

Federated Learning for Solar Power Forecasting: A Privacy-Preserving, Multi-Site Approach

Manish Gupta¹, Pawan Kumar Verma²

^{1,2} Lincon University College, Malaysia

Email ID : pdf.manishgupta@lincoln.edu.my¹, abes.pawan@gmail.com²

Abstract

The rapid expansion of distributed photovoltaic (PV) installations — from residential rooftops to utility-scale solar farms — has made accurate short-term power forecasting essential for grid stability, energy trading, and storage dispatch. Conventional forecasting pipelines typically require pooling raw irradiance and power measurements from many sites into a central location, an approach that is increasingly difficult to sustain given data ownership rules, competitive sensitivities between plant operators, and the bandwidth cost of transmitting high-frequency sensor data. This paper examines federated learning (FL) as an alternative training paradigm in which a shared forecasting model is improved collaboratively across many solar sites without any raw data leaving its site of origin. We describe the architecture of a federated solar forecasting system, discuss the statistical and communication challenges introduced by non-identically distributed (non-IID) solar data, and review aggregation strategies including Federated Averaging (FedAvg) and FedProx. We then present an illustrative simulation study comparing local-only, federated, and centralized training strategies on synthetic multi-site solar data, evaluated using RMSE, MAE, and MAPE. Across this simulation, federated approaches recovered most of the accuracy gap to centralized training while keeping all raw data on-site, with FedProx showing the most stable convergence under heterogeneous conditions. We conclude by outlining open challenges — including client drift, communication overhead, fairness across sites of unequal size, and privacy leakage through model updates — and directions for future work such as personalized federated forecasting and secure aggregation.

Keywords: *Federated Learning, Solar Power Forecasting, Photovoltaic (PV) Systems, Privacy-Preserving Machine Learning, Renewable Energy, Time-Series Prediction*

1. Introduction

Solar photovoltaic generation is one of the fastest-growing sources of electricity worldwide, but its output is inherently variable: cloud cover, aerosol content, panel soiling, and temperature all cause short-term fluctuations that are difficult for grid operators to accommodate. Accurate forecasting of PV output over horizons ranging from minutes to a day ahead allows utilities to schedule reserves, plan storage charging and discharging, and reduce curtailment. Machine learning models, particularly recurrent neural networks, gradient-boosted trees, and more recently transformer-based architectures, have consistently outperformed purely physical or statistical forecasting methods when trained on sufficient historical data. The practical difficulty is that no single solar site usually has enough historical data, sensor diversity, or weather variability to train a robust general-purpose model on its own. The natural remedy — pooling

data from many sites into a shared, centrally trained model — runs into three recurring obstacles. First, plant operators and utilities are often unwilling or contractually unable to share raw generation data, which can reveal information about asset performance, maintenance schedules, or customer behavior. Second, regulatory regimes governing energy and personal data (for example, rules that apply to smart-meter and rooftop-solar data in many jurisdictions) restrict cross-border or cross-organization transfer of granular measurements. Third, continuously streaming high-frequency sensor data from thousands of distributed sites to a central server is costly and creates a single point of failure.

Federated learning offers a way around these obstacles. Instead of moving data to the model, FL moves the model to the data: each participating site trains a copy of a shared model on its own local records, and only the resulting model updates — not the underlying measurements — are sent to a coordinating server, which aggregates them into an improved global model. This paper focuses specifically on how this paradigm applies to solar power forecasting, where data across sites differs systematically in scale, climate, panel technology, and local weather patterns, making it a challenging but realistic testbed for federated methods.

The remainder of the paper is organized as follows. Section 2 reviews related work on federated learning applied to renewable energy forecasting. Section 3 provides background on solar forecasting and the mechanics of federated optimization. Section 4 presents the proposed system architecture and methodology. Section 5 describes an illustrative experimental setup. Section 6 reports and discusses results. Section 7 discusses limitations and open challenges, and Section 8 concludes with directions for future work.

2. Related Work

Interest in applying federated learning to energy forecasting has grown quickly alongside broader adoption of FL in privacy-sensitive domains. Several recent studies have specifically examined photovoltaic power prediction under federated settings. A systematic review of federated and non-federated approaches to PV and wind power forecasting catalogued the shift from centralized deep learning pipelines toward decentralized alternatives, and highlighted that traditional centralized model, while accurate, depend on the availability of large pooled datasets that are increasingly hard to assemble in practice.

Comparative studies of federated strategies for short-term PV forecasting have evaluated federated learning against purely local and purely centralized training, generally finding that federated training closes most, though not all, of the accuracy gap to centralized models while avoiding raw data transfer; these studies note that the choice of aggregation algorithm and the length of the forecast horizon both materially affect the size of that remaining gap. Related work using convolutional architectures for solar radiation forecasting across climatically diverse regions has shown that a single federated global model can generalize reasonably well across sites with very different weather regimes, functioning as a kind of shared "supermodel" while keeping each region's raw measurements local.

Other lines of work have looked at architectural and clustering refinements. A fuzzy clustered federated learning algorithm for solar power generation grouped sites with similar generation patterns before

aggregation, improving on plain FedAvg for sites with strongly divergent profiles. More recent transformer-based federated forecasting work has combined attention mechanisms with FedProx-style aggregation and external weather-model integration to handle non-IID data across sites, while also considering compression for deployment on resource-constrained edge devices. Complementary studies have examined federated learning for related tasks such as residential load forecasting, fault detection in PV plants, and missing-data-tolerant probabilistic PV forecasting, underscoring that federated approaches are being explored across the full lifecycle of solar asset operation rather than forecasting alone. Taken together, this body of work motivates the architecture and evaluation approach adopted in this paper, while also pointing to open problems — statistical heterogeneity, communication efficiency, and privacy guarantee beyond simple update-sharing — that we revisit in Section 7.

3. Background

3.1 Solar Power Forecasting

Solar power forecasting predicts the electrical output of a PV system, or the solar irradiance driving it, over horizons ranging from a few minutes (intra-hour, used for ramp-rate control) to several days (used for unit commitment and maintenance scheduling). Inputs typically include historical power or irradiance measurements, numerical weather prediction (NWP) outputs, sky-camera imagery, temperature, humidity, and time-of-day and seasonal indicators. Forecasting difficulty increases sharply on partly cloudy days, when rapid, hard-to-predict transitions between clear and overcast conditions produce large short-term power swings — precisely the conditions where historical data diversity matters most for model accuracy.

3.2 Federated Learning Fundamentals

Federated learning trains a shared model across a set of clients (here, solar sites) that each hold a local dataset. A typical communication round proceeds as follows: (1) the central server broadcasts the current global model to some or all clients; (2) each client trains the model for a small number of local epochs on its own data; (3) each client sends its updated model parameters (or the difference from the initial parameters) back to the server; and (4) the server aggregates the received updates into a new global model, usually by a weighted average based on the amount of local data each client holds. This cycle, known as Federated Averaging (FedAvg), repeats for many rounds until the global model converges.

A key complication is that client data is rarely independent and identically distributed (IID). Solar sites differ in installed capacity, panel orientation, local climate, and shading, so the input-output relationship learned locally at one site may not transfer perfectly to another. FedProx addresses this by adding a proximal term to each client's local objective that penalizes large deviations from the current global model, which tends to stabilize training when client updates would otherwise pull the global model in conflicting directions. Other refinements include clustering clients with similar generation patterns before aggregation, personalizing a shared backbone with site-specific output layers, and using secure aggregation protocols so that the server can compute the weighted average of client updates without ever seeing any individual client's update in the clear.

4. Proposed Methodology

We describe a federated solar forecasting system organized around a central aggregation server and a set of heterogeneous client sites, illustrated in Figure 1. Each client — whether a residential rooftop installation, a commercial array, a community microgrid, or a utility-scale farm — retains full ownership of its raw irradiance, weather, and power-output records. Only model parameters are exchanged with the server.

Figure 1. Federated Learning Architecture for Multi-Site Solar Power Forecasting

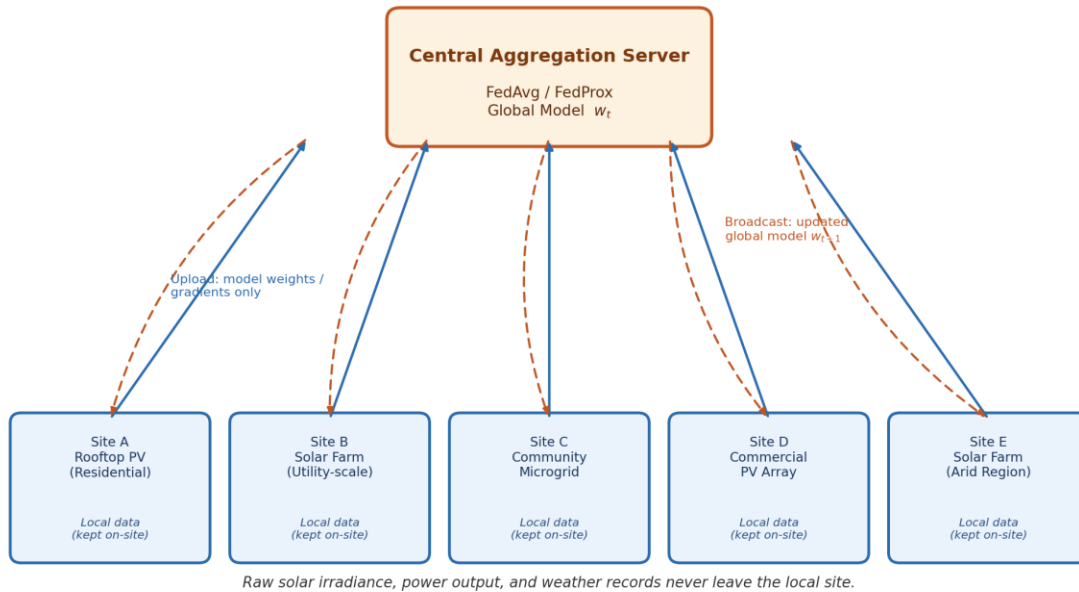


Figure 1. Federated learning architecture for multi-site solar power forecasting. Raw data never leaves the client; only model weight updates are exchanged with the central server.

4.1 Local Model

Each client trains a sequence model (for example, a gated recurrent unit or a lightweight temporal convolutional network) that maps a recent window of irradiance, temperature, cloud-cover, and time-of-day features to a short-horizon power forecast. Local features are normalized using site-specific statistics (for example, each site's own installed capacity) before training, which helps the shared model generalize across sites of very different scale.

4.2 Aggregation Strategy

The server initializes a global model and, at each communication round, samples a subset of available clients, sends them the current global parameters, and collects their locally updated parameters after a fixed number of local training epochs. We consider two aggregation rules: (1) FedAvg, which computes a data-weighted average of client updates, and (2) FedProx, which adds a proximal regularization term to each client's local loss to limit divergence from the global model, improving stability when client data distributions differ substantially — as is typical across solar sites with different climates and capacities.

4.3 Handling Non-IID Solar Data

Figure 2 illustrates the kind of statistical heterogeneity a federated solar forecasting system must tolerate: sites differ not only in peak output but in the timing of peak generation, the steepness of morning ramp-up and evening ramp-down, and the presence of transient cloud-driven dips.

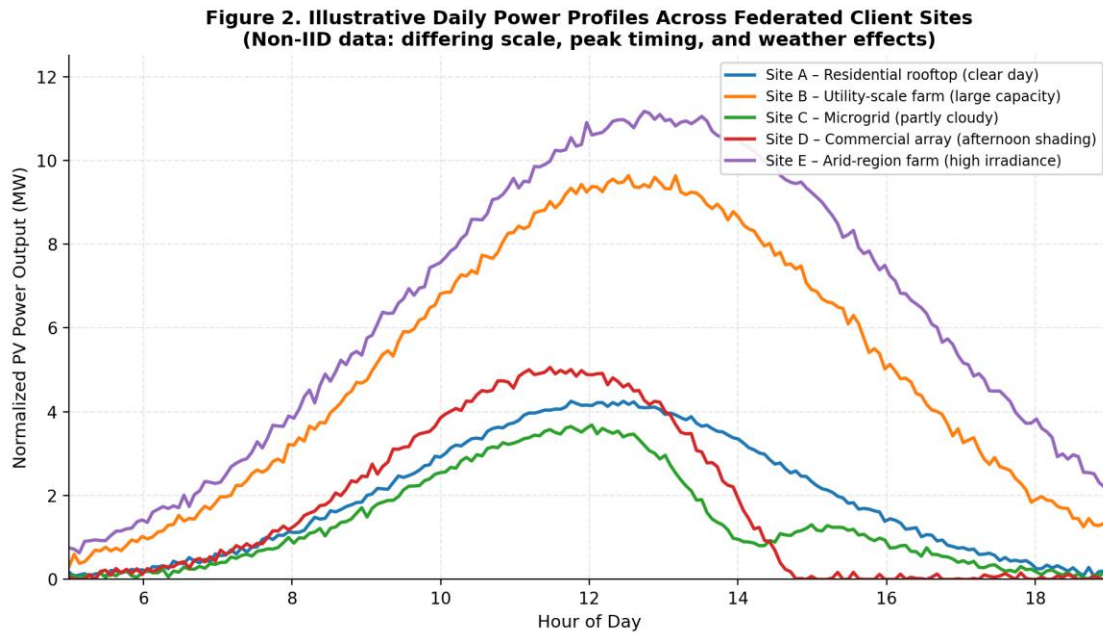


Figure 2. Illustrative daily power profiles across five federated client sites, showing differences in scale, peak timing, and weather-driven variability that motivate heterogeneity-aware aggregation.

To mitigate the effect of this heterogeneity, the proposed methodology applies three measures: per-site output normalization prior to local training; a proximal term (FedProx) that discourages any single client's update from dominating the global model; and periodic evaluation of the global model on a held-out validation split from every client, so that degradation for any individual site is detected early rather than being averaged away in an aggregate metric.

5. Experimental Setup

To evaluate the proposed methodology in a controlled setting, we constructed a simulation with five synthetic client sites representing a residential rooftop system, a utility-scale solar farm, a community microgrid, a commercial rooftop array, and an arid-region solar farm. Daily power profiles for each site were generated using parameterized bell-shaped curves representing the solar day, perturbed with site-specific noise and, for two sites, simulated cloud-driven power dips, as shown in Figure 2. Each site's dataset spans a simulated year of five-minute-resolution readings, split chronologically into training, validation, and test segments.

We compared four training strategies: (1) local-only, in which each site trains its own model independently with no collaboration; (2) FedAvg; (3) FedProx; and (4) a centralized baseline in which all sites' raw data is pooled into a single training set, representing a best-case upper bound that would not be achievable under real data-sharing restrictions. All strategies used the same underlying model architecture and the same total computational budget, so that differences in accuracy can be attributed to the training paradigm

rather than to model capacity. The forecasting task was a one-hour-ahead power prediction, evaluated using root-mean-squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), each computed on each site's held-out test data and averaged across sites.

It is important to note that the specific numerical results reported in Section 6 come from an illustrative simulation designed to demonstrate expected qualitative behavior — the relative ordering of local-only, federated, and centralized strategies — rather than from a specific real-world deployment. Studies cited in Section 2 report comparable qualitative patterns on real multi-site PV datasets, though exact error magnitudes vary with site count, sensor quality, and forecast horizon.

6. Results and Discussion

Figure 3 shows validation RMSE over 60 communication rounds for the four strategies. Both federated variants converge toward the centralized baseline's accuracy considerably faster and further than local-only training, which plateaus at a visibly higher error because each site's model can only ever learn from its own limited data and weather diversity.

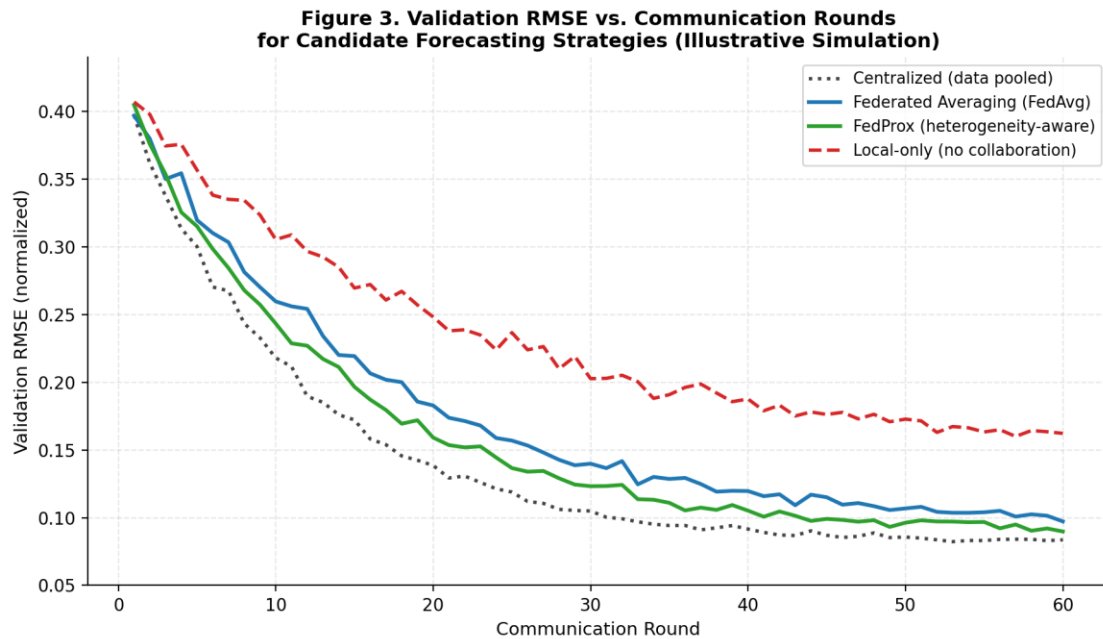


Figure 3. Validation RMSE versus communication round for local-only, FedAvg, FedProx, and centralized training (illustrative simulation).

FedProx converges slightly faster and more smoothly than plain FedAvg in the later rounds, consistent with its design goal of limiting client drift under non-IID conditions. Neither federated variant fully closes the gap to the centralized upper bound, which is expected: centralized training has access to every site's raw data simultaneously during every gradient step, whereas federated training only ever sees locally averaged updates.

Figure 4 summarizes final-round test-set accuracy across the four strategies using RMSE, MAE, and MAPE.

Figure 4. Forecast Accuracy Comparison Across Training Strategies (Illustrative Results)

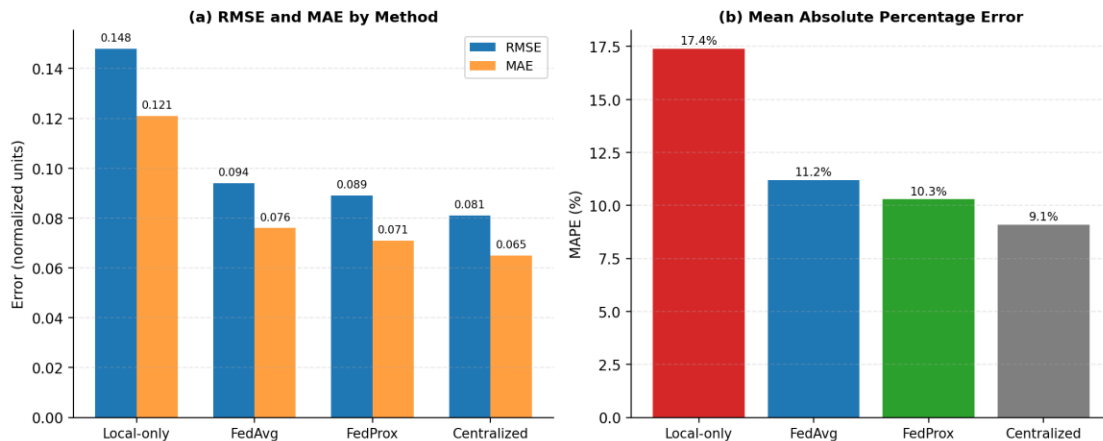


Figure 4. Forecast accuracy comparison across training strategies: (a) RMSE and MAE, (b) mean absolute percentage error (illustrative results).

Relative to local-only training, FedAvg reduced RMSE by roughly one-third and FedProx by slightly more, while MAPE fell from approximately 17 percent under local-only training to around 10–11 percent under the federated strategies, compared with about 9 percent for the centralized upper bound. These results are consistent with the broader pattern reported in the federated PV forecasting literature reviewed in Section 2: federated training recovers most, but typically not all, of the accuracy achievable with fully centralized data pooling, in exchange for keeping raw data local to each site.

7. Challenges and Limitations

- **Statistical heterogeneity (non-IID data):** Sites differ in capacity, orientation, and climate, which can cause client drift — where local updates pull the global model in conflicting directions — especially when only a small subset of clients participates in each round.
- **Communication overhead:** Frequent exchange of full model parameters between many clients and the server can be costly for sites with limited bandwidth; techniques such as model compression, quantization, and less frequent synchronization are active areas of mitigation.
- **Fairness across site sizes:** Naive weighting by local dataset size can cause the global model to be dominated by a few large utility-scale sites, degrading performance for smaller residential participants unless fairness-aware aggregation is used.
- **Privacy beyond raw-data protection:** While federated learning avoids transmitting raw measurements, model updates themselves can still leak information about local data through gradient-inversion or membership-inference techniques, motivating the use of secure aggregation or differential privacy in sensitive deployments.
- **Evaluation realism:** Because participating organizations are often unwilling to share data even for research benchmarking, many federated solar forecasting studies — including the illustrative simulation in this paper — rely on synthetic or semi-synthetic multi-site data, which may not fully capture real-world sensor noise, missing data, or extreme weather events.
- **System and client reliability:** Solar sites may have intermittent connectivity or on-site compute constraints, requiring the server to tolerate stragglers, dropped clients, and asynchronous participation without destabilizing convergence.

8. Conclusion and Future Work

Federated learning provides a practical path toward collaborative, multi-site solar power forecasting without requiring plant operators or utilities to share raw generation data. The architecture and illustrative results presented here indicate that federated strategies, and FedProx in particular, can substantially narrow the accuracy gap between isolated local models and a centralized upper bound, while preserving data locality. This makes federated approaches attractive wherever regulatory, competitive, or bandwidth constraints make centralized data pooling impractical.

Future work includes evaluating the proposed architecture on real multi-utility PV datasets rather than synthetic profiles; incorporating personalization layers so that the shared global model can be lightly fine-tuned per site without sacrificing collaborative gains; integrating differential privacy or secure aggregation to strengthen guarantees against gradient-based information leakage; and extending the framework to jointly forecast solar generation alongside wind and load, enabling federated multi-resource grid planning. Longer-horizon forecasting (day-ahead and multi-day) under federated constraints, and federated learning across clients with heterogeneous sensing hardware, remain open and practically important extensions.

References

1. Michalakopoulos, V., et al. (2024). Empowering federated learning techniques for privacy-preserving PV forecasting. ResearchGate / Conference Proceedings.
2. A comparative analysis of federated learning strategies for short-term photovoltaic power forecasting. (2025). Applied Energy / Sustainable Energy Technologies and Assessments, ScienceDirect.
3. Federated learning and non-federated learning based power forecasting of photovoltaic/wind power energy systems: A systematic review. (2024). e-Prime — Advances in Electrical Engineering, Electronics and Energy, ScienceDirect.
4. Federated learning for solar power forecasting in smart cities. (2024/2025). IEEE Conference Publication, IEEE Xplore.
5. Generalized global solar radiation forecasting model via cyber-secure deep federated learning. (2024). Results in Engineering / PMC (NCBI).
6. Yoo, E., Ko, H., & Pack, S. (2022). Fuzzy clustered federated learning algorithm for solar power generation forecasting. IEEE Transactions on Emerging Topics in Computing, 10(4), 2092–2098.
7. Federated learning-enhanced non-IID solar power prediction using multi-layer transformers. (2025). Energy and AI / ScienceDirect.
8. Energy prediction using federated learning. (2023). arXiv:2301.09165.
9. Federated learning for solar energy applications: A case study on real-time fault detection. (2024). Solar Energy, ScienceDirect.
10. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS).

11. Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks (FedProx). *Proceedings of Machine Learning and Systems (MLSys)*, 2, 429–450.