

Day-Ahead Electricity Consumption Prediction of Individual Household—Capturing Peak Consumption Pattern

Zhong Xia[✉], Member, IEEE, Ruiyuan Zhang[✉], Member, IEEE, Hui Ma[✉], Senior Member, IEEE,
and Tapan Kumar Saha[✉], Fellow, IEEE

Abstract—Day-ahead electricity consumption forecasting for individual residential consumers, especially the peak consumption time forecasting, is essential for home energy management system. However, it is considerably challenging since the single household consumption is highly volatile and stochastic and dependent on the underlying human behaviour. Further, the difficulties of accurately capturing the occurrence time of peak consumption in a house can limit the performance of existing machine learning-based load forecasting methods. In this paper, we propose a novel framework for day-ahead single-household electricity consumption forecasting by learning the peak consumption patterns of users. Instead of attempting to obtain the electricity consumption curve for the future 24 hours, the prediction of electricity consumption is achieved by combining the predicted base consumption, the predicted peak consumption occurrence time and the predicted amount of peak consumption within each time interval. The proposed framework can be used for both deterministic and probabilistic load forecasting of individual households. Case studies are conducted on hundreds of households from two different datasets. The results demonstrate that the performance of different deterministic load forecasting algorithms and probabilistic load forecasting algorithms can be improved after being integrated into the proposed framework.

Index Terms—Day-ahead, deterministic load forecasting, individual household, probabilistic load forecasting, peak load occurrence time.

I. INTRODUCTION

THE LARGE scale deployment of smart meters at the individual household-level makes it possible to collect the fine-grained electricity consumption data of residential consumers. This provides opportunities of predicting day-ahead electricity consumption (load) at single dwellings to facilitate implementing home energy management systems (HEMS) to improve energy efficiency [1], [2].

Manuscript received 15 December 2022; revised 14 May 2023 and 22 September 2023; accepted 8 November 2023. Date of publication 13 November 2023; date of current version 23 April 2024. Paper no. TSG-01878-2022. (*Corresponding author: Hui Ma*.)

Zhong Xia is with the School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, QLD 4072, Australia.

Ruiyuan Zhang is with the School of Engineering, The Hong Kong University of Science and Technology, Hong Kong, China.

Hui Ma and Tapan Kumar Saha are with the School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, QLD 4072, Australia (e-mail: huima@itee.uq.edu.au).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TSG.2023.3332281>.

Digital Object Identifier 10.1109/TSG.2023.3332281

Load forecasting at aggregated level has been extensively studied [3], [4]. Numerous algorithms using time series analysis [5], [6] and machine learning [7], [8] are implemented to provide load forecasting. However, the performance of these algorithms may decay when the level of aggregation decreases [9]. It is reported that the load forecasting accuracy (evaluated by Mean Absolute Percent Error, MAPE) of neural networks can decrease by 5 times when the aggregated level is decreased from 1000 smart meters to 10 smart meters and by 28 times as the aggregated level further decreases to a single smart meter [10], [11]. Therefore, load forecasting at the individual household level is more challenging. It is because the electricity consumption at a household largely depends on the consumer behaviour, which is with high randomness.

Different approaches have been proposed to achieve effective load forecasting of the individual household. One approach is to leverage the historic consumption data from multiple strategy households to train a federated forecasting model, which can capture the identical distribution among the consumption data of these households. Then a personalized forecasting model for each household is obtained by personalizing the federated forecasting model on the individual household's own consumption data [12], [13], [14]. Another approach is to utilise clustering techniques to obtain several prototypes (the cluster centres), each of which can represent a typical daily consumer behaviour pattern, and then the forecasting accuracy can be improved by learning the link between the future consumption and the prototypes of previous days [15]. Similarly, sparse code decomposition of previous days' load curves has been utilised as input features for individual household load forecasting [16]. The above approaches are based on the assumption that the accuracy of household-level load forecasts can be improved by discovering more stable components from the historic consumption data. Recently, deep learning techniques, especially the long short-term memory (LSTM) recurrent neural network (RNN) have been implemented to explore the temporal correlations between the future load and the hidden consumer's behaviour patterns [17], [18], [19], [20].

Even under similar influencing factors (weather conditions and time periods etc), the peak load of a household may occur at different time instances due to the changeable behaviour of the consumer. This makes it difficult to accurately predict the exact time instance of the occurrence of the future peak load.

If an algorithm ignores the peak load and instead generates a flat future load curve, its forecasting error could be small with respect to the point-wise evaluation metrics, such as mean absolute error (MAE), root mean squared error (RMSE) and MAPE. The error value can be even smaller than that of the algorithms which predict the peak load. This is because a double penalty incurs if the magnitude of the peak load is correctly predicted but displaced in the occurrence time [19]. As a result, if a single-household load forecasting model is designed to produce the shape of the future daily load curve with the objective of only minimising the point-wise evaluation metrics, its output will be conservative. In contrast, if the algorithm predicts the peak load at its approximate occurrence time, the conservative output issue can be tackled. Actually, the peak load level normally shows a higher level of consistency than the exact occurrence time of peak load for a single household. For example, a consumer always cooks in the evening (peak load occurs when cooking), but the exact time may vary between 5 pm and 7 pm depending on the consumer's daily routine and mood.

Motivated by the above discussion of the household peak load patterns, this paper proposes a novel peak load analysis-based framework for the day-ahead load forecasting of a single household. Firstly, the peak load time and valley load time of a household are identified from its historical load data. Accordingly, a day is divided into several valley-to-valley time intervals. Secondly, the total load of a household is decomposed into a base load and a peak load. The base load is predicted by the traditional load forecasting methods. For the peak load, instead of predicting its exact shape, multiple models are trained to predict the total amount of consumption and its occurrence time within each time interval. The final load forecasting is obtained by combining the predicted base load, the amount of peak consumption and its occurrence time within each time interval. Both deterministic regression models and probabilistic regression models can be integrated into this framework. Under this framework, the peak load pattern embedded in the load data can be captured, and the predicted load can deliver more useful future peak load information, which in turn can provide important support for load scheduling in HEMS. Also, a number of new metrics, including the shape-based evaluation metrics, are designed to quantitatively verify the household-level load forecasts.

The above mentioned 'double penalty' issue was illustrated in [20]. However, the main focus of [20] was on developing the new error measurement approaches instead of new load forecasts. The usage time of household's appliances varies with the consumer's immediate circumstances. To deal with this issue, in [21] a shape-based measure, Dynamic Time Warping was used for load profile clustering and generating several prototypes. However, the peak load patterns captured may still be limited, since each prototype is the presentation of a set of load profiles and can be much smoother compared to each individual load profile.

Also, numerous works have been conducted for peak load forecasting. Different machine learning methods, such as artificial neural network [22] and tree-based ensemble learning [23] have been applied and considerably accurate peak

load prediction has been achieved. Some peak load forecasting methods adopted frequency domain analysis [24], [25] or more advanced deep learning models [26]. However, the above methods mainly focus on predicting peak load at the aggregated level to facilitate power system operation. The peak load of an individual household is significantly different from the peak load at the aggregated level since the single-household load normally exhibits considerably higher uncertainty. The proposed load forecasting method in this paper is aimed at better capturing the uncertainty of peak load of an individual household. Also, the proposed method in this paper provides a day-ahead 24 hours load forecasting instead of only the peak load for the individual household.

We have also investigated single-household load forecasting in our previous work [27]. The proposed method in [27] is designed for hour-ahead load forecasting with the emphasis on revealing all the possible electricity consumption scenarios of a household in the future one hour. However, the day-ahead single-household load forecasting considered is different from the hour-ahead load forecasting, since the uncertainty lies not only in the amount of electricity consumed, but also in the time at which the electricity is consumed. The work of this paper is for an effective day-ahead single-household load forecasting with an appropriate description and the inclusion of future peak load.

The contributions of this paper are as follows:

(1) A clustering-based approach is designed to automatically identify the peak load time and valley load time for individual household and accordingly the 24 hours of a day are appropriately segmented into several valley-valley time intervals.

(2) A novel day-ahead single-household load forecasting framework is proposed to capture the peak load patterns in a targeted manner for households with highly uncertain consumption behaviour. In this framework, the task of predicting peak load is decomposed into multiple tasks of predicting the total amount of peak load and average time of peak load within each time interval. Both deterministic and probabilistic load forecasts can be integrated into this framework.

(3) New evaluation metrics instead of point-wise evaluation metrics are proposed and utilised to quantitatively assess the household level load forecasts and verifies the effectiveness of the proposed method.

II. HOUSEHOLD ELECTRICITY CONSUMPTION CHARACTERISTICS

Prediction of electricity consumption is generally achieved by learning the association between future consumption and inputs such as external variables (time, weather, etc.) and lagged consumption (Fig. 1). If the future consumption is strongly correlated with the inputs, then a trained regression model can make effective predictions given the input features. Aggregated load forecasting falls into this category. Conversely, for the single household the hour-head electricity consumption can fluctuate considerably even with very similar external conditions at the previous several hours. For the day-ahead single-household electricity consumption, its forecasting is even more difficult since the influencing factors of the

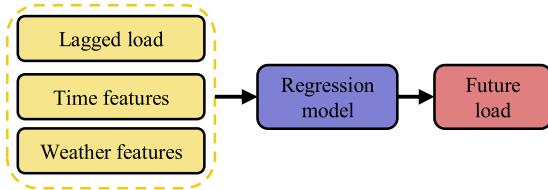


Fig. 1. Basic framework of typical load forecasting method.

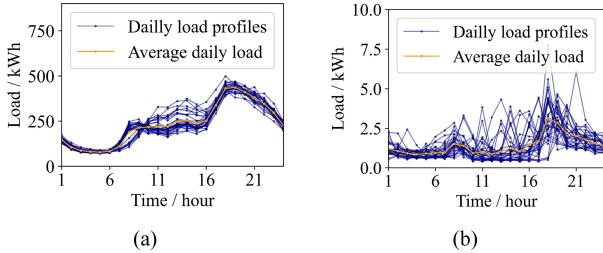


Fig. 2. Daily load profiles over 30 consecutive days (a) aggregated load of 200 households (b) an individual household.

previous several hours are no longer available as the inputs for the load forecasting of the future few hours.

The daily total load curves for 200 households over 30 consecutive days from the Irish Commission for Energy Regulation (CER) smart meter dataset are plotted in Fig. 2 (a). In the figure, the load at time stamp 1 denotes the electricity consumption during 0:00 am – 1:00 am and the load at time stamp 24 denotes the consumption during 11:00 pm – 0:00 am. The daily load profiles of one randomly chosen household over the same 30 days are plotted in Fig. 2 (b). In the figure, the load at time stamp 1 denotes the electricity consumption during 0:00 am – 1:00 am and the load at time stamp 24 denotes the consumption during 11:00 pm – 0:00 am. The daily load profiles of one randomly chosen household over the same 30 days are plotted in Fig. 2 (b).

From the above figure, we can see that the aggregated load (Fig. 2a, total load of the 200 households) is relatively consistent while the load of an individual household (Fig. 2b) is much more fluctuating in a considerably wide range. The day-ahead aggregated load prediction is easier than that of a single household. Though the magnitude of the aggregated load changes widely around noon, the fluctuations are highly correlated to the calendar features (residential loads for this time are higher on weekends than on weekdays). However, for a single household's consumption, the uncertainties are not only of the magnitudes of the load but also the time of occurrence of the load.

For the interval from time stamp 16 to time stamp 24 in Fig. 2 (b), we can observe that unlike the relatively consistent shape of the aggregated load over the 30 days, the maximum consumption (peak load) of the individual household varies considerably. Even with similar influencing factors (lagged load, weather conditions and time features), the peak load still occurs at different times. This is because the time of peak load of a single household is weakly correlated with the influencing factors but highly dependent on the consumer behaviour (which could be random). If the approach shown in Fig. 1 is utilised, the prediction provided by the trained regression model may tend to be close to the average daily load (the orange line in Fig. 2b). Such forecasts can only provide limited information on future peak loads of an individual household, and therefore it may not be able to provide sufficient support for applications like the household day-ahead load scheduling.

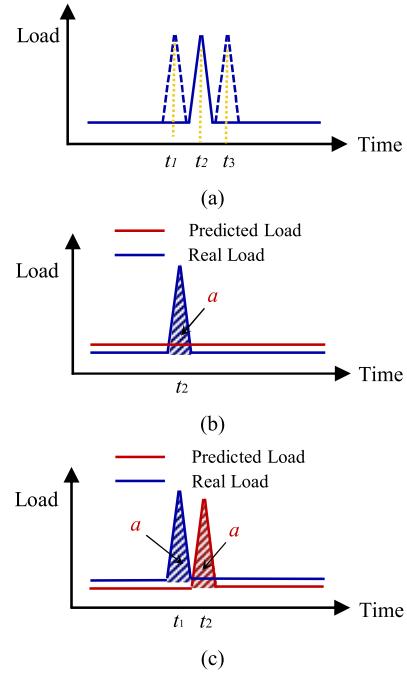


Fig. 3. (a) Real load curve with peak load at different times; (b) Predicted load by Strategy 1); (c) Predicted load by Strategy 2).

We utilise a simplified example to further illustrate this issue. Assuming the load is constant at a certain value during most of the time with only one peak load per day, as plotted in Fig. 3 (a). The peak load may occur at t_1 with a probability of $1/3$ (i.e., 33.33%), 33.33% to occur at t_2 and 33.33% to occur at t_3 . The occurrence time of peak load is highly dependent on the consumer behaviour and is weakly correlated to the input features. Thus, it is difficult to accurately predict the occurrence time of peak load. For this case, we can use two strategies to predict the future load. Strategy 1 is to predict the future daily load curve as a flat curve like Fig. 3 (b), then the prediction error will be the area a (the blue shaded area in Fig. 3b) whenever the peak load actually occurs. Strategy 2 is to predict that the future daily load curve is a flat curve with a peak load at t_2 as shown in Fig. 3 (c). If the real peak load occurs at t_1 , then the prediction error will be $a + a = 2a$ (the red shaded area plus the blue shaded area in Fig. 3c). Similarly, the error will be $2a$ if peak load occurs at t_3 . If peak load occurs at t_2 , the error will be 0. Thus, the average error for Strategy 2 is $\frac{1}{3} * 2a + \frac{1}{3} * 2a = \frac{4}{3}a > a$. We can see that by ignoring the peak load we can achieve higher prediction accuracy than when predicting the peak load, since double penalty occurs if the occurrence time of peak load is mis-predicted. Due to the double penalty, when applying the conventional load forecasting method (Fig. 1) for day-ahead load forecasting of an individual household with highly uncertain electricity consumption pattern, the forecasting results may be relatively conservative to minimise the overall error, which is closer to the outputs of Strategy 1 rather than Strategy 2.

However, Strategy 2 (predicting future peak load) can provide more useful information of future consumption than Strategy 1 (ignoring peak load to avoid double penalty) for

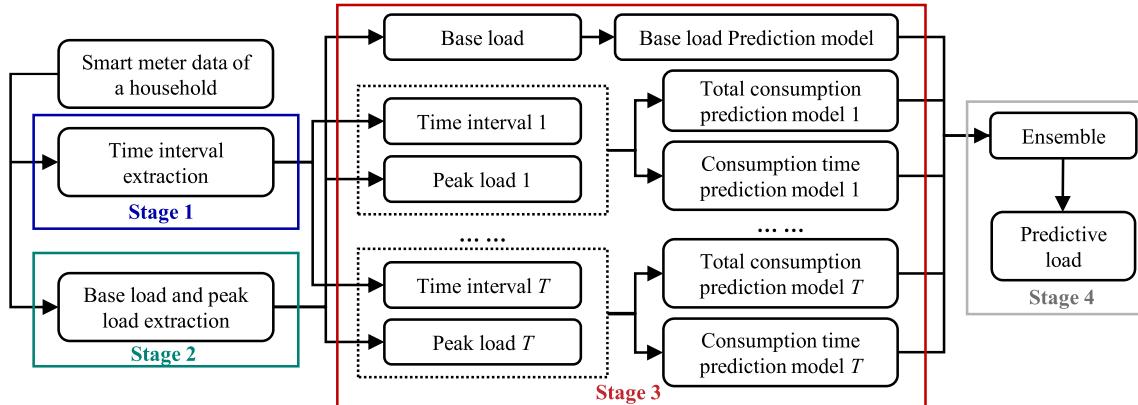


Fig. 4. Framework of day-ahead single-household load forecasting method.

home energy management systems (HEMS). As an example, if in the afternoon, the electricity price is high and the peak load is predicted to be higher than the PV output, the HEMS will charge the energy storage system (ESS) in the morning and discharge it in the afternoon. The HEMS may make wrong decisions without effective predicting future peak load, and this may lead to economic loss to the consumer.

Although it is difficult to predict the occurrence time of peak load, the occurrence of peak load normally shows a higher level of consistency than its exact occurrence time for a single household. For example, a consumer always cooks in the evening (certain) and the peak load occurs when cooking, but the exact time may vary between 5 pm and 7 pm depending on the consumer's daily routine and mood (uncertain). This is the main motivation of the proposed approach. The base load and peak load are learned and predicted by different models separately in the proposed method. The predicted base load and peak load are combined to get the final 24 numbers for future load. This framework can be regarded as 'forcing' the models to predict the peak load to fully capture the consistent part of peak load, aiming at generating outputs closer to Strategy 2 instead of Strategy 1).

Also, it can be seen that the point-wise evaluation metrics may not be good choices for such single-household load forecasting due to the double penalty. Thus, new evaluation metrics should be used to evaluate the forecasting performance of different methods (refer to Section III-D).

III. METHODOLOGY

The proposed day-ahead single-household load forecasting method is shown in Fig. 4. Its four main stages are:

- 1) Time interval extraction to identify the main time intervals of the load in a day.
- 2) Base load and peak load extraction to identify the base and peak load from the consumer's total load.
- 3) Base load and peak load prediction to make estimation of the base load and peak load of the consumer in different time intervals separately.
- 4) Load ensemble to obtain the final load prediction by combining the base load and peak load prediction obtained from stage 3.

A. Time Interval Extraction

In this stage, the 24 hours of a day are divided into several time intervals, which are determined by the consumer's electricity usage behaviour. The boundaries of two adjacent time intervals are located at the time stamps, where the consumption is relatively low and where the maximum consumption (peak load) in each of the two time intervals exhibits a unique pattern.

We use a randomly selected consumer from the Irish CER smart meter dataset as an example to present the procedure for time interval extraction.

Let $l_{d,t}$ ($t \in [1, 24]$) denote the load of a household during the t^{th} hour on the d^{th} day. Note that the first hour indicates 0 am – 1 am and the 24th hour indicates 11 pm – 12 am. The load data in 2010 for this consumer is normalised by:

$$l'_{d,t} = \frac{l_{d,t}}{\max_{i,j} l_{i,j}}, \quad i \in [1, 365], j \in [1, 24] \quad (1)$$

It should be noted that the data normalisation can be conducted either before or after conducting the time period extraction. In this paper, we perform data preparation before executing the algorithm (consisting of time period extraction and base load and peak load prediction), so the data was normalised in advance.

The normalised daily load profile of the consumer on the d^{th} day is then denoted as ' $D_L'_d = [l'_{d,t}]_{t=1,2,\dots,24}$ '. A set of normalised daily load profiles over the 365 days in 2010, $\{D_L'_d\}_{d=1,2,\dots,365}$, are plotted in Fig. 5 (a). the average daily load profile, $A_L' = [a_l']_{t=1,2,\dots,24}$, is obtained by:

$$a_l' = \frac{\sum_{d=1}^{n_d} l'_{d,t}}{n_d} \quad (2)$$

where n_d denotes the total number of days in the training dataset (equal to 365 in the illustrated example). The average daily load profile is shown in Fig. 5 (b).

From Fig. 5 (a) we can see that peak load often occurs around 8 am, 1 pm and 6 pm, and the peak load around these three times exhibits different patterns. Intuitively, the 24 hours of a day can be divided into three intervals for this consumer according to the peak load pattern. The boundaries of each time interval are located at the off-peak load time.

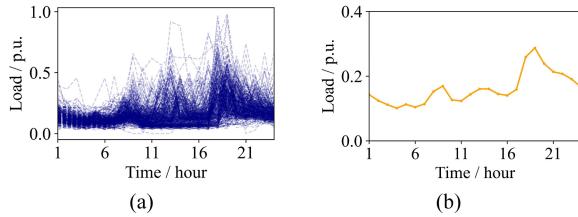


Fig. 5. (a) 365 daily load curves in 2010 (b) average daily load curve.

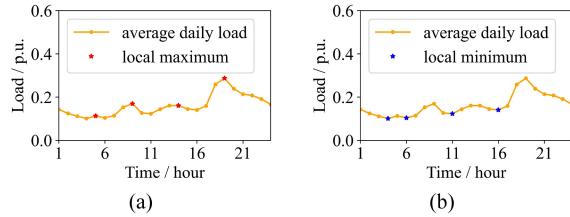


Fig. 6. (a) Local maximum points and (b) local minimum points.

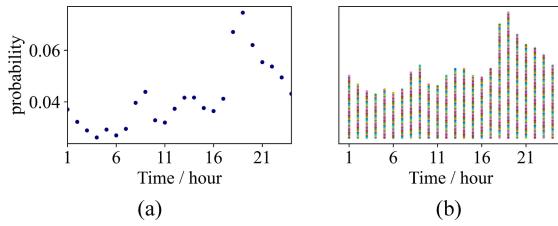


Fig. 7. (a) Discrete probability distribution; (b) 1000 samples generated from the probability distribution in (a).

We start with a simple way to extract the time periods: set the number of time intervals as the number of local maximum points in the average daily load curve and use the local minimum points in the average daily load curve as boundaries of adjacent time intervals. A local maximum point is defined as a time step, at which the average load value is higher than the load values at the adjacent time steps (refer to Fig. 6a). A local minimum point is defined as a time step, at which the average load value is lower than the load values at the adjacent time steps (refer to Fig. 6b). However, there are two issues associated with the above method:

(1) There are four local maximum points as shown in Fig. 6 (a). Thus, we get four time intervals instead of three time intervals as expected. Though time step 5 is a local maximum point in the average daily load curve, this is just due to the average load fluctuations in a small range at time step 5. It is difficult to decide whether a local maximum point should be extracted as a peak load time period or not.

(2) It is difficult to determine whether we should allocate a local minimum point to the previous time interval or to the next time interval. For example, time step 11 is a local minimum point (it indicates the electricity consumption during 10:00am-11:00am). However, we cannot decide whether this time step should be allocated to the previous time interval (6am-10am) or to the next time interval (11am-4pm) simply by considering the average daily load curve.

To overcome the above two issues, we propose a clustering-based method. Firstly, the average daily load profile $A \cdot L'$ is transformed into a discrete probability distribution (Fig. 7a). Assuming X is a discrete random variable (X can take the

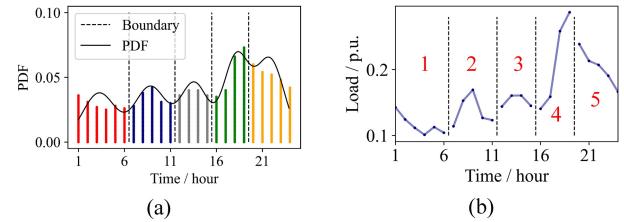


Fig. 8. Cluster boundaries plotted in (a) 1000 data samples (b) average load curve.

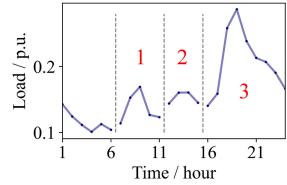


Fig. 9. The results of time interval extraction.

values 1,2 ... or 24, each representing an hour of the day) and we utilise the probability distribution in Fig. 7 (a) to generate a set of 1000 samples $[x_1, \dots, x_{1000}]$: the probability distribution is high at the peak load time and more samples are generated around these times, while fewer data points are generated around the valley load time. Instead of the total number of the samples, the relative proportion of numbers of samples at different time instances (1, 2, ..., 24) determines the results of the clustering. After the test and trial, we select 1000 as the number of samples to ensure sufficient samples are generated to represent electricity usage for each hour of the day and maintain the algorithm's efficiency. The generated samples are plotted in Fig. 7 (b).

Then, these samples are divided into several clusters by using the Gaussian Mixture Model (GMM). A GMM assumes all the samples are generated from a mixture of a finite number of Gaussian distributions. The expectation-maximization (EM) algorithm is implemented to fit the GMM model to the samples. Each sample is assigned to one of the Gaussian distributions, of which the sample has the highest probability to belong to. The samples assigned to the same Gaussian distribution are grouped into the same cluster. For the consumer in this example, the peak load normally occurs around three different time intervals. However, the number of clusters for the GMM model cannot be set as three in advance, since this number can vary for different consumers. From the analysis of average load curve above, we can see that the number of time intervals may be equal to or smaller than T_{\max} (the number of local maximum points). So, we set the number of clusters of the GMM model as $(T_{\max} + 1)$ and later some clusters may be merged to obtain the optimal number of time intervals. It should be noted that the number of clusters is set as $(T_{\max} + 1)$ instead of T_{\max} , because the first cluster (starts from 12 am midnight) may be merged with the last cluster (ends at 12 am midnight) to form an entire time interval. In this example, $(T_{\max} + 1)$ is 5, so the samples are divided into 5 clusters. The clustering result is plotted in Fig. 8 (a), where the probability density function (PDF) of the fitted model is presented and the samples belonging to different clusters are plotted in different colours. The boundaries of adjacent clusters are also plotted in the figure. By applying these cluster boundaries to the average

Algorithm 1 Time Interval Merging

Input: Number of time intervals before merging T_0 ; Boundaries of the time intervals before merging $[b_{0i}]_{i=1,2\cdots 2*N_0}$ (the boundaries of the n^{th} time interval are b_{n*2-1} and b_{n*2} , $n \in [1, T_0]$); Average daily load profile $[a_{-l'_t}]_{t=1,2\cdots 24}$

Output: Number of time intervals after merging T and boundaries of the time intervals after merging $[b_i]_{i=1,2\cdots 2*N}$

1: Initialize $T = T_0$ and $[b_i] = [b_{0i}]$, $i \in [1, N_0]$

2: **while** 1:

- for** $n = T$ **to** 1 **do**:
- if** the max value of $[a_{-l'_t}]$ in the range of time interval n is located at the left boundary of time period n **do**:
- if** $n \neq 1$ **do**: merge time interval n into time interval $n-1$; update $[b_i]$ accordingly; $T = 1$
- else if** $n = 1$ **do**: merge time interval n into time interval T ; update $[b_i]$ accordingly; $T = 1$
- end if**
- break for**
- end if**
- end for**
- if** $n = 1$ or $T = 1$ **do**:
- break while**
- end if**

3: **while** 1:

- for** $n = 1$ **to** T **do**:
- if** the max value of $[a_{-l'_t}]$ in the range of time interval n is located at the right boundary of time interval n **do**:
- if** $n \neq T$ **do**: merge time interval n into time interval $n+1$; update $[b_i]$ accordingly; $T = 1$
- else if** $n = N$ **do**: merge time interval n into time interval 1; update $[b_i]$ accordingly; $T = 1$
- end if**
- break for**
- end if**
- end for**
- if** $n = T$ or $T = 1$ **do**:
- break while**
- end if**

4: **Return** T and $[b_i]$

daily load curve, we can obtain five time periods as shown in Fig. 8 (b).

Lastly, the time period which cannot cover a typical peak load circle is merged into its adjacent clusters by Algorithm 1-Time Interval Merging. The three time intervals are obtained, as shown in Fig. 9. Time interval 1 is 4am-11am, time interval 2 is 11am-4pm and time interval 3 is 4pm-4am.

By this method, the optimal number of time intervals is obtained. This method ensures that the peak load in each of two adjacent time intervals exhibits a unique pattern. Then the pattern of peak load in different time intervals is captured by the machine learning models in Stage 3. If the number of time intervals is not the optimal number, the representativeness of each time interval cannot be ensured. Specifically, if the number of time intervals is too large, some intervals may not

be long enough cover the entire time period of a peak load; if the number of time intervals is too small, then peak loads of two adjacent intervals may be merged into one time interval.

B. Base Load and Peak Load Extraction

In this stage, the total load of a household is broken down into base load and peak load.

In our method, the boundary between the base load and the peak load is defined as a curve with a 50% probability that the actual load is above it and a 50% probability that the actual load is below it. Then the part of the load below this boundary is defined as the base load and the part of the load above this boundary is defined as the peak load.

To obtain the base-peak boundary, a straightforward way is to generate one boundary curve for all the daily load of a household. For example, the 365 load curves of a consumer over a year are plotted in Fig. 5 (a). The boundary curve is generated by calculating the median curve of these 365 curves, ensuring each hour's load is 50% above the boundary and 50% below the boundary. Then for each daily load curve, the part of consumption lower than the boundary is extracted as the base load and the higher part is extracted as the peak load.

However, the magnitude of the base load may vary under different circumstances and a single boundary may not be sufficient. For example, many residential consumers from the Irish dataset have higher electricity usage as the temperature drops in winter, possibly due to the usage of heater. Therefore, the base-peak boundary should be set at a higher magnitude level when the temperature is lower. Besides, the magnitude of the base-peak boundary in a future day may also relate to the lagging load (the electricity consumption for the previous days and the same day a week ago). The base-peak boundary extraction is aimed to learn the correlation between the base-peak boundary and influencing factors and reflect the variation of the boundary when the influencing factors change.

The procedure for the base and peak loads extraction is shown in Fig. 10. The proposed method essentially utilises a regression model to make a probabilistic forecast of the next day's profile and uses the 0.5 quantile to segment the corresponding observed profile into base and peak loads for that day. Firstly, a probabilistic regression model is trained to output the daily base-peak boundary curve. Various probabilistic regression algorithms can be used such as Gaussian Process Regressor and Quantile Regression Gradient Boosting. When training the model, the input features include:

$$X_d^{\text{train}} = [T_d, L'_{d-1}, L'_{d-2}, L'_{d-3}, L'_{d-7}] \quad (3)$$

where T_d is the average temperature on day d , L'_{d-1} , L'_{d-2} and L'_{d-3} are the normalised total daily electricity consumption of the three previous days; L'_{d-7} is the normalised total daily electricity consumption a week ago (for example, day d is Sunday, then day $d-7$ is the last Sunday).

The output is:

$$Y_d^{\text{train}} = [l'_{d,t}]_{t=1,2\cdots 24} \quad (4)$$

where $l'_{d,t}$ is the normalised electricity consumption during the t^{th} hour on the d^{th} day.

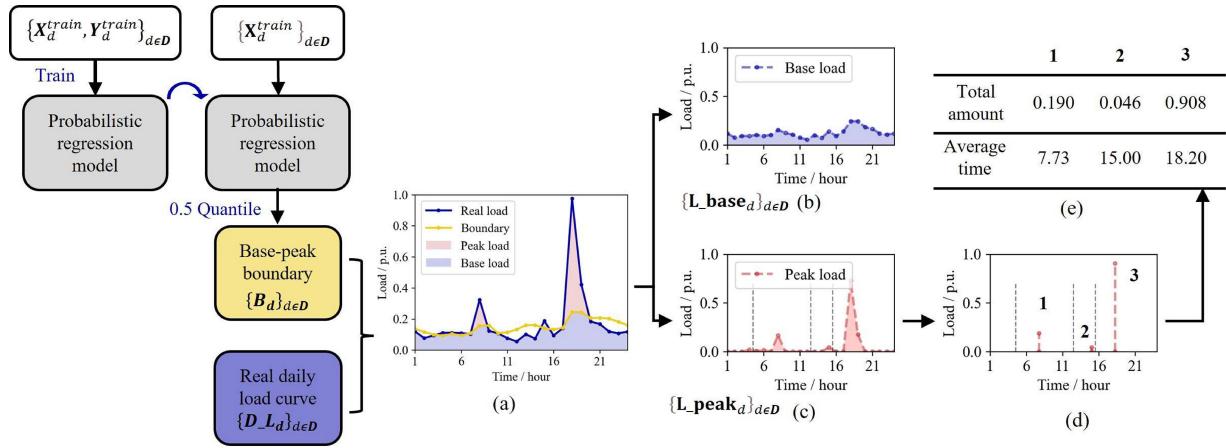


Fig. 10. Framework of base load and peak load extraction (a) real load profile and boundary (b) base load (c) peak load (d) amount and time of peak load shown in figure (e) amount and time of peak load shown in table.

The trained model can make probabilistic forecasting of the day-ahead daily load given the inputs. The output of the model can be in the form of quantiles or probability distribution, depending on the algorithm chosen. By extracting the median value of the probability distribution or the 0.5 quantile from the quantiles, we can get the expected base-peak boundary under the condition of the input features. After training the model by $\{X_d^{\text{train}}, Y_d^{\text{train}}\}_{d \in D}$, we utilise $\{X_d^{\text{train}}\}_{d \in D}$ again as an input for the trained model to get the base-peak boundaries for each day on the training dataset. The obtained boundaries are denoted as $\{B'_d\}_{d \in D}$ (D denotes a set of dates in the training dataset).

An example of the predictive boundary and the real load profile on a day in D is plotted in Fig. 10 (a). The part of the actual daily load lower than the boundary (the blue shaded area) is extracted as the base load, while the part of the actual daily load higher than the boundary (the red shaded area) is extracted as the peak load. The extracted base load and peak load of the example day are shown in Fig. 10 (b) and (c) respectively. Lastly, the amount of peak load and corresponding average time of peak load within each time interval are calculated and shown in Fig. 10 (d) and (e).

C. Base Load and Peak Load Forecasting

The base load, the amount of peak load in each time interval and the average time of peak load in each time interval are predicted separately. Final prediction is obtained by combining the predicted based load and predicted peak load.

As discussed above, it is difficult to predict the peak load of a single household accurately, since it is highly dependent on the consumer behaviour rather than the influencing factors (e.g., weather and time). To better capture the pattern of peak load, the task of predicting the exact shape of peak load is decomposed into several subtasks, namely predicting the total amount of peak load and the average time of peak load for each time interval of the day. The structure of this stage is shown in Fig. 11.

For each time interval, a regression model is trained to learn the relationship between the total amount of peak load

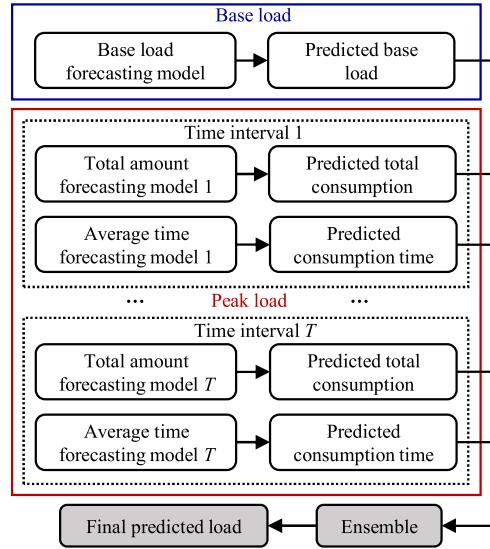


Fig. 11. Framework of base load and peak load forecasting.

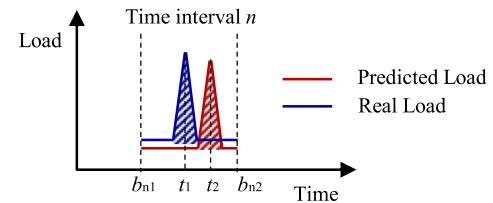


Fig. 12. Actual (real) load and predicted load in time interval n .

consumed in the time interval and the influencing factors including (1) total daily load (the total amount of electricity consumption per day) of the previous three days and the same day a week ago; (2) total peak load in the same time interval of the previous three days and the same day a week ago; and (3) temperatures. In addition, another regression model is trained for each time interval to learn the relationship between the occurrence time of the peak load and the influencing factors, including (1) total daily load of the previous 3 days and the day a week ago; (2) average time of peak load in the same

time interval of the previous 3 days and the same day a week ago; and (3) temperatures.

It should be noted that different machine learning algorithms can be integrated into the above forecasting models. If deterministic machine learning algorithms are integrated into the ‘Base load forecasting model’ and ‘Total amount forecasting models’ in Fig 11, then the final output of this framework will be deterministic load forecasting (certain values of future load). If probabilistic machine learning algorithms are integrated, then the final output will be probabilistic load forecasting (quantiles or probability distribution of future load). However, the choice of regression algorithms is not the main consideration of this paper since our main focus is on the overall performance of the proposed framework, in which multiple models work collaboratively to complete the peak load prediction task.

In this paper, we choose three typical regression algorithms for the deterministic forecasting models, namely Gaussian Process Regression (GP), Random Forest (RF) and Gradient Boosting (GB), and three algorithms for the probabilistic forecasting models, namely Gaussian Process Regression (GP), Quantile regression random Forest (QRRF) and Quantile regression gradient Boosting (QRGB). Here, the output of GP is in the form of Gaussian distribution and it can be easily transformed to a deterministic value (mean value of the distribution), so it can be used for both deterministic and probabilistic forecasting.

D. Evaluation Methods

The evaluation metrics for assessing the accuracy of deterministic forecasting results are Area error, Time error and Dynamic time warping. The evaluation metrics for assessing the probabilistic forecasting results are Interval score (IS) and Quantile score (QS).

A simplified example of real load and predicted load within a time interval are plotted in Fig. 12. The area that real load is higher than predicted load is shaded as blue, and the area that predicted load is higher than real load is shaded as red. The blue area can be regarded as the part of load that is not predicted, and the red area can be regarded as the load that is predicted to occur at a wrong time. The average time of the blue area is t_1 and the average time of the red area is t_2 .

1) Area error: the absolute difference between the area of the red region and the area of the blue region for each time interval. Area error for time interval n can be calculated by:

$$\text{area error}_n = \left| \sum_{t=b_{n1}}^{b_{n2}} (y'_t - y_t) \right| \quad (5)$$

where b_{n1} and b_{n2} are the left boundary and right boundary of the n^{th} time interval. y'_t is predicted load during the t^{th} hour and y_t is real load during the t^{th} hour.

2) Time error: the absolute difference between t_1 and t_2 . The smaller the time error, the smaller the time difference between the real peak load and the predicted peak load. Time error for

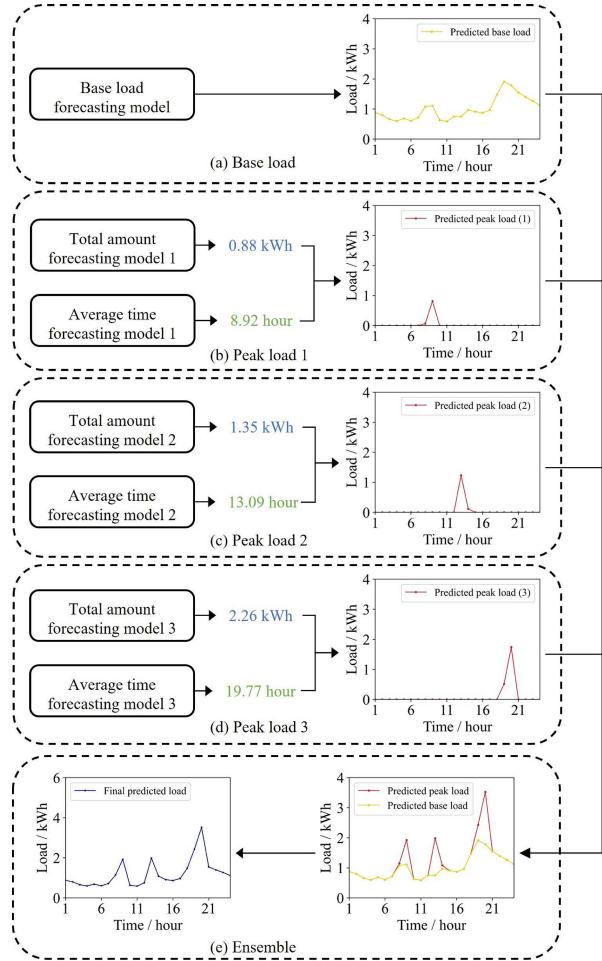


Fig. 13. The process of constructing a load forecast curve (a) base load prediction; peak load prediction of (b) time interval 1, (c) time interval 2 and (d) time interval 3; (e) final prediction.

time interval n can be calculated by:

$$t_1 = \frac{\sum_{t=b_{n1}}^{b_{n2}} [(y'_t - y_t) \cdot I(t) \cdot t]}{\sum_{t=b_{n1}}^{b_{n2}} [(y'_t - y_t) \cdot I(t)]}, I(t) = \begin{cases} 1, & \text{if } y'_t > y_t \\ 0, & \text{if } y'_t < y_t \end{cases} \quad (6)$$

$$t_2 = \frac{\sum_{t=b_{n1}}^{b_{n2}} [(y_t - y'_t) \cdot I(t) \cdot t]}{\sum_{t=b_{n1}}^{b_{n2}} [(y_t - y'_t) \cdot I(t)]}, I(t) = \begin{cases} 1, & \text{if } y_t > y'_t \\ 0, & \text{if } y_t < y'_t \end{cases} \quad (7)$$

$$\text{time error}_n = |t_1 - t_2| \quad (8)$$

3) Dynamic time warping (DTW): the similarity between two time series, which may vary in time or speed by stretching the time series under certain restrictions [20]. Details of DTW can be found in [28].

4) Mean absolute error (MAE):

$$\text{MAE} = \frac{\sum_{t=1}^n |y_t - y'_t|}{n} \quad (9)$$

5) Interval score (IS): evaluating the central $(1-\alpha)*100\%$ interval prediction quality [29]:

$$\text{IS}_{\alpha} = \frac{1}{N} \sum_{i=1}^N (U^i - L^i) + \frac{2}{\alpha} (L^i - x_{obs}^i) \mathbb{1}_{x_{obs}^i < L^i} + \frac{2}{\alpha} (x_{obs}^i - U^i) \mathbb{1}_{x_{obs}^i > U^i} \quad (10)$$

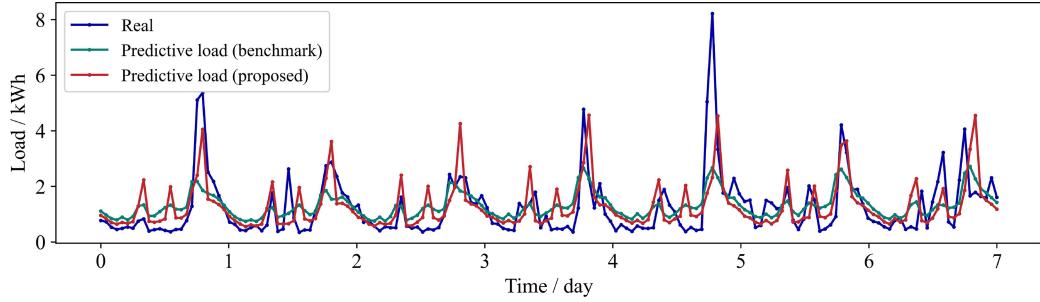


Fig. 14. Deterministic predictions and real values (Day-ahead forecasting of one household over one week).

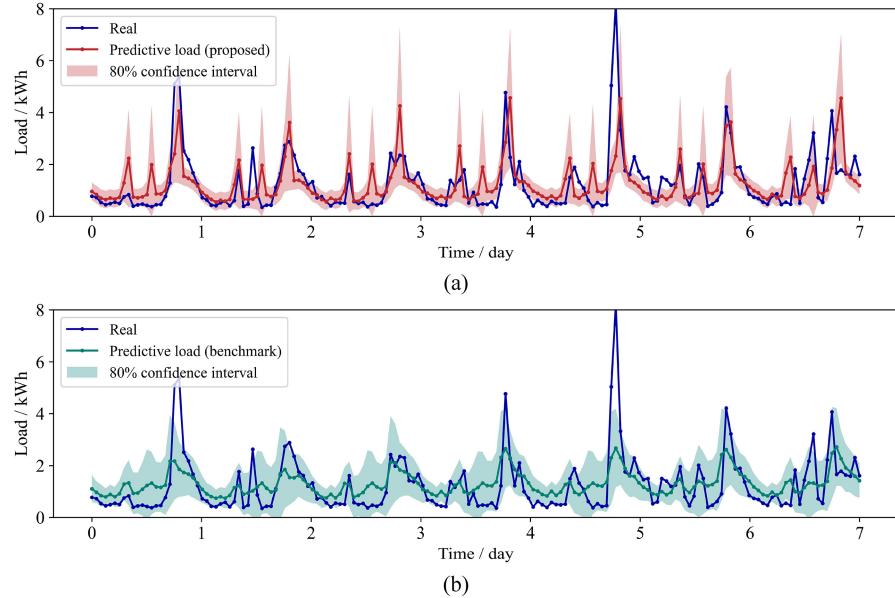


Fig. 15. Probabilistic predictions and real values (Day-ahead forecasting of one household over one week) (a) proposed method (b) benchmark method.

where L^i and U^i represent the predicted lower quantile $\hat{q}_{\tau=\alpha/2}$ and the upper quantile $\hat{q}_{\tau=1-\alpha/2}$, and x_{obs}^i represents the actual value.

When calculating the IS for 80% central interval (Section IV), x_{obs}^i is the actual load in a time interval, and L^i and U^i are the boundaries of 80% interval prediction of x_{obs}^i (i.e., there is 80% probability that x_{obs}^i residing between L^i and U^i).

6) Quantile score (QS): assessing the forecast's performance at some specific quantiles [29]. The QS of probability level τ :

$$QS = \frac{1}{N} \sum_{i=1}^N \psi_\tau (x_{obs}^i - \hat{q}_\tau^i) \quad (11)$$

$$\psi_\tau (u) = \begin{cases} \tau u & \text{if } u \geq 0 \\ (\tau - 1)u & \text{if } u < 0 \end{cases} \quad (12)$$

where \hat{q}_τ^i is the quantile forecast for probability level τ .

When calculating the QS for 0.1 probability level (Section IV), x_{obs}^i is the actual load in a time interval and \hat{q}_τ^i is the 0.1 quantile value (x_{obs}^i is expected to be lower than \hat{q}_τ^i with a probability of 10%).

IV. CASE STUDIES AND DISCUSSIONS

A. Numerical Experiment Setups

Two datasets from two different climate regions are used: (1) the data from the Smart Metering Electricity Customer Behaviour Trials (CBTs) initiated by Commission for Energy Regulation (CER) in Ireland [30]; (2) the solar home electricity data released by Ausgrid, Australia [31]. The CBTs took place during the period of 1st July 2009 - 31st December 2010 with over 5,000 Irish households and businesses participated. The full anonymized data sets comprise half-hourly sampled electricity usage data of each participant. The solar home electricity data was sourced from 300 randomly selected customers in Ausgrid's network area for the period from 1 July 2010 to 30 June 2013.

For the CER dataset, we use one and half year's electricity consumption data (from 1st July 2009 to 31st December 2010), for which each household has complete records of electricity consumption. The data is then divided into a training dataset (from 1st July 2009 to 30th June 2010) and a testing dataset (from 1st July 2010 to 31st December 2010). Similarly, we selected one and half year's electricity consumption data (from 1st January 2011 to 30th June 2012) from the Australian dataset. The first entire year of data is allocated to the training

set and the remaining half-year data is allocated to the testing set. We use the data from the entire year as the training set to ensure that forecasting models can learn different patterns in different seasons. The duration of the testing set is six months, and this allows us to assess the performance of the trained models on both hot and cold days.

The simulation is conducted on a desktop PC with a 3.6 GHz Intel i7 processor and 16 GB of memory. All the models are built in Python.

B. Numerical Experiments Results and Discussions

1) Results of single household

The day-ahead predictions of the electricity consumption of a randomly chosen household from the Irish Commission for Energy Regulation (CER) smart meter dataset on a randomly selected day are shown in Fig. 13. In the figure, the predictions are obtained from the proposed framework integrated with the Gaussian Process Regression (GP). The key steps of obtaining the predicted load are as below:

1) Utilise the trained base load forecasting model to predict the base load for the future 24 hours (i.e., the next day) as shown in Fig. 13 (a).

2) Utilise the trained total amount forecasting models and average time forecasting models to predict the total amount of peak load and the average time of peak load for each of the three time intervals for the next day, as shown in Fig. 13 (b), (c) and (d).

3) The outputs of the above two steps are ensembled to derive the final prediction (the predicted load for the future 24 hours) as shown in Fig. 13 (e).

In Fig. 13 (b), (c) and (d), the predictions of the peak load are [0.88, 1.35, 2.26] (kWh) at [8.92, 13.09, 19.77] (hour). Here, 0.88, 1.35 and 2.26 are the predicted amount of peak consumption in each of the three time intervals of the next day; and 8.92, 13.09 and 19.77 indicate the predicted occurrence time of the three peak loads.

It should be noted that the peak load occurrence time should be a time interval instead of a time instant. In Fig. 13 (b), (c) and (d), each peak load occurrence time represents the median time instant of peak consumption period. Each forecasted peak load is assigned to the adjacent two hours close to the predicted time instant. For example, the first peak load is 0.88 kWh and it is expected to occur at hour 8.92. This is transformed to 0.07 kWh consumed during the 8th hour of the day (7am-8am) and 0.81 kWh during the 9th hour (8am-9am). This allocation is weighted by the distance between 8.92 and 8 and that between 8.92 and 9. After this transformation, the predicted peak load within each of the time intervals are plotted in Fig. 13 (b), (c) and (d) respectively.

Further, the day-ahead deterministic load forecasting results of the consumer over a week is shown in Fig. 14. The forecast is constructed by integrating the GP in the proposed framework. In the figure, the blue line denotes the actual electricity consumption, the red line denotes the predicted load, and the green line denotes the predicted load obtained by the benchmark method, i.e., the original GP without integrating it into the proposed framework.

It can be seen from Fig. 14 that the green line (obtained by the benchmark method) is much flatter than the red line (obtained by integrating the GP into the proposed framework). This is because the benchmark model is trained through a pointwise approach. To avoid the double penalty of mispredicting peak load and to minimise the loss function, the trained model tends to ‘ignore’ some part of the peak load and makes a conservative estimation. On the other hand, the prediction of the proposed method provides more effective information of peak load. Though there may be a small discrepancy between the occurrence time of the actual peak load and the predicted peak load, it is better than the flat prediction without the peak load. Also, it can be observed that the proposed method may predict small peaks that do not show up in the real consumption curve. This is because that the prediction is generated based on the consumption patterns in historical data, but the real load is determined by the highly uncertain consumer behaviour. For such cases, the proposed method is still better than the benchmark method, since the benchmark method also tends to output loads higher than the real loads and the area error (the area between the predicted load curve and the real load curve) of the proposed method is smaller than the benchmark method.

The day-ahead probabilistic load forecasting results of the consumer over a week are shown in Fig. 15 (a) (the integration of the GP and the proposed method) and Fig. 15 (b) (the benchmark method, the GP without integrating the proposed method). It can be seen that the 80% confidence interval generated by the proposed method is narrower than that of the benchmark method. This implies that the proposed method outperforms the benchmark method.

2) Results of 400 households

The performance of the integration of the proposed load forecasting framework and three deterministic machine learning regression algorithms (GP, RF and GB) and three probabilistic machine learning regression algorithms (GP, QRRF and QRGB) are verified on 200 household selected from the CER dataset and 200 households selected from the Ausgrid dataset.

Table I shows the day-ahead deterministic load forecasting results of the six algorithms (Original GP, Proposed framework + GP, Original RF, Proposed framework + RF, Original GB, Proposed framework + GB) for 200 households from the CER dataset evaluated by Area Error, Time Error, DTW and MAE respectively. It can be observed from Table I that, the Proposed framework + GP, Proposed framework + RF and Proposed framework + GB outperform the original GP, RF and GB in terms of Area Error (the lower the Area Error, the better the predictor). Area error measures the similarity between the shape of the predicted peak load and real peak load. On average, the performance of GP improves 2.02% after integrating with the proposed framework. For RF and GB, the increase is 3.27% and 2.48% respectively. Time error measures the difference in timing between the predicted peak load and real peak load. In terms of Time error, the performance of GP, RF and GB improves 18.26%, 8.52% and 10.61% respectively after integrating with the proposed framework. DTW measures the overall shape-based distance

TABLE I
 (A) AREA ERROR (B) TIME ERROR (C) DTW (D) MAE OF DIFFERENT DETERMINISTIC FORECASTING METHODS TESTED ON
 200 HOUSEHOLDS FROM THE CER DATASET

Evaluation metrics	Original GP	Proposed method + GP	Improvement	Original RF	Proposed method + RF	Improvement	Original GB	Proposed method + GB	Improvement
Area Error	2.5202	2.4693	2.02%	2.6592	2.5693	3.27%	2.6825	2.6160	2.48%
Time Error	2.7917	2.2819	18.26%	2.7717	2.5356	8.52%	2.6546	2.3729	10.61%
DTW	1.1075	1.0567	4.59%	1.1127	1.0403	6.51%	1.1144	1.0635	4.30%
MAE	0.5849	0.6500	-11.13%	0.6242	0.6093	2.39%	0.6375	0.6374	0.01%

TABLE II
 (A) AREA ERROR (B) TIME ERROR (C) DTW (D) MAE OF DIFFERENT DETERMINISTIC FORECASTING METHODS TESTED ON
 200 HOUSEHOLDS FROM THE AUSTRALIAN DATASET

Evaluation metrics	Original GP	Proposed method + GP	Improvement	Original RF	Proposed method + RF	Improvement	Original GB	Proposed method + GB	Improvement
Area Error	1.4558	1.3704	5.87%	1.5446	1.4854	5.79%	1.5412	1.4981	2.80%
Time Error	2.8692	2.2106	22.95%	2.8927	2.5826	9.69%	2.8231	2.4864	11.92%
DTW	0.9966	0.9548	4.19%	1.0160	0.9568	5.83%	1.0112	0.9636	4.71%
MAE	0.3086	0.3286	-7.36%	0.3239	0.3162	2.38%	0.3267	0.3234	1.01%

TABLE III
 (A) INTERVAL SCORE (FOR 80% CENTRAL PREDICTION INTERVAL) (C) QUANTILE SCORE (FOR 0.1 PROBABILITY LEVEL) (D) QUANTILE SCORE (FOR 0.9 PROBABILITY LEVEL) OF DIFFERENT PROBABILISTIC FORECASTING METHODS TESTED ON 200 HOUSEHOLDS FROM THE CER DATA

Evaluation metrics	Original GP	Proposed method + GP	Improvement	Original QRRF	Proposed method + QRRF	Improvement	Original QRGB	Proposed method + QRGB	Improvement
IS (80%)	1.8500	1.6016	13.43%	1.5527	1.4507	6.57%	1.3855	1.3537	2.29%
QS (0.1)	0.0890	0.0735	17.40%	0.0577	0.0570	1.21%	0.0551	0.0533	3.27%
QS (0.9)	0.0960	0.0866	9.79%	0.0981	0.0871	11.21%	0.0835	0.0820	1.80%

TABLE IV
 (A) INTERVAL SCORE (FOR 80% CENTRAL PREDICTION INTERVAL) (C) QUANTILE SCORE (FOR 0.1 PROBABILITY LEVEL) (D) QUANTILE SCORE (FOR 0.9 PROBABILITY LEVEL) OF DIFFERENT PROBABILISTIC FORECASTING METHODS TESTED ON 200 HOUSEHOLDS FROM THE AUSTRALIAN DATASET

Evaluation metrics	Original GP	Proposed method + GP	Improvement	Original QRRF	Proposed method + QRRF	Improvement	Original QRGB	Proposed method + QRGB	Improvement
IS (80%)	1.4100	1.3725	2.66%	1.3846	1.3546	2.17%	1.3835	1.3538	2.15%
QS (0.1)	0.0522	0.0499	4.41%	0.0515	0.0490	4.85%	0.0514	0.0491	4.47%
QS (0.9)	0.0888	0.0873	1.69%	0.0871	0.0863	0.92%	0.0870	0.0862	1.03%

between predictions and observations. In terms of DWT, the proposed frameworks integrated with the three algorithms also outperform the original algorithms by 4.59%, 6.51% and 4.30% respectively. However, the error of proposed method + GP is higher than the original GP evaluated by MAE. This is because MAE is a point-wise approach and mis-predicting the timing of peak load leads to double penalty as discussed above.

Table II shows the day-ahead deterministic load forecasting results of the proposed methods on 200 households from the Ausgrid dataset evaluated by Area Error, Time Error, DTW and MAE respectively. It can also be seen that the proposed methods (Proposed framework + GP, Proposed framework + RF and Proposed framework + GB) outperform the original

methods (Original GP, Original RF, and Original GB), with averages of 5.87%, 5.79% and 2.8% decrease in Area Error, averages of 22.95%, 9.69% and 11.92% decrease in Time error and averages of 4.19%, 5.83% and 4.71% decrease in DTW.

Table III shows the day-ahead probabilistic load forecasting results of the six algorithms (Original GP, Proposed framework + GP, Original QRRF, Proposed framework + QRRF, Original QRGB, Proposed framework + QRGB) for 200 households from the CER dataset evaluated by Interval score (for 80% central prediction interval), Quantile score (for 0.1 probability level) and Quantile score (for 0.9 probability level) respectively. It can be observed from Table III that, the Proposed method + GP, Proposed method + QRRF and Proposed framework + QRGB outperform the original GP, QRRF and

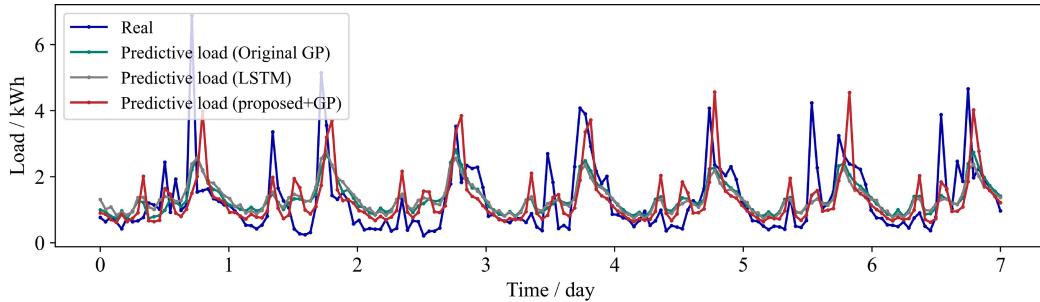


Fig. 16. Day-ahead forecasting of an individual household over one week derived from the original GP, the integration of the GP and the proposed method, and LSTM.

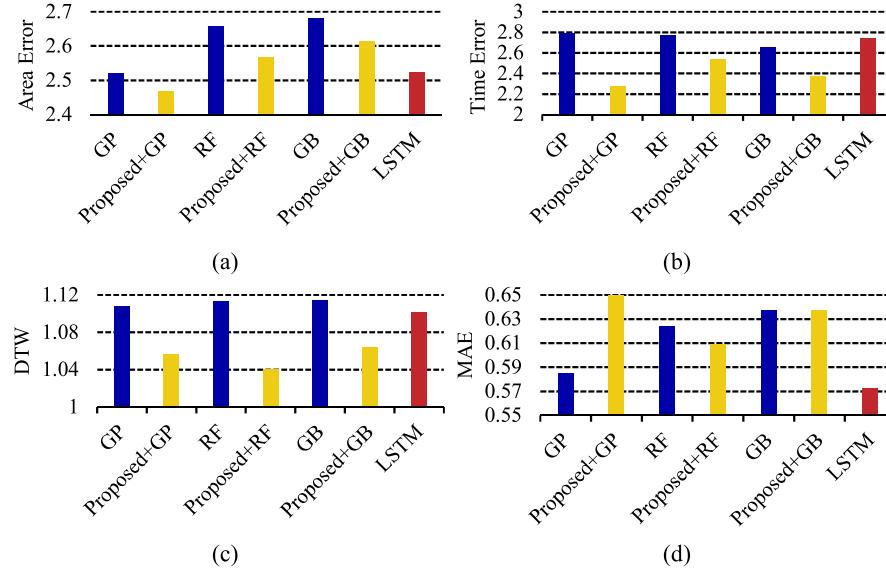


Fig. 17. (a) Area error (b) Time error (c) DTW (d) MAE of different deterministic forecasting methods testing on 200 households from the CER dataset.

QRGB in terms of all three evaluation scores. IS can evaluate the quality of the 80% central prediction interval. In terms of IS, the performance of the GP, QRRF and QRGB improves 13.43%, 6.57% and 2.29% respectively after integrating with the proposed framework. QS measures the prediction quality at a specific quantile. In terms of QS for 0.1 probability level, the proposed frameworks integrated with the three algorithms also outperform the original algorithms by 17.40%, 1.21% and 3.27% respectively, while for 0.9 probability level, the improvement reaches 9.79%, 11.21% and 1.80% respectively.

Table IV shows the day-ahead probabilistic load forecasting results of the proposed methods for 200 households from the Ausgrid dataset evaluated by IS (for 80% central prediction interval), QS (for 0.1 probability level) and QS (for 0.9 probability level) respectively. It can also be seen that the proposed methods (Proposed framework + GP, Proposed framework + QRRF and Proposed framework + QRGB) outperform the original methods (Original GP, Original QRRF, and Original QRGB), with averages of 2.66%, 2.17% and 2.15% decrease in IS (80%), averages of 4.41%, 4.85% and 4.47% decrease in QS (0.1) and averages of 1.69%, 0.92% and 1.03% decrease in QS (0.9).

3) Comparison with deep learning method

To further prove the effectiveness of the proposed method, an experiment was conducted to compare the performance of

the proposed method with that of long short-term memory networks (LSTM), a deep learning model. In [17], an LSTM recurrent neural network-based method was proposed for the short-term load forecasting of individual residential household and was proved to have good performance. In this case study, we implemented an LSTM like the one developed in [17].

The day-ahead deterministic load forecasting results derived from three algorithms (Original GP, Proposed framework + GP, Original RF, Proposed framework + RF, Original GB, Proposed framework + GB and LSTM) of one randomly selected consumer over a week are shown in Fig. 16. It can be seen that the outputs of LSTM are close to that of the original GP model. Compared to the LSTM and the original GP models, the proposed method provides more effective information of peak load. This is similar to the conclusions that are drawn from Fig. 14.

The average day-ahead deterministic load forecasting performance of the seven algorithms (Original GP, Proposed framework + GP, Original RF, Proposed framework + RF, Original GB, Proposed framework + GB and LSTM) on 200 households from the CER dataset evaluated by Area Error, Time Error, DTW and MAE are shown in Fig. 17 (a), (b), (c) and (d) respectively. In terms of point-wise evaluation metric MAE, LSTM outperforms all other load forecasting methods as shown in Fig. 17 (d) (the smaller the MAE, the better the forecasting). However, in terms of shape-based evaluation metric DTW, the proposed methods

(Proposed framework + GP, Proposed framework + RF and Proposed framework + GB) outperform LSTM as shown in Fig. 17 (c) (the smaller the DTW, the better the forecasting). Also, Proposed framework + GP outperforms LSTM in terms of Area Error, and all the proposed methods (Proposed framework + GP, Proposed framework + RF and Proposed framework + GB) outperform LSTM evaluated by Time Error. In summary, LSTM achieves the highest forecasting accuracy in terms of MAE. However, LSTM cannot well solve the uncertainty in occurrence time of peak load of an individual household and the proposed method outperforms LSTM in terms of shape-based evaluation metrics.

V. CONCLUSION

This paper proposed a new framework for day-ahead load forecasting of single household. The peak load patterns are learned in a targeted manner in this framework aiming at better capturing the peak load for single household with highly random and volatile demand. The proposed framework has been successfully integrated into existing deterministic and probabilistic load forecasting algorithms for them to achieve an improved forecasting performance. In the case studies, various metrices all proved the advantages of the proposed method over the benchmark methods.

REFERENCES

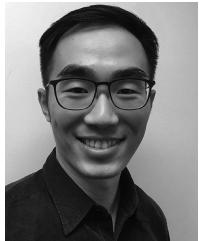
- [1] F. Luo, G. Ranzi, C. Wan, Z. Xu, and Z. Y. Dong, "A multistage home energy management system with residential photovoltaic penetration," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 116–126, Jan. 2019, doi: [10.1109/TII.2018.2871159](https://doi.org/10.1109/TII.2018.2871159).
- [2] B. Juddi, Y. Mishra, and G. Ledwich, "Differential dynamic programming based home energy management scheduler," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1427–1437, Jul. 2020, doi: [10.1109/TSTE.2019.2927237](https://doi.org/10.1109/TSTE.2019.2927237).
- [3] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3125–3148, May 2019, doi: [10.1109/TSG.2018.2818167](https://doi.org/10.1109/TSG.2018.2818167).
- [4] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy forecasting: A review and outlook," *IEEE Open Access J. Power Energy*, vol. 7, pp. 376–388, 2020, doi: [10.1109/OAJPE.2020.3029979](https://doi.org/10.1109/OAJPE.2020.3029979).
- [5] S.-J. Huang and K.-R. Shih, "Short-term load forecasting via ARMA model identification including non-Gaussian process considerations," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 673–679, May 2003, doi: [10.1109/TPWRS.2003.811010](https://doi.org/10.1109/TPWRS.2003.811010).
- [6] J. C. Lopez, M. J. Rider, and Q. Wu, "Parsimonious short-term load forecasting for optimal operation planning of electrical distribution systems," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1427–1437, Mar. 2019, doi: [10.1109/TPWRS.2018.2872388](https://doi.org/10.1109/TPWRS.2018.2872388).
- [7] D. Niu, Y. Wang, and D. D. Wu, "Power load forecasting using support vector machine and ant colony optimization," *Expert Syst. Appl.*, vol. 37, no. 3, pp. 2531–2539, Mar. 2010, doi: [10.1016/j.eswa.2009.08.019](https://doi.org/10.1016/j.eswa.2009.08.019).
- [8] M. Beccali, M. Cellura, V. Lo Brano, and A. Marvuglia, "Forecasting daily urban electric load profiles using artificial neural networks," *Energy Convers. Manage.*, vol. 45, nos. 18–19, pp. 2879–2900, Nov. 2004, doi: [10.1016/j.enconman.2004.01.006](https://doi.org/10.1016/j.enconman.2004.01.006).
- [9] B. Wang, M. Mazhari, and C. Y. Chung, "A novel hybrid method for short-term probabilistic load forecasting in distribution networks," *IEEE Trans. Smart Grid*, vol. 13, no. 5, pp. 3650–3661, Sep. 2022, doi: [10.1109/TSG.2022.3171499](https://doi.org/10.1109/TSG.2022.3171499).
- [10] T. Zufferey, A. Ulbig, S. Koch, and G. Hug, "Forecasting of smart meter time series based on neural networks," in *Proc. Data Anal. Renew. Energy Integr.*, vol. 10097, 2017, pp. 10–21, doi: [10.1007/978-3-319-50947-1_2](https://doi.org/10.1007/978-3-319-50947-1_2).
- [11] P. G. Da Silva, D. Ilic, and S. Karnouskos, "The impact of smart grid prosumer grouping on forecasting accuracy and its benefits for local electricity market trading," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 402–410, Jan. 2014, doi: [10.1109/TSG.2013.2278868](https://doi.org/10.1109/TSG.2013.2278868).
- [12] Y. Wang, N. Gao, and G. Hug, "Personalized federated learning for individual consumer load forecasting," *CSEE J. Power Energy Syst.*, vol. 9, no. 1, pp. 326–330, Jan. 2023, doi: [10.17775/CSEEPES.2021.07350](https://doi.org/10.17775/CSEEPES.2021.07350).
- [13] A. Taik and S. Cherkaoui, "Electrical load forecasting using edge computing and federated learning," in *Proc. ICC IEEE Int. Conf. Communications (ICC)*, Dublin, Ireland: 2020, pp. 1–6, doi: [10.1109/ICC40277.2020.9148937](https://doi.org/10.1109/ICC40277.2020.9148937).
- [14] Y. Wang, I. L. Bennani, X. Liu, M. Sun, and Y. Zhou, "Electricity consumer characteristics identification: A federated learning approach," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3637–3647, Jul. 2021, doi: [10.1109/TSG.2021.3066577](https://doi.org/10.1109/TSG.2021.3066577).
- [15] M. Chaouch, "Clustering-based improvement of nonparametric functional time series forecasting: Application to intra-day household-level load curves," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 411–419, Jan. 2014, doi: [10.1109/TSG.2013.2277171](https://doi.org/10.1109/TSG.2013.2277171).
- [16] C.-N. Yu, P. Mirowski, and T. K. Ho, "A sparse coding approach to household electricity demand forecasting in smart grids," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 738–748, Mar. 2017, doi: [10.1109/TSG.2015.2513900](https://doi.org/10.1109/TSG.2015.2513900).
- [17] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019, doi: [10.1109/TSG.2017.2753802](https://doi.org/10.1109/TSG.2017.2753802).
- [18] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," 2014, arXiv: 1409.3215. [Online]. Available: <http://arxiv.org/abs/1409.3215>
- [19] C. Keil and G. C. Craig, "A displacement and amplitude score employing an optical flow technique," *Weather Forecasting*, vol. 24, no. 5, pp. 1297–1308, Oct. 2009, doi: [10.1175/2009WAF222247.1](https://doi.org/10.1175/2009WAF222247.1).
- [20] S. Haben, J. Ward, D. V. Greetham, C. Singleton, and P. Grindrod, "A new error measure for forecasts of household-level, high resolution electrical energy consumption," *Int. J. Forecast.*, vol. 30, no. 2, pp. 246–256, Jun. 2014, doi: [10.1016/j.ijforecast.2013.08.002](https://doi.org/10.1016/j.ijforecast.2013.08.002).
- [21] T. Teeraratkul, D. O'Neill, and S. Lall, "Shape-based approach to household electric load curve clustering and prediction," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5196–5206, Sep. 2018, doi: [10.1109/TSG.2017.2683461](https://doi.org/10.1109/TSG.2017.2683461).
- [22] D. Sakurai, Y. Fukuyama, T. Iizaka, and T. Matsui, "Daily peak load forecasting by artificial neural network using differential evolutionary particle swarm optimization considering outliers," *IFAC-PapersOnLine*, vol. 52, no. 4, pp. 389–394, 2019, doi: [10.1016/j.ifacol.2019.08.241](https://doi.org/10.1016/j.ifacol.2019.08.241).
- [23] J. Moon, S. Park, E. Hwang, and S. Rho, "A hybrid tree-based ensemble learning model for day-ahead peak load forecasting," in *Proc. 15th Int. Conf. Hum. Syst. Interact. (HSI)*, Melbourne, Australia, 2022, pp. 1–6, doi: [10.1109/HSI55341.2022.9869440](https://doi.org/10.1109/HSI55341.2022.9869440).
- [24] I. P. Panapakidis, G. C. Christoforidis, N. Asimopoulos, and A. S. Dagoumas, "Combining wavelet transform and support vector regression model for day-ahead peak load forecasting in the Greek power system," in *Proc. IEEE Int. Conf. Environ. Electr. Eng. IEEE Ind. Commer. Power Syst. Eur. (EEEIC I CPS Europe)*, Milan, Italy, 2017, pp. 1–6, doi: [10.1109/EEEIC.2017.7977759](https://doi.org/10.1109/EEEIC.2017.7977759).
- [25] G. F. D. Apolinario, Y.-Y. Hong, Y.-D. Lee, J.-L. Jiang, and S.-S. Wang, "Comparative studies on use of discrete wavelet transform-based feature extraction for peak load forecasting using LSTM," in *Proc. IEEE 4th Int. Conf. Power Energy Appl. (ICPEA)*, Busan, Republic of South Korea 2021, pp. 88–92, doi: [10.1109/ICPEA52760.2021.9639283](https://doi.org/10.1109/ICPEA52760.2021.9639283).
- [26] Z. Yu, Z. Niu, W. Tang, and Q. Wu, "Deep learning for daily peak load forecasting—A novel gated recurrent neural network combining dynamic time warping," *IEEE Access*, vol. 7, pp. 17184–17194, 2019, doi: [10.1109/ACCESS.2019.2895604](https://doi.org/10.1109/ACCESS.2019.2895604).
- [27] Z. Xia, H. Ma, T. K. Saha, and R. Zhang, "Consumption scenario-based probabilistic load forecasting of single household," *IEEE Trans. Smart Grid*, vol. 13, no. 2, pp. 1075–1087, Mar. 2022, doi: [10.1109/TSG.2021.3132039](https://doi.org/10.1109/TSG.2021.3132039).
- [28] M. Müller, *Information Retrieval for Music and Motion*. Berlin Germany: Springer, 2007, doi: [10.1007/978-3-540-74048-3](https://doi.org/10.1007/978-3-540-74048-3).
- [29] P. Lauret, M. David, and P. Pinson, "Verification of solar irradiance probabilistic forecasts," *Solar Energy*, vol. 194, pp. 254–271, Dec. 2019, doi: [10.1016/j.solener.2019.10.041](https://doi.org/10.1016/j.solener.2019.10.041).
- [30] Commission for Energy Regulation (CER), 2012, "CER Smart Metering Project—Electricity Customer Behaviour Trial, 2009–2010 [dataset]," 1st Edition, Irish Social Science Data Archive. [Online]. Available: <https://www.ucd.ie/issda/data/commissionforenergyregulationcer/>
- [31] *Solar Home Electricity Data*, Ausgrid Co., Sydney, NSW, Australia, 2014.



Zhong Xia (Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2015 and 2018, respectively. He is currently pursuing the Ph.D. degree with the School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, Australia. His research interests include smart meter data analysis, load forecasting, and home energy management.



Hui Ma (Senior Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from Xi'an Jiaotong University, China, the M.Phil. degree in electrical engineering from Nanyang Technological University, Singapore, and the Ph.D. degree in electronic engineering from The University of Adelaide. He is currently a Senior Lecturer with the School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, Australia. He has seven years of industry experience. His current research interests include industrial informatics, condition monitoring and diagnosis, power systems, wireless sensor networks, and sensor signal processing.



Ruiyuan Zhang (Member, IEEE) received the B.Eng. degree in electrical engineering from Tongji University, Shanghai, China, in 2012, the master's degree in electrical power engineering from RWTH Aachen University, Germany, in 2015, and the Ph.D. degree from The University of Queensland, Australia, in 2021, where he was a Postdoctoral Research Fellow with the Centre for Energy Data Innovation and the School of Information Technology and Electrical Engineering. He is currently working as a Postdoctoral Research Fellow with the School of Engineering, The Hong Kong University of Science and Technology, Hong Kong, China. His research interests include predictive analysis, quantitative modeling, and deep learning.



Tapan Kumar Saha (Fellow, IEEE) received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, in 1982, the M.Tech. degree in electrical engineering from the Indian Institute of Technology Delhi in 1985, and the Ph.D. degree from The University of Queensland, Brisbane, QLD, Australia, in 1994. He is currently a Professor of Electrical Engineering and the Leader of Power, Energy and Control Engineering Discipline within the School of Information Technology and Electrical Engineering, The University of Queensland (UQ). He is also the founding Director of Australasian Transformer Innovation Center and the Leader of UQ Solar and UQ Industry 4.0 Energy TestLab. His research interests include condition monitoring of electrical assets, power systems, and renewable energy integration to the grid. He is a Registered Professional Engineer of the State of Queensland. He is a Fellow and a Chartered Professional Engineer of the Institution of Engineers, Australia.