A Stochastic Flow-Capturing Model to Optimize the Location of Fast-Charging Stations with Uncertain Electric Vehicle Flows

Fei Wu, Ramteen Sioshansi*

Integrated Systems Engineering Department, The Ohio State University, Columbus, Ohio, United States of America
Phone: +1-614-292-3932

Abstract

We develop a model to optimize the location of public fast charging stations for electric vehicles (EVs). A difficulty in planning the placement of charging stations is uncertainty in where EV charging demands appear. For this reason, we use a stochastic flow-capturing location model (SFCLM). A sample-average approximation method and an averaged two-replication procedure are used to solve the problem and estimate the solution quality.

We demonstrate the use of the SFCLM using a Central-Ohio based case study. We find that most of the stations built are concentrated around the urban core of the region. As the number of stations built increases, some appear on the outskirts of the region to provide an extended charging network. We find that the sets of optimal charging station locations as a function of the number of stations built are approximately nested. We demonstrate the benefits of the charging-station network in terms of how many EVs are able to complete their daily trips by charging midday—six public charging stations allow at least 60% of EVs that would otherwise not be able to complete their daily tours without the stations to do so. We finally compare the SFCLM to a deterministic model, in which EV flows are set equal to their expected values. We show that if a limited number of charging stations are to be built, the SFCLM outperforms the deterministic model. As the number of stations to be built increases, the SFCLM and deterministic model select very similar station locations.

Keywords: Electric vehicles, infrastructure location, stochastic, optimization

1. Introduction

Electric vehicles (EVs) are one possible solution to improving transportation energy consumption efficiency. An issue that limits consumer adoption of EVs is 'range anxiety.' Range anxiety refers to drivers' concerns with being stranded with a discharged EV battery and the associated delays to their journeys due to long recharging times. Chéron and Zins (1997); Eberle and von Helmolt (2010); Franke and Krems (2013) find that many consumers are unwilling to purchase EVs if a 'safety net' of charging stations that covers their anticipated driving needs is not available.

The EV charging infrastructure in many parts of the world is not sufficiently well developed to alleviate range anxiety, calling for added infrastructure development. An important issue in developing such an infrastructure is the type of charging technology used. EV charging technologies can be broadly broken into two categories. The first includes what are known as Level-1 and -2 charging technologies. These technologies deliver low-power charging (typically less than 6.6 kW), meaning that EV recharging times are long. For instance, a 2016 Nissan Leaf with a 30 kWh battery has a tested driving range of up to 107 miles.[1]

*Corresponding author
Email addresses: wu.1557@osu.edu (Fei Wu), sioshansi.1@osu.edu (Ramteen Sioshansi)
[1] http://www.nissanusa.com/electric-cars/leaf

Such a vehicle could take as many as 15 hours to fully recharge using Level-1 technology. Thus, Level-1 and -2 technologies are amenable to use for charging overnight or at a destination where an EV is parked for a prolonged period of time (*e.g.*, a workplace in the case of a commuter's vehicle).

The other category includes DC fast-charging technologies, which deliver high-power charging. These technologies typically deliver between 50 kW and 200 kW of charging power and could fully recharge a Nissan Leaf in as little as seven minutes.[2] Given the substantively shorter charging times, fast-charging technologies could be used for charging *en route* from origin to destination, in addition to charging at a vehicle destination. Given these fundamental differences in the use of slow- and fast-charging technologies, different techniques are appropriate to optimize the location of these two charging station types.

EV charging station location belongs to the broader class of refueling-infrastructure problems. The existing literature on refueling-infrastructure problems can be categorized as using two modeling techniques: node-serving and flow-capturing. These two methods rely on models with a similar underlying mathematical structure. In node-serving models, refueling requirements are treated as being demands at specific spatial locations (nodes). The objective of such models can include maximizing demands that are covered or minimizing refueling distances. Because of these properties, the node-serving method is mostly applied to home-based refueling-infrastructure siting. The node-serving method is used by Frade et al. (2011); Xi et al. (2013). Frade et al. (2011) introduce a maximal-covering model to locate Level-1 and -2 EV charging stations in Lisbon, Portugal. Their model forecasts charging demand from each region and maximizes the amount of EV charging demand that is satisfied by the network of stations based on a maximum charging-distance tolerance. Xi et al. (2013) develop a simulation-optimization model to locate Level-1 and -2 public charging stations and consider EV charging times. Their work treats demands as occurring at specific spatial locations when they are parked at a workplace, university, or retail shopping location. Moreover, they pay particular attention to modeling the queuing behavior of EVs at charging stations, due to the prolonged charging durations involved. One novelty of their work is that their model determines the number of chargers in each station and simulates the EV charging and queuing processes. Nicholas et al. (2004); Lin et al. (2008) introduce a median model to optimize the refueling station using convenience level. A recently applied node-serving based median method to plan refueling infrastructure is the c-means clustering approach, which is proposed by Shi and Zheng (2014). In this method, EV charging demands are represented by coordinates in a two-dimensional plane. The demands are then clustered using the fuzzy c-means method and the refueling stations are assumed to be placed at the cluster centers.

Flow-capturing models assume that the refueling demands are from vehicle flows and that the refuelings occur *en route* from origin to destination. The work of Hodgson (1990) is one of the earliest examples of a flow-capturing location model, which maximizes vehicle flows served by the network. Kuby and Lim (2005) extend this work by incorporating vehicle-range parameters and introduce additional binary variables to indicate combinations of location candidates. This model is further extended by Kim and Kuby (2012) to take shortest-time deviation into account. Tan and Lin (2014) extend the flow-capturing location model by considering uncertain vehicle flows. Wang and Lin (2009) propose a flow-capturing model to minimize the total cost of locating alternative-fuel stations. Their model constrains the facility locations to ensure that they cover the alternative-fuel vehicle flows along their shortest paths from origin to destination. Li et al. (2016) propose a dynamic multiperiod multipath refueling-location model to optimize an EV charging network. Their model focuses on EV use for intercity trips and minimizes the cost of installing new stations and relocating existing stations. He et al. (2015) introduce a bi-level model to locate EV charging stations while minimizing social cost under budget constraints. Dong et al. (2016) propose a flow-capturing based median model to design the layout of EV fast-charging infrastructure. This method is based on a shared nearest neighbor clustering algorithm for station allocation and an optimization model for station sizing. One difficulty in the use of flow-capturing models is that they are normally formulated as computationally complex mixed integer optimization models. He et al. (2015) address this computational issue using a genetic algorithm while Li et al. (2016) use a heuristic-based genetic algorithm.

[2]The 2016 Nissan Leaf sold in the United States is equipped with a SAE J1772 charging port. The DC Level-3 charging standard SAE J1772 delivers up to 240 kW.

As noted above, fast-charging infrastructure is much more amenable to use *en route* from origin to destination than slower Level-1 and -2 technologies. Because the charging demands occur as vehicles are *en route*, flow-capturing methods are more appropriate for fast-charging station location planning than node-serving. This is because node-serving methods assume that charging demands occur at fixed locations. Moreover, Kuby et al. (2013) find that 88.2% of drivers of alternative-fuel vehicles refuel their cars *en route*. In this work, we seek an optimization-based approach to station-location planning that can provide guaranteed bounds on solution quality. This is because refueling infrastructure is a capital-intensive and long-lived investment. These properties mean that charging stations must be placed judiciously to ensure that their upfront costs are recovered over their lifetimes.

EV charging station-location optimization is complicated by uncertainties in where future EV-charging demands will be geographically located within the study region. One approach to overcoming this issue, which we take, is to bootstrap EV driving patterns from tour-record data. Regional planning commissions or similar organizations typically have access to or maintain such datasets, which are used for economic and infrastructure forecasting and planning. An important limitation of tour-record data is that they do not typically specify all of the information needed for EV modeling. For instance, tour records often report the origin and destination of a vehicle trip, but not the route taken. We use a shortest-path model to translate tour records into vehicle paths for purposes of EV-charging demand modeling.

Another important limitation of existing flow-capturing models is that they do not adequately capture uncertainty in where EV-charging demands appear in the study region. Although Tan and Lin (2014) use a stochastic flow-capturing model in their analysis, they assume a small scenario tree with on the order of six EV-flow scenarios. Our solution method dynamically determines the scenario tree size to provide a desired tolerance on the error in estimating the problem's objective-function value. As a result, we solve problems with on the order of 1000 scenarios. Our method also allows us to provide statistical bounds on the objective-function value. As such, our model and solution approach can be said to provide charging station locations that are more robust to uncertainties in EV-charging demand.

The remainder of this paper is organized as follows. In Section 2 we detail the approach taken to modeling fast EV charging station location and the input data that our method uses. We model EV flows by bootstrapping from tour-record data and use a shortest-path model to translate tour records into complete EV trips. EV charging station locations are optimized using a two-stage stochastic flow-capturing location model (SFCLM). The first stage of the problem is when the charging stations are placed and the second stage determines which of the EV flows, which are unknown in the first stage, are 'captured' by the stations built. The first-stage objective is to maximize the expected number of EVs that can be captured by the stations that are built. Because the first-stage objective function cannot be directly computed, we use a sample-average approximation (SAA) method and an averaged two-replication procedure (A2RP) to solve the SFCLM and provide statistical bounds on the quality of the solution found.

Section 3 introduces the case study, based on Central Ohio, to which we apply the SFCLM. All of the inputs used to build the SFCLM are obtained from the Mid-Ohio Regional Planning Commission (MORPC). Section 4 summarizes our case study results. We find that the optimal station locations are fairly robust in the sense that as more stations are built, the optimal locations of the starting stations do not change. We then examine the benefit of the charging-station network in terms of how many EVs are able to complete their trips. We find that a six-station network can allow at least 60% of EVs that would not be able to complete their typical daily trips without the public charging stations to do so with the stations built. We also conduct simulations to estimate how many chargers would need to be installed at each charging station to allow their use without EVs having to queue. We find that installing three chargers at each station in a six-station network would result in at least an 80% probability that EVs would not have to queue to use a charger. We finally compare the SFCLM to a deterministic version of the model. We show that if there are a limited number of charging stations to be built, the SFCLM outperforms the deterministic model in terms of the expected number of EVs captured and should be used in place of a deterministic model. As the number of stations to be built increases, the SFCLM and the deterministic model choose very similar station locations. Thus, the deterministic model could be used as the SFCLM provides little incremental value and greater computational cost. Section 5 concludes.

2. Modeling and Solution Approach

Our model determines the location of a fixed number of fast EV-charging stations within a given planning region to maximize the expected volume of EVs that can be served by the station network. Our model accounts for uncertainty in EV flows, given that the charging infrastructure is being built in anticipation of future EV adoption. Moreover, our work considers limited EV range and drivers' behavior in deviating from their intended shortest paths from origin to destination to use a charging station.

Section 2.1 details the necessary inputs to our SFCLM and gives its formulation. We then describe in Section 2.2 the data sets that are readily available from most regional planning commissions. We also detail the assumptions and data processing done to convert these data into the necessary inputs to our SFCLM. Section 2.3 summarizes the challenges in solving the SFCLM and the method that we employ to obtain a near-optimal solution and statistical bounds on its quality.

2.1. Stochastic Flow-Capturing Location Model Formulation and Required Data

The key inputs to an SFCLM are: (i) EV-flow volumes on different trip chains on the road network, (ii) a set of candidate locations where charging stations can be built, and (iii) an assumption regarding how EV drivers would use charging stations that are built A trip chain is a sequence of trips constituting a daily tour for a vehicle. Thus, our definition of a trip chain is similar to that introduced in the work of Kang and Recker (2009). However, our trip chain does not include timestamps. As noted before, we formulate this problem as a two-stage stochastic integer linear optimization problem. The locations of a fixed number of charging stations are determined in the first stage. This is done with the objective of maximizing the expected number of EVs that can be 'captured' by the charging station network. 'Capturing' an EV means that its driver would be willing to deviate from his or her original intended route to one of the charging stations built, if the vehicle battery's state of charge (SoC) is sufficiently low. The second stage of the model determines which EVs can be captured by the charging stations built. The scenarios modeled in the second stage represent different possible volumes of vehicle flows along each trip chain. The scenarios are intended to capture uncertainty in where EV charging demands may appear when making charging station-placement decisions.

To formulate the SFCLM, we first introduce the sets and parameters:

Q: set of trip chain indices on the road network;

K: set of all candidate charging station locations on the road network;

N_q: set of candidate charging station locations that are capable of capturing EVs traveling on trip chain q;

p: maximum number of stations to be built in the region; and

$\tilde{f}_q$: number of EVs traveling on trip chain q.

Note that $\tilde{f}_q$ is a second-stage parameter, which is random in the first stage. We let $\tilde{f} = \left(\tilde{f}_1, \tilde{f}_2, \cdots, \tilde{f}_{|Q|}\right)$ denote a vector of all of the $\tilde{f}_q$'s. We also define the decision variables:

x_k: binary variable that equals 1 if a station is placed at location k and equals 0 otherwise; and

y_q: binary variable that equals 1 if EVs on trip chain q are 'captured' by a charging station and equals 0 otherwise.

The SFCLM is then formulated as:

$$\max_{x} \quad \mathbb{E}\left[g(x, \tilde{f})\right] \tag{1}$$

$$\text{s.t.} \quad \sum_{k \in K} x_k \leq p; \tag{2}$$

$$x_k \in \{0, 1\}, \qquad \forall k \in K; \tag{3}$$

where:

$$g(x, \tilde{f}) = \max_{y} \quad \sum_{q \in Q} \tilde{f}_q y_q \tag{4}$$

$$\text{s.t.} \quad \sum_{k \in N_q} x_k \geq y_q, \qquad \forall q \in Q; \tag{5}$$

$$y_q \in \{0, 1\}, \qquad \forall q \in Q. \tag{6}$$

The first-stage objective function (1) maximizes the expected volume of EVs that can be captured by the charging station network. This objective is implicitly defined in terms of the optimal objective-function value of the second-stage problem, which is given by (4)–(6). The first-stage problem includes constraint set (2), which limits the number of stations built, and integrality constraint set (3).

The second-stage objective function (4) maximizes the volume of EVs that can be captured by the charging station network under a particular scenario of EV flows. The different EV flow-scenarios are defined by the second-stage parameter, $\tilde{f}_q$, which is random in the first stage. As noted before, $g(x, \tilde{f})$ is defined as the optimal second-stage objective value and defines the first-stage objective function. The second-stage problem includes constraint set (5), which captures the spatial relationship between the candidate charging station locations and EV flows. This constraint set says that EVs traveling along trip chain q can only be captured if at least one charging station in the set N_q is built. Constraint set (6) imposes integrality on the second-stage EV flow-capture variables.

The SFCLM has relatively complete recourse. This means that for every feasible first-stage solution there exists at least one feasible second-stage solution. Relatively complete recourse is trivially easy to show, because $y = 0$ is always feasible in the second-stage problem. This result implies that $g(x, \tilde{f}) > -\infty$ for any realization of $\tilde{f}$. Because $g(x, \cdot)$ is measurable, we have that:

$$\mathbb{E}\left[g(x, \tilde{f})\right] > -\infty,$$

ensuring that the first-stage objective function is well defined.

2.2. Input Data and Data Processing

Regional planning commissions typically have two datasets that are valuable inputs to our SFCLM. The first is a graph model of the road network, which captures all of the major intersections and road segments. In some cases, vehicle-speed data for road segments are also available. The second is vehicle-use data, in the form of tour records. Tour-record data provide a detailed accounting of typical daily vehicle usage patterns. Tour-record data normally specify the origin and destination of, mode of transportation (*e.g.*, privately owned vehicle, public transportation, or ridesharing) used in, and departure and return time of each trip. Trips taken by privately owned vehicles, which are the ones we focus on in our analysis, are associated with vehicle identifiers. Combining all of the trips assigned to a particular vehicle identifier allows the full sequence of trips taken with that vehicle during a typical day to be determined.

We model EV trips by bootstrapping from the tour-record data. This is done by randomly sampling which vehicles in the tour-record data are EVs and which are not. In our analysis, we apply the assumed penetration rate of EVs uniformly across the vehicles in the tour-record data. More sophisticated sampling methods can be used, however. For instance, Xi et al. (2013) randomly sample EVs from tour-record data, but account for the effects of socioeconomic and demographic factors on EV adoption. They do this by adapting a regression model to determine the relative propensity of people living in different areas of their study region to adopt EVs. They then apply a geographically varying penetration rate to different areas of the study region to sample which vehicles are EVs. Alternate approaches could use spatial regression models to capture direct spatial effects on clustering of EV ownership.

Once the vehicles in the tour records are sampled to determine which subset are EVs, $\tilde{f}$, the EV flows between different trip chains, can be computed. An important limitation of tour-record data, however, is that they do not specify the route taken by a vehicle to get from origin to destination. We associate driving routes to each vehicle trip in the tour-record data by employing the following assumption that each driver takes the shortest path from origin to destination.

Assumption 1 (Shortest-Path)**.** *Each vehicle drives along a shortest path on each trip in a tour.*

Assumption 1 allows us to determine the route of each EV traveling on a trip chain by applying a shortest-path model to the graph representation of the road network. If vehicle-speed data for the road segments are available, the shortest-path problem can be formulated to provide a minimum-time path.

In addition to modeling the path that an EV takes from its origin to its destination, we must determine EV drivers' range-anxiety, a threshold SoC, below which an EV driver tends to recharge his or her vehicle. Moreover, we must also determine what charging stations each EV can be captured by on each trip. We do so by appealing to the following additional range-anxiety, *en-route* charging, and charging-deviation assumptions.

Assumption 2 (Range-Anxiety Threshold)**.** *An EV driver will seek to recharge his or her vehicle if its SoC falls below τ%.*

Assumption 2 suggests that only long-distance EV tours are considered, if we assume that each EV begins its daily tour with a fully charged battery. That is, the model only includes EVs with a sufficiently long daily tour for the battery SoC to fall below τ% in the course of the day. For other EVs, their daily tours are too short for their drivers to seek to use a charging station.

Assumption 3 (*En-Route* Recharging)**.** *Only EVs with daily tours that require a maximum of one battery recharging per day are considered.*

Assumption 3 is based on the fact that very few EVs would reasonably recharge more than once per day for daily commuting. The National Household Travel Survey reports that the longest tract-average commute distance in Ohio is 75.05 miles, which is within the range given by a single daily EV recharging. MORPC tour-record data for Central Ohio show that only 0.008% of vehicles have daily tours that exceed the range offered with one daily recharging. Thus, an extremely small number of EVs would require multiple refueling stops. At the same time, modeling multiple refueling stops (in addition to our dealing with stochastic tours) would introduce modeling and computational complexities that are unnecessary for this initial work. Extending our methodology to include longer tours that require multiple refueling stops is left for future research.

Assumption 4 (Charging Deviation)**.** *An EV can be captured by a charging station on a trip if it passes within a fixed r km radius of the charging station while on its path from its origin to destination.*

Assumption 4 says that an EV driver is willing to use an EV-charging station that is a fixed maximum distance 'out of the way' from his or her planned path from origin to destination. Otherwise, if a charging station is too far out of the way, the driver would not use it due to the inherent inconvenience of doing so. Kim and Kuby (2012) calculate the shortest time-based deviation that a driver is willing to incur to refuel a vehicle and use an exponentially decaying capturing rate in their modeling. We opt to simply use a fixed maximum distance and analyze the effect of driver deviation tolerance on model results. This is because the time complexity of calculating the shortest time-based deviation distance has exponential order with respect to the number of trip chains and the candidate pool. Li and Huang (2014) also find that the computational effort of solving multipath refueling location models exactly is exponentially increasing. Hence, a capturing circle with a fixed maximum distance is accurate enough to estimate the allowable deviation distance, especially in a metropolitan area and in light of the computational difficulties of the alternatives.

For notational convenience, we define $\mathcal{L}(q)$ as the length of trip chain q and $\mathcal{R}$ as the EV range. We also define $\mathcal{L}_k^O(q)$ as the shortest-time distance between the origin of trip chain q and the first intersection of the route and candidate location k's 'capturing' area. Combining Assumptions 1–4 with a road-network graph, Algorithm 1 determines N_q, the set of candidate charging station locations that can capture EVs traveling along different trip chains. Step 5 ensures that candidate location k can serve trip chain q if the driver has 'range anxiety' because of low SoC. Step 11 guarantees that trip chain q is able to be completed with *en-route* recharging at station candidate k.

Algorithm 1 Create N_q

for $q \in Q$ **do**
  $N_q \leftarrow \emptyset$
  if $\mathcal{L}(q) \leq \mathcal{R}$ **then**
    for $k \in K$ **do**
      if $\mathcal{L}_k^O(q)/\mathcal{L}(q) \in [\tau/100, 1]$ **then**
        $N_q \leftarrow N_q \cup \{k\}$
      end if
    end for
  else if $\mathcal{R} < \mathcal{L}(q) \leq 2\mathcal{R}$ **then**
    for $k \in K$ **do**
      if $\mathcal{L}_k^O(q) \in [\max\{\tau\mathcal{L}(q)/100, \mathcal{L}(q) - \mathcal{R}\}, \mathcal{R}]$ **then**
        $N_q \leftarrow N_q \cup \{k\}$
      end if
    end for
  end if
end for

2.3. Solution Technique

The SFCLM requires the second-stage scenarios, which are unknown when making first-stage charging station-placement decisions, representing possible EV flows. The first-stage objective function is also implicitly defined by the second-stage scenarios. As noted above, the input data used to model second-stage EV flows are generated by bootstrapping from the tour-record data. This amounts to generating numerous scenarios, in which different vehicles represented in the tour-record data are EVs. The number of such scenarios is exponential in the size of the tour-record data, however. Thus, the SFCLM would be intractable if every possible second-stage scenario is included in the model. For instance, the tour-record data for the Central Ohio case study that we introduce in Section 3 has about 1.3 million personal vehicles. This would yield approximately 9.9×10^{391338} second-stage scenarios if all possible EV/non-EV combinations are modeled.

We propose an SAA method to deal with this model intractability. Broadly speaking, an SAA method rewrites the original SFCLM and approximates the first-stage objective function by using a small subset of randomly generated second-stage problems. The second-stage problems are generated by random sampling and the sample size is dynamically determined to achieve a desired tolerance on the error in estimating the first-stage objective function.

To formally define the SAA, we first define $f^1, f^2, \ldots, f^M$ as M random samples of the random vector, $\tilde{f}$. Using these M realizations, the sample average function is defined as:

$$\bar{g}_M(x) = \frac{1}{M} \sum_{m=1}^{M} g(x, f^m).$$

The corresponding SAA problem is then defined as:

$$\max_{x \in X} \bar{g}_M(x), \tag{7}$$

where X is the first-stage feasible region, which is defined by constraint sets (2) and (3). Note, also, that X is invariant to the random samples, $f^1, f^2, \ldots, f^M$, that implicitly define the objective of the SAA problem, because of relatively complete recourse.

Algorithm 2 gives a high-level overview of the steps taken to solve the SAA problem. The SAA method begins with two inputs—an initial sample size and a desired relative tolerance. The method then randomly generates samples and solves the resulting SAA problem. If the relative optimality gap of the solution to the SAA problem is sufficiently small, then the algorithm terminates. Otherwise, M is increased and the algorithm begins with a new set of random samples. We now detail the L-shaped method used to solve the

SAA method and our approach to assessing solution quality in Steps 3 and 4 of the algorithm. We then discuss the convergence properties of Algorithm 2.

Algorithm 2 SAA Method

1: **input:** initial sample size M, relative tolerance d
2: generate i.i.d. random samples, $f^1, f^2, \ldots, f^M$
3: solve SAA problem (7) using L-shaped method with branch-and-cut
4: assess quality of SAA problem solution
5: **if** relative optimality gap $\leq d$ **then**
6: stop
7: **else**
8: increase M and return to Step 2
9: **end if**

2.3.1. Solution of SAA Problem

We solve the SAA problem in Step 3 of Algorithm 2 using a branch-and-cut method. Moreover, an L-shaped method is used to solve subproblems at each node of the branching tree. To explain our solution method, we rewrite the SAA problem as:

$$\max_{\theta,x} \quad \frac{1}{M}\sum_{m=1}^{M}\theta^m \tag{8}$$

$$\text{s.t.} \quad \sum_{k\in K} x_k \leq p; \tag{9}$$

$$x_k \in \{0,1\}, \qquad \forall k \in K. \tag{10}$$

This problem has a new set of decision variables, $\theta^1, \theta^2, \ldots, \theta^M$. The variable θ^m is defined as the optimal second-stage objective-function value in scenario m. Put another way, θ^m represents $g(x, f^m)$ in the SAA problem. The scenario-m second-stage problem is formulated as:

$$g(x, f^m) = \max_{y} \quad \sum_{q\in Q} f_q^m y_q^m \tag{11}$$

$$\text{s.t.} \quad \sum_{k\in N_q} x_k \geq y_q^m, \qquad \forall q \in Q; \tag{12}$$

$$y_q^m \in \{0,1\}, \qquad \forall q \in Q. \tag{13}$$

We explicitly place m superscripts on the y variables in this problem, to highlight the fact that the y's could take different values under different scenarios. It is straightforward to show that there is an integral optimal solution to the linear relaxation of second-stage problem (11)–(13). This can be shown using the total unimodularity of the recourse matrix corresponding to constraint set (12). Indeed, because $f^m \geq 0$, it is straightforward to explicitly derive an optimal integral solution to the second-stage problem. For each $q \in Q$, if $x_k = 1$ for at least one $k \in N_q$ then $y_q^{m*} = 1$ is optimal. Otherwise, if $x_k = 0$ for all $k \in N_q$, then only $y_q^{m*} = 0$ is feasible.

It is also straightforward to show that the second-stage problem is separable, because there are no constraints linking the y's with one another. Following from these two properties, problem (11)–(13) can be rewritten as $|Q|$ linear programs, one corresponding to each trip chain, q. The scenario-m second-stage

linear program corresponding to trip chain q is:

$$\max_{y_q^m} \quad f_q^m y_q^m \tag{14}$$

$$\text{s.t.} \quad \sum_{k \in N_q} x_k \geq y_q^m, \qquad (\pi_q^m); \tag{15}$$

$$y_q^m \leq 1, \qquad (\mu_q^m); \tag{16}$$

$$y_q^m \geq 0. \tag{17}$$

We let π_q^m and μ_q^m denote dual variables associated with constraints (15) and (16), respectively.

We can derive Benders's-type optimality cuts, which define the values of the θ's in the SAA problem in terms of dual information obtained from the second-stage subproblems. These cuts are defined as:

$$\theta^m \leq \sum_{q \in Q} \left(\mu_q^{m*} + \pi_q^{m*} \cdot \sum_{k \in N_q} x_k \right), \tag{18}$$

where π_q^{m*} and μ_q^{m*} are optimal values for the dual variables corresponding to problem (14)–(17).

These optimality cuts are iteratively added to the SAA problem, giving what we refer to as the L-shaped SAA problem. Our overall solution approach is as follows. The incumbent L-shaped SAA problem is solved to find values for the x's and θ's. The values of the x's are then input to the second-stage subproblems and the resulting objective-function values are compared to the θ's found in the SAA problem. If the θ's overestimate the true second-stage objective-function values, then optimality cuts are added and the new L-shaped SAA problem is re-solved. Otherwise, the algorithm terminates.

To give the formulation of the L-shaped SAA problem, we define $J_1, J_2, \ldots, J_M$ as the number of optimality cuts for scenarios $1, 2, \ldots, M$ that have been added to the SAA problem. We also define $\pi_q^{m,j}$ and $\mu_q^{m,j}$ as the optimal dual-variable values used to generate the jth optimality cut for scenario m. The L-shaped SAA problem with J optimality cuts is then formulated as:

$$\max_{\theta, x} \quad \frac{1}{M} \sum_{m=1}^{M} \theta^m \tag{19}$$

$$\text{s.t.} \quad \sum_{k \in K} x_k \leq p; \tag{20}$$

$$x_k \in \{0, 1\}, \qquad \forall k \in K; \tag{21}$$

$$\theta^m \leq \sum_{q \in Q} \left(\mu_q^{m,j} + \pi_q^{m,j} \cdot \sum_{k \in N_q} x_k \right), \qquad \forall m = 1, \ldots, M, j = 1, \ldots, J_m. \tag{22}$$

L-shaped SAA problem (19)–(22) is a mixed-integer program, which we solve using the branch-and-cut-based method outlined in Algorithm 3. In Step 1 we initialize the L-shaped SAA problem to have no optimality cuts and then iteratively add any optimality cuts needed in the remaining steps of the algorithm. Step 3 solves the incumbent L-shaped SAA problem with the optimality cuts found thus far. We then, in Steps 4 through 6, iteratively consider each second-stage subproblem and determine if $\hat{\theta}^m$ overestimates the true second-stage objective value that $\hat{x}$ gives. If so, a new optimality cut is generated in Steps 7 through 9 and added to the L-shaped SAA problem. This iterative procedure is repeated until the $\hat{\theta}$'s provide a proper estimate of the second-stage objective function value, in the sense that the optimality gap is less than some fixed tolerance, ρ (Step 12).

2.3.2. Assessing Quality of SAA Solution

In Step 4 of Algorithm 2 we assess the quality of the SAA solution that is found in Step 3. This is done by appealing to the results of Kleywegt et al. (2002); Bayraksan and Morton (2006). Kleywegt et al.

Algorithm 3 Branch-and-Cut Solution Method for L-Shaped SAA Problem

$J_m \leftarrow 0, \forall m = 1, \ldots, M$
repeat
  $(\hat{\theta}, \hat{x}) \leftarrow \arg\max_{\theta,x}$ (19) s.t. (20)–(22)
  for $m = 1, \ldots, M$ **do**
    $y_q^m \leftarrow \arg\max_y$ (14) s.t. (15)–(16) with f_q^m in (14) ▷ π_q^m and μ_q^m denote corresponding duals
    if $f_q^m y_q^m < \hat{\theta}^m$ **then** ▷ add a new scenario-m optimality cut
      $J_m \leftarrow J_m + 1$
      $\pi_q^{m,J_m} \leftarrow \pi_q^m$
      $\mu_q^{m,J_m} \leftarrow \mu_q^m$
    end if
  end for
until $\frac{1}{M}\sum_{m=1}^{M} \hat{\theta}^m - \frac{1}{M}\sum_{m=1}^{M} g(\hat{x}, f^m) < \rho$

(2002) show the asymptotic properties of the difference between the objective-function value of the SAA problem and the true problem. More specifically, let g^* denote the optimal objective function value of the true SFCLM:

$$\begin{aligned} g^* = \max_x \ & \mathbb{E}\left[g(x, \tilde{f})\right] \\ \text{s.t.} \ & \sum_{k \in K} x_k \leq p; \\ & x_k \in \{0,1\}, \qquad \forall k \in K; \end{aligned}$$

where $g(x, \tilde{f})$ is defined by (4)–(6). Also let:

$$X^* = \left\{ x \in X \,\middle|\, \mathbb{E}\left[g(x, \tilde{f})\right] = g^* \right\},$$

be the set of x's that are optimal in the true SFCLM, where X is (as before) defined by constraint sets (2) and (3). Also define $\hat{g}^M$ as the optimal objective-function value of the SAA problem with M random samples of the stochastic parameter:

$$\hat{g}^M = \max_{x \in X} \bar{g}_M(x).$$

Kleywegt et al. (2002) prove that for M sufficiently large we have:

$$\sqrt{M}(g^* - \hat{g}^M) \xrightarrow{d} \min_{x \in X^*} Z(x),$$

where $Z(x)$ are normally distributed random variables with mean zero and covariance equal to the covariance of $g(x, \tilde{f})$. This asymptotic property implies that the random variable, $(g^* - \hat{g}^M)$, is not normally distributed if $|X^*| > 1$ (*i.e.*, if there are multiple optimal solutions to the true SFCLM).

We can further show, by the concavity of $g(x, \tilde{f})$ (*cf.* the proof of Proposition 1 below, which shows that $g(x, \tilde{f})$ is a concave piecewise-linear function) and from Jensen's inequality, that $\hat{g}^M$ is a biased estimator of g^*. More specifically, we have:

$$\mathbb{E}\left[\hat{g}^M\right] \geq \mathbb{E}\left[\hat{g}^{M+1}\right] \geq g^*.$$

A frequent approach to estimating the optimality gap is to employ a batch-mean method. This method provides an optimality gap estimator that is asymptotically normally distributed.

Bayraksan and Morton (2006) suggest an A2RP, which is a batch-mean-based method, to estimate the optimality gap. We define the optimality gap as:

$$\eta_{x^*_M} = \mathbb{E}\left[\hat{g}^M\right] - \mathbb{E}\left[g(x^*_M, \tilde{f})\right],$$

where x_M^* is the optimal solution obtained in Step 3 of Algorithm 2 (*i.e.*, the final solution obtained in Algorithm 3). A2RP provides valid confidence intervals on the optimality gap of problems that satisfy the following three assumptions.

Assumption 5. *$g(\cdot, \tilde{f})$ is continuous on X with probability 1.*

Assumption 6. $\mathbb{E}\left[\sup_{x\in X}(g(x,\tilde{f}))^2\right] < +\infty$.

Assumption 7. *$X \neq \emptyset$ and is compact.*

We show in the following proposition that the SFCLM satisfies all three of these assumptions.

Proposition 1. *The SFCLM satisfies Assumptions 5 through 7.*

Proof. Assumption 6 follows immediately from analyzing the second-stage problem of the SFCLM, which is given by (4)–(6). More specifically, note that $g(x,\tilde{f}) \leq \sum_{q\in Q}\tilde{f}_q$ and that $\sum_{q\in Q}\tilde{f}_q$ is bounded above by the number of EVs in the study region, which is assumed to be finite. Thus, $(g(x,\tilde{f}))^2$ is guaranteed to be finite.

To show that Assumption 7 is satisfied, note that $X \neq \emptyset$ as long as $p \geq 0$. Moreover, as long as p is finite, X is bounded. We also note that X is closed because it is defined as an integer lattice. Thus, X is compact.

Finally, to show that Assumption 5 is satisfied, without loss of generality, we consider the scenario-m second-stage problem. First, we write the dual of the problem's linear relaxation and from strong duality know that:

$$
\begin{aligned}
g(x, f^m) = \min_{\pi,\mu} \quad & \sum_{q\in Q}\left(\mu_q^m + \pi_q^m \cdot \sum_{k\in N_q} x_k\right) & & \\
\text{s.t.} \quad & \pi_q^m + \mu_q^m \leq f_q^m, & \forall q \in Q; \quad & (23)\\
& \pi_q^m, \mu_q^m \geq 0, & \forall q \in Q. \quad & (24)
\end{aligned}
$$

We let Π denote the feasible region of the dual, which is defined by constraint sets (23) and (24). We know that Π is bounded by inspecting constraint sets (23) and (24). For notational convenience, define:

$$
\hat{\Pi} = \left\{(\pi^{m,1}, \mu^{m,1}), (\pi^{m,2}, \mu^{m,2}), \ldots, (\pi^{m,\Lambda}, \mu^{m,\Lambda})\right\},
$$

as the set of extreme points of polyhedron Π. We know that $\Lambda < +\infty$. Thus, we can rewrite the function $g(x, f^m)$ as:

$$
g(x, f^m) = \min_{\lambda\in\{1,\ldots,\Lambda\}} \sum_{q\in Q}\left(\mu_q^{m,\lambda} + \pi_q^{m,\lambda} \cdot \sum_{k\in N_q} x_k\right),
$$

which is a piecewise-linear concave function on $\text{Conv}(X)$. Thus, $g(\cdot,\tilde{f})$ is continuous on X. □

The A2RP works by comparing x_M^* to an optimal solution that would be obtained from solving an SAA problem with a different sample of the random vector. We outline the A2RP here and refer interested readers to the work of Bayraksan and Morton (2006) for full details. We begin by defining first- and second-moment estimators of $\eta_{x_M^*}$. This is done by defining two new random samples, $\dot{f}_1^1, \dot{f}_1^2, \ldots, \dot{f}_1^{L/2}$ and $\dot{f}_2^1, \dot{f}_2^2, \ldots, \dot{f}_2^{L/2}$, each of size $L/2$. The first-moment estimators are then defined as:

$$
G_{L/2}^i(x_M^*) = \max_{x\in X}\left\{\frac{2}{L}\sum_{l=1}^{L/2} g(x, \dot{f}_i^l)\right\} - \frac{2}{L}\sum_{l=1}^{L/2} g(x_M^*, \dot{f}_i^l), \forall i = 1, 2.
$$

Also define:

$$x^{i*}_{L/2} = \arg\max_{x \in X} \frac{2}{L} \sum_{l=1}^{L/2} g(x, \dot{f}^l_i),$$

as a maximizer of the first term defining $G^i_{L/2}(x^*_M)$. For notational convenience we define:

$$\hat{g}_{L/2}(x) = \frac{2}{L} \sum_{l=1}^{L/2} g(x, \dot{f}^l_i).$$

Thus, we can write:

$$G^i_{L/2}(x^*_M) = \hat{g}_{L/2}(x^{i*}_{L/2}) - \hat{g}_{L/2}(x^*_M), \forall i = 1, 2.$$

The second-moment estimators of the optimality gap are given by:

$$s^2_{L/2}(x^{i*}_{L/2}) = \frac{2}{L-2} \sum_{l=1}^{L/2} \left[(g(x^{i*}_{L/2}, \dot{f}^l_i) - g(x^*_M, \dot{f}^l_i)) - (\hat{g}_{L/2}(x^{i*}_{L/2}) - \hat{g}_{L/2}(x^*_M)) \right]^2.$$

Next, compute the estimates:

$$\bar{G}_{L/2}(x^*_M) = \frac{1}{2} \left[G^1_{L/2}(x^*_M) + G^2_{L/2}(x^*_M) \right],$$

and:

$$\bar{s}^2_{L/2} = \frac{1}{2} \left[s^2_{L/2}(x^{1*}_{L/2}) + s^2_{L/2}(x^{2*}_{L/2}) \right].$$

Bayraksan and Morton (2006) show that for any $\alpha \in (0, 1)$:

$$\left[0, \bar{G}_{L/2}(x^*_M) + \frac{t_{L-1,\alpha} \bar{s}_{L/2}}{\sqrt{L}} \right],$$

where $t_{L-1,\alpha}$ is a t-statistic with $L-1$ degrees of freedom is a valid one-sided $(1-\alpha)$ confidence interval on $\eta_{x^*_M}$. We state, but do not prove (interested readers are referred to their work for complete details), their formal result showing this.

Theorem 1. *Suppose Assumptions 5 through 7 hold, that $x^*_M \in X$, and that the samples, $\dot{f}^1_1, \dot{f}^2_1, \ldots, \dot{f}^{L/2}_1$ and $\dot{f}^1_2, \dot{f}^2_2, \ldots, \dot{f}^{L/2}_2$, are independent and identically distributed from the same distribution as $\tilde{f}$. Then $\forall \alpha \in (0, 1)$ we have that:*

$$\liminf_{L \to +\infty} \text{Prob} \left\{ \eta_{x^*_M} \leq \bar{G}_{L/2}(x^*_M) + \frac{z_\alpha \bar{s}_{L/2}}{\sqrt{L}} \right\} \geq 1 - \alpha,$$

where z_α is a z-score.

2.3.3. Convergence Properties of SAA Method

An important question in implementing Algorithm 2 is whether it is guaranteed to converge and, if so, the rate at which it does. We examine this in two parts. First, we examine whether Algorithm 3, which is used to solve the SAA problem in Step 3 of Algorithm 2 converges. We then examine whether we are guaranteed to find a solution with a desired optimality gap in Step 5 of Algorithm 2.

Algorithm 3 is guaranteed to converge because, at its heart, it is a branch-and-cut algorithm. The branching tree of the SAA problem will have at most 2^p nodes. Moreover, there are a finite number of optimality cuts that can be added to the L-shaped SAA problem, which is bounded by the number of extreme points of the second-stage subproblems (*cf.* the proof of Proposition 1). Nevertheless, Algorithm 3 has exponential time complexity. One method to speed up the algorithm, which we employ, is to add cuts as

necessary whenever a solution of the incumbent L-shaped SAA problem is found. That is to say, we do not have to find an integer-optimal solution to the incumbent L-shaped SAA problem in Step 3 of Algorithm 3. Rather, as solutions (either integer-feasible solutions or solutions to the linear relaxation) are found in the branching tree, optimality cuts can be added to the L-shaped SAA problem. This can significantly improve computational performance.

As for convergence of Algorithm 2 in Step 5, we rely on the results of Kleywegt et al. (2002), which can be used to show that Algorithm 2 converges and its convergence rate. We begin by defining:

$$\hat{X}_M^{\epsilon} = \left\{ x \in X \,\middle|\, |\bar{g}_M(x) - \hat{g}^M| \leq \epsilon \right\},$$

as the ϵ-optimal solution set of SAA problem (7). We similarly define:

$$X^{\epsilon} = \left\{ x \in X \,|\, |\bar{g}_M(x) - g^*| \leq \epsilon \right\},$$

as the ϵ-optimal solution set of true SFCLM. Kleywegt et al. (2002) show the following result, which we state but do not prove (interested readers are referred to their work for the details of the result).

Theorem 2. *The following two results hold:*

1. $\hat{g}^M \to g^*$ *with probability* 1 *as* $M \to +\infty$*; and*
2. *the event* $\left\{ \hat{X}_M^{\epsilon} \subset X^{\epsilon} \right\}$ *occurs with probability* 1 *for any* $\epsilon > 0$ *so long as* M *is sufficiently large.*

This theorem implies that so long as we allow for a strictly positive optimality gap and allow M to increase sufficiently, Algorithm 2 is guaranteed to converge. Kleywegt et al. (2002) also have a result showing that the algorithm has an exponential convergence rate, so long as the following assumption is satisfied.

Assumption 8. *For every* $x_1 \in X$ *and* $x_2 \in X \backslash X^{\epsilon}$*, the moment-generating function of the random variable,* $g(x_1, \tilde{f}) - g(x_2, \tilde{f})$*, is finite-valued in a neighborhood of* 0*.*

We show that the SFCLM satisfies this assumption in the following proposition.

Proposition 2. *The SFCLM satisfies Assumption 8.*

Proof. We have that:

$$-\infty < 0 \leq g(\cdot, \tilde{f}) \leq \sum_{q \in Q} \tilde{f}_q < +\infty.$$

The lower bound comes about because the lowest feasible second-stage objective value is attained by setting all of the y's equal to zero. The upper bound comes about because in a practical setting, there are a finite number of EVs. Thus, $g(x_1, \tilde{f}) - g(x_2, \tilde{f})$ is bounded and the moment-generating function is finite-valued. □

We now state, but do not prove, the convergence-rate result of Kleywegt et al. (2002).

Theorem 3. *Suppose Assumption 8 holds. Then for all* $\epsilon \geq \delta \geq 0$ *there exists a* $\gamma(\delta, \epsilon) > 0$ *such that:*

$$\text{Prob}\left\{ \hat{X}_M^{\epsilon} \not\subseteq X^{\epsilon} \right\} \leq |X \backslash X^{\epsilon}| e^{-M\gamma(\delta,\epsilon)}.$$

3. Case Study

We apply the modeling approach outlined in Section 2 to a case study based on the 12-county Central-Ohio region. This region covers roughly 6000 km^2. According to data obtained from MORPC, this region had approximately 1.3 million light-duty vehicles as of the end of 2010.

MORPC provides three datasets that we rely on: (i) geographic information system (GIS) data on the Central-Ohio region, including the road network; (ii) tour-record data specifying the use of the 1.3 million light-duty vehicles in Central Ohio on a typical workday; and (iii) average-vehicle-speed data for the road segments of the road network.

The GIS data contain three important data types for constructing our case study. One is the boundaries of 1805 traffic analysis zones (TAZs). MORPC models vehicle travel within the tour-record data as occurring between these TAZs. The second is GIS data for the Central-Ohio road network. This consists of roughly 20000 road segments and their lengths, and 26000 intersections between the road segments. Finally, MORPC provides data on the location of the 222 major commercial retail locations in the Central-Ohio region.

In our case study, we consider these 222 retailers as candidate locations for EV charging stations. We select these as candidate locations for three reasons. First, commercial retail locations are likely to have the physical space and existing electrical infrastructure to accommodate fast EV charging stations. A single fast charger may draw 200 kW of power. Thus, a station with multiple chargers may add more than 1 MW of load to the distribution system, which could not be easily accommodated by a residential circuit. Second, retail locations are likely to have a commercial rationale for EV charging station development. Recharging an EV, even with a fast charging technology, may take longer than refueling a gasoline vehicle. Because an EV driver may be parked at a charging station for several minutes while making retail purchases, this would provide a convenient opportunity for vehicle recharging. Indeed, He et al. (2013) present retailers offering EV charging as a potential business model for developing fast EV charging infrastructure. Co-locating a charging station with a retail shopping location may also provide an incentive for an EV driver to use a charging station that is out of the way from his or her intended route from origin to destination. A third rationale is that the retail locations in Central Ohio are relatively evenly placed within the study region and are concentrated around major arterial roadways that carry high traffic volumes. Thus, the candidate locations provide adequate coverage of where EVs may be driven in the future.

The tour-record data provide modeled data on the use of the 1.3 million light-duty vehicles in Central Ohio during a typical workday. The data are generated using the multistep tour-based approach described by Sener et al. (2009). The data include about 2.4 million daily tours, of which about 2.1 million involve the use of a personal vehicle, made by 688000 households. Each tour can includes multiple trips. Each trip is characterized by the vehicle's starting and ending TAZs, the departure and return time from these TAZs, and an indicator of whether the vehicle makes a sub-tour (*i.e.*, an additional trip *en route* of the main trip). If there is a sub-tour, the tour record indicates the destination TAZ of the sub-tour.

For instance, a vehicle may begin and end a tour at TAZ i and be used for commuting to TAZ j. The tour record for this particular vehicle would indicate a departure trip from TAZ i to j at its departure time and a return trip from TAZ j to i at its return time. Moreover, if the vehicle stops at a retail shopping location in TAZ k *en route* from TAZ j to i in the evening, there would be a sub-tour in the tour record data. Combining the trips and the sub-tour would indicate that the particular vehicle's tour is $i \to j \to k \to i$ with corresponding timestamps of its arrival time to and departure time from each TAZ. A trip chain is the sequence of trips that a vehicle takes (*e.g.*, $i \to j \to k \to i$) without timestamps. Although the tour-record data specify the sequence of points that a vehicle travels to, they do not specify the paths taken between the points.

Finally, MORPC has average-vehicle-speed data for a typical workday for all of the road segments in the road network dataset. MORPC reports road speeds for four time periods: 5:00 am to 9:00 am, 10:00 am to 2:00 pm, 3:00 pm to 7:00 pm, and 8:00 pm to 12:00am.

Figure 1 summarizes the steps taken to use the MORPC data to produce our case study. We begin by randomly sampling from among the 1.1 million light-duty vehicles in the tour-record data to determine which are EVs and which are conventional gasoline vehicles. Our case study assumes a 3% EV penetration rate, meaning that we randomly determine with probability 0.03 which of the modeled vehicles in the tour-record data are EVs and which are not. Note that each random sample ultimately corresponds to a realization of the exogenous random variable, $\tilde{f}$, in the SFCLM. Moreover, when we solve the SAA problem or use the A2RP we generate f^m's or f_i^l's by randomly sampling which vehicles in the tour-record data are EVs.

Next, for each realization of which vehicles are EVs and not, we rely on Assumption 1 to determine the path taken by each EV to get from its origin to destination on each trip. This is done by solving a shortest-time-path problem, using the average-vehicle-speed data for each road segment in the road network. MORPC provides vehicle-speed data for four time periods. Thus, for each EV tour we use road-speed data corresponding to the time block during which the trip takes place.

Then, we use Algorithm 1 and rely on Assumptions 2–4 to determine which EVs could be captured by

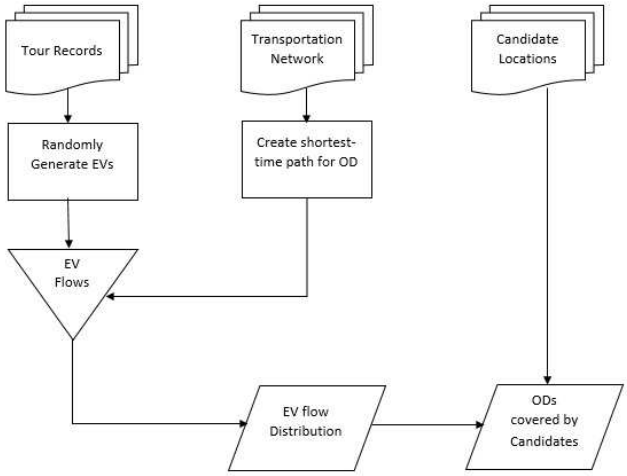

Figure 1: Case study simulation flowchart.

a charging station at each of the 222 candidate locations. In doing so, we must determine the charging-deviation radius, r, and range-anxiety threshold, $\tau\%$. Kuby et al. (2013) investigate refueling behavior of early alternative-fuel vehicle adopters. They find that 83.4% of alternative-fuel vehicle owners refuel their vehicles when the fuel tank is at less than 3/8 capacity. They also find that the median refueling deviation tolerance of alternative-fuel vehicle owners is between 0.8 miles and 1.3 miles. They further find that more than 75% of drivers are willing to deviate 9 minutes out of their intended driving paths to refuel. If we assume a 45 mile per hour average driving speed when deviating to a charging station, this 9 minutes of allowable deviation time is equivalent to about 4 miles.

Based on these findings, we use $\tau = 40$ as the range-anxiety threshold and $r = 1.6$ km (1 mile) as the charging-deviation radius. We also conduct a sensitivity analysis in which we increase the charging-deviation radius to $r = 6.4$ km (4 miles). Thus, for each EV trip in a tour, we determine its intended path from the shortest time-path model. If this intended path comes within r km of a particular candidate location and the EV SoC satisfies the range-anxiety criterion (*i.e.*, if SoC $\in [\tau, 1]$ for a trip chain, q, with $\mathcal{L}(q) < \mathcal{R}$ or SoC $\in [\max\{\tau, \mathcal{L}(q)/\mathcal{R} - 1\}, 1]$ for a trip chain, q, with $\mathcal{R} < \mathcal{L}(q) \leq 2\mathcal{R}$), then that candidate location is able to 'capture' that EV. We assume that the EVs have a driving range of 80 miles in determining their SoCs.

4. Case Study Results

We begin by summarizing the results using our SFCLM. We next present results illustrating the benefit of the public charging station network, in terms of the increase in the number of simulated EVs that can complete their typical daily tours with the charging stations present. We also provide insights into the number of chargers that must be deployed at the charging stations selected by the SFCLM for EVs to be able to use the stations. We finally contrast the results of our SFCLM with a deterministic variant of the model, in which EV flows are assumed to be known and equal to a sample mean of $\tilde{f}$.

4.1. Stochastic Flow-Capturing Location Model Results

We examine the optimal placement of charging stations as a function of the total number of charging stations built. This is done by solving the SFCLM and varying the value of p. Figure 2 shows the optimal locations of six, 11, 21, and 31 charging stations. The figure shows two important properties of the optimal locations of the charging stations. First, the optimal charging station locations are fairly robust to the total number of stations built. That is to say, each of the six locations where a charging station is built with $p = 6$ is close to an optimal charging-station location with higher values of p. This is also true of the

stations built with $p = 11$ and $p = 21$. Upchurch and Kuby (2010) note this same finding regarding nested station locations with p-median and flow-refueling models. This finding that the optimal station locations are roughly nested implies that if the region has limited resources to build an initial charging station network, it can build those stations without having to be concerned about future station deployments. If additional stations are built later on, the locations of the starting stations would still be close to optimal as the network is expanded.

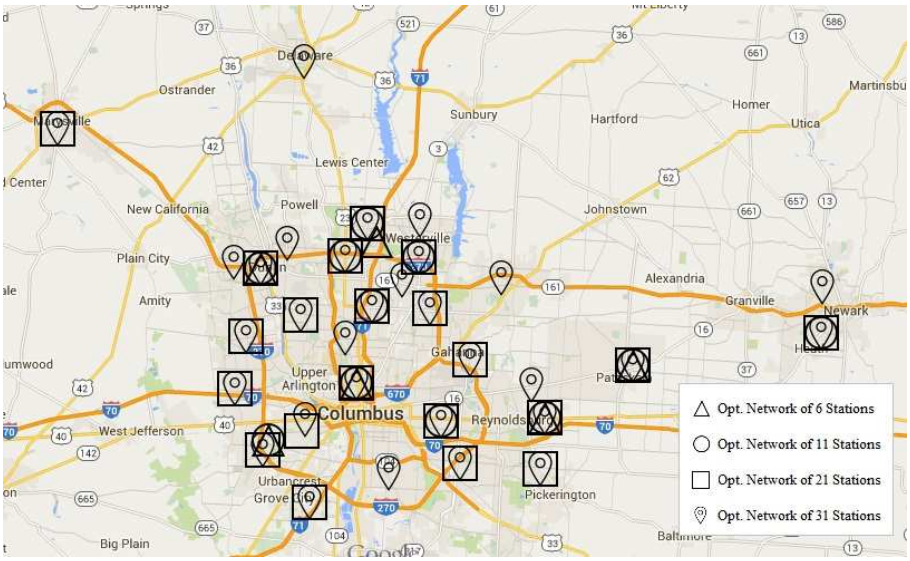

Figure 2: Optimal charging station locations as a function of p.

The second property shown in Figure 2 concerns the locations of the stations themselves. The Central-Ohio region is centered around the city of Columbus, which is relatively (compared to the surrounding region) densely populated. Columbus is surrounded by a beltway (Interstate 270). The region outside of this beltway is sparsely populated (*i.e.*, it consists of suburban and rural communities). There are also three smaller cities outside of the beltway—Marysville, Delaware, and Newark. If the network is limited to having a small number of charging stations, the stations are located on the periphery of the beltway. This provides good coverage of many EVs, which tend to use the beltway for at least a portion of one trip daily. With $p = 6$ there is only one charging station built in the center of the beltway (near the central business district of Columbus).

As the number of charging stations built increases, most of the additional stations are still built on the periphery and inside of the beltway. This is because such locations still provide more coverage of EVs than locations outside of the beltway. As the number of stations built increases further, some are placed in the three smaller cities outside of the beltway.

Figure 3 shows the expected number of EVs that can be captured as a function of p. With a 3% EV-penetration level, there would be approximately 40000 EVs (in expectation) in the study region. The figure shows that if $p = 222$ (*i.e.*, a charging station is built at every candidate location considered), a maximum of 20000 EVs are captured in expectation. Thus, approximately half of the EVs are not driven within a 1-mile radius of any of the candidate locations. Assumptions 1 and 4 are likely conservative in estimating how many EVs can be captured by the stations built. This is because drivers may adjust their destinations and driving patterns more than our assumptions account for if fast charging stations are built at retail shopping locations.

The vertical axis on the right-hand side of Figure 3 normalizes the expected number of EVs that can be captured as a percentage of the maximum number that can be captured with 222 charging stations built. This axis shows that a relatively small number of charging stations can capture a majority of the EVs. For instance, 26 charging stations, which account for 12% of the candidate locations, could capture more than

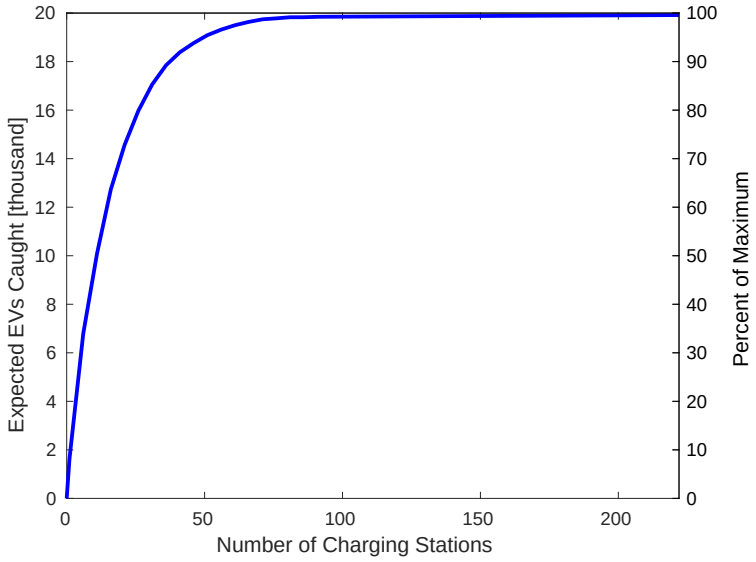

Figure 3: Expected number of EVs captured as a function of p.

80% of the maximum number of EVs captured by all 222 stations. Thirty-six stations would need to be built to be able to capture 90% of the maximum number that can be captured by all 222 stations.

Figure 4 shows a 95% confidence interval on the relative optimality gap of the SFCLM, obtained from the A2RP. This optimality gap confidence interval is defined as:

$$\frac{\bar{G}_{L/2}(x^*_M) + (z_\alpha \bar{s}_{L/2})/\sqrt{L} - \hat{g}^M}{\hat{g}^M}.$$

All of the cases (*i.e.*, with different values of p) are solved using at most $M \leq 1000$ random samples to generate the SAA problem and $L/2 = 500$ random samples for each of the two replications used in the A2RP. Figure 4 shows that these sample sizes are sufficient for obtaining a high-quality solution with an optimality gap of less than 0.0045 (0.45%). Solutions with tighter optimality gaps would require larger samples to generate the SAA problem. This would involve a commensurate increase in computational efforts. Our computational simulations are conducted on the Oakley Cluster at the Ohio Supercomputer Center (*cf.* Ohio Supercomputer Center (1987) for details of the Oakley Cluster) with a CPU-clock limit of 48 hours. Given that the SFCLM is a planning problem that can be solved long in advance of when charging stations are built, a 48-hour solution time limit is reasonable for such a problem.

4.2. Effect of Charging-Deviation Radius

We examine the effect of the charging-deviation radius, r, on the optimal charging-station layout in two ways. First, Figure 5 shows the optimal location of $p = 6$ charging stations with a radius of $r = 6.4$. Three of the charging stations are in the same location, two in very slightly different locations, and one in a very different location when compared to the $r = 1.6$ case. This sixth station is moved from far east of the city center to a location just outside of the beltway. This is because the increased charging-deviation radius allows a station at the new location to serve EVs on the beltway and on two other major arterial routes. Moreover, this station can now serve vehicles from the airport and around a major shopping center in east Columbus. This is not possible with $r = 1.6$.

Figure 6 shows the effect of the optimal charging-station network given by the SFCLM assuming $p = 6$ and $r = 1.6$ km, if drivers have a higher charging-deviation radius than assumed (*i.e.*, if r is actually greater than 1.6 km). The blue solid line in the figure shows the expected number of EVs that would not be able to complete their daily tours, as a function of the *true* value of r (*i.e.*, if r is not actually equal to 1.6 km). With

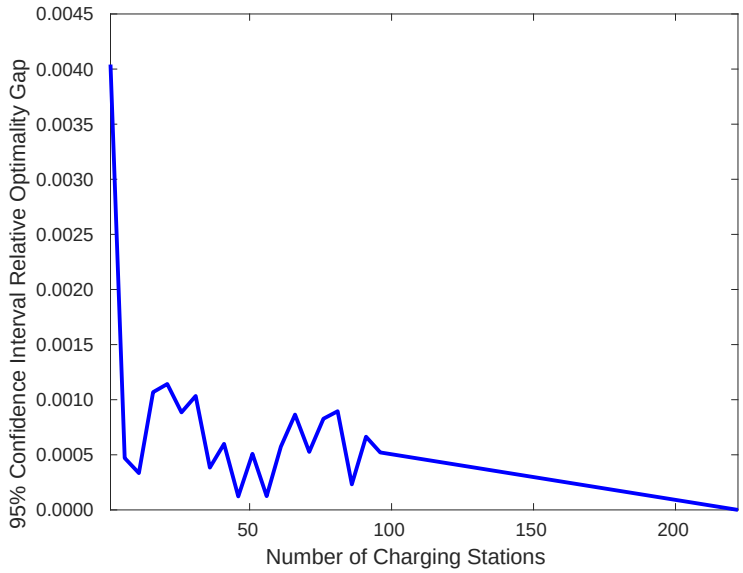

Figure 4: 95% confidence interval on relative optimality gap of solution given by SFCLM as a function of p.

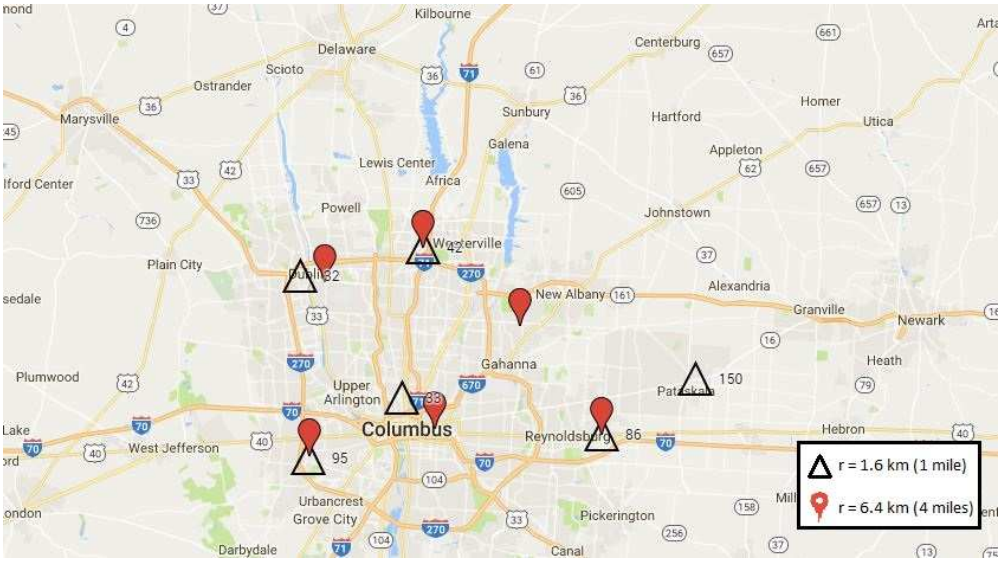

Figure 5: Optimal charging-station network with $r = 1.6$ km and $r = 6.4$ km.

the assumed 3% EV-penetration level, there are about 39000 EVs in expectation. Of these, an expected 14000 have driving patterns that exceed the assumed 80 mile range. If six charging stations are built, an expected 10000 of these 14000 EVs can complete their daily trips if the true value of r is 1.6 km. As the true value of r increases, more of these EVs are able to complete their daily trips, with all but 5% of the EVs unable to complete their trips with $r = 16$ km.

The red dashed line shows the expected number of EVs with daily trips longer than their assumed 80 mile range that can complete their daily trips with the charging stations. This is given as a percentage of the expected number of EVs with daily trips longer than the vehicle range. The figure shows that many EVs can complete their daily trips with only six charging stations, if drivers are willing to deviate from their intended path from origin to destination.

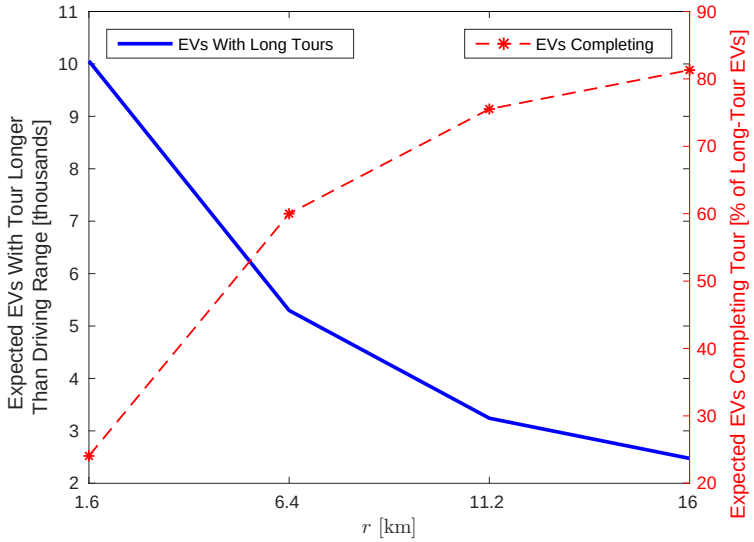

Figure 6: Expected number of EVs with typical daily driving tours longer than 80 miles, and percentage of those EVs that can complete their tours with a six-station EV-charging network as a function of r.

4.3. Benefit of Public Charging Network to EV Owners

From the perspective of EV owners, an EV-charging network can be evaluated using at least two criteria. One is whether the stations are in locations that allow them to be used by EV owners. Objective function (1) of our SFCLM maximizes the expected number of prospective EV owners that can use at least one charging station without having to drive more than a fixed distance out of the intended path from origin to destination.

A second criterion is what benefit the charging-station network provides in allowing EV owners to complete their daily tours. The expected number of EVs that can complete their daily tours can differ from the value of objective function (1). This is because some EVs that can use a charging station do not drive far enough in a single day to deplete the vehicle battery. Conversely, some EVs may have very long daily tours, which they cannot complete solely using public charging stations. Both types of EVs may still derive benefits from the charging-station network in terms of the second criterion. For instance, an EV with a very long daily tour may not be able to complete its tour using *only* the public charging stations chosen by the SFCLM. However, it may be able to complete its tour using the public charging stations and other charging infrastructure (*e.g.*, at a workplace).

We examine this second criterion by conducting a further simulation using the optimized charging-station locations determined by the SFCLM. Specifically, we fix the locations of the charging stations determined by the SFCLM and randomly generate 365 samples of $\tilde{f}$ (assuming the same 3% EV penetration), by bootstrapping from the the tour record data. For each random sample of $\tilde{f}$, we assume that each EV begins the day with a fully charged battery (*e.g.*, using an at-home charger). We next simulate the SoC of each EV battery during the course of its daily tour, assuming that any EV that passes within the r km capture radius of one of the charging stations stops at the station and fully recharges its battery.

We use this simulation to estimate the expected number of EVs that can successfully complete their daily tours (*i.e.*, the SoC of the EV battery does not reach zero before completing the tour) during the course of a typical year. We contrast this with the number of EVs that would be able to successfully complete their daily tours without any public charging stations (this is done by repeating the same simulation with no public charging stations).

Figures 7 and 8 summarize the simulation results for a six-station EV-charging network. First, the solid line in Figure 7 shows the expected number of EVs that have daily driving tours that are longer than a given EV range. Our assumed 3% EV-penetration level corresponds to about 40000 EVs in expectation. Thus, an EV with a 70-mile driving range does not provide sufficient energy for a quarter of drivers to

complete a typical daily tour without recharging midday. The stars in Figure 7 show the fraction of EVs that have a typical daily tour longer than a given EV range (*i.e.*, the fraction of the solid line in the figure) that can complete their tours if a six-station EV-charging network is built. The figure shows that a six-station network allows more than 60% of EVs that would otherwise not be able to complete their daily tours (without the charging stations) to do so.

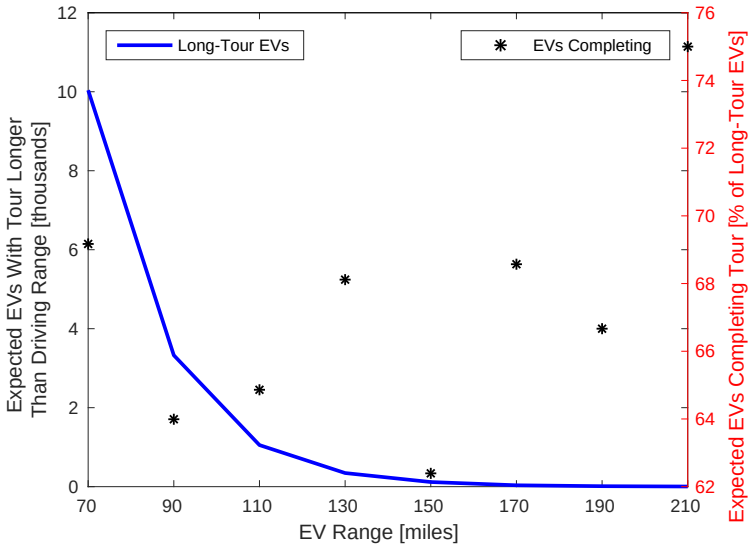

Figure 7: Expected number of EVs with typical daily driving tours longer than a given EV range, and percentage of those EVs that can complete their tours with a six-station EV-charging network.

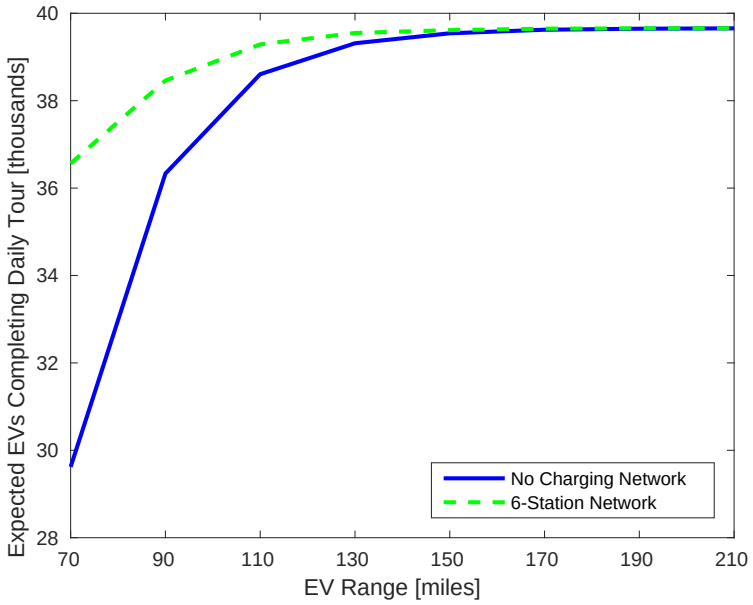

Figure 8: Expected number of EVs able to complete their daily tours with and without a six-station EV-charging network.

Figure 8 summarizes the expected number of EVs that are able to complete their typical daily tours with and without a six-station EV-charging network, as a function of the EVs' range. The figure shows that for EVs with ranges of less than 110 miles (which includes the 2016 Nissan Leaf, 2016 BMW i3, and

2017 Mitsubishi i-MiEV), a six-station EV-charging network has significant benefits in allowing more drivers to be able to complete their typical driving activities using an EV.

The estimates of the impacts of the charging-station network shown in Figures 7 and 8 are likely conservative. This is because our simulation assumes that EV owners will only deviate a fixed distance from the intended path between origin and destination to use a public charging station. In practice, EV owners may plan their driving tours based on the location of public charging stations (*e.g.*, shopping at a retailer that has a charging station available). Moreover, drivers that choose to purchase an EV will do so based on their driving patterns. Our simulations indicate that 4 EVs (in expectation) have typical daily tours that are longer than 210 miles. Drivers with such driving patterns are less likely to purchase EVs.

4.4. Charging Station Sizing

Another question in developing public EV charging stations is the number of chargers that should be deployed at each station. Xi et al. (2013) study this question in the context of deploying Level-1 and -2 charging infrastructure by simulating a queuing model of EVs waiting to use an available charger upon arrival at a charging station. We take a similar approach here, building off of the simulation conducted in Section 4.3 to estimate the number of EVs wanting to use each charging station in the network throughout the course of a typical day.

Specifically, we fix the locations of the charging stations determined by the SFCLM and use the 365 samples of $\tilde{f}$ randomly generated for the simulation conducted in Section 4.3. For each random sample, we assume that each EV begins the day with a fully charged battery and simulate the SoC of each EV battery during the course of its daily tour. Using Assumptions 2 and 4, we determine when EVs arrive to one of the public charging stations based on when the vehicle battery SoC drops below the assumed range-anxiety threshold. The charging duration, once an EV arrives to a charging station, is determined based on an assumed power capacity of the charger and the SoC of the vehicle battery upon arrival to the charging station.

Figure 9 summarizes the results of this simulation with six charging stations that have 100 kW chargers, EVs that have 85-mile driving ranges, and drivers with 20% SoC thresholds. This means that an EV driver is assumed to use a public charging station if and only if the SoC of his or her vehicle battery drops below 20% and a charging station is available within 1 mile of the planned route. Figure 9 shows the cumulative distribution of the number of vehicles that are charging at any given time. For instance, Station 1 is unoccupied 60% of the time and fewer than one EV is charging 80% of the time. The results of this simulation can be used to determine the number of chargers to deploy, based on a desired service-related criterion. As an illustrative example, the horizontal solid line in Figure 9 is at the 95th percentile. The figure shows that deploying one charger in Stations 3 and 5, two in Station 6, three in Station 1, four in Station 2, and five in Station 4 will result in a 95% probability that no EVs will have to queue to use a charger.

Figure 10 summarizes the results of the simulation assuming 200-kW chargers for the same six stations. The cumulative distributions in Figure 9 first-order stochastically dominate those in Figure 10. This dominance is because of the higher power capacity in the latter case. EVs have shorter charging times with 200-kW chargers, meaning that there will always be fewer EVs simultaneously needing to use a charger.

4.5. Comparison of Stochastic and Deterministic Flow-Capturing Location Models

An important question in developing our SFCLM is the benefit that it provides relative to a deterministic version of the model. This is because the SFCLM requires the use of Monte Carlo-based simulation and an L-shaped algorithm to find a near-optimal solution and determine its quality. If a deterministic model can provide a solution of similar quality with less work involved, the value of a SFCLM would be minimal and the deterministic model should be used instead.

We compare our SFCLM to a deterministic flow-capturing location model (DFCLM) in two ways. First, we examine the optimal locations chosen by the SFCLM and DFCLM to place charging stations. We then examine how the charging station locations given by the two models compare in terms of the expected number of EV that can be captured.

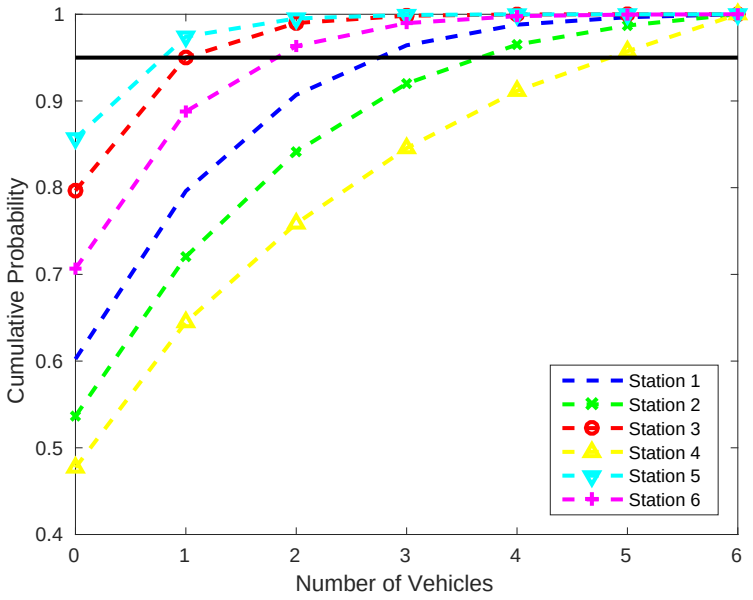

Figure 9: Cumulative distribution of the number of EVs in each of six stations with 100-kW chargers assuming 85-mile EV driving ranges and 20% SoC thresholds.

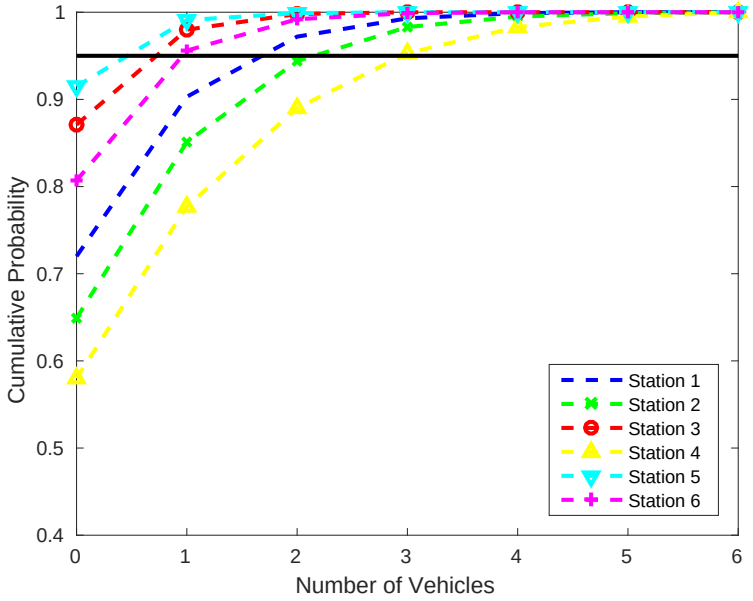

Figure 10: Cumulative distribution of the number of EVs in each of six stations with 200-kW chargers assuming 85-mile EV driving ranges and 20% SoC thresholds.

To formulate the DFCLM, we define $\bar{f}_q = \mathbb{E}\left[\tilde{f}_q\right]$ as the expected value of $\tilde{f}_q$. The DFCLM is then formulated as:

$$\max_{x,y} \quad \sum_{q \in Q} \bar{f}_q y_q \tag{25}$$

$$\text{s.t.} \quad \sum_{k \in K} x_k \leq p; \tag{26}$$

$$\sum_{k \in N_q} x_k \geq y_q, \qquad \forall q \in Q; \tag{27}$$

$$x_k \in \{0,1\}, \qquad \forall k \in K; \tag{28}$$

$$y_q \in \{0,1\}, \qquad \forall q \in Q; \tag{29}$$

where Q, K, N_q, p, x_k and y_q retain the same definitions introduced in Section 2.1. The DFCLM has a very similar structure to the SFCLM. Objective function (25) maximizes the number of EVs captured by the network of charging stations built, based on expected EV flows on each trip chain. The constraints of the DFCLM correspond to the analogous constraints of the SFCLM—constraint (26) limits the number of stations built, constraints (27) define the EV flows captured in terms of the stations built, and constraints (28) and (29) impose integrality.

As noted in Section 2.3, there are an exponentially large number of EV-flow scenarios. Thus, it is intractable to compute $\bar{f}$ exactly. For this reason, we use a sample average of $\bar{f}$ in defining objective function (25). More specifically, we use the same M random samples of $\tilde{f}$ used to define the SAA problem in solving the SFCLM to compute the sample average of $\bar{f}$.

Figures 11 and 12 show the optimal locations of six and 31 charging stations, respectively, chosen when the SFCLM and DFCLM are applied to the same underlying Central-Ohio case study data. Comparing the optimal station locations reveals some interesting properties of the station locations. First, the DFCLM results in the same general station-location strategy as the SFCLM. If there is a limited number of charging stations to be built (*e.g.*, with $p = 6$), the DFCLM chooses to locate them around the periphery of the beltway. This is because the beltway sees high expected use by EVs. As the number of stations to be built increases, the DFCLM places a greater number of stations inside the beltway and in the three smaller cities outside of Columbus.

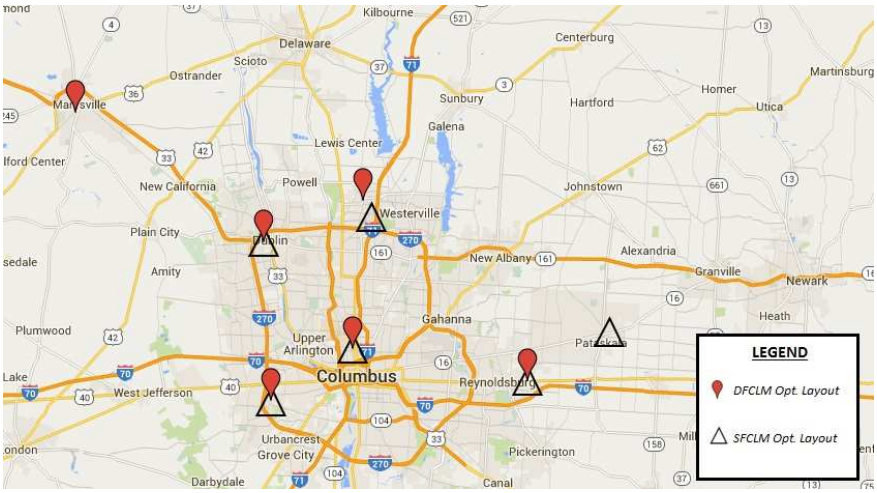

Figure 11: Optimal charging station locations from SFCLM and DFCLM with $p = 6$.

Although the DFCLM follows the same general strategy as the SFCLM, the two models nevertheless place stations at different locations. This is most notable in the $p = 6$ case shown in Figure 11. The DFCLM places one of the six stations in the small town of Marysville (northwest of Columbus), whereas the SFCLM only places a station in this town with $p \geq 21$ (*cf.* Figure 2). With $p = 31$ the locations chosen by the DFCLM and SFCLM are quite similar. Although they are not identical, each of the 31 station locations chosen by the DFCLM is in very close proximity to a station location chosen by the SFCLM.

We study the effects of building charging stations at these different locations by comparing the expected number of EVs captured by the charging station networks selected by the SFCLM and DFCLM. Specifically, we fix the station locations determined by each model and simulate the number of EVs that can be captured under 1000 randomly generated samples of $\tilde{f}$. Figure 13 shows the expected number of EVs that can be

Figure 12: Optimal charging station locations from SFCLM and DFCLM with $p = 31$.

captured when up to 31 charging stations are built. For greater numbers of charging stations the locations selected by the SFCLM and DFCLM result in virtually the same expected number of EVs able to be captured. This is consistent with what is shown in Figure 12—if 31 stations are being placed the SFCLM and DFCLM chose locations that are in extremely close proximity to one another, resulting in virtually the same expected number of EVs able to be captured. This is true for greater numbers of charging stations as well, as there is a fraction of a percent difference in the expected number of EVs that can be captured by the two station networks.

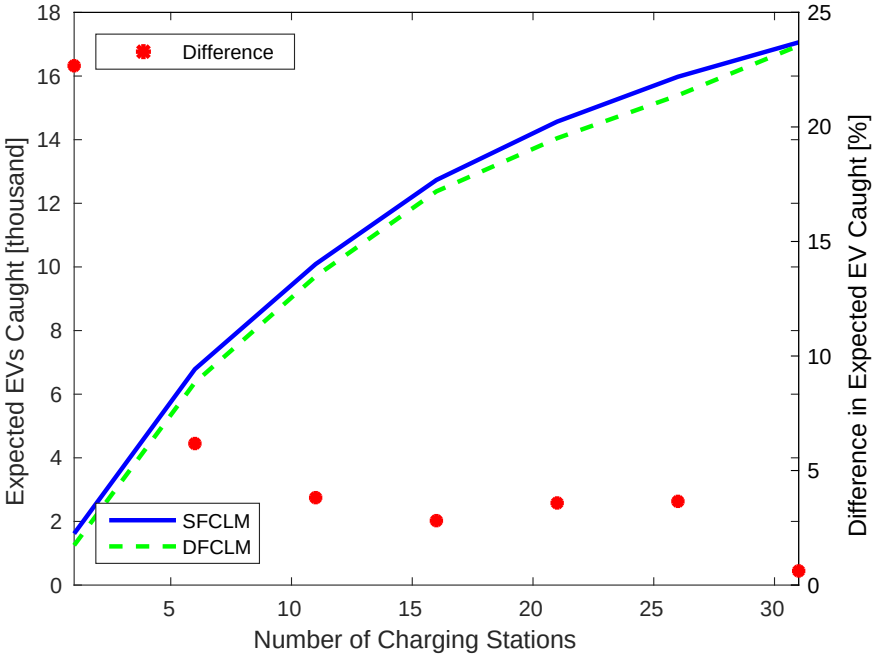

Figure 13: Expected number of EVs captured with charging station locations chosen by SFCLM and DFCLM as a function of p.

The two lines in Figure 13 show the expected number of EVs that can be captured using the station

locations determined by the SFCLM and DFCLM. The dots show the percentage difference between the expected number of EVs that can be captured by the two sets of station locations and is a measure of the value of the stochastic solution given by the SFCLM. The important finding illustrated in Figure 13 is that if a limited number of charging stations is to be built, the added effort of using a SFCLM is worthwhile. This is because the station locations chosen by the DFCLM underperform relative to those chosen by the SFCLM by 2%-22%. If there are limited resources for building a network of charging stations, the SFCLM provides a more robust set of station locations. As the available resources for building charging stations increases, the SFCLM has limited incremental value compared to the DFCLM. This is because stations are being added to provide coverage of areas with relatively low EV traffic volume and the gains from these added stations are marginal.

5. Conclusions and Future Work

This paper introduces a SFCLM to optimize the locations of a limited number of EV fast charging stations within a given study region. We formulate this problem as a two-stage stochastic integer program, in which the station locations are fixed in the first stage and EV flows between different trip chains are determined in the second stage. We apply an L-shaped SAA algorithm to decompose and solve the SFCLM and apply an A2RP to estimate the optimality gap of the final solution. A benefit of the A2RP is that it allows us to dynamically determine the number of scenarios modeled in the SAA problem to achieve a desired solution quality. Our SFCLM and solution method can be contrasted with other stochastic charging station-location models that use a very limited scenario tree size and do not provide an assessment of solution quality.

We apply our model to a case study based on Central Ohio. This is a comprehensive case study with 1.3 million vehicles, 20000 road segments, 9.9×10^{391338} potential second-stage scenarios, and 222 candidate charging station locations. Our SAA method is able to determine near-optimal station locations using less than 48 hours of CPU time, which is reasonable for a planning problem of this type. We demonstrate that the station locations are robust to the number of stations ultimately built. We further demonstrate the benefit of the public charging stations in allowing for greater use of EVs using only a small number of charging stations with a limited number of chargers. We also demonstrate the benefit of the SFCLM by contrasting it with a DFCLM. We find that if the number of stations to be built is limited, the SFCLM has value in maximizing the expected number of EVs that can be captured. As the number of stations to be built increases, the DFCLM could be used instead as the station networks determined by the two models become very similar.

Our work focuses on optimizing the location of charging stations. An important follow-on question is to how to manage charging loads at a given charging station. This can be determined and optimized using a station-operation control model. Previous works on charging control include decentralized approaches, such as those examined by Wu et al. (2012); Xi and Sioshansi (2014), and centralized approaches, such as those proposed by Hu et al. (2014); Hung and Michailidis (2015). The shortest-path based simulation that we employ can be used to generate EV flows in such a control problem. Otherwise, one can alternately use the advanced dynamic method introduced by Michailidis (2013) to create EV flows and interarrival processes. A related question, which is studied by Yudovina and Michailidis (2015); Hung and Michailidis (2015), is how to 'allocate' EVs to different charging stations to minimize queuing for access to a charger. A final area of future work, which we note in Section 2.2, is to extend our modeling methodology to include long EV tours requiring more than one refueling stop.

Acknowledgments

The authors thank the Mid-Ohio Regional Planning Commission for data collection assistance. Armin Sorooshian, the editors, and three anonymous reviewers provided helpful conversations and suggestions. This material is based upon work supported by the U.S. Department of Energy under Award Number DE-PI0000012. This work was also supported by an allocation of computing time from the Ohio Supercomputer Center.

References

Bayraksan, G., Morton, D. P., September 2006. Assessing solution quality in stochastic programs. Mathematical Programming 108, 495–514.

Chéron, E., Zins, M., May 1997. Electric vehicle purchasing intentions: The concern over battery charge duration. Transportation Research Part A: Policy and Practice 31, 235–243.

Dong, X., Mu, Y., Jia, H., Wu, J., Yu, X., October 2016. Planning of Fast EV Charging Stations on a Round Freeway. IEEE Transactions on Sustainable Energy 7, 1452–1461.

Eberle, U., von Helmolt, R., 2010. Sustainable transportation based on electric vehicle concepts: a brief overview. Energy & Environmental Science 3, 689–699.

Frade, I., Anabela, R., Goncalves, G., Antunes, A. P., 2011. Optimal Location of Charging Stations for Electric Vehicles in a Neighborhood in Lisbon, Portugal. Transportation Research Record: Journal of the Transportation Research Board 2252, 91–98.

Franke, T., Krems, J. F., February 2013. Interacting with limited mobility resources: Psychological range levels in electric vehicle use. Transportation Research Part A: Policy and Practice 48, 109–122.

He, F., Wu, D., Yin, Y., Guan, Y., January 2013. Optimal deployment of public charging stations for plug-in hybrid electric vehicles. Transportation Research Part B: Methodological 47, 87–101.

He, F., Yin, Y., Zhou, J., November 2015. Deploying public charging stations for electric vehicles on urban road networks. Transportation Research Part C: Emerging Technologies 60, 227–240.

Hodgson, M. J., July 1990. A Flow-Capturing Location-Allocation Model. Geographical Analysis 22, 270–279.

Hu, J., You, S., Lind, M., Østergaard, J., March 2014. Coordinated Charging of Electric Vehicles for Congestion Prevention in the Distribution Grid. IEEE Transactions on Smart Grid 5, 703–711.

Hung, Y.-C., Michailidis, G., 1 December 2015. Optimal routing for electric vehicle service systems. European Journal of Operational Research 247, 515–524.

Kang, J. E., Recker, W. W., December 2009. An activity-based assessment of the potential impacts of plug-in hybrid electric vehicles on energy and emissions using 1-day travel data. Transportation Research Part D: Transport and Environment 14, 541–556.

Kim, J.-G., Kuby, M., March 2012. The deviation-flow refueling location model for optimizing a network of refueling stations. International Journal of Hydrogen Energy 37, 5406–5420.

Kleywegt, A. J., Shapiro, A., de Mello, T. H., 2002. The Sample Average Approximation Method for Stochastic Discrete Optimization. SIAM Journal on Optimization 12, 479–502.

Kuby, M., Lim, S., June 2005. The flow-refueling location problem for alternative-fuel vehicles. Socio-Economic Planning Sciences 39, 125–154.

Kuby, M. J., Kelley, S. B., Schoenemann, J., December 2013. Spatial refueling patterns of alternative-fuel and gasoline vehicle drivers in Los Angeles. Transportation Research Part D: Transport and Environment 25, 84–92.

Li, S., Huang, Y., December 2014. Heuristic approaches for the flow-based set covering problem with deviation paths. Transportation Research Part E: Logistics and Transportation Review 72, 144–158.

Li, S., Huang, Y., Mason, S. J., April 2016. A multi-period optimization model for the deployment of public electric vehicle charging stations on network. Transportation Research Part C: Emerging Technologies 65, 128–143.

Lin, Z., Ogden, J., Fan, Y., Chen, C.-W., June 2008. The fuel-travel-back approach to hydrogen station siting. International Journal of Hydrogen Energy 33, 3096–3101.

Michailidis, G., 103 July 2013. Power allocation to a network of charging stations based on network tomography monitoring. In: 2013 18th International Conference on Digital Signal Processing (DSP). Institute of Electrical and Electronics Engineers, Fira, Santorini, Greece.

Nicholas, M. A., Handy, S. L., Sperling, D., 2004. Using Geographic Information Systems to Evaluate Siting and Networks of Hydrogen Stations. Transportation Research Record: Journal of the Transportation Research Board 1880, 126–134.

Ohio Supercomputer Center, 1987. Ohio Supercomputer Center. http://osc.edu/ark:/19495/f5s1ph73.

Sener, I. N., Ferdous, N., Bhat, C. R., Reeder, P., October 2009. Tour-Based Model Development for TxDOT: Evaluation and Transition Steps. Tech. Rep. FHWA/TX-10/0-6210-2, Center for Transportation Research at The University of Texas at Austin, Austin, Texas.

Shi, Q., Zheng, X., July 2014. Electric Vehicle Charging Stations Optimal Location Based on Fuzzy C-means Clustering. Applied Mechanics and Materials 556-562, 3972–3975.

Tan, J., Lin, W.-H., 8-11 October 2014. A Stochastic Flow Capturing Location and Allocation Model for Siting Electric Vehicle Charging Stations. In: 2014 IEEE 17th International Conference on Intelligent Transportation Systems (ITSC). Institute of Electrical and Electronics Engineers, Qingdao, China, pp. 2811–2816.

Upchurch, C., Kuby, M., November 2010. Comparing the p-median and flow-refueling models for locating alternative-fuel stations. Journal of Transport Geography 18, 750–758.

Wang, Y.-W., Lin, C.-C., September 2009. Locating road-vehicle refueling stations. Transportation Research Part E: Logistics and Transportation Review 45, 821–829.

Wu, C., Mohsenian-Rad, H., Huang, J., March 2012. Vehicle-to-Aggregator Interaction Game. IEEE Transactions on Smart Grid 3, 434–442.

Xi, X., Sioshansi, R., May 2014. Using Price-Based Signals to Control Plug-in Electric Vehicle Fleet Charging. IEEE Transactions on Smart Grid 5, 1451–1464.

Xi, X., Sioshansi, R., Marano, V., July 2013. Simulation-optimization model for location of a public electric vehicle charging infrastructure. Transportation Research Part D: Transport and Environment 22, 60–69.

Yudovina, E., Michailidis, G., March 2015. Socially Optimal Charging Strategies for Electric Vehicles. IEEE Transactions on Automatic Control 60, 837–842.