-LSTM model has reduced its prediction error value to 0.06 Kw.h during this period, and the Bi-LSTM model has a better prediction effect. Compared to the LSTM model, the Bi-LSTM model introduces the function of bidirectional information flow, which not only focuses on historical data before the current time step, but also focuses on future data after the current time step. The electricity load of user 18 showed an upward trend before 5:30, but during the time period of 5:30, the electricity load suddenly decreased, and the LSTM model was unable to effectively mine the time series information, resulting in the maximum prediction error value of the LSTM model during this period. However, due to the introduction of bidirectional information flow in the Bi-LSTM model, it can more comprehensively capture the dependency relationships in sequence data, which helps to improve the accuracy of prediction.

**Analysis of Evaluation Expenses for LA**

The LA is entrusted by the proxy user to declare the response volume on behalf of the user in the DR market. If the LA declares a response potential that exceeds the user's own, the LA will be punished for not being able to complete the declared response volume. This article calculates assessment fees based on the DR effectiveness evaluation standards and reward and punishment mechanisms issued by Guizhou Province, China, and calculates the profit changes of LA using the predictive model constructed in this article and other predictive models (Guizhou Provincial Energy Bureau, China, 2023). When the actual response volume of the users represented by the LA is lower than the bid winning response volume, the grid side will charge a response assessment fee to the LA, and the cost calculation formula is shown in formula 22:

$$C_t = (R_1 - R_2) * p * M \tag{22}$$

In the formula, $C_t$ represents the assessment fee that the LA should pay at the moment $t$; $R_1$ represents the amount to be completed; $R_2$ represents the actual response amount; $p$ represents the clearing price; $M$ represents the penalty factor, which defaults to 0.5 in the file.

This article calculates the difference in assessment costs between different models based on this formula, as shown in the Table 4. Through calculation, it was found that compared to other models, the LA, based on the prediction results of the model proposed in this article, delivered the least assessment fee, with a maximum reduction of 498.91 CNY. The predictive model constructed in this article can help LA better understand the changing trends and risks of user electricity consumption behavior, and timely identify potential energy supply-demand imbalances. By timely adjusting energy supply and demand, LA can reduce market transaction risks and reduce energy operating costs.

Table 4. Physical meaning and direction of indicators

| Model type | MSE | Assessment fee (CNY) |
|---|---|---|
| A predictive model based on single-layer clustering and LSTM | 9.74 | 1856.36 |
| A predictive model based on double-layer clustering and LSTM | 7.85 | 1524.58 |
| A predictive model based on double-layer clustering and Bi-LSTM | 5.94 | 1357.45 |

## CONCLUSION

Against the backdrop of the continuous popularization of DR, diversification of participating entities will become a future trend. Under this trend, the position of LA in the electricity market will be further enhanced. As the number of users participating in DR increases, LA face the challenge of accurately assessing the response potential of customers who have not participated in DR. Therefore, this article constructs a response potential prediction model that combines clustering algorithm and Bi-LSTM algorithm to improve the predictive ability of LA for customer response potential.

This article collects actual case data and evaluates the effectiveness of the model through examples. The following conclusions are drawn:

1. The existing references mainly conduct single-layer clustering based on user electricity load data. However, this article adopts a double-layer clustering model, using user electricity load data and energy-saving awareness as the clustering basis. Case analysis shows that clustering users based solely on electricity load data will affect the prediction accuracy of the model. Among users with similar electricity consumption behaviors, those with strong energy-saving awareness showed higher response rates than those with weak energy-saving awareness. Users with different energy-saving awareness showed significant differences in their response behaviors. Therefore, using user electricity load data and energy-saving awareness as clustering criteria will improve the accuracy of model prediction.

2. In order to improve the accuracy of clustering, this article introduces the DTW algorithm to improve the K-Medoids algorithm based on its shortcomings. Through case analysis, it was found that although the DTW-K-Medoids algorithm has the longest running time, it performs the best on the DBI and CH metrics compared to other models. This indicates that the DTW-K-Medoids algorithm has the closest intra cluster samples after clustering, and the distance between clusters is relatively far, which can effectively distinguish between different cluster classes. Taking into account various factors, the DTW-K-Medoids algorithm proposed in this article performs the best in terms of overall performance.

3. This article constructs a Bi-LSTM prediction model based on the LSTM model. The Bi-LSTM prediction model introduces the function of bidirectional information flow, which can deeply mine the information of time series data. By using this model, the LA can more precisely provide users favored schemes and individualized energy management services, increasing their market competitiveness and lowering potential risks.

## REFRENCES

[1] Afzalan, M., & Jazizadeh, F. (2019). Residential loads flexibility potential for demand response using energy consumption patterns and user segments. Applied Energy, 254, 113693. https://doi.org/10.1016/j.apenergy.2019.113693

[2] Ballenger, A. P., Herath, P., Caceres, M. N., Venayagamoorthy, G. K., & Corchado, J. M. (2017). Influencing behavior of electricity consumers to enhance participation in demand response. 2017 North American Power Symposium (NAPS), 1–6. https://doi.org/10.1109/NAPS.2017.8107300

[3] Chen, Z., Ma, M., Li, T., Wang, H., & Li, C. (2023). Long sequence time-series forecasting with deep learning: A survey. Information Fusion, 97, 101819. https://doi.org/10.1016/j.inffus.2023.101819

[4] Dao, F., Zeng, Y., & Qian, J. (2024). Fault diagnosis of hydro-turbine via the incorporation of bayesian algorithm optimized CNN-LSTM neural network. Energy, 290, 130326. https://doi.org/10.1016/j.energy.2024.130326

[5] Deng, N., Wang, B., He, L., Liu, J., & Wang, Z. (2023). Does electricity price reduction bring a sustainable development of business: Evidence from fine-grained industrial electricity consumption data in China. Journal of Environmental Management, 335, 117522. https://doi.org/10.1016/j.jenvman.2023.117522

[6] Duman, A. C., Gönül, Ö., Erden, H. S., & Güler, Ö. (2023). Survey- and simulation-based analysis of residential demand response: Appliance use behavior, electricity tariffs, home energy management systems. Sustainable Cities and Society, 96, 104628. https://doi.org/10.1016/j.scs.2023.104628

[7] El Amouri, H., Lampert, T., Gançarski, P., & Mallet, C. (2023). Constrained DTW preserving shapelets for explainable time-series clustering. Pattern Recognition, 143, 109804. https://doi.org/10.1016/j.patcog.2023.109804

[8] Fang, M., Zhang, F., Yang, Y., Tao, R., Xiao, R., & Zhu, D. (2024). The influence of optimization algorithm on the signal prediction accuracy of VMD-LSTM for the pumped storage hydropower unit. Journal of Energy Storage, 78, 110187. https://doi.org/10.1016/j.est.2023.110187

[9] Guizhou Provincial Energy Bureau, China. (2023). Notice of Guizhou Provincial Energy Bureau on Issuing the Implementation Plan for Guizhou Province's Electricity Demand Response (Trial).

[10] https://nyj.guizhou.gov.cn/zwgk/gzhgfxwjsjk/gfxwjsjk/202307/t20230711_80838290.html

[11] Guo, D., Jiang, K., Xu, C., & Yang, X. (2022). Industrial clustering, income and inequality in rural China. World Development, 154, 105878. https://doi.org/10.1016/j.worlddev.2022.105878

[12] Kong, X., Kong, D., Yao, J., Bai, L., & Xiao, J. (2020). Online pricing of demand response based on long short-term memory and reinforcement learning. Applied Energy, 271, 114945. https://doi.org/10.1016/j.apenergy.2020.114945

[13] Kong, X., Wang, Z., Liu, C., Zhang, D., & Gao, H. (2023). Refined peak shaving potential assessment and differentiated decision-making method for user load in virtual power plants. Applied Energy, 334, 120609. https://doi.org/10.1016/j.apenergy.2022.120609

[14] Lazzari, F., Mor, G., Cipriano, J., Gabaldon, E., Grillone, B., Chemisana, D., & Solsona, F. (2022). User behaviour models to forecast electricity consumption of residential customers based on smart metering data. Energy Reports, 8, 3680–3691. https://doi.org/10.1016/j.egyr.2022.02.260

[15] Lee, C. K. H., & Leung, E. K. H. (2023). Spatiotemporal analysis of bike-share demand using DTW-based clustering and predictive analytics. Transportation Research Part E: Logistics and Transportation Review, 180, 103361. https://doi.org/10.1016/j.tre.2023.103361

[16] Li, K., Li, Z., Huang, C., & Ai, Q. (2024). Online transfer learning-based residential demand response potential forecasting for load aggregator. Applied Energy, 358, 122631. https://doi.org/10.1016/j.apenergy.2024.122631

[17] Lu, R., Bai, R., Huang, Y., Li, Y., Jiang, J., & Ding, Y. (2021). Data-driven real-time price-based demand response for industrial facilities energy management. Applied Energy, 283, 116291. https://doi.org/10.1016/j.apenergy.2020.116291

[18] Melnykov, V., & Michael, S. (2020). Clustering Large Datasets by Merging K-Means Solutions. Journal of Classification, 37(1), 97–123. https://doi.org/10.1007/s00357-019-09314-8

[19] National Development and Reform Commission. (2021). Notice on Encouraging Renewable Energy Power Generation Enterprises to Build or Purchase Peak shaving Capacity to Increase Grid Connection Scale.

[20] https://www.beijing.gov.cn/zhengce/zhengcefagui/qtwj/202204/t20220418_2679836.html

[21] Shekari, M., Arasteh, H., Sheikhi Fini, A., & Vahidinasab, V. (2021). Demand Response Requirements from the Cultural, Social, and Behavioral Perspectives. Applied Sciences, 11(23), 11456. https://doi.org/10.3390/app112311456

[22] Shi, R., & Jiao, Z. (2023). Individual household demand response potential evaluation and identification based on machine learning algorithms. Energy, 266, 126505. https://doi.org/10.1016/j.energy.2022.126505

[23] Shirsat, A., & Tang, W. (2021). Quantifying residential demand response potential using a mixture density recurrent neural network. International Journal of Electrical Power & Energy Systems, 130, 106853. https://doi.org/10.1016/j.ijepes.2021.106853

[24] Sieranoja, S., & Fränti, P. (2022). Adapting k-means for graph clustering. Knowledge and Information Systems, 64(1), 115–142. https://doi.org/10.1007/s10115-021-01623-y

[25] Sobrinho Campolina Martins, A., Ramos De Araujo, L., & Rosana Ribeiro Penido, D. (2024). K-Medoids clustering applications for high-dimensionality multiphase probabilistic power flow. International Journal of Electrical Power & Energy Systems, 157, 109861. https://doi.org/10.1016/j.ijepes.2024.109861

[26] Sridhar, A., Honkapuro, S., Ruiz, F., Stoklasa, J., Annala, S., Wolff, A., & Rautiainen, A. (2023). Residential consumer preferences to demand response: Analysis of different motivators to enroll in direct load control demand response. Energy Policy, 173, 113420. https://doi.org/10.1016/j.enpol.2023.113420

[27] Wang, B., Cai, Q., & Sun, Z. (2020). Determinants of Willingness to Participate in Urban Incentive-Based Energy Demand-Side Response: An Empirical Micro-Data Analysis. Sustainability, 12(19), 8052. https://doi.org/10.3390/su12198052

[28] Wang, F., Li, K., Liu, C., Mi, Z., Shafie-Khah, M., & Catalão, J. P. S. (n.d.). Synchronous Pattern Matching Principle-Based Residential Demand Response Baseline Estimation: Mechanism Analysis and Approach Description.

[29] Wang, J., Qiu, R., Xu, B., Wu, H., Tang, L., Zhang, M., & Ding, M. (2023). Aggregated large-scale air-conditioning load: Modeling and response capability evaluation of virtual generator units. Energy, 276, 127570. https://doi.org/10.1016/j.energy.2023.127570

[30] Wang, K., Li, K., Zhou, L., Hu, Y., Cheng, Z., Liu, J., & Chen, C. (2019). Multiple convolutional neural networks for multivariate time series prediction. Neurocomputing, 360, 107–119. https://doi.org/10.1016/j.neucom.2019.05.023

[31] Wang, Z., Li, H., Deng, N., Cheng, K., Lu, B., Zhang, B., & Wang, B. (2020). How to effectively implement an incentive-based residential electricity demand response policy? Experience from large-scale trials and matching questionnaires. Energy Policy, 141, 111450. https://doi.org/10.1016/j.enpol.2020.111450

[32] Wu, H., Han, Y., Geng, Z., Fan, J., & Xu, W. (2022). Production capacity assessment and carbon reduction of industrial processes based on novel radial basis function integrating multi-dimensional scaling. Sustainable Energy Technologies and Assessments, 49, 101734. https://doi.org/10.1016/j.seta.2021.101734

[33] Xu, Q., Lin, L., & Liang, Q.-M. (2023). What influences industrial enterprises' willingness of demand response: A survey in Qinghai, China. Journal of Cleaner Production, 428, 139483. https://doi.org/10.1016/j.jclepro.2023.139483

[34] Yan, N., Yao, Y.-B., Jia, Z.-D., Liu, L., Dai, C.-T., Li, Z.-G., Zhang, Z.-H., Li, W., Wang, L., & Wang, P.-F. (2022). Online battery health diagnosis for electric vehicles based on DTW-XGBoost. Energy Reports, 8, 121–128. https://doi.org/10.1016/j.egyr.2022.09.126

[35] Yu, H., Zhou, C., Bao, J., Yao, H., Yan, P., & Wang, Q. (2023). Analysis and Effect Evaluation of Offshore Wind Power Output Characteristics Based on Gaussian Mixed Clustering. Procedia Computer Science, 224, 389–394. https://doi.org/10.1016/j.procs.2023.09.053

[36] Zhang, Y., Ai, Q., & Li, Z. (2020). ADMM-based distributed response quantity estimation: A probabilistic perspective. IET Generation, Transmission & Distribution, 14(26), 6594–6602. https://doi.org/10.1049/iet-gtd.2020.1380

[37] Zhou, X., Liu, X., Liu, H., Ji, Z., & Li, F. (2024). Optimal dispatching strategy for residential demand response considering load participation. Global Energy Interconnection, 7(1), 38–47. https://doi.org/10.1016/j.gloei.2024.01.004.