

Incorporating Driving Range Variability in Network Design for Refueling Facilities

Econometric Institute Report 2016-19

Harwin de Vries^{a,*}, Evelot Duijzer^a

^a*Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands*

Abstract

To stimulate and facilitate the use of alternative-fuel vehicles, it is crucial to have a network of refueling or recharging stations in place that guarantees that vehicles can reach (most of) their destinations without running out of fuel. Because initial investments in these stations are restricted, it is important to choose their locations deliberately. A fast growing stream of literature therefore analyzes the problem of locating refueling or recharging stations. The models proposed in these studies assume that the driving range is fixed, although reality shows that the driving range is highly stochastic. These models thereby misrepresent the actual coverage a network of refueling stations provides to drivers. This paper introduces two problems that do take the stochastic nature of the driving range into account. We first introduce the Expected Flow Refueling Location Problem, which is to maximize the expected number of drivers who can complete their trip without running out of fuel. The Chance Constrained Flow Refueling Location Problem is to maximize the number of drivers for which the probability of running out of fuel is below a certain threshold. We prove the problems to be strongly NP-hard, propose novel mixed-integer programming formulations for these problems, and show how these models can be extended to the case that the driving range varies during a trip. Furthermore, we extensively analyze and compare our models using randomly generated problem instances and a real life case study about the Florida state highway network. Our results show that taking the stochastic nature of the driving range into account can substantially improve the network coverage, that optimal solutions are highly robust with respect to data impreciseness, and that the potential gains of stochastic models heavily depend on the driving range distribution. Based on the results, we discuss policy implications.

Keywords: facility location, stochastic models, recourse model, chance constraint, flow refueling, electric vehicle

*Corresponding author

Email addresses: hdevries@ese.eur.nl (Harwin de Vries), duijzer@ese.eur.nl (Evelot Duijzer)

1. Introduction

The increase in oil prices and concerns about global warming have increasingly motivated the development of alternative-fuel vehicles, which, for instance, use hydrogen, ethanol, biodiesel, natural gas, or electricity as a source of energy. The number of car manufacturers presenting hybrid vehicles is going up and also pure or all-electric vehicles are more and more becoming popular. To facilitate and to stimulate this development, it is crucial to have a network of facilities in place that guarantees that vehicles can reach (most of) their destinations without any problems (i.e., without running out of fuel). Particularly in the first period after the introduction of a new type of alternative-fuel vehicle, investments in these facilities are scarce. There is little opportunity of making money on them, as the pool of potential customers is still relatively small (Romm, 2006).

To overcome this “chicken and egg problem”, governments, car manufacturers and other companies make joint efforts to establish an initial network of refueling stations that satisfies the basic needs of potential alternative-fuel vehicle users. For example, Tesla establishes a network of superchargers in Northern-America and Europe, to guarantee that the most important routes are covered sufficiently (TeslaMotors, 2015). As initial budget to place these stations is restricted, it is very important to choose the locations of new stations deliberately.

The problem of choosing the locations of refueling stations has therefore attracted considerable attention in the past few years (see e.g. Kuby and Lim, 2005; MirHassani and Ebrazi, 2012; Capar and Kuby, 2012; Capar et al., 2013). These studies model this location problem as a flow coverage problem, where a flow represents a population of electric vehicle (or, more generally, alternative-fuel vehicle) users travelling from the same origin to the same destination. Such flow is defined to be “covered” or “refueled” if the driving distance (or time) between each pair of consecutively passed refueling stations does not exceed the driving range of the vehicle. The problem of maximizing the total amount of flow that is refueled is referred to as the Flow Refueling Location Problem (FRLP) and can be solved efficiently by means of standard software (Capar and Kuby, 2012; MirHassani and Ebrazi, 2012).

The FRLP belongs to the class of facility location problems. Classical facility location problems assume non-moving demand units and a static location of the facilities. This latter assumption is relaxed in dynamic facility location problems, which allow facilities to be relocated (cf. Arabani and Farahani, 2012). The assumption of non-moving customers has been relaxed by the class of flow interception facility location problems (FIFLPs), which aim to locate facilities that capture or intercept customers along their origin-destination paths (see e.g. Hodgson, 1990; Berman et al., 1992). Examples of flow interception problems include the positioning of billboards (Averbakh and Berman, 1996), roadside healthcare facilities (De Vries et al., 2014) and refueling locations (e.g., the FRLP). Berman and Krass (1998) extend the flow interception problem by accounting for competition among different facilities.

Whereas these classical flow interception problems assume that customers follow certain pre-planned paths, Kim and Kuby (2012) formulate the location problem in which drivers are

willing to deviate from their preferred paths to refuel their vehicles. This problem is referred to as the *deviation* flow refueling location problem. Yıldız et al. (2016) propose a branch and price approach to solve this problem, which significantly reduces the computation time. Finally, whereas the FRLP and the deviation FRLP mainly focus on enabling long-distance trips, Kang and Recker (2014) optimize locations with respect to short-distance traffic, for which the driving range is assumed to be of no limitation. Instead, the authors propose models to locate the facilities based on household scheduling and routing considerations.

The flow refueling location problems discussed so far implicitly assume that the driving range of a vehicle is fixed and known in advance. However, reality shows that the driving range is highly variable. For instance, it is dependent on the age of the battery, the temperature, the amount of traffic on the road, and the driving style (Ehsani et al., 2009; Musk and Straubel, 2012; Dong and Lin, 2014). Therefore, regarding the driving range as fixed can significantly misrepresent the coverage level provided by a network of refueling stations, and potentially lead to location choices that are far from optimal in reality. Lee et al. (2014) make a first attempt to include stochasticity of the driving range into the location problem, by assuming a randomly distributed battery load at the beginning of a trip from origin to destination. The authors, however, unrealistically assume that the driving range is sufficient to cover all origin-destination combinations, such that recharging is needed at most once during a trip. Simulated annealing is used to solve the location problem for a small network.

In this paper, we investigate two ways to incorporate the stochastic nature of the vehicle driving range into the problem of locating refueling stations. We first propose a new formulation of the FRLP, which contains the driving range explicitly as a parameter (in contrast with existing formulations). Using this formulation as a starting point, we first introduce and model the Expected Flow Refueling Location Problem, which is to maximize the expected number of drivers who can complete their trip without running out of fuel. Although this model provides a natural way to deal with stochasticity, it does not consider the coverage levels provided to each of the flows separately. As a consequence, it prefers to provide two equally sized flows with 51% coverage over providing one flow with 100% coverage. This might be easily justifiable in the context of hybrid (electric) vehicles, which can switch to a different power source when the primary source is exhausted. However, since drivers of, for instance, pure electric vehicles or hydrogen vehicles will strongly dislike the large probability of running out of fuel, this solution will be far from optimal in their context. We therefore also introduce and model the Chance Constrained Flow Refueling Location Problem, which is to maximize the number of drivers for which the probability of completing their trip without of running out of fuel is at least $1 - \alpha$. We numerically analyze these models using randomly generated networks (see Capar and Kuby, 2012) and a real life case study on the Florida state highway network (see Kuby et al., 2009; Capar et al., 2013).

The remainder of this paper is organized as follows. Section 2 describes our reformulation of the FRLP. The stochastic model formulations are introduced in Section 3, and Section 4 describes our numerical results. Finally, in Section 5 we draw conclusions and state opportunities for future research.

2. Reformulation of the Flow Refueling Location Problem (FRLP)

The Flow Refueling Location Problem can be described as follows. Consider a network $G(L, E)$ where L denotes a set of locations and E a set of arcs between these locations. The set of locations L is the union of the following three sets: the set of potential facility locations, K , the set of origins, O , and the set of destinations, D . Consider a collection of drivers who travel along this network. In line with the literature (Kuby and Lim, 2005; Capar and Kuby, 2012) we use the term *flow* to refer to the subset of drivers that travels from the same origin to the same destination along the same path. Let F denote the set of all the flows and let O_f and D_f denote respectively the origin and destination of the drivers along flow f . The path travelled by the drivers in flow f is an ordered sequence of edges $e \in E$ connecting the following vertices: the start vertex $O_f \in O$, an ordered subset of potential facility locations, $K_f \subseteq K$, and the end vertex, $D_f \in D$. The volume of flow f , i.e., the number of drivers that travels from O_f to D_f , is denoted by v_f . The vehicles have a driving range of R (miles/kilometers), and can be refueled/recharged at refueling facilities along their route. We consider a flow to be *covered* if vehicles can travel from their origin to their destination and back to their origin without running out of fuel. The objective of the FRLP is to locate p facilities in the network, so as to maximize the number of drivers covered.

Existing models for the FRLP include the driving range restriction implicitly by representing the coverage of a flow or arc by means of binary parameters (Kuby and Lim, 2005; Capar and Kuby, 2012; Capar et al., 2013; MirHassani and Ebrazi, 2012). We propose a new MIP formulation that explicitly contains the driving range. Before we propose our model, let us introduce some notations. We use the binary decision variable x_k to indicate whether a facility is placed at potential facility location k ($x_k = 1$) or not ($x_k = 0$). For sake of simplicity, we assume that there are currently no facilities in the network. The model can easily be adapted in the case that there are. Furthermore, we use auxiliary variables y_f to indicate whether flow f is covered (i.e., refueled: $y_f = 1$) or not ($y_f = 0$).

In what follows we will use a slightly different vector of vertices to represent a path:

$$\pi_f = [\pi_f(1), \pi_f(2), \dots, \pi_f(n-1), \pi_f(n)]$$

Here, $\pi_f(1)$ and $\pi_f(n)$ represent the origin vertex O_f and the destination vertex D_f , respectively. The vector $[\pi_f(2), \dots, \pi_f(n-1)]$ is the sequence of locations $k \in K_f$ with a new facility passed during a trip from O_f to D_f . We call these locations new facility locations. Hence, this representation does not necessarily include all $k \in K_f$. Note that π_f depends on the decision variables x_k for all $k \in K_f$, because these variables determine the new facility locations. We make the following assumption about this path:

Assumption 1. *A driver in flow f cyclically travels from O_f to D_f to O_f to D_f et cetera, as follows: $[\pi_f(1), \pi_f(2), \dots, \pi_f(n-1), \pi_f(n), \pi_f(n-1), \dots, \pi_f(2), \pi_f(1), \pi_f(2), \dots]$*

We assume that drivers do not deviate from their route to visit another new facility location and we assume that there is always sufficient capacity at a refueling facility. We refer

to Section 5 for a discussion of these assumptions. Furthermore, we assume that a refueling facility at location k can be reached by traffic from both directions. This assumption is motivated by the fact that recharging facilities are generally located either on both sides of a highway or close to an exit of a highway, so that they are easily accessible for traffic from both directions.

Note that, in case that $|\pi_f| = 2$, there are *no* new facility locations along the path of flow f . Let us consider the case that $|\pi_f| \geq 3$. By Assumption 1 we know that a driver in flow f always returns to $\pi_f(1)$ after reaching $\pi_f(n)$, and vice versa. We refer to the path from $\pi_f(1)$ via $\pi_f(n)$ to $\pi_f(1)$ as the *flow f cycle*. This cycle could be regarded as a set of trips between successively passed new facility locations. This set of trips is given by: $\{\pi_f(h) \rightarrow \pi_f(h+1) | 2 \leq h \leq n-2\} \cup \{\pi_f(n-1) \rightarrow \pi_f(n) \rightarrow \pi_f(n-1)\} \cup \{\pi_f(h+1) \rightarrow \pi_f(h) | 2 \leq h \leq n-2\} \cup \{\pi_f(2) \rightarrow \pi_f(1) \rightarrow \pi_f(2)\}$. We refer to these trips as *cycle segments*.

Definition. A *cycle segment* of the flow f cycle is identified by two vertices k and l in the following cases:

- k is the origin and l is the first new facility location that is visited when travelling from the origin to the destination. Then vertices $k = O_f$ and $l \in K_f$ identify the cycle segment $l \rightarrow O_f \rightarrow l$.
- k is the last new facility location that is visited before reaching the destination l . Then vertices $k \in K_f$ and $l = D_f$ identify the cycle segment $k \rightarrow D_f \rightarrow k$.
- k and l are two new facility locations that are visited consecutively. Then vertices $k \in K_f$ and $l \in K_f$ identify the cycle segment $k \rightarrow l$.

Figure 1 illustrates the definition of the cycle segments.

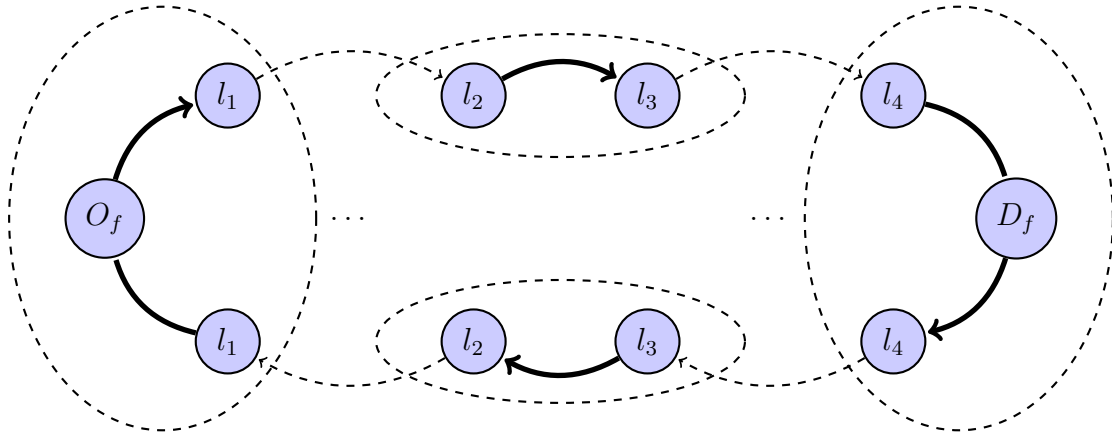


Figure 1: An illustration of the three types of cycle segments of the flow f cycle: the cycle segment identified by vertex l_1 and the origin ($l_1 \rightarrow O_f \rightarrow l_1$), the cycle segments identified by vertices l_2 and l_3 ($l_2 \rightarrow l_3$, $l_3 \rightarrow l_2$) and the cycle segment identified by vertex l_4 and the destination ($l_4 \rightarrow D_f \rightarrow l_4$).

We note that a flow f is covered if and only if the driving range exceeds the length of each cycle segment corresponding to this flow. The length of a cycle segment is represented

by the parameters t_{kl} , defined for any $k, l \in \{O_f \cup K_f \cup D_f\}$. For $k \in K_f$ and $l \in K_f$, this parameter denotes distance between locations k and l , for $k = O_f$ and $l \in K_f$ this parameter represents the distance from l , via O_f to l , and for $l = D_f$ and $k \in K_f$ this parameter represents the distance from k , via D_f to k . For later use we define $t_{O_f D_f} = \infty$ for all $f \in F$.

In the model formulation we use L_f to denote the total set of locations in the path of flow f (i.e., $L_f = \{O_f \cup K_f \cup D_f\}$) and L_{kf} denotes the set of locations that are passed *after* passing location k during a trip from O_f to D_f (including D_f). We use the binary variables i_{klf} to indicate whether locations k and l identify a cycle segment of the flow f cycle ($i_{klf} = 1$) or not ($i_{klf} = 0$). Observe that it is sufficient to define these variables only for $l \in L_{kf}$ by Assumption 1 and the definition of the distance parameters t_{kl} . Finally, M denotes a large number (e.g., $2 \cdot \max_{f \in F} \{\max_{k \in K_f} \{t_{O_f k} + t_{k D_f}\}\}$).

We summarize our notations in the following table:

<i>Sets and indices</i>		<i>Parameters (continued)</i>	
L	set of locations, indexed by k, l	R	driving range
E	set of arcs between locations	O_f	origin of flow f
K	set of potential facility locations	D_f	destination of flow f
O	set of origins	t_{kl}	length of the cycle segment identified by locations k and l
D	set of destinations	<i>Decision variables</i>	
F	set of flows, indexed by f	x_k	binary variable; 1 if a facility is placed at location k and 0 if not
K_f	set of potential facility locations along flow f	y_k	binary variable; 1 if flow f is covered and 0 if not
L_f	set of locations along flow f	i_{klf}	binary variable; 1 if locations k and l identify a cycle segment of the flow f cycle and 0 if not
L_{kf}	set of locations along flow f passed after location k on a trip from O_f to D_f		
<i>Parameters</i>			
p	number of new facilities to locate		

Using these notations, the FRLP can be formulated as the following MIP model:

$$\max \quad \sum_{f \in F} v_f y_f \quad (1)$$

$$\text{s.t.} \quad \sum_{k \in K} x_k = p \quad (2)$$

$$\sum_{l \in L_f} i_{klf} t_{kl} - (1 - y_f) M \leq R \quad f \in F, k \in \{O_f \cup K_f\} \quad (3)$$

$$\sum_{l \in L_f} i_{klf} = x_k \quad f \in F, k \in K_f \quad (4)$$

$$\sum_{l \in L_f} i_{klf} = 1 \quad f \in F, k \in O_f \quad (5)$$

$$\sum_{k \in L_f} i_{klf} = x_l \quad f \in F, l \in K_f \quad (6)$$

$$\sum_{k \in L_f} i_{klf} = 1 \quad f \in F, l \in D_f \quad (7)$$

$$i_{klf} \in [0, 1] \quad f \in F, k \in L_f, l \in L_{kf} \quad (8)$$

$$x_k, y_f \in \mathbb{B} \quad f \in F, k \in K \quad (9)$$

The objective function (1) maximizes the total flow volume covered. Constraint (2) restricts the number of new facilities that can be placed. Constraint (3) stipulates that a flow is covered if and only if the driving range exceeds the length of each of the corresponding cycle segments. The decision variables are defined in constraints (8) - (9). Finally, By Lemma 1 constraints (4) - (8) stipulate that variables i_{klf} take the values as defined above. Observe that constraints (8) define the variables i_{klf} to be continuous variables on $[0, 1]$.

Lemma 1. *Constraints (4) - (8) ensure that $i_{klf} = 1$ if and only if locations k and l identify a cycle segment of the flow f cycle and that $i_{klf} = 0$ otherwise.*

Proof. We refer to De Vries et al. (2014) for the proof of this lemma. \square

To distinguish the problem described by (1) - (9) from the stochastic variants introduced next, we refer to this problem as the Deterministic Flow Refueling Location Problem (DFRLP) from now on.

3. Stochastic Models

This section describes two ways to deal with a stochastic driving range. Without loss of generality, the driving range of a vehicle, denoted by $R(\omega)$, is defined as some function of the random variable or vector of random variables ω . This can be thought of as a set random conditions – like weather, traffic, driving style, or battery condition – influencing the driving range. We consider the same problem as in Section 2, but now assume that coverage depends on the random variables ω . Coverage thus becomes a matter of chance rather than

a binary observation, as assumed in the DFRLP. In line with this, we use binary variables y_f^ω to indicate whether or not a driver along flow f can perform its trip without running out of fuel under random conditions ω . In Section 3.1 we model and analyze the problem to maximize the *expected* flow volume covered as a two-stage recourse model. Section 3.2 describes and analyzes a model that handles driving range variability by means of chance constraints.

We make the following assumptions about the driving range:

Assumption 2. *Given the realization of ω , the driving range, $R(\omega)$, is the same at each cycle segment traversed by vehicle in flow f .*

Assumption 3. *$R(\omega)$ is randomly distributed with cumulative density function (CDF) $\mathcal{G} : \mathbb{R}_+ \mapsto [0, 1]$.*

We refer to Section 3.1.1 and Section 3.2 for a discussion on relaxing these assumptions for the two stochastic models, respectively.

3.1. Expected Flow Refueling Location Problem (EFRLP)

Let a vehicle be covered if and only if the driving distance between each pair of consecutively passed refueling stations does not exceed the driving range of the vehicle. The problem of locating refueling stations can be regarded as a two-stage stochastic optimization problem. Using the terminology from the field of stochastic optimization, the first-stage decisions are to locate the facilities (i.e., on x_k), whereas the second stage “decisions” are whether or not to cover a vehicle in flow f (i.e., on y_f^ω), given a realisation of its driving range. $Q_f(\mathbf{x}, \mathbf{i})$ denotes the expected “value” of the second stage decisions corresponding to flow f : the flow volume times the probability that a vehicle in that flow is covered. Here, $\mathbf{x} = \{x_k\}$ and $\mathbf{i} = \{i_{klf}\}$. Using these notations, the problem of maximizing the *expected* flow volume covered can be described as:

$$\max \sum_{f \in F} Q_f(\mathbf{x}, \mathbf{i}) \quad (10)$$

$$\text{s.t.} \quad \sum_{k \in K} x_k = p \quad (11)$$

$$\begin{aligned} & (3) - (8) \\ & x_k \in \mathbb{B} \quad k \in K \end{aligned} \quad (12)$$

The expected value function $Q_f(\mathbf{x}, \mathbf{i})$ is defined as follows:

$$Q_f(\mathbf{x}, \mathbf{i}) = \mathbb{E}_\omega \left[\max_{y_f^\omega \in \{0,1\}} \{v_f y_f^\omega\} \right] \quad (13)$$

$$\text{s.t.} \quad \sum_{l \in L_f} i_{klf} t_{kl} - (1 - y_f^\omega) M \leq R(\omega) \quad k \in \{O_f \cup K_f\} \quad (14)$$

The objective (10) is to maximize the *expected flow volume covered*. Constraint (11) stipulates that exactly p new facilities are located. Constraints (3) - (8) construct the cycle segments and constraints (12) define the decision variables. Finally, constraints (14) stipulate that a vehicle in flow f is covered if and only if the driving range is sufficient to traverse all segments of the flow f cycle.

Let z_f denote the probability that a vehicle in flow f is covered, also referred to as the coverage level of flow f . By Assumption 2 z_f equals the probability that the driving range exceeds the length of the *largest* cycle segment. The latter equals $\max_{(k,l)}\{i_{klf}t_{kl}\}$, so that the value of z_f is obtained by:

$$z_f = \mathcal{P} \left(R(\omega) \geq \max_{(k,l)}\{i_{klf}t_{kl}\} \right) \quad (15)$$

$$= 1 - \mathcal{G} \left(\max_{(k,l)}\{i_{klf}t_{kl}\} \right) \quad (16)$$

$$= 1 - \max_{(k,l)}\{i_{klf}\mathcal{G}(t_{kl})\} \quad (17)$$

Equation (17) holds because variables i_{klf} take the value 0 or 1, and because $\mathcal{G}(0) = 0$ (see Assumption 3). By (17) the value of $Q_f(\mathbf{x}, \mathbf{i})$ can be obtained with the following small MIP problem. Here, parameters g_{kl} denote $\mathcal{G}(t_{kl})$; the probability that the driving range is smaller than t_{kl} .

$$Q_f(\mathbf{x}, \mathbf{i}) = \max \quad v_f z_f \quad (18)$$

$$\text{s.t.} \quad z_f \leq 1 - \sum_{l \in L_f} i_{klf} g_{kl} \quad k \in \{O_f \cup K_f\} \quad (19)$$

$$z_f \geq 0 \quad (20)$$

For a given location $k \in L_f$, the right-hand side of (19) denotes the probability that the driving range exceeds the length of the segment identified by locations k and l for which $i_{klf} = 1$. Hence, z_f takes the value of the smallest probability among the cycle segments. In case that *no new facility location* is located along the path of flow f , Constraints (3) - (8) ensure that $i_{O_f D_f f} = 1$. In this case z_f equals 0, because $t_{O_f D_f f} = \infty$ by definition and thus $g_{O_f D_f f} = 1$. Substituting equations (18) - (20) in (10) - (12) yields the following model:

$$\max \sum_{f \in F} v_f z_f \quad (21)$$

$$\text{s.t.} \quad \sum_{k \in K} x_k = p \quad (22)$$

(3) – (8)

$$z_f \leq 1 - \sum_{l \in L_f} i_{klf} g_{kl} \quad f \in F, k \in \{O_f \cup K_f\} \quad (23)$$

$$z_f \geq 0 \quad f \in F \quad (24)$$

$$x_k \in \mathbb{B} \quad k \in K \quad (25)$$

We refer to the problem described by (21) - (25) as the Expected Flow Refueling Location Problem (EFRLP).

3.1.1. Non-homogeneous driving range

The EFRLP assumes that the driving range distribution is the same for all cycle segments in the flow f cycle (Assumption 2) and for every flow (Assumption 3). In this section we relax these assumptions.

Driving range per flow. Due to differences in speed restrictions and traffic volumes, it is likely that some flows are more crowded than others and thus that the driving range distribution differs per flow. Because of the separability of the second stage in the recourse formulation, this can easily be incorporated. Denote by $R_f(\omega)$ the driving range for flow f . Let $R_f(\omega)$ be randomly distributed with cumulative density function (CDF) $\mathcal{G}_f : \mathbb{R}_+ \mapsto [0, 1]$ and let g_{klf} denote $\mathcal{G}_f(t_{kl})$. The correct formulation for this case can easily be derived from the formulation for the EFRLP by replacing g_{kl} by g_{klf} in constraints (23).

Driving range per segment; complete correlation. Assumption 2 can be relaxed as follows. Let (k, l) be the locations that identify a segment in the flow f cycle. Furthermore, let $R_{(klf)}(\omega)$ denote the random driving range on this segment, with corresponding CDF $\mathcal{G}_{klf} : \mathbb{R}_+ \mapsto [0, 1]$. Next, we make the following assumption for any pair of segments of the flow f cycle, identified by the locations (k_1, l_1) and (k_2, l_2) :

Assumption 4. $R_{k_1 l_1 f}(\omega)$ and $R_{k_2 l_2 f}(\omega)$ are completely correlated. That is, for a given realization of ω , $\mathcal{G}_{k_1 l_1 f}(R_{k_1 l_1 f}(\omega)) = \mathcal{G}_{k_2 l_2 f}(R_{k_2 l_2 f}(\omega))$.

Assumption 4 is a generalization of Assumption 2 and is able to capture more realistic scenarios. It is to be expected that the average driving range differs in different parts of the network (e.g., because speed restrictions and the amount of congestion differ), but also that variations in the driving range in different parts of the network are correlated. For instance, during cold periods, the driving range will be relatively low in all parts of the network. The EFRLP can easily be adjusted to the case that Assumption 4 holds. In this case, the probability that a vehicle in flow f is covered equals the smallest probability that the driving range exceeds the length of a segment among all segments of the flow f cycle. Hence, let \tilde{g}_{klf} be $\mathcal{G}_{klf}(t_{kl})$. Then it suffices to replace g_{kl} by \tilde{g}_{klf} in constraints (23).

Driving range per segment; no complete correlation. Assumption 2 could also be relaxed to the case that the driving ranges at different segments of the flow f cycle are not or not completely correlated. This relaxation influences the calculation of the coverage probability for a flow f , as this probability is now calculated by means of the joint probability function of the driving ranges corresponding to the different segments. Obviously, this makes the model highly non-linear in general. We propose to deal with this non-linearity by using Bonferroni's inequality (Bonferroni, 1936) as an approximation for the joint probability. In our context, this inequality states the following. Let event E_j denote the event that the driving range *exceeds* the length of flow f segment $j \in \{1, \dots, m\}$, and let $P(E_j)$ denote the probability that this occurs. The probability that a vehicle in flow f is covered is then represented by $P(\cap_{j=1}^m E_j)$. Then Bonferroni's inequality states that:

$$P(\cap_{j=1}^m E_j) \geq 1 - \sum_{j=1}^m [1 - P(E_j)] \quad (26)$$

Hence, let z_f again represent the probability that a vehicle in flow f is covered. Then the expected value function is approximated by:

$$Q_f(\mathbf{x}, \mathbf{i}) = \max \quad v_f z_f \quad (27)$$

$$z_f = 1 - \sum_{k \in L_f} \sum_{l \in L_f} i_{klf} a_{klf} (1 - g_{kl}) \quad (28)$$

$$z_f \geq 0 \quad (29)$$

Here parameter a_{klf} equals 2 if $k \in K_f$ and $l \in K_f$, and equals 1 otherwise. This ensures that a segment $k \rightarrow l$ that is traversed twice during a trip at the flow f cycle (i.e., during the trip from O_f to D_f and during the return trip) is also counted twice. We did not need to consider this aspect so far, because the other problem variants consider the segments separately. Constraint (28) approximates z_f by setting it equal to Bonferroni's lower bound (Bonferroni, 1936). Note that, for a given $k \in L_f$, the summation $\sum_{l \in L_f} i_{klf} (1 - g_{kl})$ denotes the probability that the driving range exceeds the length of the segment identified by the locations k and l for which $i_{klf} = 1$.

We observe that this case generalizes the case of a driving range distribution *per arc*. Since a segment simply represents a sequence of arcs, the segment fuel consumption (which yields the driving range) has a joint probability distribution with the fuel consumption of the arcs as variables. Hence, driving range distributions per arc translates into a driving range distribution per segment.

3.1.2. Value of the Stochastic Solution

Although deterministic problems are often easier to solve, including stochasticity is more realistic. In this section we provide ways to evaluate the benefits of using a stochastic driving range. Let \bar{Q} be the optimal value of the EFRLP. Furthermore, let $Y_f(\mathbf{x}, \mathbf{i}, R(\omega)) \subseteq \{0, 1\}$

be the set of all feasible values for y_f , given the variables \mathbf{x} and some realization of $R(\omega)$, as defined by constraint (14) (note that the values of the variables \mathbf{i} are implicitly determined by the decision variables \mathbf{x}). Then \bar{Q} can be written as:

$$\bar{Q} := \max_{\mathbf{x} \in \mathbb{B}} \mathbb{E}_\omega [h(R(\omega), \mathbf{x}, \mathbf{i})] \quad (30)$$

Here, the function $h(R(\omega), \mathbf{x}, \mathbf{i})$ is defined as follows:

$$h(R(\omega), \mathbf{x}, \mathbf{i}) = \sum_{f \in F} \max\{v_f y_f \mid y_f \in Y_f(\mathbf{x}, \mathbf{i}, R(\omega))\}$$

Instead of using the distribution function of the driving range we could (naively) solve the EFRLP for the expected value of the driving range, which is equivalent to assuming that $Pr(R(\omega) = \mathbb{E}_\omega[R(\omega)]) = 1$. Let Q_{EV} denote the value of the solution to this expected value problem given in (31) and let $(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{i}})$ be the corresponding solution. Filling in this solution in the original problem provides us with the expected result of the expected value problem, denoted by Q_{EEV} , given in (32).

$$Q_{EV} := \max_{\mathbf{x} \in \mathbb{B}} h(\mathbb{E}_\omega[R(\omega)], \mathbf{x}, \mathbf{i}) = h(\mathbb{E}_\omega[R(\omega)], \bar{\mathbf{x}}, \bar{\mathbf{i}}) \quad (31)$$

$$Q_{EEV} := \mathbb{E}_\omega[h(R(\omega), \bar{\mathbf{x}}, \bar{\mathbf{i}})] \quad (32)$$

Obviously, Q_{EEV} provides a lower bound for \bar{Q} , since $(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{i}})$ is a feasible solution. This lower bound allows us to calculate the so-called Value of the Stochastic Solution $VSS = \bar{Q} - Q_{EEV}$: the costs you make by using the naive expected value solution instead of considering the stochasticity in the variables (Birge, 1982). In Section 4 we use the VSS to evaluate our stochastic models.

3.2. Chance Constrained Flow Refueling Location Problem (CCFRLP)

This section provides an alternative way to deal with the stochasticity in the driving range parameters. The proposed model defines a flow to be covered if and only if the probability that a trip along the flow f cycle can be made without running out of fuel is at least $1 - \alpha$. Hence, the probability of running out of fuel at such segment is at most α . Under Assumption 2, the probability of running out of fuel equals the smallest probability that the driving range exceeds the length of a segment among all cycle f segments. Again, let $\mathcal{G} : \mathbb{R}_+ \mapsto [0, 1]$ denote the CDF of the driving range, and let $g_{kl} = \mathcal{G}(t_{kl})$. Using these notations, the problem of maximizing the flow volume covered can be described as:

$$\max \sum_{f \in F} v_f y_f \quad (33)$$

$$\text{s.t.} \quad \sum_{k \in K} x_k = p \quad (34)$$

$$\sum_{l \in L_f} i_{klf} g_{kl} \leq \alpha + (1 - y_f) \quad f \in F, k \in \{O_f \cup K_f\} \quad (35)$$

(3) – (8)

$$x_k \in \mathbb{B} \quad k \in K \quad (36)$$

$$y_f \in \mathbb{B} \quad f \in F \quad (37)$$

The objective function maximizes the total flow volume covered, which we refer to as the *chance constrained flow volume covered*. Constraint (34) restricts the number of new facilities to place. Constraints (35) stipulate that flow f is covered if and only if the probability of running out of fuel is at most α for each of the segments of the flow f cycle. Again, constraints (3) - (8) stipulate that variables i_{klf} take the values as defined before. Finally, the decision variables are defined in constraints (36) - (37). We refer to the problem described in (33) - (37) as the Chance Constrained Flow Refueling Location Problem (CCFRLP).

Note that the CCFRLP model described in (33) - (37) bears a close resemblance to the EFRLP model described in (21) - (25). The only differences are that the latter sets $\alpha = 1$ and replaces binary variable y_f by continuous variable z_f . In Section 4 we will discuss the differences between the models in detail.

Non-homogeneous driving range. The CCFRLP model can be easily adapted when assumptions 2 and 3 are relaxed to the assumption of a completely correlated driving range per segment: it suffices to replace g_{kl} by \tilde{g}_{klf} in constraint (35) (see Section (3.1.1)). When allowing the driving ranges in different cycle segments to be not or not completely correlated, the proposed model needs to be adapted. We again propose to approximate the (joint) probability that a trip at the flow f cycle can be made without running out of fuel by means of Bonferroni's inequality. Section 3.1.1 shows that the resulting bound on this probability is given by $1 - \sum_{k \in L_f, l \in L_f} i_{klf} a_{klf} (1 - g_{kl})$. Hence, the location problem with joint chance constraints can be formulated by replacing constraint (35) in the CCFRLP model by:

$$\sum_{k \in L_f, l \in L_f} i_{klf} a_{klf} (1 - g_{kl}) \leq \alpha + (1 - y_f) \quad f \in F \quad (38)$$

4. Impact Analysis

This section numerically analyzes the performance of the models introduced in the previous sections and the impact of including stochasticity into the location problem. Specifically, Section 4.1 describes our problem instances and Sections 4.2-4.6 compare the models, analyze their sensitivity with respect to several types of data, and analyze their computational

performance. We use CPLEX 12.61 as the solver engine, on a PC with a 3.4 AMD A4-5300B APU processor and 4 GB RAM. Throughout this section, we apply the EFRLP without the multiple driving range extension and the CCFRLP with individual chance constraints. To determine the Value of the Stochastic Solution (VSS , see Section 3.1.2), we make use of the DFRLP model from Section 2 and use $\mathbb{E}_\omega[R(\omega)]$ as the value of the parameter R .

4.1. Problem instances

We test our models on the Florida state highway network described by Kuby et al. (2009) and on networks we randomly generate based on (a slightly adapted version of) the method proposed by Capar and Kuby (2012). Appendix A describes the network generation in detail and Table 4.1 describes our instances. We refer to a randomly generated instance with X potential facility locations and Y Origin-Destination nodes (corresponding to $Y(Y - 1)/2$ flows f) as $sXwY$, and refer to the instance $s80w40$ as the baseline case. Finally, we set $\alpha = 0.05$ and use the Gamma distribution for the driving range, motivated by its flexibility. The values of the shape parameter κ and scale parameter θ shape we use in the baseline situation are 50 and 5, respectively, such that $\mathbb{E}_\omega[R(\omega)] = 250$ in accordance with Capar and Kuby (2012).

Instance name	$ K $	$ F $	\bar{T}	V
Florida	302	2701	364	1.0E12
s100w50	100	1225	834	1.0E6
s80w40	80	780	822	1.0E6
s60w30	60	435	1153	1.0E6
s40w20	40	190	908	1.0E6

Note: \bar{T} : average travel distance from origin to destination per flow, V : total vehicle volume $\sum_f v_f$

Table 4.1: Characteristics of problem instances analyzed.

4.2. Comparison of the models

This section compares solutions of the EFRLP, the CCFRLP, and the DFRLP. For a given problem instance and a given value of p , we solve each of the three problems. The three solutions are then analyzed in terms of expected flow volume covered and chance constrained flow volume covered. Obviously, the EFRLP solution and the CCFRLP solution are optimal in terms of expected flow volume covered and chance constrained flow volume covered, respectively. The difference between the EFRLP and the DFRLP solutions in terms of expected flow volume covered yields the Value of the Stochastic Solution (VSS , see Section 3.1.2). We repeat this for each $p \in \{1, 2, 3, 4, 5, 10, 15, 20, 25\}$. Table 4.2 describes the resulting average VSS and average optimality gaps (i.e. the average over the VSS or gaps obtained for the 9 values of p). Appendix B describes the coverage statistics for each individual solution obtained, and Figures 2 and 3 illustrate these for the baseline case $s80w40$.

Instance	Opt. gap in terms of expected flow volume covered (%)			Opt. gap in terms of chance constr. flow volume covered (%)			VSS
	EFRLP	CCFRLP	DFRLP	EFRLP	CCFRLP	DFRLP	
	Solution	Solution	Solution	Solution	Solution	Solution	
Florida	0.00	0.11	0.03	0.08	0.00	0.21	2.90E08
s100w50	0.00	19.94	0.77	52.85	0.00	52.64	3.96E03
s80w40	0.00	5.70	1.45	17.83	0.00	29.26	1.02E04
s60w30	0.00	4.56	3.06	7.23	0.00	28.80	5.24E03
s40w20	0.00	9.14	5.55	26.42	0.00	47.36	4.20E04

Table 4.2: Average Value of the Stochastic Solution (VSS) and average optimality gap (%) in terms of expected flow volume covered and chance constrained flow volume covered over 9 EFRLP, CCFRLP and DFRLP solutions.

The results yield several interesting observations. First, the optimality gaps for the Florida case are much smaller than those for the randomly generated cases. The reason is that the length of most of the high volume vehicle routes in the Florida case is relatively small compared to the expected vehicle range (see Table 4.1), causing the differences between the solution values assigned by the different models to be small. The next section investigates this mechanism in more detail. A second observation is that the CCFRLP solutions perform relatively well in terms of expected flow volume covered, whereas EFRLP solutions perform poorly in terms of chance constrained coverage. This can be explained by the fact that EFRLP solutions benefit from providing a given flow with *partial* coverage (e.g., 50%), so that they place relatively few facilities along the path of a given flow. As a result, they fail to reach the coverage threshold $(1 - \alpha)$ for many vehicle flows. The CCFRLP, instead, induces the incentive to place relatively many facilities along a given flow. Though, in terms of expected flow coverage, corresponding solutions overly focus on a restricted set of vehicle flows, they thereby do yield substantial coverage levels (z_f) for these flows. This effect also plays a role for the DFRLP model, but it is significantly smaller, as the critical coverage distance between two facilities used by the DFRLP is much looser than the one for the CCFRLP. Nevertheless, in terms of actual expected number of vehicles covered, the average benefit of using a stochastic model instead of a deterministic model can be very large, as shown by the Value of the Stochastic Solution (see the last column of Table 4.2). For example, choosing the EFRLP solutions to instance *s40w20* instead of the DFRLP solutions yields an average increase of 42 004 expected number of vehicles covered, representing 4.2% of the total vehicle volume.

Figures 2 and 3 illustrate that adding multiple facilities to the network can bring about *synergy effects*. For instance, when increasing p from 2 to 3, the value of the EFRLP solution increases by 44 550, whereas it increases by 66 499 when increasing p from 3 to 4. Similarly, the CCFRLP solution value increases by 46 881 when p increases from 2 to 3, and increases by 53 081 when p increases from 3 to 4. This can be explained by the fact that generally multiple facilities are needed to cover a flow (CCFRLP) or to cause a substantial increase in the coverage level for a given flow (EFRLP). Increasing the value of p by 1 can be the final

push that enables this for a given flow or a network of flows, causing a jump in the objective value.

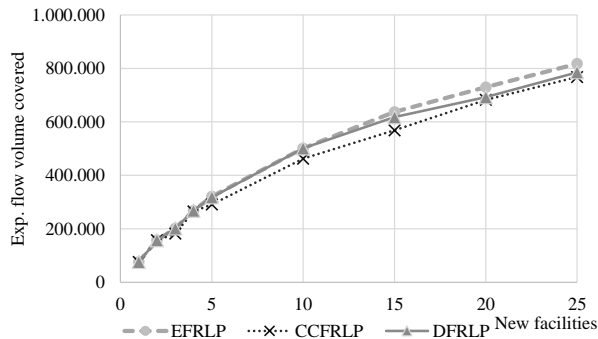


Figure 2: Expected flow volume covered for the EFRLP, CCFRLP and DFRLP solutions to the baseline case $s80w40$.

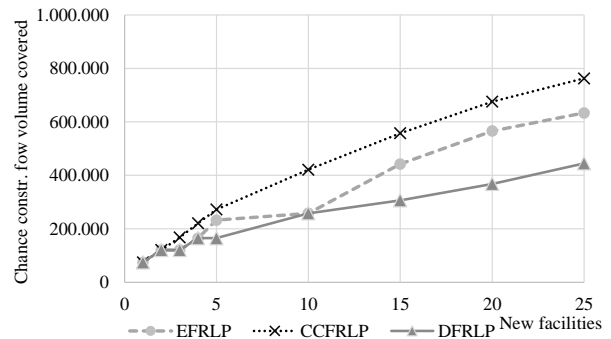


Figure 3: Chance constrained flow volume covered for the EFRLP, CCFRLP and DFRLP solutions to the baseline case $s80w40$.

4.3. Sensitivity analysis on the driving range distribution

The results shown in the previous section hint at a strong relationship between the driving range distribution and solution values. To get a better understanding of this relationship, we analyze solutions obtained for different settings of the shape parameter κ and scale parameter θ determining the driving range distribution (i.e., the Gamma distribution). First, we analyze the impact of the *expected* driving range on the solutions obtained by considering settings of low (125), medium (250), and high (500) expected driving range. Figure 4 describes the corresponding probability density functions (PDFs) and parameter values. Note that the variance is the same for each of these settings. The impact of the expected driving range on solution values for the baseline case $s80w40$ and $p = 15$ is depicted in Figures 6 and 7.

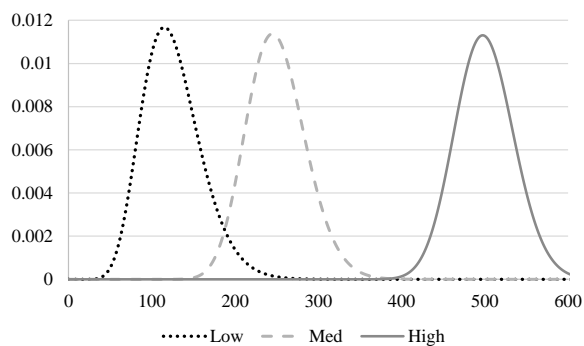


Figure 4: PDF of driving range distributions with low expected value ($\kappa = 12.5, \theta = 10, \mathbb{E}_\omega[R(\omega)] = 125$), medium expected value ($\kappa = 50, \theta = 5, \mathbb{E}_\omega[R(\omega)] = 250$), and high expected value ($\kappa = 200, \theta = 2.5, \mathbb{E}_\omega[R(\omega)] = 500$)

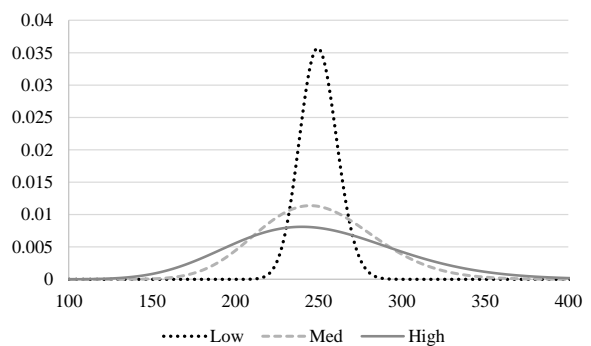


Figure 5: PDF of driving range distributions with low variance ($\kappa = 500, \theta = 0.5, \mathbb{V}_\omega[R(\omega)] = 200$), medium variance ($\kappa = 50, \theta = 5, \mathbb{V}_\omega[R(\omega)] = 1\,250$), and high variance ($\kappa = 25, \theta = 10, \mathbb{V}_\omega[R(\omega)] = 2\,500$).

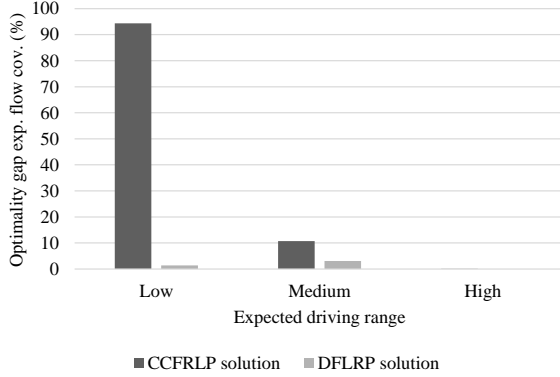


Figure 6: Optimality gap in terms of expected flow coverage for the CCFRLP and DFRLP solutions and low (125), medium (250), and high (500) expected driving range, using the baseline case $s80w40$ and $p = 15$.

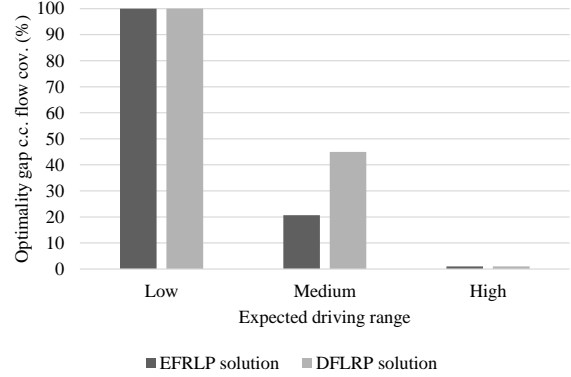


Figure 7: Optimality gap in terms of chance constrained flow coverage for the EFRLP and DFRLP solutions and low (125), medium (250), and high (500) expected driving range, using the baseline case $s80w40$ and $p = 15$.

As hypothesized, differences between the solution values tend to decrease when the expected driving range increases. The rationale underlying this trend is that the values of a given solution in the different models tend to converge when the driving range increases. For example, in case of a relatively small driving range, placing 1 facility along a given path may yield a deterministic coverage score of 1, an expected coverage score of 0.5, and a chance constrained coverage score of 0. In case of a sufficiently large driving range, however, one obtains a coverage level equal to or very close to 1, irrespective of the coverage definition used (deterministic, expected, of chance constrained).

Next, we analyze the impact of the driving range *variance* on the solutions obtained by considering settings of low variance (125), medium (1 250), and high variance (2 500). Figure 5 describes the corresponding PDFs and parameter values. Note that the expected driving range is the same for these settings. The optimality gaps for the baseline case and $p = 15$ in terms of expected flow coverage and chance constrained flow coverage are depicted in Figures 8 and 9.

We observe that a smaller driving range variance leads to smaller differences between the solutions. Again, the reason is that the value the different models assign to a given solution tend to converge when the variance decreases. Note that, in the extreme case of having zero variance, the three models are equivalent. The larger the variance, the larger the difference between the critical coverage distances between two facilities used by the CCFRLP and the DFRLP (i.e., the largest value of the distance between adjacent facilities along a path for which coverage is obtained), and hence the larger the optimality gap of the DFRLP tends to be. Similarly, the coverage probability z_f takes binary values in case of zero variance, and becomes “less binary” when the variance increases. Deterministic coverage and chance constrained coverage, instead, remain binary valued, causing the optimality gaps for the CCFRLP and DFRLP solutions in terms of expected coverage to increase.

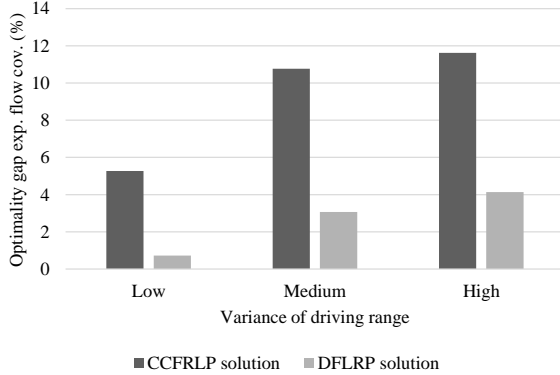


Figure 8: Optimality gap in terms of expected flow coverage for the CCFRLP and DFRLP solutions and low (125), medium (1 250), and high (2 500) driving range variance, using the baseline case $s80w40$ and $p = 15$.

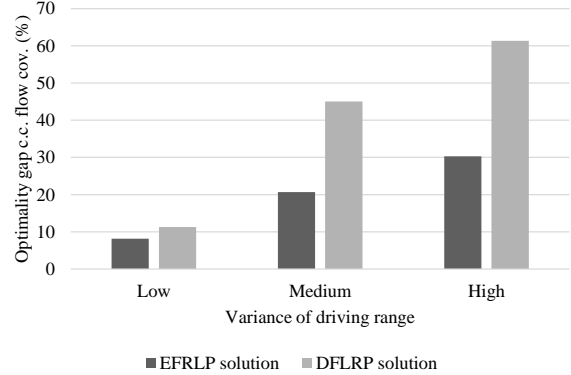


Figure 9: Optimality gap in terms of chance constrained flow coverage for the EFRLP and DFRLP solutions and low (125), medium (1 250), and high (2 500) driving range variance, using the baseline case $s80w40$ and $p = 15$.

4.4. Sensitivity analysis on vehicle volumes

Flow volume data v_f are typically estimated based on traffic counts or mobile phone data. Due to a mix of methodological limitations, data inaccuracies and traffic dynamics, these data tend to be imprecise. To assess the potential impact on optimal solutions, we randomly generate the “true” values of these parameters, v_f^* , and calculate the corresponding optimality gap: the gap between the optimal solution value and the value of the solution based on the “incorrect” parameter values. Specifically, for each f we generate a random draw from the uniform distribution on the interval $[-\delta, \delta]$, represented by $U(-\delta, \delta)$, and generate v_f^* as $v_f^* = v_f \cdot (1 + U(-\delta, \delta))$. Hence, v_f^* deviates at most the factor δ from v_f . Again, we base the analysis on the baseline case $s80w40$ and $p = 15$. Repeating this 50 times for each $\delta \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ yields the average and maximum optimality gaps depicted in Figure 10.

We observe that the EFRLP, CCFRLP, and DFRLP solutions are very robust with respect to data impreciseness. For example, the average optimality gap resulting from 50% impreciseness (i.e. $\delta = 0.5$) is only 0.16%, 1.06%, and 0.33%, respectively. This can be explained by the tendency to focus resources on a selected set of dominant vehicle flows. The 10% largest vehicle flows, which represent over 70% of the total vehicle volume, have relatively high average coverage levels: 0.72 (EFRLP), 0.68 (CCFRLP), and 0.79 (DFRLP) for the top 10% versus 0.43, 0.27, and 0.55 for the bottom 90%. The vehicle volumes are so large that, even in the extreme case that they are all substantially over-estimated, it remains attractive to dedicate resources to these flows. Consequently, the location choices based on the “incorrect” data tend to remain close to optimal in case that the “true” data are used.

4.5. Chance constrained coverage targets

Solving the CCFRLP requires a choice on the value of α , the maximum probability of running out of fuel for which a driver is still regarded as being covered. Decreasing the value

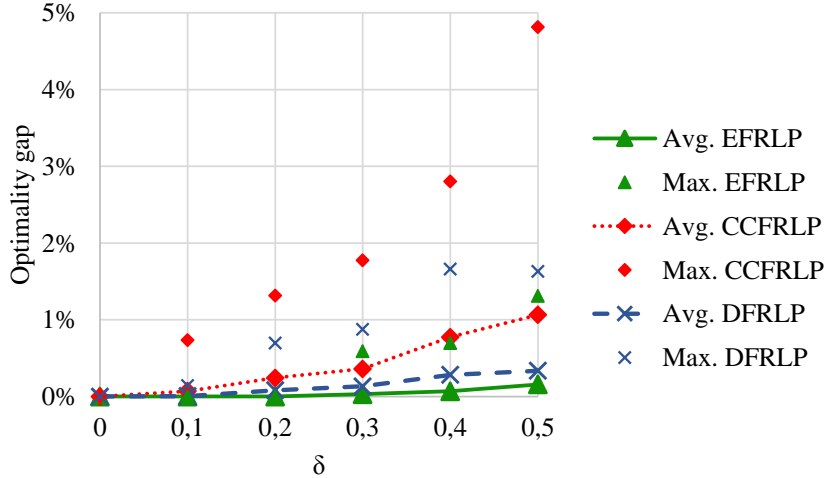


Figure 10: Average and maximum of optimality gaps for the EFRLP, CCFRLP, and DFRLP solutions resulting from 50 random draws of the “true” vehicle volumes for the baseline case $s80w40$ and $p = 15$, using different values of δ .

of α tends to decrease the number of vehicles covered, as resulting networks tend to get more concentrated. In other words, it tends to increase the number of facilities needed to reach the same target in terms of the fraction of vehicles covered, which we denote by β . To provide decision makers with insight into this relationship, this section numerically analyzes the trade-off between α , β , and the number of facilities needed by means of a modified version of the CCFRLP model:

$$\min \sum_{k \in K} x_k \quad (39)$$

$$\text{s.t. } \sum_{f \in F} v_f y_f \geq \beta \sum_{f \in F} v_f \quad (40)$$

(3) – (8), (35) – (37)

The objective function minimizes the number of facilities needed, and constraint (40) ensures that the fraction of vehicles covered is at least β . Figure 11 depicts the results for $\alpha \in \{0.10, 0.09, \dots, 0.02, 0.01, 0.001\}$ and $\beta \in \{0.1, 0.2, \dots, 0.9\}$. The number of facilities needed ranges from 2 ($\alpha = 0.1, \beta = 0.1$) to 50 ($\alpha = 0.001, \beta = 0.9$). As to be expected, it becomes more and more costly to cover additional quantities of flow, so that the marginal costs of increasing β and tend to increase substantially. For example, in case that $\alpha = 0.05$, increasing the coverage target from 0.1 to 0.2 requires 2 additional facilities, whereas increasing the target from 0.8 to 0.9 requires 7 additional facilities.

The critical coverage distance between two facilities decreases steeply with α . Consequently, the number of facilities needed to cover a flow tends to increase quickly, as can be observed from Figure 11. For example, in case that $\beta = 0.9$, decreasing α from 0.1 to 0.01

requires 6 additional facilities, whereas decreasing α from 0.01 to 0.01 requires 11 additional facilities.

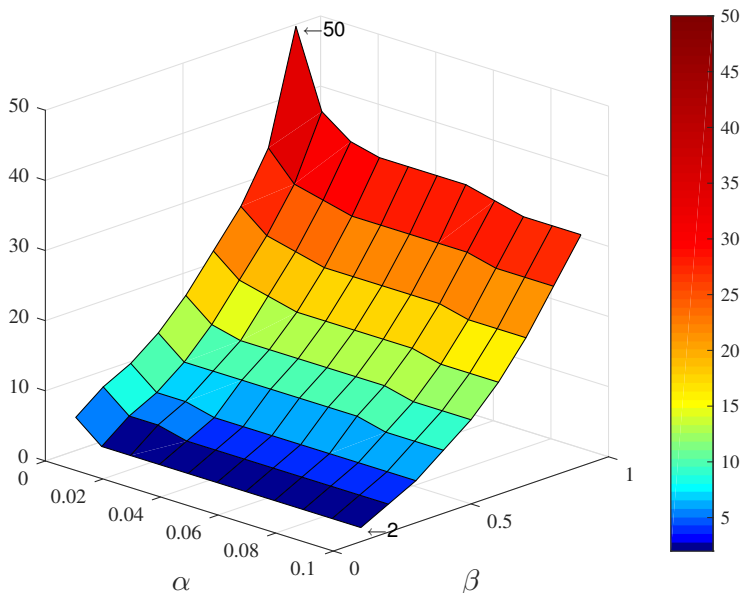


Figure 11: Number of facilities needed to cover a given target fraction β of the total vehicle volume for different values of α .

4.6. Computational results

In Appendix C we prove the following result by reduction from the problem CLIQUE:

Proposition 1. *The EFRLP, CCFRLP, and DFRLP are strongly \mathcal{NP} -hard problems.*

To gain some more insight into the tractability of the model, this section and Appendix D present several computational statistics for the EFRLP and CCFRLP models (the computational performance of models for the DRFLP has been extensively analyzed in Capar and Kuby (2012); Capar et al. (2013); MirHassani and Ebrazi (2012)). Table 4.3 presents computational results for the baseline case *s80w40*. Both models need 25 289 constraints and 14 160 variables to solve this instance. We observe that the solution times for the CCFRLP are substantially smaller than those for the EFRLP. Furthermore, solution times seem to be larger for larger values of p . This can be explained by the fact that the number of solutions equals $\binom{|K|}{p}$, which increases with p for $p \leq 40$ and $|K| = 80$. The results for the other problem instances are given in Appendix D and confirm these observations. They also show that solution times tend to grow exponentially with the problem size, as measured by $|K|$ and $|F|$. For example, the solution times for the EFRLP and $p = 25$ range from 2.4 seconds (*s40w20*) to 34 hours (Florida).

	p									
	1	2	3	4	5	10	15	20	25	
EFRLP	1.9	13.6	61.0	43.7	60.2	204.1	320.9	652.9	3118.8	
CCFRLP	2.0	15.9	23.6	32.9	36.8	67.6	143.8	138.2	159.5	

Table 4.3: Solution times (sec.) for the baseline case *s80w40*.

5. Conclusions

Existing models for locating refueling stations assume a vehicle’s driving range to be fixed. Although this assumption is easy to work with, it is highly unrealistic and potentially yields location decisions that are far from optimal. This paper sets out to develop and analyze models that relax this assumption. We first propose a novel mixed-integer programming formulation for the (deterministic) Flow Refueling Location Problem, which includes the driving range as an explicit model parameter. Next, we propose two models, which capture the stochastic nature of the driving range. The EFRLP defines the probability that a vehicle can traverse its path without running out of fuel (i.e., the probability that it is covered), and maximizes the expected number of vehicles covered. The CCFRLP defines a flow to be covered if the probability of running out of fuel during a trip along the corresponding path is at most α .

Numerical analyses reveal several key implications for decision makers. First, our analysis of the Value of the Stochastic Solution (VSS) suggests that using the EFRLP instead of its deterministic counterpart (DFRLP) can considerably improve the impact of new facilities in terms of expected coverage. Similarly, in case that coverage is defined by means of chance constraints, using the CCFRLP model instead of the DFRLP or the EFRLP model potentially leads to significant yields in terms of coverage. We show that these potential gains are largely determined by the driving range distribution. Larger values for the expected driving range and smaller values for the driving range variance tend to decrease the gaps between the different models. Third, synergy effects can be obtained by locating *multiple* facilities. Decision makers can exploit these effects in two ways. First, instead of fixing the value of p , they can investigate whether large synergy effects can be gained by (slightly) modifying its value. Second, by making long-term investment plans for the network of facilities rather than taking investment decisions sequentially, they can make optimal use of the synergy effects between all future facilities. These plans thereby avoid the sub-optimality resulting from sequential decision making. Fourth, our sensitivity analysis shows that impreciseness in the flow volume data tends to have little impact on the quality of the location decisions. Finally, there is a substantial trade-off between service level, as measured by α and β , and efficiency, as measured by the number of facilities per vehicle covered. This highlights the importance of choosing service level targets deliberately.

It is up to decision makers to discuss an appropriate definition of coverage, and thereby to choose for one of the models. Chance constrained coverage may be most intuitively appealing in the context of pure electric vehicles, since drivers will only make a trip if the chance of running out of fuel is sufficiently small. On top of that, CCFRLP solutions perform relatively

well in terms of expected coverage, whereas EFRLP solutions perform weakly in terms of chance constrained coverage. On the other hand, expected coverage might be more suitable in the context of hybrid (electric) vehicles, which can switch to a different power source when the primary source is exhausted.

One assumption of our stochastic models is that the driving range distribution is known. Though car manufacturers do analyze this distribution (see e.g. Knipe et al., 2003), technological developments on the one hand or battery deterioration on the other hand may cause the distribution to change over time. It is therefore important to forecast the evolution of this distribution, so as to ensure future effectiveness of current location decisions. This is particularly important in case that a chance constrained coverage criterion is used, since vehicle flows may lose coverage in case of a downward shift in the driving range. In case of updated information on the driving range, our methods could be used to analyze whether additional investments in the network are necessary. Another approach is to make use of “portable” facilities that can easily be relocated in response to demand or driving range fluctuations (cf. Hosseini and MirHassani, 2015).

Although the location problems considered in this paper were solved to optimality for each of the problem instances considered, our numerical experiments show that computation time increases vastly with the problem size. Developing alternative models and solution methods, like a column generation solution approach, seems to be promising (see e.g., Núñez Ares et al. (2015); Yıldız et al. (2016) for column generation methods for similar location problems). Other relevant directions for future research result from relaxing our model assumptions. One such assumption is that vehicles do not deviate from their preferred paths to a nearby refueling station. Kim and Kuby (2012) and Yıldız et al. (2016) propose models that do include routing choices into the location problem. Also the assumption that recharging facilities cannot be moved could be relaxed, as Hosseini and MirHassani (2015) do for a variant of the DFRLP. Relaxing the assumption of uncapacitated refueling stations renders queue lengths, waiting times and capacity constraints to become important location criteria. This substantially increases the complexity of the location problem, as shown in Jung et al. (2014) and Hosseini and MirHassani (2015). New types of models and solution methods are needed to solve stochastic flow refueling location problems that relax these assumptions.

References

- Arabani, A. B., Farahani, R. Z., 2012. Facility location dynamics: An overview of classifications and applications. *Computers & Industrial Engineering* 62 (1), 408–420.
- Averbakh, I., Berman, O., 1996. Locating flow-capturing units on a network with multi-counting and diminishing returns to scale. *European Journal of Operational Research* 91 (3), 495–506.
- Berman, O., Krass, D., 1998. Flow intercepting spatial interaction model: a new approach to optimal location of competitive facilities. *Location Science* 6 (1), 41–65.

- Berman, O., Larson, R. C., Fouska, N., 1992. Optimal Location of Discretionary Service Facilities. *Transportation Science* 26 (3), 201–211.
- Birge, J. R., 1982. The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming* 24 (1), 314–325.
- Bonferroni, C. E., 1936. *Teoria statistica delle classi e calcolo delle probabilita*. Libreria internazionale Seeber.
- Capar, I., Kuby, M., 2012. An efficient formulation of the flow refueling location model for alternative-fuel stations. *IIE Transactions* 44 (8), 622–636.
- Capar, I., Kuby, M., Leon, V. J., Tsai, Y.-J., 2013. An arc cover–path-cover formulation and strategic analysis of alternative-fuel station locations. *European Journal of Operational Research* 227 (1), 142–151.
- De Vries, H., Van de Klundert, J., Wagelmans, A., 2014. The Roadside Healthcare Facility Location Problem. Tech. rep., Econometric Institute, Erasmus School of Economics, report number: EI 2014-09.
URL <http://repub.eur.nl/pub/51315>
- Dong, J., Lin, Z., 2014. Stochastic modeling of battery electric vehicle driver behavior: The impact of charging infrastructure deployment on BEV feasibility. In: *Transportation Research Board 93rd Annual Meeting*. No. 14-5107.
- Ehsani, M., Gao, Y., Emadi, A., 2009. *Modern Electric, Hybrid Electric, and Fuel Cell Vehicles: Fundamentals, Theory, and Design*. CRC press.
- Hodgson, M. J., 1990. A Flow-Capturing Location-Allocation Model. *Geographical Analysis* 22 (3), 270–279.
- Hosseini, M., MirHassani, S., 2015. Refueling-station location problem under uncertainty. *Transportation Research Part E: Logistics and Transportation Review* 84, 101–116.
- Jung, J., Chow, J. Y., Jayakrishnan, R., Park, J. Y., 2014. Stochastic dynamic itinerary interception refueling location problem with queue delay for electric taxi charging stations. *Transportation Research Part C: Emerging Technologies* 40, 123–142.
- Kang, J. E., Recker, W., 2014. Strategic Hydrogen Refueling Station Locations with Scheduling and Routing Considerations of Individual Vehicles. *Transportation Science* 49 (4), 767–783.
- Kim, J.-G., Kuby, M., 2012. The deviation-flow refueling location model for optimizing a network of refueling stations. *International Journal of Hydrogen Energy* 37 (6), 5406–5420.
- Knipe, T. J., Gaillac, L., Argueta, J., 2003. 100,000-mile evaluation of the toyota rav4 ev. Southern California Edison, Electric Vehicle Technical Center.

- Kuby, M., Lim, S., 2005. The flow-refueling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences* 39 (2), 125–145.
- Kuby, M., Lines, L., Schultz, R., Xie, Z., Kim, J.-G., Lim, S., 2009. Optimization of hydrogen stations in Florida using the Flow-Refueling Location Model. *International Journal of Hydrogen Energy* 34 (15), 6045–6064.
- Lee, Y.-G., Kim, H.-S., Kho, S.-Y., Lee, C., 2014. User Equilibrium-Based Location Model of Rapid Charging Station for Electric Vehicles with Batteries That Have Different States of Charge. *Transportation Research Record: Journal of the Transportation Research Board* 2454, 97–106.
- MirHassani, S., Ebrazi, R., 2012. A Flexible Reformulation of the Refueling Station Location Problem. *Transportation Science* 47 (4), 617–628.
- Musk, E., Straubel, J., May 2012. Model S Efficiency and Range.
URL <http://www.teslamotors.com/blog/model-s-efficiency-and-range>
- Núñez Ares, J., De Vries, H., Huisman, D., 2015. A Column Generation Approach for Locating Roadside Clinics in Africa based on Effectiveness and Equity. Tech. rep., Econometric Institute, Erasmus School of Economics, report number: EI 2015-19.
- Papadimitriou, C., Steiglitz, K., 1998. *Combinatorial optimization: algorithms and complexity*. Courier Dover Publications.
- Romm, J., 2006. The car and fuel of the future. *Energy Policy* 34 (17), 2609–2614.
- TeslaMotors, 2015. Supercharger.
URL <http://www.teslamotors.com/supercharger>
- Yıldız, B., Arslan, O., Karaşan, O. E., 2016. A branch and price approach for routing and refueling station location model. *European Journal of Operational Research* 248 (3), 815–826.

Appendix A. Random network generation

The random networks are generated as follows. First, we choose the location of s nodes, by generating their coordinates uniformly in $[0, 1000]^2$. The distance between two nodes is assumed to be Euclidian. Afterwards, we generate edges between the nodes by creating the minimum spanning tree for the complete graph, and add s additional edges by connecting the s closest node pairs that were not yet connected. We end up with a network $N = G(K, E)$ with $|K| = s$ locations and $|E| = 2s - 1$ edges. Next, we duplicate w of the s nodes. These duplicates are Origin - Destination (O-D) nodes, which define $w(w - 1)/2$ O-D pairs. For a given path f , the two O-D nodes are denoted by O_f and D_f . Dijkstra's algorithm is used to determine the shortest path from O_f to D_f , the potential facility locations passed along this path (K_f), and the distance between each pair of nodes in this path (t_{kl}).

To generate the vehicle volumes v_f , we perform the following steps. First, we assign to each of the w O-D nodes a number drawn from a standard uniform distribution $\mathcal{U}(0, 1)$. Let e_f^O and e_f^D denote these numbers for the origin and destination respectively, and let T_f be the travel distance between O_f and D_f . The non-normalized vehicle volume v_f^* is calculated with equation (A.1). Compared to the network generation method proposed by Capar and Kuby (2012), this differs deviated in two ways: we do not square the denominator and we add the indicator function. These changes were made so as to reduce the impact of short distance routes on location decisions, since long distance routes constitute the main focus of this research. Next, we obtain the the vehicle volumes v_f by normalizing v_f^* in such a way that the total volume adds up to 10^6 .

$$v_f^* = \frac{e_f^O e_f^D}{T_f} \cdot 1_{T_f \geq 100} \quad (\text{A.1})$$

Appendix B. Overall results

p	Value in Solution of:	EFRLP			CCFRLP			VSS
		EFRLP	CCFRLP	DFRLP	EFRLP	CCFRLP	DFRLP	
1	2.63E+11	2.63E+11 (0.00%)	2.63E+11 (0.00%)	2.62E+11 (0.00%)	2.62E+11	2.62E+11 (0.00%)	0.00E+00	
2	4.11E+11	4.11E+11 (0.00%)	4.11E+11 (0.00%)	4.11E+11 (0.00%)	4.11E+11	4.11E+11 (0.00%)	0.00E+00	
3	5.41E+11	5.41E+11 (0.00%)	5.41E+11 (0.00%)	5.33E+11 (0.00%)	5.33E+11	5.33E+11 (0.00%)	0.00E+00	
4	6.12E+11	6.11E+11 (-0.23%)	6.12E+11 (0.00%)	6.01E+11 (0.00%)	6.01E+11	6.01E+11 (0.00%)	0.00E+00	
5	6.82E+11	6.82E+11 (0.00%)	6.82E+11 (0.00%)	6.68E+11 (0.00%)	6.68E+11	6.68E+11 (0.00%)	0.00E+00	
10	8.44E+11	8.44E+11 (-0.04%)	8.44E+11 (0.00%)	8.27E+11 (-0.03%)	8.27E+11	8.27E+11 (-0.03%)	0.00E+00	
15	9.13E+11	9.12E+11 (-0.17%)	9.12E+11 (-0.14%)	8.94E+11 (-0.25%)	8.97E+11	8.92E+11 (-0.56%)	1.24E+09	
20	9.54E+11	9.52E+11 (-0.17%)	9.53E+11 (-0.07%)	9.40E+11 (-0.17%)	9.41E+11	9.35E+11 (-0.71%)	7.07E+08	
25	9.73E+11	9.69E+11 (-0.41%)	9.72E+11 (-0.07%)	9.62E+11 (-0.22%)	9.64E+11	9.59E+11 (-0.58%)	6.69E+08	

Table B.1: Expected volume covered and chance constrained flow volume covered in the EFRLP, CCFRLP, and DFRLP solutions to the Florida case.

p	Value in Solution of:	EFRLP			CCFRLP			VSS
		EFRLP	CCFRLP	DFRLP	EFRLP	CCFRLP	DFRLP	
1		7.10E+04	2.39E+04 (-66.36%)	7.10E+04 (0.00%)	0.00E+00 (-100.00%)	2.25E+04	0.00E+00 (-100.00%)	8.30E-04
2		1.16E+05	9.17E+04 (-20.64%)	1.16E+05 (0.00%)	2.85E+04 (-57.04%)	6.62E+04	2.85E+04 (-57.04%)	0.00E+00
3		1.56E+05	1.31E+05 (-16.23%)	1.55E+05 (-1.14%)	2.85E+04 (-71.41%)	9.95E+04	5.90E+04 (-40.77%)	1.79E+03
4		1.95E+05	1.59E+05 (-18.58%)	1.95E+05 (0.00%)	5.90E+04 (-55.89%)	1.34E+05	5.90E+04 (-55.89%)	0.00E+00
5		2.25E+05	2.07E+05 (-8.15%)	2.25E+05 (0.00%)	5.90E+04 (-63.45%)	1.61E+05	5.90E+04 (-63.45%)	0.00E+00
10		3.95E+05	3.36E+05 (-15.00%)	3.95E+05 (0.00%)	1.67E+05 (-44.58%)	3.01E+05	1.67E+05 (-44.58%)	0.00E+00
15		5.41E+05	4.61E+05 (-14.86%)	5.17E+05 (-4.43%)	3.40E+05 (-21.91%)	4.36E+05	2.36E+05 (-45.83%)	2.40E+04
20		6.51E+05	5.88E+05 (-9.61%)	6.47E+05 (-0.52%)	3.87E+05 (-30.24%)	5.55E+05	3.88E+05 (-30.09%)	3.36E+03
25		7.44E+05	6.69E+05 (-10.03%)	7.38E+05 (-0.88%)	4.50E+05 (-31.12%)	6.53E+05	4.17E+05 (-36.10%)	6.52E+03

Table B.2: Expected volume covered and chance constrained flow volume covered in the EFRLP, CCFRLP, and DFRLP solutions to instance $s100w50$.

p	Value in Solution of:	EFRLP			CCFRLP			VSS
		EFRLP	CCFRLP	DFRLP	EFRLP	CCFRLP	DFRLP	
1		7.64E+04	7.64E+04 (0.00%)	7.64E+04 (0.00%)	7.48E+04 (0.00%)	7.48E+04	7.48E+04 (0.00%)	6.71E-03
2		1.57E+05	1.57E+05 (0.00%)	1.57E+05 (0.00%)	1.21E+05 (0.00%)	1.21E+05	1.21E+05 (0.00%)	0.00E+00
3		2.02E+05	1.82E+05 (-9.75%)	2.02E+05 (0.00%)	1.21E+05 (-27.96%)	1.68E+05	1.21E+05 (-27.96%)	0.00E+00
4		2.68E+05	2.65E+05 (-1.16%)	2.68E+05 (0.00%)	1.65E+05 (-25.15%)	2.21E+05	1.65E+05 (-25.15%)	0.00E+00
5		3.21E+05	2.91E+05 (-9.32%)	3.18E+05 (-0.93%)	2.32E+05 (-14.75%)	2.72E+05	1.65E+05 (-39.35%)	3.00E+03
10		5.01E+05	4.62E+05 (-7.79%)	5.01E+05 (0.00%)	2.58E+05 (-38.70%)	4.21E+05	2.58E+05 (-38.70%)	0.00E+00
15		6.37E+05	5.68E+05 (-10.76%)	6.18E+05 (-3.06%)	4.42E+05 (-20.68%)	5.58E+05	3.07E+05 (-45.01%)	1.95E+04
20		7.29E+05	6.82E+05 (-6.47%)	6.92E+05 (-5.07%)	5.66E+05 (-16.26%)	6.76E+05	3.68E+05 (-45.54%)	3.70E+04
25		8.17E+05	7.68E+05 (-6.08%)	7.85E+05 (-3.94%)	6.33E+05 (-16.94%)	7.63E+05	4.45E+05 (-41.66%)	3.22E+04

Table B.3: Expected volume covered and chance constrained flow volume covered in the EFRLP, CCFRLP, and DFRLP solutions to instance $s80w40$.

p	Value in Solution of:	EFRLP			CCFRLP			VSS
		EFRLP	CCFRLP	DFRLP	EFRLP	CCFRLP	DFRLP	
1		4.34E+04	(-100.00%)	4.34E+04 (0.00%)	0.00E+00 (0.00%)	0.00E+00	0.00E+00 (0.00%)	6.58E-03
2		8.75E+04	8.75E+04 (0.00%)	7.60E+04 (-13.13%)	7.13E+04 (0.00%)	7.13E+04	0.00E+00 (-100.00%)	1.15E+04
3		1.37E+05	1.37E+05 (0.00%)	1.31E+05 (-4.74%)	1.05E+05 (0.00%)	1.05E+05	7.13E+04 (-32.09%)	6.51E+03
4		1.81E+05	1.50E+05 (-17.06%)	1.81E+05 (0.00%)	1.05E+05 (-30.07%)	1.50E+05	1.05E+05 (-30.07%)	0.00E+00
5		2.10E+05	1.99E+05 (-4.85%)	1.93E+05 (-7.82%)	1.68E+05 (-1.67%)	1.71E+05	1.05E+05 (-38.66%)	1.64E+04
10		3.90E+05	3.64E+05 (-6.68%)	3.90E+05 (0.00%)	3.19E+05 (-9.84%)	3.54E+05	3.19E+05 (-9.84%)	0.00E+00
15		5.25E+05	5.09E+05 (-3.05%)	5.24E+05 (-0.27%)	4.51E+05 (-9.10%)	4.96E+05	3.96E+05 (-20.07%)	1.41E+03
20		6.44E+05	6.26E+05 (-2.73%)	6.42E+05 (-0.35%)	5.74E+05 (-8.30%)	6.26E+05	5.15E+05 (-17.68%)	2.27E+03
25		7.62E+05	7.46E+05 (-2.10%)	7.53E+05 (-1.19%)	6.92E+05 (-6.12%)	7.38E+05	6.58E+05 (-10.77%)	9.09E+03

Table B.4: Expected volume covered and chance constrained flow volume covered in the EFRLP, CCFRLP, and DFRLP solutions to instance $s60w30$.

p	Value in Solution of:	EFRLP			CCFRLP			VSS
		EFRLP	CCFRLP	DFRLP	EFRLP	CCFRLP	DFRLP	
1		1.53E+05	(-100.00%)	1.44E+05 (-6.03%)	0.00E+00 (0.00%)	0.00E+00	0.00E+00 (0.00%)	9.23E+03
2		2.98E+05	1.93E+05 (-35.12%)	2.98E+05 (0.00%)	7.59E+04 (-34.86%)	1.17E+05	7.59E+04 (-34.86%)	2.00E-04
3		3.82E+05	3.51E+05 (-8.13%)	3.79E+05 (-0.72%)	8.59E+04 (-69.74%)	2.84E+05	7.59E+04 (-73.24%)	2.76E+03
4		4.63E+05	4.31E+05 (-6.96%)	4.63E+05 (0.00%)	8.59E+04 (-73.59%)	3.25E+05	8.59E+04 (-73.59%)	0.00E+00
5		5.30E+05	4.80E+05 (-9.42%)	5.18E+05 (-2.19%)	1.77E+05 (-55.29%)	3.97E+05	1.16E+05 (-70.81%)	1.16E+04
10		7.86E+05	7.86E+05 (0.00%)	6.74E+05 (-14.25%)	7.62E+05 (0.00%)	7.62E+05	2.87E+05 (-62.30%)	1.12E+05
15		8.87E+05	8.45E+05 (-4.81%)	7.72E+05 (-13.01%)	8.43E+05 (0.00%)	8.43E+05	4.10E+05 (-51.36%)	1.15E+05
20		9.20E+05	8.67E+05 (-5.81%)	8.05E+05 (-12.50%)	8.47E+05 (-1.98%)	8.64E+05	4.09E+05 (-52.66%)	1.15E+05
25		9.63E+05	9.35E+05 (-2.92%)	9.51E+05 (-1.24%)	8.57E+05 (-2.36%)	8.77E+05	8.12E+05 (-7.39%)	1.19E+04

Table B.5: Expected volume covered and chance constrained flow volume covered in the EFRLP, CCFRLP, and DFRLP solutions to instance $s40w20$.

Appendix C. Strongly NP-hardness proof for the DFRLP, EFRLP and CCFRLP

Proposition 1. *The EFRLP, CCFRLP, and DFRLP are strongly NP-hard problems.*

Proof. First, note that the decision versions of the DFRLP, EFRLP and CCFRLP are in \mathcal{NP} : our MIP formulations allows us to verify a YES answer to these problems in polynomial time. Next, we show how to construct a polynomial transformation of the CLIQUE problem to the decision-version of each of the three problems. Since the CLIQUE problem is strongly \mathcal{NP} -Complete (Papadimitriou and Steiglitz, 1998), this shows that the decision-versions of the DFRLP, EFRLP and CCFRLP are strongly \mathcal{NP} -Complete too.

We consider a graph $G(V, E)$. CLIQUE corresponds to the decision problem “does G contain a complete sub-graph consisting of n nodes?”. To perform the transformation into an instance of the DFRLP, the EFRLP or the CCFRLP, we introduce for every node $v \in V$ a potential facility location $k \in K$. Furthermore, for every edge in E , we introduce a flow f and the corresponding path. This is a direct path from a potential facility location to another potential facility location. These two locations correspond to the two nodes connected by the corresponding edge. Note that this transformation can be performed in polynomial time since $|KP| = |V|$ and $|F| = |E|$. Finally, we set $p = n$ and choose the values of the remaining parameters corresponding to the three problems such that flow f is covered (DFRLP and CCFRLP) or the coverage probability equals 1 (EFRLP) if and only if facilities are placed at both potential facility locations passed by this flow, and that the coverage level or coverage probability equals 0 in all other cases. Let the two potential facility locations along the path of flow f be represented by k_{1f} and k_{2f} . Then for the DFRLP this situation can be attained by setting $t_{O_f, k_{1f}} = 0$, $t_{k_{1f}, k_{2f}} = 0$, $t_{k_{2f}, D_f} = 0$ and by setting the travel time between any other pair of locations along this path equal to $R + 1$. Similarly, for the EFRLP it suffices to set $g_{O_f, k_{1f}} = 0$, $g_{k_{1f}, k_{2f}} = 0$, $g_{k_{2f}, D_f} = 0$ and to set this probability equal to 1 for any other pair of locations along this path. Finally, for the CCFRLP the situation can be attained by using the same definition for g_{kl} and by setting $\alpha = 0$.

To complete the proof, we will show that CLIQUE has a YES-answer if and only if the corresponding instance of the decision version of the DFRLP, the EFRLP or the CCFRLP has a solution of value $n \cdot (n - 1)/2$. First, suppose that CLIQUE has a YES-answer (i.e., the graph contains a clique of size n). Note that there are at most $n \cdot (n - 1)/2$ flow paths connecting each pair of the n potential facility locations we allocate the new facilities to. Hence, at most $n \cdot (n - 1)/2$ flow paths have a coverage level or coverage probability that is larger than 0. By the construction of the instances we know that the solution value of the three problems equals the number of flow paths for which the coverage level or coverage probability equals 1, this implies that the instance has a solution value that is *at most* $n \cdot (n - 1)/2$. Next, observe that, in case that CLIQUE has a YES-answer, there are exactly $n \cdot (n - 1)/2$ flow paths connecting each pair of the n potential facility locations corresponding to the n -clique. Then allocating the new facilities to these n potential facility locations makes sure that the coverage level or coverage probability for each of the corresponding $n \cdot (n - 1)/2$ flows has the value 1. This implies that the corresponding solution value is at least $n \cdot (n - 1)/2$ and hence equals *exactly* this number.

Second, suppose that the instance corresponding to the decision version of the DFRLP, the EFRLP or the CCFRLP has a value of $n \cdot (n - 1)/2$. Then this means that *at least* $n \cdot (n - 1)/2$ flows have a strictly positive coverage level or coverage probability. Next,

observe that the fact that we are only allowed to allocate n facilities implies that *at most* $n \cdot (n - 1)/2$ flows have a strictly positive coverage level or coverage probability. Hence, *exactly* $n \cdot (n - 1)/2$ flows have a strictly positive coverage level or coverage probability. This number is only met if there exist exactly $n \cdot (n - 1)/2$ unique flows that visit two of the n allocated facilities. This implies that the n vertices in the corresponding CLIQUE instance form a clique of size n . \square

Appendix D. Computational results all instances

	Florida		s100w50		s80w40		s60w30		s40w20	
	EFRLP	CCFRLP	EFRLP	CCFRLP	EFRLP	CCFRLP	EFRLP	CCFRLP	EFRLP	CCFRLP
$ K $	302	302	100	100	80	80	60	60	40	40
$ F $	2 701	2 701	1 225	1 225	780	780	435	435	190	190
# variables	846 809	846 809	101 822	101 822	62 571	62 571	35 066	35 066	9 125	9 125
# integer variables	302	3 003	100	1 325	80	860	60	495	40	230
# constraints	172 916	172 916	40 643	40 643	25 289	25 289	14 465	14 465	4 772	4 772
	Solution time (sec.)		Solution time (sec.)		Solution time (sec.)		Solution time (sec.)		Solution time (sec.)	
1	91.7	111.6	4.9	5.0	1.9	2.0	0.8	0.9	0.2	0.1
2	878.0	554.9	145.3	76.0	13.6	15.9	6.3	3.7	0.3	0.5
3	15 169.6	580.2	152.2	87.4	61.0	23.6	17.8	6.2	0.5	0.3
4	718.8	737.7	231.3	126.1	43.7	32.9	15.9	5.8	1.6	0.5
p	30 470.4	752.4	241.4	188.6	60.2	36.8	20.9	13.3	0.8	0.5
10	41 792.3	1 890.3	814.3	503.4	204.1	67.6	34.1	8.3	1.1	0.5
15	36 249.1	10 981.3	2 217.8	569.8	320.9	143.8	61.8	20.6	3.0	0.7
20	67 820.1	30 724.4	12 872.3	486.4	652.9	138.2	113.5	38.3	9.4	0.9
25	122 198.2	79 452.4	17 341.8	818.8	3 118.8	159.5	112.4	36.1	2.4	0.5

Table D.1: Solution times (sec.) and model characteristics.