



Article Comparative Analysis of Data-Driven Algorithms for Building Energy Planning via Federated Learning

Mazhar Ali 🔍, Ankit Kumar Singh 🔍, Ajit Kumar 🔍, Syed Saqib Ali 🛡 and Bong Jun Choi * 🛡

School of Computer Science and Engineering, Soongsil University, Seoul 06978, Republic of Korea; mazhar@soongsil.ac.kr (M.A.); ankit@soongsil.ac.kr (A.K.S.); kumar@ssu.ac.kr (A.K.); kazmi@soongsil.ac.kr (S.S.A.)

* Correspondence: davidchoi@soongsil.ac.kr

Abstract: Building energy planning is a challenging task in the current mounting climate change scenario because the sector accounts for a reasonable percentage of global end-use energy consumption, with a one-fifth share of global carbon emissions. Energy planners rely on physical model-based prediction tools to conserve energy and make decisions towards decreasing energy consumption. For precise forecasting, such a model requires the collection of an enormous number of input variables, which is time-consuming because not all the parameters are easily available. Utilities are reluctant to share retrievable consumer information because of growing concerns regarding data leakage and competitive energy markets. Federated learning (FL) provides an effective solution by providing privacy preserving distributed training to relieve the computational burden and security concerns associated with centralized vanilla learning. Therefore, we aimed to comparatively analyze the effectiveness of several data-driven prediction algorithms for learning patterns from data-efficient buildings to predict the hourly consumption of the building sector in centralized and FL setups. The results provided comparable insights for predicting building energy consumption in a distributed setup and for generalizing to diverse clients. Moreover, such research can benefit energy designers by allowing them to use appropriate algorithms via transfer learning on data of similar features and to learn personalized models in meta-learning approaches.

Keywords: federated learning; building energy management; load forecasting; data-driven algorithm

1. Introduction

Climate change is a critical issue today, with an unprecedented impact on daily life. The Earth is warming beyond 1.5 °C, and energy demand is increasing [1]. Human activities tend towards indoor comfort zones, increasing the share of the building energy sector by up to 36% of global energy consumption [2]. Furthermore, this figure might increase with growing urbanization across developing countries, and energy-efficient consumption might help mitigate global carbon emissions. A policy of net-zero carbon emissions was set for newly-constructed buildings by 2030 in the Paris Agreement [3]. In a net-zero carbon building, the prosumer not only participates in providing clean resources but also earns financial savings. Building energy prediction at the early stage of construction with insufficient data is the most challenging task, as overestimation might lead to energy and monetary losses, as well as underestimation trade-offs between savings and comfort levels. Energy planners are shifting their focus towards robust models that can predict demand patterns more precisely via intelligent building automation to reduce household energy waste [4].

Currently, data-driven approaches provide exceptional outcomes in energy-demand prediction that do not require prior information, as in physical systems. A predictive mathematical model requires adequate time-series data and the respective features for concise energy predictions. Single-model machine learning (ML) and ensemble learning



Citation: 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* **2023**, *16*, 6517. https://doi.org/10.3390/ en16186517

Academic Editor: Satoru Okamoto

Received: 30 June 2023 Revised: 4 September 2023 Accepted: 7 September 2023 Published: 10 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). have been widely used to predict residential loads [5]. Precise forecasting could reduce energy consumption by up to 30 percent, and the accuracy of ML approaches depends on the selection of the model and the quantity and quality of data [6]. Utilities and third parties collect large amounts of consumer load data, store them in a central system, and make the necessary predictions. A centralized system must ensure the availability of large datasets and require multiple storage spaces. Electricity distribution companies in energy trading markets require accurate real-time predictions to balance variable demand and generation, and to achieve optimum financial gains. Similarly, energy consumers are concerned about data leakage of personal information and disclosure of incoming and outgoing schedules. The European Union formulated the General Data Protection Regulation (GDPR) policy to ensure user data privacy and collection without consent. Thus, utilities and clients are reluctant to share their data due to privacy breaches and potentially malicious attacks.

Federated learning (FL) provides a promising solution for client data privacy issues, as suggested by McMahan et al. [7]. It provides basic privacy preserving techniques, easing the computational burden and memory storage of the central server by distributing training to edge devices. The client data are not shared with the central server, and the FL platform provides training for the ML global model at the user end. The parameter server is responsible only for initializing and aggregating the model weights during training. Recently, data-driven algorithm-based approaches have been studied to analyze building energy predictions in centralized and federated settings. Several energy-forecasting studies have suggested that ML models could help mitigate greenhouse gas emissions and energy and financial waste. The accuracy of ML models depends on three important factors: model selection, quantity, and data [8]. Similarly, data collection and refinement could be more efficient for large-scale building forecasting. FL addresses these challenges by training multiple clients to preserve the inherent data privacy. There is a need to explore effective data-driven algorithms for forecasting loads based on the features of buildings in a federated setup.

This study aims to compare data-driven algorithms in FL using transfer and metalearning methods to predict building energy consumption in cross-silo settings. In cross-silo FL, clients have adequate data silos, and few clients collaborate to optimize a common task of interest. Financial and health care institutions have sufficient private data to mutually train specific tasks with fewer participants to preserve privacy. Similarly, learning patterns during the training process could be used to transfer knowledge to a new client in the same domain with less data using a transfer learning approach. Buildings have specific consumption patterns; thus, federated meta-learning provides a better opportunity to learn the personalized energy patterns of clients. This study analyzes different data-driven ML algorithms that show comparable results in centralized learning (CL), federated transfer, and federated meta-learning. The testing accuracy is better than that of the training in all settings, and the client's loss shows that the global model is generalized for adoption. Federated transfer and meta-learning approaches provide insights comparable to CL, while providing basic privacy and personalization to client local predictions. The rest of this paper is organized as follows. Section 2 provides a literature review of the tools and ML algorithms proposed for building energy planning. Section 3 presents the data-driven and FL algorithms used in the comparative analysis. Sections 4 and 5 presents the experimental results and a discussion, respectively. Finally, Section 5 concludes the study.

2. Literature Review

Conventional buildings use lighting, heating, cooling, and other auxiliary electrical equipment. Smart homes introduce an automated system with smart meters, monitoring appliances, and Internet of Things (IoT) devices coupled with TVs, cameras, intelligent lighting, and HVAC systems (heating, ventilation, and air conditioning) in home energy management networks [9–11]. Energy devices generate sub-hourly or hourly data that cloud-based central utilities could access. Physical-based modeling approaches have been used for decades to predict the loads of typical buildings. These models consider several

input factors, such as the building structure, thermal properties, orientation, insulation, HVAC system, internal occupancy load, and weather conditions [12]. Simulation tools such as EnergyPlus (https://energyplus.net/ (accessed on 23 June 2023)), DOE-2 (https://www.doe2.com/ (accessed on 23 June 2023)), and TRNSYS (https://www.trnsys.com/ (accessed on 23 June 2023)) have been used to model the performance of building HVAC systems and predict energy consumption under different scenarios. Such applications require sufficient information as input parameters to predict the energy demand as the output of the forward model. Physical models are unreliable for the general energy prediction of buildings because the required data are scarce and specific parameter collection takes time [13].

Machine learning algorithms are giving promising results on massive hourly loads, and privacy preserving federated learning could be applied to predict the energy consumption of buildings with lesser operational data. Petrangeli et al. [14] analyzed the trade-off between privacy and the accuracy of residential load prediction on an IoT-based dataset. Performance evaluation on several FL configurations with varying household numbers has been carried out with the LSTM model using the Flower framework. The result shows that there must be trade-off accuracy to achieve higher privacy in an FL setup. Shi et al. [15] recommend transfer learning in training an adaptive CNN/LSTM model on different households of UK power networks. A generalized trained model can be helpful for newly constructed buildings with limited non-IID data. A problem arises in such generalized architecture as multiple household users may have a similar pattern of load scheduling, but what if a unique user interacts and differs from the general energy usage pattern? Sater et al. [16] introduced anomalies in the energy usage of three smart buildings using IoT sensor device data. In CL, there is a delay in detecting the anomaly due to the large amount of data that can affect the pattern. However, FL gives better results in classifying the abnormalities and regressing the load pattern due to locality.

Demand response enabled consumer participation in reducing the overall energy demand by restricting the unnecessary usage of auxiliary electric appliances during peak hours. Dynamic pricing in energy markets has attracted consumer attention to incentives through the breakdown of electric equipment usage during off-peak hours. Service providers require the collection of appliance data from consumers to provide better recommendations regarding usage patterns that pose privacy and mislabeled data issues [17]. Gupta et al. [18] presented a FedAR+ algorithm to handle the noisy load data of appliances connected to wrong plugins and achieved an accuracy of 90% for 30% of mislabeled load data while preserving the privacy of three houses in an FL setting. Similarly, Gao et al. [19] presented a cloud-free, decentralized training framework for IoT-based home appliances connected to an intelligent home agent. This study achieved high accuracy in serverless FL networks to reduce the financial costs of servers in a privacy preserving manner for residence details. Gholizadeh et al. [20] proposed a clustering algorithm to analyze the performance of local and central models. Clusters were formed based on similar house attributes, and the convergence rate in privacy preserving FL settings was reduced.

Renewable energy integration at the prosumer end enabled the bidirectional flow of electricity, and distributed energy providers played a crucial role in energy trading, demandside management, load shifting, and infrastructure development. Husnoo et al. [21] proposed an FL architecture as a FedREP for retail energy providers to address the scalability issue of a centralized system through a privacy preserving distributed network. It showed compromising results with a mean square error (MSE) of 0.3 to 0.4, comparable to the centralized system with the advantages of a possible network extension and preserving the privacy of connected households. Taik et al. [22] analyzed short-term load forecasting using edge-computing FL by randomly selecting a subset of clients from a pool of 100 homes and comparing the results based on the mean average percentage error (MAPE). Similarly, Briggs et al. [23] evaluated the effect of weather features on household load forecasting model performance and the computational efficiency of different FL and CL settings. A cluster-based federated approach on load pattern has been used to evaluate the transfer learning in order to address the data shortage and privacy concerns of traditional ML [24,25]. To cluster the clients, it is essential to share information with the central server, leading to some level of trade-off on privacy and authenticity. The transfer of a trained model is only carried out within clusters, making it non-viable for clients with the same weather and other building characteristics. Moreover, the federated transfer learning approach with sensitivity analysis on data availability has been carried out in a cross-silo setup of cold climate zone offices [26]. The study analyzed the feature selection procedure and fine-tuned the hyper-parameters of the ANN model before the training process. The study demonstrated that the performance is highly variable for the testing client with operational data availability of less then 30%, making the global model not suitable for clients with limited data.

Similarly, several deep learning models were evaluated on different metrics (MAE, MSE, etc.) using an encryption cryptosystem data scheme on the real power consumption of Ausgrid [27]. The study was more focused on data privacy and did not analyze the data scarcity and personalization problem in the federated setup. In the above federated studies, the non-participating clients required few rounds to transfer the knowledge of the train model to fine-tune the client personalized layers. Thus, the data deficiency issue could be more significantly addressed in meta-learning, which evaluates the test clients on the first epoch. This means that the least amount of data can be used to evaluate the fine-tune global model. This work aims to narrow the research gap by providing a comparative analysis of the ML-based time series forecasting algorithms in cross-silo FL.

3. Comparative Analysis

Concise energy prediction is crucial for optimizing energy consumption, improving energy efficiency, and reducing environmental impact. Physical-based models require high computation and are time-consuming because they rely on detailed building characteristics, parameters, and assumptions on unavailable features, thus affecting the accuracy, leading to them being unsuitable for real-time scenarios. ML algorithms can capture the complex relationships between input variables and energy-consumption patterns from large datasets to identify nonlinearities that physical-based models might overlook. Several studies comparatively analyzed single and hybrid ML predictive algorithms on real-time energy data, depending on geographical location, weather parameters, and model hyper-parameters [28,29]. Similar works were also carried out in solar and wind energy-based forecasting on spatio–temporal variables [30,31].

Recurrent-based neural networks capture the long-range temporal dependencies as historical contexts, which is important in building energy prediction. However, apart from time-series dependency, the building physical and geographical variables also influence the consumption rate, making CNN viable to capture spatial features. Thus, a hybrid ML model better highlights the relationships between sequential, temporal, and spatial attributes for hourly energy forecasting. Our approach analyzed ML prediction with energy data scattered across clients and did not need to pile at a central system, such as distribution utilities or energy planners.

Transfer and meta-learning approaches have been applied in a cross-silo FL setting for large buildings to capture the relationship between predicted energy and features, such as previous-hour consumed energy, temperature, humidity, and building size. The model was fine-tuned and tested on a new building that was not included in the training process. Similarly, in practice, most households only have smart meter data. Thus, the hourly real-time energy consumption, temperature, and humidity parameters of the past few months were used to predict the next hour/sub-hour values using the window sliding approach. This provided a better analysis as features become the energy consumption pattern over the window size of the past hours, which could be of different sizes (e.g., 24 h, 12 h, or 6 h), depending on the short-, medium-, or long-term prediction interest. Such a method makes it easy to collect and depend on the other detailed parameters used in a physical model, as only smart meter data are used, both as samples and labels.

3.1. Data Driven Algorithms

In the last decade, ML algorithms have been used to predict time-series energy demand at the household level. Here, we discuss six basic approaches that are considered adequate for recursive time-series analysis of energy demand prediction. Using these ML algorithms has helped to analyze and support the claim of the effectiveness of data-driven approaches over physical-based methods.

3.1.1. ANN

An artificial neural network (ANN) is an ML algorithm inspired by the function of the human brain in an information processing system that attempts to mimic the communication between biological neurons [32]. It comprises a network of interconnected nodes with three primary layers that share weights based on the input data from the respective layer. ANNs have been widely used in the energy demand prediction of buildings because of their ability to generate accurate patterns based on historical input variables.

3.1.2. LSTM

LSTM is a popular recurrent neural network (RNN) variant used to process sequential data, such as time-series or natural language. LSTMs are designed to overcome the limitations of traditional RNNs, which are prone to the vanishing gradient problem [33]. This problem occurs when the gradients of the error for the weights become extremely small, making it difficult for the network to learn from the data.

3.1.3. CNN

A CNN is a type of deep learning (DL) algorithm inspired by the structure and function of the visual cortex in the human brain, which is responsible for the visual processing of information and is particularly effective in image and video processing [34]. A CNN comprises multiple layers, including convolutional, pooling, and fully-connected layers. CNNs can be useful because of their ability to learn spatial information and identify patterns in data. The convolutional layers of a CNN can learn to detect patterns in weather data, such as temperature or wind speed patterns, which could be used to make predictions regarding energy generation or consumption. In addition, CNNs can be trained with a large amount of data, thereby improving prediction accuracy.

3.1.4. Bidirectional LSTM

Bidirectional LSTMs process data in both directions, allowing it to consider both past and future contexts when making predictions [35]. A bidirectional LSTM comprises two LSTMs: one that processes the data forward and the other in the backward direction. The forward LSTM processes data from the first to the last time step, whereas the backward LSTM processes data from the last to the first time step. The outputs of the forward and backward LSTMs are concatenated at each time step, resulting in a final output that considers both the past and future contexts.

3.1.5. CNN and LSTM

A CNN with LSTM (CNN–LSTM) combines the strengths of both CNN and LSTM, which are typically used for image and video processing tasks, as it can extract features from image data by applying convolutional filters to the input. However, an LSTM is a type of RNN that is well suited for tasks involving sequential data, such as natural language processing and speech recognition. By combining the two, a CNN–LSTM network could use CNNs to apply a filter to the matrices of non-temporal building features, while using the LSTM to capture temporal dependencies in the input samples [36].

3.1.6. GRU

A gated recurrent unit (GRU) similar to LSTM was designed to improve the ability of RNNs to remember information over long periods [37]. The GRUs' unique structure allows them to store information for long periods and selectively decide which information to keep or discard. This is performed using a set of gates that control the flow of information through the network. The gates are composed of a sigmoid layer and point-wise multiplication operations. The sigmoid layer produces a value between zero and one, representing the probability of retaining or discarding information.

3.2. Federated Learning

FL is a distributed ML technique that allows multiple devices, such as smartphones, laptops, or IoT devices, to train an ML model collectively without sharing their data [38]. The data were collected, stored, and trained using traditional ML. Distributed samples rode on their respective clients and were locally trained on edge devices in FL. The model weights were shared with the parameter server responsible for aggregation. FL had several advantages over traditional ML methods. First, it allowed the training of models on a large and diverse dataset, which could improve a model's performance. Second, it provided training models for devices with limited computational resources, such as smartphones and IoT devices. Third, it protected user privacy, because the data remained on the device and were not shared with the central server.

FL has been widely applied in various fields, such as natural language processing, computer vision, and speech recognition. Smart devices make it possible to use FL for energy prediction, and further research needs to be conducted on FL energy-demand predictions. Similar to building energy demand prediction, an FL model could be trained on multiple smart meters, each with its own data, to improve the prediction accuracy. Smart meters collect data on temperature, humidity, and energy to train a local model. Then, the model parameters were sent to a central server, where they were aggregated to create a global model. The global model could be returned to the smart meters for further training. This technique had several advantages over traditional ML, including training on large and diverse datasets, training on devices with limited computational resources, and the protection of user privacy.

In this study, FL was used to analyze the local adaptation of a model trained on diverse client data, allowing personalized predictions that captured specific energy consumption patterns and characteristics. In federated personalized transfer learning, knowledge learned from one task could be transferred to another related task, such as the energy prediction of buildings having features similar to those of a trained domain but with fewer data samples. Similarly, meta-learning facilitated the rapid adaptation of a model learned from past experiences and leveraged a fine-tuned model to improve predictions for local clients. Meta-learning quickly adapted towards personalized client energy predictions after training diverse building consumers and reducing the data required for new buildings. The detailed workflow of the proposed study is given in Figure 1.



Figure 1. Overview of the flow of the study in detailed steps.

4. Experiments

4.1. Data Description

The Building Data Genome project dataset [39] was used to analyze the load prediction for the cross-silo federated settings. The open-source dataset encompasses numerous nonresidential buildings worldwide, each providing three key data types: historical electrical meter data, building metadata, and weather data. The climate zone of each building was determined using ASHRAE Standard 169-2013 and a total of eight distinct features were extracted, as given in Table 1. According to the SHAP (SHapley Additive explanations) technique, the "construction year" and "number of floors" have limited influence on the prediction model of building energy, while previous hour energy, hour, and day-type significantly influenced the predictive model [26]. Thus, in this study, day-type, hour, outdoor air temperature, relative humidity, previous hour energy, and gross floor area were selected as input variables of the ML model. Operational data from 13 office buildings were given in the cold climate zone, yielding hourly data collected from June to October. Among them, four educational buildings have been randomly selected for training, while one non-participant building from the remaining was selected for evaluation purposes.

Similarly, temporal and consumption-related attributes have a great impact on the forecasting model, and the availability of other attributes are not easily available from normal energy meters. A critical analysis of the learning process of a personalized metalearning has been carried out with only day-based features data. The training samples were selected from the hourly data of 135 days, while the model was initially validated for 16 days during testing with varying features, such as temperature, humidity, floor area, previous hour load, and day type. The load profile of the entire five months did not provide better visual insights about the variable load. Therefore, Figure 2 shows the hourly energy demand of a week for the five clients. The consumption pattern of Client 4 was a variable massive load during the four peak hours, whereas Client 2 had a lower demand during peak hours. The three remaining clients showed a moderate load curve during peak hours, and Client 3 had the maximum demand during off-peak hours. The input load profiles were all variable clients, causing heterogeneity in the sample and complicating the generalization of the time series forecasting model. In addition to the load, the temperature and humidity features were time-varying, whereas the floor areas of the clients differed.

Attributes Category Attributes Type
Temporal Day Type, Hour
Meteorological Outdoor Air Temperature, Outdoor Relative Humidity
Building structure Number of Floors, Gross Floor Area, Construction Year
Consumption-related Energy Utilization in Previous Hour

Table 1. Details of features used as input variables for ML models.



Figure 2. One-week hourly load profile of five buildings in the study.

4.2. Implementation

The TensorFlow framework simulates different environments, including cross-silo centralization, federated transfer, and meta-learning. Several data-driven models, such as ANN, LSTM, bidirectional LSTM, CNN, CNN–LSTM, and GRU, were trained individually in both cross-silo centralized and federated settings on hourly building samples. Four clients were trained in transfer learning, and the model was tested on the fifth building to analyze its effectiveness for personalization. The purpose of meta-learning was to design a model that can learn new tasks with minimal amounts of data. Essentially, the model was fine-tuned to quickly adapt the new tasks using few examples through the approach used in TensorFlow Federated (https://www.tensorflow.org/federated (accessed on 23 August 2023)). The overview of the FL process is shown in Figure 3.

In meta-learning, the number of features were reduced to temporal attributes, i.e, hourly energy consumption and types of day (weekday or weekend), and the model was fined tuned for better personalization. The trained global model was evaluated at the first epoch of the testing client, which measures how accurately a client with less data adapted to the fine-tuned model. Similarly, the local training loss across each client also showed the effectiveness of individual clients on global ML models during federated averaging. To analyze meta-learning in term of personalized learning, we increased the number of clients from different climate zones. Four clients were randomly selected from a pool of 40 clients and evaluated on two non-participating clients.



Figure 3. Federated Learning training and evaluation process in transfer and meta approach.

The study contains several data-driven algorithms with different numbers and types of sequential layers based on the model's architecture. The key hyper-parameters used in the study remain the same for each model training and are given in Table 2. All the models have one hidden dense layer with 10 neurons, while the ANN model has 50 neurons in the hidden layer. The activation function "ReLU" was used for all input and hidden layers, while for the model involving a CNN layer, the activation function was "Tanh". Similarly in the sequential algorithms, for ANN two dense layers were used: an LSTM layer was added to the LSTM model, and a convolutional layer was used to filter the input metrics in the CNN. An additional LSTM layer was added in the case of the CNN–LSTM model. In the bidirectional LSTM, the forward and backward layers of the LSTM were added, and a gated recurrent layer was added to the GRU model.

An adaptive moment estimation optimizer was selected for the training process in the centralized and federated simulations for fast and better convergence. A batch size of 80 samples was used in each round of FL, with 30 local epochs. The models were evaluated based on mean absolute error (MAE) metrics during training and validation. A detailed study of several traditional ML and DL algorithms found the MAE to be a suitable metric for comparative evaluation [5]. Moreover, to check the accuracy of the prospective model forecast for unseen samples, the test building was fed to the trained, centralized, and federated models to comparatively examine the energy forecasting result.

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

Hyperparameters	Search Space	Value
No. of Neurons	10, 20, 30, 50	10
Activation Function	ReLU, Tanh	Relu
Server Learning Rate	1.0, 0.10, 0.01	0.1
Optimizer	SGD, Adam	Adam
Batch Size	40, 60, 80, 100	80
No. of Rounds in CL	50, 100, 200, 300	200
Client Learning Rate	0.2, 0.02, 0.002	0.002
Client Epochs	5, 10, 20, 30	10
No. of Global Rounds	50, 100, 150, 200	100

4.3. Evaluation Results

This study aimed to evaluate six popular models used in time-series forecasting and check the precision of the load forecast on the unseen data of a building client. Simulations were run for both CL and FL, and the performance metrics were evaluated for all six algorithms to assess the results. Table 3 presents the evaluation of the respective algorithms for all training clients in the centralized and federated setups. The results showed that the performance of data-driven ML models was proportional to that of other data-driven and physical-based models. It gave better MAE than other ensemble learning studied by Olu-Ajayi et al. [5], such as gradient boosting, random forest, stacking, support vector machine, and decision tree. In the case of our single-model ML approaches, the ANN and CNN provided exceptional accuracy, comparable to time-series recurrent algorithms. The ANN model had a minimum loss of 0.118 during training, whereas the GRU had the least testing loss of 0.1191 in a centralized setup. In addition, the ANN yielded better results during the training period; however, it was the highest among all the models for validation. The CNN model provided comparable results during the federated and centralized network training periods. Similarly, the bidirectional LSTM showed better results, with the least MAE losses among all four variants of RNNs.

The weather profile affected the daily energy consumption of buildings; however, the pattern remained the same for clients in particular geographical regions. Client features differed between floor areas and demand patterns in the previous hours; both relied on the occupancy rate and activities inside the building. Therefore, an analysis of the evaluation of trained federated models for clients was carried out, which provided insights into the performance of individual clients who participated in FL training. Table 4 shows the results of the MAE loss on the testing data of the clients involved in privacy preserving training. A building with a highly variable time-series input load had a better convergence rate than the client with the least variability in data. Client 2 had a lesser gap between peak and base load, and the loss was extremely high in all ML models trained in FL. The energy demand of Client 4 varied between 19.75 kWh to 143.29 kWh, making it highly variable, and it gave the least MAE loss in the testing phase of CNN, BiLSTM, CNN–LSTM, and LSTM federated models.

To evaluate the model on an unseen dataset, Figure 4 shows that FL yielded better results than CL. All the data-driven algorithms were suitable for accurate hourly energy prediction (see Appendix A). However, the CNN could not accurately predict the declining hourly load. Simultaneously, the ANN provided promising results with precise hourly load curves compared to the actual load values. In physical-based tools used to profile the load scaling factor, a baseline pattern was scaled based on the average energy demand of the respective load [40]. In centralized ML, separate models were required for clients with high load variabilities, resulting in high computational and storage costs. Similarly, the trained model parameters could not be transferred to highly variable new client predictions with insufficient data. In addition, the hybrid scaling of physical and ML approaches did not work for transfer learning, as depicted in the sensitivity analysis of the availability of client load data. In FL, the model averaged the weights of all the clients instead of averaging the load data or fitting a single model to a sequential client dataset. To examine this issue, the dataset of Client 5 was scaled to 50%, 25%, and 15% and evaluated in centralized and federated transfer settings. The results in Figure 5 show that the model predicted better in FL for a 05% scaled operational dataset, whereas similar results were provided for the other cases of FL and CL (see Appendix A).

Model	Centralized		Federated Transfer		Federated Meta	
	Train	Test	Train	Test	Train	Test
ANN	0.1187	0.1498	0.1591	0.2373	0.1537	0.1567
LSTM	0.1681	0.1301	0.2539	0.2195	0.1430	0.1679
CNN	0.1227	0.1199	0.1872	0.1951	0.1516	0.2135
Bi LSTM	0.1468	0.1020	0.2221	0.1951	0.1764	0.2083
CNN– LSTM	0.1904	0.1389	0.2417	0.2148	0.2203	0.2357
GRU	0.1721	0.1037	0.2461	0.2838	0.1550	0.1898

Table 3. MAE loss of data-driven models trained and evaluated over all the clients' data in centralized and federated learning.

Table 4. MAE loss of data-driven models evaluated for each client after training in federated learning.

Model	Client 1	Client 2	Client 3	Client 4
ANN	0.2496	0.2822	0.2425	0.1750
LSTM	0.1668	0.3228	0.2356	0.1529
CNN	0.1675	0.2736	0.2140	0.1436
Bi LSTM	0.1541	0.2775	0.2142	0.1346
CNN-LSTM	0.1644	0.3207	0.2254	0.1486
GRU	0.1924	0.4681	0.2887	0.1861



Figure 4. Energy demand prediction for the test client to examine the hourly variation of load for the LSTM model trained in (**a**) Centralized Learning and (**b**) Federated Learning.



Figure 5. Energy demand prediction on LSTM model for Building 5 with five percent scaled operational data in (**a**) Centralized Learning and (**b**) Federated Learning.

5. Discussion and Future Work

The study comparatively analyzed the impact of spatial and temporal features on ML models. It highlights how global weights learned on different client data were well adaptable for unseen features in transfer and meta-learning. Moreover, it addressed the issue of data scarcity for new clients and the effectiveness of weight aggregation instead of centralized data scaling. The study does not consider the computational cost, multiple climate zones, structural attributes, occupancy rates, or residential and industrial sectors, which could provide better insights for multiple federated ML models. The evolution of big data and digital twins under two-way communication make control over client real-time data by government bodies and stakeholders more likely [41].

In smart cites, smart home users are more concerned about data leakage, privacy violation, malicious intent, and interpretability in data sharing to a centralized network. Some enterprises, such as GECKO (https://gecko-project.eu/ (accessed on 25 August 2023)), aim for transparent and explainable AI models under ethical consideration of GDPR using FL to provide secure, scalable, privacy-focused, and affordable solution for smart cities [42]. In future work, we will extend the application of FL to address the smart home with IoT-based data. The study will consider additional attributes such as consumer behaviors, occupancy rates, and other energy related attributes, under the sensitivity analysis on data heterogeneity, scalability, latency, and connectivity.

6. Conclusions

Building energy prediction relies on physical-based tools for hourly demand prediction of consumer loads. The software requires sufficient information regarding the detailed environmental and infrastructural parameters of a building for accurate predictions. The collection of such data is a challenging task; thus, ML approaches show an improvement in time-series forecasting. A comparative analysis of the ANN, LSTM, CNN, bidirectional LSTM, CNN–LSTM, and GRU models was performed to examine the accuracy of energy prediction in privacy preserving FL. Typically, centralized ML architectures incur data storage and computation costs with the prospective threat of information leakage from clients. Similarly, new clients have fewer operational samples and use a centralized model for energy prediction. This study examined how the trained model better fit the transfer of learned patterns for data-insufficient clients for CL and FL. The results showed that data-driven algorithms were best suited for energy demand prediction in federated transfer learning for clients with insufficient data. Similarly, it provided better personalized predictions for heterogeneous energy consumers. For less than 50% of the scaled data, the LSTM model trained in the FL setting outperformed the CL model. Thus, data-driven ML models are recommended for energy planners to better predict building energy consumption, while preserving the privacy of end consumers.

Author Contributions: Conceptualization, M.A. and B.J.C.; methodology, M.A. and B.J.C.; software, M.A. and A.K.S.; validation, M.A. and A.K.S.; formal analysis, A.K. and B.J.C.; investigation, M.A. and A.K.S.; resources, M.A., A.K.S. and B.J.C.; data curation, M.A. and A.K.S.; writing—original draft preparation, M.A.; writing—review and editing, S.S.A. and B.J.C.; visualization, M.A. and S.S.A.; supervision, B.J.C.; project administration, B.J.C.; funding acquisition, B.J.C. All authors have read and agreed to the published version of the manuscript.

Funding: 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.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ANN	Artificial Neural Network
CL	Centralized Learning
CNN	Convolutional Neural Network
DL	Deep Learning
FL	Federated Learning
GB	Gradient Boosting
GDPR	General Data Protection Regulation
GHG	Greenhouse gas
GRU	Gated Recurrent Unit
HVAC	Heating, Ventilation and Air Conditioning
IoT	Internet of Things
LSTM	Long Sort-term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MSE	Mean Square Error
ReLU	Rectified Linear Unit
RF	Random Forest
RNN	Recurrent Neural Network

SVM Support Vector Machine

Appendix A

Forecasting results of data-driven algorithms, i.e., ANN, CNN, Bidirectional LSTM, CNN–LSTM, and GRU, that are used in federated learning.



Figure A1. Energy demand prediction for the test client to examine the hourly variation of load for ANN model trained in FL.



Figure A2. Energy demand prediction for the test client for CNN model trained in FL.



Figure A3. Energy demand prediction for the test client for bidirectional LSTM model trained in FL.



Figure A4. Energy demand prediction for the test client for CNN-LSTM model trained in FL.



Figure A5. Energy demand prediction for the test client for GRU model trained in FL.



Figure A6. Energy demand prediction on LSTM model for Building 5 with 50 percent scaled operational data in (**a**) Centralized Learning and (**b**) Federated Learning.



Figure A7. Energy demand prediction on LSTM model for Building 5 with 25 percent scaled operational data in (**a**) Centralized Learning and (**b**) Federated Learning.



Figure A8. Energy demand prediction on LSTM model for Building 5 with 15 percent scaled operational data in (**a**) Centralized Learning and (**b**) Federated Learning.

References

- Ren, Y.Y.; Ren, G.Y.; Sun, X.B.; Shrestha, A.B.; You, Q.L.; Zhan, Y.J.; Rajbhandari, R.; Zhang, P.F.; Wen, K.M. Observed changes in surface air temperature and precipitation in the Hindu Kush Himalayan region over the last 100-plus years. *Adv. Clim. Chang. Res.* 2017, *8*, 148–156. [CrossRef]
- 2. Khalil, M.; McGough, A.S.; Pourmirza, Z.; Pazhoohesh, M.; Walker, S. Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption—A systematic review. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105287. [CrossRef]
- 3. Ohene, E.; Chan, A.P.; Darko, A. Prioritizing barriers and developing mitigation strategies toward net-zero carbon building sector. *Build. Environ.* **2022**, 223 , 109437. [CrossRef]
- 4. Dong, B.; Prakash, V.; Feng, F.; O'Neill, Z. A review of smart building sensing system for better indoor environment control. *Energy Build.* **2019**, 199, 29–46. [CrossRef]
- 5. Olu-Ajayi, R.; Alaka, H.; Sulaimon, I.; Sunmola, F.; Ajayi, S. Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. *J. Build. Eng.* **2022**, *45*, 103406. [CrossRef]
- Colmenar-Santos, A.; de Lober, L.N.T.; Borge-Diez, D.; Castro-Gil, M. Solutions to reduce energy consumption in the management of large buildings. *Energy Build.* 2013, 56, 66–77. [CrossRef]
- 7. Konečný, J.; McMahan, H.B.; Yu, F.X.; Richtárik, P.; Suresh, A.T.; Bacon, D. Federated learning: Strategies for improving communication efficiency. *arXiv* 2016, arXiv:1610.05492.
- 8. Runge, J.; Zmeureanu, R. Forecasting energy use in buildings using artificial neural networks: A review. *Energies* **2019**, *12*, 3254. [CrossRef]
- 9. Mahapatra, B.; Nayyar, A. Home energy management system (HEMS): Concept, architecture, infrastructure, challenges and energy management schemes. *Energy Syst.* 2022, *13*, 643–669. [CrossRef]
- Zafar, U.; Bayhan, S.; Sanfilippo, A. Home energy management system concepts, configurations, and technologies for the smart grid. *IEEE Access* 2020, *8*, 119271–119286. [CrossRef]
- 11. Marikyan, D.; Papagiannidis, S.; Alamanos, E. A systematic review of the smart home literature: A user perspective. *Technol. Forecast. Soc. Chang.* **2019**, *138*, 139–154. [CrossRef]
- 12. Harish, V.; Kumar, A. A review on modeling and simulation of building energy systems. *Renew. Sustain. Energy Rev.* 2016, 56, 1272–1292. [CrossRef]
- 13. Pham, A.D.; Ngo, N.T.; Truong, T.T.H.; Huynh, N.T.; Truong, N.S. Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability. *J. Clean. Prod.* **2020**, *260*, 121082. [CrossRef]
- Petrangeli, E.; Tonellotto, N.; Vallati, C. Performance Evaluation of Federated Learning for Residential Energy Forecasting. *IoT* 2022, *3*, 381–397. [CrossRef]
- 15. Shi, Y.; Xu, X. Deep Federated Adaptation: An Adaptative Residential Load Forecasting Approach with Federated Learning. *Sensors* **2022**, 22, 3264. [CrossRef] [PubMed]
- Sater, R.A.; Hamza, A.B. A federated learning approach to anomaly detection in smart buildings. *ACM Trans. Internet Things* 2021, 2, 1–23. [CrossRef]
- 17. Dutta, G.; Mitra, K. A literature review on dynamic pricing of electricity. J. Oper. Res. Soc. 2017, 68, 1131–1145. [CrossRef]
- Gupta, A.; Gupta, H.P.; Das, S.K. FedAR+: A Federated Learning Approach to Appliance Recognition with Mislabeled Data in Residential Buildings. arXiv 2022, arXiv:2209.01338.
- Gao, J.; Wang, W.; Liu, Z.; Billah, M.F.R.M.; Campbell, B. Decentralized federated learning framework for the neighborhood: A case study on residential building load forecasting. In Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems, Coimbra, Portugal, 15–17 November 2021; pp. 453–459.
- 20. Gholizadeh, N.; Musilek, P. Federated learning with hyperparameter-based clustering for electrical load forecasting. *Internet Things* **2022**, *17*, 100470. [CrossRef]
- 21. Husnoo, M.A.; Anwar, A.; Hosseinzadeh, N.; Islam, S.N.; Mahmood, A.N.; Doss, R. FedREP: Towards Horizontal Federated Load Forecasting for Retail Energy Providers. *arXiv* 2022, arXiv:2203.00219.
- 22. Taïk, A.; Cherkaoui, S. Electrical load forecasting using edge computing and federated learning. In Proceedings of the ICC 2020-2020 IEEE International Conference on Communications (ICC), Virtual, 7–11 June 2020; pp. 1–6.
- 23. Briggs, C.; Fan, Z.; Andras, P. Federated Learning for Short-term Residential Load Forecasting. *IEEE Open Access J. Power Energy* 2022, *9*, 573–583. [CrossRef]
- 24. Tang, L.; Xie, H.; Wang, X.; Bie, Z. Privacy-preserving knowledge sharing for few-shot building energy prediction: A federated learning approach. *Appl. Energy* **2023**, *337*, 120860. [CrossRef]
- 25. Dogra, A.; Anand, A.; Bedi, J. Consumers profiling based federated learning approach for energy load forecasting. *Sustain. Cities Soc.* **2023**, *98*, 104815. [CrossRef]
- 26. Li, J.; Zhang, C.; Zhao, Y.; Qiu, W.; Chen, Q.; Zhang, X. Federated learning-based short-term building energy consumption prediction method for solving the data silos problem. *Build. Simul.* **2022**, *15*, 1145–1159. [CrossRef]
- Badr, M.M.; Ibrahem, M.I.; Mahmoud, M.; Alasmary, W.; Fouda, M.M.; Almotairi, K.H.; Fadlullah, Z.M. Privacy-preserving federated-learning-based net-energy forecasting. In Proceedings of the SoutheastCon 2022, Mobile, AL, USA, 26 March–3 April 2022; pp. 133–139.
- 28. Wang, F.; Cen, J.; Yu, Z.; Deng, S.; Zhang, G. Research on a hybrid model for cooling load prediction based on wavelet threshold denoising and deep learning: A study in China. *Energy Rep.* **2022**, *8*, 10950–10962. [CrossRef]

- Alrasheedi, A.; Almalaq, A. Hybrid Deep Learning Applied on Saudi Smart Grids for Short-Term Load Forecasting. *Mathematics* 2022, 10, 2666. [CrossRef]
- Khortsriwong, N.; Boonraksa, P.; Boonraksa, T.; Fangsuwannarak, T.; Boonsrirat, A.; Pinthurat, W.; Marungsri, B. Performance of Deep Learning Techniques for Forecasting PV Power Generation: A Case Study on a 1.5 MWp Floating PV Power Plant. *Energies* 2023, 16, 2119. [CrossRef]
- Zhen, H.; Niu, D.; Yu, M.; Wang, K.; Liang, Y.; Xu, X. A hybrid deep learning model and comparison for wind power forecasting considering temporal-Spatial feature extraction. *Sustainability* 2020, *12*, 9490. [CrossRef]
- Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* 2018, 4, e00938. [CrossRef]
- 33. Van Houdt, G.; Mosquera, C.; Nápoles, G. A review on the long short-term memory model. *Artif. Intell. Rev.* 2020, *53*, 5929–5955. [CrossRef]
- 34. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, 77, 354–377. [CrossRef]
- Peng, T.; Zhang, C.; Zhou, J.; Nazir, M.S. An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. *Energy* 2021, 221, 119887. [CrossRef]
- Tian, C.; Ma, J.; Zhang, C.; Zhan, P. A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. *Energies* 2018, 11, 3493. [CrossRef]
- 37. Ke, K.; Hongbin, S.; Chengkang, Z.; Brown, C. Short-term electrical load forecasting method based on stacked auto-encoding and GRU neural network. *Evol. Intell.* **2019**, *12*, 385–394. [CrossRef]
- Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Li, J.; Poor, H.V. Federated learning for internet of things: A comprehensive survey. *IEEE Commun. Surv. Tutor.* 2021, 23, 1622–1658. [CrossRef]
- Miller, C.; Kathirgamanathan, A.; Picchetti, B.; Arjunan, P.; Park, J.Y.; Nagy, Z.; Raftery, P.; Hobson, B.W.; Shi, Z.; Meggers, F. The building data genome project 2, energy meter data from the ASHRAE great energy predictor III competition. *Sci. Data* 2020, 7, 1–13. [CrossRef] [PubMed]
- 40. Ali, M.; Wazir, R.; Imran, K.; Ullah, K.; Janjua, A.K.; Ulasyar, A.; Khattak, A.; Guerrero, J.M. Techno-economic assessment and sustainability impact of hybrid energy systems in Gilgit-Baltistan, Pakistan. *Energy Rep.* **2021**, *7*, 2546–2562. [CrossRef]
- 41. Ziosi, M.; Hewitt, B.; Juneja, P.; Taddeo, M.; Floridi, L. Smart cities: Reviewing the debate about their ethical implications. *AI Soc.* **2022.** [CrossRef]
- 42. Pandya, S.; Srivastava, G.; Jhaveri, R.; Babu, M.R.; Bhattacharya, S.; Maddikunta, P.K.R.; Mastorakis, S.; Piran, M.J.; Gadekallu, T.R. Federated learning for smart cities: A comprehensive survey. *Sustain. Energy Technol. Assess.* **2023**, *55*, 102987. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.