Reconfiguration-Robust Spatiotemporal Learning for Dynamic EV Hosting Capacity Estimation in Distribution Feeders

Md Touhidul Haque
Electrical and Computer Engineering
University of Alberta
Edmonton, Alberta, Canada
mdtouhi1@ualberta.ca

Petr Musilek
Electrical and Computer Engineering
University of Alberta
Edmonton, Alberta, Canada
pmusilek@ualberta.ca

***Abstract*—The rapid growth of electric vehicle charging loads is creating new operational stress in power distribution systems, where additional demand can trigger voltage violations and equipment overloads. To manage this risk, utilities use hosting capacity analysis to assess how much new charging load can be accommodated at each location without violating operating limits. Existing learning-based estimators can provide fast predictions, but many implicitly assume a fixed network configuration and therefore their accuracy degrades when switching operations alter feeder topology during maintenance, restoration, or load balancing. This paper introduces a topology-aware spatiotemporal learning framework that predicts time-varying hosting capacity at the bus level while remaining robust to feeder reconfiguration. The method represents the feeder as a graph with physically meaningful edge descriptors and performs joint spatial-temporal reasoning over a short history window using an edge-conditioned attention mechanism. Experiments on a real-world, smart-meter-informed distribution system under multiple reconfiguration scenarios show that the proposed approach maintains high accuracy on unseen topologies and consistently outperforms topology-agnostic baselines. These results indicate that explicitly modeling feeder topology and equipment limits in data-driven frameworks is essential for reliable, dynamic electric vehicle hosting capacity assessment in adaptive distribution networks.**

***Index Terms*—electric vehicle, hosting capacity, distribution systems, grid reconfiguration, graph neural networks.**

I. Introduction

The adoption of electric vehicles (EVs) is accelerating worldwide. In 2024, global passenger EV sales exceeded 17 million units (over 20% of total car sales), and annual sales are projected to reach 45 million by 2030 [1]. While transport electrification supports decarbonization goals, it also introduces new operational stress in power distribution systems. Unlike conventional loads, EV charging is time-varying, spatially clustered, and correlated with human activity. Uncoordinated charging can cause feeder congestion, accelerate transformer aging, and trigger voltage violations in low- and medium-voltage networks [2]. To mitigate these risks, distribution system operators (DSOs) rely on *hosting capacity* (HC) analysis to quantify how much additional EV charging can be accommodated at each bus without violating operating limits. Accurate and time-resolved EV HC assessment is therefore critical for planning, operational decision-making, and the deployment of smart charging strategies in modern distribution systems (DSs).

EV HC assessment methods are broadly categorized into model-based and data-driven approaches [3], [4]. Model-based techniques generally include deterministic, stochastic, time-series, and optimization-based formulations. Deterministic approaches prioritize speed by evaluating HC under fixed, worst-case snapshots; however, this simplification fails to capture the temporal variability and uncertainty inherent in EV charging demand [5]. To address these dynamic characteristics, stochastic and quasi-static time-series (QSTS) methods incorporate probabilistic load and generation profiles [6]–[8]. While these approaches improve fidelity, they require a large number of repetitive power-flow simulations, rendering them too computationally expensive for real-time operational deployment. Finally, optimization-based methods formulate HC estimation as a constrained maximization problem and can include multiple objectives, but their high algorithmic complexity typically restricts their application to offline planning studies [9].

In contrast, data-driven approaches leverage historical measurements or simulation-generated labels to learn direct mappings from system states to HC. Machine learning (ML) models like support vector machines (SVMs) and deep neural networks (DNNs) enable fast inference [10], [11]. Recurrent architectures, such as recurrent neural network (RNN) and long short-term memory (LSTM) network, further improve accuracy by capturing temporal correlations in load and voltage trajectories [12]. These attributes make data-driven models attractive for operational use. However, most existing data-driven techniques are topology-agnostic. They typically treat the DS as a collection of independent buses or time series, without explicitly modeling feeder topology and electrical connectivity. Consequently, they fail to generalize under grid reconfiguration, where switching operations alter power-flow paths and constraint sensitivities. In practice, grid reconfigurations are common and can occur daily; therefore, a deployable data-driven HC estimator must remain accurate when power-flow paths change while still tracking short-horizon load dynamics at the bus level. Recent advances in graph neural networks (GNNs) allow for the embedding of topology into learning models. Yet, many graph-based architectures lack explicit mechanisms to model the short-term temporal dynamics required for HC estimation at operational time scales. There remains a lack of data-driven models that jointly: (i) encode feeder topology, (ii) capture temporal dynamics via sequence modeling, and (iii) maintain robustness under grid reconfiguration and topology changes. Addressing this gap is essential for enabling reliable EV HC analytics in modern DSs.

This paper presents a spatiotemporal learning framework for dynamic, bus-level EV HC estimation that integrates temporal sequence modeling with topology-aware graph attention. The main contributions are:

- A spatiotemporal learning framework is proposed that outputs time-varying EV HC at bus-level using a short history window, while remaining accurate when feeder topology changes due to switching/reconfiguration.
- A topology-aware graph formulation is introduced that performs joint spatial–temporal message passing on a stacked 3D graph and conditions attention on electrical edge descriptors capturing coupling strength and component limits.
- A robustness evaluation protocol for topology generalization is designed, in which identical demand trajectories are maintained across multiple feeder configurations to isolate topology-induced EV HC shifts.

The paper proceeds as follows. Section II details the proposed method. Section III describes the case study, followed by the results and baseline comparisons in Section IV. Section V concludes and outlines future work. In support of open science, the source code and dataset used in this study are publicly available at https://doi.org/10.5281/zenodo.19491863.

II. Methodology

We predict bus-level dynamic EV HC from a short history of operating snapshots and the active distribution feeder topology. A distribution feeder is modeled as an attributed graph, replicated across a fixed temporal window to form a 3D spatiotemporal graph with spatial and temporal edges. A 3D edge-conditioned graph attention (3D-ECGAT) model propagates information over both edge types, conditioning attention on edge descriptors, and outputs the current-hour EV HC at each bus. An overview of the proposed spatiotemporal 3D-ECGAT for dynamic EV HC prediction in DSs is presented in Fig. 1.

A. Learning Task

A distribution feeder is represented by an attributed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where nodes correspond to buses and edges correspond to physical electrical connections. For each hour t, a node-feature matrix $\mathbf{X}_t \in \mathbb{R}^{B\times F}$ is formed from the operating snapshot, where $B = |\mathcal{V}|$ is the number of buses and F is the node-feature dimension. The target is a bus-level EV HC vector $\mathbf{y}_t \in \mathbb{R}^B$, where each entry denotes the maximum additional EV charging load that can be accommodated at that bus without violating grid operational limits. Given the feeder graph and a W-hour history of snapshots, the objective is to estimate y_t at the current hour.

B. Spatiotemporal 3D Graph Construction

A fixed-length history window of node features is used as input:

$$\mathbf{X}_t^{(W)} = \{\mathbf{X}_{t-W+1}, \dots, \mathbf{X}_t\} \in \mathbb{R}^{W\times B\times F}, \tag{1}$$

where W denotes the window length. To enable joint spatial and temporal reasoning, the input window is converted into a 3D spatiotemporal graph by stacking W copies of the spatial graph and adding temporal edges between consecutive time slices. Each original bus $i \in \mathcal{V}$ is replicated as (i,τ) for $\tau \in \{t-W+1, \dots, t\}$, producing $W \times B$ spatiotemporal nodes.

Two edge sets are used. The *spatial* edge set replicates the physical feeder edges at each time slice, connecting $(i,\tau) \to (j,\tau)$ for all $(i,j) \in \mathcal{E}$. The *temporal* edge set links each bus to itself across time, connecting $(i,\tau) \to (i,\tau+1)$ for all buses and consecutive slices. This construction preserves the original feeder configuration within each time slice, while enabling information to propagate through time via explicit temporal connectivity.

C. Edge Attributes

Each spatial edge $(i,j) \in \mathcal{E}$ is assigned a compact attribute vector $e_{ij} \in \mathbb{R}^{F_e}$ encoding electrical coupling and capacity-relevant limits (e.g., element type, phase count, voltage level, impedance/length, ampacity, and transformer ratings). Continuous attributes are standardized using training statistics. Temporal edges are assigned a lightweight attribute representation (derived from the mean of spatial attributes) with the same dimension F_e to keep the model input consistent. Overall, edge attributes are used to provide physically meaningful cues that govern voltage drop, power transfer, and constraint activation.

D. 3D-ECGAT

The graph attention network (GAT) [13] is selected to address two fundamental challenges in distribution feeder modeling. First, feeder topology can change under reconfiguration, so the model should operate on arbitrary graphs without assuming a fixed adjacency. GAT performs neighborhood-based aggregation and is therefore inductive, which supports evaluation on unseen configurations. Second, power flow is governed by specific line characteristics, not just connectivity. While standard graph convolutional networks or GCNs treat neighbors uniformly, GAT learns to weigh connections differently. By conditioning these attention weights on physical edge attributes (e.g., impedance and thermal limits), the model naturally incorporates grid physical constraints into its predictions.

Building on these advantages, the proposed 3D-ECGAT model operates directly on the spatiotemporal graph described previously. Let $\mathbf{z}_i \in \mathbb{R}^D$ denote the hidden embedding of a spatiotemporal node i (representing a specific bus at a specific time step τ), obtained by applying a learnable input projection

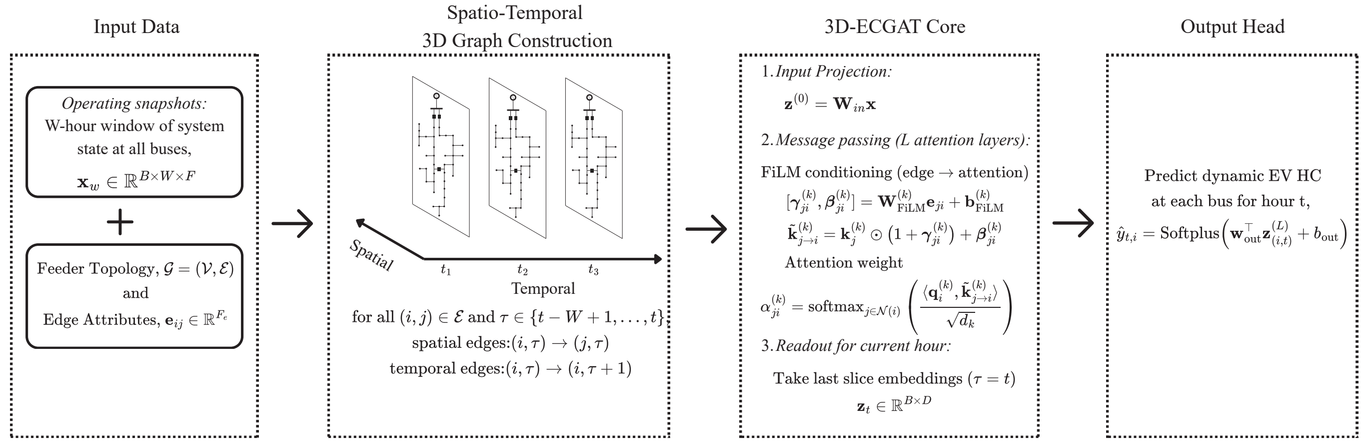


Fig. 1: Overview of the proposed spatiotemporal 3D-ECGAT for EV HC prediction.

to its feature vector. A stack of L attention layers is used to update node embeddings by aggregating information over both spatial and temporal edges.

1) Edge-Conditioned Attention: For a destination node i, messages are received from neighbors $j \in \mathcal{N}(i)$ (including spatial and temporal neighbors). For attention head k, query, key, and value vectors are computed as:

$$\mathbf{q}_i^{(k)} = \mathbf{W}_Q^{(k)}\mathbf{z}_i, \quad \mathbf{k}_j^{(k)} = \mathbf{W}_K^{(k)}\mathbf{z}_j, \quad \mathbf{v}_j^{(k)} = \mathbf{W}_V^{(k)}\mathbf{z}_j, \tag{2}$$

where $\mathbf{W}_Q^{(k)}, \mathbf{W}_K^{(k)}, \mathbf{W}_V^{(k)} \in \mathbb{R}^{d_k \times D}$ are learnable weight matrices.

Edge attributes are used to condition attention through feature-wise linear modulation (FiLM) [14]. For the edge $j \to i$ with attributes $\mathbf{e}_{ji}$, modulation parameters are produced via a linear projection:

$$[\boldsymbol{\gamma}_{ji}^{(k)}, \boldsymbol{\beta}_{ji}^{(k)}] = \mathbf{W}_{\text{FiLM}}^{(k)}\mathbf{e}_{ji} + \mathbf{b}_{\text{FiLM}}^{(k)}, \tag{3}$$

where $\mathbf{W}_{\text{FiLM}}^{(k)}$ and $\mathbf{b}_{\text{FiLM}}^{(k)}$ are learnable weights and biases. The key vector is conditioned as:

$$\tilde{\mathbf{k}}_{j\to i}^{(k)} = \mathbf{k}_j^{(k)} \odot \left(1 + \boldsymbol{\gamma}_{ji}^{(k)}\right) + \boldsymbol{\beta}_{ji}^{(k)}, \tag{4}$$

where $\odot$ denotes element-wise product. The attention weight is computed by scaled dot-product attention:

$$\alpha_{ji}^{(k)} = \text{softmax}_{j\in\mathcal{N}(i)}\left(\frac{\langle \mathbf{q}_i^{(k)}, \tilde{\mathbf{k}}_{j\to i}^{(k)}\rangle}{\sqrt{d_k}}\right), \tag{5}$$

where d_k denotes the dimensionality of the query/key vectors in head k. Messages are aggregated as:

$$\mathbf{m}_i^{(k)} = \sum_{j\in\mathcal{N}(i)} \alpha_{ji}^{(k)}\mathbf{v}_j^{(k)}. \tag{6}$$

The multi-head outputs are concatenated and projected to form the layer output:

$$\mathbf{z}_i' = \mathbf{W}_O\Big[\mathbf{m}_i^{(1)} \,\|\, \ldots \,\|\, \mathbf{m}_i^{(H)}\Big], \tag{7}$$

where $\|$ denotes concatenation, H is the number of heads, and $\mathbf{W}_O \in \mathbb{R}^{D\times(H\cdot d_k)}$ is the output projection matrix. Residual connections and layer normalization are applied to stabilize optimization.

2) Prediction at the Current Hour: After L layers, the embedding of bus i at the last time slice ($\tau = t$) is used for prediction. A node-wise linear projection followed by a Softplus activation is applied to obtain a non-negative EV HC estimate:

$$\hat{y}_{t,i} = \text{Softplus}\Big(\mathbf{w}_{\text{out}}^\top \mathbf{z}_{(i,t)}^{(L)} + b_{\text{out}}\Big), \tag{8}$$

where $\mathbf{w}_{\text{out}}$ and b_{out} are the learnable weight vector and bias of the output projection.

The model is expected to generalize under topology changes because prediction is driven by message passing over the active connectivity of $\mathcal{G}$ and by attention weights conditioned on edge attributes. When the network is reconfigured, the edge set changes and information propagation adapts accordingly. Since coupling strength and constraint sensitivity are encoded in edge attributes (e.g., impedance and ampacity), attention is influenced by physical properties rather than by memorizing a single fixed topology.

III. Case Study and Experimental Setup

A. Test System Description

The proposed approach is evaluated on a smart-meter-informed distribution system from a municipal utility in the Midwest U.S. [15]. The system comprises three primary feeders supplied from a 69/13.8 kV substation and includes 240 buses over approximately 37 kilometers of conductor, serving 1,120 customers via secondary distribution transformers (Fig. 2). Hourly customer energy measurements are available for the full year 2017 and are aggregated at the transformer secondary level (120/240 V). This testbed supports controlled topology-change experiments using documented switching configurations.

B. Dataset Preparation

1) Switching Reconfiguration Scenarios: Distribution feeders are typically reconfigured for load transfer, maintenance,

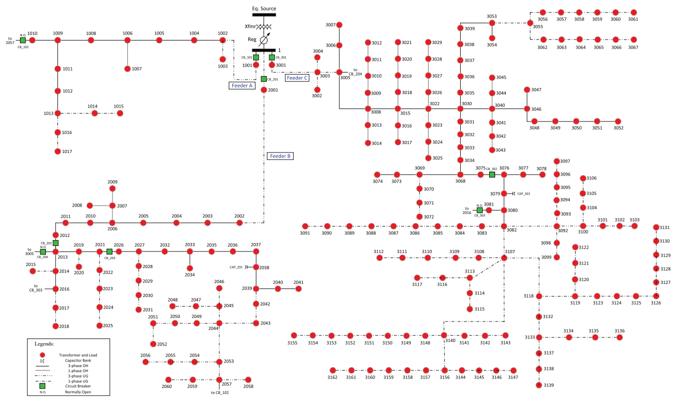


Fig. 2: Iowa distribution system one-line diagram [15].

and service restoration. Such switching changes power-flow paths, voltage drops, and thermal loading, and can therefore shift EV HC across buses even under identical demand profiles. To quantify the sensitivity of EV HC to these changes, we analyze a base configuration and five reconfiguration scenarios (S1–S5) involving feeder sectionalizing and tie-switch operations. All scenarios use identical underlying load time series to isolate the impact of topology changes. Table I details the specific switching operations.

TABLE I: Operational description of the base configuration and five reconfiguration scenarios.

Config.	**Switching description**
Base	Normal operating topology (three radial feeders; inter-feeder ties open).
S1	A–B tie enabled and Feeder A source opened (radial load transfer).
S2	Feeder B source opened; A–B and C–B ties used with sectionalizing to keep the network radial.
S3	Feeder C source opened; B–C and C–B ties used with sectionalizing to preserve radiality.
S4	A–B tie enabled; a loop-breaking sectionalizing switch opened to form a distinct radial topology.
S5	A–B and C–B ties enabled with complementary sectionalizing to form another feasible radial topology.

The complexity of these topological shifts is assessed by quantifying the deviation of the ground truth EV HC relative to the base case. Two metrics are utilized: the normalized mean absolute deviation (nMAD), representing the global average shift, and the normalized root mean squared deviation (nRMSD), by which volatility and localized extrema are captured. Scenarios are grouped into three difficulty tiers based on these deviations (Table II).

TABLE II: Quantitative complexity and difficulty level of reconfiguration scenarios relative to the base case.

Config.	**nMAD (%)**	**nRMSD (%)**	**Difficulty level**
S1	1.3	11.8	Low
S2	3.2	34.1	High
S3	2.1	15.7	Medium
S4	1.2	11.3	Low
S5	3.8	25.1	High

2) Graph Formation and Node Feature Generation: Each feeder configuration s is modeled as an attributed graph $G^{(s)} = (V, E^{(s)})$ with fixed buses V and scenario-specific edges $E^{(s)}$; physical connections are represented as bidirectional edges with standardized attributes. Hourly QSTS power-flow simulations are executed in OpenDSS to produce aligned operating snapshots and node features (e.g., per-phase minimum voltage magnitude and net active/reactive power injections), yielding $X \in \mathbb{R}^{T \times B \times F}$.

3) Ground Truth EV HC Label Generation: Ground-truth EV HC labels are generated via an integrated capacity analysis procedure [16]. For each hour t and each load bus i, EV demand is increased at the target bus while holding other loads fixed, and the maximum feasible increment is recorded. Feasibility is determined by enforcing nodal voltage bounds (0.95–1.05 p.u.) and distribution transformer thermal loading constraints. Repeating this process yields a scenario-specific label matrix $Y^{(s)} \in \mathbb{R}^{T \times B}$ for $s \in \{\text{Base}, S1, \dots, S5\}$.

Representative seasonal variability in load, voltage, and EV HC is illustrated in Fig. 3 using four weeks (winter/spring/summer/fall) at a representative bus, highlighting diurnal patterns and seasonal shifts.

C. Training Strategy

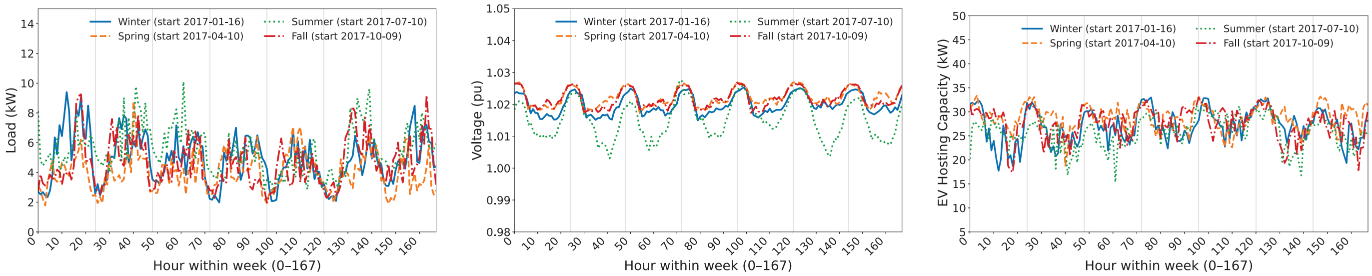
The 3D-ECGAT model is trained to map a W-hour window and the active topology to current-hour, bus-level EV HC. To test generalization, training uses Base, S1, and S2; validation uses S3 (unseen); and testing uses S4 and S5 (unseen). Within the year, data are split into contiguous train/validation/test blocks with a gap of length W to prevent temporal leakage. Optimization is performed using Adam optimizer with a robust regression loss, and early stopping is applied based on validation loss.

IV. Results and Discussion

A. Performance and Sensitivity Analysis

To rigorously evaluate the proposed 3D-ECGAT model, four performance metrics are utilized. The normalized mean absolute error (nMAE) is employed to quantify the average magnitude of prediction errors as a percentage of the true EV HC, facilitating consistent comparison across buses with varying capacity baselines. The normalized root mean squared error (nRMSE) is included to penalize larger deviations more heavily; a significant divergence between nMAE and nRMSE typically indicates the presence of outliers or localized volatility. The coefficient of determination (R^2) is calculated to assess the proportion of variance in the ground truth explained by the model, serving as an indicator of goodness-of-fit. Finally, the over-prediction rate (Over) is reported to quantify the frequency with which the model estimates a higher EV HC than the ground truth, which is a critical safety metric for grid operators.

Table III summarizes performance across Base and reconfiguration scenarios. High predictive accuracy is maintained across all topologies. In the Base case, an R^2 of 0.974 and an nMAE of 6.10% are achieved. Crucially, robustness is



(a) Load demand (kW)

(b) Voltage magnitude (p.u.)

(c) Ground truth EV HC (kW)

Fig. 3: Seasonal variability at representative Bus 3147 (Feeder C) over a 168-hour week. Profiles are shown for (a) load demand, (b) voltage magnitude, and (c) ground-truth EV HC.

demonstrated on the unseen test scenarios. For S4 (low difficulty), performance ($R^2 = 0.971$, nMAE = 6.97%) is nearly identical to the base case. Even for S5 (high difficulty), where the network topology undergoes significant modification, the correlation remains high ($R^2 = 0.971$) and the error remains bounded (nMAE $= 8.58\%$). This indicates that the topology attributes governing EV HC are successfully generalized by the graph attention mechanism.

TABLE III: Performance metrics for base and reconfiguration scenarios (test block).

Config.	nMAE (%)	nRMSE (%)	R^2	Over (%)
Base	6.10	28.03	0.974	21.0
S1	7.47	32.30	0.966	22.2
S2	8.06	34.38	0.961	24.3
S3	7.92	30.24	0.970	24.1
S4	6.97	29.58	0.971	22.4
S5	8.58	30.02	0.971	33.0

A noticeable divergence is observed in Table III between the nMAE (approx. 6–8%) and the nRMSE (approx. 28–34%). This discrepancy is elucidated by the spatial distribution of the MAE for Scenario 5, plotted in Fig. 4. The heatmap reveals that the error is highly localized: the vast majority of nodes exhibit negligible error (dark regions), while a very small number of "hotspot" nodes exhibits significantly higher deviations. These outliers disproportionately inflate the squared error term in the nRMSE calculation. This sparsity of error is consistent with the non-linear nature of grid constraints, where sensitivity to load changes is acute at specific critical locations rather than uniformly distributed. However, the low global nMAE confirms that for the majority of the system, the proposed method yields highly accurate estimates.

Regarding safety, the over-prediction rate remains within 21–24% for Base/S1–S4, and increases to 33% for S5, suggesting that the most challenging topology concentrates a larger fraction of over-estimation risk in a limited subset of buses. Such conservative behavior is desirable for DSOs, as under-estimation of available capacity is operationally safer than over-estimation, which could lead to operational limit violations.

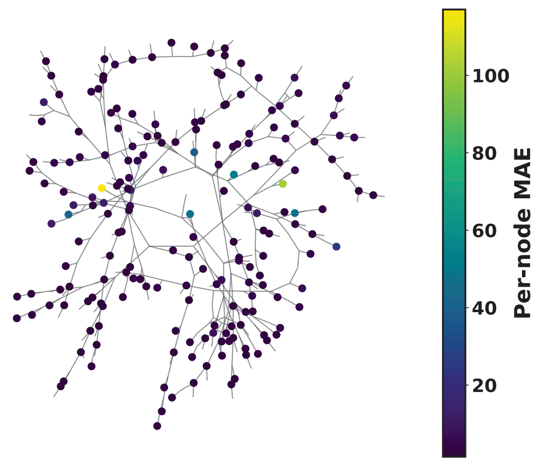


Fig. 4: Spatial distribution of MAE for Scenario 5.

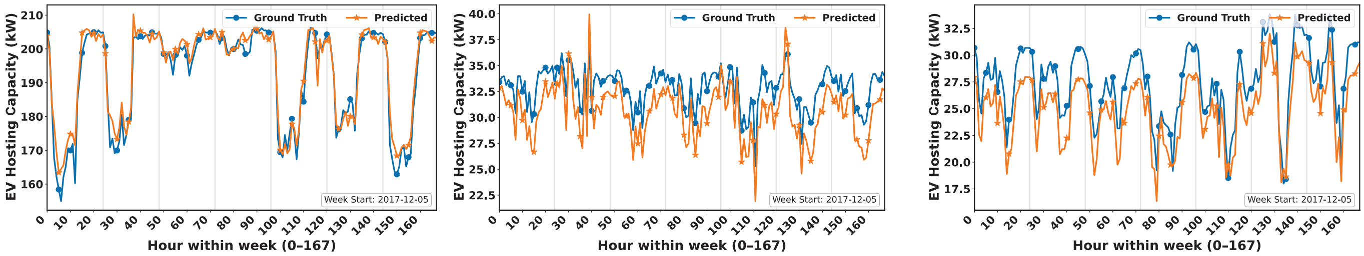
Fig. 5 shows that the model accurately tracks diurnal variability and ramping behavior in reconfiguration scenario 5 across representative buses from three feeders.

B. Comparison with Baselines

To validate the necessity of the proposed topology-aware model, we benchmark 3D-ECGAT against two established data-driven models trained on the same protocol: a DNN and a LSTM network. The DNN serves as a baseline for standard non-linear mapping without temporal memory, while the LSTM represents the state-of-the-art for capturing temporal load dynamics. Neither baseline explicitly models the varying graph topology. Table IV summarizes the performance on the training topology (Base) and two unseen test configurations (S4 and S5) for the DNN, the LSTM, and the proposed 3D-ECGAT model.

TABLE IV: Performance comparison between the proposed 3D-ECGAT and baselines (DNN and LSTM).

Model	nMAE (%)		
	Base (Seen)	S4 (Unseen)	S5 (Unseen)
DNN	16.7	36.1	43.5
LSTM	7.2	12.2	14.2
3D-ECGAT	**6.1**	**7.0**	**8.6**



(a) Bus 1011 on feeder A

(b) Bus 2028 on feeder B

(c) Bus 3147 on feeder C

Fig. 5: Weekly EV HC predictions for three representative buses from three feeders in Scenario 5.

Topology-agnostic learning fails to generalize when physical connections change. The DNN degrades sharply on unseen configurations, indicating poor transfer to new connection patterns. The LSTM preserves correlation by exploiting temporal regularities but its nMAE nearly doubles from 7.2% (Base) to 14.2% (S5). In contrast, 3D-ECGAT achieves the lowest error on all cases and reduces S5 nMAE by about 40% relative to the LSTM, highlighting the benefit of spatial-temporal message passing.

V. Conclusion

This paper presents a topology-aware spatiotemporal learning framework for estimating time-varying electric vehicle (EV) hosting capacity (HC) in distribution systems (DSs) under feeder reconfiguration. The key contribution is the explicit encoding of feeder connectivity using edge-conditioned graph attention (ECGAT) over a stacked short-horizon temporal window, enabling joint spatial–temporal reasoning that adapts to changes in active feeder topology.

Evaluation on a real-world DS with five switching-based reconfiguration scenarios shows that the proposed framework maintains stable accuracy across both seen and unseen topologies and consistently outperforms topology-agnostic baselines. Under the most challenging unseen configuration, the model substantially reduces prediction error relative to a long short-term memory (LSTM) baseline while preserving strong correlation with ground-truth HC values. Error analysis further indicates that residual inaccuracies are concentrated at a small set of electrically sensitive buses, suggesting that targeted refinements could yield additional performance gains.

Future work will focus on (i) uncertainty-aware EV HC estimation with confidence intervals, (ii) integration of distributed energy resources (DERs) and energy storage systems (ESSs), and (iii) safety-aware training objectives that penalize HC overestimation at constraint-critical locations.

Acknowledgment

This research has been supported by the Government of Alberta Major Innovation Fund project MIF-23-003 “Resilient and Clean Energy Systems Initiative”.

References

[1] International Energy Agency, “Global EV outlook 2025,” IEA, Paris, 2025, licence: CC BY 4.0. [Online]. Available: https://www.iea.org/reports/global-ev-outlook-2025

[2] A. Tayri and X. Ma, “Grid impacts of electric vehicle charging: A review of challenges and mitigation strategies,” *Energies*, vol. 18, no. 14, p. 3807, 2025.

[3] V. Umoh, I. Davidson, A. Adebiyi, and U. Ekpe, “Methods and tools for pv and ev hosting capacity determination in low voltage distribution networks—a review,” *Energies*, vol. 16, no. 8, p. 3609, 2023.

[4] H. H. Mousa, K. Mahmoud, and M. Lehtonen, “A comprehensive review on recent developments of hosting capacity estimation and optimization for active distribution networks,” *IEEE Access*, 2024.

[5] J. Then, A. P. Agalgaonkar, and K. M. Muttaqi, “Hosting capacity of an australian low-voltage distribution network for electric vehicle adoption,” *IEEE Transactions on Industry Applications*, vol. 60, no. 2, pp. 2601–2610, 2023.

[6] N. Qammar, A. Arshad, R. J. Miller, K. Mahmoud, and M. Lehtonen, “Probabilistic hosting capacity assessment towards efficient pv-rich low-voltage distribution networks,” *Electric Power Systems Research*, vol. 226, p. 109940, 2024.

[7] M. Al-Saffar, S. Zhang, A. Nassif, and P. Musilek, “Assessment of photovoltaic hosting capacity of existing distribution circuits,” in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*. IEEE, 2019, pp. 1–4.

[8] M. T. Haque and P. Musilek, “Evaluation of distributed solar pv and ev hosting capacities of standard distribution networks for off-grid remote communities in canada,” in *2025 IEEE International Humanitarian Technology Conference (IHTC)*, 2025, pp. 123–129.

[9] I. A. Soliman, V. Tulsky, H. A. Abd el Ghany, and A. M. Azmy, “Optimizing hosting capacity and spatial deployment of ev charging stations for v2g integration in zonal distribution networks,” *Applied Energy*, vol. 401, p. 126671, 2025.

[10] K. Lee, P. Zhao, A. Bhattacharya, B. K. Mallick, and L. Xie, “An active learning-based approach for hosting capacity analysis in distribution systems,” *IEEE Transactions on Smart Grid*, vol. 15, no. 1, pp. 617–626, 2023.

[11] N. Qammar, A. Arshad, R. J. Miller, K. Mahmoud, and M. Lehtonen, “Machine learning based hosting capacity determination methodology for low voltage distribution networks,” *IET Generation, Transmission & Distribution*, vol. 18, no. 5, pp. 911–920, 2024.

[12] M. T. Islam, M. J. Hossain, M. A. Habib, and M. A. Zamee, “Adaptive hosting capacity forecasting in distribution networks with distributed energy resources,” *Energies*, vol. 18, no. 2, p. 263, 2025.

[13] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio *et al.*, “Graph attention networks,” *stat*, vol. 1050, no. 20, pp. 10–48 550, 2017.

[14] E. Perez, F. Strub, H. De Vries, V. Dumoulin, and A. Courville, “Film: Visual reasoning with a general conditioning layer,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, 2018.

[15] F. Bu, Y. Yuan, Z. Wang, K. Dehghanpour, and A. Kimber, “A time-series distribution test system based on real utility data,” in *2019 North American Power Symposium (NAPS)*. IEEE, 2019, pp. 1–6.

[16] P. Paudyal, S. Ghosh, S. Veda, D. Tiwari, and J. Desai, “Ev hosting capacity analysis on distribution grids,” in *2021 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2021, pp. 1–5.