A Framework for Sequential Charging Facility Location Estimation Problem in Urban Land Use Setting
Final Report

Ashraf Uz Zaman Patwary[1], Francesco Ciari[1,*], Arsham Bakhtiari[1], and Hamed Naseri[1]

[1]Department of Civil, Geological and Mining Engineering, Polytechnique Montréal, Quebec, Canada
[*]Email: francesco.ciari@polymtl.ca

Abstract

The rapid growth of electric vehicles (EVs) necessitates efficient charging infrastructure planning, considering existing facilities. In urban contexts, EV charging times depend on activity durations rather than charging time itself. Considering both effects, this study proposes a sequential, two-step urban EV charger allocation framework. Step 1 uses a modified K-means algorithm to identify candidate locations, incorporating activity locations, participations, and durations. Step 2 employs metamodel-based optimization to allocate charger types and plug counts under setup, operational budget, and power constraints to the candidate locations. Applied to a 10% MATSim Montreal scenario with 74,542 EV users with only 1,392 public chargers, the framework reduced average peak-hour queues by 40% within 10,000 evaluations while respecting 60% increases in setup, operational, and power budgets. Results highlight a preference for deploying more slow chargers over fewer fast chargers in this high-demand scenario. Demand elasticity was observed, suggesting the need for improved behavioral modeling.

Keywords: EV, MATSim, K-Means, optimization, Location choice, Montreal

1 Introduction

With zero runtime emissions and a high energy conversion rate compared to gasoline vehicles (1.5 km/mj vs. 0.28 km/mj) (Nie & Ghamami 2013), electric vehicles (EVs) are increasingly adopted by consumers and policymakers to reduce carbon emissions. The national EV survey for Canada reports a sixfold increase in EV ownership from 2015 to 2018, driven by incentives like tax exemptions, priority lanes, and free charging (Agency 2021). Quebec targets over 90% EV penetration by 2030 (Finance 2020). Advances in battery technology, rising demand, emission goals, and regulatory pressures have made electrification a top priority for the automotive industry (Zhao et al. 2024, Csiszár et al. 2019, Li et al. 2021).

Despite advancements, the unavailability of charging infrastructure remains a critical deterrent to EV adoption (Bailey et al. 2024). Charger unavailability leads to long detours and extended waiting times (Shuai et al. 2024, Csiszár et al. 2019) discouraging potential adopters (E. Seilabi et al. 2024, Zhang et al. 2020, Ma et al. 2024). Therefore, effective planning is essential for EV infrastructure. In urban areas, Individuals' daily activities constrain them to specific network locations at particular times. Thus activity locations, durations, and start/end times govern EV users' charging patterns including charging start time and durations. This connection between charging and activity patterns, including location and duration, in urban settings is well recognized in EV simulation literature (Gharbaoui et al. 2013, Waraich 2013, Gopal et al. 2017, Liu et al. 2022, Chaudhari et al. 2018). However, very few literature applied activity-based charging logic in EV infrastructure planning.

Zhang et al. (2020) used an activity-based traffic model to generate charging requests for an electric ridesharing fleet in San Francisco, identifying charger locations through K-means clustering. However, this work primarily optimized spatial distances while neglecting key metrics such as queuing times, connection durations, and energy served. Similarly, Ma et al. (2024) developed a game-theoretical framework integrating activity networks and charging facility planning, but its scalability to large urban contexts remains limited. Csiszár et al. (2019) employed a land-use approach to optimize charger placement by identifying activity hotspots and parking availability. However, it neither incorporated temporal demand variations nor accounted for activity durations, both of which are crucial for effective urban EV infrastructure planning. Notably, while some studies integrate activity patterns, none explicitly recognize the role of activity durations in governing charging times—a critical factor in urban contexts where charging behavior is closely tied to user schedules.

Another key aspect missing in the literature is the explicit incorporation of existing charging facilities, which, in reality, should significantly influence the placement of new charger facilities and the redistribution of demand following such placements. This study introduces a sequential charger placement algorithm to address this gap. To the best of the authors' knowledge, no existing work explicitly integrates existing charging infrastructure into the charger location optimization process—a crucial step for the sequential development of an efficient charging network in urban areas.

As for the location choice, existing literature has broadly approached the problem using two primary design principles: demand-based charger allocation and flow-capturing charger allocation. Demand-based methods focus on estimating charging demand through simulations or data-driven models, aiming to allocate chargers to meet this demand efficiently. Foundational works, such as Dashora et al. (2010), Frade et al. (2011), and Chen et al. (2013), prioritize minimizing the distance between charging requests and chargers. Recent advancements have expanded these methods, with Liu et al. (2022) and Waraich (2013) developing activity-based simulation frameworks to generate spatial and temporal demand distributions, and Chaudhari et al. (2018) and Gopal et al. (2017) simulating aggregated charging loads across networks. On the other hand, flow-capturing methods prioritize strategically locating chargers to maximize accessibility and coverage. Classical studies, including Kuby & Lim (2005) and Kuby & Lim (2007), use flow-capturing approaches to intercept origin-destination flows along high-traffic corridors. Recent contributions, such as Shuai et al. (2024), optimize charger placement in high-demand areas using genetic algorithms, while Csiszár et al. (2019) adopt a land-use approach to target activity hotspots and parking availability. Similarly, Zhang et al. (2020) uses K-means clustering to identify optimal locations by minimizing the distance between chargers and activity-based charging requests. As both approaches address key aspects of EV users' behavior and charging dynamics, in this study, we combine these approaches into a multi-step framework.

Given the above literature landscape, this paper proposes a two-step, activity-driven, sequential charger allocation framework in the urban context, with existing charging infrastructures. The two steps combine both design principles i.e., capture activity trajectories with most travelers while satisfying generated charging demands.

2 Problem statement

In this research, we develop the sequential charger allocation algorithm from the perspective of a regulating body where the primary concern is to improve user experience and utilization of chargers at a given budget while satisfying power side constraints. In that respect, the charger allocation problem can be formulated as a mixed integer, constrained, non-linear queue minimization problem, where the objective is to minimize the average queue at chargers given a current EV charging infrastructure, fixed monetary setup, and operation budgets for new infrastructure placements and maintenance under certain zonal max power draw constraint. The decision variables include the integer variables determining the number and types of plugs at each charging infrastructure, for a site where no chargers should be placed the decision variables, primarily the number of plugs can be set to zero. This formulation does require the number of cites for chargers placement to be a limited discrete number. The problem dimension increases threefold to this number and the combinatorial solution space grows exponentially.

Below is the mathematical formulation of the problem to be solved. Here, $i \in I$ represents the candidate spots for new chargers and $j \in J$ represents the current charger locations. Q_i denotes the average queue and V_i denotes the power draw per plug at charger i. The decision variables denoted by x_i and p_i denote charger type and plug count at candidate location i. x_j and p_j similarly denote the charger type and plug count at the current charger locations j. $C_{s,x}$ and $C_{o,x}$ are the setup and operation cost of charger type x. B_o and B_s are budget constraints for the operation and setup budget. z represents zones and I_z and J_z are the set of candidate and current charger locations respectively.

$$
\begin{aligned}
& min_{x_i,p_i} \frac{1}{|I|+|J|} \left\{ \sum_{i \in I} Q_i + \sum_{j \in J} Q_j \right\}; \forall i \in I \\
& \quad x_i, x_j \in X \{Level\ 1,\ Level\ 2\ or\ Fast\} \\
& \quad p_i, p_j \in P \{0, 1, 2, ..., p_{max}\} \\
& s.t. \\
& \quad \sum_{i \in I} p_i \times C_{o,x_i} + \sum_{j \in J} p_j \times C_{o,x_i} \leq C_o \\
& \quad \sum_{i \in I} p_i \times C_{s,x_i} \leq C_s \\
& \quad \sum_{i \in I_z} V_i + \sum_{j \in J_z} V_j \leq V_z; \forall z \in Z
\end{aligned}
\tag{1}
$$

We use Micro Agent Traffic Simulation, i.e., MATSim (W Axhausen et al. 2016) for simulating electric vehicle charging in the proposed urban context. MATSim is an activity-based, agent-driven traffic microsimulation model where individual agents try to maximize their utility (score) by improving both their trip and activity patterns (replanning). The model reaches equilibrium when no travelers can improve their utility by unilaterally switching their activity (duration, start time, charging, etc.) or trip-related (mode, route, etc.) choices. Figure 1 shows the different steps of the MATSim loop.

Table 1: Notation Table

Notation	Description
i	Candidate charger location
j	Existing charger location
I	Set of candidate charger location
I_f	Set if candidate chargers in the choice set of facility f
I_z	Set if candidate chargers in zone z
J	Set of existing charger location
J_f	Set if existing chargers in the choice set of facility f
J_z	Set if existing chargers in zone z
f	Activity facility
F	Set of activity facilities
x	Charger type variable
X	Vector containing all $x_i; I \in I \cup J$
p	Charger plug count variable
P	Vector containing all $p_i; I \in I \cup J$
h	Hour
H	Set of hours
$\delta_{f,h}$	Incident variable, 0 if peak hour at facility f is h, 0 otherwise
Q_i	Average queue at charger i
q_f	Activities performed by EV user at facility f
$q_{f,i}$	Demand from facility f to charger i
q_i	Demand in charger i
$q_{i,h}$	Demand at charger i at hour h
$d_{f,i}$	Distance from facility f to charger i
C_o	Max operational budget for the optimization
C_s	Max setup budget for the optimization
C_{s,x_i}	Cost of setting up 1 plug for charger type x_i
C_{o,x_i}	Cost of maintaining 1 plug for charger type x_i
c_i	Cost of using charger i
b	Average battery capacity
$U_{f,i}$	Utility of choosing charger i from facility f
$\omega_{f,i}$	Probability of choosing charger i from facility f
t_i	Average charging time at peak hour at charger i
$t_{0,i}$	Average intended charging duration at charger i
$t_{0,i,h}$	Average intended charging duration at charger i at hour h
z	Power zones
Z	Set of power zones
V_i	Total energy draw at charger i during peak hour
V_z	Total energy draw in zone z during peak hour
v_{x_i}	Charger power of type x_i
β_m	Marginal utility of money
β_t	marginal utility of time
β_r	marginal utility of charger attractiveness
β_d	Marginal utility of distance
η	Scaling parameter of logit
ρ	Peak hour factor
α	Parameter of volume delay function for t_i
γ	Parameter of volume delay function for t_i

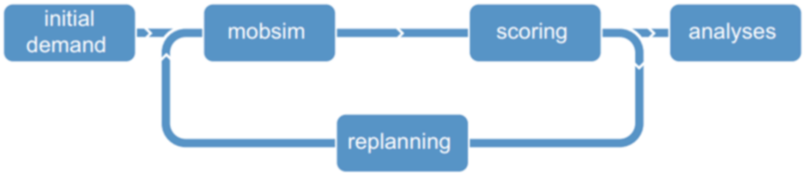


Figure 1: MATSim Equilibrium Loop

3 Methodological Framework

In this research, we propose a two-step, sequential charger location estimation framework that integrates demand-based and flow-capturing approaches while accounting for urban land-use patterns and charging behaviors. The two steps are:

- In the first step, candidate charger location set I is selected by maximizing the capture of agent activity locations and durations using a modified K-means algorithm specialized to handle existing facilities.
- In the second step, a single-shot physical metamodel-based optimization is performed to determine the optimal location and plug counts of the new charger facilities.

Finally, we analyze and evaluate our results in the original simulator. Figure 2 shows the two steps, the flow of information between them, and the corresponding data requirement at each step in the proposed charger location choice optimization framework.

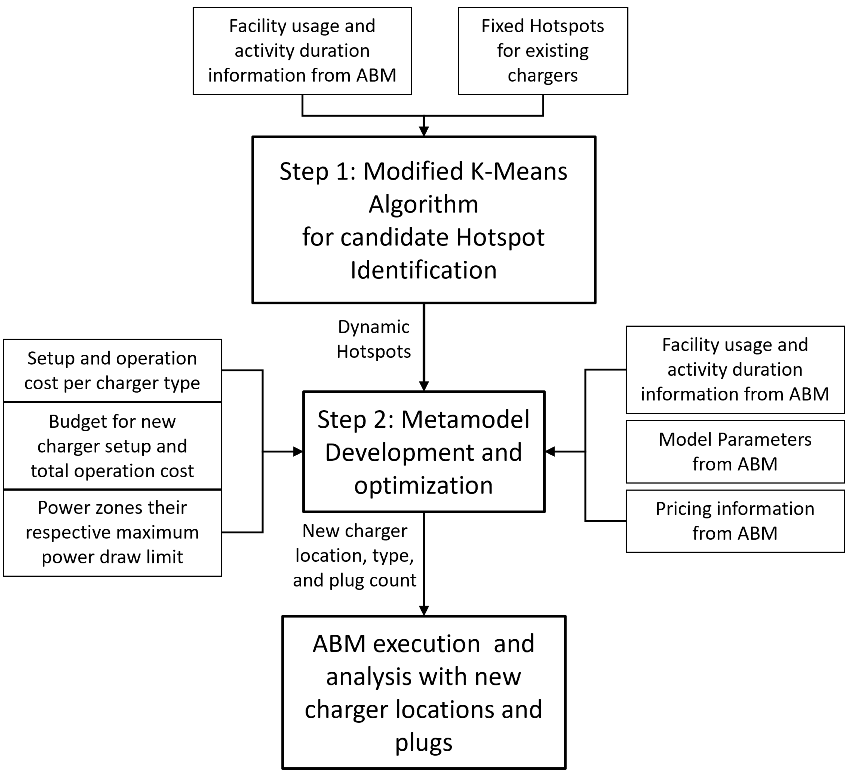


Figure 2: Schematics of the proposed charger location estimation algorithm.

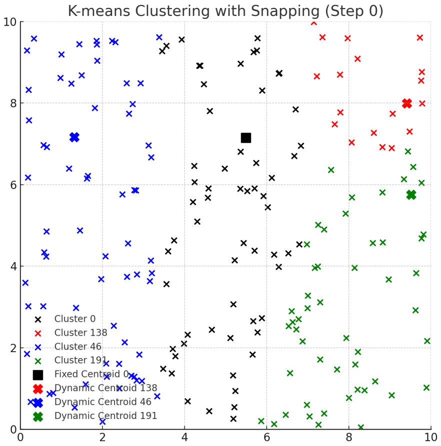
3.1 Step 1: Modified K-Means Algorithm

The first step aims to reduce the solution space by pre-determining the candidate charger locations (i.e., hotspots) using a modified k-means clustering algorithm. As K-means requires the number of clusters beforehand, the number of candidate locations to be generated from this algorithm becomes a configurable parameter. Thus the framework allows the modeler to choose the complexity level of the second step optimization problem.

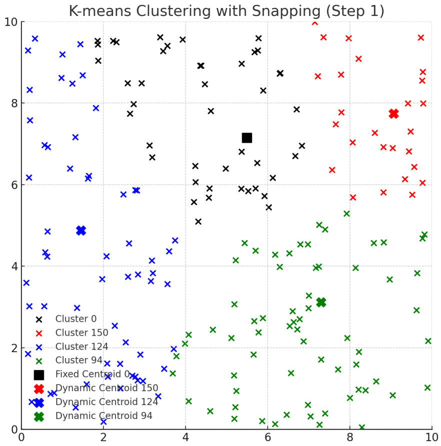
The modified k-means algorithm iterates over two types of clusters, dynamic and static, where only the dynamic cluster centroids are updated over the training process. Additionally, the centroid always snaps to the nearest activity facility location to facilitate charging while performing activities. Here, the dynamic clusters represent the candidate locations and the static clusters represent the preexisting chargers in the network. The feature vector for the modified k-means algorithm contains the location (x,y coordinates) and the number of users in each activity facility location as the feature vector for each activity facility location. As a result, the cluster centroids of the dynamic clusters will be closer in distance to the facilities with a higher number of activities performed. We further investigated including activity durations in the feature vectors of this step. The results are presented in the results section of this report. The modified algorithm is presented in algorithm 1. Figure 3 showcases a small problem with four cluster centroids (one static and three dynamic) with 100 activity facility locations in a small rectangular region.

Algorithm 1 Modified K-means with Fixed and Dynamic Clusters

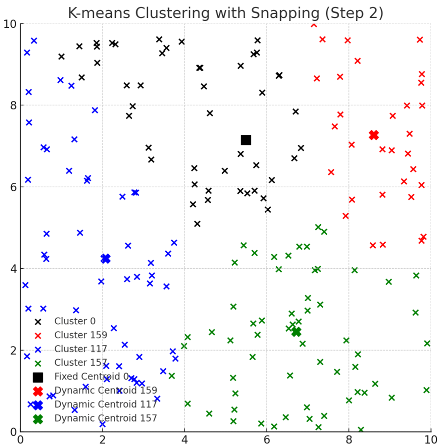
Input: Set of activity facilities F, fixed centroids $J \subseteq F$, initial dynamic centroids $I \subseteq F$, maximum iterations T, tolerance ϵ
Output: Final dynamic centroids I and facility assignments
Initialize assignments for all $f \in F$
for each iteration $k = 1, 2, \ldots, K$ **do**
    Step 1: Assign Facilities to Clusters
    for each activity facility $f \in F$ **do**
        Compute the distance from f to all centroids in $J \cup I$
        Assign f to the nearest centroid
    Step 2: Update Dynamic Centroids
    for each dynamic centroid $i \in I$ **do**
        Calculate the mean of facilities assigned to i
        Snap i to the nearest facility $f \in F$ to the calculated mean
    Step 3: Check for Convergence
    Compute the total change in centroids Δ
    if $\Delta < \epsilon$ **then**
        Stop
Return: Updated I and facility assignments



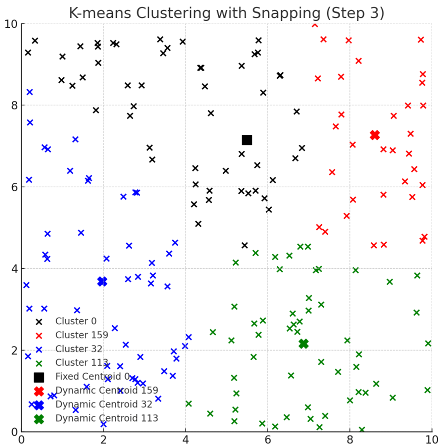
(a)



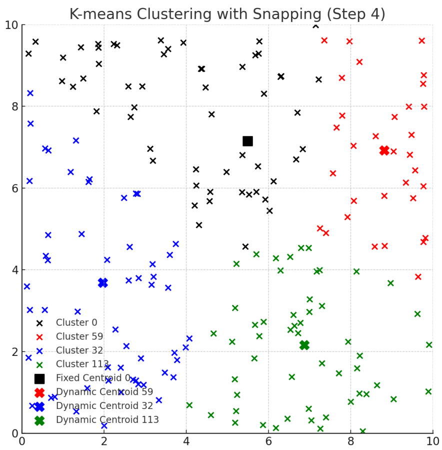
(b)



(c)



(d)



(e)

Figure 3: Iteration 1-5 (a-e) for the modified K-means clustering algorithm

3.2 Step 2: Metamodel development and optimization

We have used the urbanEV contribution of MATSim developed by Bakhtiari et al. (2024). This model simulates EV charging, usage, and discharging in the urban context where charging is coupled with activities. Assuming the mapping between charger placement to queue and charger placement to power draws is nonlinear, the problem is a combinatorial optimization problem with a nonlinear objective and mixed (linear and nonlinear) constraints. With a costly to evaluate, activity-based microsimulation model without closed-form solutions embedded in the objective, the problem quickly becomes intractable to solve directly.

In most engineering optimization problems with similar intractability, metamodeling (Khatouri et al. 2022) is used. Metamodeling is a simplified model of a complex simulation model that is developed for faster approximation of the original function. Metamodels can be categorized into two broad categories: general-purpose metamodels and problem-specific metamodels. General-purpose metamodels are surface approximations with general-purpose functions such as ridge and lasso regression (Tibshirani 1996, Hoerl & Kennard 1970), gradient boosting(Friedman 2001), and extreme gradient boosting techniques (Ester et al. 2022), etc. Problem-specific metamodels, on the other hand, are simplified mathematical representations of the original problem with simpler dynamics for faster evaluation (Khatouri et al. 2022, Razavi et al. 2012, Kianifar & Campean 2020). There are hybrid metamodeling techniques that combine both types of metamodel for more efficient training and evaluation Zhang et al. (2017), Patwary et al. (2021).

While physical metamodels with dynamics similar to the original simulation can achieve comparable performance without requiring extensive training or costly-to-generate datasets, they often incur significantly higher computational costs during optimization compared to their general-purpose counterparts. Therefore, the choice of metamodel depends on the trade-off between the simulation runtime required for dataset generation and the computational complexity and runtime of the metamodel itself.

Given the combinatorially large solution space of our problem and the high computational cost of evaluating the underlying simulation model, generating a dataset sufficiently large to cover the entire solution domain for training general-purpose or hybrid metamodels is impractical. Consequently, our framework employs a single-shot, problem-specific metamodel with simplified dynamics to approximate the outputs of the MATSim urban EV module during optimization. Despite being single-shot, the metamodel retains the behavioral parameter set of MATSim, ensuring consistency in the underlying logic. Moreover, these parameters can be manually adjusted to refine the approximation and further align the metamodel's outputs with those of the original simulator.

Therefore, step 2 solves the optimization problem presented in equation 1 except the queue and power draw for chargers at locations $i \in I$ and $j \in J$ are approximated using a problem-specific Demand Allocation metamodel.

3.3 Demand Allocation Metamodel Formulation

For a given solution, $[X, P]$ the metamodel approximates three outcomes of charger demand allocation: demand per charger (q_i,q_j), average intended charging duration per charger ($t_{0,i}, t_{0,j}$), and finally average charging time t_i, t_j required per charger for every charger location $i \in I$ and $j \in J$. The difference between intended charging duration t_0 and actual charging duration t can be understood with their parallel to the free flow and actual link travel time in static traffic assignment models. $t_{0,i}$ here is the average time the EV users want to charge

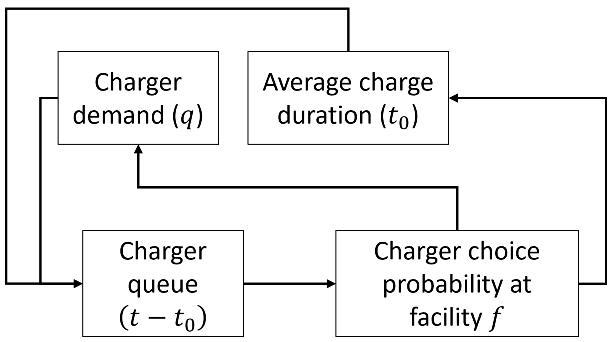


Figure 4: Wardrop's Equilibrium in the Demand Allocation Model

at facility i whereas t is the time they require to charge for that amount. So, charging time t includes charging duration t_0 and any incurred delay.

These three terms are interconnected, as the demand at a charger q governs its queue time t, which in turn influences EV users' charger choice probability. Charger choice probability from different activity facilities around a charger in turn governs the average intended duration of charging t_0, the daily peak hour of charging, and the peak hour's charging demand at that charger q. This cyclic dependency crates a Wardrops equilibrium (Wardrop & Whitehead 1952) where the equilibrium is reached when no EV user can improve their charging time by unilaterally changing their charger choice. Figure 4 shows the interdependency between different components of the proposed demand allocation model.

For this model, we assumed the choice set of chargers for EV users performing a certain activity at a facility is bound by the distance between the activity facility and the charger location. To minimize computational burden, we have further added a limit on max number of chargers to put in a facility's charger choice set. For facility f, this choice set is defined by $I_f : d_{i,f} \leq d_{max}$.

The probability of choosing charger i from facility f, $\omega_{f,i}$ is calculated using the logit model and can be written as follows. Here the utility includes queue time $(t_i - t_{i,0})$, distance $d_{f,i}$, charging cost c_i if any, and the obtained charge to battery capacity ratio $\frac{t_{0,i}\ v_{x_i}}{b}$ for determining the attractiveness of a charger.

$$\begin{aligned} \omega_{f,i} &= \frac{e^{\eta U_{f,i}}}{\sum_{i' \in I_f} e^{\eta U_{f,i'}}} \\ U_{f,i} &= \beta_t \times (t_i - t_{0,i}) + \beta_d \times d_{f,i} + \beta_m \times c_i \\ &\quad + \beta_r \times min(\frac{t_{0,i}\ v_{x_i}}{b}, 1) \end{aligned} \tag{2}$$

Queuing effects are captured while calculating t_i from $t_{0,i}$ using the volume delay function as below. Here, α and γ controls the smoothness of the curve. In our experiments, $\alpha = 0.15$ and $\gamma = 1$.

$$t_i = t_{0,i} \left\{ 1 + \alpha \times \left(\frac{q_i \times min(t_{0,i}, 3600)}{3600 \times p_i} \right)^{\gamma} \right\} \tag{3}$$

After calculating the utility and charger choice probability, we can get hourly demand for a charger from surrounding facilities using equation 4. Here, facility demand is multiplied by the facility to charger probability and ρ, the peak hour factor.

To estimate the peak hour demand for electric vehicle (EV) charging, it is important to account for two primary factors: the proportion of EV users likely to charge at a specific facility and the proportion of total daily demand occurring during peak hours. Studies indicate that, depending on the facility type, approximately 60-80% of residential EV users and 30-50% of workplace or public charging users will utilize charging services at their respective locations throughout the day (Bailey et al. 2024, Agency 2021). Additionally, charging patterns suggest that peak hour demand—typically during evening or late afternoon hours—accounts for around 15-20% of the total daily usage (Bailey et al. 2024, Finance 2020). Combining these factors allows for the calculation of a peak hour factor, which is the product of the proportion of users charging and the peak hour demand percentage. In our case, we assume 60% of users charging at a facility on average and 20% of this charging demand occurs during the peak hour, the peak hour factor is calculated as $0.6 \times 0.2 = 0.12$, meaning 12% of the daily demand occurs during the peak hour.

$$q_{i,h} = \sum_{f \in F} \rho \, \omega_{f,i} \, q_f \, \delta_{f,h} \tag{4}$$

Weighted average durations from these facilities according to their hourly demand give the hourly intended charging duration as shown in equation 5.

$$t_{0,i,h} = \frac{\sum_{f \in F} \rho \, \omega_{f,i} \, q_f \, \delta_{f,h} \, t_{0,f}}{\sum_{f \in F} \rho \, \omega_{f,i} \, q_f \, \delta_{f,h}} \tag{5}$$

Finally, the maximum of these hourly demands ($q_{i,h}$) is chosen as the design charger demand q_i and the corresponding average intended charging duration is chosen as the intended charging duration for that charger and for that demand. The process is expressed in mathematical form as shown in equation 6.

$$\begin{aligned} q_i &= max_{h \in H} \; q_{i,h} \\ t_{0,i} &= t_{0,i,h^*}; h^* = argmax_{h \in H}(q_{i,h}) \end{aligned} \tag{6}$$

The cyclic dependencies among the system of equations 2-6 create a stochastic user equilibrium. We solve this system of equations using the accelerated method of successive average (AMSA) proposed by Liu et al. (2009). Once the equilibrium is solved, the hourly energy draw per zone ($V_{z,h}$) can be calculated by summing up the hourly charger power draws $V_{i,h}$ for chargers in that zone. The maximum value among the hourly power draw is the maximum energy draw per zone V_z. This value will be used to calculate the zonal power constraints. The process is explained mathematically in equation 7.

$$\begin{aligned} V_{i,h} &= max(b, t_{0,i,h} * v_{x_i} * q_{i,h}) \\ V_{z,h} &= \sum_{i \in I_z \cup J_z} V_{i,h} \end{aligned} \tag{7}$$

3.4 Metamodel Optimization

With the above formulations, the objective in equation 1 can be reformulated as below.

$$
\begin{aligned}
&min_{x_i,p_i} \frac{1}{|I|+|J|}\left\{\sum_{i\in I}(t_i - t_{0,i}) + \sum_{j\in J}(t_j - t_{0,j})\right\};\forall i \in I \\
&\quad x_i, x_j \in X\,\{Level\ 1,\ Level\ 2\ or\ Fast\} \\
&\quad p_i, p_j \in P\,\{0,1,2,...,p_{max}\} \\
&s.t. \\
&\quad \sum_{i\in I\cup J} p_i \times C_{o,x_i} \le B_o \\
&\quad \sum_{i\in I} p_i \times C_{s,x_i} \le B_s \\
&\quad max_{h\in H}(V_{z,h}) \le V_z; \forall z \in Z
\end{aligned}
\tag{8}
$$

The decision variables in this optimization process are composite and take discrete values. The SUE demand allocation model developed in the previous section is nonlinear, making both the objective function and the zonal power constraints nonlinear. Furthermore, the high dimensionality of the problem (2400 variables) and the absence of a closed-form solution necessitate the use of derivative-free metaheuristic optimization methods. We used OPT4J a library in JAVA specialized for meta heuristics optimization for solving our problem. The evolutionary algorithm in OPT4J supports nonlinear large-dimension problems, however, with an 11s runtime of the metamodel, we had to keep the maximum budget of 10,000 evaluations. GA however, does not support constrained optimization. Hence, we moved the nonlinear i.e., the zonal power constraint to the objective as below.

$$
\begin{aligned}
&min_{x_i,p_i} \frac{1}{|I|+|J|}\left\{\sum_{i\in I}(t_i - t_{0,i}) + \sum_{j\in J}(t_j - t_{0,j})\right\} + \\
&\quad \lambda \times min\left\{max_{h\in H}(V_{z,h}) - V_z, 0\right\};\forall i \in I \\
&\quad x_i, x_j \in X\,\{Level\ 1,\ Level\ 2\ or\ Fast\} \\
&\quad p_i, p_j \in P\,\{0,1,2,...,p_{max}\}
\end{aligned}
\tag{9}
$$

λ here denotes the weight to prioritize constraints satisfaction vs objective minimization. In our experiments, we choose $\lambda = 1$. With a vast solution space most of which violates the budget and power limit constraints, we found that it is beneficial both in terms of convergence rate and objective quality to put the budget constraint in the random solution generation process rather than in the objective. Algorithm 2 shows the pseudo-code that was used to generate unbiased random solutions within the budget constraints. Manual hyperparameter tuning was performed for faster convergence.

4 Experimental Setup and Results

4.1 Scenario Description

The experimental setup for this study focuses on a detailed scenario derived from the Montreal network. A 10% sample of the population, consisting of 297,128 individuals, was used, with 1,392 existing public chargers and no provision for home charging. The scenario assumes an EV penetration rate of 25%, with 74,542 EV owners, significantly higher than Canada's 2023 national EV rate of 1.3% (proportion of car owners not proportion of total travelers) but somewhat aligned with the rapidly increasing adoption seen in Quebec, where 25% of new

Algorithm 2 Budget-Constrained Charger Configuration Initialization

Input: Number of chargers, maximum charger configurations (type and plug pairs), setup budget C_s, operation budget C_o, cost functions for creation $c_{s,x}$ and operation $c_{o,x}$
Output: Charger configurations (type and plug count)
Initialize all chargers with minimal configuration (type 0, plug count 0)
Track remaining setup and operation budgets C_s and C_o
while both $C_s > 0$ and $C_o > 0$ **do**
    Randomly select a charger
    Compute feasible configurations (type and plug count pairs) based on remaining budgets
    if a feasible configuration exists **then**
        Assign a feasible configuration (type and plug count pair) that fits within the budgets
        Update the remaining budgets C_s and C_o accordingly
    else
        Reduce feasible configuration space and retry
    if no feasible configuration can fit any charger **then**
        Break
Return: Final charger configurations (type and plug count for each charger)

vehicle registrations were electric. Additionally, no home charger availability will produce a longer queue creating a high-demand context to rigorously test the proposed charger placement framework.

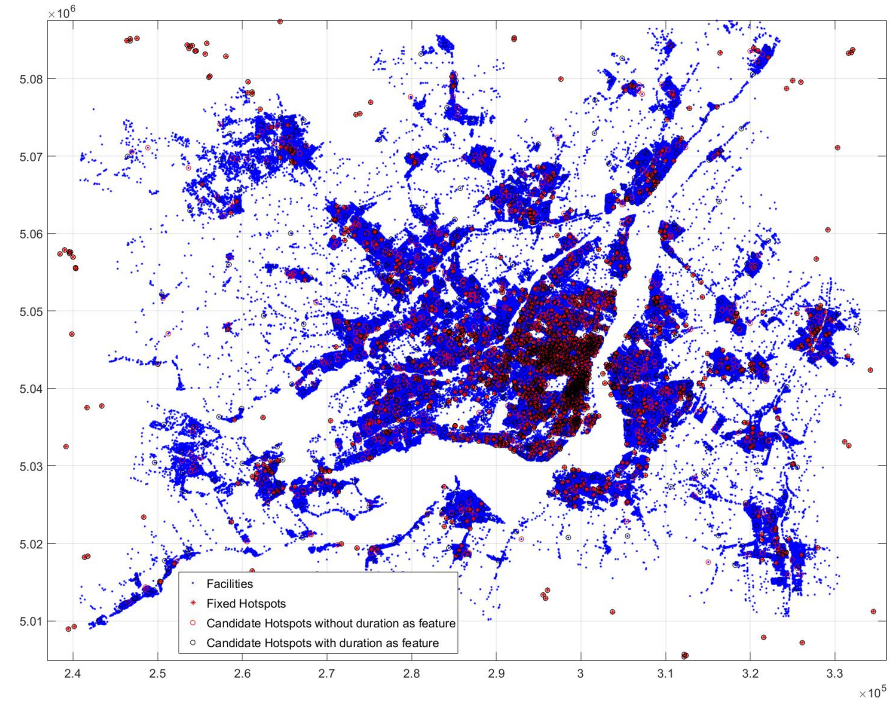
The charging infrastructure in the scenario includes chargers with a maximum plug limit of 10, with costs varying based on charger type. Level 1 chargers incur a setup cost of $5k per plug, Level 2 chargers cost $10k per plug, and Fast chargers cost $20k per plug. The operation costs are set at $200, $400, and $800 per plug for Level 1, Level 2, and Fast chargers, respectively. Budget constraints are introduced to ensure realistic deployment conditions, with setup costs capped at 60% of the cost required to establish the current facilities and operation costs set at 160% of the expense for running existing facilities.

The network is further divided into six zones, each with a power draw limit set to 1.6 times the current power draw. To address the spatial distribution of chargers, 2,500 potential hotspots were considered, including the 1,392 fixed existing charger locations. This results in a total problem dimension of 2,216 decision variables. The combination of elevated demand, budgetary constraints, and zonal power limits creates a challenging scenario for testing the proposed optimization and allocation framework.

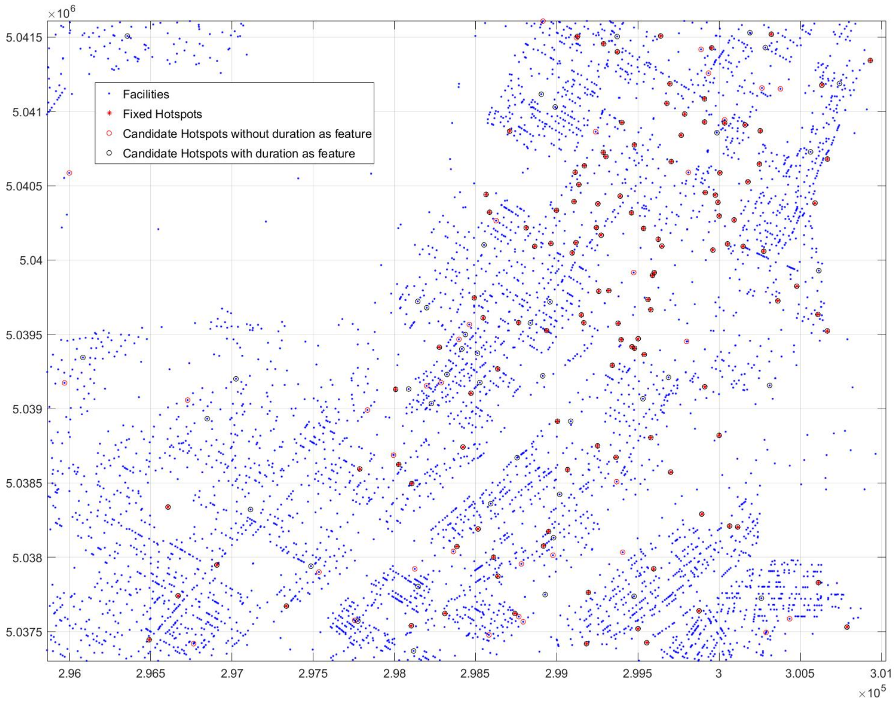
4.2 Step 1: Modified K-mean Algorithm Results

Figure 5 illustrates the outcomes of Step 1 for two different feature configurations. In the first case, the k-means algorithm used only location and facility usage data as input features. In the second case, the modified k-means algorithm incorporated location, facility usage, and the average activity duration of EV users into the feature set. While facility usage intuitively guides cluster centroids toward high-demand facilities, the impact of including activity duration remains uncertain and is investigated through this experiment. This step aims to evaluate whether prioritizing longer activity durations improves the clustering outcome or introduces unintended biases.

To further evaluate the impact of the feature vector configurations, we assigned fast chargers with 10 plugs to all candidate charger locations generated by both configurations.



(a)



(b)

Figure 5: Existing and potential charger facilities among activity facilities for different feature vector configuration (a) Montreal network and (b) Zoomed-in portion of Montreal network

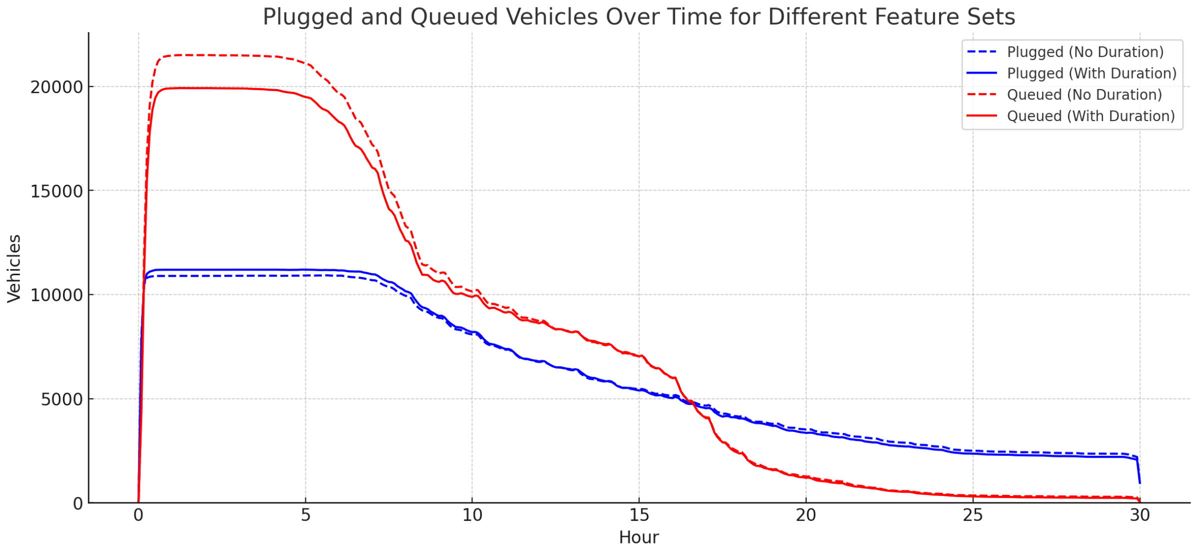


Figure 6: Vehicle plugged and queued in MATSim urban EV scenario for different feature configuration

This ensures that both configurations require the same monetary budget and, in the absence of zonal power limit constraints, would likely result in similar levels of energy draw. We then ran the MATSim urban EV scenario to observe the outcomes. Figure 6 presents the total number of vehicles plugged in and queued throughout the day for both feature configurations, providing insights into their performance under identical resource allocation.

The results indicate that, for similar resource allocation, the peak hour vehicle plug-in count was slightly higher when activity duration was included in the feature vector compared to when it was excluded. Simultaneously, the inclusion of activity duration led to a significant reduction in the number of queued vehicles, suggesting a clear benefit in incorporating activity duration as a feature for determining candidate charger locations using the proposed k-means algorithm.

To assess charger availability, we examined the number of agents unable to find chargers near their vicinity. This number was negligible in both cases—350 versus 650 out of 25,000 requests—highlighting the adequacy of charger coverage in both configurations. Interestingly, the configuration without activity duration performed marginally better in terms of charger availability. This may be attributed to the inclusion of activity duration shifting charger clusters closer to locations with longer activities, such as residential areas where home activities dominate. Consequently, fewer chargers may be allocated to commercial or office areas, where peak-hour queues typically form during work hours.

Another possible explanation is that longer activity durations can lead to fewer turnovers at chargers, as vehicles remain plugged in for extended periods. This could inadvertently lower the availability of chargers in areas with high demand during specific time periods, such as office hours. Conversely, excluding activity duration results in a more spatially distributed placement of chargers, balancing the demand between residential and non-residential areas, albeit at the cost of higher peak-hour queues.

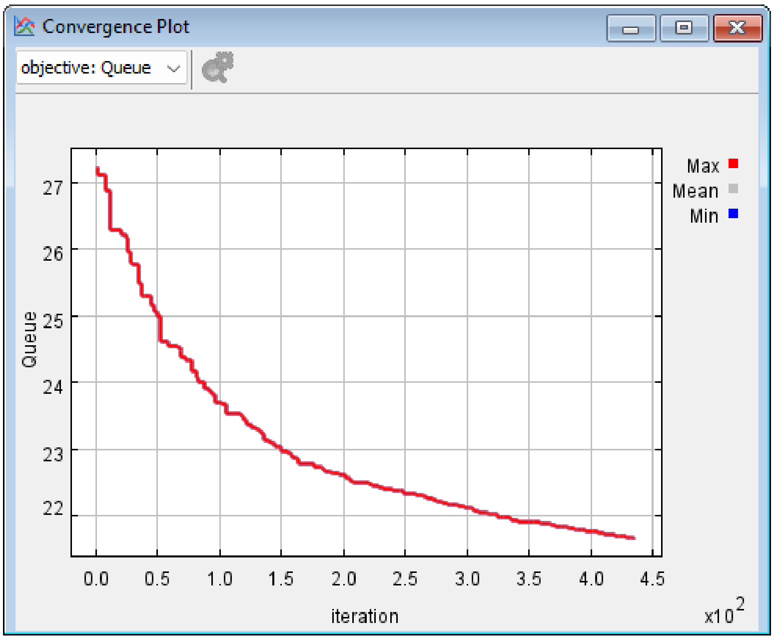


Figure 7: Convergence of Step 2: Metamodel Optimization

4.3 Step 2: Metamodel Optimization Results

In the optimization phase of the study, we began by evaluating the performance of the base scenario, which included 1,391 public chargers with a total of approximately 2,000 to 4,000 available charger plugs, servicing 75,000 EV users. This scenario exhibited an average peak-hour queue of 36.2 hours, as calculated using the volume delay function. It is important to note that this queue value does not represent practical real-world queuing times but instead reflects the severe mismatch between supply and demand. This mismatch was intentionally created by excluding home chargers, resulting in a scenario where the available infrastructure is insufficient to meet the high demand. The scenario is designed to mathematically stress test the developed charger design framework under extreme conditions, emphasizing its ability to handle high-demand scenarios effectively.

To explore potential improvements, a random assignment approach was implemented. This involved distributing plugs and charger types randomly across the network, utilizing the full budget. The results, averaged over 25 scenarios, showed a reduction in the average peak-hour queue to 27.3 hours. This reduction sets the benchmark for the optimization algorithm as it represents the bare minimum one should be able to reduce the objective even with random assignment with similar resource constraints.

Finally, the proposed optimization framework achieved a significant improvement, reducing the average peak-hour queue to 21.47 hours after 500 generations of the genetic algorithm. This optimized solution demonstrates the effectiveness of the structured optimization framework in allocating resources more efficiently. Figure 7 illustrates the convergence process over successive iterations, underscoring the robustness of the proposed methodology in addressing severe demand-supply imbalances in the absence of home charging facilities.

As shown in Figure 8, the optimized solution deployed a significantly higher number of new plugs compared to the original scenario while utilizing only 60% of the total budget. Notably, the optimization process prioritized the quantity of plugs over the installation of higher-power chargers. This strategy reflects the modeled charging behavior, where charging

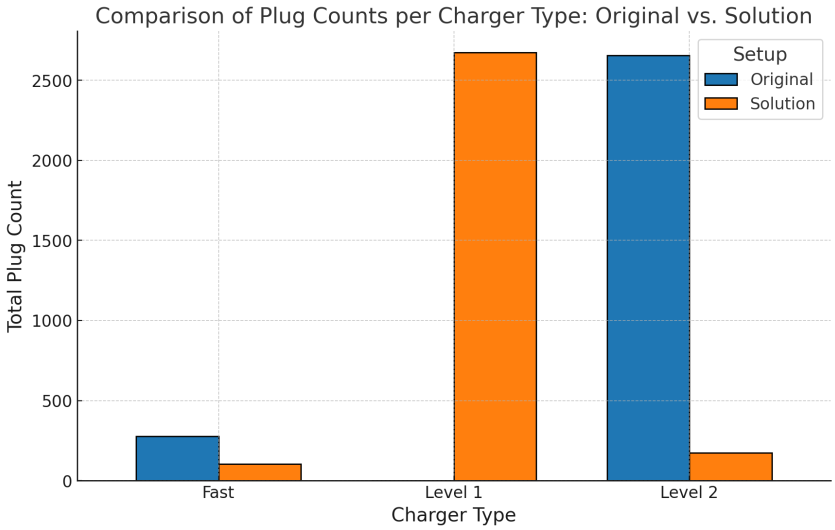


Figure 8: Charger type composition in original vs optimized new chargers

durations are governed by activities not plug types, and hence in a severely supply-demand imbalanced scenario, reducing queue requires more plugs rather than faster chargers. The optimization algorithm effectively captured this rationale, demonstrating its ability to extract the underlying dynamics of charging behavior and infrastructure requirements.

The spatial distribution of the optimized chargers is depicted in Figure 9. Here the size of the dots represents the number of plugs in that charger. The optimized solution demonstrates a clear preference for installing multiple Level 1 chargers in less congested areas rather than deploying higher-powered Level 2 or fast chargers. Fast chargers are predominantly located in high-activity zones where demand peaks, such as central business districts or other busy areas with significant throughput. Interestingly, the optimization process revealed an almost negligible deployment of Level 2 chargers. This outcome may stem from the modeled trade-off between the cost-effectiveness of Level 1 chargers and the high capacity utilization of fast chargers in busy zones. Overall, the optimization process did capture the spatial demand and activity type variability over the Montreal network.

While running the optimization solution in MATSim Montreal within the urbanEV framework, a substantial reduction in the average queue per plug was noticed compared to the base scenario. This outcome reflects the improved distribution of charging resources across the network. At the same time, the average number of vehicles plugged into chargers per plug decreased slightly, which corresponds to the addition of a larger number of slower Level 1 chargers in the optimized solution. These chargers increase coverage but have lower throughput, leading to a marginal decline in average plug utilization. Figure 10 shows the number of vehicles plugged in and queued per plug when running MATSim Montreal with the optimal charger configuration. The peak number of queued vehicles, however, increased compared to the base scenario. This result suggests the presence of demand elasticity, where improved charging infrastructure attracts additional users who previously refrained from charging due to long queues or lack of availability. In future research, we plan to include demand elasticity in our metamodel for a more precise sequential charger allocation allocation algorithm in the urban context.

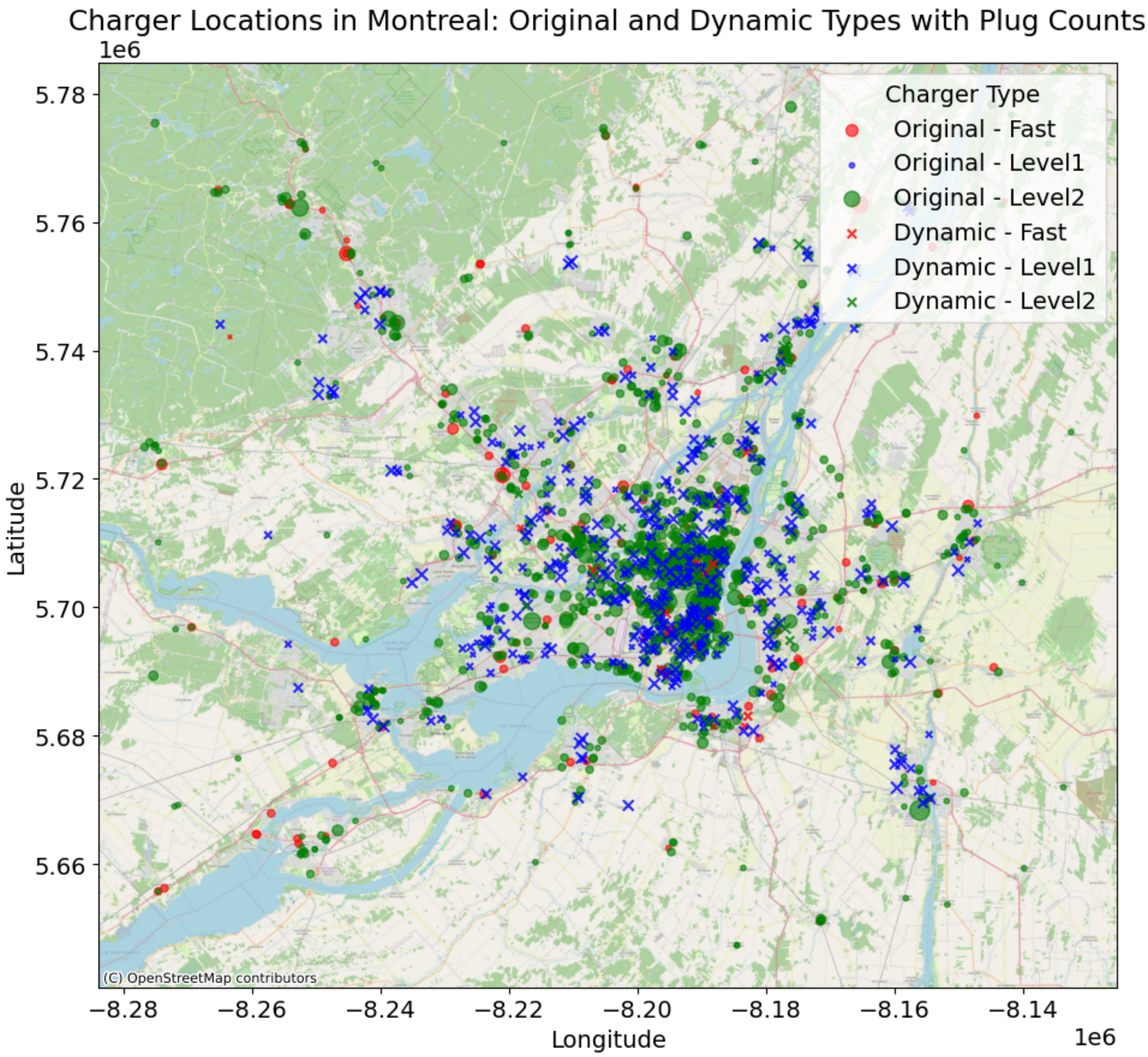


Figure 9: Spatial distribution of chargers in the optimization result

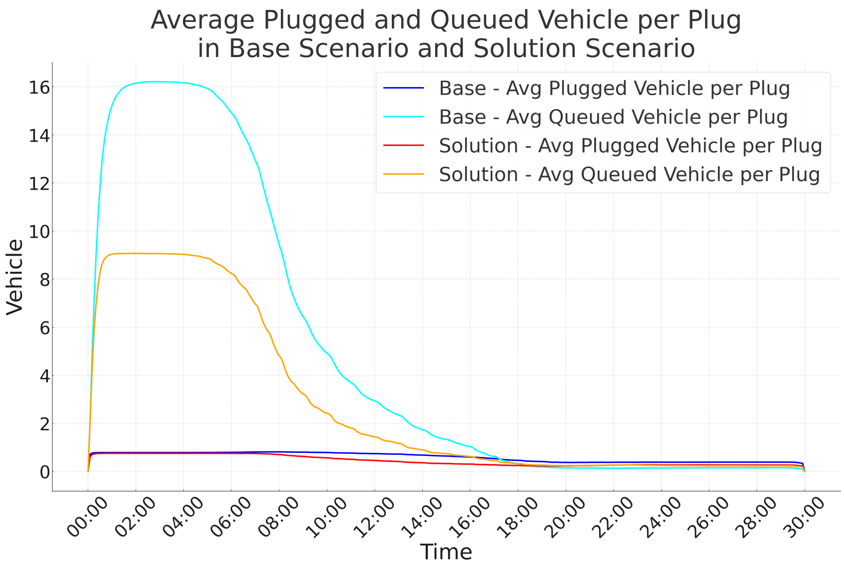


Figure 10: Normalized vehicle plugged in and queued per plug in the optimized urbanEV scenario

5 Conclusion and Future Work

This study introduced a comprehensive framework for addressing the sequential EV charger location and type allocation problem in urban contexts where charging duration is governed by activity pattern and activity duration rather than vehicle state of charge. The proposed framework combined both demand-based and flow-capturing charger allocation approaches present in the literature by developing a two-stage algorithm.

In the first step, a modified k-means clustering algorithm identified candidate charger locations by incorporating the spatial distribution of facilities, facility usage, and average activity duration. The aim of this step was to capture EV users' activity trajectories and at the same time reduce complexity and problem dimension for the second stage optimization problem. In this step, we further explored multiple feature configurations and their effects on charger placement using the proposed k-means clustering algorithms for candidate location selection. The results indicated that with activity duration incorporated into the feature vector, the proposed modified k-means algorithm works better in distributing chargers to reduce peak queue and increase peak utilization within the same resource constraints.

The second step involved a metamodel-based optimization process to allocate charger types and plugs across the identified candidate locations in the first step while satisfying operational and setup budget constraints and zonal power level constraints. A simplified stochastic user equilibrium-based charger demand allocation model was developed which approximates the charger demand, average charging duration, and charging queue given a charger type and plug configuration in the optimization process. The metamodel was quite efficient with an average equilibrium convergence of 11-seconds for the large-scale MATSim Montreal scenario. The metamodel integrated behavioral parameters from MATSim, maintaining consistency with the activity-based simulation environment. An evolutionary algorithm with a custom random sampler was created to enforce the resource constraints efficiently in the sampling method while ensuring unbiased sampling. The results revealed that, in scenarios with severe supply-demand imbalances, the optimization process favored deploying multiple low-powered chargers (Level 1) over fewer high-powered ones. This preference aligns with the modeled dynamics, where activity-driven charging durations often make throughput less critical than coverage.

The optimized solution achieved a significant reduction in queuing times compared to the baseline, despite being deployed in a heavily overloaded scenario. The spatial distribution of chargers further demonstrated the model's capacity to allocate resources strategically, prioritizing fast chargers in high-demand areas and Level 1 chargers in less congested regions. However, the peak number of queued vehicles increased due to demand elasticity, highlighting the necessity of a feedback loop between infrastructure improvement and user behavior.

Applied to the large-scale, real-world scenario of greater Montreal, characterized by significant supply-demand imbalances, the proposed framework demonstrated robustness and effectively captured the intuition of prioritizing quantity over quality in the given urban context.

Future project extensions can expand on the following ideas.

- Multi-objective optimization to minimize life-cycle emissions.
- Enhanced metamodels for better demand elasticity predictions.
- Integration of machine learning for dynamic optimization.
- Incorporation of pricing strategies to manage peak demand.

References

Agency, I. E. (2021), *Global EV Outlook 2021: Accelerating ambitions despite the pandemic*, IEA.
URL: *https://www.iea.org/reports/global-ev-outlook-2021*

Bailey, M. R., Brown, D. P., Shaffer, B. C. & Wolak, F. A. (2024), 'Shifting electric vehicle owners to off-peak charging', *NBER Working Paper* (31630).
URL: *https://www.nber.org/papers/w31630*

Bakhtiari, A., Patwary, A. U. Z., Ciari, F., Moeini, A. & Hajebrahimi, A. (2024), 'A comparative analysis of time-based and hybrid pricing models for electric vehicle charging', *Procedia Computer Science* **238**, 757–762.

Chaudhari, K., Kandasamy, N. K., Krishnan, A., Ukil, A. & Gooi, H. B. (2018), 'Agent-based aggregated behavior modeling for electric vehicle charging load', *IEEE Transactions on Industrial Informatics* **15**(2), 856–868.

Chen, T. D., Kockelman, K. M. & Khan, M. (2013), 'Locating electric vehicle charging stations: Parking-based assignment method for seattle, washington', *Transportation research record* **2385**(1), 28–36.

Csiszár, C., Csonka, B., Földes, D., Wirth, E. & Lovas, T. (2019), 'Urban public charging station locating method for electric vehicles based on land use approach', *Journal of Transport Geography* **74**, 173–180.

Dashora, Y., Barnes, J. W., Pillai, R. S., Combs, T. E., Hilliard, M. & Chinthavali, M. S. (2010), 'The phev charging infrastructure planning (pcip) problem', *International Journal of Emerging Electric Power Systems* **11**(2).

E. Seilabi, S., Pourgholamali, M., Miralinaghi, M., Homem de Almeida Correia, G., Li, Z. & Labi, S. (2024), 'Sustainable planning of electric vehicle charging stations: A bi-level optimization framework for reducing vehicular emissions in urban road networks', *Sustainability* **17**(1), 1.

Ester, M., Kriegel, H. & Xu, X. (2022), 'Xgboost: A scalable tree boosting system. in proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (vol, pg 785, 2016)', *GEOGRAPHICAL ANALYSIS* .

Finance, B. N. E. (2020), 'Electric vehicle outlook 2020', *Bloomberg* .
URL: *https://about.bnef.com/electric-vehicle-outlook/*

Frade, I., Ribeiro, A., Gonçalves, G. & Antunes, A. P. (2011), 'Optimal location of charging stations for electric vehicles in a neighborhood in lisbon, portugal', *Transportation Research Record* **2252**(1), 91–98.

Friedman, J. H. (2001), 'Greedy function approximation: a gradient boosting machine', *Annals of statistics* pp. 1189–1232.

Gharbaoui, M., Martini, B., Bruno, R., Valcarenghi, L., Conti, M. & Castoldi, P. (2013), Designing and evaluating activity-based electric vehicle charging in urban areas, *in* '2013 IEEE International Electric Vehicle Conference (IEVC)', IEEE, pp. 1–5.

Gopal, A., Sheppard, C., Waraich, R., Campbell, A. & Pozdnukov, A. (2017), 'Modeling plug-in electric vehicle charging demand with beam, the framework for behavior energy autonomy mobility'.

Hoerl, A. E. & Kennard, R. W. (1970), 'Ridge regression: Biased estimation for nonorthogonal problems', *Technometrics* **12**(1), 55–67.

Khatouri, H., Benamara, T., Breitkopf, P. & Demange, J. (2022), 'Metamodeling techniques for cpu-intensive simulation-based design optimization: a survey', *Advanced Modeling and Simulation in Engineering Sciences* **9**(1), 1.

Kianifar, M. R. & Campean, F. (2020), 'Performance evaluation of metamodelling methods for engineering problems: towards a practitioner guide', *Structural and Multidisciplinary Optimization* **61**(1), 159–186.

Kuby, M. & Lim, S. (2005), 'The flow-refueling location problem for alternative-fuel vehicles', *Socio-Economic Planning Sciences* **39**(2), 125–145.

Kuby, M. & Lim, S. (2007), 'Location of alternative-fuel stations using the flow-refueling location model and dispersion of candidate sites on arcs', *Networks and Spatial Economics* **7**, 129–152.

Li, C., Dong, Z., Chen, G., Zhou, B., Zhang, J. & Yu, X. (2021), 'Data-driven planning of electric vehicle charging infrastructure: A case study of sydney, australia', *IEEE Transactions on Smart Grid* **12**(4), 3289–3304.

Liu, H. X., He, X. & He, B. (2009), 'Method of successive weighted averages (mswa) and self-regulated averaging schemes for solving stochastic user equilibrium problem', *Networks and Spatial Economics* **9**, 485–503.

Liu, Y. S., Tayarani, M. & Gao, H. O. (2022), 'An activity-based travel and charging behavior model for simulating battery electric vehicle charging demand', *Energy* **258**, 124938.

Ma, Z., Liu, X., Wang, W., Yang, S., Yang, Y., Zhao, Y., Xia, H. & Wang, Y. (2024), 'An activity network design and charging facility planning model considering the influence of uncertain activities in a game framework', *World Electric Vehicle Journal* **15**(11), 537.

Nie, Y. M. & Ghamami, M. (2013), 'A corridor-centric approach to planning electric vehicle charging infrastructure', *Transportation Research Part B: Methodological* **57**, 172–190.

Patwary, A. U., Huang, W. & Lo, H. K. (2021), 'Metamodel-based calibration of large-scale multimodal microscopic traffic simulation', *Transportation Research Part C: Emerging Technologies* **124**, 102859.

Razavi, S., Tolson, B. A. & Burn, D. H. (2012), 'Review of surrogate modeling in water resources', *Water Resources Research* **48**(7).

Shuai, C., Ruan, L., Chen, D., Chen, Z., Ouyang, X. & Geng, Z. (2024), 'Location optimization of charging stations for electric vehicles based on heterogeneous factors analysis and improved genetic algorithm', *IEEE Transactions on Transportation Electrification* .

Tibshirani, R. (1996), 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society Series B: Statistical Methodology* **58**(1), 267–288.

W Axhausen, K., Horni, A. & Nagel, K. (2016), *The multi-agent transport simulation MATSim*, Ubiquity Press.

Waraich, R. A. (2013), Agent-based simulation of electric vehicles: design and implementation of a framework, PhD thesis, ETH Zurich.

Wardrop, J. G. & Whitehead, J. I. (1952), 'Correspondence. some theoretical aspects of road traffic research.', *Proceedings of the institution of civil engineers* **1**(5), 767–768.

Zhang, C., Osorio, C. & Flötteröd, G. (2017), 'Efficient calibration techniques for large-scale traffic simulators', *Transportation Research Part B: Methodological* **97**, 214–239.

Zhang, H., Sheppard, C. J., Lipman, T. E., Zeng, T. & Moura, S. J. (2020), 'Charging infrastructure demands of shared-use autonomous electric vehicles in urban areas', *Transportation Research Part D: Transport and Environment* **78**, 102210.

Zhao, A. P., Li, S., Li, Z., Wang, Z., Fei, X., Hu, Z., Alhazmi, M., Yan, X., Wu, C., Lu, S. et al. (2024), 'Electric vehicle charging planning: A complex systems perspective', *IEEE Transactions on Smart Grid* .