

Personalized Federated Learning With Cost-Oriented Load Forecasting for Home Energy Management Systems

Sara Barja-Martinez¹, Fei Teng², *Senior Member, IEEE*, Adrià Junyent-Ferré¹, *Senior Member, IEEE*, and Mònica Aragüés-Peñalba¹, *Member, IEEE*

Abstract—Accurate day-ahead demand forecasting is crucial for optimizing the performance of home energy management systems. Traditional forecasting methods often decouple the forecasting task and the subsequent decision making, resulting in imbalanced economic penalties from load deviations. Furthermore, the rise of digitization has led to a massive increase in fine-grained smart meter data stored daily, posing significant challenges to customers' data privacy and security. To address these technical challenges, this study proposes a personalized federated learning methodology that incorporates a cost-oriented loss function. This methodology is designed to learn end-user-specific patterns, reduce penalization costs, and preserve customer privacy. Comparative analyses reveal that the proposed method, which utilizes a cost-oriented loss function and L^2 regularization, outperforms traditional symmetric loss functions in terms of efficiency and economic benefits. The results confirm that this personalized federated learning approach consistently achieves the lowest error rates and penalization costs compared to other methods. Additionally, sensitivity analyses indicate that even households with limited historical consumption data can achieve accurate load predictions using the personalized federated learning approach.

Index Terms—Cost-oriented, imbalances, privacy preservation, federated learning, load forecasting.

I. INTRODUCTION

THE energy sector is rapidly transforming towards a distributed, digitized and decarbonized system [1]. The advent

of digitization has revolutionized household energy management, providing consumers with real-time granular electricity consumption and self-generation data through smart meters, enabling the development of data-driven energy services [2]. However, the increasing volume of data brings challenges in ensuring data privacy and security, especially under strict regulations like the General Data Protection Regulation (GDPR) enforced by the European Union [3].

Federated Learning (FL) has emerged as a promising solution to these challenges, offering a framework for training artificial intelligence models across multiple decentralized devices without the need to exchange the local data samples with a central server. This approach preserves data privacy, reduces cloud computing and storage costs, and enables advanced data analytics [4], [5]. General FL concepts are reviewed in [6], which defines FL architectures and applications and provides a comprehensive survey of existing works on this subject. In contrast, [7], [8] focus on the challenges and future directions.

In the context of energy systems, [9] proposes a novel federated reinforcement learning approach for managing energy in smart homes with solar photovoltaic systems and energy storage. Similarly, [10] reviews FL algorithms and their applications in renewable energy, including case studies on energy generation and consumption, evaluating the potential and challenges of FL in power contexts and outlining future research directions. FL is particularly beneficial for addressing customer data privacy and security concerns when using end-users smart meter data [11]. One of the most important tasks in an energy management system is to accurately predict consumption and generation. In this regard, [12] reports one of the first uses of FL in predicting load for microgrid energy management systems. Various studies have examined short-term load prediction using FL, such as [13], which suggests various federated aggregation methods for load forecasting, with FedAvg achieving the best results. In contrast, [14] focuses on the privacy and security aspects of data using FL. Additionally, [15] proposes a federated transfer learning approach that uses a deep federated adaptation approach.

Developing a common global FL model for a non-independent and non-identically distributed (non-iid) group of customers is challenging, as the model may not effectively capture all customers' patterns. To address this issue, Personalized FL (PFL) has emerged as a solution. In literature, recent advancements in PFL further enhance model performance by

Received 15 February 2024; revised 16 June 2024; accepted 29 July 2024. Date of publication 17 September 2024; date of current version 31 January 2025. Paper 2023-KDSEM-1541.R1, approved for publication in the IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS by the Knowledge-and Data-Driven Smart Energy Management in Distribution Networks of the IEEE Industry Applications Society. This work was supported in part by the I+D+i project ATLAS Digitalization using novel data analytic methods and toolboxes for secure, renewable, and flexible grids, with reference PID2021-128101OB-I00 funded by the Ministry of Science, Innovation, and Universities of Spain under Grant MCIN/AEI/10.13039/501100011033, and in part by the ERDF as a way of making Europe. (Corresponding author: Sara Barja-Martinez.)

Sara Barja-Martinez is with the Universitat Politècnica de Catalunya, Electrical Engineering, 08034 Barcelona, Spain (e-mail: sara.barja@upc.edu).

Fei Teng is with Imperial College London, London SW7 2AZ, U.K. (e-mail: f.teng@imperial.ac.uk).

Adrià Junyent-Ferré is with the Imperial College London Department of Electrical and Electronic Engineering, London SW7 2AZ, U.K. (e-mail: adria.junyent-ferre@imperial.ac.uk).

Mònica Aragüés-Peñalba is with the Universitat Politècnica de Catalunya, 08034 Barcelona, Spain (e-mail: monica.aragues@upc.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TIA.2024.3462668>.

Digital Object Identifier 10.1109/TIA.2024.3462668

customizing and fine-tuning the global FL model to individual users' data. PFL addresses the non-iid nature of household energy consumption data by allowing local retraining on each edge device, significantly improving load forecast accuracy [16]. Several reviews, including those by [17] and [18], provide detailed studies on PFL model architectures, their benefits, limitations, and applicability in different scenarios. In [19], a novel PFL approach balances global data utilization with local data distribution, enhancing forecasting accuracy while protecting privacy. Another study by [20] introduces a PFL user-level load forecasting system that incorporates a generative adversarial network-based differential privacy algorithm to ensure data privacy. Similarly, [19] presents a PFL for individual consumer load forecasting, emphasizing improved accuracy and privacy protection. Furthermore, [21] proposes a short-term photovoltaic power forecasting method that incorporates a personalization technique and a semi-asynchronous aggregation strategy to handle data heterogeneity. Another PFL approach for load forecasting is presented in [22], comparing its performance against centralized and localized methods.

Accurate load forecasting is essential for the efficient operation of energy markets, as forecasting errors can lead to significant financial penalties. Electricity retailers are penalized for deviations between forecasted and actual consumption, with over-predictions resulting in excess energy purchases and under-predictions leading to insufficient supply. These imbalance costs are asymmetric and vary with market conditions and regulations [23]. This work considers the penalization costs assumed by the electricity retailer companies in household load forecasting. By integrating cost-oriented loss functions into forecasting models, the economic impact of these errors can be better captured. Previous studies, such as those by [24], [25], and [26], have developed models using cost-oriented loss functions for various energy forecasting applications. However, these studies operated with centralized models and focused on demand prediction at a state or regional level, rather than at the household level.

This work addresses a gap in existing methodologies by introducing a cost-oriented loss function into FL methods for household load prediction. By tackling both the technical challenges of data privacy and the economic implications of forecasting errors, this approach improves the overall performance and cost-effectiveness of load forecasting models in the energy market. Unlike traditional methods that use symmetric loss functions like Mean Squared Error (MSE), the proposed approach accounts for the asymmetry in the costs of over-prediction and under-prediction. To the best of our knowledge, no prior studies have integrated a cost-oriented loss function within FL methods for household load prediction purposes. While [12] reports the first use of FL to support microgrid energy management systems in load prediction using personalization, the cost-oriented approach is not included in any global or personalized FL models. Similarly, [19] introduces a personalized FL approach for individual customer load forecasting but does not incorporate a cost-oriented loss function.

This research makes the following original contributions to the field of personalized federated learning for energy management systems' load forecasting, specifically addressing economic impacts through cost-oriented loss functions. Novel contributions include:

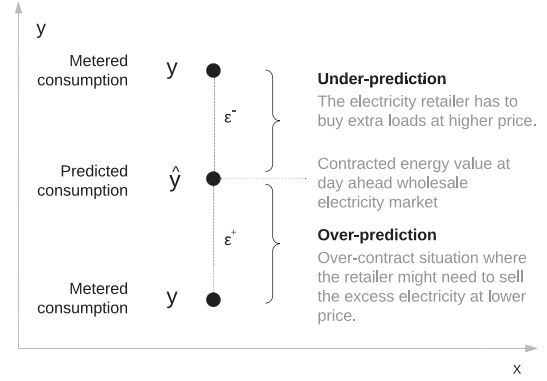


Fig. 1. Financial impacts of prediction errors in the context of load consumption.

- 1) A cost-oriented asymmetric loss function is integrated into FL methods for household load forecasting. This addresses the economic impact of forecasting errors by prioritizing the minimization of costs associated with under and over-predictions.
- 2) The study compares the error performance and cost efficiency of different FL methods against benchmark models. It demonstrates and validates cost savings in imbalance fees achieved by applying cost-oriented FL models.
- 3) A new personalization approach is introduced, combining two state-of-the-art methods. This approach includes a regularization component during both the training of the local model in communication rounds and the final step of local personalization retraining, aiming to enhance the personalization and overall performance of the FL model.
- 4) A sensitivity study identifies the method that performs best in terms of error and cost with a set of limited historical data available for training.

This work is organized as follows. Section II formulates the cost-oriented loss function for load forecasting. Section III presents the local learning and federated learning methods that integrate the cost-oriented loss function. Section IV describes the case studies and provides results and validation of the proposed approach. Finally, Section V concludes the work, summarizing the findings and suggesting potential directions for future research.

II. COST-ORIENTED LOSS FUNCTION

Traditional symmetric loss functions, such as MSE or Mean Absolute Error (MAE), impose equal penalties on over-predictions and under-predictions. Nevertheless, it is essential to distinguish between these errors when forecasting electricity load consumption, as emphasized by [27]. Fig. 1 illustrates the distinct economic consequences of these forecasting inaccuracies. An under-prediction occurs when the predicted consumption ($\hat{y}$) is less than the actual metered consumption (y), denoted as (ϵ^-). This scenario forces the electricity retailer to purchase additional loads at higher prices or incur penalties for not meeting the demand. In contrast, over-prediction, represented as (ϵ^+), the forecast exceeds the actual consumption, leading to surplus electricity that might be sold at potentially lower prices. The figure highlights the differing financial impacts

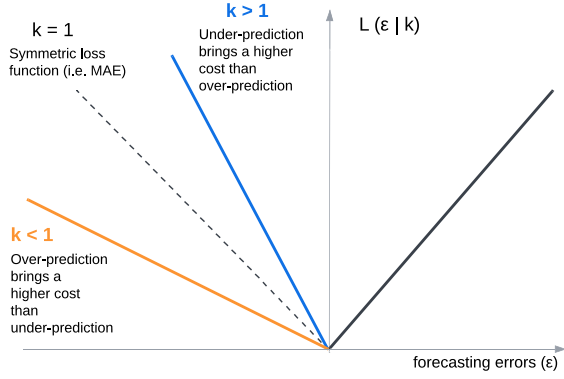


Fig. 2. Piecewise asymmetric cost-oriented loss function for household load forecasting.

of these errors: under-predictions lead to significantly higher procurement costs, whereas over-predictions reduce the resale value of excess electricity, as also discussed by [26].

To address this, [26] proposes a piecewise loss function tailored to minimize the real economic costs associated with load forecasting errors. It is formulated as follows:

$$L(|\epsilon|k_1, k_2) = \begin{cases} k_1\epsilon, & \text{if } \epsilon > 0, \\ k_2\epsilon, & \text{if } \epsilon \leq 0 \end{cases} \quad (1)$$

where ϵ represents the prediction error $\epsilon = \hat{y} - y$, with $\epsilon > 0$ indicating an over-prediction and $\epsilon \leq 0$ under-prediction. The coefficients k_1 and k_2 are positive and penalize over-predictions and under-predictions, respectively. The cost function $L(|\epsilon|k_1, k_2)$ is designed to meet these specific conditions [23]:

- 1) $L(0) = 0$. There is no cost if the forecasting error ϵ is zero.
- 2) $\min L(\epsilon) = 0$. The cost must be greater than or equal to zero.
- 3) $L(\epsilon)$ is monotonically non-decreasing as ϵ moves away from zero. In other words, it is a convex function, satisfying $f''(x) \geq 0$ for all x .

The relationship between k_2 and k_1 is critical, as illustrated in Fig. 2. The normalization constant $k = \frac{k_2}{k_1}$ indicates the relative cost severity of under-predictions compared to over-predictions. A k value greater than 1 suggests that under-predictions are costlier, while with a k value less than 1, over-predictions become more expensive. Thus, the piecewise cost-oriented loss function is defined as:

$$L(|\epsilon|k) = \begin{cases} \epsilon, & \text{if } \epsilon > 0, \\ k\epsilon, & \text{if } \epsilon \leq 0 \end{cases} \quad (2)$$

The cost associated with the prediction errors is defined in (3). Over-prediction leads to excess electricity bought at wholesale market prices k_1 . Contrarily, under-prediction requires purchasing additional power at the higher downward imbalance price k_2 :

$$C(\epsilon) = k_1|\epsilon| \text{ (if } \epsilon \geq 0) + k_2|\epsilon| \text{ (if } \epsilon < 0) \quad (3)$$

This cost-oriented approach emphasizes the importance of having a prediction mechanism in electricity load forecasting to effectively optimize financial outcomes.

III. LOAD FORECASTING METHODS WITH A COST-ORIENTED APPROACH

This section provides an overview of the local learning and federated learning configurations, including novel personalization strategies focused on home energy management system's load forecasting tasks. Fig. 3 visualizes these methodologies, while Fig. 4 contrasts the proposed cost-oriented loss function with traditional symmetric cost functions across these configurations. Detailed discussions on each method are presented in subsequent subsections.

A. Local Learning (LL)

Local Learning (LL) employs a state-of-the-art model, the Long Short-Term Memory (LSTM), due to its effectiveness in capturing temporal dependencies in time-series data [28]. In this approach, each household develops its own prediction model using its historical consumption data. The main advantage of this method is its ability to customize the model to fit individual usage patterns. However, a major drawback is the need for extensive historical data to ensure the model's accuracy, reliability and robustness. To ensure a fair comparison among all the methods, the LSTM model will be consistently used across all methodologies in this study.

B. Federated Learning Methods

FL includes several configurations based on the structure of the data involved, including horizontal FL, vertical FL, and transfer FL. In this study, household load forecasting uses the horizontal configuration, which involves datasets that share identical features but targets are different and distributed across households.

In the process of aggregating the model weights from clients, various strategies can be adopted. This study employs the Federated Averaging (FedAvg) method, one of the most common and widely used. Findings from [13] have demonstrated the superior performance of FedAvg in load forecasting. The FedAvg aggregation function is mathematically represented as:

$$\theta_{r+1}^S \leftarrow \sum_{i \in S_r} \frac{n_i}{n_{S_r}} \theta_r^i \quad (4)$$

where r represents the communication round, n is the number of randomly selected households per FL round, S_r is the subset of households selected in round r , θ_{r+1}^S are the global model parameters at round r , θ_r^i are the local model parameters of household i after local training in round r and n_i is the number of data points held by client i .

1) *Global Federated Learning*: This approach employs a centralized server to coordinate federated learning across multiple households, utilizing distributed data as depicted in Algorithm 1. The process begins with the initialization of a global model using randomly assigned weights. Subsequently, the model is trained and refined iteratively over a series of communication rounds R . During each round, a randomly selected subset of households, S_r , receives the latest weights of the global model. Each participating household i performs a local training, where the model is updated using its local data

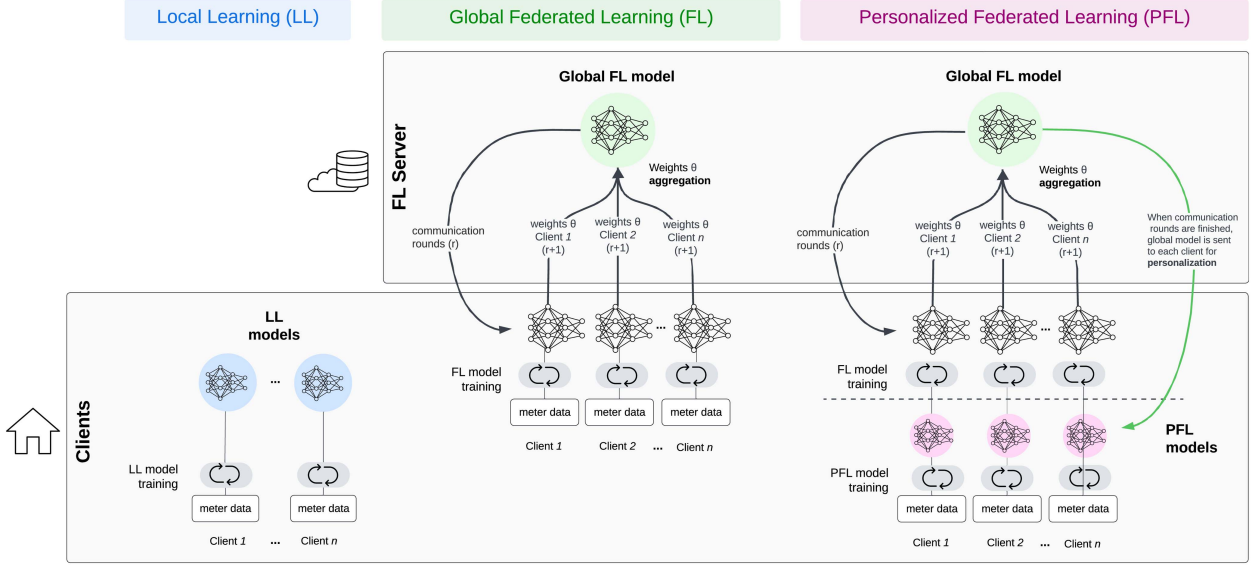
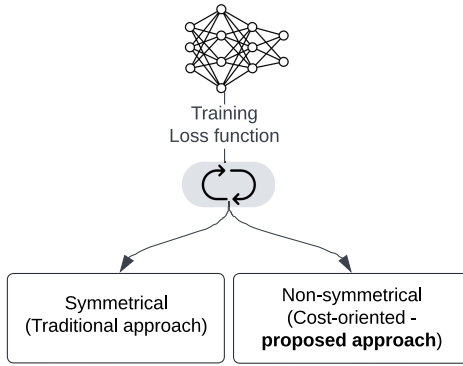
Fig. 3. Personalized FL methodology for load forecasting. Cluster i example.

Fig. 4. Loss functions used in load forecasting.

batches to minimize the loss function, denoted as $\mathcal{L}(\theta; X, y)$. After local training, the households send their updated weights θ_{r+1}^i back to the central server. The server aggregates these weights using FedAvg, taking into account the volume of data each household contributed to training. This aggregation step ensures that the global model incorporates the knowledge gained from each household's local data. This iterative process of local training and aggregation continues for the specified number of communication rounds R . The final output of the algorithm is a trained global model, θ^S , which effectively integrates the diverse data from all participating households, thereby improving its generalization capabilities across the entire dataset.

Despite its strengths, this method faces challenges due to the non-independent and identically distributed nature of household consumption profiles. These differences can introduce biases and affect the global FL model's generalization performance across different household datasets. To mitigate these effects and improve robustness, a personalization method is proposed in the following subsection.

2) *Personalized Federated Learning (PFL)*: This method combines elements from previous methodologies by utilizing

Algorithm 1: Global Federated Learning Method.

- 1: **Input:** N : total households, θ^S : global FL model, $\{\theta^i\}_{n=1}^N$: local models, R : communication rounds between server and households, E : local epochs, η : learning rate, b batch size.
 - 2: **Output:** Trained global FL model θ^S
 - 3:
 - 4: **Algorithm:**
 - 5: Initialize global model weights randomly $\theta_{r=0}^S$
 - 6: **for** each communication round $r = 0$ to $R - 1$ **do**
 - 7: $\mathcal{S}_r \leftarrow$ random subset of i households $\in N$
 - 8: **for** each household $i \in \mathcal{S}_r$ **in parallel do**
 - 9: $\theta_r^i \leftarrow$ Receive server weights θ_r^S
 - 10: $\theta_{r+1}^i \leftarrow \text{LocalTraining}(\theta_r^i, E, \eta, b)$
 - 11: Upload updated local model θ_r^i to the server
 - 12: **Server** computes and updates the global model:
 - 13: $\theta_{r+1}^S \leftarrow \text{FedAvgAggregate}(\{\theta_{r+1}^i\}_{i \in \mathcal{S}_r})$
 - 14: **return** θ_R^S
 - 15:
 - 16: **function** LocalTraining θ_r^i, E, η, b
 - 17: Update local model
 - 18: **for** each local epoch $e = 1$ to E **do**
 - 19: **for** each batch b of local data **do**
 - 20: $\theta_{r+1}^i \leftarrow \theta_r^i - \eta \nabla \mathcal{L}(\theta_r^i, b)$
 - 21: **return** θ_{r+1}^i
 - 22: **function** FedAvgAggregate $\{\theta_{r+1}^i\}_{i \in \mathcal{S}_r}$
 - 23: Weighted average of local models
 - 24: $\theta_{r+1}^S \leftarrow \sum_{i \in \mathcal{S}_r} \frac{n_i}{n_{\mathcal{S}_r}} \theta_{r+1}^i$
 - 25: **return** θ_{r+1}^S
-

user-specific data to locally retrain the global FL model, creating a customized model for each client. This retraining, limited to a few epochs, utilizes exclusively the user's data, ensuring that

the personalized model remains within the user's domain to maintain data privacy, as detailed in [16].

This work proposes two PFL approaches. The first, referred to as PFL-retrain (PFL-rt), involves local retraining of the global model at each household over a few epochs to maintain the core knowledge of the global model. The second approach, PFL-l2, incorporates the L^2 regularization to both the training of the global model and the local retraining phase. Previous studies, such as [19], have applied L^2 regularization but only to the global FL model. Algorithm 2 explains the methodology for PFL-l2, whereas PFL-rt follows a similar process but omits the regularization term for the model loss function.

The L^2 regularization is incorporated into the loss function to prevent overfitting and enhance the generalization capabilities of the models. This technique ensures that the models do not become too customized to the specific training data, enabling them to identify broader patterns applicable to new, unseen data. The loss function of the LSTM model with L^2 parameter norm penalty is a combination of the selected loss function $\mathcal{L}$ and the regularization term $\Omega(\theta) = \frac{1}{2}\|\theta\|_2^2$, also known as weight decay [29]. We denote the regularized objective loss function by $\mathcal{L}_\lambda$:

$$\mathcal{L}_\lambda(\theta; X, y) = \mathcal{L}(\theta; X, y) + \lambda\Omega(\theta) \quad (5)$$

where $\lambda \in [0, \infty)$ is a hyperparameter that controls the strength of our preference for smaller weights. When $\lambda = 0$, we impose no preference, whereas larger values force the weights to become smaller. $\Omega(\theta)$ helps in controlling overfitting by penalizing large weights.

Both PFL-rt and PFL-l2 are evaluated using two types of cost functions: the traditional symmetric cost function (MAE) and the asymmetric cost function referenced in Section II. These evaluations aim to determine the efficacy and cost of each personalization strategy under different loss functions.

The implementation process for PFL with L^2 regularization for load forecasting is outlined in Algorithm 2. Starting similarly to the global federated learning method described in Algorithm 1, this method introduces the L^2 regularization term into the loss function during local training of the global FL model. After completing the communication rounds, the trained global model is sent to all households for personalization. During the personalization phase, each household further trains the received global model on its own data to tailor it to its specific needs. This additional training, which also includes L^2 regularization, allows each household to fine-tune the model, resulting in a set of personalized models. The algorithm concludes with the return of the personalized models, which now better capture the individual characteristics of the data from each household. This combination of global collaboration and local customization enhances the overall performance and applicability of the federated learning model.

IV. CASE STUDY AND RESULTS

This section aims to validate the effectiveness and economic benefits of the proposed cost-oriented personalized FL methodology. For this purpose, this study uses actual household data,

Algorithm 2: Personalized Federated Learning With Retraining and L^2 Regularization.

```

1: Input:  $N$ : total households,  $\mathbf{D} = [D_1, D_2, \dots, D_N]$ :
   historical data of  $N$  households,  $\theta^S$ : global FL model,
    $\{\theta\}_{n=1}^N$ : local models,  $\{\theta^p\}_{n=1}^N$ : personalized local
   models,  $\mathcal{L}$ : chosen loss function {cost-oriented, mae},
    $R$ : communication rounds between server and
   households,  $E$ : local epochs,  $P$ : personalization
   epochs,  $\eta$ : learning rate,  $b$ : batch size,  $\lambda$ : regularization
   parameter
2: Output: Set of personalized FL models with
   regularization  $\{\theta_1^P, \theta_2^P, \dots, \theta_N^P\}$ , trained global FL
   model  $\theta^S$ 
3:
4: Algorithm:
5: Initialize global model weights randomly  $\theta_{r=0}^S$ 
6: for each communication round  $r = 0$  to  $R - 1$  do
7:    $S_r \leftarrow$  random subset of  $i$  households  $\in N$ 
8:   for each household  $i \in S_r$  in parallel do
9:      $\theta_r^i \leftarrow$  Receive server weights  $\theta_r^S$ 
10:     $\theta_{r+1}^i \leftarrow \text{LOCALTRAINING}(\theta_r^i, E, \mathcal{L}, \eta, b, \lambda)$ 
11:    Upload updated local model  $\theta_{r+1}^i$  to the server
12:   Server computes and updates the global model:
13:    $\theta_{r+1}^S \leftarrow \text{FEDAVGAGGREGATE}(\{\theta_{r+1}^i\}_{n \in S_r})$ 
14:   return  $\theta^S$ 
15:
16: Send trained global FL model weights  $\theta^S$  to all
   households for personalization
17: for each household  $i \in N$  in parallel do
18:    $\theta_{p=0}^{i,P} \leftarrow$  Receive global server weights  $\theta^S$ 
19:   Household  $n$  personalization
20:    $\theta_p^{i,P} \leftarrow \text{LOCALTRAINING}(\theta_{p=0}^{i,P}, P, \mathcal{L}, \eta, b, \lambda)$ 
21:   return  $\theta_p^{i,P}$ 
22:
23: function LocalTraining $\theta_r^i, E, \mathcal{L}, \eta, b, \lambda$ 
24:   Update local model
25:   for each local epoch  $e = 1$  to  $E$  do
26:     for each batch  $b$  of local data do
27:        $\theta_{r+1}^i \leftarrow \theta_r^i - \eta(\nabla \mathcal{L}(\theta_r^i, b) + \lambda \theta_r^i)$ 
28:   return  $\theta_{r+1}^i$ 
29: function FedAvgAggregate $\{\theta_{r+1}^i\}_{n \in S_r}$ 
30:   Weighted average of local models
31:    $\theta_{r+1}^S \leftarrow \sum_{i \in S_r} \frac{n_i}{n_{S_r}} \theta_{r+1}^i$ 
32:   return  $\theta_{r+1}^S$ 

```

employing different cost functions in FL models to compare the results with benchmark machine learning models trained locally.

A. Dataset Description and Experimental Settings

This study uses the publicly available Low Carbon London dataset [30], consisting of electricity consumption data collected from 5,567 households in London between November 2011 and February 2014, with a 30-minute sampling interval. The features used for the load forecast include past consumption data from 1,

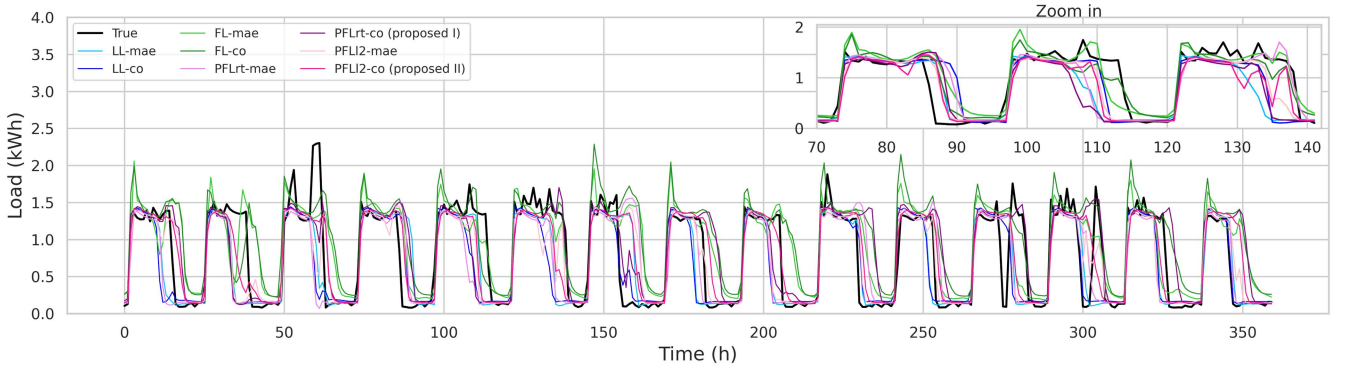


Fig. 5. LL, global FL and PFL prediction results for household #2 over two weeks.

TABLE I
CASE STUDY CONFIGURATION DETAILS

Parameter	Configuration
Dataset	
Source	Low Carbon London dataset [30]
Number of households	$N = 10$
Training period	Jul 1, 2013 to Jan15, 2014
Test period	Jan 15, 2014 to Feb 15, 2014
Width of sliding window	{24, 72, 168 }
Forecasted period	24h ahead
LSTM architecture	
LSTM layers	3
Multi-FC layers	1
Hidden nodes per layer	64
Training	
Input data	Past load observations, work day, temp, rain, hour
Communication rounds R	{50, 100 , 200}
Local epochs E before averaging	{5, 10, 20}
Local epochs for local learning	{50, 100}
Proportion of participants	{0.5}
Batch size b	{32, 64, 128, 256 }
Dropout	0.2
L^2 factor λ	0.00001
Initial learning rate η	0.001
Optimizer	Adam
Normalization	Minmax scaler
Computational information	
FL library	Flower
CUDA version	11.3
GPU	NVIDIA GeForce GTX 1050 Ti

2, and 7 days before, working days and hours, and also daily average weather data such as temperature and precipitation. Table I outlines the setup of the experimental framework used in the study. It includes the dataset specifications, evaluation periods for the case study, the architecture of the recurrent neural network, detailed training specifications, and computational parameters. During the simulation of results, various configurations and parameters were tested. The values that yielded the best results are highlighted in bold. This comprehensive Table ensures the reproducibility and clarity of the research conditions. According to the U.K. energy regulator *Ofgem* and the Balancing and Settlement Code Company, *Elaxon*, during the first half of early June 2024, the wholesale electricity price was (£75/MWh) and the downward imbalance price (£132/MWh). These average values are utilized in this study to define cost

function parameters to build the cost-oriented loss function. Therefore, the k factor for this case study is approximately 1.8, suggesting that the cost for under-predict is 1.8 times higher than the magnitude of over-prediction. It is important to clarify that, although the historical household consumption data is from 2014, electricity market prices from 2024 have been used due to the lack of available data for 2014. This decision does not impact the results; on the contrary, it ensures that the outcomes reflect the current state of prices.

B. Results

Figs. 5 and 6 illustrate the load prediction performance across two weeks for eight different methods, as introduced in Fig. 3, for two households. The x-axis records time in hours, while the y-axis compares the predictions of each method (each represented by a unique color for clarity) against the true consumption, shown in black. Additionally, Tables II and III detail the error metric MAE and the total associated costs for each method, respectively.

The global FL method shows the worst performance in terms of error and costs, ranking last in both categories as detailed in Tables II–III. This sub-optimal performance is primarily due to the global FL model’s inability to generalize effectively across the different households’ energy consumption patterns. To compensate this limitation, two PFL approaches (PFL-rt and PFL-I2) are developed and evaluated. These models improve prediction errors by tailoring the learning process to the characteristics of each household while keeping the insights gained from the global FL model training.

It is noteworthy that the methods utilizing a specialized cost function generally offer lower total costs compared to those optimizing for MAE alone. Particularly, PFL-I2 emerges as the most effective model, optimizing both the error rates and associated costs, closely followed by PFL-rt. Methods such as LL-mae and LL-co also show competitive performance in terms of error and cost, respectively.

As expected, the methods using MAE as the loss function achieve the lowest errors (see Table II), with PFL-I2-mae marking the best performance at an average error of 0.2566kWh. PFL-rt is a close second, reflecting the effectiveness of personalized training regimes. Table III shows that PFL-I2-co has the lowest

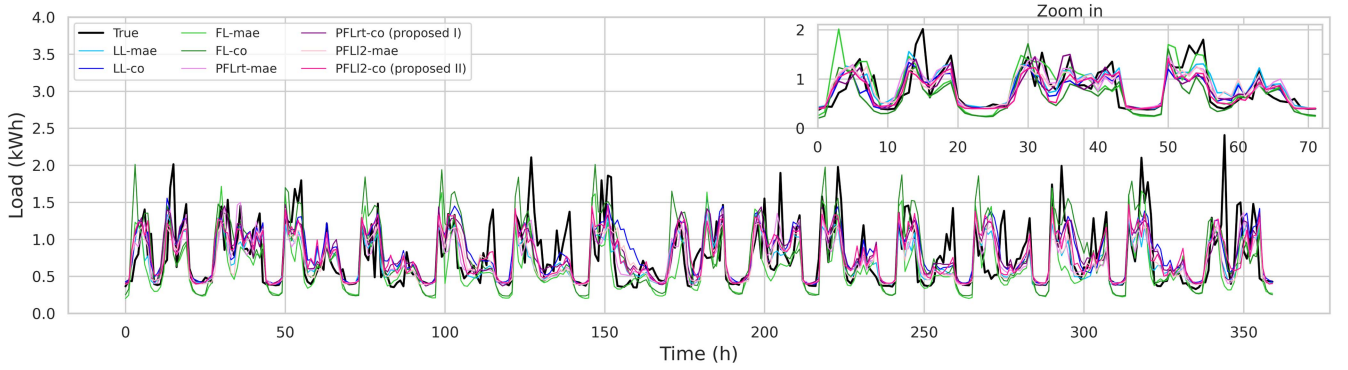


Fig. 6. LL, global FL and PFL prediction results for household #3 over two weeks.

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT LEARNING METHODS FOR HOUSEHOLD LOAD PREDICTION

House ID	Local Learning (LL)		Federated Learning (FL)		Personalized FL retrain		Personalized FL regularization	
	LL-mae	LL-co	FL-mae	FL-co	PFLrt-mae	PFLrt-co	PFLl2-mae	PFLl2-co
#1	0.6917	0.6725	0.6954	0.7051	0.6688	0.6425	0.6574	0.6515
#2	0.2666	0.2648	0.3178	0.3641	0.2672	0.2900	0.2566	0.2575
#3	0.2142	0.2340	0.2684	0.2932	0.2050	0.2245	0.2114	0.2202
#4	0.2443	0.2318	0.2602	0.2824	0.2170	0.2396	0.2201	0.2541
#5	0.2949	0.3783	0.2913	0.3183	0.2892	0.3699	0.2914	0.3149
#6	0.1689	0.1969	0.1841	0.2391	0.1693	0.2267	0.1675	0.1944
#7	0.1408	0.1289	0.2076	0.1740	0.1270	0.1290	0.1315	0.1235
#8	0.2409	0.2326	0.2779	0.2760	0.2467	0.2430	0.2421	0.2474
#9	0.2257	0.2444	0.2994	0.3379	0.2340	0.2448	0.2298	0.2481
#10	0.1570	0.1618	0.2642	0.3042	0.1591	0.1853	0.1578	0.1679
Average (MAE)	0.2645	0.2746	0.3066	0.3294	0.2583	0.2795	0.2566	0.2680
Rank	3	5	7	8	2	6	1	4

MAE error [kwh].

TABLE III
COST COMPARISON OF DIFFERENT LEARNING METHODS FOR HOUSEHOLD LOAD PREDICTION

House ID	Local Learning (LL)		Federated Learning (FL)		Personalized FL Retrain		Personalized FL Regularization	
	LL-mae	LL-co	FL-mae	FL-co	PFL-rt-mae	PFL-rt-co	PFL-l2-mae	PFL-l2-co
#1	64.35	56.47	63.67	61.61	59.36	52.98	59.60	54.84
#2	24.11	23.89	23.15	24.39	24.30	23.02	22.36	21.15
#3	17.89	16.75	23.94	22.10	17.19	16.75	17.36	16.54
#4	19.62	17.36	22.71	21.99	17.93	17.46	18.45	18.39
#5	22.51	25.23	23.51	24.06	21.41	24.64	21.46	22.19
#6	12.38	13.06	13.78	15.83	12.76	14.55	12.63	12.75
#7	13.12	11.60	16.63	13.00	10.68	11.26	11.73	10.69
#8	19.12	16.92	22.98	20.21	20.71	17.82	18.81	17.69
#9	18.91	18.16	24.26	24.64	19.72	18.62	19.10	18.63
#10	13.55	12.95	20.22	21.71	13.93	14.33	13.75	13.44
Sum (£)	225.55	212.39	254.85	249.54	218.00	211.44	215.24	206.31
Rank	6	3	8	7	5	2	4	1

Total cost [£].

total cost of £206.31. The next best method in terms of cost is PFL-rt, with a total cost of £211.44. Both these personalized approaches significantly reduce the total cost compared to local and global FL methods, which have higher total costs of £225.55 and £254.85, respectively.

There is a noticeable trade-off between costs and prediction accuracy. Methods achieving lower costs do not necessarily have the lowest MAE and vice versa. This highlights the true value

and effectiveness of the cost-oriented loss function when applied to predictions. Overall, the results highlight that personalized federated learning with regularization not only improves the accuracy of benchmark load predictions but also enhances cost efficiency, making it the most effective method among those evaluated.

The box plots presented in Figs. 7 and 8 offer a comparative visualization of the average prediction errors and penalization

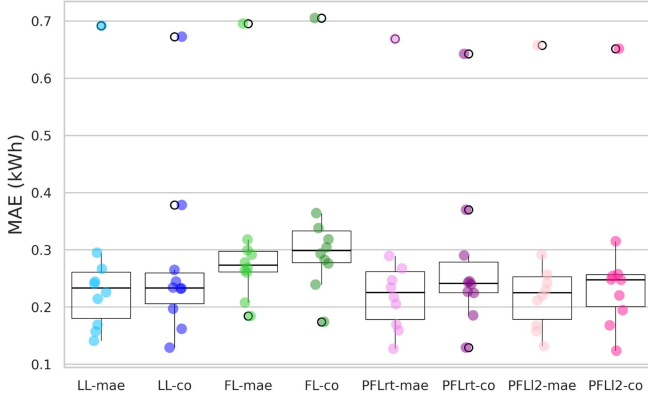


Fig. 7. Distribution of averaged load predictions errors for LL, FL and PFL methods.

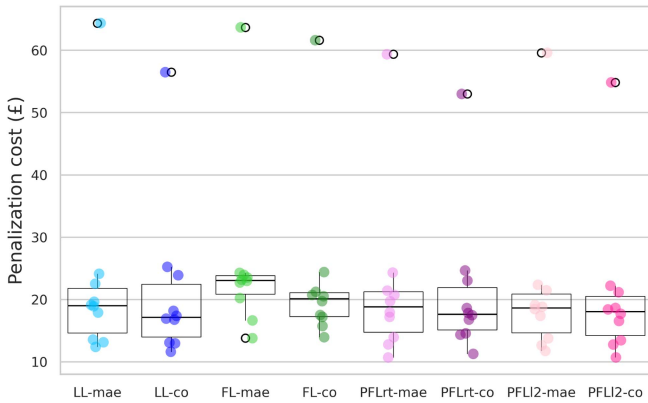


Fig. 8. Distribution of total penalization cost for LL, FL and PFL methods.

costs across the evaluated methods, respectively. Each data point within the box plots represents the average value for an individual household. The wider the box plot, the greater the difference in performance between households and the poorer the generalization of the model. As already mentioned above, Fig. 7 demonstrates that personalization approaches using MAE as a loss function offer lower median MAE values. Regarding the costs in Fig. 8, the trends are similar. Personalization methods using a cost-oriented loss function achieve the lowest costs, along with LL-co. Outliers were prevalent in all methods, due to household #1 results.

C. Sensitivity Analysis

1) Historical Data Availability and Its Impact on Performance and Costs: The results discussed in the previous section were obtained using six and a half months of training data. How would the three learning frameworks —LL, global FL, and PFL— as well as different loss functions, perform in terms of cost and error then trained with different amounts of historical data? To answer this question, a sensitivity analysis is conducted to explore the change in MAE and its associated cost penalization across different historical data availabilities. Each scenario is evaluated using the same one-month test dataset

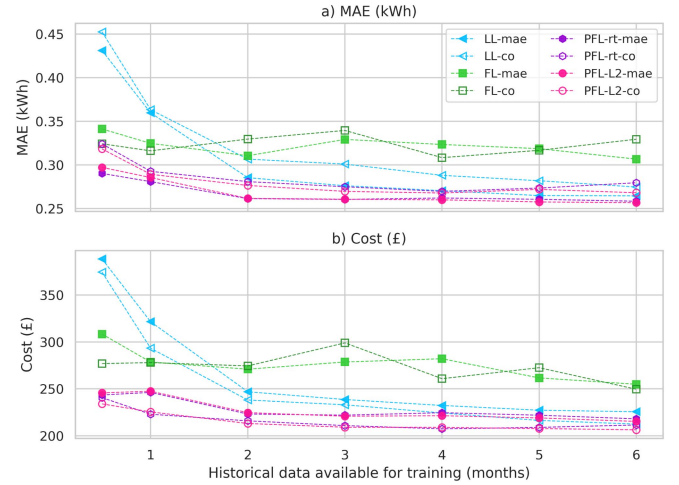


Fig. 9. Sensitivity analysis for different historical data training availability in terms of (a) MAE and (b) penalization cost.

(from January 15 to February 15, 2014), ensuring consistent comparisons of results.

Fig. 9 displays the performance of LL, global FL and PFL utilizing traditional MAE and cost-oriented loss functions with historical data ranging from 0.5 to 6 months. For Local Learning methods (LL-co and LL-mae), there is a noticeable improvement in both cost and error performance as the volume of available data increases, corroborating the hypothesis that more data typically enhances model training. Notably, when two months of historical data are available, LL models begin to achieve cost and error metrics comparable to global FL and PFL methods.

PFL methods demonstrate the most robust and stable results. With only 15 days of available data, PFL can achieve an error rate nearly half that of LL models, demonstrating the advantage of shared knowledge among users. This suggests that in scenarios with few historical data samples, PFL models outperform those trained only on isolated historical data (LL). On the other hand, the global FL model shows consistent performance but fails to effectively generalize the model and adapt to each household due to non-iid household data behavior.

Fig. 10 demonstrates the general relationship between cost and performance across methods. The x-axis represents the cost (£), while the y-axis represents the error (MAE). The lighter the color of the data points, the less data is available for training. Filled data points represent methods trained using the traditional MAE loss function, while outlined ones were trained using the cost-oriented loss function proposed in this study. PFL methods, particularly those integrating regularization strategies like PFL-I2, achieve an optimal balance between cost and error. This graph demonstrates that methods trained with the proposed cost function generally attain lower costs, even though they may exhibit slightly higher errors compared to those trained with MAE.

In summary, this sensitivity analysis demonstrates that reliable load predictions and low penalization costs can be achieved for households, despite having limited available data.

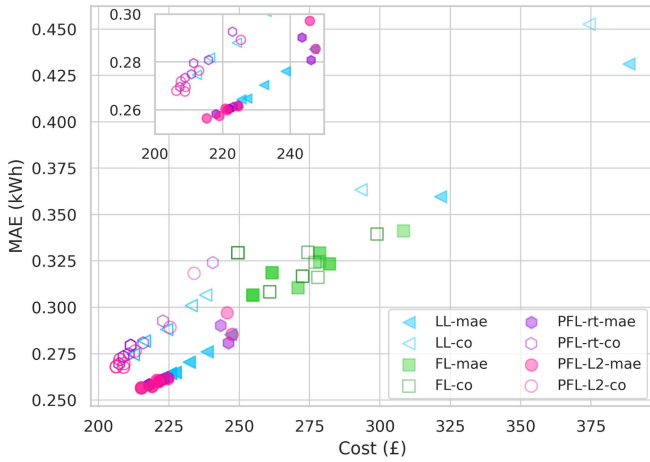


Fig. 10. Sensitivity analysis with varying historical data availability. Relationship between penalization cost and error metrics.

V. CONCLUSIONS AND CHALLENGES

This study has effectively integrated a cost-oriented loss function within FL methods for household energy management system load forecasting, addressing the technical challenges of data privacy, ensuring low prediction errors and considering the economic penalties due to forecasting errors. Load forecasting methods trained with the asymmetric cost-oriented loss function consistently incur lower costs compared to those using traditional symmetric loss functions. Incorporating personalized strategies and L^2 regularization in FL improves prediction accuracy and cost efficiency for household load forecasting. The PFL approach, particularly the PFL L^2 regularization method, has demonstrated superior performance, followed by LL and global FL methods. PFL effectively personalizes and adapts the global FL model to individual household consumption patterns, minimizing the risk of cost penalties and contributing to the stability of the electricity grid. Moreover, a sensitivity analysis highlighted the robustness of FL frameworks in delivering reliable load predictions across households with varying historical data availability. Unlike LL models, FL frameworks maintain consistent performance regardless of the amount of available data. The PFL methods, especially those incorporating regularization (PFL-L2), offer an optimal balance between minimizing prediction errors and managing costs, making them highly effective for energy management applications.

Despite the promising results of the proposed FL frameworks, its implementation in energy systems faces several technical challenges related to data heterogeneity, privacy and communication constraints. In response to these challenges, future research could propose solutions. One promising direction is refining the cost function approach by incorporating predictions for day-ahead over-predictions and under-predictions. This enhancement could significantly improve the economic efficiency of the models, making them more responsive and aligned with real-world energy pricing dynamics. Although this article does not explore clustering, existing literature has demonstrated that this technique can notably improve the performance and computational efficiency of non-iid data models while also preserving

data privacy. Thus, incorporating clustering techniques into our framework could be a future line of research. Additionally, personal information leakage can occur when sharing the model weights. To mitigate this risk, the adoption into our model of some advanced techniques such as differential privacy, secure multi-party computation, and encryption should be considered.

REFERENCES

- [1] Hitachi ABB Power Grids, "Power Grids drives digital transformation of the power sector with new Smart Digital Substation," Press release, Feb. 2021.
- [2] S. Barja-Martinez, M. Aragüés-Peñalba, Í. Munné-Collado, P. Lloret-Gallego, E. Bullich-Massagué, and R. Villafafila-Robles, "Artificial intelligence techniques for enabling Big Data services in distribution networks: A review," *Renewable Sustain. Energy Rev.*, vol. 150, 2021, Art. no. 111459.
- [3] European Parliament and Council of the European Union, "Regulation (EU) 2016/679 General Data Protection Regulation (GDPR)," *Official J. Eur. Union*, vol. L119, pp. 1–88, Apr. 2016.
- [4] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Federated learning of deep networks using model averaging," 2016, *arXiv:1602.05629*.
- [5] U. Zafar, S. Bayhan, and A. Sanfilippo, "Home energy management system concepts, configurations, and technologies for the smart grid," *IEEE Access*, vol. 8, pp. 119271–119286, 2020.
- [6] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, 2019.
- [7] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, "Federated Learning," *Synth. Lectures Artif. Intell. Mach. Learn.*, vol. 13, no. 3, pp. 1–207, 2020.
- [8] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik, and Y. Jararweh, "Federated learning review: Fundamentals, enabling technologies, and future applications," *Inf. Process. Manag.*, vol. 59, no. 6, 2022, Art. no. 103061.
- [9] S. Lee and Dae-Hyun Choi, "Federated reinforcement learning for energy management of multiple smart homes with distributed energy resources," *IEEE Trans. Ind. Inform.*, vol. 18, no. 1, pp. 488–497, Jan. 2022.
- [10] A. Grataloup, S. Jonas, and A. Meyer, "A review of federated learning in renewable energy applications: Potential, challenges, and future directions," *Energy AI*, vol. 17, 2024, Art. no. 100375.
- [11] P. Kairouz et al., "Advances and open problems in federated learning," *Found. Trends Mach. Learn.*, vol. 14, no. 12, pp. 1–210, Jun. 2021. [Online]. Available: <http://dx.doi.org/10.1561/22000000083>
- [12] R. Firouzi, R. Rahmani, and T. Kanter, "Federated learning for distributed reasoning on edge computing," *Procedia Comput. Sci.*, vol. 184, pp. 419–427, 2020.
- [13] M. N. Fekri, K. Grolinger, and S. Mir, "Distributed load forecasting using smart meter data: Federated learning with recurrent neural networks," *Int. J. Elect. Power Energy Syst.*, vol. 137, 2022, Art. no. 107669.
- [14] J. Fernández, S. P. Menci, C. M. Lee, A. Rieger, and G. Fridgen, "Privacy-preserving federated learning for residential short-term load forecasting," *Appl. Energy*, vol. 326, 2022, Art. no. 119915.
- [15] Y. Shi and X. Xu, "Deep federated adaptation: An adaptive residential load forecasting approach with federated learning," *Sensors*, vol. 22, no. 9, 2022, Art. no. 3264.
- [16] A. Taik and S. Cherkaoui, "Electrical load forecasting using edge computing and federated learning," in *Proc. IEEE Int. Conf. Commun.*, 2020, pp. 1–6.
- [17] A. Z. Tan, H. Yu, L. Cui, and Q. Yang, "Towards personalized federated learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 12, pp. 9587–9603, Dec. 2023.
- [18] F. Sabah, Y. Chen, Z. Yang, M. Azam, N. Ahmad, and R. Sarwar, "Model optimization techniques in personalized federated learning: A survey," *Expert Syst. Appl.*, vol. 243, 2024, Art. no. 122874.
- [19] Y. Wang, "Personalized federated learning for individual consumer load forecasting," *CSEE J. Power Energy Syst.*, vol. 9, no. 1, pp. 326–330, Jan. 2022.
- [20] X. Qu et al., "Personalized federated learning for heterogeneous residential load forecasting," *Big Data Mining Analytics*, vol. 6, no. 4, pp. 421–432, 2023.
- [21] W. Zhang et al., "Semi-asynchronous personalized federated learning for short-term photovoltaic power forecasting," *Digit. Commun. Netw.*, vol. 9, no. 5, pp. 1221–1229, 2023.

- [22] C. Briggs, Z. Fan, and P. Andras, "Federated learning for short-term residential load forecasting," *IEEE Open Access J. Power Energy*, vol. 9, pp. 573–583, 2022.
- [23] W. J. C. Granger, "Outline of forecast theory using generalized cost functions," *Spanish Econ. Rev.*, vol. 1, pp. 161–173, 1999.
- [24] J. Zhang, Y. Wang, and G. Hug, "Cost-oriented load forecasting," *Elect. Power Syst. Res.*, vol. 205, 2022, Art. no. 107723.
- [25] G. Li and H. D. Chiang, "Toward cost-oriented forecasting of wind power generation," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2508–2517, Jul. 2018.
- [26] J. Wu, Y.-G. Wang, Y.-C. Tian, K. Burrage, and T. Cao, "Support vector regression with asymmetric loss for optimal electric load forecasting," *Energy*, vol. 223, 2021, Art. no. 119969.
- [27] A. Zellner, "Bayesian estimation and prediction using asymmetric loss functions," *J. Amer. Stat. Assoc.*, vol. 81, pp. 446–451, 1986.
- [28] A. Taik and S. Cherkaoui, "Electrical load forecasting using edge computing and federated learning," in *Proc. IEEE Int. Conf. Commun.*, 2020, pp. 20–25.
- [29] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: <http://www.deeplearningbook.org>
- [30] "Smart meters in London." Kaggle, 2017. Accessed: Jan. 10, 2023. [Online]. Available: <https://www.kaggle.com/datasets/jeanmidev/smart-meters-in-london>



Sara Barja-Martínez received the M.Sc. degree in industrial engineering (major in energy) and the Ph.D. degree in electrical engineering from Universitat Politècnica de Catalunya, Barcelona, Spain, in 2018 and 2023, respectively. During her Ph.D., she completed a doctoral stay with Imperial College London, London, U.K. Since 2017, she has been an integral part of the CITCEA-UPC research group, contributing and leading several innovative projects. Her research interests include energy management system algorithms, machine learning techniques applied to the energy sector, electricity and flexibility markets, optimization, and demand response programs.



Fei Teng (Senior Member, IEEE) received the B.Eng. degree in electrical engineering from Beihang University, Beihang, China, in 2009, and the M.Sc. and Ph.D. degrees in electrical engineering from Imperial College London, London, U.K., in 2010 and 2015, respectively. He is currently a Senior Lecturer with the Department of Electrical and Electronic Engineering, Imperial College London. His research interests include the power system operation with high penetration of inverter based resources (IBRS) and the cyber-resilient and privacy-preserving cyber-physical power grid.



Adrià Junyent-Ferré (Senior Member, IEEE) received the Industrial Engineering degree from the School of Industrial Engineers of Barcelona, Polytechnic University of Catalonia (UPC), Barcelona, Spain, in 2007, and the Ph.D. degree in electrical engineering from UPC, in 2011. He was a Researcher with CITCEA-UPC from 2006 to 2012, and a Lecturer with the Barcelona College of Industrial Engineering, Barcelona, in 2012. In 2013, he joined the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K., where he is currently a Senior Lecturer. He investigates how converters can be resilient to network transients with minimal oversizing. His research interests include power electronic converter applications for wind power generation, VSC-HVDC transmission, and new forms of low-voltage distribution.



Mònica Aragüés-Peñalba (Member, IEEE) received the M.Sc. degree in industrial engineering (major in electricity) and the Ph.D. degree in electrical engineering from the School of Industrial Engineering of Barcelona, UPC, Barcelona, Spain, in 2011 and 2016, respectively. Since 2010, she has belonged to CITCEA-UPC, Electrical Engineering Department, UPC. Since April 2018, she has been a Lecturer with the Electrical Engineering Department, UPC (Serra Hunter Fellow). She is currently the Project Coordinator of BD4OPEM H2020. Her research interests include renewable integration in power systems, transmission and distribution power systems, wind and solar power plant operation and control, microgrids operation and control, optimization, and data science applications to power systems.