

Prediction of Overnight Truck Parking Availability in the U.S.

Troy Choi, P.E. Ph.D.¹

¹ Research Lead at Trucking Lab, LLC.

ABSTRACT

Safe overnight truck parking in the United States is in critical shortage, jeopardizing road safety, logistics efficiency, and driver well-being. This study proposes a data-driven framework to forecast real-time parking availability by integrating nationwide facility data, traffic patterns, and crowdsourced reports. Key techniques include clustering via the K -Prototypes algorithm, matrix factorization for missing data, and time-series forecasting using long short-term memory (LSTM) networks. Through this approach, over 10,000 parking facilities across the country are grouped into five operationally meaningful clusters—Premium Private, Budget Large Full-Service, Budget Medium Mixed, Budget Medium Full-Service, and Budget Small Mixed—supporting more precise availability estimation at the facility level.

Ground truth validation using sensor and camera data demonstrates the model’s ability to capture macro-level utilization trends, despite challenges such as sparse user input and dynamic demand shifts. The study emphasizes the need for a unified national truck parking information system to consolidate fragmented state-level efforts, facilitate cluster-based performance tracking, and support targeted investments, such as capacity expansion. This study recommends a hybrid approach—leveraging predictive models and user-generated reports—as a scalable, low-cost substitute for sensor-dependent infrastructure.

Keywords: *Overnight Truck Parking, Real-Time Truck Parking Availability, Truck Parking Information Management System, LSTM Forecasting*

1 INTRODUCTION

2 The United States faces a critical shortage of safe overnight truck parking, directly threatening
3 road safety, driver well-being, and freight efficiency. With only 313,000 designated parking spaces
4 nationwide—averaging one spot per eleven trucks—drivers waste 56 minutes daily searching for
5 parking, resulting in \$4,600–\$5,600 in lost wages per driver each year [1, 2, 3]. This scarcity forces
6 drivers to violate federal Hours-of-Service (HOS) regulations or park in unsafe locations, such as
7 highway shoulders, increasing crash risks. Multiple factors have worsened the problem: rising e-
8 commerce freight volumes, limited infrastructure investment, and urban development constraints,
9 including community resistance [4]. While efforts such as Jason’s Law and recent federal grants
10 (e.g., \$80 million under the Biden-Harris Administration) aim to expand capacity [2, 5, 6], new
11 construction alone cannot address the dynamic and time-sensitive nature of parking demand.

12 Accordingly, there is growing interest in data-driven solutions that can help drivers locate
13 available parking in real time and help drivers make informed decisions without depending solely
14 on physical sensors. To pursue this direction, the study sets out to answer two core questions:

15 **1. Can we accurately predict truck parking availability without installing expensive** 16 **sensors at each location?**

- 17 • For example, if facilities in Maryland and Colorado share key attributes (e.g., capacity,
18 amenities, traffic), can usage patterns from one be used to predict the other?

19 **2. Is it feasible to build a unified, nationwide truck parking information system?**

- 20 • What technical and institutional mechanisms are needed for interstate drivers to access
21 consistent, real-time parking data across jurisdictions?

22 To address these questions, this study proposes a scalable framework that integrates national
23 truck parking data with crowdsourced availability reports, traffic volumes, and facility
24 characteristics. Using a combination of clustering (via K -Prototypes), matrix factorization, and
25 Long Short-Term Memory (LSTM)-based time series forecasting, the model estimates real-time
26 parking availability across over 10,000 facilities nationwide.

27 The following chapters first analyze existing studies to identify key factors influencing truck
28 parking demand and evaluate current predictive models. The methodology section details the
29 process of parking data collection, preprocessing, and integration, followed by the development of
30 a predictive framework. Finally, the results and implications—such as policy recommendations
31 for scalable infrastructure planning or technological strategies for low-cost parking availability
32 systems—are discussed in depth.

33 LITERATURE REVIEW

34 This section provides an overview of key variables, statistics, and truck driver preferences
35 affecting truck parking, as identified in various existing surveys. It further examines
36 methodologies for predicting truck parking availability, ranging from real-time sensor data to
37 advanced modeling. Finally, the discussion explores how these insights can help government
38 agencies and stakeholders address truck parking challenges by optimizing parking infrastructure,
39 mitigating operational inefficiencies, and developing data-driven policies.

1 **Factors Influencing Truck Drivers' Parking Decisions**

2 A 2020 survey of truck drivers revealed that amenities at parking sites are prioritized in the
3 following order: restrooms, anticipated parking availability, food services, parking space size,
4 showers, and refueling services. Among these, parking availability was particularly important due
5 to Hours of Service (HOS) regulations, which mandate drivers to pre-plan stops and occasionally
6 end their trips before their maximum driving hours to comply with rest requirements [7].

7 Truck drivers' routing and parking decisions involve a careful balance of time, cost, regulations,
8 personal preferences, and distance to their destination. When choosing where to park, drivers
9 primarily consider HOS, available amenities, and their target destination. Location is the top
10 priority, with drivers favoring parking areas that offer easy access to highways and major road
11 networks or are close to their delivery points. Safety, particularly at night, is another crucial factor.
12 Interestingly, the physical condition of parking lots—such as potholes, lane markings, or
13 lighting—has a minimal impact on parking decisions. In addition, drivers generally exhibit two
14 distinct preferences when selecting parking. Some prioritize convenience and reliability, seeking
15 easily accessible locations that allow them to maximize driving time while adhering to HOS
16 regulations. Others, particularly younger drivers, favor parking facilities with amenities and
17 services that enhance comfort during rest breaks [8].

18 Boris and Brewster analyzed truck parking log data to identify primary factors influencing truck
19 drivers' parking site selection. The top priorities were proximity to route (96.5%), restroom and
20 shower facilities (79.8%), and expected parking availability (75.5%), followed by spacious parking
21 with easy access (31.9%) and access to restaurants (30.5%). The study also highlighted operational
22 inefficiencies: drivers frequently had to park with 31–60 minutes of remaining drive time under
23 HOS regulations, a pattern estimated to result in an annual loss of approximately 9,300 miles in
24 drivable distance [9].

25 Anderson et al. conducted a survey among truck drivers in the Pacific Northwest to identify
26 key factors influencing truck parking. The study found that truck drivers prioritize convenience
27 and safety, favoring parking facilities equipped with showers, internet access, ample space, and
28 bright lighting. Additionally, the presence of nearby infrastructure (e.g., restaurants, convenience
29 stores, restrooms, and repair shops) plays a role in their choice of parking locations. However, the
30 study also emphasized that improving amenities alone would not directly resolve parking issues,
31 as the primary issue remains a fundamental shortage of parking spaces. Furthermore, the lack of
32 real-time parking availability prevents drivers from knowing where open spaces are available,
33 making it difficult for them to find suitable parking even if high-quality facilities exist [10].

34 **Truck Parking Classification and Availability Prediction**

35 Nevland et al. analyzed GPS data from long-haul commercial vehicle stop endpoints to classify
36 truck parking locations into nine categories. These included five legal types—1) public rest areas
37 and gas stations, 2) weigh stations, 3) open-access heavy commercial vehicle (HCV) parking, 4)
38 limited-access HCV parking, and 5) authorized roadside parking—and four unauthorized types,
39 such as unauthorized roadside parking, illegal highway ramp parking, and unauthorized parking
40 on public or private property. This framework helps planners understand parking supply by
41 showing where shortages occur and supporting solutions for parking management [11]

1 Regarding truck parking availability, existing studies predict availability based on
2 observational data. Using cameras, embedded sensors beneath pavement within parking spaces, or
3 truck entry/exit records, availability can be predicted from as short as 16 minutes to as long as 6
4 hours [12, 13, 14]. While short-term predictions offer accurate availability information from the
5 moment they are displayed on Variable Message Signs (VMS) until the driver arrives at the
6 parking facility, their benefit can be minimal. Since VMS displays are seen close to arrival, they
7 provide little practical value for drivers who must plan rest breaks well in advance, often early in
8 their shifts. As stated in the Kansas study, truck drivers make parking decisions dynamically at
9 various points during their journeys—hundreds of miles from their destination, near exits to known
10 facilities, or at parking lots when a lack of spaces forces them to seek alternatives [8]. Therefore,
11 extending prediction capabilities to cover the entire 11-hour driving window could enable drivers
12 to use GPS navigation or smartphone apps for proactive trip planning, potentially reducing parking
13 search time and enhancing operational efficiency.

14 **Efforts to Address Truck Parking Challenges**

15 State agencies periodically develop State Freight Plans and supplemental Truck Parking Studies
16 to assess critical issues such as undesignated parking hotspots, gaps in amenities, and future
17 parking demand projections. These studies inform strategies like capacity expansion or Truck
18 Parking Information Management System (TPIMS) recommendations. To date, approximately 24
19 states have conducted truck parking studies, utilizing various data sources, such as GPS telematics,
20 crash and traffic datasets, truck stop inventories, driver surveys, and trajectory analyses.

21 Several states, including Florida, Utah, and Rhode Island, prioritize real-time truck parking
22 availability information to address identified challenges [15, 16, 17]. Colorado, for instance, has
23 developed a traffic information app providing truck drivers with updates on road closures,
24 incidents, traffic camera feeds, and rest area inventories [18]. While useful, such tools are often
25 limited to state-specific data, thus reducing their utility for interstate drivers. Similarly,
26 Connecticut is currently developing a state-specific parking inventory app [19]. Prozzi et al. [20]
27 conducted a pilot study at nine rest areas along the I-80 and I-94 corridors in Iowa and Wisconsin,
28 using a mobile-based system to provide real-time parking availability information. The study
29 found that voluntary participation resulted in low app usage, thereby reducing the reliability of
30 parking data. Beyond apps, states utilize DOT-operated 511 systems to share truck-related updates,
31 while the Mid-America Association of State Transportation Officials (MAASTO)—a coalition of
32 eight Midwestern states (Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Ohio,
33 Wisconsin)—has launched a TPIMS web application to centralize regional parking data [21].
34 While state-level efforts exist, their fragmentation indicates the need for a unified national
35 framework.

36 In brief, key variables influencing truck parking decisions include availability, location,
37 amenities, safety, and convenience. The existing studies suggest that parking availability—
38 particularly long-term prediction capabilities—is a key variable for truck driver’s decision-making,
39 as it directly impacts operational efficiency and compliance with driving hour regulations. While
40 real-time parking availability information is essential, current state-specific systems (e.g., mobile
41 apps, 511 alerts) face limitations in addressing the needs of interstate drivers, who often traverse
42 multiple jurisdictions. The findings emphasize the need for standardized, multi-state data

1 integration to overcome fragmented information sources. A unified national truck parking
2 application, supported by interoperable data protocols and cross-agency collaboration, could
3 bridge this gap, ensuring drivers access reliable, real-time occupancy data regardless of state
4 boundaries.

5 **METHODOLOGY**

6 The methodology proceeds in the following sequence. It begins with consolidating existing
7 database information with newly reported parking locations from truck drivers. Satellite imagery
8 and street view are used to estimate parking capacity. When necessary, this preliminary capacity
9 assessment is validated through cross-verification with parking facility operators to confirm the
10 actual number of available spaces. Truck volume and speed limit data are sourced from the
11 Highway Performance Monitoring System (HPMS), a national roadway database [22].

Structured Data Workflow

[1. Data Collection]

a. Capacity Data

- Sources: Jason's law truck parking surveys, truck stop directory
- Methods: Manual count using satellite imagery and street view
- Verification: Cross-validation with facility owners or operators

b. Traffic Data

- Metrics: Truck traffic volume from Highway Performance Monitoring System (HPMS)

c. Availability Data

- Source: Crowdsourced user-reported availability data

[2. Data Processing]

a. Parking Lot Classification

- Technique: Clustering using k-Prototype algorithm

b. Missing Data Imputation

- Method: Matrix factorization via Singular Value Decomposition (SVD)

c. Availability Prediction

- Model: LSTM for 24/7 availability forecasting

[3. Data Provision]

- Delivery: Real-time data visualization via mobile application
- Feedback Loop: User-reported data from the app → integrated into [2. Data Processing]

12 Real-time availability and amenity data are collected via user reports. Facilities are grouped
13 into clusters based on attributes such as traffic, amenities, and parking fees. Within each cluster,
14 missing values in time-series data are imputed. Final availability predictions incorporate temporal
15 patterns and are visualized on a mobile map. User feedback continuously improves model accuracy.

16

1 **Data Collection**

2 *Capacity Data*

3 This section proposes a standardized yet adaptable framework for capacity estimation across
4 diverse parking environments, integrating satellite imagery, area-based calculations, and crowd-
5 sourced data.

6 Types of Truck Parking Facilities

7 Truck parking facilities include a wide range of public and private options. Public
8 infrastructure—such as rest areas, welcome centers, weigh stations, and designated roadside
9 zones—offers basic amenities and regulated spaces, though drivers also utilize informal areas
10 including shoulders and industrial streets. Private facilities range from large full-service truck stops
11 (e.g., Pilot, TA, Love’s) to smaller fuel stations. Some restaurants, retail stores, and commercial
12 lots permit overnight parking under informal arrangements. This diverse and often improvised
13 network reflects both the flexibility and fragmentation of the nation’s truck parking landscape.

14 Public Parking Facilities

15 Public parking areas typically have striped spaces that allow for straightforward capacity
16 estimation. Researchers can manually count these spaces using high-resolution satellite imagery,
17 as illustrated in Figure 1(a). Supplementary data from state-maintained inventories and freight
18 plans are also useful. However, these sources may overlook informal parking along shoulders or
19 curbs and may not reflect real-time usage or recent layout changes.



21 (a) Striped Lots



(b) Unmarked Gravel/Dirt Lots

22 **Figure 1.** Satellite Imagery Examples of Manual Counts and Area-Based Estimates

23 Private Truck Stops

1 Private truck stops often publish capacity data online, but these figures require scrutiny.
2 Terminology may be ambiguous, as spaces for tractors without trailers (i.e., bobtails) are often
3 conflated with full semi-truck spaces, leading to overestimated capacity. Satellite imagery may
4 help, though tree cover and outdated views limit its usefulness. Verifying capacity thus requires
5 direct confirmation from facility operators.

6 Smaller and Undesignated Sites

7 Smaller or informal sites—such as restaurant lots or ramp/shoulder areas—often lack marked
8 spaces, as shown in Figure 1(b). In such cases, capacity is estimated by measuring total area via
9 GIS tools and dividing by 400–500 square feet per truck, based on an 8.5 ft width, 65–70 ft length,
10 and buffer space for standard semi-trailers. Observed patterns show typical spacing of 13–14 feet
11 laterally and 80–90 feet linearly.

12 Capacity Reduction Factors

13 Capacity estimates should account for real-world factors that reduce usable space. Physical
14 constraints—such as potholes, uneven surfaces, and trailer drop areas—directly limit parking
15 capacity, as illustrated in Figure 2. In addition, drivers often avoid poorly lit or obstructed areas at
16 night, further lowering effective capacity. Crowdsourced reviews (e.g., Google Maps, Yelp) can
17 help identify such usability issues. Area-based calculations also assume dry, stable ground;
18 unpaved lots may lose capacity during inclement weather.



19 (a) Dropped Trailers

20 (b) Potholes

21 **Figure 2.** Capacity Reduction Factors

22 While no universal formula exists due to site-specific variation, combining satellite imagery,
23 area-based methods, and user-generated data enables more accurate inventory estimation.

24 *Traffic Data*

25 Truck-related roadway data help characterize parking facility usage and context. This study
26 uses HPMS data, with truck volume selected as the primary input—representing the average daily

1 traffic of combination trucks (vehicle classes 8–13). While HPMS contains broader attributes and
2 projections, truck volume is emphasized due to its direct relevance to parking demand.

3 *Availability Data*

4 Parking availability is defined as a ratio representing the proportion of unoccupied spaces
5 relative to a facility’s total capacity, calculated using the equation (1):

$$6 \quad \text{Availability} = 1 - \left(\frac{\text{Reported or Predicted Number of Parked Trucks}}{\text{Total Parking Capacity}} \right) \quad (1)$$

7 This metric ranges from 0 to 1 and is derived either from model-based predictions or, when
8 available, from user reports. Thresholds are interpreted as follows:

- 9 • <0.25: Near or exceeding capacity
- 10 • 0.25–0.5: Mostly full
- 11 • 0.5–0.75: Half full
- 12 • ≥ 0.75 : Mostly empty

13 Availability data is organized into 15-minute intervals (672 per week per cluster) to support
14 real-time prediction and display. Rather than relying on costly on-site sensors (e.g., induction loops,
15 cameras) or video analytics [12, 14], this study uses a mobile app to visualize availability on an
16 interactive map, helping drivers quickly identify options based on updated occupancy.

17 **Data Processing**

18 The collected data—capacity, traffic, and availability—are integrated to assess parking
19 availability. Using the Spatial Join tool in ArcGIS Pro, each facility is georeferenced to adjacent
20 HPMS highway segments within a 0.5-mile radius. Truck traffic volumes from nearby segments
21 are then aggregated, and availability records appended. The resulting dataset combines facility
22 attributes, traffic metrics, and availability into a unified geospatial resource for analysis.

23 *Parking Lot Classification*

24 Variables influencing truck parking decisions—such as safety, parking type, fees, amenities,
25 and traffic—are first tested using Analysis of Variance (ANOVA) to identify statistically
26 significant inputs for clustering. Next, clustering is performed using the K-Prototypes algorithm,
27 which supports both numerical and categorical variables and is well-suited for interpretable,
28 mixed-type clustering. The algorithm proceeds as follows:

- 29 1. **Cluster Initialization:** Randomly select K cluster centroids.
- 30 2. **Distance Calculation:**
 - 31 ○ *Numerical data:* Use Euclidean distance.
 - 32 ○ *Categorical data:* Use Hamming distance (1 if different, else 0).
 - 33 ○ A weight parameter (λ) balances the importance of both distances.

- 1 3. **Cluster Assignment:** Assign each data point to the nearest cluster centroid.
- 2 4. **Centroid Update:**
- 3 ○ *Numerical variables:* Update using the mean.
- 4 ○ *Categorical variables:* Update using the mode.
- 5 5. **Iteration:** Repeat the process until convergence is achieved.

6 The optimal number of clusters (K) is determined using the Elbow Method, which evaluates
7 clustering quality by computing the Within-Cluster Sum of Squared Distances (WCSS), as shown
8 in Equation (2). Lower WCSS values indicate more cohesive clusters, but as K increases, the
9 marginal gain diminishes. The point where the WCSS curve begins to flatten represents the optimal
10 balance between model simplicity and clustering accuracy.

$$11 \qquad \qquad \qquad WCSS = \sum_{i=1}^K \sum_{x \in C_i} (x - u_i)^2 \qquad (2)$$

12 where K : number of clusters, C_i : set of data points in the i^{th} cluster, x : each data point in cluster
13 C_i , and u_i : centroid of the i^{th} cluster

14 As the number of clusters increases, WCSS naturally decreases since data points are closer to
15 their assigned centroids. However, beyond a certain point, the reduction in WCSS slows, forming
16 an “elbow” in the plot of WCSS versus K . This elbow point reflects the optimal trade-off between
17 under-segmentation (too few clusters) and overfitting (too many), ensuring meaningful and
18 generalizable results. To apply the Elbow Method, a range of K values (e.g., 1 to 10) is tested. The
19 K -Prototypes algorithm is executed for each K , and the corresponding WCSS is calculated. These
20 values are plotted to visually identify the elbow point, where the rate of WCSS reduction
21 diminishes significantly. The selected K is then used for final clustering.

22 In addition to the Elbow Method, the silhouette score is used to assess clustering quality. The
23 analysis evaluates how well each data point is clustered by measuring the cohesion within clusters
24 and the separation between clusters, using Equation (3). Scores closer to 1 indicate well-defined
25 clusters, while lower values suggest overlap or poor separation.

$$26 \qquad \qquad \qquad S = \frac{b-a}{\max(a,b)} \qquad (3)$$

27 where a is the average distance between a data point and all other points within the same cluster
28 (cohesion), and b is the average distance between the data point and all points in the nearest
29 neighboring cluster (separation).

30 *Missing Data Imputation*

31 Parking availability data sourced from crowdsourced user reports often includes missing values
32 at certain time intervals. To address this, a two-step imputation strategy is applied. First, a cluster-
33 based pattern extraction method is used. Facilities with similar characteristics—such as amenities
34 and traffic volume—are grouped ($K=5$). Within each cluster, shared time-dependent availability
35 patterns are used to estimate missing values. This approach offers:

1 Parking availability data sourced from crowdsourced user reports often includes missing values at
2 certain time intervals. To address this, a two-step imputation strategy is applied. First, the
3 previously defined five clusters are used to group facilities with similar characteristics (e.g.,
4 amenities, traffic volume). Within each cluster, missing values are estimated based on typical time-
5 dependent availability patterns (e.g., lots in the same group may consistently fill up by early
6 evening on weekdays). This approach offers:

- 7 1. Improved prediction accuracy by reducing outlier effects,
- 8 2. Faster model training by processing smaller subsets, and
- 9 3. Clearer insights for planners and policymakers.

10 Second, within each cluster, matrix factorization is applied using Singular Value
11 Decomposition (SVD), a technique widely used in recommendation systems, such as suggesting
12 movies based on past viewing history. It decomposes the facility–time matrix into latent patterns
13 (e.g., peak congestion, daily cycles) to estimate missing values. For example, if data for “Parking
14 Lot #7 at 12:00 PM” is missing, the model uses similar time patterns from other facilities in the
15 same cluster to impute the value.

16 In summary, combining cluster-based modeling with latent factor decomposition enables
17 context-aware, scalable imputation. However, this step does not yet account for external temporal
18 factors (e.g., weather or events), which are addressed later using LSTM-based time-series models.

19 *Availability Prediction*

20 The refined dataset supports prediction of truck parking availability over a weekly cycle,
21 segmented into 672 intervals of 15 minutes each. An LSTM model—a type of recurrent neural
22 network designed to learn temporal dependencies—is applied to cluster-level data. This model
23 captures complex time-series dynamics, including day/night patterns, weekday–weekend shifts,
24 and seasonal trends.

25 The model uses the first six intervals (90 minutes) as input to predict the remaining 666. This
26 sequence length was chosen after testing multiple configurations. Shorter sequences, such as 2 or
27 4 intervals (30–60 minutes), enable faster training but often miss essential daily or cyclical patterns.
28 Longer sequences, such as 8 or 10 intervals (120–150 minutes), may capture more comprehensive
29 trends but risk overfitting and require longer training times. The 6-interval setting achieves a
30 practical balance, capturing meaningful patterns like morning peaks while avoiding excessive
31 noise or computational cost.

32 **Data Provision and User Feedback Loop**

33 Predicted availability is shared through two channels. First, TPIMS can display near-term
34 availability on highway Variable Message Signs (VMSs). Second, a mobile app visualizes
35 forecasts that cover periods longer than 11-hour HOS shifts to help drivers plan rest breaks.
36 Considering that driver-submitted reports via the mobile app can be subjective, the system
37 encourages drivers to provide descriptive input; for instance, if only a few parking spaces are
38 available, the status might simply be labeled “mostly full.” However, when many parking lots are

1 involved, the system calculates the expected number of available spaces relative to total capacity
 2 and sets thresholds accordingly—for example, in a lot with a capacity of 120 vehicles, a "mostly
 3 full" condition might indicate about 30 available spaces (25% of capacity), while a "mostly empty"
 4 condition might correspond to roughly 90 available spaces (75% of capacity). Additionally, the
 5 system averages multiple user reports within the same interval, filtering extreme values to reflect
 6 realistic availability. This user-reported data is then cycled back into the data processing stage to
 7 enable regular model retraining, continuously improving the model’s predictive accuracy.

8 **RESULT**
 9 **Clustering Result**

10 Table 1 summarizes the ANOVA test results, ranking variables by statistical significance. Five
 11 variables were selected for clustering based on operational relevance and data quality: TruckSpace,
 12 AADT_COMBI, Shower, Fee, and Type. While variables like Wi-Fi and Food were also
 13 significant, they were excluded due to multicollinearity (e.g., showers often co-occur with
 14 food/Wi-Fi) and data sparsity.

15 **Table 1. ANOVA Test Summary**

Rank	Variable	Description	P-Value
1	TruckSpace	Parking lot capacity	p < 0.001
2	AADT	Traffic volume	p < 0.001
3	AADT_COMBI	Heavy truck traffic volume	p < 0.001
4	Shower	Shower facility availability (0 = No, 1 = Yes)	p < 0.001
5	Food	Food availability (0 = No, 0.5 = Vending machine, 1 = Hot food)	p < 0.001
6	Wifi	Wi-Fi availability (0 = No, 1 = Yes)	p < 0.001
7	Type	Public or private parking (0 = Public, 1 = Private)	p < 0.001
8	Fee	Overnight parking fee (e.g., \$15)	p = 0.002
9	Safety	Safety score based on user reports (0 = Safe, 1 = Caution, 2 = Dangerous)	p = 0.033

16

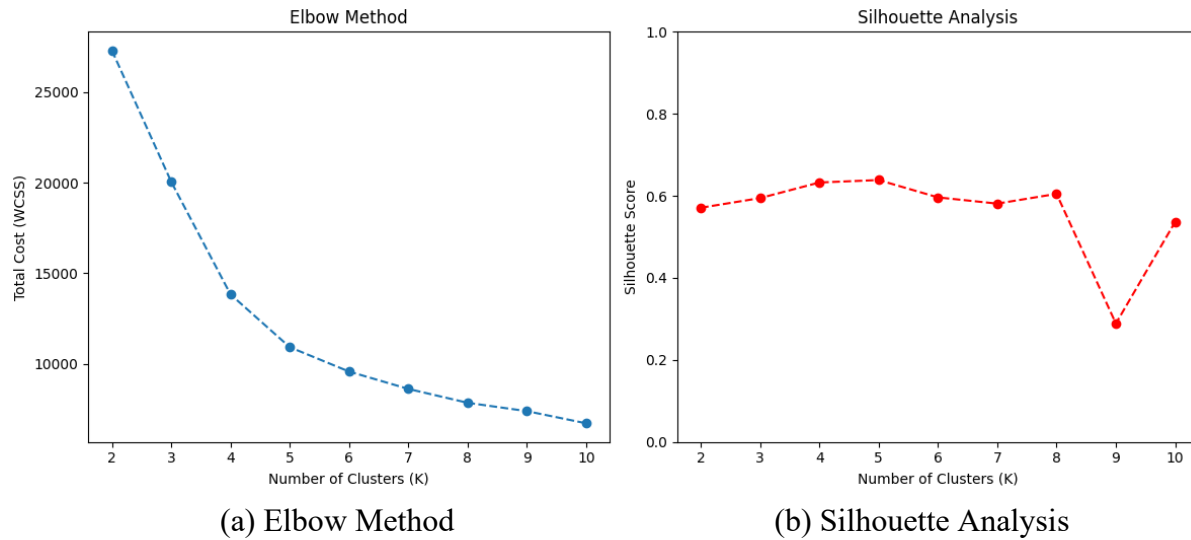


Figure 3. Optimal Cluster Determination

Figure 3 shows the elbow method and silhouette analysis used to determine the optimal number of clusters. The elbow curve reveals a notable drop in WCSS between $K=4$ and $K=6$, after which the rate of decrease becomes marginal. Silhouette scores remain above 0.5 up to $K=8$, but drop sharply at $K=9$, suggesting overfitting. Combining both metrics, the optimal range is narrowed to $K=4-6$.

These patterns become more evident in how the clustering structure evolves as K increases. As K increases, parking facilities are progressively divided based on operational characteristics. At smaller K values, clusters are first separated by parking cost—for example, facilities with high overnight fees versus those offering free or low-cost parking. As the cluster count increases further, the grouping becomes more specific in the following order: Fee \rightarrow TruckSpace \rightarrow Type \rightarrow Shower. This hierarchical structure reflects increasing differentiation of facility types, especially among lower-cost sites, which exhibit wider variation in size, services, and ownership. In contrast, higher-cost facilities tend to remain consistently grouped across different K values. However, when K becomes too large, distinctions between some groups blur, and clusters with nearly identical profiles start to appear, reducing interpretability. Considering these factors, this study selects $K=5$ as the final cluster count, balancing granularity with practical utility.

Table 2. Summary of Optimal Truck Parking Lot Clusters

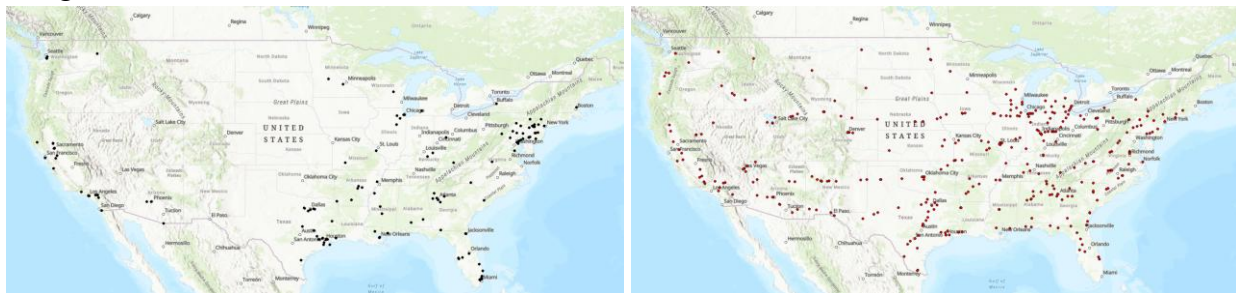
Cluster Category	Count	Shower	Fee	Truck Space	AADT_COMBI	Type	Description
1. Premium Private Parking	140	Moderate	Very High	Average	Average	Mostly Private	Private lots near urban areas with high fees, average space and showers (e.g., truck storage sites)

2. Budget Large Full-service Parking	363	Very High	Average	Very Large	Slightly Low	Mostly Private	Large private truck stops with full services and shower availability. (e.g., full-service large truck stops).
3. Budget Medium Mixed Parking	1,155	Very Low	Average	Mostly Small	Very High	Mostly Public	Public facilities near interstates with high traffic and limited amenities (e.g., rest areas).
4. Budget Medium Full-service Parking	1,895	Very High	Average	Above Average	Slightly Low	Mostly Private	Mid-sized private stops in suburban areas with full services.
5. Budget Small Mixed Parking	6,454	Very Low	Average	Small	Slightly Low	Partially Public	Small public facilities with limited services and low cost (e.g., turnouts, pullouts).

1 Table 2 summarizes the five optimal clusters derived through *K*-Prototypes analysis. Key
2 numerical variables are translated into descriptive categories. For the optimal cluster count,
3 representative examples include:

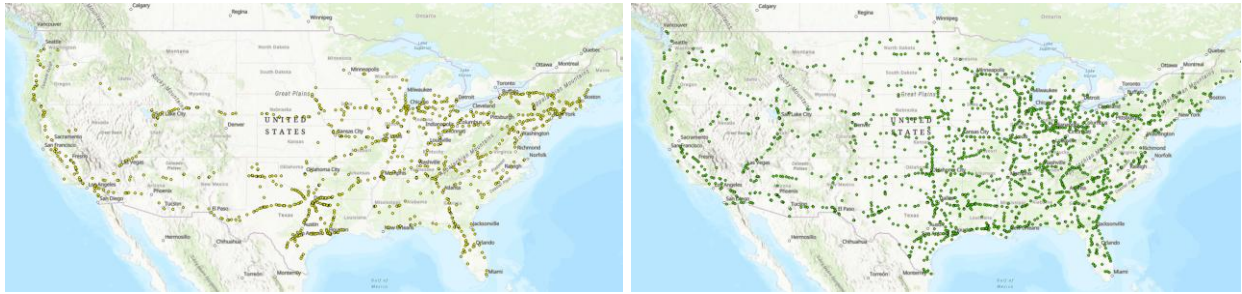
- 4 • **Fee**
5 ○ Very High: \$17–\$50 per night
6 ○ Average: Free to \$17 per night
7 • **Truck Space**
8 ○ Very Large: ≥ 130 spaces
9 ○ Above Average: 34–125 spaces
10 ○ Small: ≤ 50 spaces

11 Note that these thresholds are not strictly mutually exclusive. For example, a facility with 90
12 spaces (possibly classified as *Above Average*, *Average*, or *Small*) and \$16/night fees (*Average*)
13 could align with multiple categories depending on other variables. The *K*-Prototypes algorithm
14 resolves such ambiguity by evaluating combined attribute similarity and assigning each facility to
15 a single cluster.



16 #1 Premium Private Parking

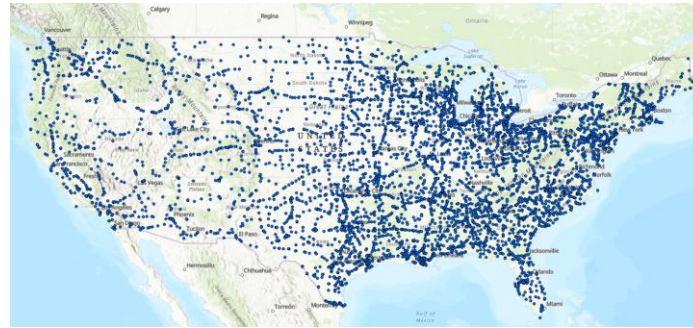
17 #2 Budget Large Full-service Parking



1
2

#3 Budget Medium Mixed Parking

#4 Budget Medium Full-service Parking



3
4

#5 Budget Small Mixed Parking

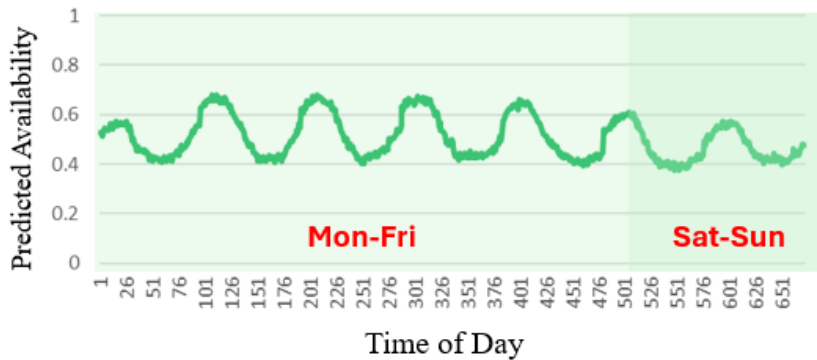
Figure 4. Geographic Distribution of Optimal Truck Parking Lot Clusters

5
6
7

Figure 4 visualizes the geographic distribution of clusters at K=5.

User Feedback Loop and Usage

Coldwater Welcome Center (MI)



8
9

Figure 5. Sample Weekly Availability Data for an Individual Facility

10 Availability data is collected via a mobile app, where drivers submit real-time reports for
 11 specific facilities. These are sent to a central database organized by location. The data—whether
 12 user-reported or algorithm-predicted—is used to generate weekly forecasts, as visualized in Figure
 13 5. For instance, a public welcome center in Mississippi that accommodates up to 24 trucks shows
 14 weekday vs. weekend patterns, with more vacancies on weekends.

15 **Table 3.** State-Level Summary of Truck Parking Inventory and Average Utilization

No.	State	Private Facilities			Public Facilities		
		Locations	Spaces	Avg. Utilization	Locations	Spaces	Avg. Utilization
1	Alabama	336	7,815	57.0%	27	716	54.0%
2	Arkansas	218	6,406	56.8%	27	650	54.1%
3	Arizona	129	6,847	55.9%	29	713	50.1%
4	California	265	13,002	56.8%	128	3,072	53.1%
5	Colorado	110	3,552	60.4%	35	385	52.5%
6	Connecticut	15	862	58.8%	21	534	58.9%
7	Delaware	10	189	61.1%	22	105	46.6%
8	Florida	216	7,834	62.1%	86	3,041	54.9%
9	Georgia	307	11,449	60.5%	48	1,836	53.5%
10	Iowa	205	6,722	55.6%	48	735	51.5%
11	Idaho	65	2,340	55.4%	14	316	57.6%
12	Illinois	274	12,218	55.4%	97	1,805	57.6%
13	Indiana	248	11,518	56.0%	42	2,960	51.9%
14	Kansas	142	4,689	53.0%	40	475	53.6%
15	Kentucky	119	5,270	53.0%	38	1,188	53.6%
16	Louisiana	267	9,191	55.3%	23	396	52.2%
17	Massachusetts	26	524	59.3%	41	466	56.1%
18	Maryland	41	1,773	62.1%	24	622	58.7%
19	Maine	31	718	55.9%	11	106	52.8%
20	Michigan	238	5,757	57.0%	73	1,843	44.7%
21	Minnesota	171	3,563	55.8%	54	789	51.9%
22	Missouri	171	9,886	57.9%	59	1,441	49.3%
23	Mississippi	183	5,283	56.6%	55	653	51.2%
24	Montana	70	2,757	51.8%	70	870	49.5%
25	N. Carolina	192	4,676	60.2%	51	812	55.4%
26	N. Dakota	71	2,998	51.4%	21	224	50.0%
27	Nebraska	123	4,029	56.3%	20	327	60.5%
28	New Hampshire	24	376	56.6%	13	191	52.6%
29	New Jersey	79	2,458	61.8%	26	940	58.8%
30	New Mexico	118	5,673	61.8%	55	573	58.8%
31	Nevada	71	4,412	51.9%	50	596	48.8%
32	New York	140	3,929	56.8%	77	1,826	51.7%
33	Ohio	298	10,985	52.4%	97	2,547	57.5%
34	Oklahoma	246	7,957	56.4%	18	298	48.8%
35	Oregon	79	3,460	55.0%	44	674	55.6%
36	Pennsylvania	213	7,301	59.5%	79	1,729	57.9%
37	Rhode Island	2	135	65.1%	7	66	51.8%
38	S. Carolina	193	6,231	60.1%	40	936	52.1%
39	S. Dakota	94	2,599	53.0%	40	509	49.2%
40	Tennessee	170	6,806	60.7%	59	980	56.7%
41	Texas	966	33,952	56.7%	149	2,441	53.2%

42	Utah	80	3,348	56.7%	46	595	53.7%
43	Virginia	155	7,148	61.0%	36	802	60.2%
44	Vermont	18	196	55.4%	44	270	51.1%
45	Washington	69	2,646	56.3%	100	1,028	49.9%
46	Wisconsin	280	6,891	54.7%	40	984	52.0%
47	W. Virginia	35	1,203	55.8%	24	718	57.4%
48	Wyoming	75	3,769	54.8%	80	1,105	49.3%

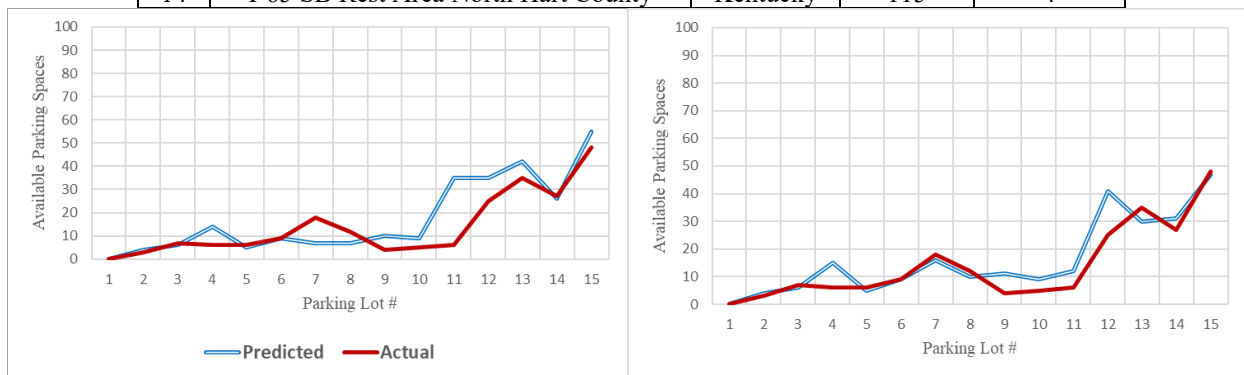
1 Availability data can also be aggregated by regions (e.g., county or state), supporting macro-
2 level analyses. As shown in Table 3, such regional aggregation facilitates comparisons of average
3 and peak-hour utilization by parking type, offering insight into regional demand patterns and
4 guiding infrastructure planning.

5 Assessing Model Accuracy Using Real-Time Occupancy Data

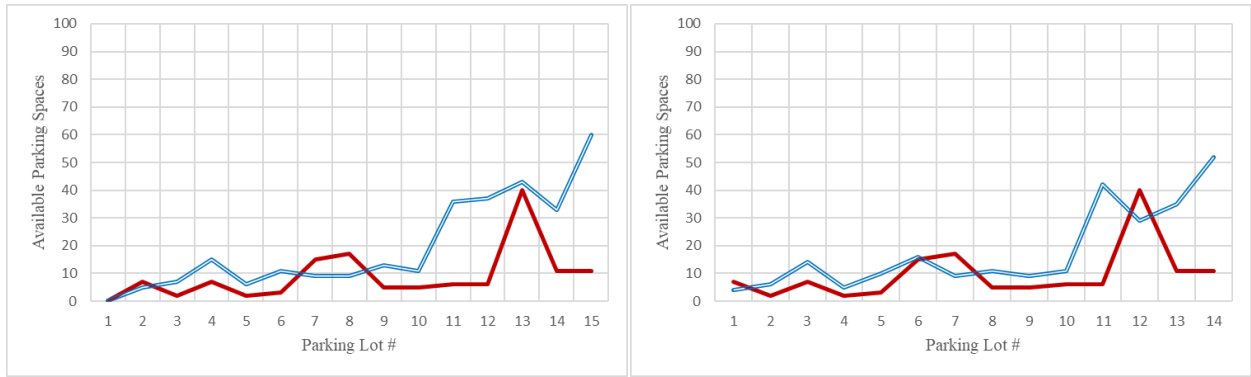
6 Model accuracy is evaluated by comparing predictions with ground truth data from 14 public
7 facilities equipped with sensors, cameras, and entry/exit logs (see Table 4). Data was collected
8 over four weeks (Feb–Mar 2025). Initial baseline availability was estimated using early app user
9 reports and processed per the methodology in the prior section.

10 **Table 4.** List of Public Parking Locations with Real-Time Sensor Validation

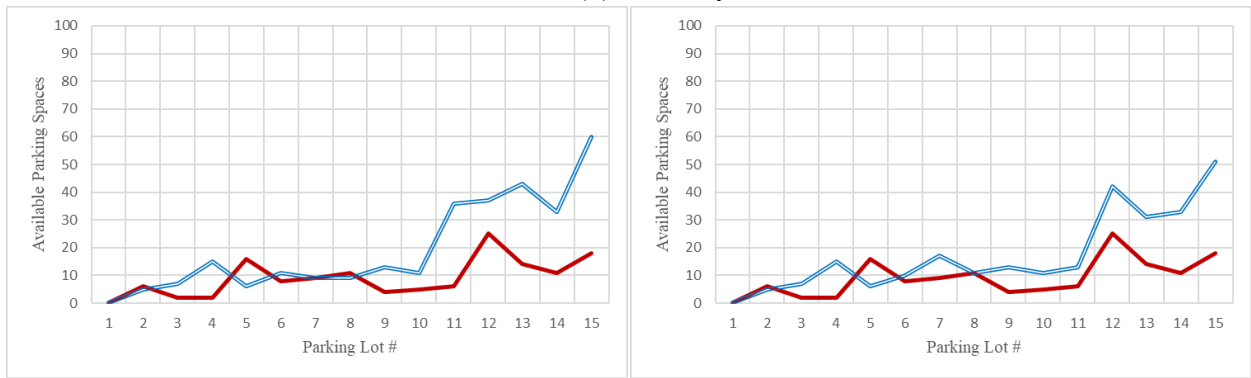
No.	Facility Name	State	Capacity	Cluster Category
1	I-70 WB Grainfield Rest Area	Kansas	10	2
2	I-94 Rest Area WB Dunn County	Wisconsin	24	4
3	I-90 Rest Area NB Juneau County	Wisconsin	25	2
4	I-94 Rest Area EB Jefferson County	Wisconsin	27	2
5	I-75 Welcome Center SB Hamilton	Florida	39	4
6	I-94 Rest Area EB Jackson County	Wisconsin	41	2
7	I-94 Rest Area WB Jackson County	Wisconsin	42	2
8	I-71 SB Weigh Station	Kentucky	45	4
9	Welcome Center - St Croix	Minnesota	53	2
10	I-90 Rest Area Beloit	Wisconsin	60	3
11	I-90 Rest Area SB Columbia County	Wisconsin	68	4
12	I-90 Rest Area Rock County	Wisconsin	76	3
13	I-65 NB Rest Area North Hart County	Kentucky	115	4
14	I-65 SB Rest Area North Hart County	Kentucky	115	4



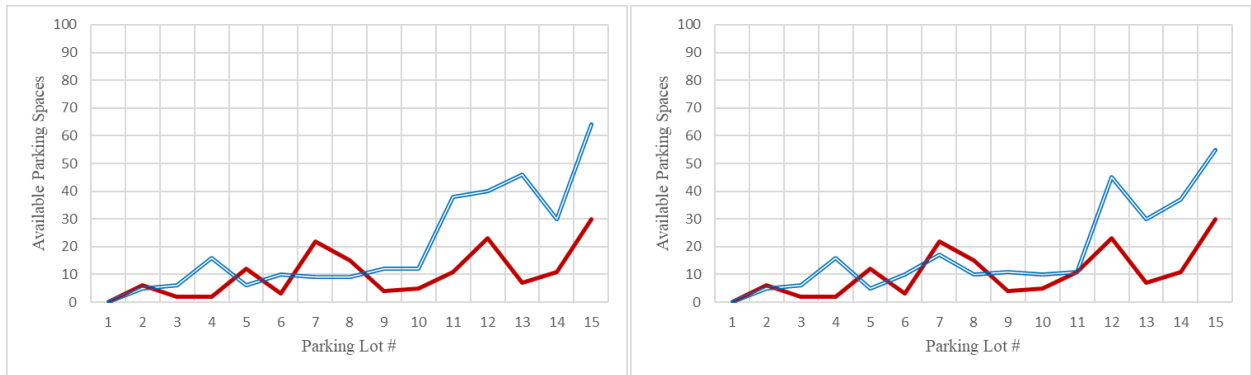
1 (a) Baseline Model (Left Panel) vs. Retrained Model (Right Panel) – Monday



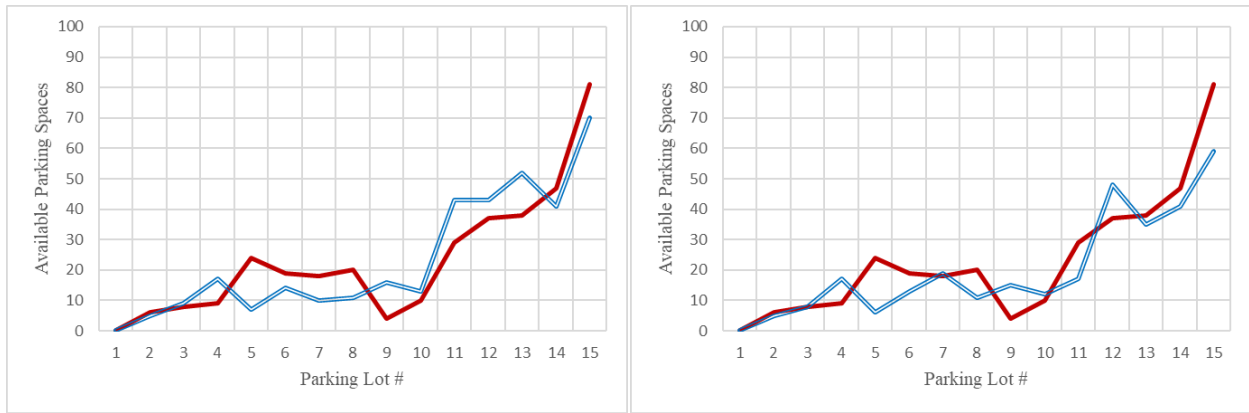
2
3 (b) Tuesday



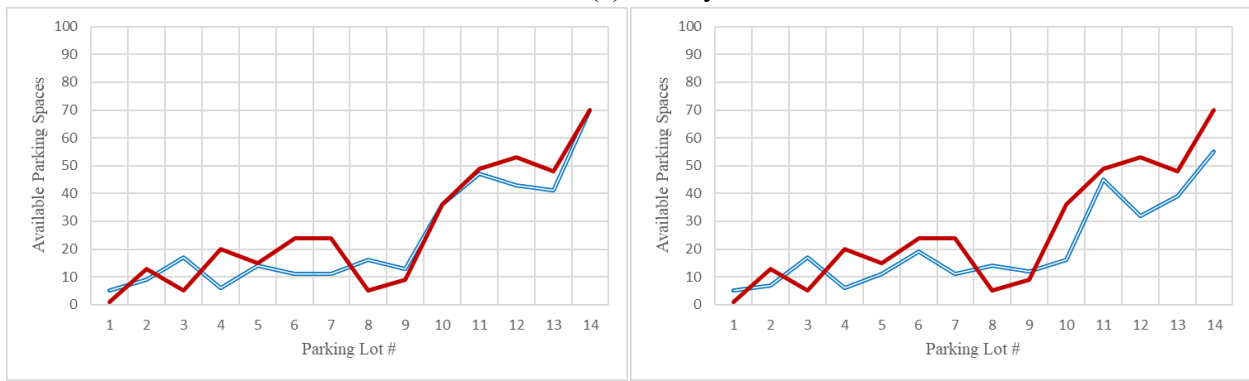
4
5 (c) Wednesday



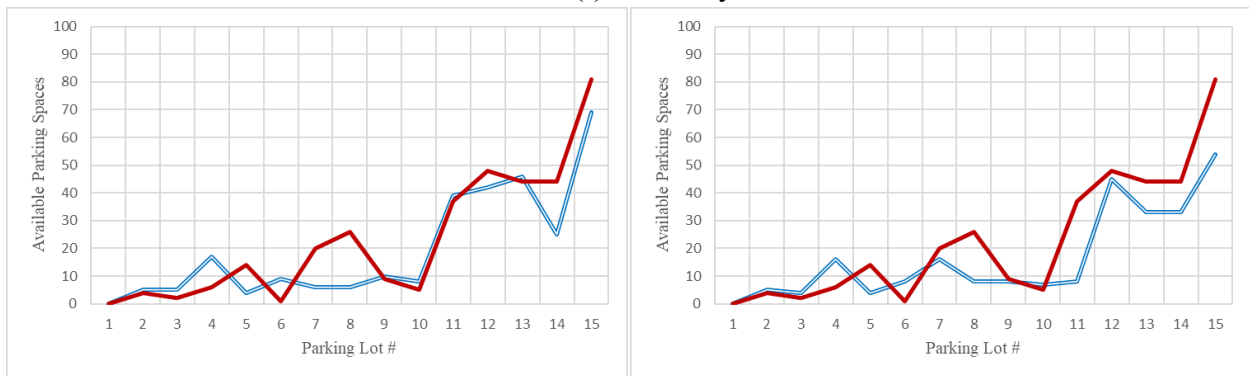
6
7 (d) Thursday



(e) Friday



(f) Saturday



(g) Sunday

Figure 6. Impact of Model Retraining on Parking Availability Prediction

Figure 6 shows how retraining affects parking availability predictions. The baseline (left panel) and retrained models (right panel) were evaluated using one week of new ground-truth data. At some facilities, sensors report availability in binary terms (e.g., "low") once capacity drops below a certain threshold, without providing precise vehicle counts. To address this limitation, availability was assumed to be 10% of total capacity when labeled as "low." Under this assumption, the retrained model achieved a lower mean absolute error (MAE) of 18.6%, compared to 20.5% for the baseline.

1 Despite these improvements, some limitations persist. First, reliance on sparse user-reported
2 data reduces accuracy at facilities with low engagement, where real-time dynamics are poorly
3 captured. Second, the model requires frequent retraining to adapt to changing demand patterns. To
4 address this, the model can be fine-tuned on a short cycle (e.g., weekly or bi-weekly), with
5 retraining frequency dynamically adjusted based on the volume of new user data—balancing
6 accuracy and computational efficiency.

7 **INSTITUTIONAL IMPLICATIONS**

8 **Performance Monitoring with Cluster-Based Classifications**

9 Current state-level classifications of truck parking facilities, which mainly distinguish between
10 public and private types, lack the granularity needed for effective planning and performance
11 tracking. This study’s five-cluster framework offers more actionable categories. For example,
12 states could focus capacity expansion on high-demand clusters (e.g., Cluster 3), which often
13 includes public rest areas with high utilization. Similarly, long-term trends in facility condition
14 and usage could be monitored at the cluster level to better inform infrastructure maintenance and
15 investment. Agencies may pilot cluster-based scorecards to track key metrics—search time,
16 utilization, and safety concerns—for improved accountability. Such scorecards could incentivize
17 data-driven investments and improve transparency in addressing the truck parking challenges.

18 **Interstate Collaboration and System Integration**

19 The need for real-time parking information, as emphasized by Anderson et al. [10] and state
20 truck parking studies [15, 16, 17], remains unmet. While some states have implemented mobile
21 applications and 511 alerts, these state-specific solutions do not adequately address the challenges
22 faced by interstate drivers. Therefore, scaling such initiatives is crucial, ideally to a unified national
23 TPIMS similar to MAASTO, or at least to the corridor or regional level. To achieve this, a federally
24 coordinated task force (e.g., FMCSA, National Coalition on Truck Parking) may unify data
25 standards, integrate public/private inventories, and issue consistent TPIMS guidelines. This system
26 would consolidate real-time availability, overnight parking fees, and amenity data into a unified
27 platform. Standardized protocols for sensor inputs (e.g., camera feeds, embedded sensors,
28 entry/exit logs) would improve predictive accuracy and ensure seamless trip planning across states.

29 **Cost-Effective Data Solutions with Advanced AI Techniques**

30 Physical sensors are costly, particularly for public agencies. A hybrid model—using LSTM-
31 based predictions with crowdsourced reports—offers a scalable alternative. Real-time availability
32 can be collected via mobile apps and refined through feedback loops. Participation may be
33 increased through gamification (e.g., rewards for frequent reporting) or incentives for private
34 partners (e.g., tax incentives for data sharing). In locations with low reporting activity, reliability
35 can be improved through data-driven calibration, which imputes estimates by utilizing time-series
36 patterns from similar facilities within the same cluster.

37 **SUMMARY**

38 The ongoing shortage of safe overnight truck parking poses risks to roadway safety, driver
39 productivity, and freight efficiency in the United States. This study proposes a nationwide

1 framework that utilizes public and private facility data, traffic conditions, infrastructure features,
2 and crowdsourced availability reports. Using clustering techniques, over 10,000 truck parking
3 facilities are categorized into five operationally meaningful groups based on attributes such as
4 amenities, truck traffic, and parking fees. A predictive model then forecasts parking availability
5 across a weekly cycle. Validation against real-time sensor data shows the model's capacity to
6 capture facility-level utilization trends, with a mean absolute error (MAE) of 18.6%.

7 The findings emphasize the potential of a unified national system to integrate fragmented state-
8 level data and provide seamless support for interstate drivers. Addressing the parking crisis
9 requires a coordinated, data-driven approach. Cluster-based performance metrics, shared data
10 standards, and cost-effective AI tools can help transform parking availability into a scalable and
11 safety-enhancing asset. This integrated framework aligns infrastructure governance, technology
12 deployment, and user needs, providing a roadmap for sustainable freight mobility.

13 Several limitations should be noted. First, some important location-based factors, such as
14 proximity to delivery destinations, were excluded due to a lack of granular data (e.g., truck
15 trajectory logs). Second, the accuracy of real-time availability predictions depends on mobile app
16 adoption and consistent user participation. Low reporting rates may limit both data reliability and
17 prediction performance, particularly at smaller or low-traffic facilities. Third, even when reports
18 are available, crowdsourced estimates tend to be subjective and imprecise. A structured reporting
19 interface may be needed to reduce ambiguity. Finally, while the LSTM model captures weekly
20 temporal trends, it does not dynamically adjust to unexpected events such as weather events or
21 road closures. These disruptions can affect parking availability but were not modeled.

22 Future work may explore integrating external sources (e.g., weather APIs), implementing
23 adaptive retraining frequency, and developing layered models to capture both short-term
24 fluctuations and long-term contextual changes, such as seasonal or even-driven shifts.

25

1 **AUTHOR CONTRIBUTION STATEMENT**

2 The authors confirm contributions to the paper as follows: study conception and design: Y. Choi;
3 analysis and interpretation of results: Y. Choi; data collection: Y. Choi; draft manuscript
4 preparation; Y. Choi. The manuscript was prepared with assistance from the Grok 3 and ChatGPT
5 o3 model for code development and language improvement.

6

7 **Declaration of Conflicting Interest**

8 The authors declared no potential conflicts of interest with respect to the research, authorship and
9 publication of this article.

10

11 **Funding**

12 The authors disclosed no financial support for the research, authorship, and publication of this
13 article.

1 REFERENCES

- 2 [1] Boris, C., and Brewster, R. Managing Critical Truck Parking Case Study – Real World
3 Insights from Truck Parking Diaries. American Transportation Research Institute. 2016.
4 <https://trid.trb.org/View/1440565>
- 5 [2] *Jason's Law Truck Parking Survey Results and Comparative Analysis*. Federal Highway
6 Administration, Washington, D.C., 2015.
7 https://ops.fhwa.dot.gov/freight/infrastructure/truck_parking/jasons_law/truckparkingsurvey/jasons_law.pdf
- 8 [3] *Jason's Law Commercial Motor Vehicle Parking Survey and Comparative Assessment*.
9 Federal Highway Administration, Washington, D.C., 2020.
10 https://ops.fhwa.dot.gov/Freight/infrastructure/truck_parking/coalition/2020/mtg/jasons_law_results.pdf
- 11 [4] *Truck Parking Development Handbook*. Federal Highway Administration. 2022.
12 https://ops.fhwa.dot.gov/freight/infrastructure/truck_parking/docs/Truck_Parking_Development_Handbook.pdf
- 13 [5] *Biden-Harris Administration Announces More Than \$80 Million in Grants to Improve
14 Highway Safety Including Better Access to Truck Parking*, Federal Motor Carrier Safety
15 Administration. 2023. <https://www.fmcsa.dot.gov/newsroom/biden-harris-administration-announces-more-80-million-grants-improve-highway-safety>
- 16 [6] *Truck Parking Safety Improvement Act*. U.S. Congress. 2024.
17 <https://www.congress.gov/bill/118th-congress/house-bill/2367/text>
- 18 [7] *Ohio Truck Parking Study*, Ohio Department of Transportation. 2021.
19 <https://www.transportation.ohio.gov/programs/transport-ohio/transport-ohio-respository/ohio-truck-parking-study>
- 20 [8] *Kansas Statewide Freight Network Truck Parking Plan*. Kansas Department of
21 Transportation. 2016.
22 https://ops.fhwa.dot.gov/Freight/infrastructure/truck_parking/coalition/state_reg_lgov_coord/duct/freight_plans.pdf
- 23 [9] Boris, C., and Brewster, R. A Comparative Analysis of Truck Parking Travel Diary Data.
24 *Transportation Research Record*. Vol. 2672(9). 2018.
25 <https://doi.org/10.1177/0361198118775869>
- 26 [10] Anderson, J., Hernandez, S., Jessup, E., North, E. Perceived Safe and Adequate Truck
27 Parking: A Random Parameters Binary Logit Analysis of Truck Driver Opinions in the Pacific
28 Northwest. *International Journal of Transportation Science and Technology*. Vol. 7(1). pp. 89-
29 102. 2018. <https://doi.org/10.1016/j.ijtst.2018.01.001>
- 30 [11] Nevland, E., Gingerich, K., and Park, P. A Data-Driven Systematic Approach for
31 Identifying and Classifying Long-Haul Truck Parking Locations. *Transport Policy*. Vol 96. pp.
32 48-59. 2020. <https://doi.org/10.1016/j.tranpol.2020.04.003>

- 1 [12] Vital, F., and Ioannou, P. Survey on Intelligent Truck Parking: Issues and Approaches.
2 *IEEE Intelligent Transportation Systems Magazine*. Vol. 13(4). pp. 31-44. 2020.
3 <https://doi.org/10.1109/MITS.2019.2926259>
- 4 [13] Sadek, B., Martin, E., Shaheen, S. Forecasting Truck Parking Using Fourier
5 Transformations. *Journal of Transportation Engineering, Part A: Systems*. Vol. 146(8). 2020.
6 <https://doi.org/10.1061/JTEPBS.0000397>
- 7 [14] Yang, H., Liu, C., Zhuang, Y., Sun, W., Murthy, K., and Pu, Z. Truck Parking Pattern
8 Aggregation and Availability Prediction by Deep Learning. *IEEE Transactions on Intelligent*
9 *Transportation Systems*. Vol. 23(8). pp. 12778-12789. 2022.
10 <https://doi.org/10.1109/TITS.2021.3117290>
- 11 [15] *Statewide Truck Parking Study*. Florida Department of Transportation. 2020.
12 <https://www.fdot.gov/rail/studies/truck-parking>
- 13 [16] *Utah Interstate 15 Truck Parking Study*. Utah Department of Transportation. 2012.
14 [https://udot.utah.gov/connect/wp-content/uploads/sites/50/2019/11/Utah-I-15-Truck-Parking-](https://udot.utah.gov/connect/wp-content/uploads/sites/50/2019/11/Utah-I-15-Truck-Parking-Draft-Study.pdf)
15 [Draft-Study.pdf](https://udot.utah.gov/connect/wp-content/uploads/sites/50/2019/11/Utah-I-15-Truck-Parking-Draft-Study.pdf)
- 16 [17] *Rhode Island Statewide Truck Parking Study*. Rhode Island Department of Transportation.
17 2022. [https://planning.ri.gov/sites/g/files/xkgbur826/files/2023-](https://planning.ri.gov/sites/g/files/xkgbur826/files/2023-06/Appendix%20%20-%20Rhode%20Ialand%20Statewide%20Truck%20Parking%20Study.pdf)
18 [06/Appendix%20%20-%20Rhode%20Ialand%20Statewide%20Truck%20Parking%20Study.pdf](https://planning.ri.gov/sites/g/files/xkgbur826/files/2023-06/Appendix%20%20-%20Rhode%20Ialand%20Statewide%20Truck%20Parking%20Study.pdf)
19 [f](https://planning.ri.gov/sites/g/files/xkgbur826/files/2023-06/Appendix%20%20-%20Rhode%20Ialand%20Statewide%20Truck%20Parking%20Study.pdf)
- 20 [18] COTrip. Colorado Department of Transportation. <https://maps.cotrip.org/home>
- 21 [19] *Reducing Big Rig Parking Problems*. UConn Today. University of Connecticut. 2024.
22 <https://today.uconn.edu/2024/05/reducing-big-rig-parking-problems/>
- 23 [20] Prozzi, J., Ajith, A., Holstead, K., Perry, E., Srivastava, A., Hayes, S., and Senthis Ros, C.
24 Assessing Truck Parking Capacity Usage to Inform Truck Parking Needs Assessments and
25 Determine the Feasibility and Benefits of Truck Parking Capacity Management Platforms.
26 Federal Motor Carrier Safety Administration. 2024. <https://rosap.nhtl.bts.gov/view/dot/79224>
- 27 [21] Moore, D., Ivy, G., Comer, B., DeMent, M., Junak, M., and Miller, C. Creating a Roadmap
28 for Successfully Planning, Implementing, and Administering Complex Multi-Jurisdictional
29 Transportation Technology Projects. *Transportation Research Record*. Vol. 2673(11). pp. 764-
30 770. 2019. <https://doi.org/10.1177/0361198119855340>
- 31 [22] *Highway Performance Monitoring System Field Manual*. Federal Highway Administration.
32 2023. https://downloads.regulations.gov/FHWA-2023-0014-0003/attachment_1.pdf