



## OPEN ACCESS

## EDITED BY

Hao Wang,  
Monash University, Australia

## REVIEWED BY

Lv Chaoxian,  
China University of Mining and Technology,  
China  
Yang Lei,  
Anhui University, China  
Yongliang Liang,  
Shandong University, China

## \*CORRESPONDENCE

Mengdi Wei,  
✉ wmd18054369586@tju.edu.cn

RECEIVED 05 December 2023

ACCEPTED 17 January 2024

PUBLISHED 07 February 2024

## CITATION

Chen B, Liu R, Wei M, Wang X, Sun Y and Sun D (2024), PV output estimation method of power distribution station area based on federated learning framework and improved transformer neural network.

*Front. Energy Res.* 12:1349995.

doi: 10.3389/fenrg.2024.1349995

## COPYRIGHT

© 2024 Chen, Liu, Wei, Wang, Sun and Sun. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# PV output estimation method of power distribution station area based on federated learning framework and improved transformer neural network

Bo Chen<sup>1</sup>, Rui Liu<sup>1</sup>, Mengdi Wei<sup>2\*</sup>, Xian Wang<sup>1</sup>, Yi Sun<sup>1</sup> and Donglei Sun<sup>1</sup>

<sup>1</sup>Economic and Technological Research Institute, State Grid Shandong Electric Power Company, Jinan, China, <sup>2</sup>School of Electrical Automation and Information Engineering of Tianjin University, Tianjin, China

In contemporary distribution networks (DNs), characterized by extensive integration of distributed energy, the photovoltaic (PV) power output data from the distribution station areas become crucial for system planning and operational optimization. Since many PVs are installed behind-the-meter (BTM), it is difficult to directly obtain PV power data through measurement devices. Therefore, it is important to estimate the BTM PV power from the aggregating data that can be directly obtained. However, the existing estimation methods usually require centralized large-scale data training, which brings certain privacy leakage risks. In order to solve these problems, we propose a federated learning-based improved Transformer Neural Network strategy to estimate BTM PV generation at the community level with data privacy protection. Initially, enhanced Transformer neural networks, employing a fused-attention mechanism, are deployed to precisely delineate the solar power generation pattern. Subsequently, federated learning principles facilitate the sharing of specific parameters among multiple edge endpoints and a central server. This model bifurcates into two layers: an individual layer, where parameters are retained locally, and an exchange layer, where parameters are collectively shared and conveyed through momentum aggregation. This dual-layer structure effectively synchronizes the capture of both unique and common characteristics. The test on the Australian residential load dataset verifies the effectiveness of the proposed method.

## KEYWORDS

PV Estimation, BTM, improved Transformer, federated Learning, momentum aggregation

## 1 Introduction

As the energy landscape evolves, the adoption of distributed photovoltaics, particularly on residential and commercial building rooftops, has surged. By the end of 2020, approximately 2.66 million Australian homes were equipped with rooftop solar systems (Australian Government, 2020). However, due to budgetary limitations, most of these installations lack dedicated metering devices, making direct power measurement challenging. The growing prevalence of such unmonitored distributed PV systems poses

significant risks to the economic and safe operation of distribution station areas. These include issues like voltage instability (Rosado and Khadem, 2019), inaccurate load forecasting (Wang et al., 2017), and suboptimal fault recovery (Chen et al., 2018). Consequently, accurately estimating behind-the-meter (BTM) PV power generation using robust methodologies has become increasingly crucial.

The methodological core of estimating BTM PV is to establish a link between other data available and unknown PV data. It can be divided into two main categories: physical model-based and data-driven methods.

The physical model-based approaches employ PV array performance models to represent physical PV arrays. In (Chen and Irwin, 2017), PV arrays are integrated with a clear-sky model to estimate customer-level solar power generation. Reference (Wang et al., 2017) employs a virtual equivalent PV plant model to represent the collective power generation of regional behind-the-meter (BTM) PV systems. Reference (Cheung et al., 2018) introduces an unsupervised consumer mixture model for PV estimation. In (Kabir et al., 2019; Kabir et al., 2023), physical PV models and statistical models are respectively utilized to estimate BTM solar power generation and local demand. A major drawback of these physical model-based approaches is the need for detailed PV array parameters or accurate meteorological data. However, in practice, these parameters are usually not available to the utility. In addition, obtaining meteorological data may impose additional costs on utilities. In (Li et al., 2019a), the net demand under heterogeneous weather conditions is used to estimate the BTM PV capacity, which is multiplied with the standard solar power time series to infer the BTM PV generation. Reference (Sossan et al., 2018) proposed an estimation method on transformer level PV generation based on the variation difference of load and solar power.

Data-driven methodologies predominantly utilize artificial intelligence techniques for deep learning model training. As detailed in reference (Lin et al., 2022), a method for real-time energy decomposition is proposed, enabling the segregation of behind-the-meter (BTM) solar energy from a substation's total energy consumption, employing partially labeled aggregated data for model training. Reference (Shaker et al., 2016) introduces a data-driven strategy for estimating BTM solar power generation, encompassing data dimensionality reduction and mapping functions, and involves selecting a limited number of representative sites for model training. Since support vector machine has the ability to learn the complex hidden relationships in the energy decomposition problem, reference (Li et al., 2019b) proposed a supervised machine learning model based on multiple Support Vector Regression (SVR) to regression the solar power generation on the extracted feature input. However, these methods fail to make full use of historical data to comprehensively reflect the features of different resources and mine the intrinsic relationship of historical data. At the same time, many time-series regression models with excellent performance are not considered.

In addition, traditional data-driven methods are centralized, and their performance largely depends on the quality and quantity of training data. However, this centralized approach may raise a series of concerns about data privacy, especially in a competitive electricity market (Li et al., 2023). Reference (Yang et al., 2023) solves the

potential privacy problem by introducing the framework of federated learning. In literature (Lin et al., 2022), the problem of privacy protection and the uncertainty of various types of data are considered, and the Bayesian neural network is used for PV estimation.

Building on the aforementioned analysis, this study introduces an enhanced Transformer neural network approach, underpinned by federated learning, designed to estimate community-level BTM PV generation. Federated learning mechanism can share training parameters and private training data, which can not only improve data quality but also protect user privacy. Considering the successful application of the self-attention mechanism in the electric power field in recent years (Azam and Younis, 2021; Wang et al., 2023; Zhou et al., 2023), a decentralized improved Transformer model is designed, and the model is further divided into two parts: individual and center, so as to adapt to the specific needs of each community (Tzeng et al., 2014; Chen et al., 2020). Furthermore, the study implements a hierarchical parameter strategy for model updating. Experimental evaluations demonstrate the efficacy of this method on real-world datasets.

In summary, this paper makes the following contributions.

1. The refined Transformer model, applied to BTM PV estimation, incorporates a fused attention mechanism to thoroughly extract PV-related information from net load data.
2. The federated learning mechanism is used to realize parameter sharing between different communities while considering privacy protection, so the receptive field and generalization ability of the model are improved.
3. A novel hierarchical parameter update mechanism is introduced in the enhanced Transformer model, consisting of an exchange layer and an individual layer. This configuration facilitates the sharing of exchange layer parameters through momentum aggregation while retaining individual layer parameters, thereby augmenting the model's capacity to address both individual and common feature challenges.

The remaining structure of the paper is as follows: Section II introduces the main models involved in this paper; Section III introduces the data set used and the numerical example verification. Section IV provides a summary and outlook.

## 2 Main models

### 2.1 Community PV estimation model

Since most of the user-level distributed PVs are deployed behind the meters, only the net load data can be directly observed for most user. A small number of users can directly obtain the photovoltaic output information. The net load power value is equal to the difference between the load and the PV power, as shown in the following equation:

$$P_{Ni,t} = P_{Li,t} - P_{Gi,t} \quad (1)$$

Where  $P_{Ni,t}$  is the observed net load of the corresponding user of the  $i$ th community at time  $t$ ;  $P_{Li,t}$ ,  $P_{Gi,t}$  is the corresponding actual

load value and the BTM PV power generation value. Since there is no illumination intensity at night, 7–17 h are selected as the reference in this paper.

Through the above analysis, the data-driven based BTM PV estimation strategy can be abstracted and modeled as:

$$P_G = h(q, P_N, S) \quad (2)$$

Where  $P_G$  is the PV estimated value,  $h$  is the transfer function of the network,  $q$  is the network parameter,  $P_N$  is the available net load data,  $S$  is relevant weather information.

## 2.2 Federated learning framework

Data-driven BTM PV decomposition strategies rely on large amounts of data for model training. It is difficult for a community to collect a large amount of labeled data. On the one hand, only a small number of users can measure the PV behind the meter. On the other hand, collecting data from neighboring communities has the risk of privacy leakage. Centralized data collection and optimization models are not advantageous in dealing with the BTM PV estimation problem.

This paper addresses the challenge by employing a federated learning framework. Federated learning offers a mechanism enabling multiple entities to collaboratively train and refine a model, without the need for direct raw data sharing. This approach not only safeguards data privacy and security but also facilitates the distributed training of the BTM solar power generation estimation model across multiple data sources, thereby enhancing the model's generalization capability. A supervised model for photovoltaic estimation is constructed under the federated learning framework, in the form of [Formula 3](#):

$$[P_{G,1}, P_{G,2}, \dots, P_{G,n}] = H(q_1^p(P_{N,1}, S_1), q_2^p(P_{N,2}, S_2), \dots, q_n^p(P_{N,n}, S_n)) \quad (3)$$

Where  $i$  is the index of the community,  $P_{G,i}$  is the estimated PV output,  $H$  is the global shared parameters,  $q_i^p$  is the individual parameter,  $P_{N,i}$  is the net load data as input, and  $S_i$  is the meteorological information.

It is worth noting that there are differences in the relationship between net load and solar power generation in different local communities due to geographical location and weather conditions. In order to capture this difference, refer to the idea of reference ([Azam and Younis, 2021](#)), the model layer is divided into individual layer and exchange layer, where:

$$q_i = [q_i^p; q_i^c] \quad (4)$$

Where  $q_i$  is the learnable parameters of the  $i$ th community,  $q_i^p$  is the individual parameter,  $q_i^c$  is the exchange parameter.

In the implementation of the federated learning framework, a momentum aggregation strategy is selectively employed, particularly suitable for the complex nonlinear dynamics and potential non-convex optimization challenges inherent in photovoltaic (PV) estimation. This strategy expedites model convergence towards an optimal solution by integrating gradient information from preceding iterations, a critical factor in handling large-scale, intricate datasets. Additionally, when

processing data from diverse buildings and regions, the model often faces oscillatory challenges. Momentum aggregation addresses this by smoothing the model's update process, thereby enhancing the stability and reliability of the training procedure. The parameter update process based on momentum aggregation is formulated as follows:

$$H^{t+1} = \beta H^t + (1 - \beta) \left( \frac{1}{N} \sum_{i=1}^N q_i^{c,t+1} \right) \quad (5)$$

Where  $t$  is the number of iterations index,  $\beta$  is the parameter of momentum,  $N$  is the number of communities.

The specific structure of BTM PV estimation based on the federated learning framework is shown in [Figure 1](#).

## 2.3 Improved transformer network

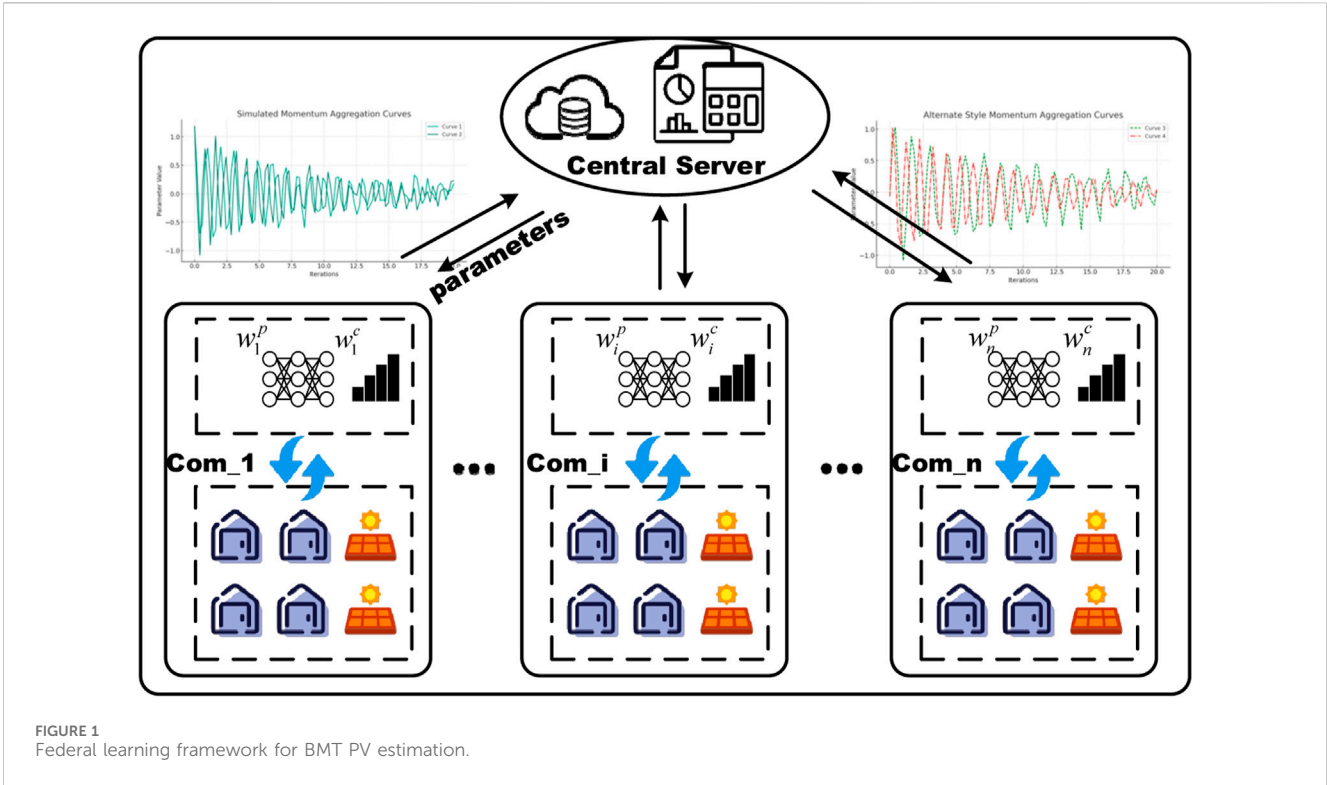
Transformer is a network architecture based on self-attention mechanism, which inherits the advantages of self-attention mechanism in dealing with long distance sequences with volatility. At the same time, it integrates residual network and other structures to solve the problems of gradient disappearance caused by depth deepening, so that the model has higher plasticity. All these advantages enable it to achieve better results in dealing with long-distance serial regression problems.

The complexity of PV power data, characterized by its time-series attributes, poses challenges for traditional attention mechanisms, especially in capturing short-term (e.g., intraday) and long-term (e.g., seasonal) variations. The integration of local and global attention mechanisms offers a solution. The local mechanism concentrates on accurately predicting critical periods, while the global mechanism discerns long-term trends and impacts, thereby enriching the overall comprehension of factors influencing PV output. This dual approach not only bolsters the model's capacity to navigate data nonlinearity and variability, thereby enhancing prediction accuracy and robustness, but also augments model interpretability by delineating distinct temporal dependencies. Crucially, this method also elevates computational efficiency, a vital aspect for processing large-scale PV data. The expression is as follows:

$$Ac(Q, K, V) = \lambda \cdot M \left( \text{soft max} \left( \frac{QK}{\sqrt{d_k}} \right) V \right) + \gamma \cdot \text{soft max} \left( \frac{QK}{\sqrt{d_k}} \right) V \quad (6)$$

Where  $AC$  is the output of fused self-attention,  $M$  is the window selection function,  $\lambda$ ,  $\gamma$  is the weight parameters, respectively.  $Q$ ,  $K$  and  $V$  is the query, key and value, respectively.  $d_k$  is the dimension of  $K$ .

When training the model, the inputs of the encoder and decoder are timing data and training labels respectively. Both are mapped to the high-dimensional space through a linear network, and the high-dimensional data are endowed with location characteristics by the location coding function. Then, the input data passes through  $N$  cyclic units composed of multi-head fusion attention structure and feedforward neural network respectively, and the output is finally obtained.



The feedforward neural network is composed of fully connected layers with the transfer function as follows:

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (7)$$

Where,  $FFN(x)$  is the output of the fully connected layer,  $W_1, b_1$  and  $W_2, b_2$  are the parameters of the two fully connected layers, and non-linear transformation is performed by retaining non-negative data to improve the network expression ability.

Due to the existence of the residual network structure, the actual output of each layer in the network is the sum of the theoretical output and the residual:

$$O_i = Norm(in_i + o_{idea}) \quad (8)$$

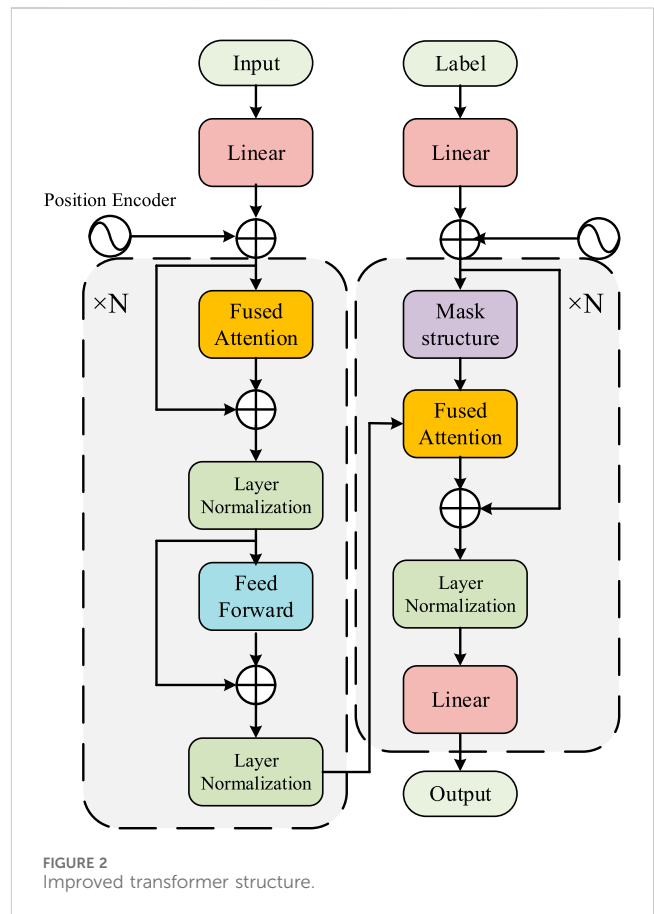
Where  $in_i$  is the input of the  $i$ th layer structure in the model;  $o_{idea}$  is the theoretical output of this layer.  $Norm$  is the layer normalization operation.  $o_i$  is the actual output of layer  $i$ .

The attention mechanism, by applying attention operations uniformly across all data, results in an output matrix devoid of original positional information. This leads to the model's inability to learn the data's order information, necessitating the infusion of positional features into the input sequence.

The sine and cosine functions with different frequencies are used to encode the position, so as to give the absolute and relative position information to the time series data. The position-encoding function is as follows:

$$P_{x,2i} = \sin\left(\frac{x}{10000^{2i/dm}}\right) \quad (9)$$

$$P_{x,2i+1} = \cos\left(\frac{x}{10000^{2i/dm}}\right) \quad (10)$$



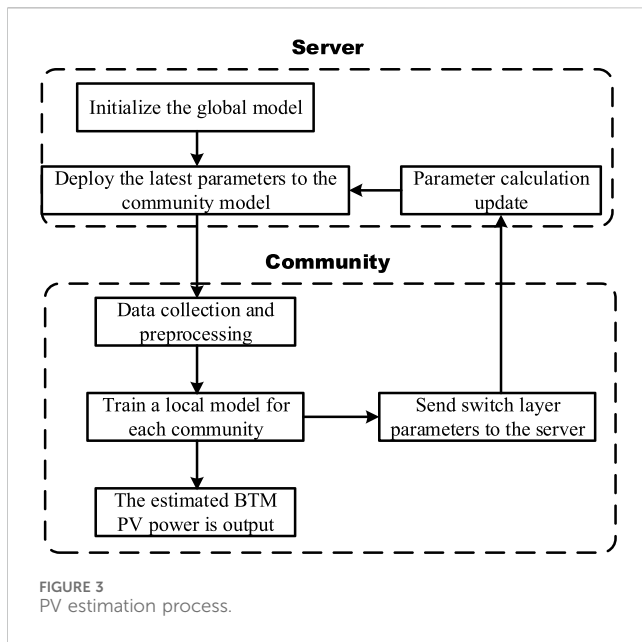


FIGURE 3 PV estimation process.

Where  $d_m$  is the dimension of time series data,  $i$  is the dimension index,  $x$  is the absolute position,  $P_x$  is the position encoding.

In the decoder structure, the initial output reaches the same dimension as the training label through a normalization layer and a linear layer, and the final output is the estimated result.

The improved Transformer structure is shown in Figure 2:

The enhanced Transformer network serves as both the sub-network and global network for each community. It undergoes hierarchical processing and parameter updates in the individual and exchange layers, aligned with the federated learning architecture. This culminates in the creation of a cloud-edge collaborative supervised regression model, where a central server facilitates interaction, and multiple communities function as distinct training units. The overall process is shown in Figure 3.

### 3 Results analysis

#### 3.1 Introduction of the dataset and description of the experimental platform

The performance of the proposed model is evaluated using the Ausgrid Solar Home Electric (ASHE) dataset (Ratnam et al., 2017). Comprising actual smart meter data from households in and around Sydney, Australia, this publicly accessible dataset encompasses controlled load consumption for each household, general consumption over the entire period, and solar power generation data at 30-min intervals.

After data preprocessing, the data of 300 users in 4 communities from 1 July 2010 to 30 June 2011 were selected for analysis. The actual load of a user can be directly determined by the sum of controlled and general consumption. The corresponding net demand is calculated by subtracting the solar generation from the actual load. Subsequently, by aggregating the measurements from the consumer communities, we can obtain the net load and BTM solar generation at the community level.

TABLE 1 Hyperparameter setting.

Hyperparameter description	Value
Number of parameter exchanges	20
Number of sub-model training rounds	100
Learning rate	1e-3

Experimental platform setup: Intel i5-13600k and NVIDIA GeForce RTX 3070 are used as the core processors, Python 3.7 is used as the programming language for the algorithm model, and the network model is built based on the open source machine learning framework Pytorch.

The hyperparameter Settings of the proposed model are shown in Table 1.

#### 3.2 Evaluation function

Root Mean Square Error (RMSE), Standardized Root Mean Square Error (NRMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are used as indicators to measure the prediction results of the model, and the expressions are as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{i\_tru} - y_i)^2} \tag{11}$$

$$NRMSE = \frac{RMSE}{y_{i\_tru\_max} - y_{i\_tru\_min}} \tag{12}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{i\_tru} - y_i| \tag{13}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{i\_tru} - y_i}{y_{i\_tru}} \right| \tag{14}$$

Where,  $y_{i\_tru}$  is the real value of photovoltaic power generation at time  $i$ ;  $y_i$  is the predicted value of photovoltaic power generation at time  $i$ ;  $n$  is the total number of sampled data.

#### 3.3 Error comparison

To assess the decomposition performance of the proposed model, it is compared with different learning frameworks and state-of-the-art methods. These include Local Learning (LL), where communities independently train and evaluate the model without interaction; Centralized Learning (CL), which involves aggregating data from all communities for unified model training by a central server; and Federated Learning (FL), where each community trains the model individually, exchanges layer parameters with a central server for updating global model parameters, and then conducts gradient descent training. We call the attention-integrated Transformer model FTransformer.

Specifically, FL-FTransformer, FL-BNN, FL-LSTM, FL-Transformer, LL-FTransformer, and CL-FTransformer are compared.

The historical net load data of the current moment and the same time of the previous week and the corresponding meteorological

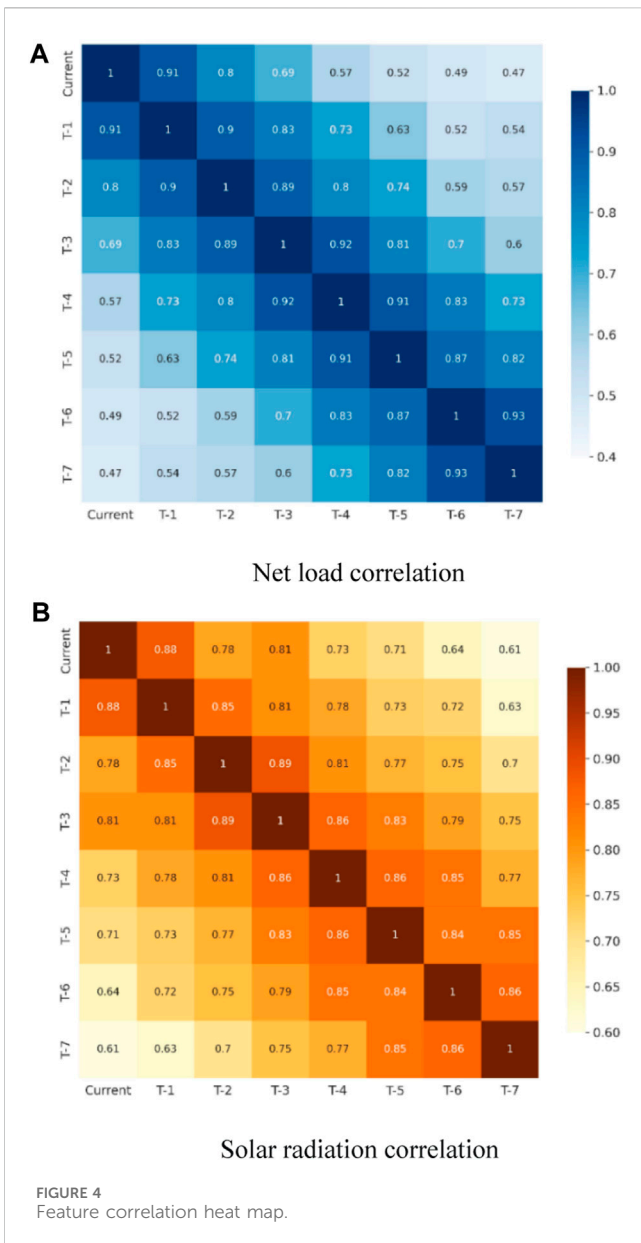


FIGURE 4 Feature correlation heat map.

data are selected as the input features of the model. Figure 4 illustrates the correlation between the net load value and solar radiation at the historical time and the PV power at the current time.

According to the heat map, it can be found that the closer to the current time, both the net load data and the solar radiation intensity have a greater impact on the PV power at the current time, which proves the effectiveness of feature selection.

Figure 5 selects part of community 1 to visually compare the estimated PV power output of FL-FTransformer network with the real value. It can be found that even in the case of large data fluctuations, the proposed model can still achieve accurate estimation of PV power. Since the visual difference between many models in the comparison diagram is not large, and it is easy to stack together, it is impossible to intuitively compare the used model with the real value. Therefore, the power curve of other models is not added in the comparison diagram, and the effectiveness of the proposed model is proved through the comparison of error indicators.

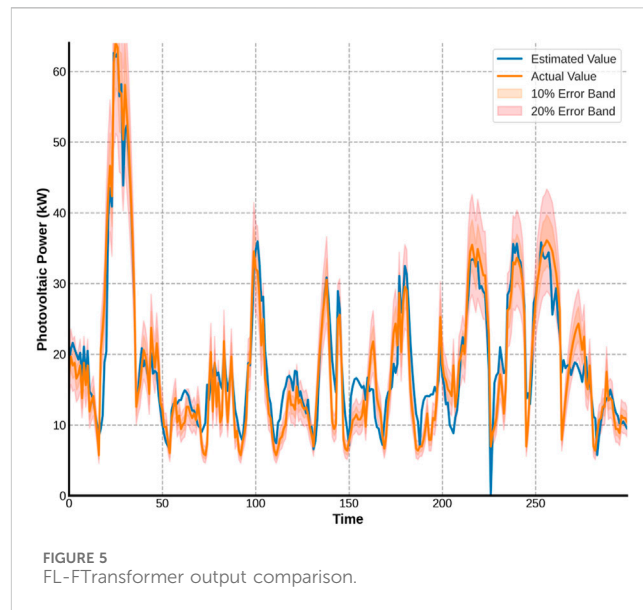


FIGURE 5 FL-FTransformer output comparison.

TABLE 2 Error comparison of different models.

	RMSE (kW)	NRMSE (%)	MAE (kW)	MAPE (%)
FL-FTrans	1.52	2.43	4.56	11.72
FL-BNN	2.39	3.37	5.22	13.47
FL-LSTM	2.15	2.88	4.93	12.58
FL-Trans	2.21	2.75	5.04	12.43
LL- FTrans	3.62	5.47	6.52	15.88
CL- FTrans	1.79	2.74	4.88	12.14

Table 2 shows the error index situation of the proposed FL-FTransformer model and the comparison model. Under the same FL framework, the improved FTransformer improves RMSE and MAPE respectively compared with BNN, LSTM and Transformer 0.87kW, 0.63kW, 0.69kW, 1.65%, 0.86%, and 0.71%. Under the same improved FTransformer network, FL, compared with LL and CL, improves 2.1kW, 0.27kW, 4.16% and 0.42%.

In a horizontal analysis, the augmented FTransformer model, integrated within the federated learning (FL) framework, outperforms across four key metrics. This superiority is attributed to its combined attention and parallel processing capabilities, which provide robust and versatile model for temporal regression tasks like BTM PV estimation. In a vertical comparison, the FL architecture, also employing the enhanced FTransformer as its core network, demonstrates superior performance. This is due to FL's ability to expand data scope and enhance model generalization compared to local learning. Additionally, unlike centralized learning, FL's two-layer training mechanism not only emphasizes the unique characteristics of each community but also safeguards user data privacy while broadening the data pool.

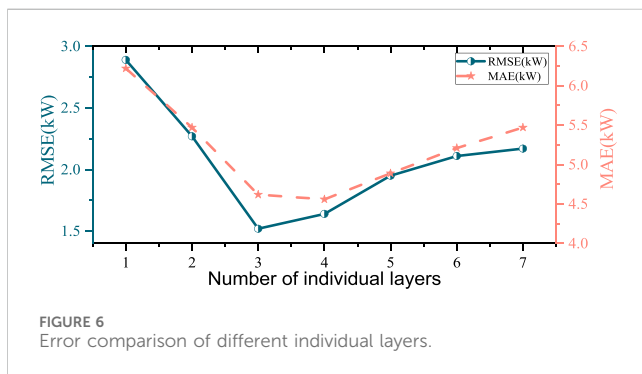
Table 3 compares the performance of different models in different seasons, and the FL-FTransformer model achieves the

TABLE 3 Error comparison of different seasons.

	Spring	Summer	Autumn	Winter
FL-FTrans	1.55	1.61	1.58	1.65
FL-BNN	2.37	2.44	2.35	2.40
FL-LSTM	2.13	2.32	2.21	2.18
FL-Trans	2.25	2.28	2.64	2.21
LL- FTrans	3.60	3.85	3.69	3.71
CL- FTrans	1.82	1.85	1.71	1.77

TABLE 4 Error Comparison of Different communities.

	com1	com2	com3	com4
Estimated value (kW-h)	62.55	70.42	65.84	72.49
True value (kW-h)	65.72	67.49	71.34	70.06
Deviation (kW-h)	3.17	2.93	5.50	2.43
Proportion of deviation (%)	4.82	4.34	7.71	3.47



best performance in different seasons. Due to the fluctuation difference of illumination radiation in different seasons, the specific error indicators are also different. The incorporation of the fused attention mechanism enables the enhanced FTransformer network to adeptly manage varying degrees of PV fluctuations across different seasons.

The estimated value of PV generation in each community is compared with the true value from 12:00 to 13:00 when the solar radiation is sufficient. Since the PV power does not fluctuate much in the time scale of 1 hour, the PV power at 12:30 is selected as the average power in this period for the calculation of power generation, and the summary results are shown in Table 4.

As can be seen from Table 4, the error of PV generation of each community in the selected period is about 5%, which has a high accuracy, indicating that the proposed model has the ability to accurately estimate the regional PV power generation.

Figure 6 illustrates the error comparison for different number of individual layers. In summary, when the number of individual layers is 3, the model evaluation effect is the best, and it is used as the final output model. It shows that the model has a compromise between focusing on community characteristics and global commonality.

## 4 Conclusion

This study addresses regional photovoltaic (PV) power estimation by proposing a deep learning model that leverages an improved Transformer within a federated learning framework. The model's efficacy is validated through experiments, yielding notable findings:

- 1) The FL-Transformer model effectively utilizes community-level data in a distributed manner for behind-the-meter (BTM) PV estimation, enhancing the model's generalizability, estimation accuracy, and user privacy protection.
- 2) The proposed two-layer training mechanism not only pays attention to the characteristics of a single community data, but also pays attention to the commonality between multiple communities, which significantly improves the performance of the model. The method of updating global parameters by momentum aggregation improves the robustness and solution efficiency of the model.
- 3) Enhancements to the traditional Transformer model, incorporating both global and local attention mechanisms, markedly improve the model's capacity to discern long and short cycle features, thereby refining decomposition accuracy.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

## Author contributions

BC: Conceptualization, Writing–review and editing. RL: Investigation, Writing–review and editing. MW: Methodology, Writing–original draft. XW: Data curation, Writing–review and editing. YS: Supervision, Writing–review and editing. DS: Formal Analysis, Writing–review and editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research is funded by the Science and Technology Project of Economic and Technological Research Institute, State Grid Shandong Electric Power Company (52062522000N). This research was funded by the Spatial and Temporal Layout of Shandong New Energy Power Project considering development potential and system constraints.

## Conflict of interest

Authors BC, RL, XW, YS, and DS were employed by the State Grid Shandong Electric Power Company.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Australian Government (2020). Solar PV installation information in Australia. [Online]. Available: <https://www.energy.gov.au/households/solar-pv-and-batteries>.
- Azam, M. F., and Younis, M. S. (2021). Multi-horizon electricity load and price forecasting using an interpretable multi-head self-attention and EEMD-based framework. *IEEE Access* 9, 85918–85932. doi:10.1109/access.2021.3086039
- Chen, B., Chen, C., Wang, J., and Butler-Purry, K. L. (2018). Sequential service restoration for unbalanced distribution systems and microgrids. *IEEE Trans. Power Syst.* 33 (2), 1507–1520. doi:10.1109/tpwrs.2017.2720122
- Chen, D., and Irwin, D. (2017). "SunDance: black-box behind-the-meter solar disaggregation," in Proc. 8th Int. Conf. Future Energy Syst, Shatin Hong Kong, May, 2017, 16–19.
- Chen, Y., Sun, X., and Jin, Y. (2020). Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation. *IEEE Trans. Neural Netw. Learn. Syst.* 31 (10), 4229–4238. doi:10.1109/tnnls.2019.2953131
- Cheung, C. M., Zhong, W., Xiong, C., Srivastava, A., Kannan, R., and Prasanna, V. K. (2018). "Behind-the-meter solar generation disaggregation using consumer mixture models," in Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids, Aalborg, Denmark, October, 2018, 1–6.
- Kabir, F., Yu, N., Yao, W., Yang, R., and Zhang, Y. (2019). "Estimation of behind-the-meter solar generation by integrating physical with statistical models," in IEEE Int. Conf. Commun. Control Comput. Technol. Smart Grids, Beijing, China, October, 2019, 1–6.
- Kabir, F., Yu, N., Yao, W., Yang, R., and Zhang, Y. (2023). Joint estimation of behind-the-meter solar generation in a community. *IEEE Trans. Sustain. Energy* 12 (1), 682–694. doi:10.1109/tste.2020.3016896
- Li, D., Guo, Q., and Feng, J. (2023). Private settlement model of distributed power transactions based on blockchains. *Power Syst. Technol.*, 1–19.
- Li, K., Wang, F., Mi, Z., Fotuhi-Firuzabad, M., Duić, N., and Wang, T. (2019a). Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation. *Appl. Energy* 253, 113595. doi:10.1016/j.apenergy.2019.113595
- Li, K., Wang, F., Mi, Z., Fotuhi-Firuzabad, M., Duić, N., and Wang, T. (2019b). Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation. *Appl. Energy* 253, 113595. doi:10.1016/j.apenergy.2019.113595
- Lin, J., Ma, J., and Zhu, J. (2022). A privacy-preserving federated learning method for probabilistic community-level behind-the-meter solar generation disaggregation. *IEEE Trans. Smart Grid* 13 (1), 268–279. doi:10.1109/tsg.2021.3115904
- Ratnam, E. L., Weller, S. R., Kellett, C. M., and Murray, A. T. (2017). Residential load and rooftop PV generation: an Australian distribution network dataset. *Int. J. Sustain. Energy* 36 (8), 787–806. doi:10.1080/14786451.2015.1100196
- Rosado, S. P., and Khadem, S. K. (2019). Development of community grid: review of technical issues and challenges. *IEEE Trans. Industry Appl.* 55 (2), 1171–1179. doi:10.1109/tia.2018.2883010
- Shaker, H., Zareipour, H., and Wood, D. (2016). A data-driven approach for estimating the power generation of invisible solar sites. *IEEE Trans. Smart Grid* 7 (5), 2466–2476. doi:10.1109/tsg.2015.2502140
- Sossan, F., Nespoli, L., Medici, V., and Paolone, M. (2018). Unsupervised disaggregation of photovoltaic production from composite power flow measurements of heterogeneous prosumers. *IEEE Trans. Industrial Inf.* 14 (9), 3904–3913. doi:10.1109/tii.2018.2791932
- Tzeng, E., Hoffman, J., Zhang, N., Saenko, K., and Darrell, T. (2014). Deep domain confusion: maximizing for domain invariance. [Online]. Available: <https://arxiv.org/abs/1412.3474>.
- Wang, W., Feng, B., Huang, G., Liu, Z., and Ji, W. (2023). Short-term net load forecasting based on self-attention encoder and deep neural network. *Proc. CSEE*, 1–13.
- Wang, Y., Zhang, N., Chen, Q., Kirschen, D. S., Li, P., and Xia, Q. (2017). Data-driven probabilistic net load forecasting with high penetration of behind-the-meter PV. *IEEE Trans. Power Syst.* 33 (3), 3255–3264. doi:10.1109/tpwrs.2017.2762599
- Yang, T., Qin, X., Xiangwei, F., and Xu, Z. (2023). Federated learning electric vehicle short-term charging load prediction accounting for user charging behavior and privacy protection. *High. Volt. Eng.*, 1–8.
- Zhou, S., Li, Y., Guo, Y., Qiao, X., and Liu, Y. (2023). Ultra-short-term load forecasting based on temporal convolutional network considering temporal feature extraction and dual attention fusion. *Automation Electr. Power Syst.*, 1–21.