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

Kshitij Nikhal[1], Luke Ackerknecht[1], Benjamin S Riggan[2], Phil Stahlfeld[1]
Alpha Grid[1], University of Nebraska-Lincoln[2]

MOTIVATION

- Electric vehicles (EVs) are central to achieving net zero emissions in the United States.
- Existing research is limited by small datasets, proximity-based clustering, and poor generalization to newly deployed sites with sparse data, failing to capture temporal and contextual dependencies across diverse usage environments.
- Forecasting EV charging demand informs infrastructure planning, grid stability, dynamic pricing, and climate policy.

KEY CONTRIBUTIONS

- **Nationwide Dataset:** Comprehensive coverage of nearly all U.S. public Level-3 (DC fast) chargers, enabling large-scale, high-resolution analysis of temporal and behavioral patterns.
- **Forecast-Guided Clustering:** Integration of clustering and few-shot forecasting to uncover behaviorally consistent site archetypes.
- **Semantic Interpretation:** Linking archetypes to geography, amenities, and usage context, providing interpretable insights into charging behavior.
- **Knowledge Transfer:** Demonstrating that archetype-specific expert models generalize better than global baselines for new or data-sparse sites.

METHODOLOGY

1. Input time series dataset aggregated at a day level.

$$X = \{x_1, x_2, \ldots, x_n\}$$
$$x_i = \{(t_{ij}, y_{ij})\}_{j=1}^{T_i}$$

2. Characterize site-level demand into weekly utilization profile and catch22 features.

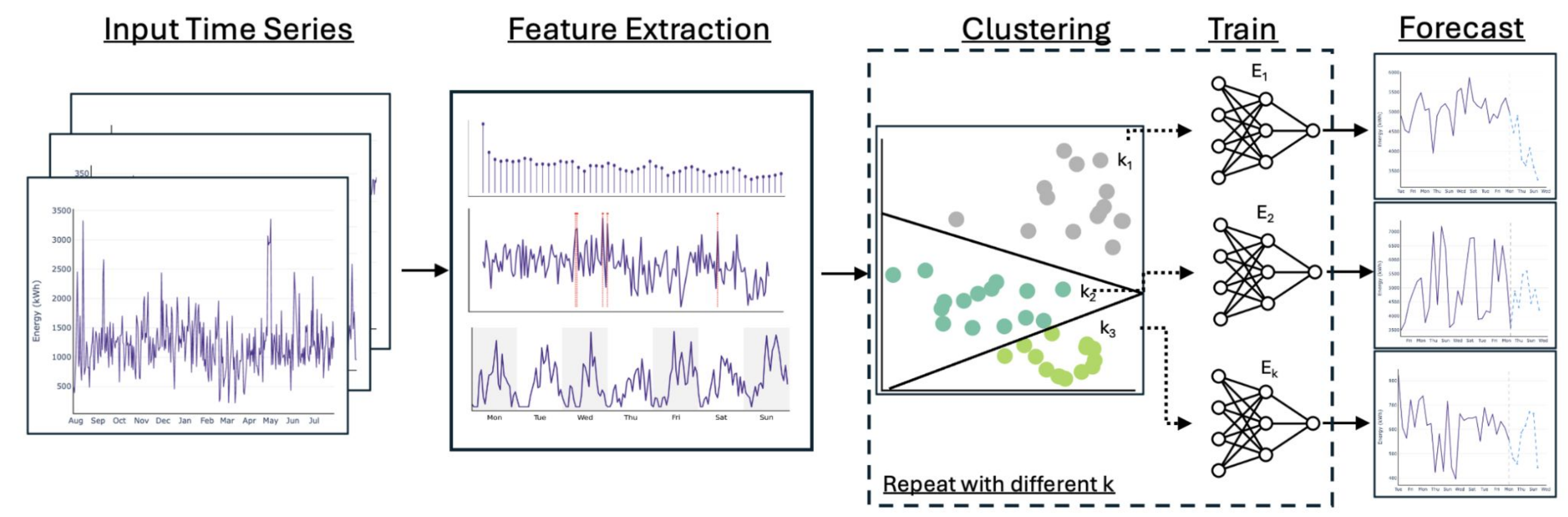
$$\Phi_i = \left[\, u_i,\ \phi(u_i) \,\right]$$

3. Cluster and train k-models for each cluster.

$$\arg\min_{k} \|\Phi_i - \mu_k\|_2^2$$
$$\Theta = \{\theta_1, \ldots, \theta_k\}$$

4. Forecast the next H days using past L days.

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

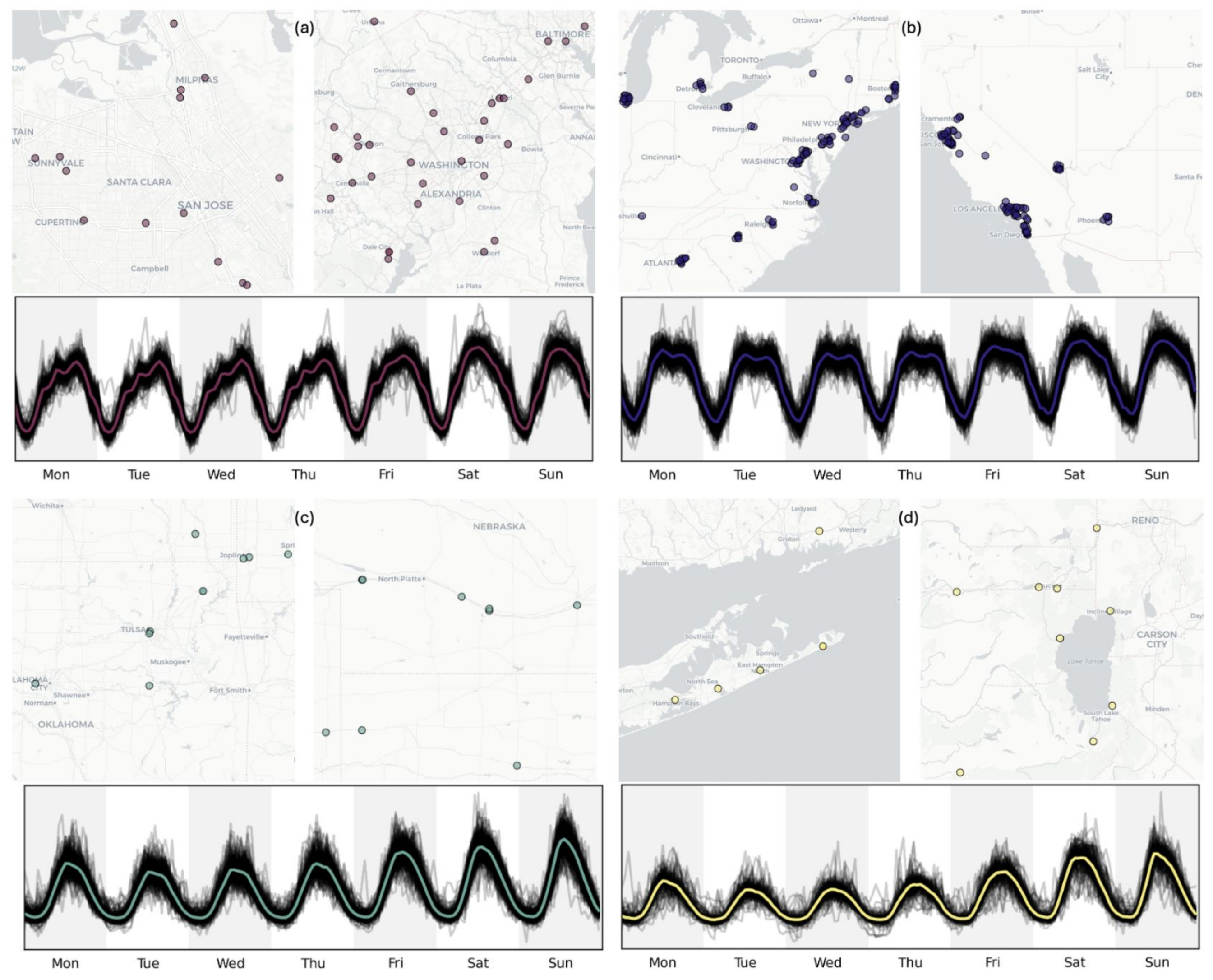


★ The framework leverages historical demand data to generate features for clustering, with expert models trained on resulting site clusters to forecast demand. This is repeated for different k.

DISCOVERED ARCHETYPES

★ Archetype patterns range from regular weekday commuter flows to volatile and event-driven surges.

- (a) Recurring weekday patterns across downtown cores.
- (b) Stable daytime demand in metro areas adjacent to retail chains.
- (c) Utilization concentrated along major highways.
- (d) Elevated weekend demand near leisure hubs such as Lake Tahoe and Hampton Bays.

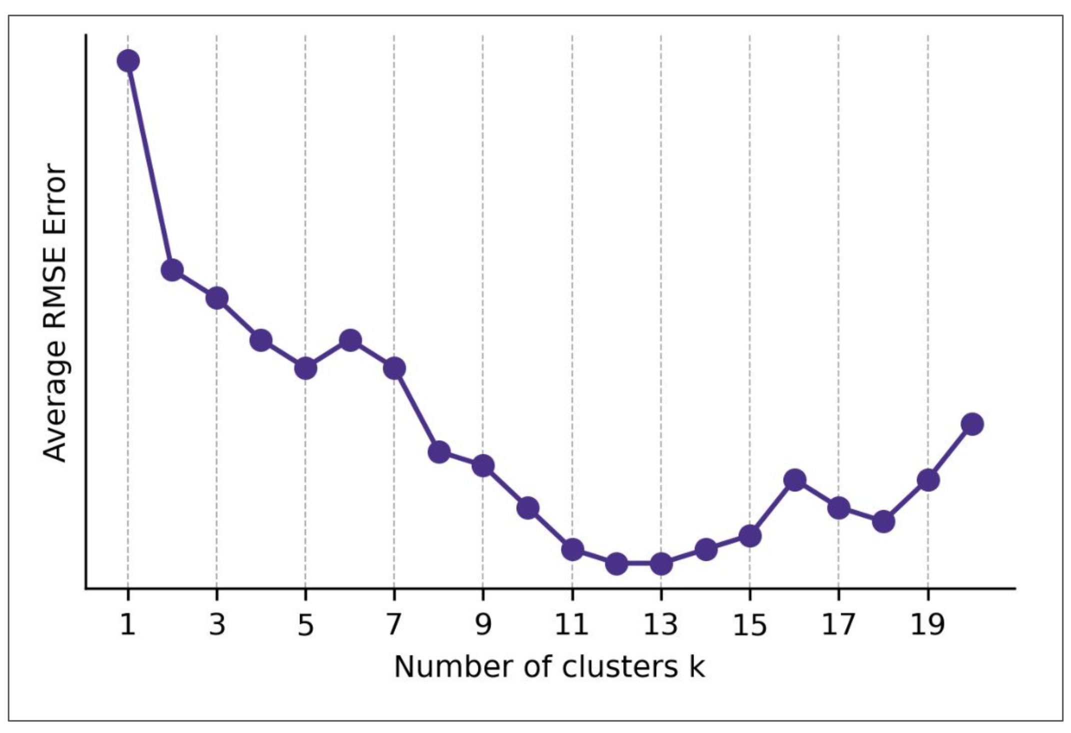


EXPERIMENTAL RESULTS

★ Archetypes and performance of global (G) vs. expert (E) models. Expert models consistently outperform, especially in volatile clusters.

Archetype (% of sites) and Description	sMAPE		RMSE	
	G	E	G	E
A1. Steady Retail (7%): Stable all-week 9AM-5PM demand (↑ 1.3σ), 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 (↑ 2.6σ) co-located with travel plazas near popular destinations and metros.	24.17	**20.76**	728	**654**
A3. Regional Corridors (12%): Similar peaks (↑ 2.1σ) 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 (↑ 1.5σ), 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 (↑ 1.21σ), and deep troughs (↓ 1.75σ).	23.46	**21.80**	650	**608**
A6. Commuter Corridors (15%): Weekday AM/PM ramps exhibit smoother peaks (↑ 1.40σ) compared to sharper weekend peaks (↑ 1.9σ), reflecting work–home travel rhythms.	24.50	**23.23**	614	**575**
A7. Mega Metro (4%): Consistent, saturated (↑ 1.0σ) 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 (↑ 1.5σ), seen in San Francisco, Dallas and Seattle.	12.72	**8.69**	1232	**1131**
A9. Suburban Shopping (13%): Spread across most states with midday peaking (↑ 1.55σ), 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.(↓ 1.76σ and ↑ 1.21σ).	33.67	**32.89**	682	**664**
A11. Seasonal Leisure (4%): Stronger weekend/holiday surges (↑ 2.3σ) than weekdays (↑ 1.3σ), near resorts and vacation homes.	34.48	**31.73**	762	**705**
A12. Erratic (2.5%): Low baseline demand with irregular peaks (↑ 1.24σ) and troughs (↓ 0.94σ).	65.46	**61.16**	457	**377**

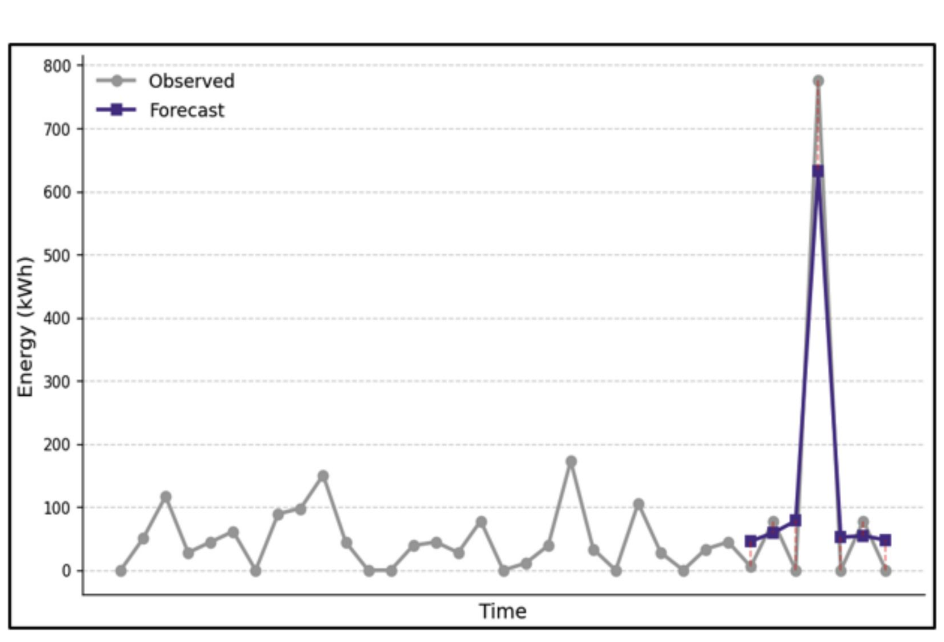
- The global model (k=1) provides a strong baseline, while specialized models yield substantial gains up to an optimal k=12.

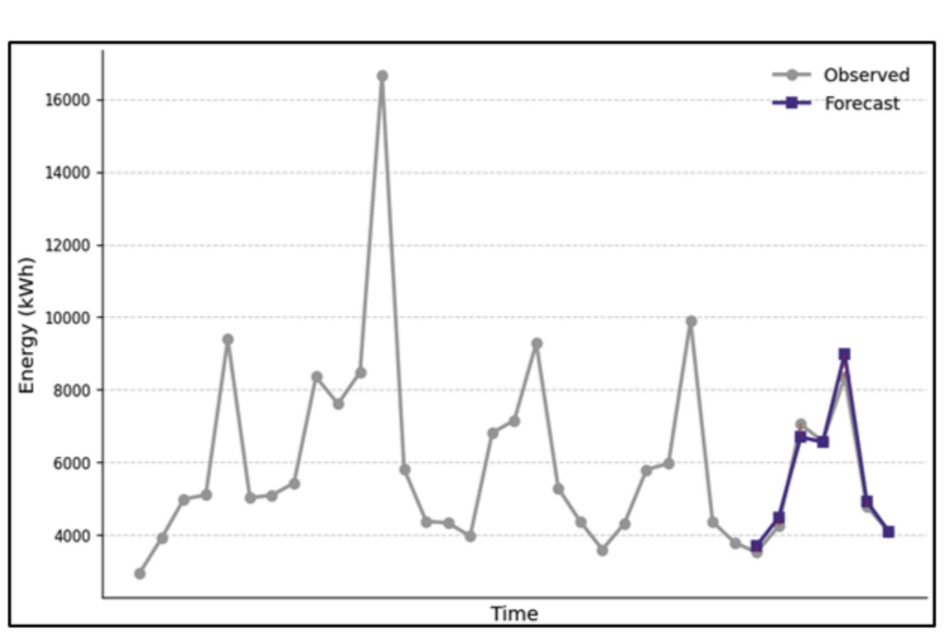


- Beyond this point, excessive clustering reduces predictive power, highlighting the importance of balancing cluster size.
- This suggests that **twelve** distinct site archetypes capture dominant demand patterns and enable robust few-shot inference.

FORECASTING

★ Effectively transfers archetype-specific knowledge to forecast both peak surges and recurring demand trends.





CONCLUSION

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.