A Novel Fuzzy Inference System-Based Federated Learning Approach

Köksal Erentürk

College of Engineering, Department of Computer Engineering, Atatürk University, Erzurum, Türkiye
0000-0001-7536-1351 keren@atauni.edu.tr

Abstract

Federated learning enables distributed training of machine learning models across a network of decentralised computational nodes, ensuring data privacy by retaining sensitive data on local devices. This paradigm supports the construction of high-performance models by leveraging non-i.i.d. and heterogeneous data dispersed across multiple clients, thereby removing the dependency on centralized data aggregation. Federated learning frameworks utilize a range of model aggregation algorithms—such as FedAvg (federated averaging), FedProx (proximal gradient methods), and FedOpt (adaptive federated optimization)—to achieve convergence by integrating locally updated model parameters into a global model, all while upholding data locality constraints. Fuzzy logic, characterized by its capacity to model vagueness and imprecision through fuzzy sets and linguistic variables, provides a robust mechanism for approximate reasoning in uncertain environments. This study proposes FedFIS, a novel fuzzy logic-based aggregation scheme embedded within the federated learning architecture. The FedFIS methodology circumvents the reliance on gradient-based optimization and computationally intensive mathematical formulations by leveraging fuzzy inference mechanisms, thus offering a computationally lightweight and privacy-preserving alternative for federated parameter aggregation.

***Keywords*:** federated learning, fuzzy logic, aggregation, machine learning

Received:	Revised:	Accepted:	Published:
27/05/2025	02/06/2025	06/06/2025	14/06/2025

1. Introduction

While distributed training is applicable to traditional machine learning algorithms such as k-nearest neighbors, support vector machines, Bayesian networks, and decision trees, it frequently requires the transmission of raw data, which introduces significant privacy concerns [1]. To address these challenges, Federated Learning (FL) has emerged as a decentralized learning framework [2–4]. It allows multiple clients to collaboratively train a shared model without transferring their local data to a central repository [5]. Each client independently trains a local model on its private dataset and shares only model parameter updates with a central server or aggregator [6, 7]. This server integrates the received updates to refine the global model iteratively, enabling collaborative learning while maintaining data privacy. The flexibility and privacy guarantee of FL have led to its adoption in sensitive domains such as healthcare [8], finance [4], and automotive systems [9], where confidentiality is critical.

Although distributed training is feasible for traditional machine learning algorithms such as k-nearest neighbors, support vector machines, Bayesian networks, and decision trees, it typically requires the exchange of raw data, which can lead to significant privacy concerns [10]. To address these issues, FL has emerged as a decentralized learning framework [11–13]. FL enables multiple clients to collaboratively train machine learning models while keeping their data local and private [14]. Each client trains a model on its own dataset and shares only model parameter updates with a central server [15]. This server aggregates updates to improve the global model, allowing for joint model development without compromising data privacy [16]. The flexibility and privacy-preserving nature of FL have led to its growing adoption across industries where data sensitivity is critical [12].

In healthcare, FL facilitates collaborative model development for disease diagnosis by allowing hospitals and research institutions to contribute to model training while maintaining patient data confidentiality [17]. In the finance sector, FL supports robust fraud detection by enabling data collaboration among financial institutions without exposing customer data [13]. Similarly, in the automotive industry, FL is used to develop autonomous driving systems collaboratively, enabling manufacturers to share knowledge derived from driving data without disclosing proprietary information [18].

While existing building energy forecasting models are effective, there remain key areas for improvement. Most current approaches focus on training individual models for each building, often overlooking the advantages of collaboration and shared data. Federated Learning, as a privacy-preserving and distributed learning paradigm, has recently gained significant traction in the energy domain. Applications include probabilistic solar power generation decomposition [19], building heating load forecasting [20], electricity consumption clustering [21], reinforcement learning-based voltage control [22], and voltage forecasting enhanced with differential privacy [23].

Recent reviews have introduced various FL paradigms, such as Horizontal FL, Vertical FL, Transfer FL, Cross-Device FL, and Cross-Silo FL, and emphasized the importance of secure aggregation and encryption techniques to mitigate associated challenges [24]. Traditional centralized machine learning methods used in building energy analysis often neglect privacy concerns and are susceptible to limitations arising from data scarcity. Additionally, differences in energy consumption patterns across buildings present challenges in developing a generalized model. Personalized federated learning has shown potential in addressing both data heterogeneity and privacy but remains underexplored in the context of building energy analytics.

To address these challenges, recent efforts have proposed novel architectures, such as a Mixture of Experts deep learning model, to enhance personalization and better handle the non-i.i.d. nature of building energy datasets [25, 26].

Federated Learning relies on model aggregation to enhance global model performance across distributed clients. Various aggregation algorithms have been proposed to tackle critical challenges, particularly communication efficiency and privacy preservation. Among them, FedAvg [27] remains one of the most widely adopted techniques. It selects a subset of clients in each training round and computes a weighted average of their local model parameters based on dataset size. FedProx [28] extends FedAvg by introducing a proximal term to mitigate the impact of local optima during Stochastic Gradient Descent (SGD) training in federated settings.

In contrast to fixed aggregation, adaptive aggregation methods dynamically adjust to update strategies. For example, a temporally weighted aggregation method [29] enhances global model accuracy by factoring in previous local model contributions. Inverse Distance Aggregation (IDA) [30] is another technique that assigns weights based on model similarity. Additionally, adaptive learning algorithms adjust key training parameters, such as learning rates, in response to changing data distributions, resource constraints, or environmental factors. These enhancements often lead to improved convergence and accuracy. Since SGD is not always optimal in environments with heavy-tailed gradient noise, adaptive optimizers such as

Adagrad, Adam, and Yogi have been integrated into federated learning pipelines to improve training dynamics [31].

This research proposes a novel fuzzy logic-based aggregation method for federated learning. Unlike conventional approaches that rely on gradient-based updates and derivative computations, the proposed method uses fuzzy inference to guide the aggregation process. This technique aims to reduce computational complexity, address challenges arising from data heterogeneity, and mitigate model drift. The fuzzy logic framework provides a flexible and interpretable mechanism for integrating client updates, offering a privacy-aware and computationally efficient alternative for global model optimization.

2. Federated Learning

FL, initially developed by Google researchers in 2016 to address privacy challenges in deep learning, has rapidly expanded its scope of applications. Early applications focused on image classification and mobile-based applications such as keyboard suggestions and language models. Recognizing its potential in various domains, FL has been increasingly adopted in sectors like healthcare, finance, and smart grids, where the need for privacy-preserving data collaboration aligns well with the decentralized nature of the FL framework.

One of the primary aims of FL is to determine the optimal global model of θ that can minimize the aggregated local loss function $lf_k(\theta^k)$.

$$lf_k(\theta^k) = \frac{1}{N_k}\sum_j^{N_k} klf\left(x_j, y_j, \theta^k\right) \tag{1}$$

where x is the input data feature, y is the output data label, N_k is the local data size, klf is the loss function and k is the client index.

$$\min_\theta lf(\theta) = \sum_{k=1}^{Cl\times K} \frac{N_k}{N} lf_k(\theta^k) \tag{2}$$

where Cl is the participation ratio and $N = \sum_{k=1}^{Cl\times K} N_k$.

g_k is the gradient of the k^{th} client that will be sent to the central server for aggregation. The gradient of (1) has been given as below:

$$g_k = \nabla lf_k(\theta^k) \tag{3}$$

In order to update the global model, (4) and (5) have been used to determine global model's weight for $(t+1)^{th}$ instant.

$$\forall k, \theta^{k+1} \longleftarrow \theta^k - \eta . g_k \tag{4}$$

$$\theta^{t+1} \longleftarrow \sum_{k=1}^{Cl\times K} \frac{N_k}{N} \theta^{k+1} \tag{5}$$

In general, FL can be classified as into horizontal and vertical FL based on their characteristics of data distribution among the connected clients.

3. FedFIS: Fuzyy Aggregation

Conventional aggregation methodologies, including those prevalent in existing literature, typically employ historical data to derive update algorithms. In contradistinction to these established paradigms, this research proposes a novel aggregation method predicated on fuzzy logic controller design principles.

Fuzzy logic-based controllers offer a robust and adaptive framework for the management of complex systems characterized by inherent uncertainty and imprecision. Unlike classical control systems reliant on crisp, binary logic, fuzzy logic controllers utilize linguistic variables and fuzzy sets to emulate human-like reasoning. This facilitates the representation of imprecise or ambiguous information, rendering it particularly efficacious in applications where precise mathematical models are intractable. A canonical fuzzy logic controller comprises three core modules: a fuzzification stage, a rule-based inference engine, and a defuzzification stage. The fuzzification stage translates crisp input values into fuzzy sets, while the inference engine applies a predefined set of fuzzy rules to determine appropriate control actions. Finally, the defuzzification stage transforms the fuzzy output back into a crisp value suitable for system control.

This approach engenders controllers that exhibit enhanced resilience to noise, disturbances, and operational variations, thereby improving system performance and stability.

In this study, the current weight value received from each client and the differential between this received weight value and the preceding weight value serve as inputs to the fuzzy controller. The resulting output determines the final weight update magnitude for each client. The architectural schema of the proposed methodology is depicted in Figure 1.

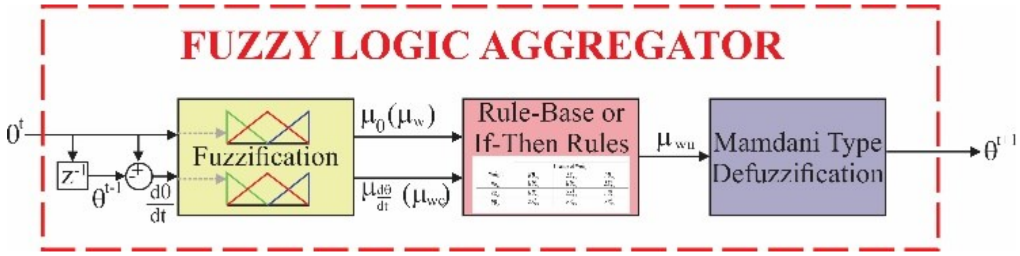

Figure 1. Proposed aggregation method similar to fuzzy logic controller.

In Figure 1, the symbol z^{-1} represents a zero-order hold or unit-time delay element, while θ and $\frac{d\theta}{dt}$ denote the weight value received from each client and the differential between the current and preceding weight values, respectively. The resultant weight update value is computed independently for each client at the termination of each epoch. This methodology effectively mitigates client drift and provides a robust and resilient approach to noise and significant variations. By incorporating both instantaneous and historical weight change magnitudes, the proposed system establishes a more dynamic architectural paradigm. Furthermore, this approach obviates the requirement for complex mathematical operations, such as derivatives, and implements a predictive aggregation methodology by integrating the rate of weight change.

	Change of Weight		
Weight	NB_{wc}	ZZ_{wc}	PB_{wc}
NB_w	NB_{wu}^1	NB_{wu}^2	ZZ_{wu}^3
ZZ_w	NB_{wu}^4	ZZ_{wu}^5	PB_{wu}^6
PB_w	ZZ_{wu}^7	PB_{wu}^8	PB_{wu}^9

Table 1. 9-rule configuration.

The rule-based inference mechanisms employed within this framework are detailed in Tables 1, 2, and 3. The rationale for these rule sets mirrors the conventional methodology used in the development of standard fuzzy logic controllers. Table 1 delineates the rule set for the 9-rule configuration utilizing three membership functions (MFs). In these tables, NB, NS, CZ, PS, and

PB represent Negative Big, Negative Small, Closure to Zero, Positive Small, and Positive Big, respectively. The subscripts, andcorrespond to the new weight value received from each client, the weight change, and the final weight update computed by fuzzy logic, respectively.
This study provides comparative evaluations of various aggregation methodologies within federated learning frameworks to assess the utility of federated learning for photovoltaic power prediction for 14 clients, as shown in Figure 2.

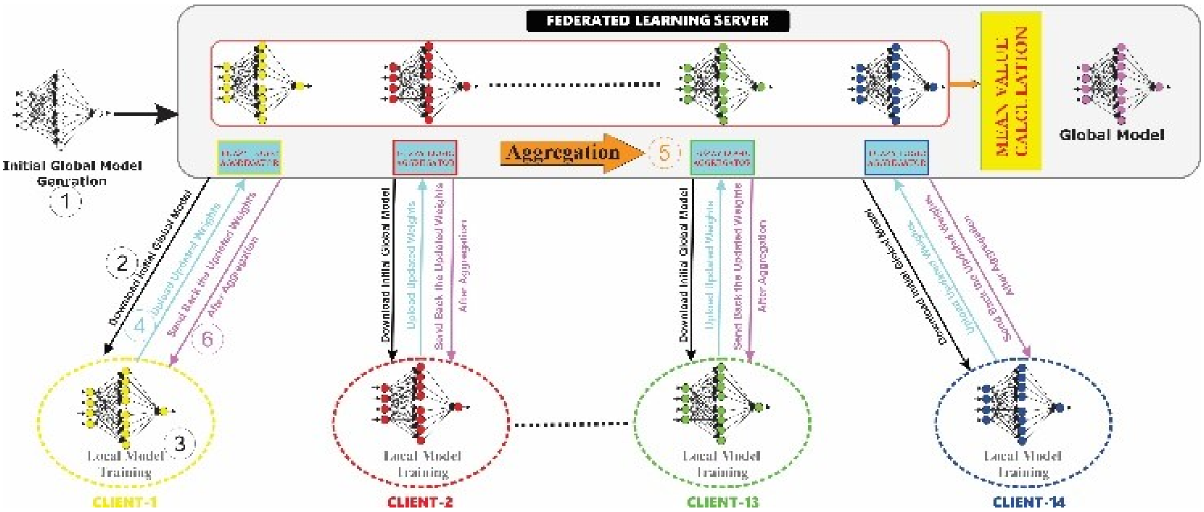

Figure 2. Proposed FedFZY system applied to 14 client case.

The FL process is characterized by a sequence of distinct operational stages:

① *Initialization and ②Model Distribution:* The central server initiates the process by generating a preliminary model, typically through the random initialization of weights within an artificial neural network. This initial model is subsequently disseminated to all participating electrical energy generation forecasting clients.
③ *Decentralized Local Model Training:* Each client engages in local model training utilizing its private dataset. This process avoids the transmission of sensitive raw data. Local training employs gradient descent optimization algorithms to refine model parameters, minimizing a predefined loss function, such as Mean Squared Error (MSE) for regression tasks, thereby enhancing the model's predictive accuracy on local data.
④*Model Parameter Upload to Server:* Upon completion of local training, each client transmits the updated model parameters to the central server for aggregation. Critically, only model parameters, encapsulating the learned knowledge, are shared, preserving the privacy of the underlying datasets.
⑤ *Model Aggregation and ⑥ Global Model Distribution:* The central server aggregates the received model updates to construct an enhanced global model. This study employs a fuzzy logic-based controller for aggregation, utilizing client weight values and their differentials as inputs to determine final weight updates. The refined global model is then distributed back to each client.
Local Model Update and Integration: Each client updates its local model using the aggregated global model parameters, leveraging the collective intelligence of the network. This process facilitates collaborative model training while maintaining client data privacy. The optimal global model is achieved by minimizing the aggregated loss function as defined in Equation 2, across all participating clients.
Iterative Convergence: The aforementioned cycle is repeated until the model converges to an optimal global state or a predefined number of iterations is reached. The FL architecture is depicted in Figure 2.

The FL procedure was executed using a distributed system comprising 15 computational nodes: 14 clients and 1 central server. Each node was equipped with an Intel Core i7 processor (2.40 GHz), 8 GB RAM, 512 GB SSD, and Windows 10 Pro operating system. All training and FL operations were performed within the MATLAB 2022b environment. Inter-node communication was facilitated through an Ethernet protocol via a shared router.

4. Experiments

The experimental results demonstrate that the proposed FedFZY aggregation method achieved the most favorable performance, producing the lowest values for Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Although FedOpt showed performance metrics close to those of FedFZY, the remaining aggregation methods, including FedAvg, FedProx, and TWA, performed comparatively worse.

For comparative analysis, five aggregation methods were implemented within the federated learning framework: FedAvg, FedProx, TWA, FedOpt, and the proposed FedFZY. These methods were used to forecast energy generation characteristics for individual clients. The evaluation was carried out using MAE, RMSE, and MAPE metrics across 14 buildings located in 7 districts of Erzurum.

On average, the FedFZY method outperformed the other aggregation techniques, delivering improvements of 2.77% in MAE, 1.71% in RMSE, and 3.64% in MAPE. These results highlight the effectiveness of FedFZY in the given application.

5. Results and discussion

To facilitate a comprehensive comparison, the aggregation methods FedAvg, FedProx, TWA, FedOpt, and the proposed FedFZY were applied individually to each client within the designed federated learning architecture to forecast energy generation characteristics. The performance of each method was evaluated using MAE, RMSE, and MAPE metrics across 14 distinct buildings located in 7 districts of Erzurum. The minimum values for these metrics are presented in Tables 2 to 4.

		Aggregation Method				
		FedAvg	FedProx	TWA	FedOpt	FedFZY
District	Yakutiye	138.5692	138.6880	129.7358	126.6989	122.0647
	Aziziye	182.8269	196.2342	197.7295	176.9135	171.8654
	Palandoken	253.4424	250.8306	237.0463	229.6824	226.1908
	Tortum	164.2067	165.0764	167.3689	152.7521	143.7930
	Ispir	228.5725	224.9069	220.8055	198.8149	197.4416
	Askale	79.2471	79.8421	80.8624	74.5186	72.1975
	Cat	108.4894	110.1439	112.4134	100.6183	97.8732
	Average	*165.0506*	*166.5317*	*163.7088*	*151.4284*	*147.3466*

Table 2. Model performance in terms of MAE

This study presents a comparative evaluation of various aggregation methods within FL frameworks to assess their effectiveness in photovoltaic (PV) power prediction. A novel FL-based approach is proposed to develop a generalized model capable of accurate forecasting across multiple individual PV farms. Unlike localized learning paradigms, FL enhances generalizability by enabling collaborative model training without sharing raw data. Unlike centralized learning approaches, FL preserves data privacy by ensuring that training data remains on each client's device.

		Aggregation Method				
		FedAvg	FedProx	TWA	FedOpt	FedFZY
District	Yakutiye	160.2125	158.1006	142.7451	142.2913	140.8243
	Aziziye	213.9827	224.8338	210.2695	198.0675	197.3896
	Palandoken	290.8992	272.9868	262.0130	249.8793	245.7975
	Tortum	189.8992	185.5666	176.3057	175.5641	171.8751
	Ispir	258.6381	258.0029	236.5154	233.0638	225.4206
	Askale	189.8992	194.8306	87.9216	83.5781	81.3138
	Cat	92.7525	94.5550	117.0463	116.1255	115.8217
	Average	*199.4690*	*198.4109*	*176.1166*	*171.2242*	*168.3489*

Table 3. Model performance in terms of RMSE

		Aggregation Method				
		FedAvg	FedProx	TWA	FedOpt	FedFZY
District	Yakutiye	0.3635	0.3740	0.3689	0.3466	0.3299
	Aziziye	0.5132	0.5514	0.5750	0.5073	0.4764
	Palandoken	0.6775	0.6521	0.6356	0.6270	0.6037
	Tortum	0.4268	0.4337	0.4434	0.4031	0.3953
	Ispir	0.5855	0.5629	0.5745	0.5619	0.5580
	Askale	0.2206	0.2237	0.2318	0.2184	0.2054
	Cat	0.2980	0.3076	0.3154	0.2857	0.2774
	Average	*0.4407*	*0.4436*	*0.4492*	*0.4214*	*0.4066*

Table 4. Model performance in terms of MAPE

Model generalization capability is evaluated using a client holdout strategy, where the data from one client is reserved for testing, and the remaining clients' data is used for training. After the training and validation phases, a separate holdout subset is utilized to provide a final estimate of model performance. This approach supports the development of models that are robust and applicable to future, unseen data.

Experimental results indicate that the proposed FedFZY aggregation method achieved the best performance in terms of MAE, RMSE, and MAPE. The closest performance to FedFZY was observed with the FedOpt method, while the other evaluated aggregation methods yielded comparatively poorer results, as shown in Tables 2 to 4.

4. Conclusion

This study addresses the complex challenge of distributed photovoltaic (PV) energy forecasting, which is further complicated by strict data privacy requirements across PV farm deployments. To tackle these issues, a privacy-preserving architecture for PV power prediction is proposed. The architecture incorporates a fuzzy logic-based aggregation mechanism within a federated learning (FL) framework, allowing collaborative model training among geographically dispersed PV farm clients while maintaining data privacy. This research introduces a novel contribution by investigating the use of fuzzy logic-based aggregation strategies in the context of federated learning for data-driven PV energy forecasting. By leveraging FL, the approach significantly reduces inter-client and client-server communication bandwidth since raw data is never shared with a central aggregator. As a result, data confidentiality is maintained, and no centralized data repository of individual PV farm observations is created. Comprehensive experimental results confirm that the FedFZY algorithm, which integrates fuzzy logic into the FL paradigm, achieves superior performance compared to conventional aggregation techniques. Moreover, FedFZY demonstrates strong generalization capability on out-of-distribution datasets, highlighting its robustness, adaptability, and suitability for real-world applications.

Authors' Declaration

Conflicts of Interest: The authors declare no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

References

1. D. Mariano-Hernández, L. Hernández-Callejo, A. Zorita-Lamadrid, O. Duque-Pérez and F. S. García, "A review of strategies for building energy management system: model predictive control, demand side management, optimization, and fault detection & diagnosis," Journal of Building Engineering, vol. 33, p. 101692, 2021.
2. D.-S. Lee, C.-W. Lai and Shih-Kai Fu, "A short- and medium-term forecasting model for roof pv systems with data pre-processing," Heliyon, vol. 10, no. 6, p. e27752, 2024.
3. T. Ahmad, H. Zhang and B. Yan, "A review on renewable energy and electricity requirement forecasting models for smart grid and buildings," Sustainable Cities Soc., vol. 55, p. 102052, 2020.
4. R. Selvaraj, V. M. Kuthadi and S. Baskar, "Smart building energy management and monitoring system based on artificial intelligence in smart city," Sustainable Energy Technol. Assess, vol. 56, p. 103090, 2023.
5. S. Goubran, T. Walker, C. Cucuzzella and T. Schwartz, "Green building standards and the United Nations' sustainable development goals," Journal of Environmental Management, vol. 326, p. 116552, 2023.
6. S. Mousavi, M. G. V. Marroquín, M. Hajiaghaei-Keshteli and N. R. Smith, "Data-driven prediction and optimization toward net-zero and positive-energy buildings: a systematic review," Building Environment, p. 110578, 2023.
7. D. Lee, H.-Y. Huang, W.-S. Lee and Y. Liu, "Artificial intelligence implementation framework development for building energy saving," International Journal of Energy Research, vol. 44, no. 14, p. 11908–11929, 2020.
8. Y. Sun, F. Haghighat and B. C. Fung, "A review of the-state-of-the-art in data-driven approaches for building energy prediction," Energy Building, vol. 221, p. 110022, 2020.
9. D. Lee, Y. Chen and S. Chao, "Universal workflow of artificial intelligence for energy saving," Energy Reports, vol. 8, p. 1602–1633, 2022.
10. X. Liu, Y. Deng, A. Nallanathan and M. Bennis, "Federated Learning and Meta Learning: Approaches, Applications, and Directions," IEEE COMMUNICATIONS SURVEYS & TUTORIALS, vol. 26, no. 1, pp. 571-618, 2024.
11. L. Tang, H. Xie, X. Wang and Z. Bie, "Privacy-preserving knowledge sharing for few-shot building energy prediction: a federated learning approach," Applied Energy, vol. 337, p. 120860, 2023.
12. Q. Yang, Y. Liu, T. Chen and Y. Tong, "Federated machine learning: concept and applications," ACM Transaction on Intelligent Systems Technology, vol. 10, no. 2, p. 1–19, 2019.
13. L. Li, Y. Fan, M. Tse and K. Lin, "A review of applications in federated learning," Computers & Industrial Engineering, vol. 149, p. 106854, 2020.
14. C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li and Y. Gao, "A survey on federated learning," Knowledge Based Systems, vol. 216, p. 106775, 2021.
15. Y. Tang, S. Zhang and Z. Zhang, "A privacy-preserving framework integrating federated learning and transfer learning for wind power forecasting," Energy, p. 129639, 2023.
16. Z. Zhang, Y. Wang, X. Ruan and X. Zhang, "A federated transfer learning approach for lithium-ion battery lifespan early prediction considering privacy preservation," Journal of Energy Storage, vol. 102, p. 114153, 2024.
17. S. K. Satapathy, S.-B. Cho, S. Mishra, S. Sah and S. N. Mohanty, "A federated learning approach for classifying chest diseases from chest X-ray images," Biomedical Signal Processing and Control, vol. 100, p. 107107, 2025.

18. M. M. E. Kishawy, M. T. A. El-Hafez, R. Yousri and M. S. Darweesh, "Federated learning system on autonomous vehicles for lane segmentation," Scientific Reports, vol. 4, p. 25029, 2024.
19. J. Lin, J. Ma and J. Zhu, "A privacy-preserving federated learning method for probabilistic community-level behind-the-meter solar generation disaggregation," IEEE Trans. Smart Grid, vol. 13, no. 1, p. 268–279, 2022.
20. H. M. A. Moradzadeh, B. Mohammadi-Ivatloo, A. P. Aguiar and A. Anvari-Moghaddam, "A secure federated deep learning-based approach for heating load demand forecasting in building environment," IEEE Access, vol. 10, p. 5037–5050, 2021.
21. Y. Wang, M. Jia, N. Gao, L. V. Krannichfeldt, M. Sun and G. Hug, "Federated clustering for electricity consumption pattern extraction," IEEE Trans. Smart Grid, vol. 13, no. 3, p. 2425–2439, 2022.
22. H. Liu and W. Wu, "Federated reinforcement learning for decentralized voltage control in distribution networks," IEEE Trans. Smart Grid, vol. 13, no. 5, p. 3840–3843, 2022.
23. J.-F. Toubeau, F. Teng, T. Morstyn, L. V. Krannichfeldt and Y. Wang, "Privacy-preserving probabilistic voltage forecasting in local energy communities," IEEE Trans. Smart Grid, vol. 14, no. 1, p. 798–809, 2023.
24. F. ElRobrini, S. M. S. Bukhari, M. H. Zafar, N. Al-Tawalbeh, N. Akhtar and F. Sanfilippo, "Federated learning and non-federated learning based power forecasting of photovoltaic/wind power energy systems: A systematic review," Energy and AI, vol. 18, p. 100438, 2024.
25. R. Wang, L. Bai, R. Rayhana and Z. Liu, "Personalized federated learning for buildings energy consumption forecasting," Energy & Buildings, vol. 323, p. 114762, 2024.
26. D. Qin, C. Wang, Q. Wen, W. Chen, L. Sun and Y. Wang, "Personalized Federated DARTS for Electricity Load Forecasting of Individual Buildings," IEEE TRANSACTIONS ON SMART GRID, vol. 14, no. 6, pp. 4888-4901, 2023.
27. B. McMahan, E. Moore, D. Ramage, S. Hampson and B. y. Arcas, "Communication-efficient learning of deep networks from decentralized data," Proceedings of the 20th Artificial Intelligence and Statistics, p. 1273–1282, 2017.
28. T. Li, A. Sahu, M. Zaheer, M. Sanjabi and e. al., "Federated optimization in heterogeneous networks," in Proceedings of Machine Learning and Systems, 2020.
29. Y. Chen, X. Sun and Y. Jin, "Communication-efficient federated deep," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 10, p. 4229–4238, 2020.
30. Y. Yeganeh, A. Farshad, N. Navab and S. Albarqouni, "Inverse distance aggregation for federated learning with non-IID data," arxiv:2008, p. 07665, 2020.
31. S. J. Reddi, Z. Charles, M. Zaheer, Z. Garrett, K. Rush, J. Konecný, S. Kumar and H. B. McMahan, "Adaptive Federated Optimization," in ICLR 2021, 2021.