Reinforcement Learning Optimization of Placement of Public Vehicle Charging Stations using a Novel Parking Demand Model

Introduction

A modest 1.39% of light-duty vehicles in Connecticut are fully electric vehicles (EVs) or plug-in hybrid electric vehicles (PHEVs), yet they collectively mitigate more than 238 million pounds of carbon dioxide annually compared to their gasoline counterparts, based on estimates from the Department of Energy. The escalating awareness of environmental concerns and government incentives are propelling a swift rise in the adoption of EVs. In tandem with this trend, there is an increase in the demand for an adequate charging infrastructure to support EVs, analogous to the role of gas stations for gasoline vehicles. A critical need emerges for strategically optimizing the placement of charging stations to enhance accessibility, convenience, and overall efficiency in sustainable transportation networks. Traditional optimization algorithms are reliant on scarce and often proprietary travel survey data to obtain potentially comprehensive information on parking behavior; consequently, they face reproducibility challenges due to data format inconsistencies and lack of availability. Addressing this challenge, this project leverages publicly available map details in OpenStreetMap to estimate parking data for a more efficient and reproducible approach. Charging station optimization in Connecticut poses a unique challenge, as the state has yet to be extensively studied for charger placement despite active efforts to encourage EV adoption. The intricacies of the optimization process involve a delicate balance of budget constraints, material and installation costs, user convenience, and charging times, among other factors. In tackling this geographical placement challenge, this project employs reinforcement learning (RL), a methodology that has shown promising capabilities in optimizing charging station locations, particularly in urban settings. Bridgeport, characterized by the highest population density among Connecticut cities, serves as a fitting testbed for implementing the parking demand model within the RL framework.

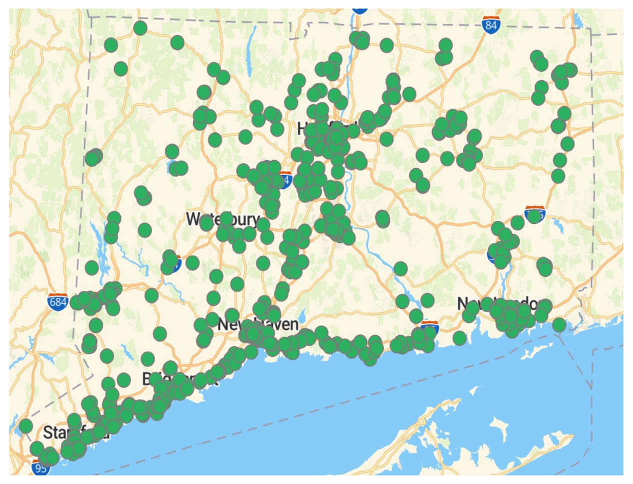

Figure 1: This figure shows the 751 electric charging stations in Connecticut as of January 2024. These stations contain 2169 ports, yet the number of ports in the state would have to triple to 6777 ports, based on EVI-Pro (the US government's EV Infrastructure Projection tool), to support and achieve the state's goal of 150,000 EVs by 2020. *Credit: Department of Energy*

Engineering Goal

This research aims to develop a method to optimize the geographical locations of public charging stations by implementing an innovative demand estimation model with reinforcement learning (RL). This study will then analyze the effectiveness of the model on dense urban street networks versus rural areas. This research will also investigate the impact of at-home private charging on the RL framework.

Data for Reinforcement Learning

The data for this project comes entirely from the free global maps of OpenStreetMap (OSM) and Open Charge Map (OCM), downloaded through their respective Application Programming Interfaces (APIs).

Street network

Street network data was procured from OpenStreetMap (OSM) using a central address and radius as input, forming the foundational framework for the reinforcement learning (RL) environment. Subsequently, intersections, called "nodes," on the street network serve as the "gameboard" where the AI agent plays the game of placing the optimal plan of stations on the nodes. While stations would not always be on intersections in practice, the distance between their real-world location and the node are very small, especially in urban settings where nodes are typically tightly packed.

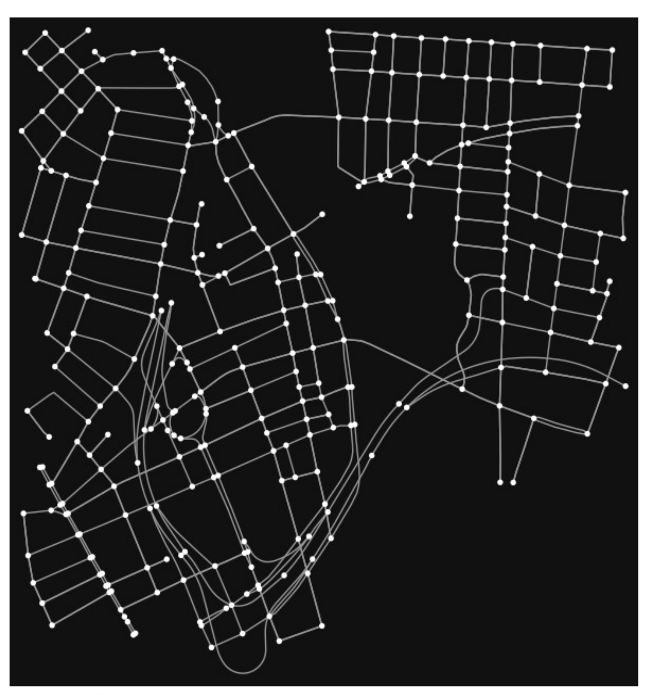

Figure 2: The street network of Bridgeport with a radius of 1 km. Each white dot represents one of 312 nodes in the region.

Features

Features on OSM are tagged with both a name and a specific attribute. The names represent overarching categories such as 'amenity' and 'building', while the attributes denote more specific subcategories. For instance, amenities may be further characterized with attributes like 'restaurant' or 'hospital', and buildings with descriptors such as 'hotel' or 'bungalow'. This research takes interest to places where individuals would commonly park their vehicles, so features irrelevant to the problem (like 'bench') were excluded from acquisition. This data becomes the basis for my novel parking demand model. During reproduction of this work, the names and attributes of desired features to download can be easily adjusted if needed.

Existing charging infrastructure

Acquisition of data from Open Charge Map (OCM) encompassed details on the existing charging infrastructure, including charger locations (longitude, latitude) and the count of charging ports at each station (reflecting capacity for EVs). The assessment of the current charging infrastructure serves as the initial condition of the RL model as well as a benchmark for evaluating its efficacy. Charging stations were assigned to their closest node because the RL algorithm deals with a node graph, not geospatial latitudes and longitudes like the raw data from OCM. Figure 3 shows the street network of Bridgeport along with its four existing charging stations represented by orange dots.

Figure 3: Existing charging infrastructure of Bridgeport overlayed onto the street network.

Unless otherwise noted, all images and graphs were taken or created by the student researcher.

Estimation of Parking Demand using Publicly Available Feature Data in OpenStreetMap

Gasoline cars can refuel within minutes, allowing drivers to stop at gas stations en route. Electric charging, on the other hand, takes a minimum of 20 minutes, and it therefore necessitates strategic placement of charging stations near parking locations. OpenStreetMap features signify these final endpoints, with certain amenities or buildings naturally attracting higher parking demand. For instance, supermarkets are intuitively expected to have higher parking demand than cafes. To reflect these variations, the model uses a flexible weighting scale, assigning each feature a 0 to 5 weight, indicating its influence on parking demand. Researchers can adjust this scale, facilitating nuanced evaluation and prioritization of features based on their predicted significance for parking demand.

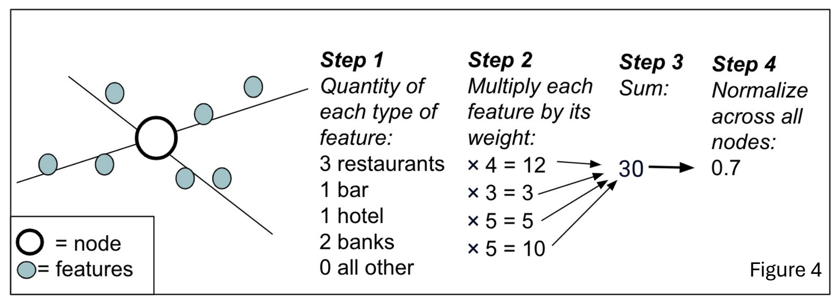

Figure 4

Figure 4 illustrates an example node of how feature data is integrated into the RL environment. First, the model finds the closest node to each feature. Then, it determines the count of each type of feature at the respective node (Step 1). Next, it factors in the weighted influence (Step 2). Subsequently, it aggregates the total parking demand at the node (Step 3). To ensure uniformity and comparability across nodes, the total demand per node is normalized from 0 to 1 (Step 4). This allows the RL model to consistently interpret and respond to the demand, irrespective of the original scale used. A heat map of the normalized demand per node is shown in Figure 6.

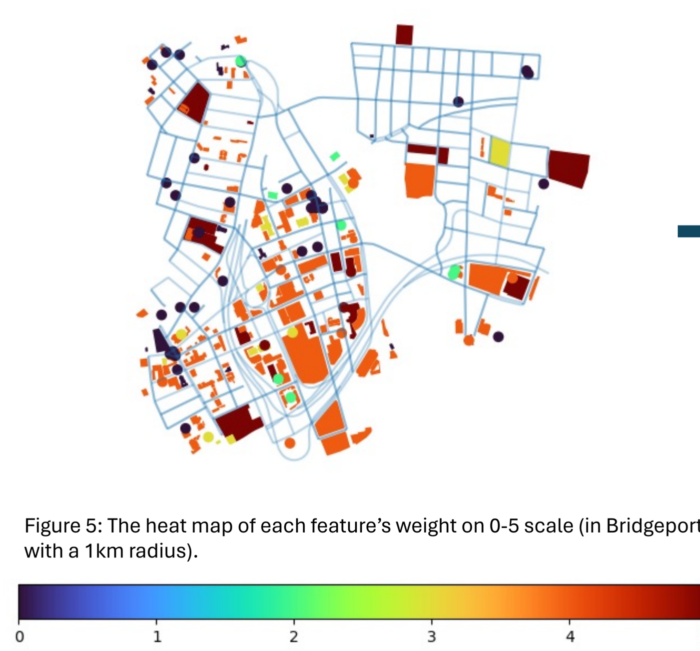

Figure 5: The heat map of each feature's weight on 0-5 scale (in Bridgeport with a 1km radius).

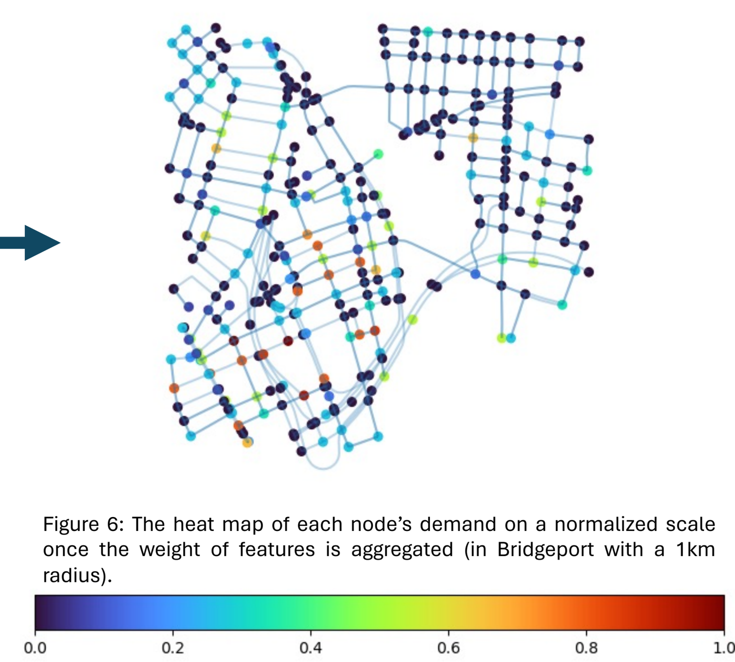

Figure 6: The heat map of each node's demand on a normalized scale once the weight of features is aggregated (in Bridgeport with a 1km radius).

Placement of Chargers by Reinforcement Learning (PCRL) Model

In this project, Reinforcement Learning (RL) optimally places charging stations using strategic node selection. Previous research (Figure 7) has demonstrated that PCRL (Placement of Chargers by Reinforcement Learning) surpassed baselines and existing infrastructure, by forming charger miniclusters for wider city distribution. The baseline Charger-Based Greedy approach prioritized low charging times at the expense of costly chargers. Bounding&Optimising+ clustered stations where node density was highest, inconvenient for distant EV owners. In further evaluation, the PCRL demonstrated substantial improvement from existing charging infrastructure by minimizing the inconvenience associated with the maximum waiting time, the duration spent waiting at an occupied charging station before a charger becomes available. The PCRL diminished the maximum waiting time by 92.89%, whereas the best baseline only achieved 59.57% reduction.

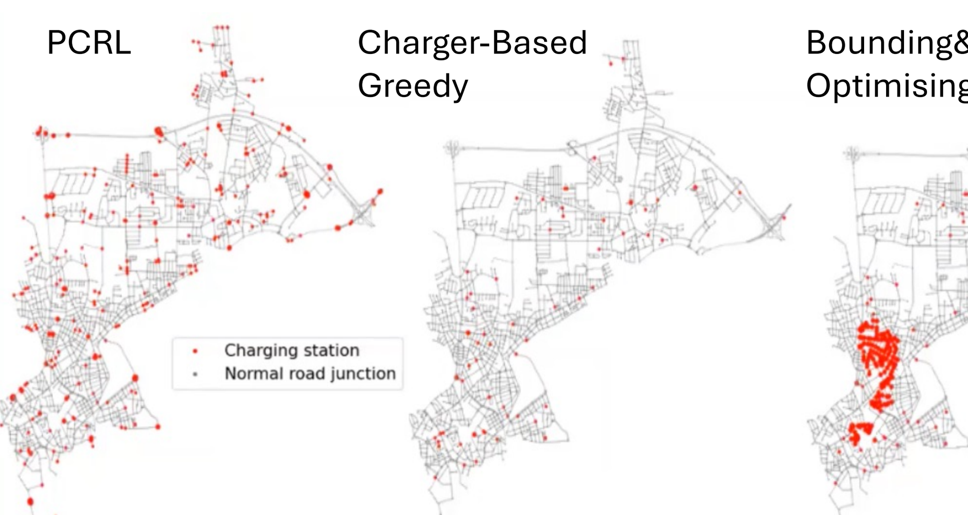

Figure 7: Charger plans of Hannover *Credit: modified from von Wahl et al.*

$$p^* = \overbrace{argmax_{p\in P}\{\lambda \cdot benefit(p) - (1-\lambda)\cdot cost(p)\}}^{score(p)} \quad (1)$$

s.t.

$$\sum_{s\in p^*} fee(s) \leq B \quad (2)$$

$$\rho(s) < 1 \ \forall s \in p^* \quad (3)$$

$$\sum_{i=1}^{m} t_i \leq K \ \forall s \in p^* \quad (4)$$

Figure 8: The optimizations and constraints within the RL model. Equation 1 articulates the goal: determining the optimal plan $p*$ with the highest score, a combination of benefit and cost functions. The weighting parameter $\lambda \in[0,1]$ balances the trade-off between benefit and cost. The benefit function accounts for station capacity, where stations with more chargers yield higher benefits. Conversely, proximity to an existing charging station diminishes station benefit due to partially satisfied parking demands. The cost function evaluates user discomfort through travel time, charging time, and waiting time. Constraints, outlined in Equations 2-4, ensure that the total charger installation cost(s) remains below the budget B, the number of ports per station is limited to K, and waiting time is both well-defined and positive. *Credit: modified from von Wahl et al*

The RL model integrates data concerning the road network, existing charging infrastructure, and parking demand into its environment (Figure 9). The agent, driven to enhance the charging plan, continuously takes actions within this environment. After each action, the agent gathers observations regarding the current charging plan and the evolving environment. A reward is then assigned based on the score, a composite of benefit and cost, following each action. This iterative process of taking actions, observing, and receiving rewards is repeated through multiple episodes. Episodes conclude under specific conditions: when the budget B is exhausted, the maximum number of chargers is deployed (an improbable scenario wherein every node is covered with the maximum number K of charging ports at each station), or when the maximum number of iterations (i_{max}) is reached. This structured framework ensures a systematic and adaptive approach to optimizing the charging infrastructure.

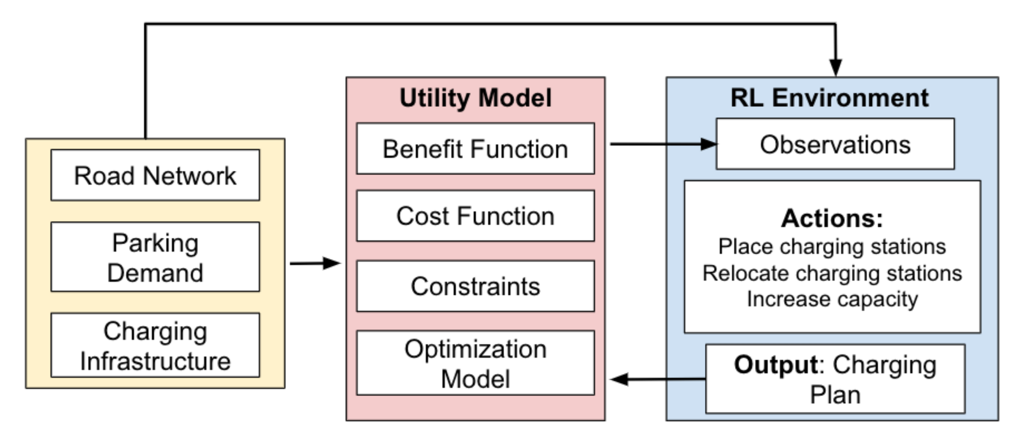

Figure 9: Design process of RL model

Implementation of Real-World Data into RL Model

The RL model was implemented in the Bridgeport region with a 1km radius, encompassing an area of approximately 1.2 square miles. To facilitate training, this project gives values to parameters within the optimization and constraints of the model (see Figure 8 above). The weighting parameter $\lambda\in[0,1]$ has the chosen value of 0.5 to equalize the importance between the benefit and cost functions (Equation 1). The constraints in this study allow for a maximum of K=10 chargers per station (Equation 4) while considering three types of chargers with varying outputs and installation costs: $4,000 for a 15 kW output, $28,000 for a 50 kW output, and $75,000 for a 150 kW output. The cumulative installation costs cannot exceed the budget set to B = $15 million, and reaching this budget concludes an episode. While tailored for the Bridgeport study, these values can be adjusted based on specific charger types and regional costs chosen by the researcher. The maximum number of iterations per episode is set to half the total number of nodes, ensuring a thorough exploration of the charging infrastructure landscape in the region. After training for 200,000 episodes, the agent finalized the optimal plan (Figure 9). The visualization reaffirms previous findings that RL strategically forms miniclusters of chargers across the city to achieve a maximum score and ensure the most convenient layout.

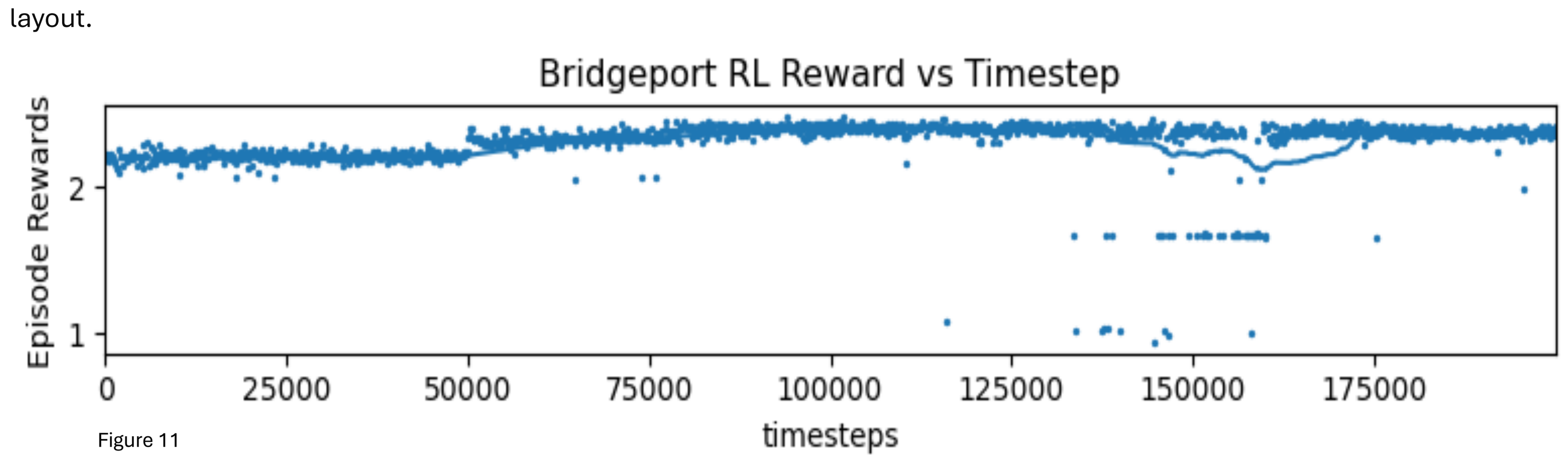

Figure 11

Figure 10: Optimal charging plan for Bridgeport

The chart depicting the score over the number of episodes or "timesteps" (Figure 11) demonstrates the RL agent's dynamic interplay between exploitation and exploration. Initially, the agent explores various strategies to discover optimal charging configurations. As training progresses, a shift towards exploitation occurs, refining the charging plan based on gained insights. The upward trend in the score signifies continuous improvement. Despite advanced training, occasional points significantly below the trend indicate intermittent exploration phases. This balance is essential for the RL model to evolve towards an increasingly optimal charging infrastructure configuration.

Urban vs Rural

The rural area of Warren, Connecticut, was selected for comparison with urban Bridgeport. It underwent the same data collection process with consistent parameters, recognizing that adjustments will be necessary for real-world application in specific regions. Figure 12 illustrates the optimal plan for Warren with a 5 km radius after 200,000 training episodes. The large number of chargers compared to the number of nodes in the plan can be attributed to the use of the same budget (B = $15 million), allowing for a greater allocation of chargers.

Figure 12

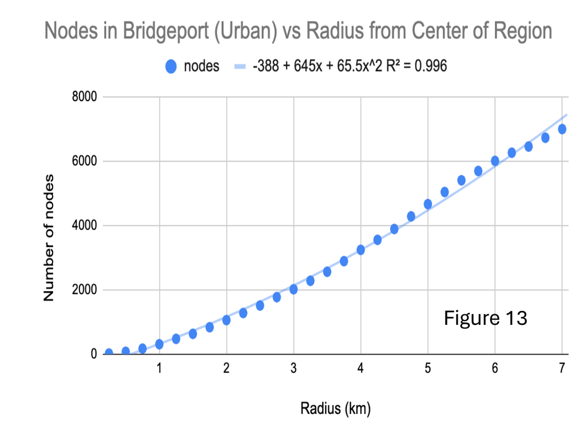

Figure 13

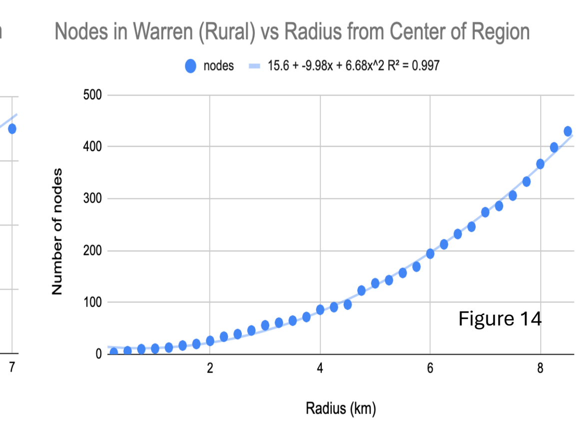

Figure 14

Figures 13 and 14 demonstrate that the number of nodes rises quadratically in both the urban and rural settings as the km radius expands. It is crucial to note that the y-axis scale for the number of nodes is significantly higher for Bridgeport, the dense urban network. Extended training times and heightened computational demands are observed because there are an increased number of charging plan combinations and higher iterations per episode (equal to half the number of nodes) when using an urban network or a larger radius.

The RL model assumes that while stations may not always align with intersections in practice, the small real-world distance between their locations and nodes provides an excellent approximation. However, this method's transferability to real-life situations is less effective in rural environments, where the large distances between nodes impact the applicability of using nodes as an environment and approximation. Despite longer training times in urban settings, RL is a better method for urban environments than rural.

Private Home Charging

Many EV owners prefer at-home charging for its overnight convenience, reducing reliance on public charging stations. The graph (Figure 15) illustrates the number of charging ports needed to support the current EV population in Connecticut, categorized into Level 2 and direct current (DC) chargers, depending on the percentage of EVs utilizing at-home private chargers. Level 2 ports require several hours of charging and have an output of about 6kW, whereas DC chargers can charge to 80% from empty in 20-30 minutes with an output of around 50 kW, albeit being more expensive to install. Each charger type shows a very well-fit linear regression: state-wide, 27.3 less Level 2 charging ports and 3.35 less DC ports are needed to support the EV population per 1% increase in the percentage of EV owners who charge at home.

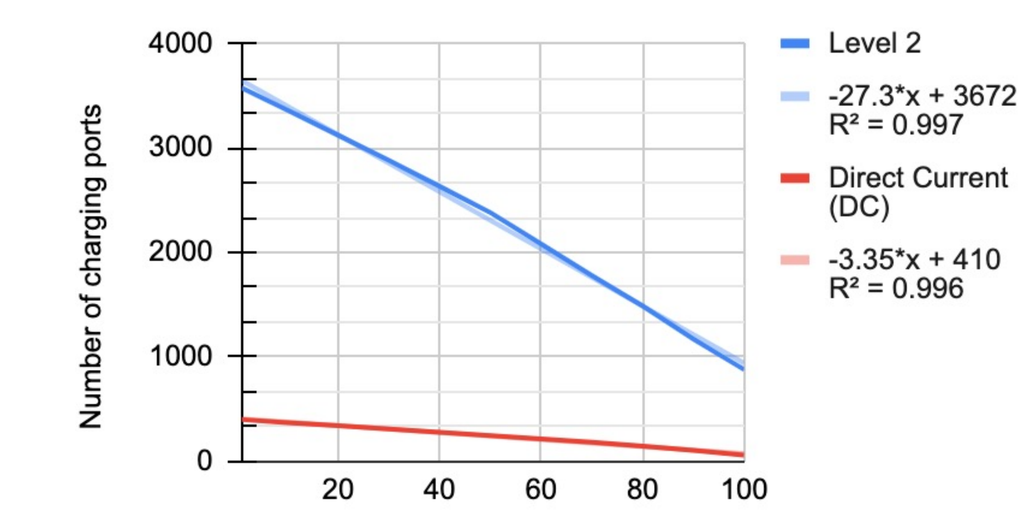

Figure 15

However, the importance of public chargers persists, especially in urban settings and low-income areas where accessibility challenges may hinder at-home charging. At-home charging options allows for a more efficient utilization of charging infrastructure budgets for owners who depend on it the most. This is relevant to the RL model since budget, which decreases with demand of chargers, is the main restrictor in how many stations and ports the agent is able to place on the street network of nodes.

Discussion and Conclusions

The methodology employed in this project underscores the simplicity and accessibility of the implemented charging station optimization model. The process merely requires the input of the address and radius for the target region, offering a user-friendly approach for researchers seeking to enhance any charging infrastructure. This research revealed that increasing the radius results in a quadratic increase in the number of nodes, impacting the training time for the reinforcement learning (RL) agent. This result highlights the necessity for a balance between the required geographical coverage and the processing needs.

The parking demand estimation method uses exclusively publicly available data sources, thus serving as a versatile tool. It is flexible in that researchers can adjust the relative importance of features to align with the unique needs of their target region. Importantly, it extends the applicability of charging infrastructure optimization to virtually anywhere globally, compared to the previous availability with the few regions having comprehensive travel data. Crucially, it broadens the scope of charging infrastructure optimization to virtually any region worldwide, surpassing the limited applicability seen in only those with comprehensive travel data.

Future Research

Future research will include:

- Consideration of EVs commuting into the region of interest
- Impact of the RL output plan on the electrical grid