Power Quality Forecasting of Microgrids Using Adaptive Privacy-Preserving Machine Learning

Mazhar Ali, Ajit Kumar, and Bong Jun Choi(✉)

School of Computer Science and Engineering, Soongsil University, Seoul 06978, Republic of Korea
{mazhar,davidchoi}@soongsil.ac.kr, kumar@ssu.ac.kr

Abstract. Microgrids face challenges in monitoring and controlling the power quality (PQ) of integrated electrical systems to make timely decisions. Inverter-based technologies handle small-scale smart grids' power quality parameters (PQPs) and play an important role in condition monitoring. Accurate forecasting of such parameters is difficult due to the stochastic nature of demand, distributed generation, and weather conditions. Moreover, energy clients have concerns over growing privacy and security breaches for collaboration involving data exchanges. This study aims to predict PQPs indices of home microgrids using ANN, LSTM, and CNN-LSTM models. To preserve users' privacy, federated learning has been applied with some adaptive differential privacy on the global model and clients' data. Comparative analysis of the ML model and DP parameters shows that the LSTM model gives better results with adequate privacy parameters to predict the PQPs of five distributed microgrids. LSTM model gives the least MAE of 0.2323 for FL without privacy and 0.3256 test loss for appropriate DP level.

Keywords: Machine Learning · Microgrid · Federated Learning · Power Quality

1 Introduction

Integration of small renewable energy (RE) sources at the user end eases environmental degradation and climate change. Intermittent RE makes power grid stability less reliable, leading to cascading failure due to prolonged disturbances. The growing integration of distributed energy resources (DER), enormous electronic devices such as controllers, power management units (PMUs), relays, and household appliances deteriorate the power quality of modern intelligent grids [8]. Intelligent control and monitoring systems are vital for appropriate, timely decisions to protect sensitive equipment in PQ management activities. Electric appliance operations will be affected due to severe voltage deviation, frequency changes, power factor variations, transients, and current imbalances. Accurate prediction of the PQPs is an emerging problem in intelligent grid

M. Andreoni (Ed.): ACNS 2024 Workshops, LNCS 14586, pp. 235–245, 2024.
https://doi.org/10.1007/978-3-031-61486-6_14

dynamics and stable system operations. It can be helpful for better and quicker responses in case of PQ standards violations.

Microgrids (MG) require intelligent control systems for steady-state operation and monitoring in case of minor disturbances such as PQ parameter fluctuations. The general parameters involve voltage (U), frequency (f), total harmonic distortion of voltage (THD_u), and total harmonic distortion of current (THD_i) [5]. These parameters rapidly fluctuate with the power demand and supply imbalance. Such variance is a significant problem in modern microgrids with highly variable distributed solar and wind energy. Microgrids with long-lasting transient states can lead to the collapse of the whole distribution network. Thus, these parameters are directly or indirectly affected by renewable generations and load patterns, which are influenced by weather conditions. This study forecasts PQ parameters according to the weather patterns such as wind speed, solar irradiance, temperature, humidity, etc.

Previous studies focus on statistical and linear ML models to forecast the PQ parameters in centralized and local setups. In centralized learning, clients share the data with the server; thus, information leakage concerns from the clients. Similarly, in local learning, users face data scarcity issues that need to be improved for ML training. This study uses a time series regression model to predict the PQ parameters in a federated setting to preserve the privacy and data islands.

Moreover, a differential privacy (DP) approach is also adapted to address the issue of poisoning and model inference attacks. The literature needs to include the application of FL and adaptive federated DP in forecasting the PQ parameters of MG. This study opens the research toward distributed secure learning on the regression tasks of PQ forecasting and the tradeoff between model degradation and privacy. The contribution of this research study is summarized as follows:

- Comparatively analyze three data-driven models (ANN, LSTM, CNN-LSTM) as a PQP forecaster in a federated setup to address MG clients' privacy and data scarcity.
- Evaluate the federated ML models based on test loss and use the most appropriate forecasting approach to analyze the DP mechanism in the distributed setting of MG.
- Apply the adaptive differential privacy approach in a federated setup to secure the server and client models against poisoning and inference attacks. Also, compute the threshold of security that does not severely degrade the models during the training.

The paper continues with Sect. 2 as a literature review, which provides insights into the past related studies. Section 3 discusses the proposed method of the study. Section 4 analyzes the simulation setup, data processing, and results of the research work. Lastly, Sect. 5 concludes the study by highlighting key findings and gaps in the current study.

2 Literature Review

Power Quality remains a significant problem in microgrids, and it deteriorates further with multiple intelligent devices and highly variable local renewable generation. PQ parameter prediction is critical for early warning and preparedness in transient disturbances. I.S Jahan et al. [5] predicted five PQ indices, i.e., frequency, voltage, flicker, total harmonics distortion of current and voltage with decision tree and neural network approaches. DT was found to be a suitable model for the off-grid system experiment based on the test loss for six days. Jakub Kosmal and Stranislav Misak [7] analyzed PQ management of a decentralized microgrid predominantly with PV generation and active demand side management (ADSM). The three PQ parameters included ficker severity, frequency, and THD_u. The ADSM controls the consumption plan based on the predicted PQ parameters, which would lead to equipment damage outside limits. Similarly, Ibrahim Jahan et al. [6] carried out clustering approaches for the same data based upon several features like appliance (AC. heating, light, fridge, TV) states with weather variables (temperature, pressure, GHI, U.V., wind speed) to predict five PQ indices (U, PF, PL, THD_u, THD_i). Four forecasting models (DT, KNN, BGDT, BODT) were used for each cluster node and evaluated based on RMSE. All the models better forecast the power factor (PL) and load, while BODT gives the least RMSE for all the parameters except higher error of 6.736 for THD_i.

Federated learning is a new paradigm of machine learning where multiple clients collaborate to learn a global model without sharing their data with the central server. The computation is done at edge devices where client data resides. Thus, FL provides a better solution in cases of data scarcity, privacy, and security concerns. FL applications have been seen in intelligent grids for anomaly detection, energy trading, EV scheduling, NILM, and RE forecasting. Several FL studies have been conducted to accurately predict the demand and generation of different building setups and energy resources, such as solar and wind [1,3].V. Venkatesh et al. [9] analyzed the distributed energy forecasting using the BuildFL framework on IoT-based pecan street datasets. FL prediction gives similar load patterns when compared with GridLAB-D generated consumption profile. Similarly, Zhang and Wang [10] performed distributed aggregation of sub-parameters of the probabilistic wind forecasting model. The ADMM algorithm decomposes the problem into sub-parts and evaluates the probabilistic regression models of 10 wind farms based on the quantile score. ADMM and mirror-descent algorithms have been studied in distributed setup for measuring the PQ variables, and the literature still needs to include FL [4].

The probabilistic ADMM approach, in a distributed setup, concatenates the cost function into sub-problems in which clients share their information. Such a technique has limitations over non-convex models and lacks privacy guarantees in collaborative learning. Literature has thorough FL studies on energy demand prediction and renewable generation forecasting. Data-driven ML approaches have provided reasonable solutions in smart grids, and implementing FL is more straightforward than traditional probabilistic methods. However, research on

minute time series PQP prediction needs for federated and centralized ML. Thus, current research aims to analyze the application of FL and DP in power quality forecasting in distributed microgrid networks. The study compares non-linear ML models in privacy-preserving distributed learning to address the privacy and security issue in microgrid PQ parameter predictions, which is lacking in the literature.

3 Methodology

Modern power systems aim to be more resilient towards energy security, climate change, cyber-physical attacks, power disturbances, and cascading outages. DERs at the consumer end increase smart grid resiliency, energy, and cyber security by providing energy in case of catastrophic power outages and disturbances. Moreover, flexible energy markets encourage prosumers in cost-effective demand response (DR) tasks through home energy monitoring and control systems. Such intelligent home energy systems make incremental usage of power electronics and IoT devices for energy conversion, storage, monitoring, and control of power quality variables at the user end. The smart home system collects sensitive data from these devices and electrical appliances. We applied distributed ML and differential privacy in a federated setup to preserve the privacy and security of home microgrids. The methodology of the study is provided in detail below (Fig. 1).

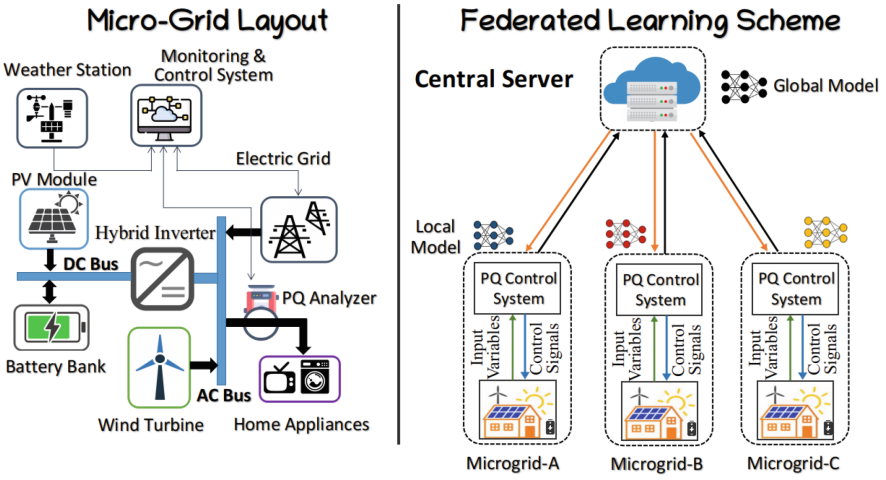

Fig. 1. (Left) Home MG system with DERs, Storage, Inverters, Appliances, and Control System. (Right) Privacy-preserving federated setup for Home MG clients for collaborative ML training without sharing data.

3.1 Microgrid System

The Microgrid concept has been practiced for decades at the distribution, community, and user levels. Prosumers with integrated RE, PEV, and battery storage made themselves small-scale microgrids that involved different tasks, like energy trading, demand response, load management, and protection schemes via monitoring and control systems [2]. Our study considers five home microgrids with PV, a small wind turbine, and a battery bank. The system model has been used in a home hybrid system test platform by Smart Grid Lab of VSB Technical University, Czech Republic [6]. Two buses are connected to two 2 kW PV modules and four 115 Ah lead-acid batteries with respective inverters. The voltage across the DC bus varies from 40.5–64 V (V) due to variable charging and discharging. A 240 V and 50 Hz frequency AC bus is connected with a load, grid, wind turbine, and hybrid inverter responsible for converting DC supply to AC for end usage. The load consists of several electrical appliances used for daily household activities, producing high noise in the AC system due to the appliances' inductive, capacitive, and resistive nature. An energy management system has been used to monitor and control the microgrid operation, which has several input signals from the weather station, electric grid PQ analyzer, etc.

3.2 Data-Driven Model

Prediction has been carried out by analyzing linear and non-linear ML approaches. Comparative analysis on ANN, LSTM, and CNN-LSTM hybrid models has been conducted to predict the PQ variables. ANN models are relatively simple to implement as they better approximate any continuous function but can be problematic for data scarcity and temporal dependencies. LSTM better captures the temporal features but has a complex model and lacks spatial feature extraction. CNN-LSTM is a hybrid model in which the CNN layer extracts spatial features, and LSTM layers handle the time-series patterns. Comparing these three models gives a better understanding of the relationship between single and hybrid ML models for recurrent tasks.

3.3 Privacy Preserving Method

Clients have concerns about data leakage, which can lead to misuse of personal information, malfunction of devices, and potential attacks on microgrids to disturb the whole power system. Federated learning, which can better preserve users' privacy, has been used in the study. In FL, the data reside on clients, and models are trained on edge devices; thus, no information has been shared with a central server. As the goal of FL is to learn a general global model, there is a threat of poison and model inversion attacks. Differential privacy adds noise in the client model weights to protect the user information from a poison attack. Similarly, noise is added to the server model weights to protect the global model from inversion attacks by malicious clients. However, if the noise or security is high, the accuracy declines, and the prediction task will be affected. So, we evaluated different privacy parameters using an adaptive approach.

Algorithm 1. Pseudo Code of Proposed Method

Initialize: Model (M^0), Clients (K), NoiseValue (N), Batch($d \in D$)
for $t = 1$ to T client($i \in K$) **do**:
  Client Updates: Δ_i^t, $b_i^t \leftarrow$ **FedAVG**(i, M^t, S^t)
  Server Updates:
    $\bar{\Delta}^t = Agg.(D - i^t) + Noise(N)$
    $M^{t+1} = M^t + n_s\bar{\Delta}^t$
    $S^{t+1} = \text{Adaptive}(S^t)$
end for
LocalUpdate: FedAVG(i, M^t, S^t)
  $M \leftarrow M_0, \quad M \leftarrow SGD(M, n_l, d)$
  $\Delta \leftarrow M - M^0, \quad b \leftarrow ClippingNorm(\|\Delta\| \leq S)$
  $\Delta' \leftarrow \Delta \cdot \min(1, S/b)$
return (Δ', b)

Federated Differential Privacy. In federated learning, the model poses a threat from malicious actors to manipulate the raw local information. Encryption schemes present viable protective measures but pose a possibility of cryptographic breach and incur high computational costs. A nascent and promising alternative comes from DP, which offers privacy guarantees during training. FL process starts with the server initializing the forecasting model (M) to the clients (K). Each client ($i \in K$) locally updates the global model (M) on their private data (D_i) and sends it back to the server with noise bit b_i. The server aggregates the client update at each round with the additional noise under the FedAVG and DP mechanism, as shown in Algorithm 1. Any randomized learning algorithm satisfies $(\epsilon, \delta)-$DP for any adjacent input data d and d', by adding noise function as given.

$$M(d) = f(d) + Noise(S_f)$$

where, S_f is the maximum l_2-distance norm $\|f(d) - f(d')\|_2$ and ϵ is the privacy loss parameter with the failure probability $\delta \in [0, 1]$. In the above equation, the *M(d)* can achieve $(\epsilon, \delta)-$DP privacy by adding Gaussian noise $N(0, S_f^2\sigma^2)$ with $\epsilon \leq 1$ and $\delta \geq 0.8 \cdot exp(-(\sigma\epsilon)^2/2)$ in the function *f(d)*. Here, σ is the noise multiplier that controls the trade-off between privacy and model degradation during the federated training process.

Adaptive Clipping. To ensure better privacy, the FedAVG algorithm made two levels of DP mechanism in a federated setup. In the client updates, local model parameters must be clipped before sending to the server, while the server adds enough noise to the aggregated weights. These measures provide enough security for poisoning and inference attacks from the malicious adversary. It has been seen in past studies that a clipping norm with too small a value will slow the model converge process, while a larger value adds too much noise, which degrades the model performance. Thus adaptive clipping approach

$S \leftarrow S \cdot exp(-n_c S(\bar{b} - \gamma)$, which start will low value S^0 and gradually increase with the learning rate n_c(=0.2) to the target quantile γ(=0.5).

4 Simulation and Results

4.1 Dataset

The dataset used for our work is obtained from experimental results of a simulated test bed environment by Smart Grid Lab in the Czech Republic. It consists of several temporal and spatial features, as shown in Table 1. The dataset consists of every 5-min reading of the respective variables for the June and July months of 2019. The power quality parameters have been collected using a PQ analyzer at the AC bus connected to the household load under EN 50160 and EN 61000-2-20 European standards. Minute-wise power load consists of different household appliances such as TV, boiler, kettle, fridge, microwave, lights, etc. These are inductive, capacitive, and resistive loads, thus fluctuating the minute variation. Similar time series weather datasets have been collected from the periphery of the study site in Ostrava, Czech Republic.

Power quality and meteorological datasets used in the study have been collected via a test-bed of a home hybrid system by Smart Grid Lab of VSB Technical University, Czech Republic [5]. The PQ analyzer collected the power load, voltage, frequency, power factor, THD_u, and THD_i parameters from the AC bus. The dataset contains 5-min intervals of input (GHI, WS, Pressure, Temperature, and PL) and output parameters (frequency, voltage, THD_u and THD_i) for two months.

Table 1. Input and Output Parameters used in Power Quality Analysis

Symbol	Description	Range
Weather Parameters		
GHI	Global Horizon Irradiance (W/m^2)	0–1033
W_s	Wind Speed (m/s)	0–5.7
P	Atmospheric Pressure (hPa)	976.4–995.3
T	Atmospheric Temperature (°C)	9.2–32.2
Power Quality Parameters		
PL	Power Load (kW)	0.6–2.61
f	Frequency (Hz)	49.9–50.08
U	Voltage (V)	223.96–245.64
THD_u	THD of Voltage (%)	0.51–5.75
THD_i	THD of Current (%)	4.48–61.68

4.2 Data Preprocessing

The multivariate time series study has several features that are used to predict the desired output variable. In our research, we aim to forecast four parameters that have directly or indirectly influenced the input features and the output variable. The principal component analysis (PCA) approach is used to analyze the correlation among all the input and output parameters. Based on the feature correlation matrix, the respective parameters have been dropped before the training process. Similarly, as time series forecasting depends upon its past trends, a lookback is also given as an input feature. Data normalization is crucial in training the machine learning model to access the optimized weights and connections between neurons. Thus, we normalized the data during preprocessing for flexible training, a robust model, and better prediction results.

4.3 Experiment

Tensorflow federated (TFF) framework is used to perform simulations in a federated setup. The dataset has been used for centralized machine learning, and to address the federated setup, we divide the data among five client modules. It is assumed to be a cross-silo setup, which means the amount of data is the same for each client, but the tabular data are highly different in temporal nature. LSTM and CNN-LSTM models have one dense layer with respective LSTM and CNN/LSTM layers, while only two layers are used for the ANN model. The number of neurons for these layers has been kept the same, i.e., $n = 50$. SGD optimizer has been used for the federated experiment. The MAE metric has been used throughout the simulations to evaluate the model. The model has been trained for global round $R = 500$ and evaluated on the test datasets. Table 2 gives the details about the hyperparameters of ML models used in the experiment of federated learning. Similarly, in the differential privacy, several noise values have been added in global model weights and client model weights to secure the model from malicious attacks and privacy leakage.

Table 2. Details of hyperparameters used in the implementation of ML models.

Hyperparameters	Search Space	Value
No. of Neurons	10, 20, 30, 50	50
Activation Function	ReLU, Tanh	Relu
Server Learning Rate	1.0, 0.10, 0.01	0.1
Batch Size	40, 60, 80, 100	60
Client Learning Rate	0.2, 0.02, 0.002	0.02
Client Epochs	5, 10, 20, 30	10
No. of Global Rounds	200, 300, 500, 750	500

4.4 Results and Discussion

The privacy-preserving FL approach has been analyzed to evaluate the three ML models based on the MAE loss, as shown in Table 3. We only considered the distributed setup of microgrids. We did not analyze the local learning as our main aim is to evaluate the better model in FL due to privacy constraints. The three trained models have been assessed on individual test datasets of microgrids after selecting appropriate hyperparameters during the training stage. After extensive experiments, it has been shown that MG_4 gives a better result for all the ML models used in the experiment. Still, the LSTM model has the lowest average MAE value, i.e., 0.3467. MAE value for MG_3 is 0.2323, depicting that the LSTM model is better learned on client 3. As the LSTM model gives better results than ANN and hybrid models, it is used to analyze the impact of the differential privacy approach in the federated setup. Similarly, the differential privacy insights on clients lead to better secure training results.

Table 3. MAE loss of the ANN, CNN-LSTM and LSTM models on test datasets.

Clients	ANN	CNN-LSTM	LSTM
MG_1	0.4293	0.3296	0.2378
MG_2	0.4148	0.3709	0.3918
MG_3	0.2562	0.2617	**0.2323**
MG_4	**0.2335**	**0.2461**	0.2383
MG_5	0.8538	0.9798	0.6333
Average	0.4355	0.4376	**0.3467**

Different DP parameters have been given for the federated training to learn in a secure environment. During the training, a noise ratio is added to clients and server models to secure them from poison and model inference attacks. Different noise multiplier values [0.0, 0.25, 0.5, 0.65, 0.75, 1.0] have been given to find the tolerance range of the LSTM model from degradation. Extensive simulations have been carried out to determine the optimum noise parameter through the search space approach. Figure 2 shows these noise parameters' results for a training round of 500 rounds. It depicts that the LSTM model can tolerate a noise value of up to 0.5 without degrading model quality. A noise value of 0.65 slightly deviates the model from the optimum, while a higher value of 0.75 and 1.0 significantly degrades the model. That's why we stopped the training for noise = 1.0 after rounds = 200. Though the noise secures the user's privacy during the training process, there is a trade-off of accurate prediction, which is crucial in power quality parameter forecasting.

The noise values that mimic the global model without DP have been evaluated on the test datasets, as shown in Table 3. It shows the tradeoff between privacy and model precision, as the LSTM model with lesser noise values has

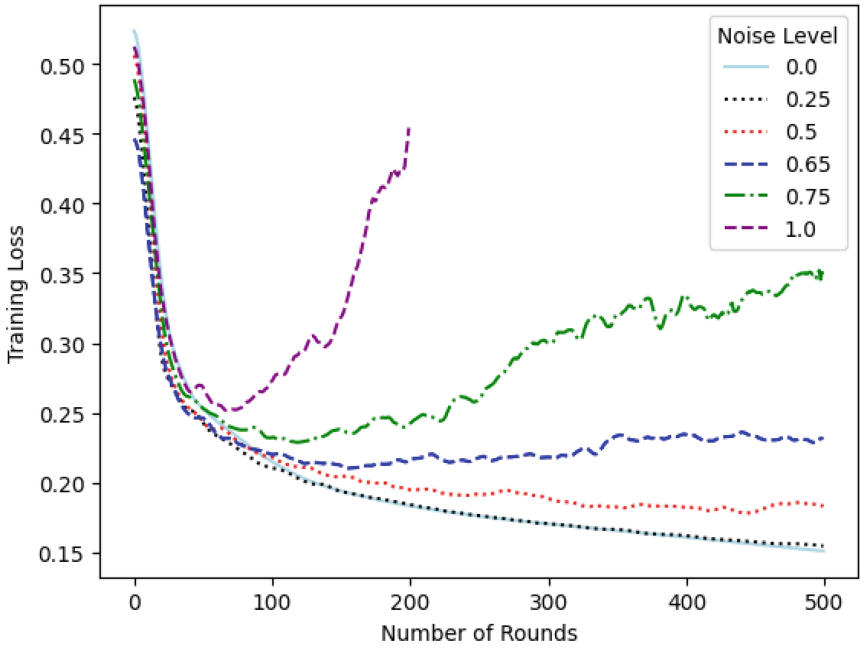

Fig. 2. MAE of LSTM model during the FL training with different level of DP Noise.

the least MAE loss of 0.3809 average. The study addresses the privacy issue in distributed learning through an adaptive DP mechanism. It can be the baseline for future studies that tackle secure aggregation techniques in intelligent grid PQ forecasting applications. However, to address the data heterogeneity and accurate model adaptation, personalized federated learning will give better analysis in the future. Future studies must incline towards a statistical approach or tolerance factor to mitigate the impact of the clipping approach and change the time series forecasting to an anomaly or error detection problem (Table 4).

Table 4. MAE loss of the LSTM model on different noise levels on clients test data.

	MG_1	MG_2	MG_3	MG_4	MG_5	Avg
Noise = 0.25	0.2595	0.4201	0.2941	0.2539	0.6767	0.3809
Noise = 0.5	0.3256	0.4614	0.3701	0.3202	0.7291	0.4332
Noise = 0.65	0.5128	0.5417	0.4989	0.4156	0.9528	0.5844

5 Conclusion

Distributed energy generation made microgrids more intelligent with the proliferation of power electronic devices and monitoring systems. Intelligent ML

operation of green microgrids faces privacy and security issues due to sharing power quality parameters. Federated learning is a suitable approach to learning the patterns of predictive ML models to preserve privacy via edge training. Comparative analysis on ANN, LSTM, and CNN-LSTM models evaluate better prediction models in distributed settings. The LSTM model has the least MAE test loss of 0.2323, making it most appropriate for federated predictive learning. FL faces the challenge of potential model inversion and poison attacks at the server and client end. Thus, the study provides an adaptive differential privacy technique to secure the microgrid in such an FL setup. The results showed that a privacy parameter of 0.5 value gave a better solution to secure the server and home microgrid clients. In future studies, we aim to analyze the personalized FL approach with DP under IEEE standards for hybrid energy systems.

Acknowledgment. This research was supported by the MSIT Korea under the NRF Korea (NRF-2022R1A2C4001270) and the Information Technology Research Center (ITRC) support program (IITP-2022-2020-0-01602) supervised by the IITP.

References

1. Ali, M., Singh, A.K., Kumar, A., Ali, S.S., Choi, B.J.: Comparative analysis of data-driven algorithms for building energy planning via federated learning. Energies **16**(18), 6517 (2023)
2. Ali, M., et al.: Techno-economic assessment and sustainability impact of hybrid energy systems in Gilgit-Baltistan, Pakistan. Energy Rep. **7**, 2546–2562 (2021)
3. Cheng, X., Li, C., Liu, X.: A review of federated learning in energy systems. In: 2022 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia), pp. 2089–2095 (2022)
4. Gholizadeh, N., Musilek, P.: Distributed learning applications in power systems: a review of methods, gaps, and challenges. Energies **14**(12), 3654 (2021)
5. Jahan, I., Misak, S., Snasel, V.: Power quality parameters analysis in off-grid platform. In: Kovalev, S., Tarassov, V., Snasel, V., Sukhanov, A. (eds.) IITI 2021. LNNS, vol. 330, pp. 431–439. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-87178-9_43
6. Jahan, I.S., Blazek, V., Misak, S., Snasel, V., Prokop, L.: Forecasting of power quality parameters based on meteorological data in small-scale household off-grid systems. Energies **15**(14), 5251 (2022)
7. Kosmák, J., Mišák, S.: Power quality management in an off-gtrid system. In: 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), pp. 1–5. IEEE (2018)
8. Luo, A., Xu, Q., Ma, F., Chen, Y.: Overview of power quality analysis and control technology for the smart grid. J. Mod. Power Syst. Clean Energy **4**(1), 1–9 (2016)
9. Venkataramanan, V., Kaza, S., Annaswamy, A.M.: Der forecast using privacy-preserving federated learning. IEEE Internet Things J. **10**(3), 2046–2055 (2022)
10. Zhang, Y., Wang, J.: A distributed approach for wind power probabilistic forecasting considering spatio-temporal correlation without direct access to off-site information. IEEE Trans. Power Syst. **33**(5), 5714–5726 (2018)