This article has been accepted for publication in IEEE Transactions on Industry Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIA.2022.3200352

Adaptive Optimal Greedy Clustering-based Monthly Electricity Consumption Forecasting Method

Yuqing Wang, Ziyang Fu, Fei Wang, *Senior Member, IEEE*, Kangping Li, Zhenghui Li, Zhao Zhen, Payman Dehghanian, *Senior Member, IEEE*, Mahmud Fotuhi-Firuzabad, *Fellow, IEEE* and João P. S. Catalão, *Fellow, IEEE*

***Abstract*—Accurate monthly electricity consumption forecasting (MECF) is important for electricity retailers to mitigate trading risks in the electricity market. Clustering is commonly used to improve the accuracy of MECF. However, in the existing clustering-based forecasting methods, clustering and forecasting are independently performed and lack coordination, which limits the further improvement of forecasting accuracy. To address this issue, an adaptive optimal greedy clustering-based MECF method is proposed in this paper. Firstly, a metric of predictability is defined based on the goodness of fit and cluster's average electricity consumption. Under a pre-defined cluster number, the greedy clustering algorithm achieves the optimal division of individuals with the goal of maximizing predictability. Then, an adaptive method is designed to select the optimal cluster number from a variety of clustering scenarios according to the prediction accuracy on the validation dataset. The effectiveness and superiority of the proposed method have been verified on a real-world dataset.**

Keywords—Monthly electricity consumption forecasting; Electricity retailer; Greedy clustering; Predictability

NOMENCLATURE

A. Sets and Indices

i, j, k	Index for clusters
m	Index for substations
t	Index for timeslot
R	Index for the number of clusters updating

This work was supported by the National Key R&D Program of China (2018YFE0122200) and the Science & Technology Project of State Grid Hebei Electric Power Co., Ltd (SGHEYX00SCJS2000037). *(Corresponding author: Fei Wang)*

Yuqing Wang, Fei Wang and Zhao Zhen are with the Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China; Fei Wang is also with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China, and also with the Hebei Key Laboratory of Distributed Energy Storage and Microgrid (North China Electric Power University), Baoding 071003, China (e-mail: wangyuqing@ncepu.edu.cn; feiwang@ncepu.edu.cn).

Zhiyang Fu is with the Department of Mathematics and Physics, North China Electric Power University, Baoding 071003, China (e-mail: fzy@ncepu.edu.cn).

Kangping Li is with College of Smart Energy, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: kangpingli@sjtu.edu.cn).

Zhenghui Li is with State Grid Anhui Electric Power Company, Bengbu Power Supply Company, Bengbu 233000, China (e-mail: lizhenghui960128@163.com).

Payman Dehghanian is with the Department of Electrical and Computer Engineering, The George Washington University, Washington, DC, USA (e-mail: payman@gwu.edu)

Mahmud Fotuhi-Firuzabad is with the Center of Excellence in Power System Management and Control, Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran (e-mail: fotuhi@sharif.edu)

J.P.S. Catalão is with the Faculty of Engineering of the University of Porto, and INESC TEC, 4200-465 Porto, Portugal (e-mail: catalao@fe.up.pt).

S	Set of substations in a region
C	Set of clusters
S^{se}	Set of selected substations during clustering
S^{re}	Set of remaining substations during clustering

B. Parameters

K	The number of clusters
N	The number of substations
M	The number of clusters updating
T	The number of timeslots
s_i	The substation i
c_j	The cluster j
c_j^{temp}	The temporary cluster j
$d_m(t)$	The real electricity consumption data of the m-th substation at time t
$\hat{d}_m(t)$	The forecasting electricity consumption data of the m-th substation at time t
$\boldsymbol{d}_{c_j}$	The aggregated electricity consumption series of each element from the cluster j.
$\bar{d}_{c_i}$	The average of $\boldsymbol{d}_{c_j}$
$\mathbf{D}$	The electricity consumption matrix of substations
$\mathbf{D}^{tr}$	The electricity consumption matrix for training
$\mathbf{D}^{va}$	The electricity consumption matrix for tuning
$\mathbf{D}^{te}$	The electricity consumption matrix for testing
p_j	The predictability of cluster j
ϕ_{c_i}	The goodness of fit of the forecasting algorithm on the cluster j
λ	The optimal placement of cluster label
$K_{\min}, K_{\max}$	The upper and lower boundaries of the selection range of cluster number
$K_{optimal}$	The optimal cluster number
$\boldsymbol{X}, \boldsymbol{Y}$	The input and output matrix of the forecasting model

I. INTRODUCTION

With the deregulation of the electricity market, an increasing number of electricity retailers are being to directly participate in electricity trading. Monthly electricity consumption forecasting (MECF) can help the electricity retailers to master the customers' medium and long term electricity consumption [1]. In this way, the electricity retailers can make reasonable purchasing decisions to trade in the wholesale market [2] and mitigate the potential price risk caused by the huge fluctuations in the spot market [3]. Meanwhile, accurate MECF can also guide the planning department to reasonably arrange the mid-term operation and maintenance plan, reduce the cost of power supply and improve the reliability of the power grid [4].

In recent years, many methods have been applied to MECF. From the perspective of forecasting algorithms, these methods can be classified into two categories: statistical methods and artificial intelligence (AI) methods [5]. The statistical methods mainly include autoregressive integrated moving average (ARIMA) [6], the grey model [7], etc. The AI methods mainly include artificial neural network (ANN) [8], support vector regression (SVR) [9] and recurrent neural network (RNN) [10], etc. From the perspective of spatial scale, it can be divided into regional forecasting which achieves the aggregated future electricity consumption of a region, and individual forecasting which achieves electricity consumption of a lower aggregated level [11] (e.g., substation level, feeder level, and customer level). Generally, there are three ways to achieve the regional (aggregated) electricity consumption forecasting. The first way is to aggregate the electricity consumption data of each lower aggregated level object in this region to construct a regional level electricity consumption time series. Then, forecasting algorithms are performed on the regional time series to obtain the future regional electricity consumption [12]-[13]. However, simply aggregating the lower level data cannot make full use of the spatial correlation information [14]. Some papers have investigated the second way, forecasting the electricity consumption data of each lower level object separately, and then aggregating each forecasting result to obtain the final regional electricity consumption [15]. However, due to the randomness and volatility of these individual series [16], this approach cannot achieve an ideal forecasting result.

To effectively utilize spatial correlation information and reduce the negative effect of randomness and volatility of individual series on electricity consumption forecasting, some existing studies investigate the adoption of the third way, the clustering-based forecasting methods. In these methods, the clustering algorithm is adopted to group all the individuals into several clusters first, then the forecasting algorithm is performed on each cluster separately to obtain forecasting results of each cluster, and eventually, the forecasting results of each cluster will be aggregated to obtain the final forecasting result. In Ref. [17], the individuals with the same load characteristic were grouped together by K-means and separate forecasting models were constructed based on each cluster. In order to make full use of the data in different clustering scenarios, an ensemble forecasting model was proposed in Ref. [18]. By combining the forecasting results under multiply clustering numbers, the method shows better forecasting performance than all clustering scenarios. It is worth noting that when the number of individuals meets a certain threshold, the clustering-based forecasting methods can achieve ideal forecasting results [19]-[20]. However, in existing clustering-based forecasting methods, the two steps of clustering and forecasting are executed separately and lack a direct linkage. The reason is that most of the current clustering algorithms, including K-means [21], hierarchical clustering [22], and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), take maximizing the intra-cluster similarity and minimizing the inter-cluster similarity under a predefined cluster number as the clustering objective function [23], but not the improvement of the final forecasting accuracy. In this way, forecasting and clustering are considered as two independent steps without coordination. Therefore, the improvement of the forecasting accuracy achieved by existing clustering based forecasting methods is restricted. To deal with this limitation, a data-driven linear clustering method is proposed in Ref. [24], which aims to make clusters' load curves smoother by putting complementary individual load curves into the same cluster. In this way, the prediction accuracy of the linear load forecasting model can be improved. A closed-loop clustering algorithm is proposed in Ref. [25]. In this algorithm, the K-means is used to generate initialized clusters and the forecasting models are built based on each cluster. Then, each individual is tested on all forecasting models to determine the optimal position according to the best prediction accuracy. However, this algorithm takes into account the effect of clusters on the accuracy of individual predictions, rather than the individual contribution to the predictability of clusters. This may cause the effect of clusters update to be limited by the initial partition of the cluster. At the same time, the influence of different individual load levels on the final overall predictability was not considered.

To solve the above limitations, a greedy clustering-based electricity consumption forecasting model is proposed which can directly achieve joint optimization of clustering and forecasting. As shown in Fig. 1, in the greedy clustering algorithm, firstly, K lower aggregated level objects are selected as the initial center of the K clusters. Then, ordinally allocate the remaining objects s_m into the above K clusters according to certain criteria. The criteria is grouping the object into one of the clusters where it can achieve greater improvement of the predictability $\boldsymbol{p}$. For each object, only its optimal allocation in the current situation is taken into consideration. Thus, the optimal partition of all individuals under a predefined K will be completed and the consistency of clustering target and forecasting target can be achieved.

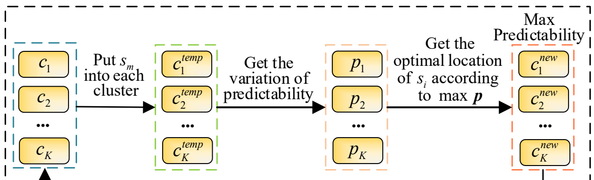


Fig. 1. The general process of greedy clustering for each object.

In addition, the clustering algorithms are usually applied under a predefined cluster number K, so it is crucial to provide an appropriate cluster number to achieve the optimal partition of individuals [26]. In order to deal with this difficulty of choosing cluster number, some performance metrics, e.g., Davies-Bouldin index (DBI), Ratio of the within-cluster sum of squares to between cluster variation (WCBCR), were used to find the appropriate cluster number [27]. However, these metrics reflecting the intra-cluster compactness and inter-cluster separation are designed to optimize the conventional clustering algorithms, which also aim to maximize the similarity of each cluster but not the predictability. To address this problem, instead of determining the cluster number directly, a range of cluster number K which may include the optimal cluster number is set to generate multiple clustering scenarios. Then, the optimal cluster number is adaptively selected based on the MECF accuracy of various clustering scenarios, which is performed on the validation set.

The contributions of this paper are summarized as follows:

1) A greedy clustering-based MECF method is proposed in this paper, which takes the prediction accuracy as the guide to adjusting clustering by iteration. In this way, the proposed method can coordinate clustering and forecasting well, thus significantly improving the accuracy of MECF.

2) An adaptive optimal cluster number selection method is designed for the proposed greedy clustering, which chooses the optimal cluster number according to the forecasting accuracy under multiple clustering scenarios. In this way, the difficulty of pre-defining the cluster number for greedy clustering can be avoided.

This paper is an extended version of our conference paper [28]. The major revisions have been made as follows: 1) The abstract and introduction have been rewritten to make the innovation clearer. 2) The case study has been improved. The impacts of initialization strategy and dataset partition on the performance of adaptive optimal greedy clustering-based forecasting are explored.

The rest of this paper is organized as follows. The greedy clustering-based forecasting method is proposed in Section II. The verification of this method on the real-world data is presented in Section III. The impact of two factors on the performance of the proposed method is analyzed and discussed in Section IV. Section V highlights the conclusion and future work.

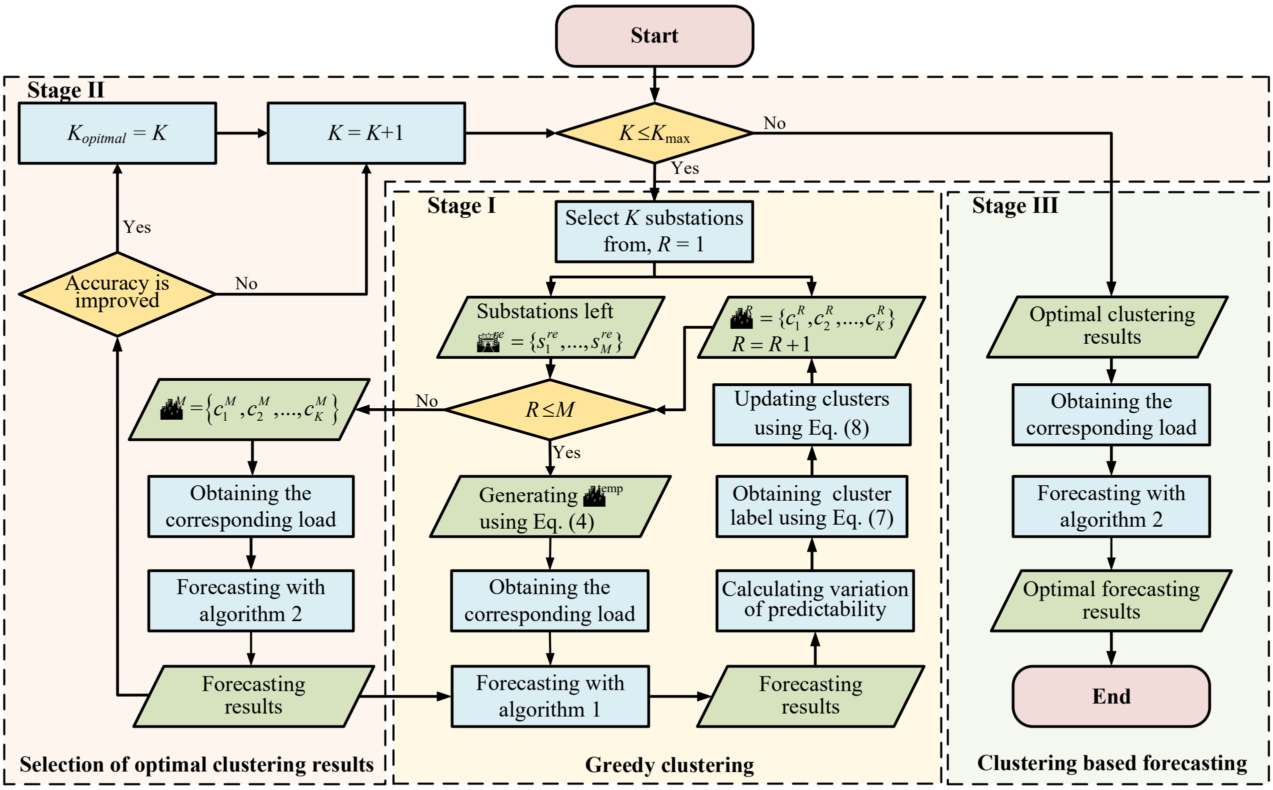


Fig. 2. The flowchart of greedy clustering-based forecasting method.

II. Methodology

This paper proposes a greedy clustering-based electricity consumption forecasting method illustrated in Fig. 2, which aligns the clustering target with the improvement of forecasting accuracy. Furthermore, the optimal cluster number K can be adaptively selected. The whole process can be divided into three main stages, including the greedy clustering (stage I), the adaptive selection of optimal cluster number (stage II), and the clustering-based forecasting (stage III).

To explain the method more clearly, the substations level data of a region is taken as an example. First, presume that a region contains N substations $S = \{s_1, s_2, \ldots, s_N\}$. $\mathbf{D} \in {}^{N\times T}$ denotes the electricity consumption matrix of N substations in T time length. $d_m(t)$ denotes the electricity consumption data of the m-th substation s_m at time t ($t \in [1,2,\ldots,T]$). Then, the substations data matrix $\mathbf{D}$ is segmented into three parts: $\mathbf{D}^{tr}$, $\mathbf{D}^{va}$, $\mathbf{D}^{te}$ in chronological order for training, tuning and testing the proposed method.

In stage I, the greedy clustering algorithm is applied under a predefined cluster number, which maximizes the predictability of the *Forecasting Algorithm 1(FA1)* on the data of each cluster to improve the overall predictability. Then, in stage II, the *Forecasting Algorithm 2 (FA2)* is applied based on different cluster number K to obtain the corresponding total electricity consumption forecasting results. After that, the optimal clustering results can be adaptively selected among all clustering scenarios. The two stages both are performed on $\mathbf{D}^{tr}$ and $\mathbf{D}^{va}$.

Finally, in stage III, based on the optimal clustering results, the *FA2* will be employed on each cluster. Then, all of the forecasting results obtained from each cluster will be

This article has been accepted for publication in IEEE Transactions on Industry Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIA.2022.3200352

aggregated to get the final regional monthly electricity consumption. This stage is conducted on $\mathbf{D}^{te}$.

A. The Proposed Greedy Clustering

The target of the greedy clustering algorithm is to put the substation $s_i \in \mathrm{S}$ into the cluster $c_j \in \{c_1, c_2, ... c_K\}$ where it can achieve a larger improvement of predictability. The predictability is mainly measured by the goodness of fit of *FA1* on $\mathbf{D}^{va}$. With respect to the goodness of fit, many metrics and techniques have been presented in numerous publications. In this paper, the R^2 is used as the metric of it.

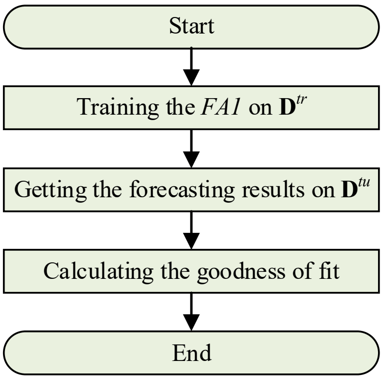


Fig. 3. The process of calculating the goodness of fit.

Algorithm 1: **Greedy clustering algorithm**

Input: The cluster number *K*, all substations $\mathrm{S} = \{s_1, s_2, ..., s_N\}$, electricity consumption data $\{\boldsymbol{d}_1, \boldsymbol{d}_2, ..., \boldsymbol{d}_N\}$

Output: The clusters $\mathrm{C} = \{c_1^M, c_2^M, ..., c_K^M\}$

$\mathrm{S}^{se} = \{s_1^{se}, ..., s_K^{se}\} \leftarrow$ Randomly select *K* substations from S;

$\mathrm{S}^{re} = \{s_1^{re}, ..., s_M^{re}\} \leftarrow$ Remaining substations after selection;

$M = N - K$, $R = 1$, $m = 1$;

for $j \in \{1, ..., K\}$ **do**

  $c_j^0 \leftarrow s_j^{se}$;

end for

for $R \le M$ **do**

  for c_j^{R-1} in C^{R-1} **do**

    $c_j^{temp} = c_j^{R-1} \bigcup s_m^{re}$;

    $\boldsymbol{d}_{c_j^{temp}} = \boldsymbol{d}_{c_j^{R-1}} + \boldsymbol{d}_m$;

    $p_j = (\phi_{c_j^{temp}} - \phi_{c_j^{R-1}}) * \bar{d}_{c_j^{temp}}$;

  end for

  $\lambda = \arg\max_{j \in \{1, ... k\}} \boldsymbol{p}$;

  $c_\lambda^R \leftarrow c_\lambda^{R-1} \bigcup s_m^{re}$;

  $m = m + 1$;

  $R = R + 1$;

end for

The process to calculate the goodness of fit is shown in Fig. 3. Take the consumption data of c_j for example. Firstly, the *FA1* is trained with the data from $\boldsymbol{d}_{c_j}^{tr}$. After obtaining the forecasting results on $\boldsymbol{d}_{c_j}^{va}$, the goodness of fit of the forecasting algorithm on c_j can be calculated according to Eqs. (1) and (2).

$$\phi_{c_j} = \frac{1}{T} \sum_{t=1}^{T} \frac{(\hat{d}_{c_j}^{va}(t) - \bar{d}_{c_j}^{va})^2}{(d_{c_j}^{va}(t) - \bar{d}_{c_j}^{va})^2} \tag{1}$$

$$\bar{d}_{c_j}^{va} = \frac{1}{T} \sum_{t=1}^{T} d_{c_j}^{va}(t) \tag{2}$$

where ϕ_{c_j} denotes the goodness of fit of the forecasting algorithm on c_j; $\boldsymbol{d}_{c_j}$ represents the aggregated electricity consumption series of each element from the *j*-th cluster c_j; $d_{c_j}^{va}(t)$ and $\hat{d}_{c_j}^{va}(t)$ denote the true value and forecasting value of $\boldsymbol{d}_{c_j}^{va}$ at time *t*; $\bar{d}_{c_j}^{va}$ denotes the average of $\boldsymbol{d}_{c_j}^{va}$; *T* represents the length of time series $\boldsymbol{d}_{c_j}^{va}$.

The whole procedures of greedy clustering are presented in **Algorithm 1** and the specific process of greedy clustering is as follows:

1) *Initialization:* K $(1 \le K \le N)$ substations $\mathrm{S}^{se} = \{s_1^{se}, s_2^{se}, ... s_K^{se}\}$ are randomly selected from the substations set S and the remaining substations set $\mathrm{S}^{re} = \{s_1^{re}, s_2^{re}, ..., s_M^{re}\}$ is produced at the same time. S^{re} and S^{se} both satisfy the following constraints:

$$\begin{aligned} \mathrm{S}^{se} \bigcup \mathrm{S}^{re} &= \mathrm{S} \\ \mathrm{S}^{se} \bigcap \mathrm{S}^{re} &= \varnothing \end{aligned} \tag{3}$$

Making the S^{se} as the initial elements of *K* clusters $\mathrm{C}^0 = \{c_1^0, c_2^0, ..., c_K^0\}$, i.e. $s_j^{se} \in c_j^0$. Then, the initial goodness of fit on each cluster can be calculated to get $\boldsymbol{f}^0 = [\phi_{c_1}^0, \phi_{c_2}^0, ..., \phi_{c_K}^0]$ using Eq. (1).

2) *Calculating the variation of predictability:* Arrange s_m^{re} from the remaining substations S^{re} into each cluster $c_j^0 \in \{c_1^0, c_2^0, ..., c_K^0\}$ successively to generate C^{temp} as Eq. (4). And the electricity consumption data $\boldsymbol{d}_{c_j^{temp}}$ of c_j^{temp} can be obtained by Eq. (5)

$$\begin{aligned} C^{temp} = \{ & c_1^{temp} = c_1^0 \cup s_m^{re}, \\ & c_2^{temp} = c_2^0 \cup s_m^{re}, \\ & \dots \\ & c_K^{temp} = c_K^0 \cup s_m^{re} \} \end{aligned} \tag{4}$$

$$\boldsymbol{d}_{c_j^{temp}} = \boldsymbol{d}_{c_j^0} + \boldsymbol{d}_m \tag{5}$$

Then, the new goodness of fit $\boldsymbol{f}^1 = [\phi_{c_1}^1, \phi_{c_2}^1, ..., \phi_{c_K}^1]$ of cluster $c_j^{temp} \in C^{temp}$ can be obtained by using Eq. (1) again. The variation of the goodness of fit before and after clustering s_m^{re} into c_j^0 could be a part of a description of predictability variation. Meanwhile, the cluster with a larger magnitude of electricity consumption contributes more to the final forecasting result. Therefore, in the process of measuring the variation of the predictability, the impact of the magnitude of each cluster's electricity consumption data is considered by multiplying the variation of the goodness of fit with the average electricity consumption of the cluster c_j^{temp}. The variation of predictability $[p_1, p_2, ..., p_K]$ before and after clustering s_m^{re} into $c_j^0 \in C^0$ can be obtained by Eq. (6).

$$p_j = (\phi_{c_j^{temp}} - \phi_{c_j^0}) * \bar{d}_{c_j^{temp}} \tag{6}$$

where $\phi_{c_j^0}$ and $\phi_{c_j^{temp}}$ represents the goodness of fit of consumption data of c_j^0 and c_j^{temp}; c_j^0 represents the j-th cluster of C^0; $\bar{d}_{c_j^{temp}}$ represents the average electricity consumption of c_j^{temp}.

3) *Updating the clusters:* According to step 2), the vector $\boldsymbol{p} = [p_1, ..., p_j, ..., p_K]$ can be calculated to denote the variation of predictability. The clustering target is to group s_m^{re} into the cluster c_λ^0 where s_m^{re} can achieve higher improvement of predictability. Therefore, after the comparison of all the situations (put s_m^{re} into each cluster $\{c_1^0, ..., c_K^0\}$ and calculating the variation of predictability) according to Eq. (7), the corresponding cluster label λ of s_m^{re} can be obtained, and then the cluster c_λ^0 can be updated by using Eq. (8) while other clusters of C^0 keep the same. Thus, once the updating is completed, the new clusters $C^1 = C^0$ can be obtained.

$$\lambda = \underset{j \in \{1,...k\}}{\arg\max}\, \boldsymbol{p} \tag{7}$$

$$c_\lambda^1 \leftarrow c_\lambda^0 \cup s_m^{re} \tag{8}$$

where λ denotes the subscript corresponding to the minimum of $\boldsymbol{p}$.

4) *Termination mechanism:* Repeating steps 2) and 3) $M = N - K$ times until there are no substations left in S^{re}, and then the final clustering result $\{c_1^M, c_2^M, ..., c_K^M\}$, an optimal allocation of all substations under a predefined cluster number K can be obtained. It is considered to be the clustering result with the highest overall predictability under a predefined K.

B. The Adaptive Selection of Optimal Cluster Number

The purpose of greedy clustering is to achieve the maximum predictability of each cluster under a predefined cluster number K, while the selection of the optimal cluster number is to select the K which helps the forecasting method to maximize the forecasting accuracy.

To overcome the key challenge of predefining the optimal cluster number, this stage generates multiple clustering scenarios under different K. Based on these scenarios, *FA2* was used for each cluster to get K consumption forecasting results, and then the K results are added together to obtain the forecast value of total electricity consumption which is performed on $\mathbf{D}^{va}$. The Mean Absolute Percentage Error (MAPE) of the forecasting results under various clustering scenarios, $K_{optimal}$ can be obtained by Eqs. (9) and (10).

$$K_{optimal} = \underset{K \in [K_{\min}, K_{\max}]}{\arg\min} \{\text{MAPE}\} \tag{9}$$

$$\text{MAPE} = \frac{100\%}{T} \sum_{t=1}^{T} \frac{|d^{va}(t) - \hat{d}^{va}(t)|}{d^{va}(t)} \tag{10}$$

where {MAPE} represents the set of all MAPE calculated under multiple forecasting scenarios; T represents the number of forecast values; $d^{va}(t)$ and $\hat{d}^{va}(t)$ represents the total true value and total forecast value of electricity consumption at time t on $\mathbf{D}^{va}$.

In other words, aiming to determine the optimal cluster number, the greedy clustering is applied according to cluster number K which increases within this range $[K_{\min}, K_{\max}]$. Then, *FA2* is employed on $\mathbf{D}^{tr}$ and $\mathbf{D}^{va}$ to obtain the total consumption forecasting results based on each kind of clustering scenario. As K increases, if the forecasting accuracy is improved (in other words, MAPE decreases), the optimal cluster number becomes the corresponding K. In this way, the optimal clustering results can be finally achieved. When this range is large enough, it will include the globally optimal K.

C. The Selection of Forecasting Algorithm

Forecasting algorithms are used in all three stages where the greedy clustering stage uses the *FA1*, the selection of optimal K stage, and the final forecasting stage use the *FA2.* Due to the fact that the forecasting algorithm is not the main concern of this paper, the commonly used forecasting algorithm, Extreme Learning Machine (ELM) [29], is applied to this greedy clustering based MECF method.

In this paper, the *FA1* and *FA2* both apply the same ELM, but it is worth noting that the forecasting algorithm applied in the first stage is only used to calculate the goodness of fit, so it definitely can be inconsistent with *FA2* but the input features and out features should stay same with *FA2*. As for the configuration, the hidden layer of ELM is set to 20 neurons, the kernel function is sigmoid, and no other optimizations are performed on it.

Considering the seasonal periodicity of monthly electricity consumption, the input data is the electricity consumption data for the preceding 12 months of the *t*-th month to be forecasted, as shown in Eq. (11).

$$\begin{aligned} \boldsymbol{X} &= [d(t-12), d(t-11), ..., d(t-1)] \\ \boldsymbol{Y} &= [d(t)] \end{aligned} \tag{11}$$

Note that employing other factors (e.g., temperature) as input features also can help to improve forecasting accuracy. However, to highlight the proposed method, only the historical electricity consumption data was used in this paper.

III. Case Study

A. Pre-configuration of Computational Experiments

1) The Description of Dataset

The data used in this paper is collected from Ausgrid [30]. From all the substations operating continuously from May 1st of 2012 to April 30th of 2019, the 105 substations that have relatively complete data are selected. After filling up the missing data through linear interpolation, the original data with 15-min sampling interval was aggregated into monthly data. Data from May 1st of 2012 to April 30th of 2017, from May 1st of 2017 to April 30th of 2018, and from May 1st of 2018 to April 30th of 2019 are used as the training set, tuning (validation) set and testing set.

2) The Environment of Experiment

The computational experiments in this paper are performed using MATLAB (R2019b) and Python 3.8 on a laptop equipped with AMD Ryzen 7-4800H 2.90 GHz, 16GB usable RAM and Microsoft Windows 10 Home Edition.

3) The Evaluation Metrics

In this paper, three commonly used evaluation metrics are adopted to quantify the forecasting performance of the proposed method, including Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE).

$$\text{MAE} = \frac{1}{T}\sum_{t=1}^{T} | d^{te}(t) - \hat{d}^{te}(t) | \tag{12}$$

$$\text{MAPE} = \frac{100\%}{T}\sum_{t=1}^{T} \frac{| d^{te}(t) - \hat{d}^{te}(t) |}{d^{te}(t)} \tag{13}$$

$$\text{RMSE} = \frac{1}{T}\sqrt{\sum_{t=1}^{T} | d^{te}(t) - \hat{d}^{te}(t) |^2} \tag{14}$$

where T represents the number of forecast values; $d^{te}(t)$ and $\hat{d}^{te}(t)$ respectively represent the true value and forecast value of electricity consumption at time t on $\mathbf{D}^{te}$.

B. The Configuration of Test

There are 105 substations in this dataset. For greedy clustering, if K is too large, it will increase the calculation burden; on the contrary, if K is too small, the obtained clustering results may not meet the requirements of improving accuracy. Considering both the forecasting accuracy and the complexity of model training, the range of K is set from 2 to 11. It is noticed that this range may not be large enough to include the globally optimal cluster number, but it is large enough to show the performance of clustering.

To verify the superiority of the proposed method, three other MECF methods are set as comparisons:

M1: The proposed greedy clustering-based MECF method.

M2: K-means based MECF method, i.e., K-means is used as the clustering method.

M3: Forecast after aggregation, i.e. the electricity consumption data of 105 substations are aggregated to construct the regional electricity consumption time series, and then ELM is used to forecast the accumulated time series directly.

M4: Forecast before aggregated, i.e. the electricity consumption data of all substations are forecasted separately by ELM, and then the 105 forecasting results are aggregated to get the final results.

TABLE I. Error comparison of different methods

Metrics	M1	M2	M3	M4
MAE/MWh	**45721.70**	57551.64	76960.82	226944082
MAPE/%	**3.30**	4.19	4.84	14.30
RMSE/MWh	**62973.07**	68230.40	76960.82	226944.82

Table I shows the forecasting metrics of these four methods, while the best metrics are bolded. Obviously, according to these metrics, compared with the forecasting methods without clustering optimization (M3, M4), the clustering-based forecasting methods (M1, M2) achieve better forecasting results, which is well known. Among them, the M1 achieves the best MAPE at 3.30%, while the M4 has the worst MAPE at 14.30%. Compared with the commonly used K-means clustering-based forecasting (M2), the accuracy is improved by 21.24%. It can also be seen that the proposed method M1 achieves lower RMSE and MAE than M2, whose clustering target is inconsistent with the forecasting target.

Fig. 4 shows the curve between the real values and the forecast values of the four methods. It can be seen that the M1 also has a good fitting effect.

The forecast results of all methods are recorded in Table II, and the optimal results for each month are in grey. It clearly shows that the proposed M1 method achieves the best forecast results for most months.

C. The Comparison of Different Greedy Algorithms

In the previous section, both the *FA1* and *FA2* of the proposed method use the same ELM algorithm. It is worth noting that the *FA1* is only used to calculate the goodness of fit, so it does not have to stay the same with *FA2*.

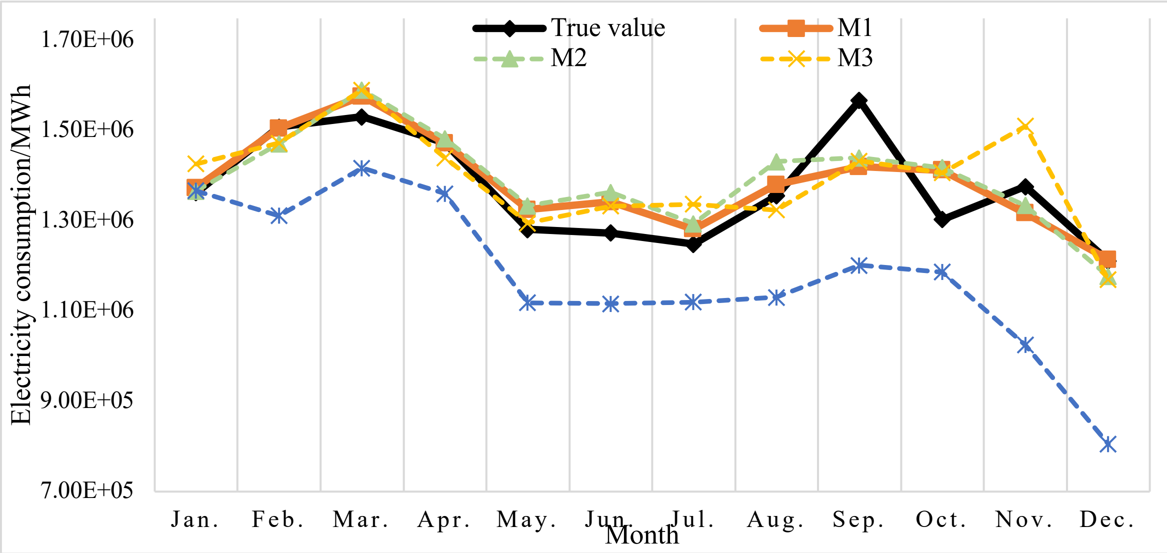


Fig. 4. Forecasting result curves of different methods.

TABLE II. THE FORECASTING RESULTS OF DIFFERENT METHODS

Month	True value/MWh	Forecast value/MWh			
		M1	M2	M3	M4
Jan.	1363355.10	1373771.44	1366895.59	1426803.23	1367405.87
Feb.	1508640.66	1505798.08	1472059.42	1473205.76	1312456.51
Mar.	1531246.07	1577306.12	1590790.33	1591456.02	1418131.13
Apr.	1473714.19	1473621.34	1483263.07	1440980.88	1361458.59
May.	1281689.57	1325807.08	1334473.09	1295797.31	1119057.17
Jun.	1274009.60	1342828.34	1363985.81	1333729.76	1116763.46
Jul.	1248956.83	1282358.08	1294413.55	1337497.11	1120441.15
Aug.	1356315.36	1381294.00	1431986.29	1325209.08	1130942.53
Sep.	1567507.58	1420791.25	1440533.81	1433768.04	1202105.22
Oct.	1303756.19	1413003.76	1418604.89	1407527.14	1187200.77
Nov.	1376441.71	1317650.39	1335018.02	1511111.00	1024986.70
Dec.	1211899.12	1215076.29	1177627.81	1170250.55	805651.67

TABLE III. ERROR COMPARISON OF DIFFERENT METHODS

Metrics	M1	G1	G2	G3	G4
MAE/MWh	45721.70	52395.67	49200.50	47000.20	43960.46
MAPE/%	3.30	3.79	3.51	3.40	3.19
RMSE/MWh	62973.07	64778.59	63832.63	59057.08	54927.85

In this part, the effectiveness of this method when using different forecasting algorithms as *FA1* is investigated and *FA2* keeps the same. Four kinds of greedy clustering based forecasting methods are designed as follows:

G1: Linear Regression (LR) is used as *FA1*.

G2: Support Vector Regression (SVR) with a linear kernel is used as *FA1*.

G3: ELM which has 50 hidden layers of neurons is used as *FA1*.

G4: ELM which has 100 hidden layers of neurons is used as *FA1*.

The metrics of forecasting performance are recorded in TABLE III. Compared with the forecasting results in the previous part, these four greedy clustering based forecasting methods all achieve better forecasting accuracy than those of conventional methods. Meanwhile, the results also indicate that the selection of the *FA1* may influence the final forecasting accuracy when *FA2* keeps the same. Only based on the forecasting performance metrics of these methods, it can be inferred that when *FA1* and *FA2* are the same types, the M1 can achieve better forecasting results. But the truth of the problem and its reasons need to be further explored in future research.

IV. DISCUSSION

The key point of the proposed greedy clustering-based forecasting method is the performance of the clustering algorithm. In this section, the impact of two factors on the greedy clustering algorithm is investigated and discussed.

A. The Impact of Initialization Strategy

For traditional clustering algorithms, like K-means, the initialization of the cluster center has a great influence on the clustering results and finally affects the forecasting accuracy [31]. To explore the effect of clusters' initialization on the forecasting performance, the three initialization strategies are adopted in the proposed greedy clustering:

I1: Randomly select K substations as the initialized cluster center, which is the initialization strategy used in this paper.

I2: Select K substations with the smallest similarity (biggest Euclidean distance) as the initialized cluster center, which is the strategy to improve the performance of K-means.

This article has been accepted for publication in IEEE Transactions on Industry Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIA.2022.3200352

I3: Select K substations with the biggest similarity (smallest Euclidean distance) as the initialized cluster center.

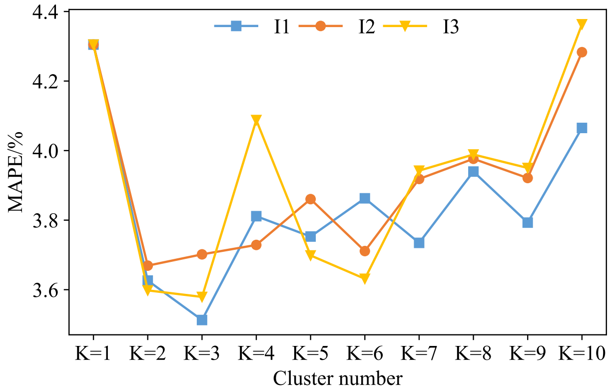


Fig. 5. The MAPE of three strategies under different cluster numbers.

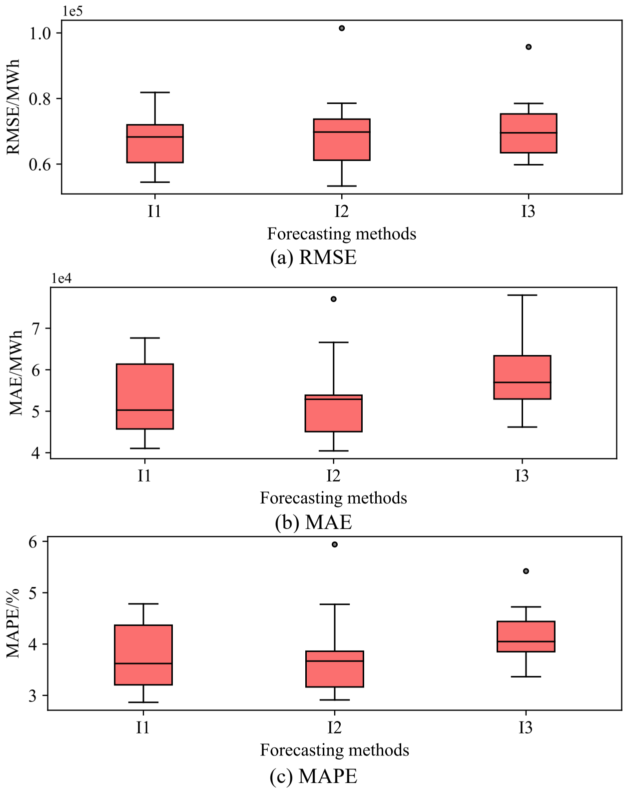


Fig. 6. The accuracy metrics of three initialization strategies for adaptive optimal greedy clustering based forecasting.

For the robustness of the results, after repeated calculation a hundred times using the above three strategies. The average MAPE of the forecasting results under several clustering scenarios is shown in Fig. 5. From the obtained results, it can be seen that the initialization strategy to achieve the optimal forecasting accuracy varies with cluster number. Among three strategies, strategy I1 showed higher forecasting accuracy in more clustering scenarios and achieved the highest forecasting accuracy when the cluster number K=3. It is worth noting that s large increase in forecasting accuracy occurs from K=1(no clustering) to K=2. When K=10, the accuracy of clustering-based forecasting is even worse than no clustering The reason is that the total number of individuals participating in the clustering is small, so a larger accuracy improvement can be achieved with a small cluster number under all strategies. However, when the cluster number is large, the division of individuals is too fine to improve the forecasting accuracy. The accuracy metrics of the adaptive optimal greedy clustering forecasting using three initialization strategies are shown in Fig. 6. The strategy I2 and I3 perform worse than I1, which means that random initialization can give full play to the superior performance of greedy clustering.

B. The Impact of Dataset Partition

In the proposed adaptive optimal greedy clustering-based forecasting method, the *FA1* is trained on the training set and the calculation of predictability and the selection of optimal cluster number are performed on the validation set. Therefore, the length of the training set and validation set both have an influence on the forecasting results. In this part, the testing set is fixed to 12 months, the validation set and the training set are 72 months in total. The length of the training set varied from 36 months to 68 months in 4-month steps and the cluster number K ranged from 2 to 10.

After one hundred rounds of repeated calculation, the MAPE under several clustering scenarios is shown in Fig. 7. As the length of the training set increases, the forecasting accuracy under these clustering scenarios generally shows an upward trend. In addition, it is worth noting that forecasting accuracy varies with the different cluster numbers and length of the validation set. It can be seen from Fig. 8 that as the length of the training set increases, the reduction of the validation set leads to a decrease in the probability of finding the real optimal number of clusters (the probability that the cluster number is the same when the highest accuracy on the validation set and the highest accuracy on the testing set), which limits further improvements in forecasting accuracy. Therefore, finding a suitable training set and validation set division ratio is also very important for the proposed method.

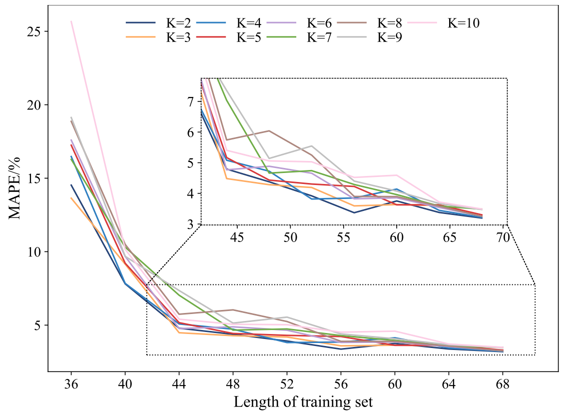


Fig. 7. The variation of MAPE with the length of the training set under the different cluster numbers.

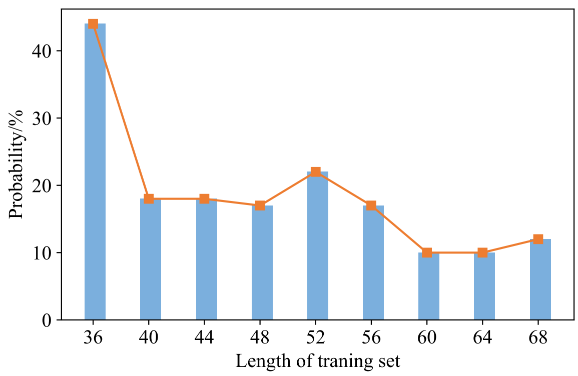


Fig. 8. The probability of being selected to the optimal validation set varies with the length of the training set.

V. Conclusion And Future Work

An adaptive optimal greedy clustering-based MECF method is proposed in this paper, which can achieve the coordination of clustering and forecasting to further improve the MECF accuracy. The basic idea is to find an optimal clustering to maximize the predictability. Case studies show that compared with the existing clustering algorithms that aim to maximize intra-class similarity, the proposed method can achieve more accurate forecasting. The proposed method can help electricity retailers mitigate the trading risks in the electricity market and also can provide more accurate basis for power planning.

The future works of this research are listed as follows:

1) The initialization strategy of the cluster's center will have a certain impact on the performance of greedy clustering. How to design the initialization strategy of the greedy clustering will be further explored.

2) The increasing penetration of distributed photovoltaic (DPV) in the distributed network has significant impacts on the MECF. The MECF methods under high penetration of DPV will be investigated in our future work to support more practical applications, such as peer-to-peer trading [32].

References

[1] D. Hadjout, J.F. Torres, A. Troncoso, et al, "Electricity consumption forecasting based on ensemble deep learning with application to the Algerian market," *Energy*, vol. 243, p. 123060, 2022.

[2] F. Wang, X. Ge, K. Li, and Z. Mi, "Day-ahead market optimal bidding strategy and quantitative compensation mechanism design for load aggregator engaging demand response," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp. 5564–5573, 2019.

[3] X. Lu, K. Li, H. Xu, F. Wang, Z. Zhou, and Y. Zhang, "Fundamentals and business model for resource aggregator of demand response in electricity markets," *Energy*, vol. 204, no. 117885, 2020.

[4] W. Jiang, X. Wu, Y. Gong, W. Yu, and X. Zhong, "Holt–winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption," *Energy*, vol. 193, p. 116779, 2020.

[5] Z. Li, K. Li, F. Wang, Z. Xuan, Z. Mi, W. Li, P. Dehghanian, and M. Fotuhi-Firuzabad, "Monthly Electricity Consumption Forecasting: A Step-Reduction Strategy and Autoencoder Neural Network," *IEEE Ind. Appl. Mag.*, vol. 27, no. 2, pp. 90–102, 2021

[6] C. Yuan, S. Liu, and Z. Fang, "Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model," *Energy*, vol. 100, pp. 384–390, 2016.

[7] L. Wu, X. Gao, Y. Xiao, Y. Yang, and X. Chen, "Using a novel multi-variable grey model to forecast the electricity consumption of Shandong province in China," *Energy*, vol. 157, no. 2018, pp. 327–335, 2018.

[8] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen, and K. Li, "Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting," *Appl. Sci.*, vol. 8, no. 8, pp. 1–29, 2018.

[9] Z. Hu, Y. Bao, R. Chiong, and T. Xiong, "Mid-term interval load forecasting using multi-output support vector regression with a memetic algorithm for feature selection," *Energy*, vol. 84, pp. 419–431, 2015.

[10] Z. Li, K. Li, F. Wang, Z. Mi, W. Li, and P. Dehghanian, "Auto-encoder neural network-based monthly electricity consumption forecasting method using hourly data," *2020 IEEE/IAS 56th Industrial and Commercial Power Systems Technical Conference (I&CPS)*, Las Vegas, NV, USA, 2020, pp. 1-8.

[11] K. Li, F. Wang, Z. Mi, M. Fotuhi-Firuzabad, N. Duić, and T. Wang, "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Appl. Energy*, vol. 253, p. 113595, 2019.

[12] T. Hong, J. Xie, and J. Black, "Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting," *Int. J. Forecast.*, vol. 35, no. 4, pp. 1389–1399, 2019.

[13] L. Peng, L. Wang, D. Xia and Q. Gao, "Effective energy consumption forecasting using empirical wavelet transform and long short-term memory," *Energy*, vol. 238, p. 121756, 2022.

[14] M. Somu, G. Raman M R, K. Ramamritham, "A deep learning framework for building energy consumption forecast," *Renew. Sustain. Energy Rev.*, vol. 137, p. 110591, 2022.

[15] B. Goehry, Y. Goude, P. Massart, et al., "Aggregation of multi- scale experts for bottom-up load forecasting," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 1895–1904, 2020.

[16] F. Wang, B. Xiang, K. Li, X. Ge, H. Lu, J. Lai, P. Dehghanian, "Smart households' aggregated capacity forecasting for load aggregators under incentive-based demand response programs," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1086-1097, 2020.

[17] F. Han, T. Pu, M. Li and G. Taylor, "Short-term forecasting of individual residential load based on deep learning and K-means clustering," *CSEE J. Power Energy Syst.*, vol. 7, no. 2, pp. 261-269, 2021.

[18] Y. Wang, Q. Chen, M. Sun, C. Kang, and Q. Xia, "An Ensemble Forecasting Method for the Aggregated Load with Subprofiles," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3906–3908, 2018.

[19] T. K. Wijaya, M. Vasirani, S. Humeau, and K. Aberer, "Individual, aggregate, and cluster-based aggregate forecasting of residential demand," Lausanne, Switzerland, Tech. Rep. *EPFL-REPORT-198477*, 2014.

[20] X. Wang, W. J. Lee, H. Huang, R. L. Szabados, D. Y. Wang, and P. V. Olinda, "Factors that impact the accuracy of clustering-based Load forecasting," *IEEE Trans. Ind. Appl.*, vol. 52, no. 5, pp. 3625–3630, 2016.

[21] X. Ge, F. Xu, Y. Wang, et al., "Spatio-Temporal Two-Dimensions Data Based Customer Baseline Load Estimation Approach Using LASSO Regression," *IEEE Trans. Ind. Appl.*, vol. 58, no. 3, pp. 3112-3122, 2022.

[22] A. M. Alonso, F. J. Nogales, and C. Ruiz, "Hierarchical clustering for smart meter electricity loads based on quantile autocovariances," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4522–4530, 2019.

[23] K. Li, X. Cao, X. Ge, F. Wang, X. Lu, M. Shi, R. Yin, Z. Mi, and S. Chang, "Meta-heuristic optimization-based two-stage considering intra-cluster compactness and inter-cluster separation," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3375–3384, 2020.

[24] Y. Li, D. Han and Z. Yan, "Long-term system load forecasting based on data-driven linear clustering method," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 2, pp. 306-316, 2018.

[25] C. Zhang and R. Li, "A Novel Closed-Loop Clustering Algorithm for Hierarchical Load Forecasting," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 432–441, 2021.

[26] F. Han, T. Pu, M. Li and G. Taylor, "Short-term forecasting of individual residential load based on deep learning and K-means clustering," *CSEE J. Power Energy Syst.*, vol. 7, no. 2, pp. 261-269, 2021.

[27] F. Wang, K. Li, C. Liu, Z. Mi, M. Shafie-Khah, and J. P. S. Catalao, "Synchronous pattern matching principle-based residential demand response baseline estimation: Mechanism analysis and approach description," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6972–6985, 2018.

[28] Y. Wang, Z. Li, F. Wang, et al., "Greedy Clustering-based Monthly Electricity Consumption Forecasting Model," *2021 IEEE Industry Applications Society Annual Meeting (IAS AM)*, Vancouver, BC, Canada, 2021, pp. 1-8.

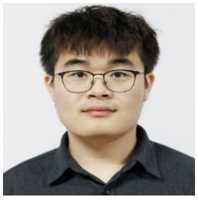
[29] G. Huang, "What are Extreme Learning Machines ? Filling the gap between Frank Rosenblatt's dream and John von Neumann's puzzle," *Cognit. Comput.*, vol. 7, no. 3, pp. 263–278, 2015.

[30] Ausgrid. *Distribution Zone Substation Information Data to Share*. Accessed: Aug. 31, 2020. [Online]. Available: http://www.ausgrid.com.au/Common/About-us/Corporate-information/Data-to-share/DistZonesubs.aspx#.WYD6KenauUl.

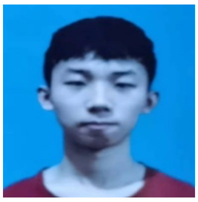
[31] Y. Li, D. Han and Z. Yan, "Long-term system load forecasting based on data-driven linear clustering method," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 2, pp. 306-316, 2018.

This article has been accepted for publication in IEEE Transactions on Industry Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIA.2022.3200352

[32] C. Huang, M. Zhang, C. Wang, N. Xie, and Z. Yuan, "An Interactive Two-stage Retail Electricity Market for Microgrids with Peer-to-Peer Flexibility Trading," *Appl. Energy*, vol. 315, p. 119085, 2022.



Yuqing Wang received the B.Sc. degree in Electrical Engineering from North China Electric Power University, Baoding, China, in 2021. Currently, he is pursuing the Ph.D. degree in the Department of Electrical Engineering at North China Electric Power University, Baoding, China. His research interests include demand response, renewable power and load forecasting and power system operation.



Zhiyang Fu received the B.Sc. degree in Mathematics from Shandong University, Jinan, China, in 2021. Currently, he is pursuing the M.Sc. degree in the Department of Mathematics and Physics at North China Electric Power University, Baoding, China. His research interests include machine learning and data mining and load forecasting.

Fei Wang (Senior Member, IEEE) received the B.Sc. degree from Hebei University, Baoding, China in 1993, the M.Sc. and Ph.D. degree in Electrical Engineering from North China Electric Power University (NCEPU), Baoding, China, in 2005 and 2013, respectively. He is currently a Professor with the Department of Electrical Engineering, NCEPU and the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, Baoding and Beijing, China. He is the Director of Institute of Power System Automation, the Director of Smart Energy Network Integrated Operation Research Center (SENIOR) and the Leader of "Double First-Class" scientific research team project at NCEPU. He was a Visiting Professor with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA, from 2016 to 2017. He was a Researcher with the Department of Electrical Engineering, Tsinghua University, Beijing, China, from 2014 to 2016.

He is an Associate Editor of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, the IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS, the IEEE POWER ENGINEERING LETTERS, the IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY, the IET RENEWABLE POWER GENERATION, the FRONTIERS IN ENERGY RESEARCH, the PROTECTION AND CONTROL OF MODERN POWER SYSTEMS (Springer), and the E-PRIME JOURNAL (Elsevier). He was the *Guest Editor-in-Chief* for the Research Topic "*Source-Grid-Load-Storage Collaborative and Interactive Optimization Control Technology of New Types of Active Distribution Network*" of the FRONTIERS IN ENERGY RESEARCH, and the *Guest Editor* for the special issue "*Demand Side Management and Market Design for Renewable Energy Support and Integration*" of the IET RENEWABLE POWER GENERATION. He is an IEEE Senior Member and the Expert Member of *IEC SC8A/WG2*. He supervised more than 80 Postdocs, Ph.D. and M.Sc. students. He has authored or coauthored more than 260 publications, including 100 international journal papers.

He was the recipient of the 2021 Elsevier China Highly Cited Scholar, the *2020 Science and Technology Progress First Award* of Hebei Province, *2018 Technical Invention First Award* of Hebei Province, the *2018 Patent Third Award* of Hebei Province, the *2014 Natural Sciences Academic Innovation Achievement Award* of Hebei Province, the *2018 China Electric Power Science and Technology Progress Third Award*, and the *2014 Outstanding Doctoral Dissertation Award* of NCEPU. He was the General Chair of the 2017 International Seminar of Renewable Energy Power Forecasting and Absorption Technology and 2018 International Seminar of Integrated Energy and Smart Microgrid Technology. He was the member of Series Steering Committee and Program Committee of 1^{st} to 5^{th} *International Conference on Smart Energy Systems and Technologies* (SEST) from 2018 to 2022. He was also the member of Scientific Advisory Board of 14^{th} to 17^{th} *International Conference on Sustainable Development of Energy, Water and Environment Systems* (SDEWES) from 2019 to 2022.

His research interests include renewable energy power/electricity price/electricity load forecasting, electricity market, demand response, smart grid, microgrid and integrated energy system.

Kangping Li received the B.Sc. and Ph.D. degree both in electrical engineering from North China Electric Power University, Baoding, China, in 2015 and 2020. Currently, he is an assistant professor with the College of Smart Energy, Shanghai Jiao Tong University, Baoding, China. His research interests include demand response, electricity market, renewable power and load forecasting and power system optimization.

Zhenghui Li received the B.Sc. and M.Sc. degree both in Electrical Engineering from Zhengzhou University, Zhengzhou, China, and North China Electric Power University, Baoding, China, in 2018 and 2021 respectively. Currently, he is an engineer with State Grid Anhui Electric Power Company, Bengbu Power Supply Company, Bengbu 233000, China. His research interests include load forecasting and demand response.

Zhao Zhen received the B.Sc. and Ph.D. degrees both in Electrical Engineering from North China Electric Power University (NCEPU), Baoding, China, in 2012, 2018, respectively. Currently, he is an assistant professor with the Department of Electrical Engineering at NCEPU, Baoding, China. His research interests include solar irradiance and PV power forecasting, wind speed and power forecasting.

Payman Dehghanian (S'11, M'17) is an Assistant Professor at the Department of Electrical and Computer Engineering in George Washington University, Washington, D.C., USA. He received the B.Sc., M.Sc., and Ph.D. degrees all in Electrical Engineering respectively from University of Tehran, Tehran, Iran, in 2009, Sharif University of Technology, Tehran, Iran, in 2011, and Texas A&M University, College Station, Texas, USA in 2017. His research interests include power system online situational awareness, real-time decision making, power system reliability and resiliency, asset management, and smart electricity grid applications.

Dr. Dehghanian is the recipient of the 2013 IEEE Iran Section Best M.Sc. Thesis Award in Electrical Engineering, the 2014 and 2015 IEEE Region 5 Outstanding Professional Achievement Awards, and the 2015 IEEE-HKN Outstanding Young Professional Award.

Mahmud Fotuhi-Firuzabad (F'14) received the B.Sc. and M.Sc. degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran, and Tehran University, Tehran, Iran, in 1986 and 1989, respectively, and M.Sc. and Ph.D. degrees in electrical engineering from the University of Saskatchewan, Saskatoon, SK, Canada, in 1993 and 1997, respectively. He is a Professor with the Department of Electrical Engineering and President of the Sharif University of Technology, Tehran, Iran. He is a member of center of excellence in power system control and management in the same department. His research interests include power system reliability, distributed renewable generation, demand response, and smart grids. Dr. Fotuhi-Firuzabad is a Visiting Professor with Aalto University, Finland. He serves as the Editor-In-Chief of the IEEE Power Engineering Letters, the Deputy Editor-in-Chief of Protection and Control of Modern Power Systems and also the Editor of Journal of Modern Power Systems and Clean Energy. He is the recipient of several national and international awards including World Intellectual Property Organization Award for the Outstanding Inventor, 2003, and PMAPS International Society Merit Award for contributions of probabilistic methods applied to power Systems in 2016.

This article has been accepted for publication in IEEE Transactions on Industry Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIA.2022.3200352

João P. S. Catalão (Fellow, IEEE) received the M.Sc. degree from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2003, and the Ph.D. degree and Habilitation for Full Professor ("Agregação") from the University of Beira Interior (UBI), Covilha, Portugal, in 2007 and 2013, respectively. Currently, he is a Professor at the Faculty of Engineering of the University of Porto (FEUP), Porto, Portugal, and Research Coordinator at INESC TEC. He was also appointed as Visiting Professor by North China Electric Power University (NCEPU), Beijing, China.

He was the Primary Coordinator of the EU-funded FP7 project SiNGULAR ("Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration"), a 5.2-million-euro project involving 11 industry partners. He was also the Principal Investigator of 3 funded projects by FCT (Portuguese National Funding Agency for Science, Research and Technology) and FEDER (European Regional Development Fund). Moreover, he has authored or coauthored more than 975 publications, including 485 international journal papers (more than 165 IEEE TRANSACTIONS/JOURNAL papers, 218 *Elsevier* and 22 *IET* journal papers), 443 international conference proceedings papers (vast majority co-sponsored by IEEE), 4 books and 43 book chapters, with an *h*-index of 77, an *i10*-index of 419, and more than 23,000 citations (according to Google Scholar), having supervised more than 110 post-docs, Ph.D. and M.Sc. students, and other students with project grants.

He was the inaugural Technical Chair of SEST 2018 — *1st International Conference on Smart Energy Systems and Technologies* (technically co-sponsored by IEEE IES), the General Chair of SEST 2019 (technically co-sponsored by IEEE PES and IEEE IES), the General Co-Chair of SEST 2020 (technically co-sponsored by IEEE PES, IEEE IES and IEEE IAS), and the Honorary Chair of SEST 2021 (technically co-sponsored by IEEE PES, IEEE IES, IEEE IAS and IEEE PELS). He was also the Editor of the books entitled *"Electric Power Systems: Advanced Forecasting Techniques and Optimal Generation Scheduling"* and *"Smart and Sustainable Power Systems: Operations, Planning and Economics of Insular Electricity Grids"* (Boca Raton, FL, USA: CRC Press, 2012 and 2015, respectively).

He is an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON CLOUD COMPUTING, and the IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS. He was the Senior Editor of the IEEE TRANSACTIONS ON SMART GRID and the Promotion and Outreach (Senior) Editor of the IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY, 2020-2021, being also a member of the IEEE PES Publications Board. Furthermore, he was an Associate Editor of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, 2011-2018, an Associate Editor of the IEEE TRANSACTIONS ON SMART GRID, 2013-2020, an Associate Editor of both the IEEE TRANSACTIONS ON POWER SYSTEMS and the IEEE POWER ENGINEERING LETTERS, 2017-2021, and an Associate Editor of both the IEEE SYSTEMS JOURNAL and the IEEE ACCESS, 2020-2022.

He was the *Guest Editor-in-Chief* for the Special Section on "Real-Time Demand Response" of the IEEE TRANSACTIONS ON SMART GRID, published in December 2012, the *Guest Editor-in-Chief* for the Special Section on "Reserve and Flexibility for Handling Variability and Uncertainty of Renewable Generation" of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, published in April 2016, the *Corresponding/Lead Guest Editor* (*Guest Editor-in-Chief*) for the Special Section on "Industrial and Commercial Demand Response" of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, published in November 2018, the *Guest Co-Lead Editor* for the Special Section on "Invited Papers on Emerging Topics in the Power and Energy Society" of the IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY, published in October 2020, the *Guest Co-Lead* Editor for the Special Section on "Invited Papers in 2021 on Emerging Topics in the Power and Energy Society" of the IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY, published in November 2021, and the *Guest Editor-in-Chief* for the Special Section on "Demand Response Applications of Cloud Computing Technologies" of the IEEE TRANSACTIONS ON CLOUD COMPUTING, published in March 2022.

He was the recipient of the 2011 Scientific Merit Award UBI-FE/Santander Universities, the 2012 Scientific Award UTL/Santander Totta, the 2016-2020 (five years in a row) FEUP Diplomas of Scientific Recognition, the 2017 Best INESC-ID Researcher Award, and the 2018 Scientific Award ULisboa/Santander Universities (with an Honorable Mention in 2017). He was recognized as one of the Outstanding Associate Editors 2020 of the IEEE TRANSACTIONS ON SMART GRID, and also one of the Outstanding Associate Editors 2021 of the IEEE TRANSACTIONS ON POWER SYSTEMS. He has multiple Highly Cited Papers in Web of Science. He is a *Top Scientist* in Research.com, which lists only scientists having *h*-index equal or greater than 30. He is also among the *0.5% Top Scientists*, according to a study published by a team at Stanford University. Furthermore, he has won 5 *Best Paper Awards* at IEEE Conferences (MELECON, POWERENG, SEGE, EDST, SCEMS) and the MPCE *Best Paper Award* 2019. Moreover, his former M.Sc. and Ph.D. students have won the National Engineering Award in 2011, the 1st Prize in the REN (Portuguese TSO) Award in 2019, and the 1st Prize in the Young Engineer Innovation Award in 2020.

His research interests include power system operations and planning, power system economics and electricity markets, distributed renewable generation, demand response, smart grid, and multi-energy carriers.