Adaptive Smart Parking Prediction: Enhancing Accuracy and Resilience in Dynamic Environments

Michael Craig[1], Carter Dixon[1], Dillon Rose[1], Prajakta Belsare[1]*, and Samy El-Tawab[1]
[1] Computer Science Department, College of Integrated Science & Engineering
James Madison University, Harrisonburg, VA, USA
*Corresponding Author: belsarpp@jmu.edu

***Abstract*—As smart parking systems become more prevalent, they increasingly depend on data-driven prediction models trained on historical data to forecast parking availability. However, many existing implementations operate under the assumption of static model configurations and fixed retraining schedules. This approach can be problematic in dynamic settings like university campuses, where parking behaviors fluctuate frequently due to shifting academic calendars, seasonal changes, and special events. Such variability can lead to diminished prediction accuracy if models do not adapt accordingly. This study investigates smart parking prediction in the presence of both routine variability and disruptive conditions by analyzing parking patterns across space and time. Using over 18 months of high-resolution parking data collected from five campus garages at a mid-atlantic university, we built a three-way LightGBM ensemble that routes predictions based on academic calendar context. The model is trained on a substantially richer feature set than prior work, incorporating cyclical time encodings, zone metadata, peak-hour indicators, historical occupancy statistics, and campus event flags, alongside a seven-step automated cleaning pipeline applied to over 22 million raw records. A spatial Inverse Distance Weighting post-processing layer captures cross-zone demand spillover. Prediction performance is evaluated using MAPE, RMSE, and R^2 across 8.15 million held-out future observations spanning a 10-month test window. This work improves a data leakage flaw present in the prior baseline and demonstrates a 33.6% reduction in MAPE over the prior-year result, with over three-quarters of all prediction errors are within 5 spots, and 91% are within 25 spots. The findings provide practical insights into how academic calendar structure, event-driven demand, and zone-level capacity interact to shape prediction accuracy, and offer a replicable framework for parking prediction in dynamic environments.**

***Keywords*—Smart Parking, Machine Learning, LightGBM, Adaptive Parking Prediction, Intelligent Transportation Systems (ITS), Ensemble Learning.**

I. Introduction

The proliferation of smart city technologies and the exponential growth of urban mobility data have ushered in a new era of intelligent transportation systems. Among the most persistent and pervasive challenges in modern urban and institutional settings is the efficient management of parking resources. Inadequate parking management contributes to increased traffic congestion, elevated fuel consumption, unnecessary carbon emissions, and significant losses in user productivity. These consequences are particularly acute in densely populated academic environments where thousands of commuters converge within narrow time windows dictated by class schedules and institutional events [1], [2]. The rapid advancement of data-driven technologies has, in recent years, facilitated the integration of Machine Learning (ML) into smart parking solutions, offering a promising pathway toward significantly optimizing parking efficiency in both urban and academic contexts [3], [4]. Despite notable progress, parking behavior remains inherently complex. Demand is shaped by dynamic, interdependent factors such as class schedules, special events, time of day, seasonal calendars, and environmental conditions [5], [6]. This complexity renders traditional rule-based and statistical methods inadequate, as they fail to capture the non-linear temporal dependencies and contextual variability of real-world parking patterns [7]. Thus, there is a compelling need for robust, data-driven predictive frameworks capable of modeling these intricate relationships.

Our prior work has explored the application of Machine Learning (ML) techniques for parking occupancy prediction, drawing on data from parking sensors, university parking garages, and IoT-enabled infrastructure [3], [4]. Studies have used a variety of approaches, including regression-based models, tree ensemble methods, and deep learning architectures, each demonstrating varying degrees of predictive performance depending on the characteristics of the target environment and the features used [8], [9].

Despite growing research in parking prediction, critical gaps remain. Few studies target university environments, where parking dynamics differ fundamentally from urban settings due to structured yet frequently disrupted schedules. Additionally, the isolated contributions of temporal metrics, event schedules, and historical occupancy trends are underexplored, and practical deployment considerations such as traffic limitations, inference speed, long-horizon validity, and retraining cadence, are rarely addressed together. This study addresses these gaps by comparing multiple ML models for predicting parking availability in a large university setting, leveraging over 18 months of real-world occupancy data among the most longitudinally extensive campus parking evaluations to date.

Model performance was rigorously assessed using a multi-metric evaluation framework encompassing Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and coefficient of determination (R^2), providing a comprehensive and multidimensional view of predictive fidelity. Building on the inference speed testing and future-date prediction validation conducted in prior work [4], this study focuses on long-horizon generalization under a strict temporal evaluation protocol. Recognizing that campus parking patterns are subject to continuous temporal drift arising from changes in enrollment, scheduling practices, and institutional policy, we also empirically determined the optimal model retraining frequency required to sustain predictive accuracy over extended operational periods, a dimension of applied ML system design that is frequently neglected in academic investigations.

The remainder of this paper is structured as follows: Section II describes the data collection methodology, and preprocessing pipeline; Section III details the ML models and evaluation framework; Section IV reports and discusses experimental results; and Section V concludes with a summary of findings and directions for future research.

II. Data Collection and Pre-processing

Parking occupancy data was collected from five garages at James Madison University (JMU), extending the infrastructure established in prior work [4]. Garages are divided into four zone types: Accessible, Commuter, EV, and Faculty/Staff, each with distinct capacity allocations. Campus traffic is shaped by gate closures along the central corridor (Mon–Thu: 7AM–7PM; Fri: 7AM–1PM), creating asymmetric demand patterns across garages based on proximity to academic buildings. Data was collected via a Python script interfacing with JMU's public parking API, storing records in a MySQL database on AWS at one-minute intervals. Each record captures Zone ID, Timestamp, and Current Availability. Collection spans February 16, 2024 to March 12, 2026, yielding over 22 million raw records across approximately two academic years, enabling inter-annual comparisons and reliable modeling of semester-scale temporal drift. The five garages studied, Ballard, Champions, Chesapeake, Grace Street, and Warsaw, collectively offer over 4,000 parking spaces.

A. Automated Data Cleaning Pipeline

The prior implementation relied on ad hoc manual cleaning [4]. This work introduces a seven-step automated pipeline that reduces 22.4 million rows to 20.2 million clean rows (9.7% removed), with each step targeting specific sensor failure modes identified through systematic audit:

1) **Unknown zone removal:** 718,686 records belonging to Zone 41, which does not correspond to any physical zone in the parking system, were removed.
2) **Deduplication:** 1,434,514 duplicate (Zone, Timestamp) pairs were removed to eliminate repeated sensor writes within the same minute.
3) **Sensor overflow correction:** Zone 6 (Grace Faculty, capacity 55) exhibited values reaching 65,336, a 16-bit unsigned integer wraparound artifact, affecting 20,483 records. These unrecoverable values were dropped. Zone 6 experienced complete sensor failure on July 18, 2025 and all subsequent Zone 6 records were excluded from the test set.
4) **Negative value correction:** 202,089 records reporting negative availability were clamped to zero.
5) **Over-capacity correction:** 1,867,850 records exceeding the documented zone maximum capacity were capped at the respective zone ceiling.
6) **Maintenance artifact removal:** 7,290 records corresponding to systematic 4:00 AM sensor resets were removed.
7) **Sensor spike removal:** 554 isolated minute-to-minute spike records inconsistent with physically plausible occupancy transitions were removed.

Test set cleaning was handled separately to preserve ground truth integrity. Commuter zone zero-availability readings during peak hours (8:00 AM–4:00 PM) were retained as legitimate full-capacity observations. System-wide outages, defined as ten or more zones simultaneously reporting zero availability, were excluded, along with prolonged zero-availability streaks on non-commuter zones and overnight Saturday-to-Sunday sensor resets on large-capacity zones.

III. Feature Engineering and Model Architecture

The feature set was substantially expanded from the six features used in prior work: ZoneID, cyclic time, day of week, month, and a binary event flag [4]. The rebuilt pipeline produces 21 features for non-event sub-models and 36 for the events sub-model, organized into five categories.

Cyclical temporal encodings: Minute-of-day sine and cosine features were retained from the prior work and extended with day-of-year encodings (`doy_sin`, `doy_cos`) to capture seasonal drift across semesters, and week-of-year encodings (`woy_sin`, `woy_cos`) to represent semester-scale rhythm. Separate integer hour and minute features were added, as tree-based models benefit from sharp decision boundaries at operationally significant thresholds such as 8:00 AM and 5:00 PM that smooth continuous encodings alone cannot represent.

Academic calendar phase: A 10-level categorical feature (`campus_phase`) encodes the academic calendar period, distinguishing summer, move-in week, the first two weeks of semester, regular session, fall break, Thanksgiving break, exam week, winter break, and spring break. This replaces the prior binary school-in-session indicator and resolves a documented August transition failure in the predecessor model, which routed late-August rows to the school year sub-model despite campus being largely unoccupied, producing MAPE of 37% during that period.

Zone metadata: Each zone's physical capacity, permit type (Commuter, Faculty, Accessible, EV), and garage identifier (Ballard, Champions, Chesapeake, Grace, Warsaw) were integer-encoded to allow the model to reason about relative fullness and to learn garage-specific occupancy patterns.

Peak hour indicators: A binary `is_peak` flag marking hours 10:00 to 15:00, and a linear ramp feature `hours_since_8am`, were included to support learning of the steep morning fill-up curve and class-change transition dynamics characteristic of campus environments.

Historical statistical lookups: For each combination of Zone, Hour, and Day of Week in the training data, mean and standard deviation of historical availability were computed and merged as `hist_mean` and `hist_std`. A coarser Zone-by-Hour lookup (`hist_mean_zh`, `hist_std_zh`) was also included to provide a more stable prior for day-of-week combinations with sparse training samples.

Event flags: One-hot encoded binary columns derived from the JMU academic and athletic master calendar indicate the presence of specific campus events including home football games, exam periods, orientation, and institutional breaks. These flags feed exclusively into the events sub-model.

A. Ensemble Architecture and Routing

All three sub-models are implemented using LightGBM, a gradient boosting framework employing leaf-wise tree growth, histogram-based binning, and built-in early stopping. LightGBM replaces the KNN and Random Forest Regressor sub-models used in the prior implementation [4], offering iterative residual correction, automatic feature importance weighting, and training throughput approximately 10 to 20 times faster than Random Forest at this data scale.

Each prediction row is routed to one of three sub-models based on academic calendar context:

- **Events sub-model:** Any row where at least one event flag is active (highest priority routing).
- **Summer sub-model:** Non-event rows within May 12 to August 20, 2024, or May 15 to August 19, 2025.
- **School Year sub-model:** All remaining non-event rows.

Hyperparameters for each sub-model were optimized independently using Optuna with Tree-structured Parzen Estimators (TPE) across 100 trials per sub-model, minimizing MAE on a temporal validation split. The search space spanned 300 to 1,500 estimators, learning rates of 0.005 to 0.15 on a log scale, 20 to 400 leaves, tree depths of 4 to 18, and L1 and L2 regularization coefficients from 10^{-8} to 10 (Table I). This represents a substantial methodological advancement over the GridSearchCV approach used in [4], which examined only 8 parameter combinations per model.

TABLE I
OPTUNA HYPERPARAMETER SEARCH SPACE

Hyperparameter	Search Range
`n_estimators`	300 – 1500 (int)
`learning_rate`	0.005 – 0.15 (log scale)
`num_leaves`	20 – 400 (int)
`max_depth`	4 – 18 (int)
`min_child_samples`	5 – 100 (int)
`subsample`	0.5 – 1.0
`colsample_bytree`	0.5 – 1.0
`reg_alpha` (L1)	10^{-8} – 10 (log)
`reg_lambda` (L2)	10^{-8} – 10 (log)
Early stopping	30 rounds

A key improvement from the prior work is the adoption of a strict temporal train/test split: training data precedes May 18, 2025, and test data spans May 18, 2025 through March 12, 2026. The prior random 80/20 split placed adjacent one-minute timestamps on opposite sides of the boundary, inducing data leakage that inflated sub-model scores (Summer RFR: $R^2 = 0.9999$). The corrected prior-year baseline on held-out data was MAE $= 10.87$ and MAPE $= 30.68\%$. The temporal split adopted here eliminates this bias and provides a more honest basis for comparison.

B. Spatial Post-Processing

A spatial gravity layer applies Inverse Distance Weighting (IDW) across zones of the same permit type following ensemble inference, a modeling dimension absent from the prior publication. When a neighboring zone exceeds 90% occupancy, a pressure term is applied to reduce the predicted availability in the proximate zones:

$$\hat{a}_i^* = \hat{a}_i - \alpha \cdot C_i \sum_{j \neq i} \frac{\mathbf{1}[f_j > \theta]}{d_{ij}^p} \quad (1)$$

C. Inference Optimization

To reduce runtime, the ensemble evaluates only zones matching the user-specified permit type. If the selected garage reports less than 15% availability, ranked alternatives of the same type with at least 10% availability are returned, with spatial adjustment applied before results reach the application layer.

IV. RESULTS AND DISCUSSION

A. Overall Ensemble Performance

The rebuilt ensemble was evaluated on 8,152,097 held-out future records spanning May 18, 2025 through March 12, 2026, covering summer, fall 2025, all break periods, and spring 2026 opening weeks, across 21 zones over a window 67% longer than the prior test period. Predictions were generated for 21 of 23 active zones; Warsaw Zones 2 and 3 were excluded due to insufficient early-period training data.

Table II presents the complete performance comparison between the prior year implementation and the rebuilt ensemble. The prior ensemble, based on KNN and Random Forest sub-models, reported MAPE $= 21.09\%$ and MAE $= 6.75$ under a random 80/20 split that induced data leakage. The corrected prior-year true baseline on genuinely held-out data was MAPE $= 30.68\%$ and MAE $= 10.87$. The rebuilt LightGBM ensemble achieves MAPE $= 20.37\%$, MAE $= 11.42$, RMSE $= 45.85$, and $R^2 = 0.9816$ under a strict temporal split across 21 zones. The MAPE reduction of 33.6% relative to the corrected prior baseline represents a substantive improvement in proportional prediction accuracy. Figure 1 shows over three-quarters of all prediction errors are within 5 spots, and 91% are within 25 spots, indicating a practically meaningful level of precision for user-facing deployment.

B. Performance by Permit Type

Prediction error is strongly stratified by zone capacity and occupancy volatility (Table III), a level of diagnostic resolution not present in the prior publication. Faculty zones perform well (MAE $= 11.19$, $R^2 = 0.9828$). Commuter zones account

TABLE II
PERFORMANCE COMPARISON: PRIOR YEAR VS. REBUILT ENSEMBLE

Model Version	RMSE	MAE	MAPE	R^2
Prior year (Spring 2025) — 11 zones, 6-month test				
Original Ensemble (KNN/RFR)	28.79	6.75	21.09%	0.9934
Current year (Spring 2026) — 21 zones, 10-month test				
Rebuilt LightGBM Ensemble	45.85	11.42	20.37%	0.9816

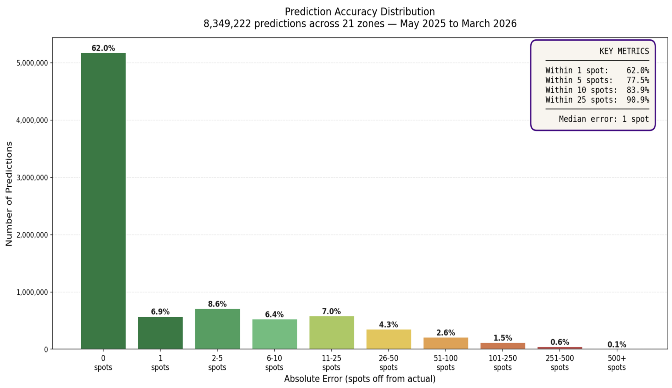

Fig. 1. Prediction Accuracy Distribution for Parking Spot Availability

for 81.4% of total absolute error (MAE = 37.33, RMSE = 88.97, MAPE = 31.62%, $R^2 = 0.9532$), driven by large zone capacities, rapid class-change transitions, and the sensitivity of percentage-based metrics to near-full conditions. The low R^2 for EV zones (0.4787) reflects the near-constant behavior of these 2 to 4 spot zones rather than model inadequacy.

C. Temporal Error Distribution

Table IV shows the hour-of-day error decomposition for commuter zones, providing granular diagnostic insight that extends beyond the aggregate metrics reported in [4]. Overnight hours (00:00 to 06:00) yield MAE = 15.76 and MAPE = 3.59%, reflecting the stable, near-empty baseline occupancy during off-hours. Early morning hours (07:00 to 09:00) show elevated error (MAE = 36.95, MAPE = 36.78%) as garages begin filling in advance of the first major class period. The peak window (10:00 to 15:00), when multiple class periods begin and end within minutes of each other, produces the highest error: MAE = 61.09, RMSE = 121.23, and MAPE = 74.29%, with R^2 declining to 0.9168.

TABLE III
PREDICTION ERROR BY PERMIT TYPE

Permit Type	MAE	RMSE	MAPE	R^2	Error %
Commuter	37.33	88.97	31.62%	0.9532	81.4%
Faculty	11.19	29.12	30.56%	0.9828	15.1%
Accessible	1.00	2.15	10.94%	0.9164	2.6%
EV	0.36	0.78	15.40%	0.4787	0.9%

TABLE V
ERROR CONCENTRATION BY ZONE

Zone	Deck	Type	Cap	MAE	Error %
22	Ballard	Commuter	1,462	68.59	29.9%
42	Warsaw	Commuter	599	34.88	15.2%
13	Champions	Commuter	451	34.43	15.0%
19	Chesapeake	Commuter	630	24.45	10.7%
4	Grace	Commuter	389	24.30	10.6%
All remaining zones (16 zones)				< 11.19	18.6%

TABLE IV
HOUR-OF-DAY ERROR FOR COMMUTER ZONES

Hour Window	MAE	RMSE	MAPE	R^2
Overnight (00:00–06:00)	15.76	59.87	3.59%	0.9762
Early AM (07:00–09:00)	36.95	74.03	36.78%	0.9707
Peak (10:00–15:00)	61.09	121.23	74.29%	0.9168
Afternoon (16:00–18:00)	49.72	98.25	38.05%	0.9382
Evening (19:00–23:00)	31.54	78.70	17.62%	0.9594

Prediction difficulty is structurally tied to the rate of occupancy change rather than the time of day itself. Outside the 10:00–15:00 window, commuter MAPE falls below 38%, confirming that the model generalizes well under stable demand conditions. Fig. 3 plots the hourly mean actual and predicted availability averaged across all commuter zones, showing that the model tracks the observed occupancy curve closely throughout the day. The largest visible divergence occurs during the midday trough between hours 9 and 14, consistent with the peak-window error reported in Table IV.

D. Zone-Level Error Concentration

Table V presents the five zones responsible for 81.4% of total prediction error, an analysis that was not conducted in the prior work but is essential for guiding targeted model improvement. Zone 22 (Ballard Commuter, capacity 1,462) accounts for 29.9% of total absolute error with MAE = 68.59, consistent with its position as the largest and most volatile commuter zone on campus. Warsaw Zone 42 (15.2%, MAE = 34.88) and Champions Zone 13 (15.0%, MAE = 34.43) follow. The remaining 16 zones collectively Error scales with zone capacity, as larger zones exhibit wider absolute availability swings. Commuter zone proportional accuracy is consistent with prior campus parking studies reporting peak-hour MAPE of 20 to 35% at minute-level resolution.

E. Day-of-Week and Weekend Patterns

The model performs best on Mondays (MAE = 10.63) and Fridays (MAE = 11.19), with peak error occurring on Saturdays (MAE = 16.93). Saturday commuter MAPE reaches 56.19% compared to a weekday average of 29.65%. This degradation reflects the irregular, event-driven demand patterns associated with weekend activities including intercollegiate athletics and campus events, for which the training set contains limited and temporally irregular examples. The campus gate schedule, which removes the weekday traffic restriction on weekends, further alters circulation patterns relative to weekday baselines, reducing the predictive leverage of features calibrated to weekday behavior. Saturday performance represents a concrete target for future sub-model development.

F. Spatial Layer Analysis

The IDW layer yields a modest MAPE reduction without meaningful impact on MAE, suggesting that LightGBM sub-models already capture inter-zone dependencies through shared historical lookup features. Future work should explore learned graph-based correction mechanisms, such as graph attention networks, that can better adapt to non-stationary spatial relationships during atypical demand periods.

G. Longitudinal Occupancy Trends and Retraining Frequency

Analysis of hourly mean occupancy across 2024 to 2026 (Tables. VI–VII) indicates that inter-year occupancy patterns are broadly stable, with all the garages exhibiting consistent peak timing between late morning and early afternoon. Champions Deck consistently exceeds 90% occupancy during peak hours across all three years. Grace Street Deck approached full capacity in 2025 with a marginal reduction in 2026. Ballard Deck shows a gradual increase in midday utilization, while Warsaw Deck exhibits a slight peak-hour decline and Chesapeake shows moderately elevated early-morning utilization in 2026. Fig. 2 further illustrates this stability by comparing Ballard Deck availability on a representative Tuesday in October 2024 and October 2025. The two curves track closely throughout the day, with the most notable divergence occurring during the midday peak window, consistent with the gradual utilization increase observed in Fig. 3.

TABLE VI
2024–2026 MEAN PARKING OCCUPANCY PER HOUR: CHAMPIONS PARKING DECK.

Year	7 AM	8 AM	9 AM	10 AM	11 AM	12 PM	1 PM	2 PM	3 PM	4 PM
2024	34%	81%	94%	94%	94%	93%	95%	94%	87%	73%
2025	31%	80%	99%	99%	98%	98%	100%	99%	92%	79%
2026	40%	86%	96%	96%	94%	94%	96%	96%	92%	84%

TABLE VII
2024–2026 MEAN PARKING OCCUPANCY PER HOUR: BALLARD PARKING DECK.

Year	7 AM	8 AM	9 AM	10 AM	11 AM	12 PM	1 PM	2 PM	3 PM	4 PM
2024	2%	5%	19%	32%	37%	40%	45%	47%	38%	28%
2025	1%	5%	19%	32%	38%	41%	48%	50%	41%	30%
2026	6%	9%	20%	31%	38%	45%	52%	53%	46%	37%

TABLE VIII
2024–2026 MEAN PARKING OCCUPANCY PER HOUR: CHESAPEAKE PARKING DECK.

Year	7 AM	8 AM	9 AM	10 AM	11 AM	12 PM	1 PM	2 PM	3 PM	4 PM
2024	2%	10%	38%	60%	61%	62%	68%	69%	58%	41%
2025	2%	11%	41%	62%	66%	70%	77%	76%	62%	44%
2026	15%	22%	44%	64%	67%	67%	71%	72%	62%	48%

The multi-year stability of these patterns supports advancing the prior work's recommendation of monthly retraining [4], suggesting annual retraining is unnecessary under stable campus conditions, provided no significant changes occur in parking infrastructure or enrollment. This has direct operational relevance for institutions seeking to minimize computational and administrative overhead. Daily occupancy distributions confirm pronounced day-of-week structure, with Fridays exhibiting suppressed demand and Wednesday–Thursday sustaining the highest mid-week loads, patterns reliably captured by the `campus_phase` and day-of-week features.

V. CONCLUSION

This paper presented a data-driven parking prediction system evaluated on over 18 months of high-resolution occupancy data from five multi-zone parking garages at a mid-sized university campus, extending and improving upon the prior publication [4] through four principal contributions.

First, a seven-step automated cleaning pipeline resolves sensor overflow, duplication, maintenance artifacts, and out-of-bounds readings across 22 million raw records, replacing the prior ad hoc manual approach with a reproducible procedure. Second, an expanded feature engineering framework produces up to 36 features per prediction row, incorporating cyclical temporal encodings, a 10-level academic calendar phase variable, zone metadata, peak-hour indicators, and historical statistical lookups, representing a threefold increase over the prior six-feature baseline. Third, a three-way LightGBM ensemble with calendar-based routing and Optuna Bayesian hyperparameter optimization over 100 trials replaces the prior KNN and Random Forest sub-models, under a strict temporal train/test split that eliminates data leakage. Fourth, an IDW spatial post-processing layer introduces cross-zone demand spillover modeling absent from prior work.

Evaluated on 8.15 million genuinely held-out future observations across a 10-month test window covering summer, fall semester, all academic breaks, and spring 2026, the rebuilt ensemble achieves MAPE $= 20.37\%$, representing a 33.6% reduction over the corrected prior-year baseline of 30.68%. The median absolute prediction error is one parking space across all predictions, reflecting the practical precision attainable at one-minute temporal resolution. Disaggregated analysis confirms that commuter zones during the 10:00 to 15:00 class-change window constitute the primary prediction challenge, driven by rapid availability transitions that exceed the resolution of static feature-based models. Weekend performance, particularly on Saturdays where commuter MAPE reaches 56.19%, is explicitly characterized as a concrete target for future improvement. Accessible and EV zones are predicted with sub-spot accuracy. Longitudinal occupancy analysis across 2024 to 2026 demonstrates sufficient year-over-year stability to extend the prior work's monthly retraining recommendation, supporting multi-year model reuse absent structural campus changes.

The spatial IDW post-processing layer provides marginal improvement, motivating future exploration of learned graph-

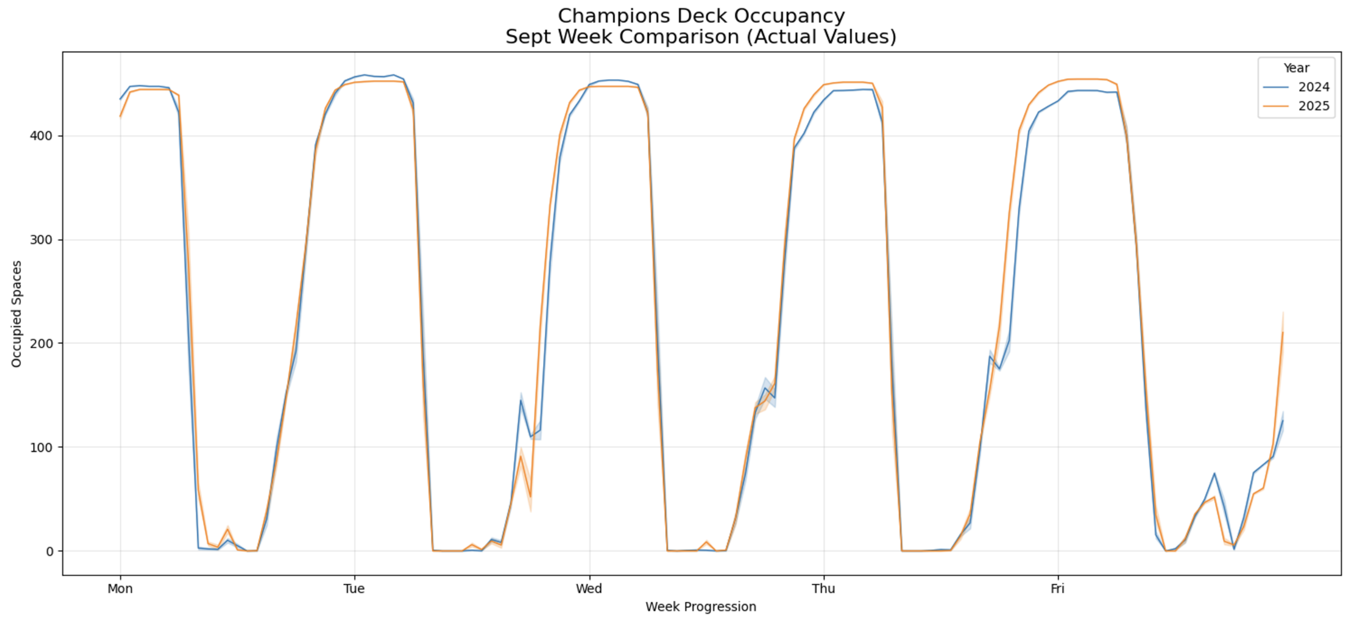

Fig. 2. Champion Parking Deck availability on a representative Week of September 2024 vs 2025. The close alignment of the two curves illustrates year-over-year occupancy stability, supporting the case for reduced retraining frequency.

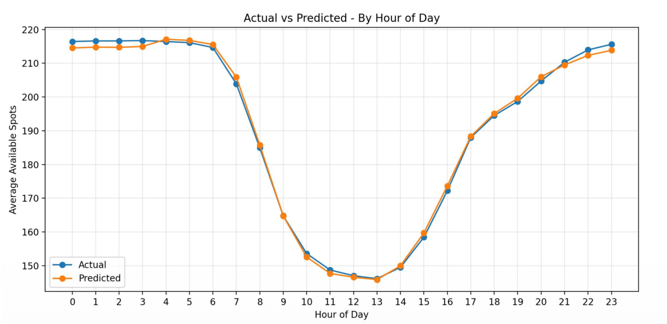

Fig. 3. Average Parking Prediction across Hours of the Day.

based spatial correction architectures for modeling inter-zone dependencies. Future work includes dedicated sub-models for weekend and event-driven demand, per-garage ensemble specialization, and automated event calendar ingestion to reduce reliance on manual flag tables. The framework has been deployed as a live web application providing campus commuters with ML-based parking forecasts, real-time occupancy, and contextual weather and traffic information, demonstrating the operational viability of the proposed approach in an active campus mobility environment [4].

REFERENCES

[1] H. Chen and H. A. Rakha, "Developing and field testing a green light optimal speed advisory system for buses," *Energies*, vol. 15, no. 4, p. 1491, 2022.

[2] S. Paiva, M. A. Ahad, G. Tripathi, N. Feroz, and G. Casalino, "Enabling technologies for urban smart mobility: Recent trends, opportunities and challenges," *Sensors*, vol. 21, no. 6, p. 2143, 2021.

[3] A. Alghoniemy, J. Susko, D. Kahle, L. Saunders, P. Belsare, and S. El-Tawab, "Real-time cloud-based data analysis using machine learning for smart parking," in *2024 International Conference on Computer and Applications (ICCA)*. IEEE, 2024, pp. 1–7.

[4] J. Susko, P. Belsare, and S. El-Tawab, "Enhancing smart parking management: a machine learning approach to predicting parking availability in university environments," in *2025 Systems and Information Engineering Design Symposium (SIEDS)*. IEEE, 2025, pp. 66–71.

[5] S. Zhang and E. Frimpong Boamah, "Managing campus parking demand through course scheduling–an approach to campus sustainability," *International Journal of Sustainability in Higher Education*, vol. 22, no. 4, pp. 909–930, 2021.

[6] M. Garcia, P. Rose, R. Sung, and S. El-Tawab, "Secure smart parking at james madison university via the cloud environment (space)," in *2016 IEEE Systems and Information Engineering Design Symposium (SIEDS)*. IEEE, 2016, pp. 271–276.

[7] M. Farag, A. Hilal, and S. El-Tawab, "Parking occupancy prediction and traffic assignment in a university environment," in *2022 10th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC)*. IEEE, 2022, pp. 186–190.

[8] X. Xiao, Z. Peng, Y. Lin, Z. Jin, W. Shao, R. Chen, N. Cheng, and G. Mao, "Parking prediction in smart cities: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 10, pp. 10 302–10 326, 2023.

[9] J. Barker and S. ur Rehman, "Investigating the use of machine learning for smart parking applications," in *2019 11th International Conference on Knowledge and Systems Engineering (KSE)*. IEEE, 2019, pp. 1–5.