IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 November 2025, accepted 16 November 2025, date of publication 24 November 2025, date of current version 4 December 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3636433

RESEARCH ARTICLE

Spatial–Temporal Deep Learning for Electric-Vehicle Charging Demand: An Exploratory Study of Graph Convolutional and LSTM Networks Performance

MAHER ALARAJ[1], (Member, IEEE), CARLA MARTINS[2], MOHAMMED RADI[3], (Member, IEEE), MOHAMED DARWISH[4], AND MUNIR MAJDALAWIEH[1]

[1]Department of Information Systems and Technology Management, College of Technological Innovation, Zayed University, Dubai 19282, United Arab Emirates
[2]ULS de Sao Joao, Alameda Professor Hernani Monteiro, University of Porto, Porto 4200-319, Portugal
[3]UK Power Networks, SE1 6NP London, U.K.
[4]Department of Electronic and Electrical Engineering, College of Engineering, Design and Physical Sciences, Brunel University of London, UB8 3PH Uxbridge, U.K.

Corresponding author: Maher Alaraj (maher.alaraj@zu.ac.ae)

This work was supported by the Office of Research, Zayed University, through the Research Incentive Fund under Grant R23079.

ABSTRACT Electric-vehicle (EV) charging is a localized, time-varying load that challenges distribution networks. This study offers practical insights into when spatial graph structure adds value beyond temporal context, utilizing real-world data and a transparent evaluation. We compare Long Short-Term Memory (LSTM) and Graph Convolutional Network (GCN) models for hourly EV-charging energy forecasting, based on 145,778 sessions recorded in Boulder, Colorado (2018–2023). After preprocessing and temporal alignment, temporal covariates (hour, day, month, year) and, when applicable, ZIP-code indicators were engineered. LSTMs were trained with 1 h and 24 h input windows, with or without ZIP features, and evaluated through 5-fold cross-validation. GCNs operated on hourly node–time tensors with a dense adjacency (no self-loops) and were trained on an 80/20 temporal split, using a 1% subsample for tractability. All models used Adam (lr = 0.005), early stopping, and ReduceLROnPlateau. Temporal context was the main driver of LSTM accuracy: 24 h inputs outperformed 1 h, while ZIP features improved only the shorter window. For GCNs, depth and node features shaped performance: a 2-layer GCN with ZIP features achieved the lowest RMSE (4.75 kWh), whereas a 6-layer GCN without node features reached the lowest MAE (2.46 kWh). Despite higher computational cost, GCNs captured spatial coupling effectively. Overall, 24 h LSTMs provide a strong and efficient baseline, while GCNs add value when spatial correlations are relevant. A hybrid GCN–LSTM architecture is a promising next step to jointly leverage spatial and temporal dependencies.

INDEX TERMS Electric vehicle charging, energy consumption forecasting, graph convolutional networks (GCN), long short-term memory (LSTM), smart grid management, spatio-temporal deep learning.

I. INTRODUCTION

In recent years, there has been a significant global increase in electricity demand, a phenomenon the International Energy Agency (IEA) has described as the "New Age of Electricity." [1] This growth is particularly evident in emerging economies and countries with high gross domestic product (GDP), where electric vehicles (EVs) are becoming one of the primary drivers of electricity demand, currently ranking among the top five contributors, alongside heat pumps, electrolyzers, air conditioning systems, and data centers. EVs now account for approximately 20% of all vehicle sales worldwide, reflecting their rapid market penetration and the growing pressure they place on electricity systems [2]. As a result, power grids

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaosong Hu.

are under increasing pressure to meet peak demand levels, posing significant challenges for energy infrastructure and planning [1], [2]. This growing demand for electricity poses significant challenges to the reliability of power supply, particularly due to the need for uninterrupted energy and the geographic concentration of high energy consumption. One notable example of such localized energy demands includes electric vehicle (EV) charging hubs or residential areas with a high density of home-charging EVs. These localized loads can create stress points in the power grid, potentially leading to service disruptions and network instability [2], [3]. Power grid stability refers, among other aspects, to the ability of the system to return to an equilibrium state after being subjected to a physical disturbance, such as a fault, lightning strike, or sudden change in generation or load. A more stable grid also enhances system reliability, as a reliable power system is more resilient to failures and less likely to experience blackouts or cascading outages. Moreover, such systems improve safety, since unexpected power fluctuations can potentially damage electrical equipment [4]. One of the potential strategies to mitigate instabilities in power supply systems is to understand supply and demand patterns, thereby enabling supply adjustments based on consumption needs. The most appropriate method for forecasting energy consumption involves time series modeling. Time series databases are collections of sequential data points recorded at regular time intervals, representing observations of phenomena that evolve. Analyzing these data points allows for the identification of various patterns, such as long-term trends, seasonality, cyclicality, and irregularities. These patterns provide valuable insights for forecasting future values and detecting changes during critical periods [5]. Another key characteristic of the datasets used in time series analysis, and particularly for energy consumption forecasting, is their inherent complexity. These datasets are typically high-dimensional, publicly available, and collected from multiple locations. They often exhibit nonlinear patterns, high dimensionality, and both temporal and spatial dependencies, which pose additional challenges for accurate modeling and prediction. Traditional statistical methods were the most used approaches for time series forecasting before the widespread adoption of deep learning techniques. One of the most widely applied models was the Autoregressive Integrated Moving Average (ARIMA) [6]. However, although statistical methods remain relevant in certain scenarios, they have largely been outperformed by deep learning approaches in terms of predictive accuracy and adaptability to complex data patterns [6], [7]. In recent years, various deep learning models and architectures have been explored for time series forecasting. However, there is still no clear consensus on which model or architecture performs best in each context or for specific types of data. Among the deep learning models applied to time series forecasting, Recurrent Neural Networks (RNNs) and their more advanced variant, Long Short-Term Memory (LSTM) networks, have shown significant promise. RNNs are specifically designed to process sequential data by maintaining a hidden state that captures information from previous time steps. However, traditional RNNs suffer from limitations such as vanishing and exploding gradients, which hinder their ability to learn long-term dependencies. LSTM networks were developed to address these issues by incorporating memory cells and gating mechanisms that regulate the flow of information [8]. As a result, LSTMs are particularly effective in modeling complex, nonlinear, and long-range temporal dependencies, making them a preferred choice in many time series forecasting applications, including energy consumption prediction [9]. Nevertheless, the LSTM architecture is primarily designed to handle data with an underlying Euclidean structure, such as images or regular tabular datasets. In contrast, when dealing with data characterized by complex and irregular structures, such as those found in power grids, more advanced approaches are required. This has led to the emergence of Graph Convolutional Networks (GCNs), which are specifically designed to model spatial dependencies within graph-structured data [10]. In a GCN model, there are two fundamental components: (1) the nodes, which represent entities (such as power plants or charging stations), and (2) the edges, which represent the connections between these nodes. To model spatial dependencies, GCNs can learn patterns within this graph structure, enabling tasks such as flow prediction, vulnerability identification, anomaly detection, and analyzing the impact of disturbances [11]. Thus, while LSTMs offer the primary advantage of modeling temporal dependencies in sequential data, GCN models excel in capturing spatial dependencies. These two types of models have become among the most favored choices for forecasting energy consumption data, particularly for the specific characteristics of electric vehicle (EV) charging. EV charging can place significant pressure on the power grid at specific locations, times, or power levels, making these models highly suitable for addressing such challenges. The goal of this paper is to analyze the performance of Graph Convolutional Networks (GCN) and Long Short-Term Memory (LSTM) networks for energy consumption forecasting in the context of predicting the behavior of EV charging.

II. RELATED WORK

Previous studies, published within the last four years, have sought to develop models capable of addressing the challenge of energy consumption forecasting in the context of electric vehicle charging demands, particularly by employing artificial intelligence and deep learning methods such as Long Short-Term Memory (LSTM) networks and Graph Convolutional Networks (GCNs). The results obtained so far are promising, with improvements and the proposal of new models and frameworks that enhance forecasting accuracy. Some of these works are discussed in the following section.

A. LONG SHORT-TERM MEMORY (LSTM)

Long Short-Term Memory (LSTM) networks are well-suited for time series prediction and have been employed in numerous studies. In a study by Shangmuganatan [12] published in 2022, an LSTM model was optimized using the classical Arithmetic Optimization Algorithm (AOA) and time series data. The proposed model was tested on the Georgia Tech electric vehicle charging dataset and achieved a prediction accuracy of 97.14%, with a Mean Absolute Error (MAE) of 0.1083 and a Root Mean Square Error (RMSE) of 2.0628×10.5, demonstrating its capabilities. In another study, the performance of the LSTM model was validated using real-world data from fast-charging stations on Jeju Island, South Korea. The proposed LSTM model achieved strong predictive performance, with an RMSE of 61.63, a Normal- ized Mean Absolute Error (NMAE) of 5.15%, and a Normal- ized Root Mean Square Error (NRMSE) of 6.65%. It outper- formed other deep learning models, including Bi-LSTM and RNN, further confirming its effectiveness in forecasting elec- tric vehicle charging demand [13]. Another study employed LSTM networks to forecast electric vehicle (EV) charging demand at the station level from 1 to 5 hours, using a dataset comprising trajectories of over 76,000 private EVs in Beijing. The LSTM model was compared with the ARIMA model, while also examining the influence of different input data structures, sample sizes, intervals, and time periods. The results indicated that the LSTM model outperformed ARIMA, achieving a Mean Absolute Percentage Error (MAPE) of 6.83%. Moreover, the time period and interval were found to have a greater impact on LSTM's accuracy than data structure or sample size, with shorter periods and intervals (e.g., 1 hour) yielding better predictive performance [14]. Although results generally favor the use of LSTM over ARIMA for time series forecasting, this advantage is not consistent across all contexts. For example, a study comparing different forecasting models for the increasing energy consumption associated with the growing popularity of electric vehicles (EVs) across multiple geographic scales over two years, using both ARIMA and LSTM models, showed nuanced findings. For macro-level data (national and city scales), where patterns are more linear and stable, the ARIMA model with regressors yielded the best performance. In contrast, for micro-level data (individual charging stations), where individual user behavior plays a more significant role, the LSTM model outperformed ARIMA, demonstrating superior predictive capacity [15]. Interestingly, another study found that the performance of LSTM models can be enhanced when multi-feature inputs are incorporated, such as historical weather observations, including wind speed, temperature, and humidity. This multi-feature approach assumes that weather conditions influence the driving behavior and mobility patterns of EV users, thereby affecting their charging habits. The implemented model achieved a mean absolute error (MAE) of 3.29%, which was lower than that of the baseline LSTM model, demonstrating improved forecast accuracy [16].

B. GRAPH CONVOLUTIONAL NEURAL NETWORKS (GCN)

Studies employing only GCNs are not available to our knowledge. Although certain studies highlight the strong potential of GCNs for spatiotemporal forecasting, their effective use typically involves integration with other techniques, resulting in hybrid models. This combination is often necessary to fully capture both spatial relationships and temporal dynamics inherent in such data. In one study conducted by Wang et al [17], an innovative model, trained and validated using an extensive GPS dataset encompassing over 76,000 EVs charging sessions in Beijing, demonstrated superior performance compared to existing spatiotemporal approaches. This superiority was observed in forecasting accuracy across various regional scales (1km x 1km, 2km x 2km, 4km x 4km) and for short-term prediction horizons, specifically 3 and 6 hours, underscoring the model's ability to capture the complex dynamics of charging demand in real urban scenarios. Another investigation, led by Fahim et al. [18], focused on developing a hybrid model to forecast demand at EV charging stations, integrating an optimized fusion Graph Convolutional Network (GCN) with the Berkeley Wavelet Transform (BWT). This approach achieved a notable evaluation efficiency of 98%, signaling significant improvements over previous techniques in modeling charging demand. Furthermore, the Transformer architecture has been explored in other studies for temporal forecasting, often in combination with GCNs to incorporate the understanding of spatial dependencies. For instance, Zhang [19] developed a model to predict the usage of EV charging stations using data from the city of Dundee, Scotland. The results showed that this hybrid model, combining GCN and Transformer, attained a forecasting accuracy exceeding 80%, reinforcing the effectiveness of integrating graph processing for spatial relationships with the Transformer's sequential data modeling capabilities for temporal data. The low use of GCN alone for the temporal forecast may be related to the possible inability of the model to deal with temporal features and the high computational cost that limits the possible implementations for simpler scenarios or smaller data samples, when researchers do not have access to supercomputers.

III. AIM

The primary aim of this research is to develop and test long-term forecasting models for hourly energy consumption, with a particular focus on predicting the charging behavior of electric vehicles (EVs) in terms of kilowatt-hours (kWh). The study seeks to analyze the performance of Graph Convolutional Networks (GCN) and Long Short-Term Memory (LSTM) architectures to determine their respective strengths and weaknesses in capturing spatial and temporal patterns.

IV. METHODS

A. DATA

The dataset is publicly accessible as open source, and its use for research purposes does not require prior authorization. The dataset comprises 145,778 electric vehicle charging sessions recorded in Boulder, Colorado, between January 2018 and December 2023. Each record corresponds to an individual charging session and includes information on the charging station (name, address, and ZIP code), temporal variables (start and end date and time of the session), utilization metrics (session duration, effective charging time, and energy consumption in kWh), and environmental indicators (greenhouse gas emission reductions and gasoline/diesel savings). The dataset encompasses five ZIP codes associated with charging activity. ZIP code 80302 accounts for the largest share, with approximately 66,000 sessions and 547 MWh of energy consumed, characterized by predominantly nighttime charging profiles and higher average energy per session (9 kWh). In ZIP code 80303, total consumption is lower (213 MWh), but sessions are on average longer and exhibit a wider temporal dispersion. ZIP codes 80304 and 80301 (162 and 160 MWh, respectively) show more stable patterns, with usage distributed throughout the day and reduced seasonal variability. In contrast, ZIP code 80305 (105 MWh) is marked by lower average consumption (7 kWh) and a stronger concentration during daytime hours, suggesting usage profiles more closely linked to short-distance travel. Fig. 1 shows the total energy consumption during the data time frame.

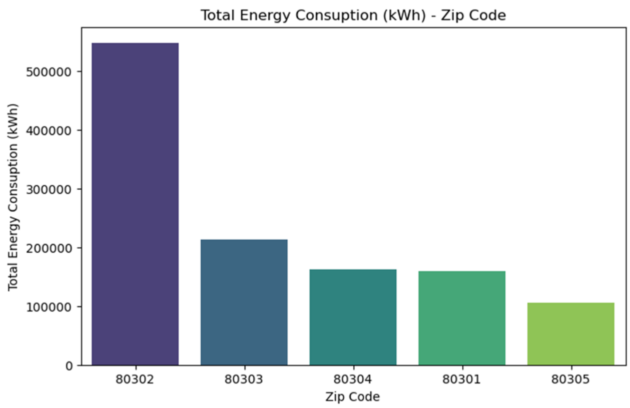

FIGURE 1. **Total energy consumption (kWh) by ZIP code.**

From a temporal perspective, there is a clear upward trend in average energy consumption per session, increasing from 7.9 kWh in 2019 to 8.7 kWh in 2022, followed by a slight decline in 2023, this trend can be seen in Fig. 2. Seasonal variation is evident, with higher values observed in winter and lower values in summer. The hourly analysis reveals a typical residential pattern, characterized by nighttime peaks in charging activity and reduced consumption during the afternoon (Fig. 3).

B. PREPROCESSING

Data preprocessing pipelines were implemented separately for the different architectures, ensuring that the input features, temporal structure, and normalization procedures were aligned with the specific requirements of each model.

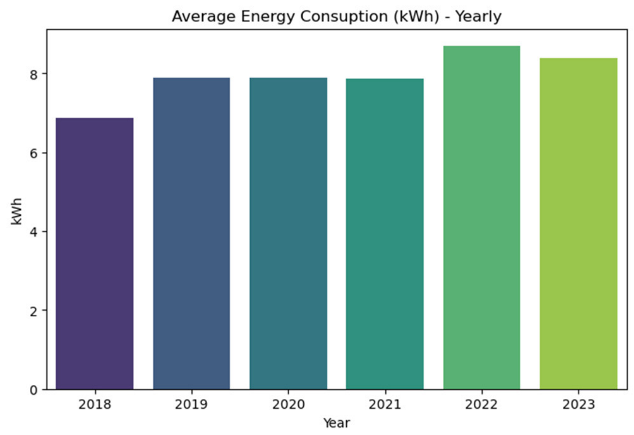

FIGURE 2. **Average yearly energy consumption (kWh) from 2018 to 2023.**

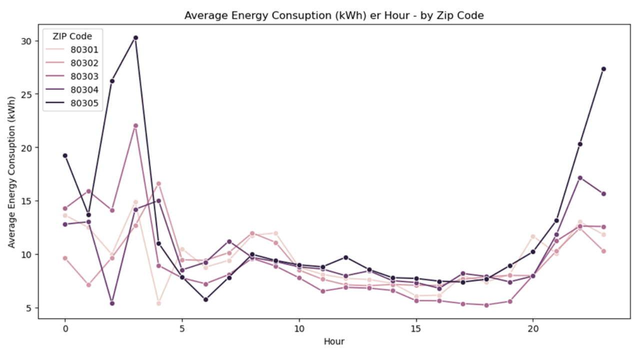

FIGURE 3. **Average energy consumption (kWh) per Hour by ZIP code.**

1) LSTM MODELS

After loading the database, the date column was converted to a datetime format and decomposed into derived temporal attributes: hour, day, month, and year. The main numerical variables (energy, savings, temporal attributes) were coerced into numeric format, with non-convertible values treated as NaN. Subsequently, observations with missing values in the critical columns were removed. The reference node (Zip Code) was then defined, and the data were aggregated on an hourly basis. Each combination of hour and postal code generated a row containing the total energy consumption and the corresponding temporal attributes. To avoid gaps, a complete grid was created with all possible combinations of postal codes and available hours. The features selected for modelling were energy consumption (the target variable to be predicted) and the temporal attributes (hour, day, month, year). Two datasets were prepared: one containing only these numerical variables, and another extending the first with a one-hot encoding of the postal code (allowing the models to account for location as a categorical variable). Finally, a function was defined to create temporal windows: the data were organized into fixed-length sequences (one with a length of 24 hours and another with a length of 1 hour), with the target variable (energy) placed at the subsequent time step.

2) GCN MODELS

Data loading, datetime conversion, and NaN handling were performed consistently with the LSTM preprocessing pipeline. To ensure the computational feasibility of GCN training on a personal computer, the dataset was downsampled to a random 1% subset. Records were then aggregated at an hourly resolution by postal code, summing energy consumption within each interval, and a complete grid of all postal code–hour combinations was generated to avoid missing entries. The node variable (postal code) was one-hot encoded, yielding a feature set that combined scaled temporal attributes with a categorical representation of location. The dataset was partitioned strictly along the temporal axis into training (80%) and validation (20%) sets, with no randomization, thereby preserving the chronological order of events. To prevent data leakage, the standardization of nu- numerical features was performed using StandardScaler fitted exclusively on the training data; the same transformation was then applied to the validation set. This ensured that future information was never incorporated into the training process. The dataset was subsequently ordered by time and node and reshaped into 3D tensors (time × nodes × features), with the target variable (energy) structured in the same format. Fi- Finally, a complete adjacency matrix (dense graph without self-loops) was defined and replicated across time, ensuring that the GCN model had access to node connectivity information at each time step.

C. STATISTICAL ANALYSIS

To evaluate whether the observed differences in model performance were statistically significant, non-parametric significance tests were performed using the fold-level results from the five-fold cross-validation. For comparisons involving two groups (LSTM architectures with different input window lengths), the Wilcoxon signed-rank test was applied. For comparisons among three groups (the 2-, 4-, and 6-layer GCN architectures), the Friedman Chi-Square test was used. These non-parametric tests were selected because they do not assume normality and are appropriate for small paired samples. Statistical significance was set at $p < 0.05$.

V. MODELS

A. LSTM

Long-short-term memory networks (LSTMs), are a type of recurrent neural network capable of capturing sequential dependencies in time-series data. To model the temporal dependencies present in energy consumption patterns, we employed a two-layer LSTM architecture, with each layer consisting of 64 hidden units. A schematic representation is shown in fig. 4. This configuration allows the network to retain long-range temporal information and detect complex temporal trends, such as fluctuations in energy usage across different hours, days, or seasons, enhancing the model's capacity to learn meaningful time-based patterns. Feature selection was conducted through an integrated approach combining correlation analysis and empirical model performance evaluation. Four different LSTM models were trained: in two of the models, time steps were set to 24h, one model containing only time features (hour, day, month, and year), and in the other, node features (ZIP code) were added to the model. Another set of two models with the same entry features was used, where time steps were set to 1h.

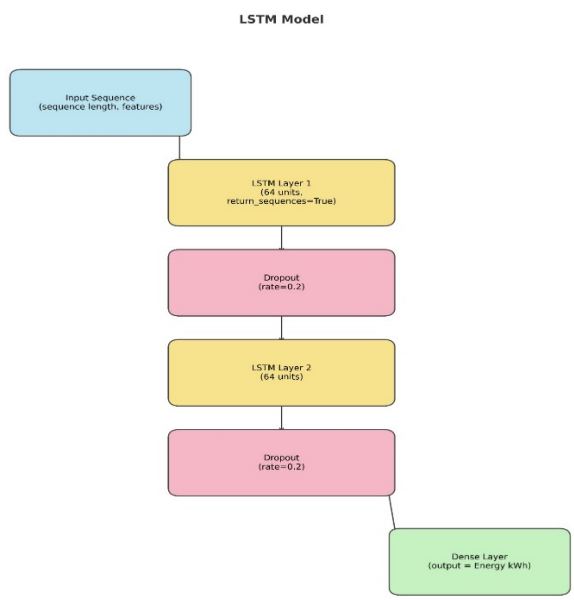

FIGURE 4. LSTM models architecture.

B. GCN

Graph Convolutional Networks (GCNs) are a class of neural networks specifically designed to operate on graph-structured data, making them well-suited for modeling spatial relationships among charging stations in an energy system. In this study, the energy system is represented as a graph, where nodes correspond to ZIP codes of the charging stations and edges denote the connections or relationships between them. To assess the efficacy of various graph-based learning strategies, 6 configurations of Graph Convolutional Network (GCN) architectures were implemented, each tailored to accommodate different degrees of graph complexity and levels of feature integration. A basic two-layer GCN served as a lightweight baseline model; a GCN with four layers was designed to provide a balanced architecture, rendering it suitable for modeling moderately complex spatial relationships; and a six-layer GCN introduced greater representational capacity through increased depth, thereby enhancing the model's ability to learn intricate graph patterns; however, this complexity also heightened the risk of overfitting or noise introduction, particularly in scenarios involving smaller datasets. An additional variant of the three GCN with node features was introduced, where node features were introduced as an adjacency matrix (dense graph without self-loops) and replicated across time, ensuring that these GCN models had access to node connectivity information at each time step. The architecture of the GCN models is shown in fig 5.

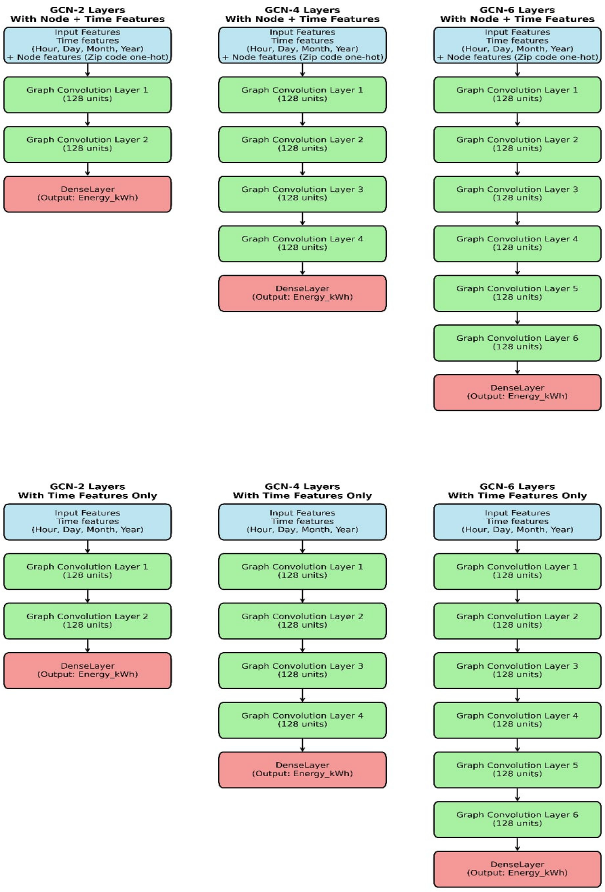

FIGURE 5. GCN models' architecture.

VI. TRAINING AND OPTIMIZATION

A. LSTM

Supervised learning sequences were constructed with a fixed input window of 1 hour and 24 hours; at each sample, the model consumed a length-1 or length-24 sequence of feature vectors and produced a one-step-ahead forecast of Energy_kWh at t+1. Models were optimized to minimize mean squared error (MSE) using the Adam optimizer with a constant learning rate of 0.005. We monitored RMSE and MAE as auxiliary metrics during training. Each training run used a batch size of 64 and a maximum of 500 epochs. To regularize training and prevent overfitting, we applied early stopping on the validation loss with patience of 5 epochs and restoration of the best weights. In addition, we used a reduced learning rate in the plateau scheduler in validation loss with a factor of 1/2 and a minimum learning rate of 0.001, enabling adaptive step size reductions when validation improvement plateaued. Within the network, a standard dropout of 0.2 was used after each recurrent block. Generalization performance was estimated via five-fold cross-validation. For each fold, the model was trained on the corresponding training split and validated on the held-out split; loss, RMSE, and MAE were recorded on the validation sequences and summarized as mean ± standard deviation across folds. Training, early-stopping, and learning-rate scheduling decisions were based exclusively on the validation split of each fold.

B. GCN

In the model without spatial inputs, GCN forecasters were trained on an hourly node–time panel with Energy_kWh as the target, using only temporal covariates (Hour, Day, Month, Year) as inputs; no node features were provided to the model. Source records were aggregated by node (Zip_Postal_Code) at hourly resolution, and a complete panel (all nodes × all hours) was constructed. In the model with the spatial inputs, categorical node identifiers were one-hot encoded and concatenated to the calendar covariates (Hour, Day, Month, Year). To prevent leakage, we used a strict temporal split with the earliest 80% of timestamps for training and the latest 20% for validation (no shuffling). Standardization was employed by fitting the StandardScaler exclusively to the training interval and applying it to the validation interval. Only the temporal features were standardized, while the target remained in physical units for loss computation. In the model with node features, one-hot node indicators remained as float32. This ensures that scaling parameters and stopping criteria depend solely on past information. Optimization minimized mean squared error (MSE) using Adam with a fixed learning rate of 0.005. We monitored RMSE and MAE as auxiliary metrics. Each configuration was trained for up to 500 epochs with a batch size of 64. Generalization was promoted via early stopping on validation loss with patience of 10 and best-weight restoration, alongside a ReduceLROnPlateau scheduler on validation loss with factor 1/2, patience of 5, and minimum learning rate of 0.001 to adapt step sizes when progress plateaued. A dropout rate of 0.2 was applied after each graph block. Model selection and all reported metrics (loss, RMSE, MAE) refer to the held-out temporal validation split.

VII. FRAMEWORK

The models were implemented using the TensorFlow library, suited for both sequence modeling with LSTMs and graph-based learning with GCNs.

VIII. RESULTS

The implementation of GCN and LSTM models for energy consumption forecasting enabled a detailed analysis of the key strengths and weaknesses of each approach. The primary objective was to accurately predict energy consumption and the charging demand of electric vehicles.

A. COMPARISON OF LSTM MODELS

We evaluated four LSTM configurations that differed in input context and the inclusion of location information: models using only temporal covariates (Hour, Day, Month, Year) versus models augmented with one-hot ZIP codes ("with node"), each trained with input windows of either 1 hour or 24 hours. Performance was assessed and summarized as mean (± SD) across five folds. Table 1 reports validation loss (MSE), RMSE, and MAE together with the average number of epochs to early stop.

TABLE 1. LSTM performance for hourly energy forecasting using 1 h and 24 h input windows, with and without ZIP (node) features.

Model	Epochs	Loss	RMSE	MAE
LSTM w/o node 1h	9.4	0.0037 (0.0001)	0.0611 (0.0007)	0.0398 (0.0003)
LSTM w/o node 24h	14.4	0.0033 (0.0001)	0.0577 (0.0007)	0.0348 (0.0005)
LSTM w/ node 1h	10.2	0.0034 (0.0001)	0.0585 (0.0006)	0.0367 (0.0005)
LSTM w/ node 24h	13.4	0.0033 (0.0001)	0.0575 (0.0006)	0.0349 (0.0005)

Reported metrics (Loss/MSE, RMSE, MAE) are mean ± SD over 5 folds; Epochs is the average of early stopping.

Across both feature sets, extending the temporal context from 1h to 24h improved accuracy. For the model without node features, RMSE decreased from 0.0611 to 0.0577, a relative reduction of 5.6%, with MAE dropping from 0.0398 to 0.0348, 12.6%. The "with node" variant showed a smaller but consistent gain from longer context, with RMSE improving 1.7% and MAE 4.9%. Although these improvements in RMSE and MAE indicate better predictive accuracy, they were not statistically significant according to the Wilcoxon signed-rank test. However, a significant reduction in the loss metric was observed in favor of the 24-hour forecasts, both with and without node features ($p = 0.03125$), supporting the benefit of incorporating longer temporal context for LSTM performance. The effect of adding ZIP-code indicators depended on the context length. With a 1h window, node features yielded a small numerical improvement over time-only inputs (RMSE improving by 4.3%), while with a 24h window, both variants performed nearly identically. However, statistical testing revealed that these differences were not significant in either case, indicating that the inclusion of node features did not provide a measurable advantage for either temporal configuration. This suggests that temporal information alone is sufficient to capture most of the variability in hourly energy consumption for the studied dataset. Training converged rapidly under early stopping, with average epochs ranging from about 9–10 (1h) to approximately 13–14 (24h), indicating that longer sequences required slightly more iterations but did not introduce instability. These results indicate that longer temporal windows (24h) are easier to predict than 1h windows, and the incremental benefit of node features is only pronounced when the temporal context is short (1h). A visualization of the performance of each model is provided in fig. 6 and fig. 7. In Fig. 6, the performance of the 1-hour predictor is shown without node features (top) and with node features (bottom). The model incorporating node features demonstrates superior performance in certain regions; for instance, at approximately 250 time steps, a spike is correctly captured by the model with node features but not by the model without them. In Fig. 7, a similar pattern is observed for the 24-hour predictors, particularly just before 1750 time steps, where the model with node features predicts higher energy consumption, in contrast to the model without node features.

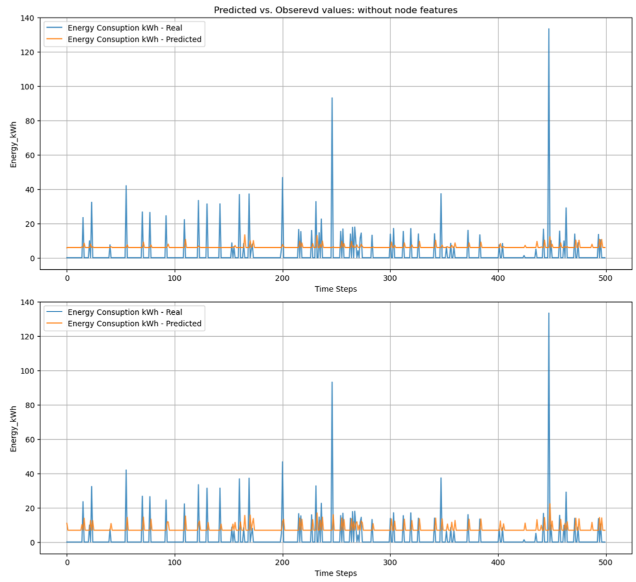

FIGURE 6. Predicted vs. Observed 1h energy consumption on the validation set for the LSTM with and without node features. Blue = ground truth; orange = model prediction.

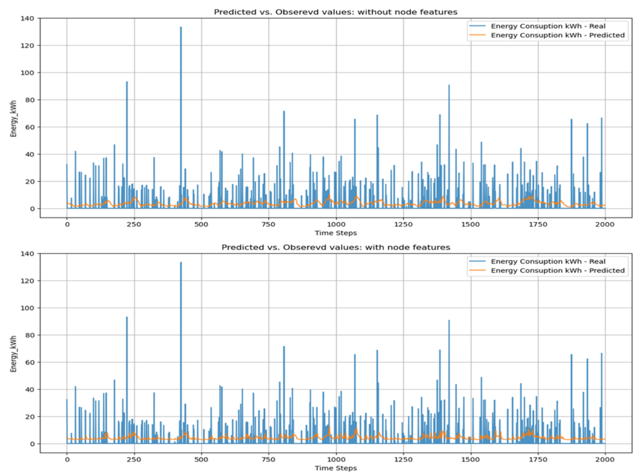

FIGURE 7. Predicted vs. Observed 24h energy consumption on the validation set for the LSTM with and without node features. Blue = ground truth; orange = model prediction.

B. COMPARISON OF GCN MODELS

Different versions of the Graph Convolutional Network (GCN) model were tested to evaluate the robustness of the architecture under varying configurations. The results obtained for each of these different architectures are presented in Table 2.

TABLE 2. GCN performance for hourly energy forecasting, with and without ZIP (node) features.

Model	Epochs	Loss	RMSE	MAE
GCN 2L w/o node	15	23.3725	4.8345	2.7905
GCN 4L w/o node	11	23.3009	4.8271	2.6652
GCN 6L w/o node	14	23.2620	4.8231	2.4587
GCN 2L w/ node	28	22.5858	4.7524	2.6956
GCN 4L w/ node	16	23.2004	4.8167	2.5287
GCN 6L w/ node	24	23.2455	4.8214	2.7269

Reported metrics (Loss/MSE, RMSE, MAE) and Epochs are the ones observed during training.

Without node features, increasing depth consistently reduced absolute error: MAE fell from 2.7905 (2 layers) to 2.4587 (6 layers), an 11.9% reduction, and convergence was faster (11 to 15 epochs). Adding one-hot ZIP features shifted the trade-off: the 2-layer model achieved the lowest RMSE overall and the lowest loss, improving RMSE by approximately 1.7% relative to the best time-only GCN. However, none of these differences in RMSE or MAE were statistically significant according to the Friedman Chi-Square tests. These results favour the simpler 2-layer configuration, supporting the choice of the shallower architecture as a more efficient alternative, balancing predictive accuracy and computational cost. Models with node features tended to require more iterations to converge (16 to 28 epochs), suggesting a modest optimization overhead. If we want to ponder a trade-off between accuracy and lighter computational cost, the GCN with nodes, 2 layers, is the strongest. If the goal prioritizes pointwise absolute accuracy, the GCN without nodes, 6 layers, yields the best results. Across settings, differences in RMSE are small, whereas MAE is more sensitive to depth and the inclusion of node features. Visualizations are provided in Figures 8 and 9.

In Fig. 8, the performance of the three models is presented, showing that they behave very similarly, indicating that a deeper architecture does not necessarily yield better performance.

In Fig. 9, node features are incorporated into all three models, which alters the behavior of the 2-layer model, resulting in improved loss, RMSE, and MAE compared with the 4-layer and 6-layer models.

C. COMPARISON OF LSTM AND GCN MODELS

Direct numeric comparison of error magnitudes between LSTM and GCN models is not meaningful because their

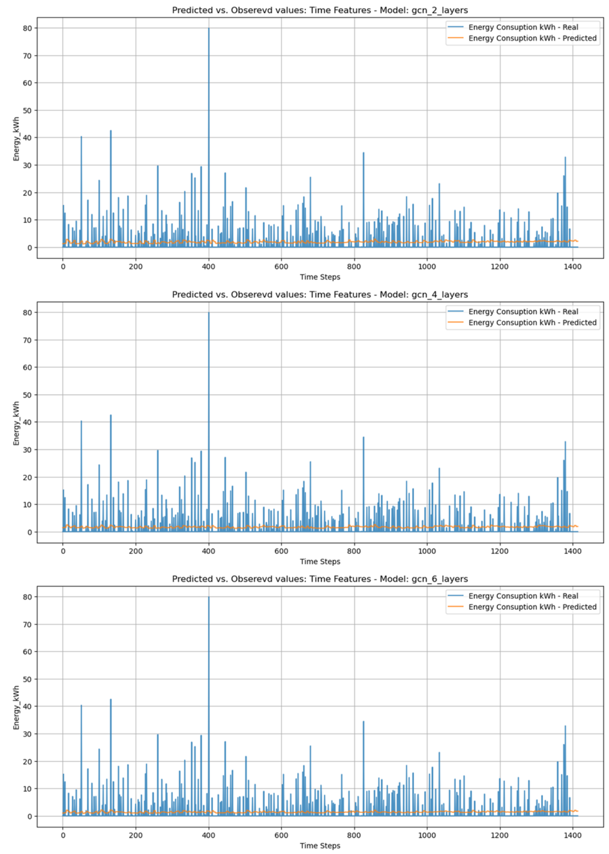

FIGURE 8. Predicted vs. observed energy consumption on the validation set for the GCN models with 2 layers, 4 layers and 6 layers without node features. Blue = ground truth; orange = model prediction.

architectures and implementations are different. We therefore contrast models on their patterns of improvement and modeling trade-offs. For LSTMs, node features were most beneficial when the context was short (1h), but their advantage largely vanished with a 24h history. Similarly, in GCN trained with 1-hour time steps, we observed improvements after adding node features. In our setting, GCNs are markedly more expensive to train than LSTMs. Practically, our GCNs (especially with node features) needed more epochs to converge than LSTMs and showed higher wall-clock time per epoch due to repeated sparse-to-dense-like like multiplications on the graph, poorer GPU/CPU cache locality than batched matrix multiplies in LSTMs, and the need to build and feed full (time $\times$ nodes $\times$ features) tensors together with (nodes $\times$ nodes) adjacencies. Consequently, GCNs deliver competitive accuracy but at significantly higher compute and memory cost relative to LSTMs in this node–time forecasting task.

IX. DISCUSSION

Accurate energy consumption forecasting is essential for smart grid management. This study compares Graph Convolutional Networks (GCN) and Long Short-Term Memory (LSTM) models, discussing their relative performance,

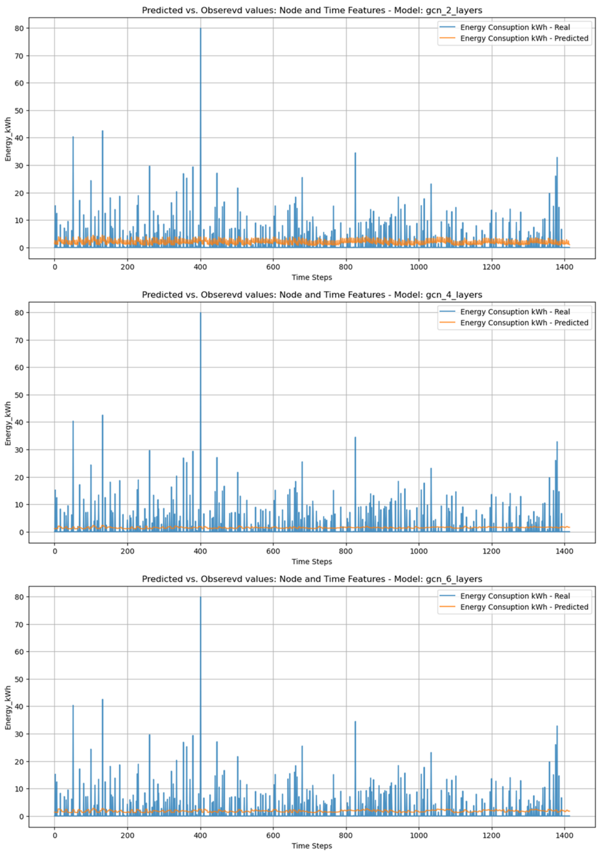

FIGURE 9. Predicted vs. observed energy consumption on the validation set for the GCN models with 2 layers, 4 layers and 6 layers with node features. Blue = ground truth; orange = model prediction.

strengths, and limitations in capturing spatiotemporal consumption patterns.

In our hourly forecasting setting, structuring the data as a graph helped the GCN exploit spatial coupling between charging locations, but its advantage was nuanced. Incorporating ZIP (node) information yielded the lowest RMSE among GCNs with a shallow (2-layer) architecture, indicating that modest message passing can stabilize dispersion errors when locations are correlated. A similar study by Campagne et al., published in 2024 [20], proposed an approach based on Graph Neural Networks (GNNs) for energy consumption forecasting. The authors emphasized that graph-based modeling enables the capture of spatial relationships across different regions, thereby improving predictive accuracy. This approach closely aligns with the GCN model implemented in our study, which also explored connections among spatial and temporal variables. However, Campagne et al. employed GraphSAGE, a variant of GCN that enables adaptive neighborhood learning, whereas our work implemented standard GCN architectures with varying network depths. The main difference between the two studies lies in how the graphs were constructed: in our study, the connections were based on energy consumption similarity, while Campagne et al. utilized decentralized power grid structures to define graph connectivity. The LSTM model, on the other hand, stood out for its ability to capture sequential patterns in energy consumption. LSTMs benefited most from longer temporal context (24h), which largely compensated for the absence of node features and reduced the marginal value of adding location encodings. These results imply that temporal history can substitute for some spatial signal. The study by Hora et al. [21] proposed a model based on LSTM optimized with metaheuristic algorithms for energy consumption forecasting. The authors employed the Butterfly Optimization Algorithm (BOA) to fine-tune LSTM hyperparameters, thereby enhancing its sequential learning capabilities. In comparison with our study, both approaches applied feature selection to optimize LSTM performance. The findings suggest that combining the strengths of both models in a hybrid GCN- LSTM model tends to offer superior performance by combining spatial and temporal learning. In the present study, the models were tested separately; however, the results indicate that a combined approach could further enhance forecasting accuracy. In a study by Vontzos [22], published in 2024, which applied GCN to forecast energy consumption in smart buildings. The authors combined GCN and LSTM models to capture spatial and temporal dependencies, respectively. This hybrid method resembles the current study, where GCN and LSTM models were tested individually. The results reported by Vontzos et al. indicate that combining GCN and LSTM improves forecasting accuracy, particularly in environments with high consumption variability. Another study by R. Ahmad [23] proposes a hybrid model that combines LSTM and GCN for electricity demand forecasting. The authors employ the Wavelet Transform to decompose time series data, thereby enhancing the model's ability to capture consumption patterns. This hybrid approach is similar to the one proposed by Vontzos et al., but differs from the present study, in which GCN and LSTM were evaluated separately. The results reported by Rawal Ahmad indicate that combining GCN for spatial learning with LSTM for temporal learning improves forecasting accuracy compared to using either model individually.

A study by Straub et al. [24] represents an important advance in spatiotemporal forecasting of EV charging demand, extending predictive scope beyond aggregated consumption patterns toward high-resolution mobility modeling. Their approach combines large-scale traffic simulations with detailed variables such as travel distance, speed, car-access attractiveness, parking availability, and route assignment, enabling fine-grained forecasts of when and where charging events are likely to occur in fully electrified urban settings. In contrast, the present work focuses on forecasting hourly energy consumption based solely on observable charging-session data, without access to dynamic traffic or mobility information. While our models (LSTM and GCN) capture temporal and spatial dependencies directly from real-world charging behavior, Straub et al.'s framework integrates exogenous mobility determinants, offering greater predictive granularity but requiring data not typically available in open datasets.

The present study has limitations that inform future work. LSTM metrics were reported on a normalized scale, whereas GCN metrics were in kWh; a unit-consistent evaluation would enable tighter head-to-head comparisons. The GCN pipeline used a 1% subsample for tractability. This decision was taken based on hardware limitations, and future work may rely on more potent hardware. In the present study, only endogenous time and location features were modeled. Exogenous drivers such as weather, tariffs, and events could further enhance accuracy. Cross-validation is unsuitable for GCNs because of temporal or relational leakage. Rolling-origin evaluation over multiple horizons offers a more thorough assessment but demands longer training times.

X. CONCLUSION

Temporal context is the dominant driver of accuracy in LSTMs, while architectural depth (and to a lesser extent node features) governs accuracy in GCNs. If dispersion-focused metrics (RMSE/MSE) are prioritized, a GCN with nodes (2 layers) is preferred. Among LSTMs, the 24 h window is consistently strongest and largely obviates the need for node features. Methodologically, the study challenges the ''deeper-is-better'' intuition around GNNs: moving beyond two convolutional hops does not significantly improve results. Future work should focus on a hybrid GCN–LSTM that couples graph-based spatial encoding with sequence-based temporal modeling, potentially delivering superior accuracy with calibrated computational cost.

REFERENCES

[1] *Eletricity 2025—Analysis and Forecast to 2027*, International Energy Agency, Paris, France, 2025.

[2] R. Poudineh, ''Global electricity demand: What's driving growth and why it matters?'' OIES Energy Comment, Oxford Inst. Energy Stud., Jan. 2025. [Online]. Available: https://www.oxfordenergy.org/wpcms/wp-content/uploads/2025/01/Global-electricity-demand.pdf

[3] *Power Systems in Transition: Challenges and Opportunities Ahead for Electricity Security*, International Energy Agency, Paris, France, 2020.

[4] A. M. Saleh, V. István, M. A. Khan, M. Waseem, and A. N. A. Ahmed, ''Power system stability in the era of energy transition: Importance, opportunities, challenges, and future directions,'' *Energy Convers. Manag., X*, vol. 24, Oct. 2024, Art. no. 100820, doi: 10.1016/j.ecmx.2024.100820.

[5] J. Kim, H. Kim, H. Kim, D. Lee, and S. Yoon, ''A comprehensive survey of deep learning for time series forecasting: Architectural diversity and open challenges,'' *Artif. Intell. Rev.*, vol. 58, no. 7, pp. 1–95, 2025, doi: 10.1007/s10462-025-11223-9.

[6] V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, ''A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks,'' *Future Internet*, vol. 15, no. 8, p. 255, Jul. 2023, doi: 10.3390/fi15080255.

[7] Y. Dou, S. Tan, and D. Xie, ''Comparison of machine learning and statistical methods in the field of renewable energy power generation forecasting: A mini review,'' *Frontiers Energy Res.*, vol. 11, Jul. 2023, Art. no. 1218603, doi: 10.3389/fenrg.2023.1218603.

[8] S. M. Al-Selwi, M. F. Hassan, S. J. Abdulkadir, A. Muneer, E. H. Sumiea, A. Alqushaibi, and M. G. Ragab, ''RNN-LSTM: From applications to modeling techniques and beyond—Systematic review,'' *J. King Saud Univ. Comput. Inf. Sci.*, vol. 36, no. 5, Jun. 2024, Art. no. 102068, doi: 10.1016/j.jksuci.2024.102068.

[9] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, ''A survey on long short-term memory networks for time series prediction,'' *Proc. CIRP*, vol. 99, pp. 650–655, Jan. 2021, doi: 10.1016/j.procir.2021.03.088.

[10] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, ''A review of graph neural networks and their applications in power systems,'' 2021, *arXiv:2101.10025*.

[11] D. Beinert, C. Holzhüter, J. M. Thomas, and S. Vogt, ''Power flow forecasts at transmission grid nodes using graph neural networks,'' *Energy AI*, vol. 14, Oct. 2023, Art. no. 100262, doi: 10.1016/j.egyai.2023.100262.

[12] J. Shanmuganathan, A. A. Victoire, G. Balraj, and A. Victoire, ''Deep learning LSTM recurrent neural network model for prediction of electric vehicle charging demand,'' *Sustainability*, vol. 14, no. 16, p. 10207, Aug. 2022, doi: 10.3390/su141610207.

[13] M. Chang, S. Bae, G. Cha, and J. Yoo, ''Aggregated electric vehicle fast-charging power demand analysis and forecast based on LSTM neural network,'' *Sustainability*, vol. 13, no. 24, p. 13783, Dec. 2021, doi: 10.3390/su132413783.

[14] S. Wang, C. Zhuge, C. Shao, P. Wang, X. Yang, and S. Wang, ''Short-term electric vehicle charging demand prediction: A deep learning approach,'' *Appl. Energy*, vol. 340, Jun. 2023, Art. no. 121032, doi: 10.1016/j.apenergy.2023.121032.

[15] Y. Kim and S. Kim, ''Forecasting charging demand of electric vehicles using time-series models,'' *Energies*, vol. 14, no. 5, p. 1487, Mar. 2021, doi: 10.3390/en14051487.

[16] P. Aduama, Z. Zhang, and A. S. Al-Sumaiti, ''Multi-feature data fusion-based load forecasting of electric vehicle charging stations using a deep learning model,'' *Energies*, vol. 16, no. 3, p. 1309, Jan. 2023, doi: 10.3390/en16031309.

[17] S. Wang, A. Chen, P. Wang, and C. Zhuge, ''Predicting electric vehicle charging demand using a heterogeneous spatio-temporal graph convolutional network,'' *Transp. Res. C, Emerg. Technol.*, vol. 153, Aug. 2023, Art. no. 104205, doi: 10.1016/j.trc.2023.104205.

[18] S. R. Fahim, R. Atat, C. Kececi, A. Takiddin, M. Ismail, K. R. Davis, and E. Serpedin, ''Forecasting EV charging demand: A graph convolutional neural network-based approach,'' in *Proc. 4th Int. Conf. Smart Grid Renew. Energy (SGRE)*, Jan. 2024, pp. 1–6, doi: 10.1109/SGRE59715.2024.10428726.

[19] R. Zhang, ''GCN-TRN: Efficient transformer based electric vehicle charging demand forecasting system,'' in *Proc. 5th Int. Conf. Comput. Sci. Softw. Eng.*, Oct. 2022, pp. 527–533, doi: 10.1145/3569966.3570101.

[20] E. Campagne, Y. Amara-Ouali, Y. Goude, and A. Kalogeratos, ''Leveraging graph neural networks to forecast electricity consumption,'' 2024, *arXiv:2408.17366*.

[21] S. K. Hora, R. Poongodan, R. P. de Prado, M. Wozniak, and P. B. Divakarachari, ''Long short-term memory network-based metaheuristic for effective electric energy consumption prediction,'' *Appl. Sci.*, vol. 11, no. 23, p. 11263, Nov. 2021, doi: 10.3390/app112311263.

[22] G. Vontzos, V. Laitsos, A. Charakopoulos, D. Bargiotas, and T. E. Karakasidis, ''Estimating spatio-temporal building power consumption based on graph convolution network method,'' *Dynamics*, vol. 4, no. 2, pp. 337–356, May 2024, doi: 10.3390/dynamics4020020.

[23] K. Rawal and A. Ahmad, ''Hybrid LSTM-graph convolutional neural network with wavelet transform and correlation analysis for electrical demand forecasting,'' *Social Netw. Comput. Sci.*, vol. 5, no. 4, p. 412, Apr. 2024, doi: 10.1007/s42979-024-02765-w.

[24] Z. Guan, P. Li, Y. Wen, Y. Du, and G. Wang, ''Bubble energy harvesting suitable for weak gas sources using bubble stream release scheme,'' *Appl. Energy*, vol. 349, Nov. 2023, Art. no. 121620, doi: 10.1016/j.apenergy.2023.121620.

MAHER ALARAJ (Member, IEEE) received the Ph.D. degree in machine learning from Brunel University of London. He is currently an Associate Professor with the College of Technological Innovation, Zayed University. He is also an IBM-Certified Artificial Intelligence Analyst. Professionally, he has worked in the consulting sector, focusing on the development and validation of data-driven decision-support models aligned with industry standards and best practices. In academia, he has contributed to various projects in intelligent systems, financial modeling, sustainability, and healthcare analytics across U.K. and UAE. He has been an active researcher; his work can be seen in prestigious journals and international conferences, such as *Journal of Big Data*, *Expert Systems with Applications*, and *Annals of Operational Research*. Similarly, he is a Regular Reviewer for leading journals, such as *Expert Systems with Applications* and *Knowledge-Based Systems*.

CARLA MARTINS received the B.Sc. degree in health technology. She is pursuing the Ph.D. degree in public health with the University of Porto. She completed advanced training in biostatistics, clinical research, and research methods through ISPUP and EAACI. Since 2010, she has worked as a Senior Respiratory Physiologist at Centro Hospitalar Universitário de São João/ULS de Sao Joao, specializing in respiratory assessment and airway inflammation. As an Integrated Researcher at ISPUP, she contributes to the Generation 21 cohort study. She also lectures on Allergology and Pulmonology in the Sports Medicine postgraduate program at the University of Porto and for the Portuguese Society of Sports Medicine. In addition, she develops AI-based educational content for Korbit Technologies, focusing on machine learning and data science. Her research interests include machine learning applications in public health, particularly in predictive and decision-support models. She has published and presented widely, earning several Travel Grants at international conferences.

MOHAMMED RADI (Member, IEEE) received the M.Sc. degree (Hons.) in sustainable electrical power and the Ph.D. degree in electrical power systems from Brunel University of London, U.K., in 2012 and 2017, respectively. He was a Senior Researcher with Brunel University of London, working on the EU Horizon 2020 TDX-ASSIST project, focusing on the coordination of transmission and distribution data exchanges for renewable energy integration. He is currently a Network Data Modeling Engineer (CIM) with UK Power Networks (UKPN), where he works on various projects addressing network data interoperability. He is also a Network Data Modeling Expert with a strong background in power systems and smart grid technologies. His research interests include power systems, the common information model (CIM), renewable and EV integration, smart power networks, and the application of machine learning and artificial intelligence in energy systems. He is a member of CIGRE, IET, Horizon 2020-Bridge, the Brunel Institute of Power Systems, and the Institute of Energy Futures, U.K.

MOHAMED DARWISH is currently the Course Director for M.Sc. in advanced electronic and electrical engineering and a Professor in power electronics and systems with the College of Engineering, Design and Physical Sciences, Brunel University of London, U.K. He has more than 35 years of experience in research and teaching in the area of power electronics and power systems. He has supervised more than 30 Ph.D. research students and more than 90 M.Sc. students. He has published more than 120 journal and conference papers within the area of power electronics and renewable energy systems. He was a Principal Investigator for an EPSRC grant (Programmable Tuned Filters for Retrofit Applications) and for a TCS grant (Power Quality Solutions). He has obtained several British Council Newton Researcher Link Workshop Grants; the next one will be in Egypt, U.K.–Egypt enhanced research network in solar energy system, in February 2017. His research interests include active power filters, uninterruptible power supply (UPS) systems, power quality, and renewable energy systems. He is a Chartered Engineer.

MUNIR MAJDALAWIEH received the M.Sc. degree in computer science and applied mathematics from Northeastern University, Boston, MA, USA, and the Ph.D. degree in IT and the E.M.B.A. degree from George Mason University, Fairfax, VA, USA. He is currently the Acting Director with the Institute of Community Engagement and a Professor with the College of Technological Innovation, Zayed University. Prior to joining Zayed University, in 2012, he was with American University of Sharjah for six years and Booz Allen Hamilton, HP, Compaq, and DEC, USA, for more than 22 years. He has been teaching in the undergraduate and graduate programs. He has published several peer-reviewed papers in international journals and conference proceedings. He received several funds to carry out research on several topics related to IS/IT and IS management.