Black Sea Journal of Engineering and Science
doi: 10.34248/bsengineering.2004291

Open Access Journal
e-ISSN: 2619 – 8991

Research Article
Volume 9 - Issue 5: 2817-2834 / September 2026

CONGESTION-TRIGGERED ELECTRIC-VEHICLE CHARGING STATION ASSIGNMENT USING SHORT-TERM CUMULATIVE DEMAND FORECASTS

Oğuzkağan ALIÇ[1]*

[1]*Eskişehir Technical University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 26555, Eskişehir, Türkiye*

Abstract: Short-term congestion at public electric-vehicle charging stations depends on both near-future demand and how vehicles are distributed across the network. We forecast cumulative station-level charging initiations and use the forecasts in congestion-triggered station assignment. The analysis reuses 441077 publicly released charging transactions from 13 stations in Jiaxing, China, aggregated at 15-minute resolution. Daily and weekly seasonal baselines are compared with Poisson, Tweedie, hurdle, and quantile LightGBM models for cumulative 15-, 30-, and 60-minute demand. Because physical arrivals and queueing times are not observed, recorded charging starts served as simulation-entry events in the counterfactual analysis. The forecasts are embedded in a counterfactual capacity-and-routing simulation with transaction-level service durations, inbound reservations for redirected vehicles, partial recommendation acceptance, an eight-minute switching-time penalty, queue abandonment, demand growth, and charger outages. All routing policies are evaluated on immutable scenarios with common random numbers over 21 test days stratified by weekday and charging-initiation volume. Relative to the stronger daily seasonal baseline, the Tweedie model reduces mean absolute error by 17.6-19.8% and root-mean-square error by 21.7-26.8%. Under normal demand and full capacity, the routing policies remain inactive and introduce no unnecessary switching. Under the evaluated stress scenarios and full recommendation acceptance, dynamic assignment reduces simulated mean total delay by 75.0-78.7% and simulated abandonment by approximately 95-100%, while rerouting fewer than 10% of vehicles. Risk-aware routing has the numerically lowest aggregate simulated delay, but under the combined demand-and-outage scenario it differs from current-load routing by only 0.001 minutes per vehicle, and the pairs 95% confidence interval included zero. Within these counterfactual scenarios, most of the reduction therefore come from the shared congestion-triggered routing architecture rather than from the forecast-risk terms.

Keywords: Electric-vehicle charging, Cumulative demand forecasting, Sparse count prediction, Probabilistic forecasting, Charging-station assignment, Congestion management

***Corresponding author:** Eskişehir Technical University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 26555, Eskişehir, Türkiye
E mail: oguzkaganalic@eskisehir.edu.tr (O. ALIÇ)
Oğuzkağan ALIÇ https://orcid.org/0000-0001-8611-7987

Received: July 27, 2026
Accepted: August 27, 2026
Published: September 15, 2026

Cite as: Alıç, O. (2026). Congestion-triggered electric-vehicle charging station assignment using short-term cumulative demand forecasts. *Black Sea Journal of Engineering and Science,* 9(5), 2817-2834.

1. Introduction

The continued electrification of road transport increases the need for reliable and efficiently utilized public charging infrastructure. Charging demand varies considerably across locations, user groups, time periods, and infrastructure types, creating challenges for both power system operation and service quality (Kim et al., 2026; Alaraj et al., 2025; Unterluggauer et al., 2021; Elahe et al., 2022). Even when the overall network has sufficient charging capacity, uneven spatial and temporal demand can overload individual stations, increase waiting times, and leave nearby chargers underutilized (Schoenberg and Dressler, 2023; Widmann and Sawodny, 2026; Zhang et al., 2025a) These localized imbalances become particularly critical during periods of high demand or charger outages, when congestion at a small number of stations can spread throughout the network.

Driver behavior further complicates the management of public charging networks. Station choice depends on several factors, including waiting time, charging speed, price, remaining driving range, and detour distance, while users differ in both their tolerance for waiting and their willingness to accept alternative charging stations (Hoen et al., 2023; Hanni et al., 2024; Brückmann and Bernauer, 2023). Consequently, a routing policy that minimizes station queues without considering additional travel or user inconvenience may reduce queueing delays while increasing the overall burden on drivers. Likewise, a recommendation that initially appears beneficial can become counterproductive if many vehicles are redirected to the same station at the same time. Operational station assignment must consequently account for current occupancy, near-future charging initiations, already redirected vehicles, recommendation

This work is licensed (CC BY-NC 4.0) under Creative Commons Attribution 4.0 International License

acceptance, and switching cost.

Short-term demand forecasting provides valuable information for making proactive assignment decisions. Existing studies predict charging power, energy consumption, station occupancy, session duration, regional demand, and future load profiles using statistical methods, gradient boosting, recurrent neural networks, convolutional architectures, transformers, and ensemble models (Unterluggauer et al., 2021; Yi et al., 2021; Mohammad et al., 2023; Aghsaee et al., 2023; Hussain et al., 2025; Mansour et al., 2026). Although these forecasting targets are valuable for infrastructure planning and grid operation, they do not directly represent the number of vehicles expected to initiate charging within a short queue-formation period. In addition, station-level charging counts at fine temporal resolutions are discrete, non-negative, sparse, and often overdispersed, making count-based and probabilistic forecasting approaches more appropriate than conventional point regression models.

Charging-station recommendation evolves alongside advances in charging-demand forecasting. Existing approaches combine factors such as travel distance, travel time, charging price, charger availability, expected waiting time, reservations, and announced charging intentions with optimization techniques, reinforcement learning, and adaptive routing strategies (Liu et al., 2018; Lee et al., 2020; Basmadjian et al., 2020; Schoenberg and Dressler, 2023; Nie et al., 2023; Dudkina et al., 2025; Widmann and Sawodny, 2026). Collectively, these studies demonstrate that coordinated station assignment can effectively reduce congestion. However, their practical performance largely depends on how future vehicle commitments are represented. If vehicles that accept a recommendation but have not yet reached their assigned station are excluded from the system state, successive routing decisions may redirect multiple vehicles to the same station, creating new congestion. Reservation- and intention-aware approaches address this issue by explicitly representing future arrivals. However, they often rely on formal reservation systems, detailed trip information, full user compliance, or repeated global optimization, which may limit their practical applicability (Basmadjian et al., 2020; Schoenberg and Dressler, 2023; Dudkina et al., 2025; Widmann and Sawodny, 2026).

A second methodological issue concerns policy evaluation. Simulation outcomes can be distorted when competing policies receive different random event streams, service durations, outages, or behavioral draws. This issue becomes particularly important when the performance differences among advanced routing strategies are much smaller than their shared improvement over the no-routing baseline. A fair comparison therefore requires all policies to be evaluated under identical stochastic conditions, together with paired statistical analyses at the scenario level.

This study addresses these challenges by integrating short-term demand forecasting with congestion-aware station assignment in a unified simulation framework based on the publicly available Jiaxing charging-transaction dataset introduced by (Zhang et al., 2025a). The study relies exclusively on secondary analysis of the published dataset, and no primary charging data are collected. The dataset records charging-start times but does not provide reliable physical vehicle-arrival or observed queueing times; accordingly, the simulation is formulated as a counterfactual capacity-and-routing analysis based on observed charging initiations. Station-level charging initiations are forecasted as cumulative counts over the next 15, 30, and 60 minutes using seasonal baselines, Poisson and Tweedie gradient boosting, a two-stage hurdle model, and quantile regression. These forecasts drive a congestion-triggered assignment mechanism that preserves immediately available station capacity, accounts for redirected vehicles through inbound reservations, applies a minimum-benefit threshold before rerouting, models partial user acceptance of recommendations, and incorporates a fixed switching penalty into the total-delay objective. Four policies, no routing, current-load routing, predicted-wait routing, and risk-aware routing, are evaluated under observed demand, a 20% demand increase, a 10% charger-outage probability, and their combination using immutable scenarios and common random numbers (CRN).

This study makes three main contributions. First, it aligns the forecasting task with the operational decision by predicting cumulative charging-start counts over short queue-formation windows rather than isolated future time slots or energy demand. Second, it combines these forecasts with a selective, inbound-aware routing mechanism that reduces unnecessary station switching and mitigates recommendation-induced herding. Third, it evaluates the proposed routing strategies using a reproducible paired simulation framework with fixed scenario inputs, integrity checks, and day-level bootstrap confidence intervals. The analysis is designed to determine not only whether dynamic routing reduces the simulated outcomes relative to no routing under the stated scenarios, but also whether forecast and uncertainty information provides a meaningful incremental reduction beyond a simpler current-load policy.

The experimental analysis addresses the following research questions:

RQ1: Do count-oriented models improve cumulative station-level point forecasts relative to seasonal baselines, and how well calibrated are the corresponding probabilistic forecasts?

RQ2: Does congestion-triggered, inbound-aware assignment reduce simulated total delay and queue abandonment under demand and capacity stress without causing unnecessary routing under normal conditions?

RQ3: Does risk-aware forecast information provide a practically meaningful advantage over current-load and predicted-wait routing?

1.1. Charging-Demand Forecasting and Data Foundations

EV charging-demand forecasting spans statistical, machine-learning, deep-learning, and ensemble approaches, with model choice depending on prediction horizon, charging context, spatial scale, data availability, and operational objective (Kim et al., 2026; Alaraj et al., 2025). Statistical and seasonal models remain useful as transparent benchmarks, whereas recurrent, convolutional, transformer, and hybrid architectures are increasingly used to represent nonlinear temporal and spatial structure (Unterluggauer et al., 2021; Mohammad et al., 2023; Hussain et al., 2025; Mansour et al., 2026; Yu et al., 2026). For example, prior studies have forecast interval-level charging load, energy demand, monthly regional demand, session duration, connector occupation, and multi-step load sequences (Yi et al., 2021; Unterluggauer et al, 2021; Aghsaee et al., 2023; Orzechowski et al., 2023). These targets are valuable for planning and grid operation, but they do not directly represent the number of charging initiations that may compete for limited charger capacity during an imminent routing window. Gradient-boosting methods offer a comparatively lightweight alternative to large sequence models. LightGBM, XGBoost, CatBoost, and random forests can represent nonlinear interactions among station identity, calendar variables, recent demand, and lagged observations without requiring long recurrent training pipelines (Ke et al., 2017; Sreekumar and Lekshmi, 2024; Aghsaee et al., 2023; Mansour et al., 2026; Keskin et al., 2024). Previous-slot, previous-day, and real-time station-state features have been shown to contribute materially to charging-related forecasts (Sreekumar and Lekshmi, 2024; Aghsaee et al., 2023). These findings support evaluating count-oriented gradient boosting for short operational horizons, while retaining seasonal baselines to determine whether the additional model complexity produces meaningful improvement. Forecast comparisons are also strongly affected by dataset construction and evaluation design. Public charging resources vary in duration, station diversity, temporal resolution, transaction volume, and the availability of contextual and operational variables (Elahe et al., 2022; Zhang et al., 2025a). Chronological holdout, walk-forward validation, and cross-site testing are therefore preferable to random splitting when repeated seasonal patterns and temporally adjacent observations could otherwise appear in both training and evaluation sets (Unterluggauer et al., 2021; Orzechowski et al., 2023; Mansour et al., 2026; Ali et al., 2026). Differences in preprocessing, target construction, aggregation, and partition boundaries nevertheless continue to limit direct comparisons across studies (Kim et al., 2026; Alaraj et al., 2025). The present analysis uses the public Jiaxing transaction dataset introduced by (Zhang et al., 2025a). It contains 441077 charging sessions from 13 stations and includes charging-post identifiers, charging-start and end timestamps, delivered energy, transaction attributes, weather variables, and station categories. Its temporal precision supports station-level charging-initiation counts and empirical service-duration estimates, while its multiple stations preserve operational heterogeneity. However, the dataset does not provide reliable physical vehicle-arrival times, observed queueing times, route-specific travel times, or rejected charging attempts. It therefore supports a counterfactual capacity-and-routing analysis based on observed charging initiations, not reconstruction of historical queues.

1.2. Sparse Counts and Probabilistic Forecasting

Station-level charging-initiation counts at 15-minute resolution are non-negative, discrete, and frequently zero. Most directly comparable EV forecasting studies instead model continuous power or energy using conventional point-regression objectives (Unterluggauer et al., 2021; Elahe et al., 2022; Mohammad et al., 2023; Sreekumar and Lekshmi, 2024; Hussain et al., 2025; Mansour et al., 2026). For sparse count targets, Poisson and Tweedie objectives provide non-negative conditional-mean estimates and can represent demand-dependent variance, while hurdle models separate the probability of positive demand from its conditional magnitude. Because the EV literature does not establish a universally preferable count formulation, empirical comparison among these alternatives is more appropriate than assuming a single distributional model (Kim et al., 2026; Alaraj et al., 2025). Probabilistic forecasts provide information that is not available from a conditional mean alone. Multi-quantile charging-demand models have been evaluated using coverage and interval-width measures, including applications involving site transfer and locations with limited local data (Ali et al., 2026). More general electrical-load research likewise uses quantile-based methods to evaluate reliability and sharpness jointly (Dang et al., 2022). These studies support the use of lower, median, and upper conditional quantiles, but they also require careful interpretation when the target is zero-heavy. In particular, a conditional median can achieve low absolute error while underestimating aggregate expected demand when the distribution contains many zeros and a long positive tail. Median and upper-quantile forecasts should therefore complement, rather than automatically replace, expected-demand estimates in operational control. This distinction matters for station assignment because a routing score may require both expected near-future pressure and an upper-tail risk margin. Prior work has incorporated forecast uncertainty into pricing and energy-procurement decisions (Matkovic et al., 2025), but its use in short-horizon station assignment with finite capacity, current queues, inbound commitments, and switching costs remains less developed (Schoenberg and Dressler, 2023; Dudkina et al., 2025; Widmann and Sawodny, 2026).

1.3. Queue-Aware Station Assignment and Behavioral Assumptions

Charging stations can be represented as multi-server systems in which vehicles compete for finite service capacity. Analytical queueing models offer computational tractability but depend on assumptions about arrival and service processes (Heting et al., 2021; Widmann and Sawodny, 2026). Discrete-event and agent-based simulations can retain heterogeneous charger counts, empirical or vehicle-specific service durations, traffic concentration, and operational disturbances (Mishra et al., 2022; Rupnik et al., 2025; Schoenberg and Dressler, 2023). Across these studies, local congestion can occur even when network-wide capacity is adequate because demand and service times vary substantially among stations. Station-assignment methods consequently combine current availability with travel time, price, expected waiting, reservations, announced intentions, or centralized coordination (Liu et al., 2018; Lee et al., 2020; Basmadjian et al., 2020; Schoenberg and Dressler, 2023; Nie et al., 2023; Dudkina et al., 2025; Widmann and Sawodny, 2026). Reservation- and intention-aware methods are particularly relevant because a vehicle that has accepted a recommendation but has not yet reached the destination already represents a future capacity commitment. Ignoring such commitments can cause several decisions to select the same apparently uncongested station, whereas explicitly representing planned demand can reduce congestion migration (Schoenberg and Dressler, 2023; Widmann and Sawodny, 2026). However, existing approaches may require formal bookings, detailed trip plans, full compliance, accurate arrival-time estimates, repeated optimization, or interoperable reservation systems (Basmadjian et al., 2020; Dudkina et al., 2025; Widmann and Sawodny, 2026). Behavioral assumptions also affect apparent policy performance. Waiting time, detour distance, price, charging speed, amenities, remaining range, and user characteristics influence station choice and waiting tolerance (Hoen et al., 2023; Hanni et al., 2024; Brückmann and Bernauer, 2023). A policy that minimizes destination queueing without charging for additional travel can therefore appear beneficial while increasing the user's total burden. Partial recommendation acceptance and an explicit switching cost provide a more conservative evaluation than universal compliance and costless redirection, although fixed recommendation-acceptance rates, patience thresholds, and switching parameters remain simplifications when individual-level behavioral and route data are unavailable. Stress conditions and evaluation design are equally important. Demand growth, charger loss, seasonal concentration, and infrastructure failures can reveal differences that remain dormant under uncongested operation (Mishra et al., 2022; Rupnik et al., 2025; Zhang et al., 2025b). Yet comparisons can be confounded when policies receive different event streams, service durations, outages, or behavioral draws. Common random numbers and paired day-level inference are therefore especially useful when differences among dynamic policies are expected to be much smaller than their common difference from no routing. The reviewed assignment literature uses Monte Carlo and large-scale simulation extensively, but immutable scenario objects, input-integrity checks, and paired bootstrap comparisons are not usually presented together as a policy-evaluation framework (Schoenberg and Dressler, 2023; Dudkina et al., 2025; Widmann and Sawodny, 2026).

1.4. Research Gap and Study Positioning

Three connected gaps motivate the present study. First, much of the forecasting literature focuses on electrical load, energy, occupation, charging duration, or isolated future intervals rather than cumulative charging-initiation counts over the next 15-60 minutes (Unterluggauer et al., 2021; Yi et al., 2021; Aghsaee et al., 2023; Mohammad et al., 2023; Orzechowski et al., 2023; Hussain et al., 2025). Such cumulative counts are more directly aligned with the number of observed charging initiations that may compete for finite station capacity during a short decision window. Second, probabilistic demand forecasts are commonly evaluated separately from real-time station assignment. Existing reservation- and intention-aware methods demonstrate the importance of future commitments, but frequently rely on formal reservations, submitted travel plans, idealized compliance, or iterative global optimization (Basmadjian et al., 2020; Schoenberg and Dressler, 2023; Dudkina et al., 2025; Widmann and Sawodny, 2026). A lighter-weight alternative is to combine current station state with short-horizon expected demand, an upper-tail risk signal, and explicit counters for vehicles already redirected but not yet at the destination. Third, operational conclusions require controlled comparison. Stress scenarios, partial recommendation acceptance, switching costs, abandonment, and charger failures each affect results, while small differences among dynamic policies can be obscured by independently generated simulation inputs. The present framework therefore evaluates no routing, current-load, predicted-wait, and risk-aware assignment on immutable common-random-number scenarios with integrity checks and paired day-level inference. The contribution is an integration and evaluation framework rather than a new standalone forecasting architecture. It links cumulative count forecasts to selective inbound-aware routing and tests whether expected-demand and uncertainty terms provide material value beyond a simpler current-load policy. This positioning allows the study to assess both the shared effect of congestion-triggered routing and the incremental value of forecast complexity without claiming novelty for the individual forecasting, queueing, or reservation components.

2. Materials and Methods

3.1. Problem Formulation and Study Design

Consider a network of charging stations indexed by $j \in \mathcal{J}$ where station j has effective capacity c_j. Time is divided into 15-minute intervals indexed by t. Let $A_{j,t}$ denote the observed number of charging initiations at station j during interval t, and let $Y_{j,t}^{(h)}$, denote cumulative demand during the next h intervals. The forecasting task estimates the conditional distribution of $Y_{j,t}^{(h)}$ for $h \in \{1,2,4\}$ from station identity, calendar variables, lagged demand, and rolling demand statistics.

At each simulation-entry event, corresponding to an observed charging initiation, an assignment policy selects either the historically observed station or an alternative station. The decision uses current station occupancy, queue length, inbound reservations, service-duration information, and, for forecast-aware policies, cumulative demand forecasts. A recommendation is applied only when the vehicle accepts it and when the estimated score improvement meets the minimum routing-gain threshold after switching time has been included. The primary objective is to minimize mean total user delay (equation 1),

$$\overline{T}^{\text{delay}} = \frac{1}{N}\sum_{i=1}^{N}(W_i^{\text{queue}} + T_i^{\text{switch}}), \quad (1)$$

where N is the number of simulated vehicles, W_i^{queue} is the queueing time of vehicle i, and T_i^{switch} is the additional time incurred when the vehicle is redirected. Equation 1 defines the primary system objective, subject to station-capacity constraints and non-negative inbound reservations. Queue abandonment and rerouting frequency are treated as complementary outcomes.

The study uses a two-stage counterfactual design. Forecast accuracy is evaluated first; the resulting predictions are then supplied to the discrete-event simulation. The published transactions provide station-level charging-initiation histories, empirical service durations, and station-capacity proxies. Because physical vehicle-arrival times are unavailable, each recorded charging initiation is used as a simulation-entry event. Each routing policy is evaluated on the same CRN scenario, so differences in simulated outcomes arise from policy decisions rather than different charging-initiation event streams, service durations, or outage realizations.

3.2. Data Source and Preparation

The empirical analysis reuses the high-resolution electric-vehicle charging transaction dataset originally collected, anonymized, and publicly released by Zhang et al. (2025b). The present study did not participate in primary data collection, anonymization, or construction of the source dataset. According to the dataset descriptor, the source contains 441077 charging transactions recorded at 13 charging stations in Jiaxing, China, over approximately two years, from December 2019 to December 2021. The published records include anonymized user and charging-post identifiers, order-creation time, charging start and end times, delivered energy, charging-price information, transaction status, termination reason, weather variables, and station-location information.

For the present secondary analysis, the published transaction-level records are used to derive station-level demand time series and empirical charging-duration distributions. A total of 92 distinct charging-post identifiers are observed among records with valid service information. For each station j, the capacity proxy c_j is calculated as the number of distinct charging-post identifiers associated with that station in the valid service records. This procedure yields station capacities of four to eight posts. It assumes that each identifier represents one independently usable charger and that the set of identifiers observed over the study period is an appropriate proxy for simultaneous station capacity; temporal installation, removal, or prolonged unavailability of individual posts cannot be reconstructed from the transaction records.

The source dataset has been anonymized before public release by its original providers and authors (Zhang, et al., 2025). No new personal data are collected, and all analyses are conducted on the publicly available anonymized records.

Because demand forecasting and service-duration estimation require different validity criteria, two analysis-specific subsets are derived from the published records. The demand-analysis subset includes all 441077 transactions with valid station identifiers and charging-start timestamps. The charging-start timestamp is selected as the operational demand-event timestamp because the order-creation timestamp is later than the charging-start timestamp in 248588 records, corresponding to 56.36% of the raw dataset. This pattern indicates that neither available timestamp can be interpreted consistently as the physical vehicle-arrival time.

Accordingly, the modeled demand represents observed charging initiations, not physical station arrivals. The source data do not identify vehicles that arrived but left without initiating a recorded charging session.

For the service-duration subset, the duration of transaction i was calculated as (equation 2),

$$D_i = t_i^{\text{end}} - t_i^{\text{start}}, \quad (2)$$

where t_i^{start} and t_i^{end} denote the recorded charging start and end times, respectively.

Transactions are excluded from service-duration estimation when the calculated duration is non-positive or shorter than one minute. Sessions longer than 24 hours are also treated as invalid. After applying these criteria, 422053 transactions remain for service-duration modeling and simulation. The resulting transaction-level duration D_i is retained as the service requirement of that charging-initiation event in the simulation.

Energy values are retained for descriptive analyses when

they are between 0.01 and 200 kWh. Energy validity is not required for constructing the charging-initiation time series because the forecasting target is the number of charging initiations rather than delivered energy.

The order-to-start difference is inspected only as a data-quality diagnostic. Because it may include authentication, connection, communication, payment, or recording processes, it is not interpreted as physical queueing time and was not used as a simulation outcome.

3.3. Forecasting Design and Evaluation

Charging demand is aggregated into fixed 15-minute intervals. Let $A_{j,t}$ denote the number of charging sessions starting at station *j* during interval *t*.

The forecasting targets are defined as cumulative future demand. For a forecasting horizon of *h* intervals, the target was (equation 3),

$$Y_{j,t}^{(h)} = \sum_{k=1}^{h} A_{j,t+k}\,, h \in \{1,2,4\} \quad (3)$$

The cumulative target in equation 3 used *h=1*, *h=2*, and *h=4* to represent the total number of charging initiations expected during the next 15, 30, and 60 minutes, respectively. This cumulative formulation is selected because the total number of near-future charging initiations is more directly related to counterfactual capacity competition than demand in a single isolated future interval.

Predictor variables are generated independently for each station. Historical demand lags are constructed at 1, 2, 4, 8, 96, 192, and 672 intervals. At a 15-minute resolution, these corresponded to 15 minutes, 30 minutes, one hour, two hours, one day, two days, and one week.

Rolling demand statistics are calculated over windows of 4, 12, 24, and 96 intervals, corresponding to one, three, six, and 24 hours. The rolling variables include historical mean and dispersion measures calculated only from observations available at or before the forecast origin.

Calendar variables include hour of day, 15-minute slot of day, day of week, month, and weekend status. Cyclical time variables are represented using sine and cosine transformations. Calendar features associated with the prediction target are derived from the target timestamp, whereas lagged and rolling demand features are derived exclusively from the forecast-origin timestamp.

Station identity is included as a categorical predictor to account for differences in location, infrastructure, and historical utilization.

The 15-minute station panel extends from 31 December 2019 at 21:15 to 1 January 2022 at 00:00. The observations are partitioned chronologically using 65% of the panel duration for training, 15% for validation, and 20% for testing. The validation period begins on 20 April 2021 at 02:15, and the test period begins on 7 August 2021 at 18:30.

Assignment to each subset is based on the end timestamp of the cumulative target window rather than the feature timestamp. Target windows ending before 20 April 2021 at 02:15 are assigned to training; windows ending from that timestamp to, but not including, 7 August 2021 at 18:30 are assigned to validation; and windows ending at or after the test boundary are assigned to testing. This rule prevents observations whose predictors occur before a partition boundary but whose targets extend beyond it from leaking future information across the training, validation, or test sets.

The validation subset is used as the evaluation set for early stopping during every LightGBM fit. Model parameters are estimated using only the training subset, and the validation observations are not subsequently merged into model fitting. The test subset remains untouched until final forecast generation and is not used for early stopping, hyperparameter adjustment, or selection of the reported model families. No random train-test split is used because random splitting would allow temporally adjacent observations and repeated seasonal patterns to occur in both training and evaluation subsets.

The model comparison includes seasonal reference methods, count-oriented gradient boosting, a two-stage hurdle formulation, and quantile regression. Daily and weekly seasonal-naive models serve as the reference methods. For horizon *h*, the daily seasonal prediction is aligned with the target interval as (equation 4),

$$\hat{Y}_{j,t}^{(h,\text{daily})} = Y_{j,t-96}^{(h)} \quad (4)$$

and the weekly seasonal prediction is (equation 4),

$$\hat{Y}_{j,t}^{(h,\text{weekly})} = Y_{j,t-672}^{(h)} \quad (5)$$

This target-aligned definition ensures that each cumulative future window is compared with the equivalent window one day or one week earlier.

For count-based prediction, a LightGBM regression model with a Poisson objective is trained to estimate conditional expected demand. The Poisson model is included because the response variable is a non-negative count. All LightGBM models use a maximum of 700 boosting iterations, a learning rate of 0.04, 31 leaves, a minimum of 40 observations per child node, a row-subsampling fraction of 0.90, a feature-subsampling fraction of 0.90, and an L2 regularization coefficient of 0.20. These common tree settings are fixed in the analysis configuration rather than chosen through an automated validation-set search. For each horizon and objective, LightGBM monitors its objective-specific validation loss, retains the best boosting iteration, and stops after 50 consecutive rounds without improvement. The global model seed is 2026.

Because the charging counts exhibit excess zeros and variance greater than the mean, a LightGBM Tweedie model is also evaluated. The Tweedie variance power is fixed at 1.35 in the analysis configuration; it is not selected by optimizing test performance. The Poisson, Tweedie, hurdle-classification, hurdle-positive-count, and quantile models otherwise use the same common

tree settings. The quantile objectives use levels 0.10, 0.50, and 0.90. All prespecified model families are retained for final test-set comparison rather than selecting a single forecasting family on the validation results. Table 1 consolidates these forecasting settings with the simulation and routing parameters and states the basis on which each group is fixed.

In addition, a hurdle model represents the demand-generation process in two stages. First, a LightGBM classifier with the binary objective estimated the probability of at least one charging initiation (equation 6):

$$\hat{p}_{j,t}^{(h)} = \Pr\left(Y_{j,t}^{(h)} \middle| > 0 \boldsymbol{x}_{j,t}\right). \tag{6}$$

Second, a LightGBM regressor with the Poisson objective is fitted only to positive-count training observations, $Y_{j,t}^{(h)} > 0$, to estimate the conditional positive magnitude. Its validation evaluation set likewise contains only positive-count validation observations. The classifier and positive-count regressor use the common LightGBM settings reported above: at most 700 boosting iterations, learning rate 0.04, 31 leaves, minimum child size 40, row and feature subsampling fractions of 0.90, L2 regularization 0.20, seed 2026, and early stopping after 50 non-improving validation rounds. The expected demand is calculated as (equation 7),

$$\hat{Y}_{j,t}^{(h,\text{hurdle})} = \hat{p}_{j,t}^{(h)} \widehat{\mathbb{E}}\left[Y_{j,t}^{(h)} \middle| Y_{j,t}^{(h)} \middle| > 0 \boldsymbol{x}_{j,t}\right]. \tag{7}$$

The area under the receiver operating characteristic curve (AUC) is used to evaluate the binary occurrence component. The hurdle estimate is retained as the expected-demand signal for the routing implementation because its explicit occurrence-magnitude decomposition matched the zero-heavy charging-initiation target and produces an interpretable non-negative expected count. This is a structural and interpretability-based implementation choice rather than selection of the model with the lowest validation or test error; the Tweedie model is evaluated independently and achieves slightly better point-forecast accuracy. The one-hour cumulative horizon is the longest of the three horizons declared in the analysis configuration and represents demand pressure over four consecutive 15-minute intervals, rather than an isolated near-term slot.

To characterize predictive uncertainty, separate LightGBM quantile models are trained for the 0.10, 0.50, and 0.90 conditional quantiles. These models generate a lower demand bound $\hat{Q}_{0.10}$, a median forecast $\hat{Q}_{0.50}$, and an upper-risk forecast $\hat{Q}_{0.90}$. Quantile predictions are monotonically reordered when necessary to ensure $\hat{Q}_{0.10} \le \hat{Q}_{0.50} \le \hat{Q}_{0.90}$.

The conditional median is evaluated as a point forecast, while the difference between the upper and median quantiles is used as an uncertainty indicator in the risk-aware station-assignment policy.

For n_f forecast observations, point forecasts are evaluated using mean absolute error, root-mean-square error, weighted absolute percentage error, the coefficient of determination, and mean forecast bias (equation 8):

$$\begin{aligned}
\text{MAE} &= \frac{1}{n_f}\sum_{i=1}^{n_f}|y_i - \hat{y}_i|, \\
\text{RMSE} &= \sqrt{\frac{1}{n_f}\sum_{i=1}^{n_f}(y_i - \hat{y}_i)^2}, \\
\text{WAPE} &= \frac{\sum_{i=1}^{n_f}|y_i - \hat{y}_i|}{\sum_{i=1}^{n_f} y_i}, \\
R^2 &= 1 - \frac{\sum_{i=1}^{n_f}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_f}(y_i - \bar{y})^2}, \\
\text{Bias} &= \frac{1}{n_f}\sum_{i=1}^{n_f}(\hat{y}_i - y_i).
\end{aligned} \tag{8}$$

WAPE is not bounded above by one. A value greater than one indicates that the sum of absolute forecast errors exceeds the total number of observed charging initiations over the evaluation set. For the nominal 80% interval $[L_i, U_i] = [\hat{Q}_{0.10,i}, \hat{Q}_{0.90,i}]$, calibration is assessed using empirical coverage, mean interval width, coverage deviation from 0.80, and the interval score (equation 9),

$$\begin{aligned}
\text{IS}_{0.20,i} &= (U_i - L_i) + 10(L_i - y_i)\mathbf{1}(y_i < L_i) \\
&+10(y_i - U_i)\mathbf{1}(y_i > U_i).
\end{aligned} \tag{9}$$

The reported interval score is the mean, $n_f^{-1}\sum_{i=1}^{n_f}\text{IS}_{0.20,i}$. Lower mean interval scores indicate a preferable combination of sharpness and calibration.

3.4. Discrete-Event Charging Simulation

A discrete-event simulation represents the charging stations as parallel-server queueing systems. Each station j contains c_j parallel service resources, where c_j is the station-specific capacity proxy defined as the number of distinct charging-post identifiers observed for that station in the valid service records.

Each simulated vehicle is characterized by a simulation-entry timestamp, an original station, a charging-service duration, a fixed waiting tolerance, a recommendation-acceptance draw, and an optional alternative assigned station.

The recorded charging-start timestamp is interpreted as a charging-initiation event and used as the simulation-entry time. It is not treated as the vehicle's observed physical arrival time. Consequently, the simulation is a counterfactual capacity-and-routing analysis based on observed charging initiations, not a reconstruction of historical station arrivals, rejected charging attempts, or real-world queues.

A vehicle immediately enters service when a charger is available. Otherwise, it joins the station queue. Vehicles abandon the queue when their waiting time exceeds 30 minutes. The 30-minute threshold is used as a transparent finite waiting-tolerance scenario for identifying severe service delay; it is not estimated from individual driver behavior and is not interpreted as a universal patience value.

Each original vehicle retains its transaction-level observed service duration D_i even when it is redirected to another station. Redirection changes the destination and adds the modeled switching interval, but it does not replace D_i with the destination station's median duration. The redirected vehicle therefore occupies one destination charger for the same service duration as its source transaction. Additional vehicles introduced in the demand-growth scenarios inherit the service duration of the sampled source transaction. This implementation assumes that required charging time is unaffected by destination-specific charger power, connector compatibility, or charging conditions, which are unavailable in the dataset. The system state is monitored at 15-minute intervals.

3.5. Immutable Scenario Generation and Common Random Numbers

To ensure fair policy comparisons, simulation scenarios are generated independently of the assignment policy. For each combination of simulation date, demand multiplier, charger-outage probability, and random seed, an immutable scenario instance is generated.

Each scenario stores the complete charging-initiation event stream used for simulation entry, original station assignments, service durations, charger-outage realizations, waiting tolerances, and fixed recommendation-acceptance draws.

For the 20% demand-growth condition, an additional number of transactions equal to 20% of the original daily vehicle count is sampled with replacement from that day's records. The copied charging-initiation times are independently perturbed by an integer offset between -7 and +7 minutes and clipped to the boundaries of the same day. Under the 10% charger-outage-probability condition, each charger is independently retained with probability 0.90; effective station capacity is constrained to remain at least one charger. Thus, 10% denotes the Bernoulli probability applied to each charger, not a requirement that every scenario lose exactly 10% of network capacity. These stochastic draws are generated once per immutable scenario and reused by every policy.

All station-assignment policies use the same scenario without modifying its stochastic inputs. Policy names and recommendation-acceptance parameters are not used when generating charging-initiation event streams, service durations, or outages. This common-random-number design reduces simulation variance and ensures that paired differences between policies are not confounded by different random inputs.

Scenario hashes, vehicle counts, capacity constraints, and inbound-reservation counters are automatically checked after each experiment. Experiments are considered valid only when all policies used identical scenario inputs, the number of vehicles in service never exceeds effective station capacity, and no inbound counter becomes negative.

3.6. Assignment Policies and Routing Controls

Four station-assignment policies are evaluated. Let $n_{j,t}^{\mathrm{srv}}$, $n_{j,t}^{\mathrm{que}}$, and $n_{j,t}^{\mathrm{in}}$ denote the numbers of vehicles in service, waiting, and committed inbound at station j, respectively. Effective load was (equation 10),

$$L_{j,t} = \frac{n_{j,t}^{\mathrm{srv}} + n_{j,t}^{\mathrm{que}} + n_{j,t}^{\mathrm{in}}}{c_j} \tag{10}$$

Let $\widetilde{D}_j$ denote the station-specific median service duration. The median is used in the routing-score approximation because service durations are positively skewed: the station-specific mean exceeds the median at all 13 stations, indicating that a small number of long sessions can disproportionately inflate a mean-based queue proxy. The median therefore provides a robust typical-duration multiplier for comparing candidate stations. It is used only in the routing score; actual resource occupancy and service completion in the discrete-event simulation use each vehicle's transaction-level duration D_j. Current backlog delay is approximated by (equation 11),

$$W_{j,t}^{\mathrm{backlog}} = \max(0, n_{j,t}^{\mathrm{srv}} + n_{j,t}^{\mathrm{que}} - c_j + 1)\frac{D_j}{c_j} \tag{11}$$

The simulation uses the 60-minute cumulative hurdle expectation $\hat{Y}_{j,t}^{(4,\mathrm{hurdle})}$ as the expected-demand input. The 60-minute horizon is specified in the analysis configuration to summarize charging-initiation pressure over the next four decision intervals, a window more closely aligned with queue formation and charging-service evolution than a single 15-minute slot. It is also the most predictable evaluated horizon, with higher R^2, lower WAPE, and higher occurrence AUC than the shorter horizons, although those test results are not used to select among model families. The hurdle signal is chosen because it explicitly combines the estimated probability of a positive charging-initiation count with the expected magnitude conditional on positive demand. It is not selected as the numerically best point forecast: at the 60-minute horizon, the Tweedie model achieves an MAE of 1.490 and an RMSE of 2.263, compared with 1.507 and 2.269 for the hurdle model. Thus, the horizon choice and the model-family choice serve different purposes: the former defines the operational accumulation window, whereas the latter provides an interpretable expected-count decomposition for the routing score. The routing results characterize this implemented hurdle-based signal and do not imply that the hurdle formulation is generally superior to Tweedie forecasting. Forecast pressure and the resulting predicted wait are (equation 12),

$$\begin{aligned} W_{j,t}^{\mathrm{forecast}} &= \lambda_f(n_{j,t}^{\mathrm{in}} + \max(0, \hat{Y}_{j,t}^{(4,\mathrm{hurdle})}))\frac{D_j}{c_j} \\ \hat{W}_{j,t} &= W_{j,t}^{\mathrm{backlog}} + W_{j,t}^{\mathrm{forecast}} \end{aligned} \tag{12}$$

with $\lambda_f = 0.50$. The upper-tail risk margin and load-balance term are (equation 13),

$$R_{j,t} = \max(0, \hat{Q}_{0.90,j,t} - \hat{Q}_{0.50,j,t}) \frac{D_j}{c_j} B_{j,t} = L_{j,t} D_j \quad (13)$$

Under the no-routing policy, each vehicle remains assigned to its historically observed station. For the remaining policies, all 13 stations are eligible candidates. Let $C_{i,j}$ equal zero for the original station and eight minutes for an alternative station. A hard barrier $H_{j,t} = 10^6$ is added for an alternative station when its inbound count reaches c_j. The implemented scores are (equation 14),

$$\begin{aligned} S_{i,j,t}^{\text{load}} &= L_{j,t} \widetilde{D}_j + C_{i,j} + H_{j,t}, \\ S_{i,j,t}^{\text{wait}} &= \widehat{W}_{j,t} + C_{i,j} + H_{j,t}, \\ S_{i,j,t}^{\text{risk}} &= \widehat{W}_{j,t} + \lambda_r R_{j,t} + \lambda_b B_{j,t} + C_{i,j} + H_{j,t}. \end{aligned} \quad (14)$$

Here, $\lambda_r = 0.60$ and $\lambda_b = 0.25$. equation 14 defines the risk-aware score used throughout the paper. Relative to current-load routing, this policy changes several elements jointly: it uses the forecast-augmented waiting estimate $\widehat{W}_{j,t}$, adds the upper-tail margin $R_{j,t}$, and adds the load-balance term $B_{j,t}$. The comparison between risk-aware and current-load routing therefore evaluates this bundled score and does not identify the separate causal contribution of forecasted expected demand, predictive uncertainty, or load balancing.

To prevent recommendation-induced herding, a station-specific inbound counter represents vehicles that have been redirected to a station but have not yet begun charging. When a vehicle is assigned to an alternative station, the corresponding counter is immediately increased and is decreased only after the vehicle completes the modeled switching interval and enters the destination system.

Inbound vehicles are included in the effective-load and forecast-pressure calculations used by subsequent assignment decisions. This prevents multiple vehicles with simulation-entry events in the same decision interval from independently selecting the same apparently uncongested station.

The inbound count of an alternative station is capped at its effective charger capacity. A vehicle is not redirected when the effective occupancy of its original station, vehicles in service, waiting, and inbound, is below capacity. This capacity condition is a common gate for all three dynamic policies. Consequently, the expected-demand and upper-quantile terms can influence the ranking of candidate destinations only after the original station has reached effective capacity; they cannot initiate a pre-emptive reassignment while immediate original-station capacity remains available. Otherwise, the lowest-score station is selected only when its score improves upon the original-station score by at least two minutes. Because the station score already includes the eight-minute switching cost for an alternative, this additional threshold suppresses recommendations whose estimated advantage is only marginal. The two-minute value is used as a fixed decision safeguard rather than an empirically optimized threshold.

Because reliable station-to-station travel times are not available in the dataset, all 13 stations are evaluated using a fixed eight-minute travel and inconvenience time for an alternative assignment. This value serves as a stylized combined proxy for additional travel and switching inconvenience and prevents rerouting from being treated as costless. It is specified before the policy comparisons as a moderate non-zero scenario penalty: an alternative station has to overcome a meaningful switching burden, while rerouting can still occur when the original station is sufficiently congested. The value is not selected from the validation or test results, tuned to maximize the reported policy reductions, or estimated from observed Jiaxing origin-destination trips. Consequently, the reported policy effects are conditional on this assumption. A redirected vehicle is held in the inbound state during this eight-minute travel interval before joining the destination queue. No travel time is applied when the vehicle remains at its original station. For each vehicle, total delay is defined as,

$$T_i^{\text{delay}} = W_i^{\text{queue}} + T_i^{\text{switch}} \quad (15)$$

where W_i^{queue} is the simulated queueing time and T_i^{switch} is the modeled eight-minute switching time for a redirected vehicle. Equation 15 therefore accounts for both destinations queueing and the modeled switching time. Total delay, rather than queueing time alone, is used as the primary policy outcome because a policy can otherwise appear beneficial by reducing queues while imposing excessive travel costs.

Finally, the experiments consider two recommendation-acceptance levels are a ϵ {0.75, 1.00}.

Each vehicle is assigned a fixed uniform recommendation-acceptance draw during immutable scenario generation. A vehicle follows a routing recommendation when its stored draw is less than or equal to the evaluated recommendation-acceptance rate. The 0.75 and 1.00 levels represent scenarios with 75% and 100% recommendation acceptance rather than behaviorally estimated compliance rates. Because the draw is fixed across policies, differences between recommendation-acceptance scenarios do not alter the charging-initiation event stream or other stochastic inputs.

3.6. Experimental Scenarios and Statistical Analysis

Day selection: Twenty-one test days are selected deterministically from the test period using stratification by day of week and daily charging-initiation volume. Within each weekday group, the days nearest the 0.15, 0.50, and 0.85 quantiles of daily charging-initiation volume are selected. This produces three low-, medium-, and high-volume representatives for each weekday and avoided over-representing a particular weekday or demand level. Table 1 reports the exact dates.

Demand and capacity stressors: The assignment policies are evaluated under combinations of two demand levels

and two infrastructure-availability levels.

The demand multiplier is $m_d \in \{1.0,1.2\}$, representing observed demand and a 20% demand-increase scenario.

The charger-outage probability is $m_o \in \{0.0,0.1\}$, representing full capacity and a scenario in which each charger is independently unavailable with probability 0.10. The 20% demand increase is used to raise the number of charging-initiation events while preserving the empirical day, station, and service-duration structure through resampling. The 10% charger-outage probability represents partial charger-level unavailability rather than complete station closure. Across the 42 unique outage realizations, 21 simulation days at each of the two demand levels, the mean realized network-capacity reduction is 9.91%, corresponding to a mean effective capacity of 82.88 of 92 chargers. The mean reduction is 9.73% without demand growth and 10.09% with the 20% demand increase; individual realizations range from 3.26% to 18.48%. Neither stress value is a forecast of future demand or an estimate of historical outage frequency. Independent charger failures provide a reproducible capacity stressor, while the requirement that each station retain at least one charger avoids a degenerate all-closed station.

The forecasting, simulation, and routing settings are consolidated in Table 2. The table distinguishes configuration choices fixed before final test evaluation, validation-controlled early stopping, structural model definitions, scenario-design assumptions, and numerical implementation safeguards. The validation subset is used only for early stopping; it is not used to optimize the fixed tree settings, Tweedie variance power, routing weights, or behavioral scenario values. The magnitude of the reported forecasting and simulation results remains conditional on these choices.

Table 1. Simulation dates selected from the chronological test period by weekday and daily charging-initiation volume

Weekday	Low-volume date	Median-volume date	High-volume date
Monday	2021-11-22	2021-11-29	2021-12-13
Tuesday	2021-08-24	2021-08-31	2021-11-23
Wednesday	2021-08-25	2021-09-15	2021-10-06
Thursday	2021-09-23	2021-11-04	2021-11-18
Friday	2021-10-22	2021-12-03	2021-12-17
Saturday	2021-12-04	2021-12-18	2021-12-25
Sunday	2021-10-03	2021-10-24	2021-12-19

Table 2. Key forecasting and routing parameters used in the analysis

Parameter	Value	Selection basis
Tweedie variance power	1.35	Fixed analysis setting
Forecast-pressure weight λ_f	0.50	Fixed policy setting
Risk weight λ_r	0.60	Fixed policy setting
Load-balance weight λ_b	0.25	Fixed policy setting
Minimum routing gain	2 min	Decision safeguard
Switching time	8 min	Scenario assumption
Waiting tolerance	30 min	Scenario assumption

Together with the two recommendation-acceptance rates and four assignment policies, the full experimental design comprises $21 \times 2 \times 2 \times 2 \times 4 = 672$ simulation runs.

Simulation performance measures: The primary simulation outcome is mean total delay. Mean queueing time, queue-abandonment rate, and the proportion of redirected vehicles are reported as complementary outcomes. Simulation-integrity outcomes include vehicle-count and scenario-hash equality across policies, capacity violations, negative inbound counts, non-negative forecasts, and quantile ordering.

The percentage reduction in waiting or total delay relative to the no-routing policy is calculated only when the baseline value is greater than a small numerical tolerance (equation 16):

$$\text{Reduction} = 100 \cdot \frac{T^{\text{baseline}} - T^{\text{policy}}}{T^{\text{baseline}}} \tag{16}$$

When the baseline delay is zero, the percentage reduction is treated as undefined and only the absolute difference is reported.

Statistical analysis: Policy comparisons are performed at the simulation-day level using paired observations. Because all policies are evaluated with identical scenario inputs, the daily policy outcomes are directly paired with the corresponding no-routing outcomes.

For the inferential comparisons reported below, recommendation acceptance is fixed at 1.00. Nonparametric paired bootstrap resampling with 20,000 resamples is used to estimate the mean policy difference and its 95% confidence interval. Resampling is performed over the 21 simulation days rather than over

individual vehicles to preserve within-day dependence.

A policy is interpreted as providing consistent improvement over no routing when the 95% bootstrap confidence interval for the difference in total delay is entirely below zero. The same paired day-level bootstrap procedure is also applied to three direct contrasts among the dynamic policies: risk-aware minus current-load, predicted-wait minus current-load, and risk-aware minus predicted-wait. These comparisons are interpreted using both the confidence interval and the absolute difference in minutes and seconds per vehicle, because a numerical ranking alone does not establish a practically meaningful advantage.

Reproducibility: The forecasting models are implemented with LightGBM (Ke et al., 2017), and the discrete-event simulation is implemented with SimPy. The model configuration is fixed before final test evaluation; the validation subset is used for early stopping only, and all prespecified forecasting families are reported on the untouched test subset. The analysis uses a global seed of 2026 for model fitting, bootstrap resampling, recommendation-acceptance draws, demand-growth resampling, temporal jitter, and charger-outage draws. Scenario-specific seeds are generated deterministically from the base seed together with the simulation date, demand multiplier, outage rate, and the label “scenario” using a SHA-1-based stable-seed function. The scenario identifier is likewise derived from the pipeline version, date, stress settings, and base seed. This procedure ensured that every routing policy receives the same immutable charging-initiation events, service durations, outage realization, and recommendation-acceptance draws.

The archived numerical outputs identify pandas 2.2.2, PyArrow 18.1.0, and Matplotlib 3.10.8. The archived Colab notebook specifies a Python 3 kernel and installation constraints of LightGBM $\geq$ 4.0 and SimPy $\geq$ 4.1; however, the exact resolved Python, LightGBM, and SimPy versions are not preserved in the available runtime artifacts. Numerical values underlying the reported figures are retained together with the experiment-level outputs.

5. Results

5.1. Source-Data Characteristics and Station-Level Heterogeneity

The demand-analysis subset contains 441077 charging-initiation events distributed across 13 stations. Of these published records, 422053 have valid charging-service durations and are retained for service-duration estimation and discrete-event simulation. The source records contained 92 distinct charging-post identifiers. Counting distinct identifiers within each station produce the four-to-eight-post capacity proxies used in the simulation.

The observed charging demand is unevenly distributed across the station network. The three highest-demand stations, Bus Station, Park B, and shopping mall, accounted for 39.6% of all charging initiations, while the five highest-demand stations accounted for 59.3%. Bus Station records the largest number of demand events, with 60809 sessions, whereas Expressway Service District B records the smallest number, with 12401 sessions.

The station-specific mean service duration ranges from 27.3 minutes at Expressway Service District C to 75.4 minutes at Wholesale Market, while the station-specific median ranges from 23.0 minutes at Expressway Service District A to 48.1 minutes at Park B. At every station, the mean exceeds the median, consistent with positively skewed service-duration distributions. The capacity of individual stations ranged from four to eight charging posts. Across the complete source dataset, approximately 7.25 GWh of energy is recorded. Table 3 reports both descriptive measures so that the observed means can be distinguished from the medians used in the routing-score approximation.

Note that mean and median durations are descriptive summaries of valid transaction-level service times. The station median, not the mean, is used only as the typical-duration multiplier in the routing-score approximation; simulated service completion uses each vehicle's observed duration. Capacity is proxied by the number of distinct charging-post identifiers observed within each station. Stations are ordered by charging-initiation volume.

The observed values indicate substantial heterogeneity in both demand volume and service duration. The consistent mean-median gap also shows why the mean is retained for descriptive reporting but not used as the typical-duration multiplier in the routing score. Consequently, a network-wide average demand or service-duration assumption would not adequately represent the operational differences among stations.

5.2. Cumulative Charging-Demand Forecasting

Table 4 presents the overall forecasting results for cumulative 15-, 30-, and 60-minute demand. The machine-learning count models consistently outperform the daily seasonal baseline in terms of MAE and RMSE. At the 15-minute horizon, the Tweedie model achieves an MAE of 0.633 and an RMSE of 0.974, corresponding to reductions of 17.6% and 26.8%, respectively, relative to the daily seasonal model. At the 30-minute horizon, the corresponding reductions are 19.8% for MAE and 24.9% for RMSE. At the 60-minute horizon, the Tweedie model reduces MAE by 19.0% and RMSE by 21.7%.

Table 3. Station-level charging demand, service-duration distribution, and capacity proxy

Station	Charging initiations	Network share (%)	Valid service records	Mean duration (min)	Median duration (min)	Capacity proxy
Bus Station	60809	13.8	58480	69.0	43.0	8
Park B	57294	13.0	56294	65.1	48.1	8
Shopping Mall	56605	12.8	55275	64.2	38.5	8
Expressway Service District C	50187	11.4	46260	27.3	23.2	8
Government Agency	36551	8.3	35775	74.2	45.4	8
Wholesale Market	34491	7.8	30765	75.4	31.9	8
Technology Park	28867	6.5	28620	56.3	38.5	4
Park A	27845	6.3	27406	41.8	23.3	8
Tourist Attraction	23659	5.4	22723	42.3	23.4	8
Financial Industrial Park	18923	4.3	18232	53.2	42.1	8
Industrial Park	18370	4.2	17902	69.0	36.6	8
Expressway Service District A	15075	3.4	12976	27.8	23.0	4
Expressway Service District B	12401	2.8	11345	30.5	25.9	4

Table 4. Point-forecast performance for cumulative charging-initiation demand in the chronological test period.

Horizon (min)	Model	MAE	RMSE	WAPE	R^2	Bias
15	Daily seasonal	0.769	1.330	1.112	-0.390	-0.001
15	Weekly seasonal	0.775	1.341	1.120	-0.413	-0.008
15	Poisson	0.635	0.974	0.918	0.255	-0.053
15	Tweedie	0.633	0.974	0.916	0.255	-0.057
15	Hurdle	0.636	0.975	0.920	0.253	-0.053
15	Quantile Q50	0.577	1.078	0.834	0.087	-0.348
30	Daily seasonal	1.203	1.956	0.870	-0.117	-0.003
30	Weekly seasonal	1.224	1.987	0.885	-0.152	-0.016
30	Poisson	0.966	1.469	0.698	0.370	-0.110
30	Tweedie	0.964	1.469	0.697	0.370	-0.112
30	Hurdle	0.970	1.471	0.701	0.368	-0.103
30	Quantile Q50	0.941	1.562	0.680	0.288	-0.399
60	Daily seasonal	1.840	2.889	0.665	0.171	-0.005
60	Weekly seasonal	1.899	2.980	0.686	0.118	-0.032
60	Poisson	1.494	2.262	0.540	0.492	-0.226
60	Tweedie	1.490	2.263	0.539	0.491	-0.240
60	Hurdle	1.507	2.269	0.545	0.489	-0.205
60	Quantile Q50	1.499	2.372	0.542	0.441	-0.552

The daily seasonal model outperforms the weekly seasonal model at every horizon and is therefore used as the primary reference for relative-improvement calculations. At the 15-minute horizon, the daily and weekly seasonal WAPE values are 1.112 and 1.120. These values mean that the summed absolute errors are 111.2% and 112.0%, respectively, of the total observed charging-initiation count; they do not mean that more than 100% of individual predictions are incorrect. Values above one are possible because WAPE is unbounded above. In this sparse station-interval setting, many targets are zero or small, so timing a charging initiation

one interval early or late can contribute absolute error in two low-count intervals while the denominator remains the total observed count. The relative forecasting performance improves as the cumulative horizon increases. The R^2 value of the Tweedie model increases from 0.255 at 15 minutes to 0.491 at 60 minutes, while its WAPE decreases from 0.916 to 0.539. Cumulative one-hour demand is therefore more predictable than demand within an individual 15-minute interval.

The Poisson and Tweedie models produce nearly identical performance. The Tweedie model achieves the lowest MAE and WAPE at all three horizons, whereas the Poisson model produces a marginally lower RMSE at 60 minutes. The hurdle model performs similarly to the two count-regression models but do not improve their point-forecasting accuracy. At 60 minutes, its MAE and RMSE exceed the Tweedie values by 0.017 and 0.006, respectively. Nevertheless, its binary occurrence component shows AUC values of 0.808, 0.848, and 0.889 at the 15-, 30-, and 60-minute horizons, respectively. Its use in routing is therefore motivated by the interpretability of its occurrence--magnitude decomposition, not by superior point-forecast performance. The Q50 model produces the lowest MAE at 15 and 30 minutes. However, this result is accompanied by substantial systematic underprediction. Its predicted demand totals are 50.4%, 28.8%, and 20.0% below the observed totals at the three respective horizons. By comparison, the Tweedie model underestimates aggregate demand by approximately 8.1-8.7%. Therefore, the Q50 model represents typical or median demand more accurately but is less suitable as an estimator of expected network load.

Station-level results show that predictability differs across locations. At the 60-minute horizon, the Tweedie model achieves its highest station-level R^2 at Technology Park ($R^2 = 0.614$), whereas its lowest value is observed at Expressway Service District A ($R^2 = 0.167$). Thus, the improvement obtained from the cumulative target definition does not eliminate spatial differences in forecasting difficulty.

5.3. Probabilistic Forecast Calibration

The empirical performance of the Q10-Q90 prediction intervals is reported in Table 6. The nominal interval coverage is 80%. Empirical coverage exceeds the nominal level at all horizons, reaching 89.8% at 15 minutes, 89.5% at 30 minutes, and 83.2% at 60 minutes. The prediction intervals are therefore conservative, particularly at the shorter horizons. Interval width increases from 1.70 charging events at 15 minutes to 4.53 events at 60 minutes, reflecting the larger magnitude and uncertainty of cumulative demand over longer time windows.

The Q10 model collapses to zero for all 15- and 30-minute observations and for 51.8% of the 60-minute observations. This behavior is consistent with the high proportion of zero-demand station-interval combinations. Quantile crossing is not observed at the 15- or 30-minute horizons and occurs in only 0.0016% of observations at 60 minutes before monotonic reordering.

Figure 1 summarizes the point and probabilistic forecasting results. Figure 1 (a) shows that the learned models consistently reduce MAE relative to the daily seasonal baseline, although the low Q50 error is accompanied by systematic underprediction. Figure 1 (b) shows that empirical interval coverage remains above the nominal 80% level at every horizon, indicating conservative uncertainty estimates, particularly at 15 and 30 minutes.

Table 5. Overall Q10-Q90 prediction-interval performance

Horizon (min)	Coverage (%)	Width	Coverage deviation (pp.)	Interval score	Q10 zero rate (%)
15	89.85	1.700	9.85	2.702	100.00
30	89.45	2.939	9.45	4.288	100.00
60	83.15	4.528	3.15	6.892	51.77

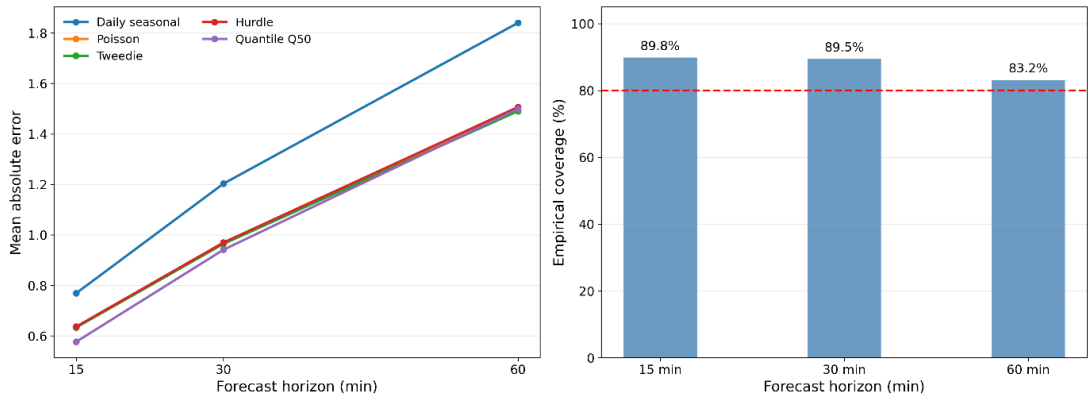


Figure 1. Forecast accuracy and probabilistic calibration for cumulative 15-, 30-, and 60-minute demand. (a) compares mean absolute error across the principal point-forecast models. (b) compares empirical Q10-Q90 coverage with the nominal 80% level.

5.4. Simulation Integrity and Experimental Isolation

The corrected discrete-event simulation completes all 672 planned experiments. These experiments represent 21 simulation days, two demand levels, two charger-outage levels, two recommendation-acceptance levels, and four station-assignment policies. A total of 260 automated integrity tests are performed, and all tests pass. The common-random-number design is confirmed by 84 successful vehicle-count equality checks and 84 successful scenario-hash equality checks. The no-routing policy is invariant to the recommendation-acceptance parameter in all 84 relevant comparisons. No capacity violations or negative inbound-reservation counters are observed. Forecasts remain non-negative, and quantile-ordering checks pass for all forecasting horizons. These results confirm that the policy comparisons are based on identical vehicle streams, service durations, and outage realizations. Consequently, differences among policies are attributable to assignment decisions rather than to variation in the simulated input scenarios.

5.5. Performance under Normal Operating Conditions

Under observed demand and full charger availability, the baseline system does not develop queues. All four assignment policies produce zero mean waiting time, zero total delay, zero abandonment, and zero rerouting across the 21 simulation days. The dynamic policies therefore remain inactive when the original station has sufficient capacity. For example, on the representative normal-condition day of 24 August 2021, all 721 simulated vehicles are served at their original stations under every policy, and no rerouting decision is made.

This behavior shows that the congestion trigger and minimum-gain conditions successfully prevent unnecessary station switching. Because the no-routing delay is already zero, percentage delay reductions are undefined for the normal scenario and are not reported.

5.6. Performance under Demand and Capacity Stress

Table 7 summarizes policy performance in the three stress scenarios at full routing acceptance. Total delay includes both simulated queueing time and the eight-minute switching penalty assigned to rerouted vehicles.

When 10% of charger capacity is unavailable, the no-routing policy produces a mean total delay of 0.814 minutes per vehicle and an abandonment rate of 0.527%. The current-load, predicted-wait, and risk-aware policies reduce mean total delay to 0.183, 0.180, and 0.174 minutes, respectively. These values correspond to reductions of 77.6%, 77.9%, and 78.7%. Only approximately 2.1-2.3% of vehicles required rerouting.

Under the 20% demand-increase scenario, mean total delay increases to 1.845 minutes without routing. The current-load, predicted-wait, and risk-aware policies reduce this value to 0.442, 0.461, and 0.427 minutes, representing improvements of 76.0%, 75.0%, and 76.9%, respectively. The dynamic policies reroute between 4.8% and 5.5% of vehicles.

The largest baseline delay occurs under combined demand and capacity stress. Without routing, mean total delay reaches 3.505 minutes, and 2.962% of vehicles abandon the queue. The three dynamic policies reduce total delay to between 0.770 and 0.850 minutes. Current-load and risk-aware assignment achieves nearly identical reductions of approximately 78.0%, while predicted-wait assignment achieves a reduction of 75.8%. Figure 2 (a) visualizes the corresponding reduction in mean total delay. The reductions are achieved using selective rather than system-wide rerouting. Even under combined stress, fewer than 10% of vehicles are assigned to an alternative station. Thus, the performance improvement is not produced by transferring the majority of demand away from historically observed stations.

Table 6. Policy performance under stress scenarios with full routing acceptance. Abandonment and rerouting rates are reported as percentages of all vehicles

Scenario	Policy	Mean queue wait (min)	Mean total delay (min)	Abandonment (%)	Rerouted (%)
10% capacity loss	No routing	0.814	0.814	0.527	0.00
	Current load	0.000	0.183	0.000	2.28
	Predicted wait	0.010	0.180	0.005	2.12
	Risk-aware	0.000	0.174	0.000	2.17
20% demand increase	No routing	1.845	1.845	1.120	0.00
	Current load	0.002	0.442	0.000	5.50
	Predicted wait	0.081	0.461	0.056	4.76
	Risk-aware	0.013	0.427	0.015	5.17
Combined stress	No routing	3.505	3.505	2.962	0.00
	Current load	0.017	0.771	0.000	9.43
	Predicted wait	0.175	0.850	0.106	8.44
	Risk-aware	0.069	0.770	0.036	8.76

5.7. Queue Abandonment

Dynamic assignment also reduces the number of vehicles exceeding the 30-minute waiting tolerance. With full routing acceptance, the current-load policy eliminates abandonment in all three stress scenarios. The risk-aware policy eliminates abandonment under isolated capacity loss and reduces it by 98.7% and 98.8% under the demand-increase and combined-stress scenarios, respectively. The predicted-wait policy also produces large reductions but retains small residual abandonment rates. Under combined stress, its abandonment rate is 0.106%, compared with 2.962% for no routing. Figure 2 (b) shows the corresponding reductions in queue abandonment. At 75% routing acceptance, the dynamic policies remain effective, although their reductions are smaller. Depending on the policy and stress condition, total delay is reduced by approximately 70.4-73.2% relative to no routing. This indicates that the operational benefits do not depend on universal compliance with routing recommendations.

5.8. Paired Statistical Comparison

Table 7 reports paired day-level bootstrap comparisons of mean total delay at full routing acceptance. Negative differences indicate improvement relative to no routing. All confidence intervals are entirely below zero, indicating that each dynamic assignment policy reduces mean total delay relative to no routing in all three stress scenarios. The current-load and risk-aware policies improve total delay on all 21 simulated days in every stress condition. The predicted-wait policy improves 20 of 21 days under isolated capacity loss and all 21 days under the remaining two stress conditions.

Differences among the dynamic policies are substantially smaller than their differences from no routing. Risk-aware assignment achieves a lower mean total delay than current-load assignment under isolated capacity loss and the demand-increase scenario. However, under combined stress their aggregate performance is effectively identical: 0.770 minutes for risk-aware assignment and 0.771 minutes for current-load assignment.

Predicted-wait assignment is slightly less effective than the other dynamic policies as stress increases. Under combined stress, it produces approximately 0.079 minutes more total delay than current-load assignment. Nevertheless, its performance remains substantially better than the no-routing baseline.

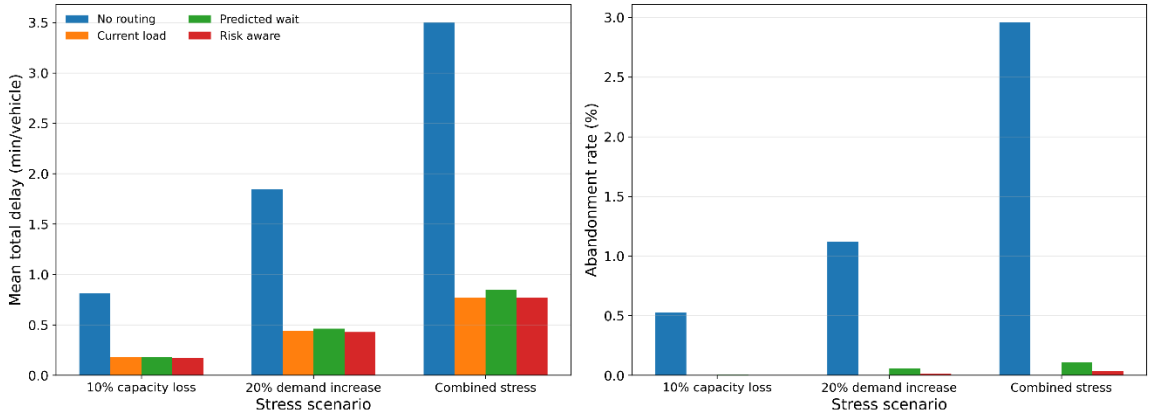


Figure 2. Simulation outcomes under the three stress scenarios with full recommendation acceptance. (a) reports mean total delay, including queueing time and the station-switching penalty. (b) reports queue-abandonment rates. All dynamic routing policies substantially reduced delay and nearly eliminated abandonment relative to no routing.

Table 7. Paired bootstrap comparison of mean total delay against the no-routing policy. Confidence intervals were obtained by resampling the 21 paired simulation days

Scenario	Policy	Policy mean (min)	Reduction (%)	Mean difference (min)	95% CI
10% capacity loss	Current load	0.183	77.6	-0.632	[-1.003, -0.351]
	Predicted wait	0.180	77.9	-0.635	[-1.011, -0.350]
	Risk-aware	0.174	78.7	-0.641	[-1.017, -0.356]
20% demand increase	Current load	0.442	76.0	-1.403	[-1.709, -1.111]
	Predicted wait	0.461	75.0	-1.384	[-1.688, -1.090]
	Risk-aware	0.427	76.9	-1.418	[-1.722, -1.128]
Combined stress	Current load	0.771	78.0	-2.734	[-3.228, -2.256]
	Predicted wait	0.850	75.8	-2.655	[-3.141, -2.188]
	Risk-aware	0.770	78.0	-2.735	[-3.231, -2.259]

6. Discussion and Limitations

The forecasting results show that cumulative charging-initiation counts can be predicted more accurately than the daily and weekly seasonal baselines, although the improvement is moderate rather than transformative. Poisson and Tweedie gradient boosting produce nearly identical point forecasts, and their relative performance improve as the cumulative horizon increased. At 60 minutes, the models reach an R^2 of approximately 0.49 and a WAPE of approximately 0.54, whereas the 15-minute target remains more difficult because station-level counts are sparse. This horizon-dependent pattern is consistent with prior charging-load studies showing that temporal aggregation and target construction materially affect forecast accuracy (Unterluggauer et al., 2021; Hussain et al., 2025; Mansour et al., 2026). The hurdle formulation does not improve point accuracy, while the Q50 model achieves low MAE at the cost of substantial aggregate underprediction. The Q10-Q90 intervals are conservative, and the lower quantile frequently collapses to zero. The probabilistic outputs should therefore be interpreted as broad indicators of upper-tail demand rather than as a finely calibrated predictive distribution.

The simulation results support the operational value of dynamic assignment under the evaluated stress conditions. The policies remain inactive under observed demand and full capacity, indicating that the immediate-capacity rule and minimum-gain threshold prevent unnecessary switching. Under demand growth, capacity loss, and combined stress, the dynamic policies reduce mean total delay by approximately 75-79% at full recommendation acceptance and nearly eliminate abandonment while redirecting fewer than 10% of vehicles. Because total delay includes the eight-minute switching time, these gains are not obtained by shifting queueing costs outside the objective. The CRN design and integrity checks further strengthen this comparison by ensuring that every policy receives identical arrivals, service durations, outages, and acceptance draws.

The main comparative finding is that forecast-aware routing offers little practical advantage over the simpler current-load policy. Under combined stress, mean total delay is 0.770 minutes for risk-aware routing and 0.771 minutes for current-load routing. Thus, nearly all of the observed benefit is already captured by the shared congestion trigger, switching-cost rule, and representation of committed inbound vehicles. Forecast pressure may become more useful when near-future arrivals differ sufficiently across stations to alter decisions before congestion is visible in current occupancy, but the present scenarios do not produce a large incremental gain. Deployment should therefore balance any forecast-based improvement against the additional requirements for model training, calibration, monitoring, and maintenance.

The results suggest three practical principles. Routing should be selective, redirected vehicles should be included in the destination state before they arrive, and policy evaluation should include travel-related delay and abandonment rather than queueing time alone. The benefit also persists at 75% acceptance, suggesting that universal compliance is not required. However, the acceptance rate is an abstract behavioral parameter; actual compliance will depend on expected benefit, trust, route constraints, charger compatibility, price, and user preferences (Hoen et al., 2023; Hanni et al., 2024; Brückmann and Bernauer, 2023). Moreover, the independent contribution of inbound reservations is not isolated through ablation, so the present experiments demonstrate the performance of the complete routing architecture rather than the causal effect of that component alone.

Several limitations constrain the interpretation of the findings. The source data do not contain a reliable physical arrival timestamp or directly observed queue-waiting time. Charging start is therefore used as the demand event and simulated arrival time, which means that the study models counterfactual congestion among observed charging initiations rather than reconstructing historical queues or including vehicles that left without charging. The fixed 30-minute patience threshold and eight-minute switching time also simplify heterogeneous behavior and geography. Actual waiting tolerance and detour cost depend on the driver, trip purpose, origin-destination pair, traffic, charger compatibility, battery state of charge, and station amenities.

The experimental scope is also limited. The analysis uses one charging network from Jiaxing during 2019-2021, including the COVID-19 disruption, and the simulation covers 21 weekday-balanced test days rather than the full test period. All 13 stations are treated as feasible alternatives because route-specific coordinates and travel times are unavailable. The demand-growth and outage scenarios are stylized, and the routing weights, inbound cap, minimum gain, and switching time are fixed without a dedicated sensitivity study. Forecasting is evaluated using one chronological split, while the bootstrap analysis establishes improvement over no routing but does not test pairwise superiority among the dynamic policies. Consequently, the results establish the potential of coordinated assignment under the stated assumptions, not the realized performance of a deployable navigation service or general superiority across charging networks.

Future work should evaluate the framework using observed arrival and waiting-time data, route-specific travel costs, geographically restricted candidate stations, heterogeneous patience and acceptance models, rolling-origin forecast evaluation, component-level ablations, parameter sensitivity, and additional charging networks. These extensions are particularly important for determining when forecast uncertainty provides enough operational value to justify its complexity beyond current-load routing.

7. Conclusion

This study investigated whether short-term cumulative charging-demand forecasts can support congestion-aware assignment in a public electric-vehicle charging network. Station-level charging initiations were forecasted over 15-, 30-, and 60-minute windows using seasonal, count-regression, hurdle, and quantile models. The forecasts were integrated with a discrete-event simulation that incorporated station-specific service durations, partial compliance, switching cost, queue abandonment, inbound reservations, and common-random-number scenario control.

Poisson and Tweedie gradient boosting produced moderate but consistent improvements over daily seasonal forecasting, with the strongest relative performance at the 60-minute cumulative horizon. In the simulation, the dynamic policies introduced no unnecessary rerouting under normal conditions and substantially reduced total delay and abandonment under demand and capacity stress. These improvements were achieved by redirecting only a small share of vehicles. At full recommendation acceptance, risk-aware routing produced the lowest aggregate delay, but its advantage over the simpler current-load policy was operationally small.

The current-load policy achieved nearly the same benefit as the risk-aware policy, so the additional forecast-risk terms were not operationally decisive in these experiments. The independent effect of inbound reservations remains to be measured through component-level ablation. Further evaluation should use observed vehicle-arrival and waiting-time data, route-specific travel costs, heterogeneous behavioral models, broader sensitivity analyses, and networks from additional cities.

Author Contributions

The percentages of the author' contributions are presented below. The author reviewed and approved the final version of the manuscript.

	O.A.
C	100
D	100
DCP	100
DAI	100
L	100
W	100
CR	100
SR	100
PM	100

C= concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management.

Conflict of Interest

The author declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because there was no study on animals or humans.

Acknowledgements

This study did not generate or collect a new dataset. It reuses the publicly available Jiaxing electric-vehicle charging transaction dataset introduced by (Zhang et al., 2025).

References

Aghsaee, R., Hecht, C., Schwinger, F., Figgener, J., Jarke, M., & Sauer, D. U. (2023). Data-Driven, Short-Term Prediction of Charging Station Occupation. *Electricity, 4*, 142–160. https://doi.org/10.3390/electricity4020009

Alaraj, M., Radi, M., Alsisi, E., Majdalawieh, M., & Darwish, M. (2025). Machine Learning-Based Electric Vehicle Charging Demand Forecasting: A Systematized Literature Review. *Energies, 18*, 4779. https://doi.org/10.3390/en18174779

Ali, M. W., Mustafa, M. A., Shuvo, M. A., & Sick, B. (2026). Location-Based Probabilistic Load Forecasting of Electric Vehicle Charging Sites: Deep Transfer Learning with Multi-quantile Temporal Convolutional Network. In *Architecture of Computing Systems* (Vol. 15839). https://doi.org/10.1007/978-3-032-03281-2_14

Basmadjian, R., Kirpes, B., Mrkos, J., & Cuchý, M. (2020). A Reference Architecture for Interoperable Reservation Systems in Electric Vehicle Charging. *Smart Cities, 3*, 1405–1427. https://doi.org/10.3390/smartcities3040067

Brückmann, G., & Bernauer, T. (2023). An experimental analysis of consumer preferences towards public charging infrastructure. *Transportation Research Part D: Transport and Environment, 116*, 103626. https://doi.org/10.1016/j.trd.2023.103626

Dang, S., Peng, L., Zhao, J., Li, J., & Kong, Z. (2022). A Quantile Regression Random Forest-Based Short-Term Load Probabilistic Forecasting Method. *Energies, 15*, 663. https://doi.org/10.3390/en15020663

Dudkina, E., Scarpelli, C., Apicella, V., Ceraolo, M., & Crisostomi, E. (2025). Optimised Centralised Charging of Electric Vehicles Along Motorways. *Sustainability, 17*, 5668. https://doi.org/10.3390/su17125668

Elahe, M. F., Kabir, M. A., Mahmud, S. M., & Azim, R. (2022, December). Factors Impacting Short-Term Load Forecasting of Charging Station to Electric Vehicle. *Electronics, 12*, 55. https://doi.org/10.3390/electronics12010055

Hanni, U. e., Yamamoto, T., & Nakamura, T. (2024). Modeling of the Acceptable Waiting Time for EV Charging in Japan. *Sustainability, 16*, 2536. https://doi.org/10.3390/su16062536

Heting, Q., Shuihai, D., Huayan, S., & Jun, Z. (2021, June). Charging station planning based on the accumulation prospect theory and dynamic user equilibrium. *Complex & Intelligent Systems, 9*, 2521–2539. https://doi.org/10.1007/s40747-021-00414-w

Hoen, F. S., Díez-Gutiérrez, M., Babri, S., Hess, S., & Tørset, T. (2023). Charging electric vehicles on long trips and the willingness to pay to reduce waiting for charging: Stated preference survey in Norway. *Transportation Research Part A: Policy and Practice, 175*, 103774.

https://doi.org/10.1016/j.tra.2023.103774

Hussain, A., Lu, Q.-C., Rizvi, S. S., Wang, S., & Kwon, S. J. (2025, October). Short term demand forecasting of electric vehicle charging stations using context aware temporal transformer model. *Scientific Reports, 15*. https://doi.org/10.1038/s41598-025-20557-x

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: a highly efficient gradient boosting decision tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 3149–3157). Red Hook, NY, USA: Curran Associates Inc.

Keskin, F., Yıldız, M., & Arslannur, B. (2024, July). A Data-Driven Framework for Driving Cycle Generation and Analysis. *Transportation Research Record: Journal of the Transportation Research Board*. https://doi.org/10.1177/03611981241260700

Kim, M., Park, W. Y., Moon, H. S., & Kim, Y.-S. (2026). A Review of Electric Vehicle Charging Demand Forecasting Models by Prediction Horizon: A Multi-Criteria Decision Analysis Approach. *Journal of Electrical Engineering & Technology, 21*, 1457–1476. https://doi.org/10.1007/s42835-025-02556-0

Lee, K.-B., Ahmed, M. A., Kang, D.-K., & Kim, Y.-C. (2020). Deep Reinforcement Learning Based Optimal Route and Charging Station Selection. *Energies, 13*, 6255. https://doi.org/10.3390/en13236255

Liu, H., Yin, W., Yuan, X., & Niu, M. (2018). Reserving Charging Decision-Making Model and Route Plan for Electric Vehicles Considering Information of Traffic and Charging Station. *Sustainability, 10*, 1324. https://doi.org/10.3390/su10051324

Mansour, H. S., Mohamed, A. S., & Abdel-Aziz, M. (2026). Electric vehicles charging stations load forecasting based on hybrid XGBoost-BiLSTM model. *Scientific Reports, 16*, 374. https://doi.org/10.1038/s41598-025-29739-z

Matkovic, D., Pilski, T. M., & Capuder, T. (2025, August). Participation of electric vehicle charging station aggregators in the day-ahead energy market using demand forecasting and uncertainty-based pricing. *Energy, 328*, 136299. https://doi.org/10.1016/j.energy.2025.136299

Mishra, P., Miller, E., Santhanagopalan, S., Bennion, K., & Meintz, A. (2022). A Framework to Analyze the Requirements of a Multiport Megawatt-Level Charging Station for Heavy-Duty Electric Vehicles. *Energies, 15*, 3788. https://doi.org/10.3390/en15103788

Mohammad, F., Kang, D.-K., Ahmed, M. A., & Kim, Y.-C. (2023). Energy Demand Load Forecasting for Electric Vehicle Charging Stations Network Based on ConvLSTM and BiConvLSTM Architectures. *IEEE Access, 11*, 67350–67369. https://doi.org/10.1109/access.2023.3274657

Nie, J., Xia, S., Liu, Y., Ding, S., Hu, L., Zhao, M., Fan, Y., Abdel-Aty, M., Preindl, M., & Jiang, X. (2023). A Data-Driven and Human-Centric EV Charging Recommendation System at City-Scale. *Proceedings of the 14th ACM International Conference on Future Energy Systems*, (pp. 427–438). https://doi.org/10.1145/3575813.3597350

Orzechowski, A., Lugosch, L., Shu, H., Yang, R., Li, W., & Meyer, B. H. (2023). A data-driven framework for medium-term electric vehicle charging demand forecasting. *Energy and AI, 14*, 100267. https://doi.org/10.1016/j.egyai.2023.100267

Rupnik, B., Wang, Y., & Kramberger, T. (2025). Hybrid Model for Motorway EV Fast-Charging Demand Analysis Based on Traffic Volume. *Systems, 13*, 272. https://doi.org/10.3390/systems13040272

Schoenberg, S., & Dressler, F. (2023). Reducing Waiting Times at Charging Stations With Adaptive Electric Vehicle Route Planning. *IEEE Transactions on Intelligent Vehicles, 8*, 95–107. https://doi.org/10.1109/TIV.2022.3140894

Sreekumar, A. V., & Lekshmi, R. R. (2024, December). Electric vehicle charging station demand prediction model deploying data slotting. *Results in Engineering, 24*, 103095. https://doi.org/10.1016/j.rineng.2024.103095

Unterluggauer, T., Rauma, K., Järventausta, P., & Rehtanz, C. (2021). Short-term load forecasting at electric vehicle charging sites using a multivariate multi-step long short-term memory: A case study from Finland. *IET Electrical Systems in Transportation, 11*, 405–419. https://doi.org/10.1049/els2.12028

Widmann, J., & Sawodny, O. (2026). Intention-Aware Electric Vehicle Routing: A Markovian Queuing Model Approach to Reduce Wait Times at Charging Stations. *IET Intelligent Transport Systems, 20*. https://doi.org/10.1049/itr2.70180

Yi, Z., Liu, X. C., Wei, R., Chen, X., & Dai, J. (2021, August). Electric vehicle charging demand forecasting using deep learning model. *Journal of Intelligent Transportation Systems, 26*, 690–703. https://doi.org/10.1080/15472450.2021.1966627

Yu, L., Ge, X., Cai, C., & Wang, S. (2026, September). Short-term load forecasting for electric vehicle charging stations based on an adaptive multi-module cascaded framework. *Energy and AI, 25*, 100829. https://doi.org/10.1016/j.egyai.2026.100829

Zhang, T., Tang, D., Fan, P., Wang, Q., & Wang, P. (2025a). A Probabilistic Graphical Model for Predicting Cascade Failures of Electric Vehicle Charging Networks Caused by Hurricanes. *IEEE Transactions on Smart Grid, 16*, 627–639. https://doi.org/10.1109/TSG.2024.3458932

Zhang, Y., Xu, T., Chen, T., Hu, Q., Chen, H., Hu, X., & Jiang, Z. (2025b). A high-resolution electric vehicle charging transaction dataset with multidimensional features in China. *Scientific Data, 12*. https://doi.org/10.1038/s41597-025-04982-1