Discovering EV Charging Site Archetypes Through Few Shot Forecasting: The First U.S.-Wide Study

Kshitij Nikhal
Alpha Grid
Palo Alto, CA 94306
technikhal@alphagrid.ai

Luke Ackerknecht
Alpha Grid
Palo Alto, CA 94306
luke@alphagrid.ai

Benjamin S. Riggan
University of Nebraska-Lincoln
Lincoln, NE, 68588
briggan2@unl.edu

Phil Stahlfeld
Alpha Grid
Palo Alto, CA 94306
phil@alphagrid.ai

Abstract

The decarbonization of transportation relies on the widespread adoption of electric vehicles (EVs), which requires an accurate understanding of charging behavior to ensure cost-effective, grid-resilient infrastructure. Existing work is constrained by small-scale datasets, simple proximity-based modeling of temporal dependencies, and weak generalization to sites with limited operational history. To overcome these limitations, this work proposes a framework that integrates clustering with few-shot forecasting to uncover site archetypes using a novel large-scale dataset of charging demand. The results demonstrate that archetype-specific expert models outperform global baselines in forecasting demand at unseen sites. By establishing forecast performance as a basis for infrastructure segmentation, we generate actionable insights that enable operators to lower costs, optimize energy and pricing strategies, and support grid resilience critical to climate goals.

1 Introduction

The global transportation sector is undergoing a paradigm shift, with electric vehicles (EVs) central to achieving net zero emissions in the United States [1, 2]. Achieving this transition requires not only rapid deployment of charging infrastructure, but also a nuanced understanding of EV charging behavior. Such insights directly inform investment decisions [3], grid stability [4, 5], improve energy management [6] and dynamic pricing strategies [7, 8], and shape policies [9, 10], allowing a reliable and cost-effective transition to electric mobility. Forecasting charging demand is therefore not only a matter of operational efficiency, but also a critical lever for accelerating decarbonization.

However, despite growing interest [11], three persistent challenges remain. First, most prior work relies on small-scale datasets that do not capture the intricate complexity of real-world charging patterns across locations and usage scenarios [12–14]. Second, existing models often treat charging sites in isolation [12], forecast regional demand [15–18], or cluster by proximity [19], unable to model complex temporal and spatial relationships between sites. Third, there is limited understanding of how global knowledge can be transferred to forecast demand at newly deployed sites where there is little to no operational history [20]. Inaccurate demand modeling not only inflates operational costs, but also slows EV adoption—challenges that can be mitigated through informed interventions such as curtailment, peak-shaving through dynamic pricing, and optimized battery dispatch.

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2025.

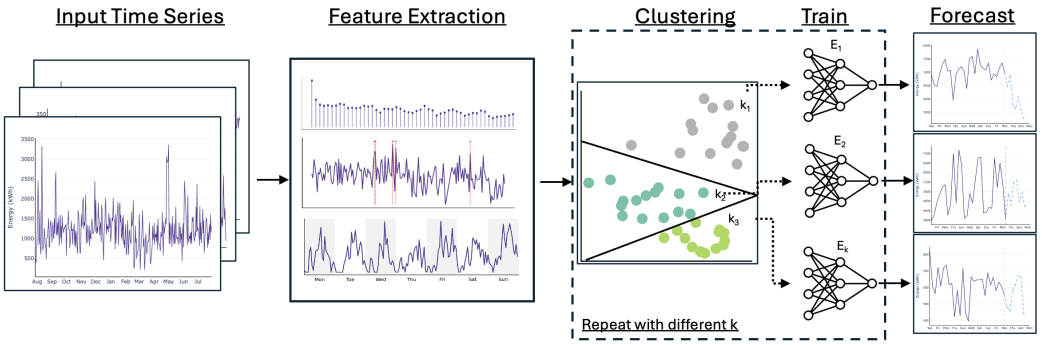

Figure 1: The framework leverages historical demand data to generate features for clustering, with expert models trained on resulting site clusters to forecast demand. This is repeated for different k.

This work leverages a novel industry-level dataset covering most U.S. public Level 3 (DC fast) charging sites, enabling a fine-grained analysis of market behavior and temporal patterns previously unattainable. Building on this foundation, three key contributions are made: (1) Site archetypes are identified by combining clustering with few-shot forecasting, with the number of clusters selected based on predictive performance at unseen sites. (2) Semantic characterizations of the archetypes using a nationwide dataset, linking clusters to geography, surrounding amenities, and usage contexts. (3) Demonstrating knowledge transfer to new sites with limited history, showing that archetype-specific models consistently outperform global models in few-shot forecasting.

2 Methodology

Consider a dataset represented as $X = \{x_1, x_2, \ldots, x_n\}$, where each time series $x_i = \{(t_{ij}, y_{ij})\}_{j=1}^{T_i}$ corresponds to the site i, t_{ij} denotes the calendar date of observation j, and y_{ij} is the total energy sold (in kWh) at the site i on day t_{ij}. The dataset is partitioned into training and test sets, X_{train} and X_{test}, using a standard 80%–20% split. For X_{test}, the historical context is restricted to $L = 28$ days (four weeks) preceding the forecast horizon. For X_{train}, the complete history per site is used (ranging from several weeks to multiple years, depending on availability). The objective is to learn a predictive mapping such that, for each unseen site, the model (or models) produces a forecast:

$$\hat{\mathbf{y}}_{t+1:t+H|t} \;=\; f_\theta(\mathbf{y}_{t-L+1:t})\,, \tag{1}$$

where $y_{t-L+1:t} = (y_{t-L+1}, \ldots, y_t)$ denotes the observed demand in the context window, and f_θ is the forecasting function parameterized by θ. We set the forecast horizon $H = 7$ days to capture both weekday and weekend demand while aligning with weekly cycles in energy management, dynamic pricing, and operational planning. This setup evaluates "few-shot" inference, as lag $L = 28$ is statistically insignificant to capture seasonality and variability in demand.

To achieve this, we characterize site-level demand by aggregating into weekly utilization profiles. In addition, we extract a set of canonical features (e.g., distributional characteristics, autocorrelation structure, and outlier dynamics), summarizing key statistical properties of each time series [21]. Using these representations, we apply k-means [22] to group sites into clusters of similar behavior. For each candidate number of clusters $k \in \{1, \ldots, 20\}$, cluster assignments are used to train a Temporal Fusion Transformer [23] model specialized for that cluster, resulting in a set of forecasting models $\Theta = \{\theta_1, \ldots, \theta_k\}$. The performance of the model is then evaluated across k to identify the optimal clustering. Symmetric Mean Absolute Percentage Error (sMAPE) and Root Mean Squared Error (RMSE) are used to capture complementary aspects of predictive performance, where sMAPE measures relative accuracy across scales, and RMSE emphasizes magnitude and sensitivity to large deviations. Note that similar performance gains are observed with other architectures such as TCN [24] and N-BEATS [25] with varying baseline (or global) performance.

This experimental design enables us to evaluate: (i) generalization to previously unseen sites, (ii) the relative performance gains of cluster-specialized "expert" models compared to a global baseline, and (iii) the appropriate number of clusters (archetypes) that best capture charging behavior. Figure 1 shows an overview of the framework.

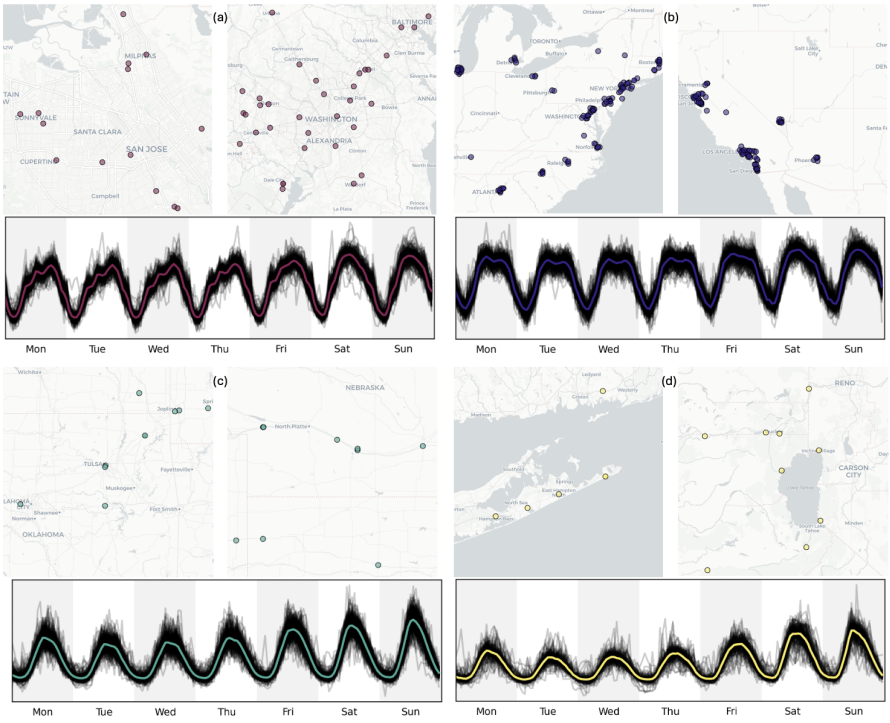

Figure 2: **(a) A8**: Recurring weekday patterns across downtown cores; **(b) A1**: Stable daytime demand in metro areas adjacent to retail chains; **(c) A3**: Utilization concentrated along major highways; **(d) A11**: Elevated weekend demand near leisure hubs such as Lake Tahoe and Hampton Bays. See Appendix A.1 for complete geospatial distribution and demand profiles.

3 Key Results

Dataset: The AGCharging dataset contains session data from 8000 DC Fast Charging (Level 3) sites filtered to the top five charging networks by utilization to capture dominant patterns. The charging sessions are aggregated at the daily level to produce a dataset containing each site's daily energy consumption. Sites with fewer than 35 days of history are excluded, as the model requires a 28-day lag and a 7-day prediction window. To support research, a slice of this dataset will be released.

Forecasting Performance: We evaluate accuracy on unseen sites in X_{test} by first predicting cluster memberships using the k-means predictor. The expert models are then trained on different cluster counts k to determine the optimal granularity.

Figure 3 demonstrates the trade-off between global and expert models. The global model ($k = 1$) provides a strong baseline, while specialized models yield substantial gains up to an optimal $k = 12$. Beyond this point, excessive clustering reduces predictive power, highlighting the importance of balancing cluster size. This suggests that twelve distinct site archetypes capture dominant demand patterns and enable robust few-shot inference.

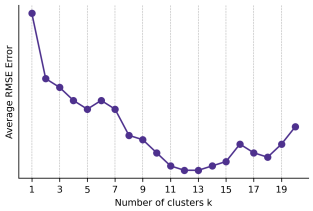

Figure 3: Global $(k = 1)$ vs expert $(k > 1)$ performance.

To interpret these clusters, we attach semantic labels and visualize them on a raster map. Figure 2 shows accurate clustering results with clear behavioral patterns: the downtown sites exhibit weekday ramps driven by commuters, the travel corridor sites cluster together with weekend spikes and high variance, the amenity-adjacent sites show consistent retail-linked usage, and the vacation destination sites display deep weekday troughs and pronounced weekend peaks.

Table 1 summarizes the characteristics of each archetype along with comparative forecast performance. The results show that expert models consistently outperform the global model, confirming that local

Table 1: Archetypes and performance of global (G) vs. expert (E) models. Expert models consistently outperform, especially in volatile clusters. σ denotes deviation from baseline demand.

Archetype (% of sites) and Description	sMAPE		RMSE	
	G	E	G	E
A1. Steady Retail (7%): Stable all-week 9AM-5PM demand ($\uparrow 1.3\sigma$), predominantly located near retail chains in cities such as Los Angeles, Chicago, Denver, and Washington DC.	16.15	**15.82**	**614**	616
A2. Urban Corridors (12%): Large weekend peaks ($\uparrow 2.6\sigma$) co-located with travel plazas near popular destinations and metros.	24.17	**20.76**	728	**654**
A3. Regional Corridors (12%): Similar peaks ($\uparrow 2.1\sigma$) to A2 but lower baseline and higher variance, reflecting sparser corridors.	30.23	**28.60**	638	**608**
A4. Mixed Urban (12%): Blend of commuter and leisure usage ($\uparrow 1.5\sigma$), observed in dense downtowns with diverse land use.	24.77	**22.87**	799	**735**
A5. Balanced Urban (8%): Similar to A1, but cleaner day/night split, moderate peaks ($\uparrow 1.21\sigma$), and deep troughs ($\downarrow 1.75\sigma$).	23.46	**21.80**	650	**608**
A6. Commuter Corridors (15%): Weekday AM/PM ramps exhibit smoother peaks ($\uparrow 1.40\sigma$) compared to sharper weekend peaks ($\uparrow 1.9\sigma$), reflecting work–home travel rhythms.	24.50	**23.23**	614	**575**
A7. Mega Metro (4%): Consistent, saturated ($\uparrow 1.0\sigma$) profile characteristic of large EV markets such as Southern California, reflecting large-scale adoption.	**12.30**	12.91	1126	**1005**
A8. Weekday Ramps (7%): Pronounced weekday ramps with evening plateaus ($\uparrow 1.5\sigma$), seen in San Francisco, Dallas and Seattle.	12.72	**8.69**	1232	**1131**
A9. Suburban Shopping (13%): Spread across most states with midday peaking ($\uparrow 1.55\sigma$), tied to grocery/Big Box stores.	17.86	**17.19**	829	**792**
A10. Emerging Metro (3.5%): Found in Houston, Atlanta, Orlando, and Dallas, where workplace, retail, and leisure usage combine but baseline demand remains volatile.($\downarrow 1.76\sigma$ and $\uparrow 1.21\sigma$).	33.67	**32.89**	682	**664**
A11. Seasonal Leisure (4%): Stronger weekend/holiday surges ($\uparrow 2.3\sigma$) than weekdays ($\uparrow 1.3\sigma$), near resorts and vacation homes.	34.48	**31.73**	762	**705**
A12. Erratic (2.5%): Low baseline demand with irregular peaks ($\uparrow 1.24\sigma$) and troughs ($\downarrow 0.94\sigma$).	65.46	**61.16**	457	**377**

specialization improves predictive accuracy. The improvement is especially pronounced for highly variable sites, such as travel corridors and vacation destinations, where shared temporal patterns (e.g., holiday and event-driven spikes) allow the expert models to generalize effectively.

Overall, these results provide strong evidence that site-level archetypes are not only semantically interpretable, but also predictively useful, demonstrating the value of cluster-aware modeling for scalable forecasting of EV charging demand. See Appendix A.2 for detailed profiles.

4 Conclusion and Future Work

We present the first nationwide assessment of the U.S. public fast-charging market, highlighting utilization patterns and site-level heterogeneity. Using forecast-guided evaluation, we identify the optimal number of archetypes and show that cluster-aware expert models outperform global models in few-shot inference. Beyond accuracy gains, these insights strengthen interconnection studies, guide incentives toward underserved areas, and inform underwriting through archetype membership. Future work will extend to longer-term horizons, soft clustering, and external signals (e.g., events). These directions will further enhance our ability to anticipate and manage the rapidly evolving EV charging ecosystem.

References

[1] Pouria Ahmadi. Environmental impacts and behavioral drivers of deep decarbonization for transportation through electric vehicles. *Journal of cleaner production*, 225:1209–1219, 2019.

[2] Maxwell Woody, Gregory A Keoleian, and Parth Vaishnav. Decarbonization potential of electrifying 50% of us light-duty vehicle sales by 2030. *Nature Communications*, 14(1):7077, 2023.

[3] Zhang Linjuan, Fu Han, Zhou Zhiheng, Wang Shangbing, and Zhang Jinbin. Site selection and capacity determination of charging stations considering the uncertainty of users' dynamic charging demands. *Frontiers in Energy Research*, 11:1295043, 2024.

[4] CH Dharmakeerthi, N Mithulananthan, and Tapan Kumar Saha. Impact of electric vehicle fast charging on power system voltage stability. *International Journal of Electrical Power & Energy Systems*, 57:241–249, 2014.

[5] Lisa Calearo, Andreas Thingvad, Kenta Suzuki, and Mattia Marinelli. Grid loading due to ev charging profiles based on pseudo-real driving pattern and user behavior. *IEEE Transactions on Transportation Electrification*, 5(3):683–694, 2019.

[6] Pei Huang and Zhenliang Ma. Unveiling electric vehicle (ev) charging patterns and their transformative role in electricity balancing and delivery: Insights from real-world data in sweden. *Renewable energy*, 236: 121511, 2024.

[7] Abed Kazemtarghi, Ayan Mallik, and Yan Chen. Dynamic pricing strategy for electric vehicle charging stations to distribute the congestion and maximize the revenue. *International Journal of Electrical Power & Energy Systems*, 158:109946, 2024.

[8] L. Bernard, A. Hackett, R. Metcalfe, L. Panzone, and A. Schein. The impact of dynamic prices on electric vehicle public charging demand: Evidence from a nationwide natural field experiment. Working paper, Centre for Net Zero, 2025. URL `https://cdn.sanity.io/files/lrxd4jqj/production/a597e342e5b87e57dfb22b3ac464d1d5f9b8a18e.pdf`.

[9] Muhammad Shahid Mastoi, Shenxian Zhuang, Hafiz Mudassir Munir, Malik Haris, Mannan Hassan, Muhammad Usman, Syed Sabir Hussain Bukhari, and Jong-Suk Ro. An in-depth analysis of electric vehicle charging station infrastructure, policy implications, and future trends. *Energy Reports*, 8:11504–11529, 2022.

[10] Ying Li, Chris Davis, Zofia Lukszo, and Margot Weijnen. Electric vehicle charging in china's power system: Energy, economic and environmental trade-offs and policy implications. *Applied energy*, 173: 535–554, 2016.

[11] AKM Yousuf, Zhanle Wang, Raman Paranjape, and Yili Tang. An in-depth exploration of electric vehicle charging station infrastructure: A comprehensive review of challenges, mitigation approaches, and optimization strategies. *IEEE access*, 12:51570–51589, 2024.

[12] Tijmen van Etten, Victoria Degeler, Ding Luo, et al. Large-scale forecasting of electric vehicle charging demand using global time series modeling. In *VEHITS*, pages 40–51, 2024.

[13] Ana Martins, João Lagarto, Hiren Canacsinh, Francisco Reis, and Margarida GMS Cardoso. Short-term load forecasting using time series clustering. *Optimization and Engineering*, 23(4):2293–2314, 2022.

[14] Frederik B Hüttel, Filipe Rodrigues, Inon Peled, and Francisco Pereira. Deep spatial temporal forecasting of electrical vehicle charging demand. In *ICML 2021 Workshop on Tackling Climate Change with Machine Learning*, 2021. URL `https://www.climatechange.ai/papers/icml2021/64`.

[15] Henry M Louie. Time-series modeling of aggregated electric vehicle charging station load. *Electric Power Components and Systems*, 45(14):1498–1511, 2017.

[16] Zhiyan Yi, Xiaoyue Cathy Liu, Ran Wei, Xi Chen, and Jiangpeng Dai. Electric vehicle charging demand forecasting using deep learning model. *Journal of Intelligent Transportation Systems*, 26(6):690–703, 2022.

[17] Shijun Wang, Guobin Xue, Chang Ping, Dinggang Wang, Feng You, and Tao Jiang. The application of forecasting algorithms on electric vehicle power load. In *2018 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 1371–1375. IEEE, 2018.

[18] Yunsun Kim and Sahm Kim. Forecasting charging demand of electric vehicles using time-series models. *Energies*, 14(5):1487, 2021.

[19] Marek Miltner, Jakub Zíka, Daniel Vašata, Artem Bryksa, Magda Friedjungová, Ondřej Štogl, Ram Rajagopal, and Oldřich Starý. Towards using machine learning to generatively simulate ev charging in urban areas. In *NeurIPS 2024 Workshop on Tackling Climate Change with Machine Learning*, 2024. URL `https://www.climatechange.ai/papers/neurips2024/22`.

[20] Mamunur Rashid, Tarek Elfouly, and Nan Chen. A comprehensive survey of electric vehicle charging demand forecasting techniques. *IEEE Open Journal of Vehicular Technology*, 2024.

[21] Carl H. Lubba, Sarab S. Sethi, Philip Knaute, Simon R. Schultz, Ben D. Fulcher, and Nick S. Jones. catch22: CAnonical Time-series CHaracteristics. *Data Mining and Knowledge Discovery*, 33(6):1821–1852, November 2019. ISSN 1573-756X. doi: 10.1007/s10618-019-00647-x. URL `https://doi.org/10.1007/s10618-019-00647-x`.

[22] Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

[23] Bryan Lim, Sercan Ö Arık, Nicolas Loeff, and Tomas Pfister. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International journal of forecasting*, 37(4):1748–1764, 2021.

[24] Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.

[25] Boris N Oreshkin, Dmitri Carpov, Nicolas Chapados, and Yoshua Bengio. N-beats: Neural basis expansion analysis for interpretable time series forecasting. *arXiv preprint arXiv:1905.10437*, 2019.

[26] Uber Technologies. H3: A geospatial indexing system. `https://h3geo.org/`, 2025. Accessed: 2025-08-18.

A Appendix

A.1 Cluster spread across regions

Figure 4 visualizes the geospatial distribution of charging site archetypes across the US. Individual sites are aggregated into hexagonal cells [26], and each hexagonal is colored by the dominant utilization cluster within its boundary. The result highlights regional demand patterns, from dense urban corridors to steady retail hubs and emerging suburban clusters. By abstracting sites into cluster-level patterns, the map provides a high-level view of infrastructure heterogeneity that informs siting, forecasting, and pricing strategies.

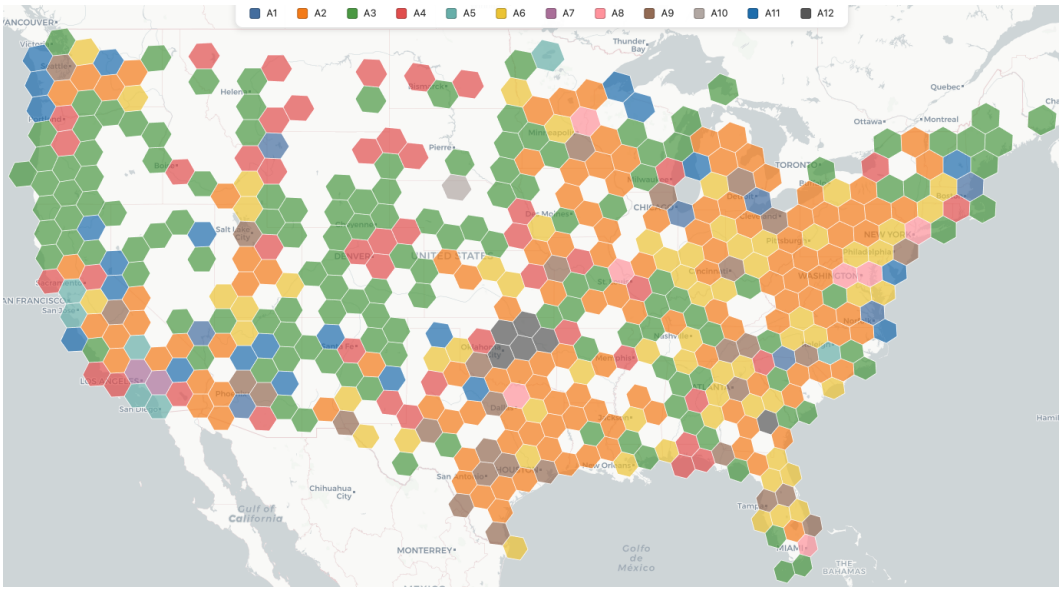

Figure 4: Geospatial distribution of charging site archetypes across the United States. Sites are aggregated into hexagonal cells, with each hexagon colored by the dominant utilization cluster.

A.2 Cluster Demand Profiles

Figure 5 highlights 12 distinct demand profiles that capture the diversity of EV charging behaviors throughout the US. Patterns range from regular weekday commuter flows (e.g., A8 and A10) to volatile and event-driven surges (A11). Stable archetypes such as A1 and A5 exhibit smooth and predictable demand, while clusters like A2, A3, A4, and A6, blend variability from both commuter and leisure segments. At the extremes, A7 reveal unique large-market dynamics, while A9 illustrates widespread midday–oriented demand tied to retail activity. Together, these profiles provide a structured lens on how location type and land-use context shape charging demand.

A.3 Forecasting Performance

The model demonstrates strong generalization across sites with distinct temporal patterns. As seen in Figure 6, the model accurately identifies the timing and shape of peaks, though occasional deviations in magnitude remain. In some cases, the model predicts the correct demand shape while slightly deviating in scale. In particular, the model successfully infers peaks in settings where no such peaks were observed in the historical data, highlighting its transferability and robustness.

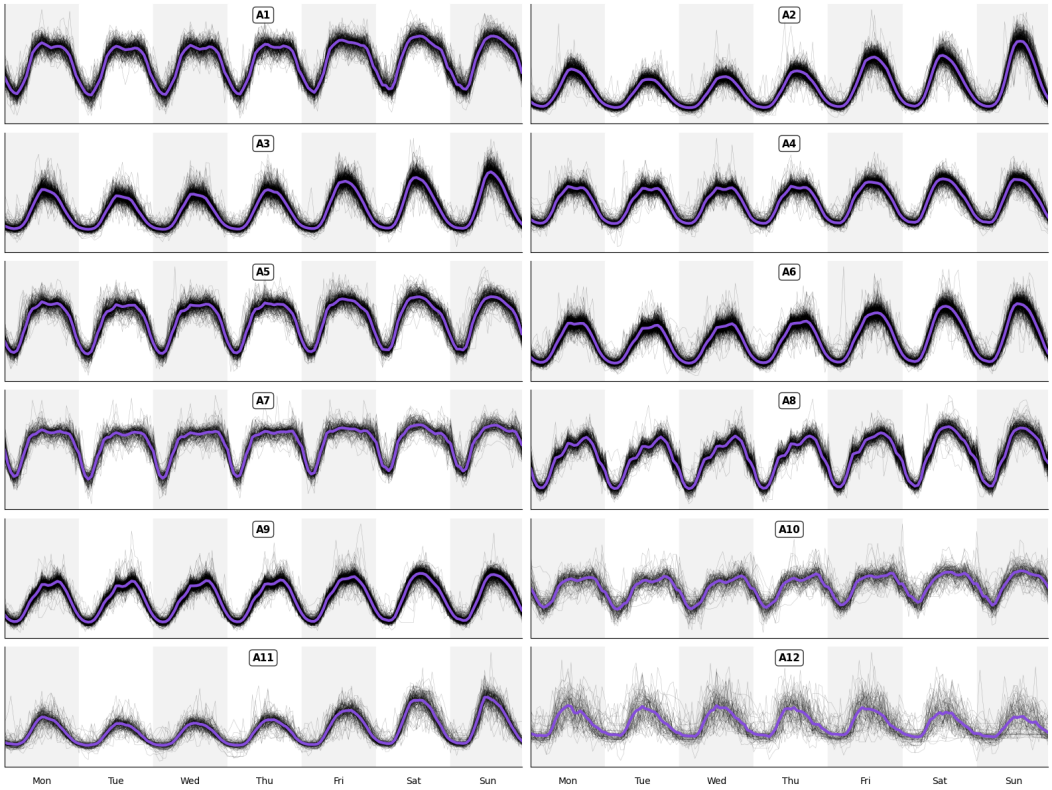

Figure 5: Representative demand profiles for the identified archetypes. Each profile shows the barycenter of clustered sites, illustrating characteristic temporal patterns ranging from stable weekday cycles to highly volatile weekend or tourism-driven surges.

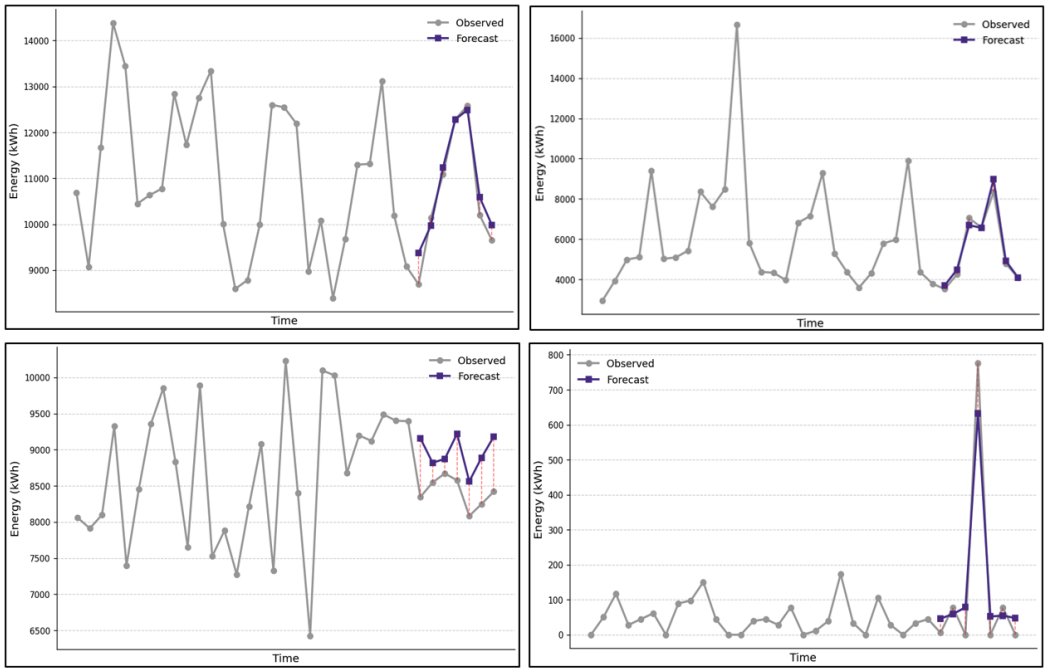

Figure 6: Examples of model forecasts across diverse site profiles.