Erasmus University Rotterdam - Erasmus School of Economics

Bachelor Thesis BSc^2 Econometrics/Economics

An Analysis of Synergy Effects in a Flow Refueling Location Problem with Capacity Constraints

Supervisor: T. Breugem

Second assessor: L. E. Westerink-Duijzer

D.T.J. van der Schaft

431738

July 8, 2019

Abstract

In this research, we investigate several instances of the Flow Refueling Location Problem (FRLP). First, we examine the formulation of De Vries & Duijzer (2017) and their extensions of it (the Expected FRLP and Chance Constrained FRLP) that model the driving range as a stochastic parameter. Next, we treat a new FRLP formulation of Boujelben & Gicquel (2019) and use it as a basis for our newly introduced Unit-Capacitated FRLP, in which capacity is represented as interchangeable modular units. We show that the models of De Vries and Duijzer outperform the deterministic FRLP in terms of expected coverage and chance constrained coverage, and that they inhibit synergy effects. Moreover, it is shown that the flexibility of a modular unit structure for capacity yields similar efficiency gains, especially when the driving range is relatively low compared to the size of the network and units are split up into smaller pieces multiple times. Finally, using the capacitated model as opposed to the uncapacitated deterministic model results in substantial coverage gains.

Contents

1 Introduction

Technological innovations and the urge for environmentally friendly modes of transportation have lead to the rise of alternative-fuel vehicles (AFVs). Electricity, hydrogen and bio-ethanol have the potential to replace petrol, diesel and gas as energy sources in the near future. However, their soar is inhibited by the absence of a sufficient network that facilitates large-scale use.

There are many obstacles involved in creating an entirely new infrastructure for AFVs. Melendez (2006) identifies four major barriers: the limited availability of refueling infrastructure, the high costs of constructing refueling facilities, the high costs of AFVs and their limited driving range. These factors altogether underline the importance for a sophisticated method that determines where recharging stations can be located most efficiently.

As vehicles can only drive a limited time on a full tank or battery, refueling stations must be located in such a way that the distance between two consecutive facilities is not too large. The research of Duijzer and De Vries (2017) responds to this need by investigating several instances of the Flow Refueling Location Problem (hereafter referred to as the FRLP), which aims to maximise the number of vehicles that is covered by a refueling facility. Only when a vehicle encounters enough facilities on its route such that it will never run out of fuel, it is considered covered.

The main disadvantage of the standard FRLP is that it assumes the driving range of vehicles to be exactly the same on every segment of the trip. In reality, the range is highly dependent on the speed at which the driver travels (either due to congestion or legally enforced speed limits), weather conditions and driving behaviour (McIlroy & Stanton, 2017; X. Wu et al., 2015). Duijzer and De Vries acknowledge this and come up with two models that incorporate the stochasticity of the driving range into the FRLP.

Firstly, they introduce the Expected Flow Refueling Location Problem (EFRLP), which aims to maximise the expected number of drivers that can finish their trip without running out of fuel. Instead of setting the driving range to a fixed number, it is modelled as a random variable with the same expectation. Secondly, the Chance Constrained Flow Refueling Location Problem (CCFRLP) uses this random element in a similar manner, but takes a different approach when it comes to maximising coverage. In this model, a flow is considered covered if the probability of running out of fuel during the entire trip is less than α.

One of the main assumptions that they make throughout their research is the unlimited capacity of the recharging stations. Although this assumption reduces the theoretical and computational complexity of the models, it is unlikely to hold when the adoption of AFVs becomes more widespread in the future. It is reasonable to assume unlimited capacity if the number of vehicles that uses the network is small and maximum capacity will barely ever be reached. However, future growth of the AFV market will lead to higher vehicles volumes and hence the risk of congestion at charging stations.

Our research extends the existing literature in this field by applying the notion of capacity to the brand new formulation of (Boujelben & Gicquel, 2019). This formulation allows us to include an element of capacity into the framework, based on the idea of Upchurch et al. (2009). They model capacity as interchangeable modular units that can be distributed in any way over the potential facility locations. Considering capacity of recharging stations has not been extensively investigated in the existing literature and adds an unprecedented degree of realism to the FRLP.

Furthermore, we look at the extent to which synergy effects can be utilised by picking smaller units of capacity, without increasing the total capacity. De Vries and Duijzer discovered synergy effects when incrementing the number of uncapacitated facilities. They argue that these effects arise because multiple facilities are usually necessary to cover a flow. This line of reasoning is another strong advocate in favor of our approach of capacity. It will not only give insight into the suboptimality of the solutions to the uncapacitated model, but also into the real-world benefit of flexibly adjusting the size of the facilities to the structure of the network.

In this paper, the deterministic FRLP as well as the two stochastic models by De Vries and Duijzer are investigated and replicated. In a similar fashion as in their research, this is followed by a numerical analysis of the results that are found using randomly generated networks. Furthermore, the formulation of Boujelben & Gicquel (2019) is introduced and used as the basis for our capacitated model. This model is finally adopted to test whether synergy effects can be established.

Section 2 dives into the existing literature regarding refueling location problems and its relevance. Moreover, Section 3 elaborates upon the technical aspects of the models of De Vries & Duijzer (2017) and Boujelben & Gicquel (2019). Our new capacitated model is introduced in Section 4. Section 5 presents the results of our analysis and lastly, Section 6 draws the conclusions of our research and discusses limitations and recommendations for future research.

2 Literature Review

Around a quarter of worldwide CO_2 emissions are directly caused by the transportation sector, of which 75% is contributed by road transport (Chapman, 2007). A reduction in this number requires both changes in travel behaviour and technologically advanced alternative transportation methods (Choi et al., 2013). In terms of technology, electricity, hydrogen and bio-ethanol powered vehicles have presented themselves as potential eco-friendly alternatives for petrol, diesel and gas. Nonetheless, a so-called 'chicken and egg' problem avoids widespread implementation of these AFVs in the short run (Lim & Kuby, 2010). On the one hand, a sufficient number of AFVs must be in use to make the construction of recharging station worthwile. On the other hand, the existence of an extensive network of recharging stations is often a prerequisite for the demand for AFVs. It is therefore highly relevant for decision-makers to possess methods that allow them to construct a recharging network with maximal coverage at minimal cost.

Models that optimally allocate facilities to locations in order to serve a certain type of demand have existed for many years. In the early stages of research in this field, demand was usually modelled as being an asset of a location, such as in the maximal covering location problem (MCLM). In this model, a number of facilities is installed with the objective to maximise coverage, with coverage being defined as the number of people that can reach a facility within a certain distance or time span (Church & Velle, 1974). In other words, demand is regarded as a static object.

However, there are many instances where demand is dynamic and it makes more sense to represent is as a flow. A good example of such an instance is our case of refueling stations. People are commonly not interested in a close proximity of a gasoline station from their house or work, but more in the possibility to refuel on the most common trips that they take. Demand is hence more realistically represented as a traffic flow.

The main problem of a static demand model arises in the form of cannibalisation. It may occur that several facilities are located close together in a network due to high demand in that area and therefore cover busy paths multiple times, whereas other traffic remains uncovered (Hodgson, 1990). Additional facilities therefore take away the demand of existing recharging stations, instead of increasing overall coverage. Hodgson (1990) acknowledged the cannibalisation problem and attempts to solve it by alternatively considering network flows as demand objects in his the Flow-Capturing Location Problem (FCLP).

Several interesting extensions of the FCLP have been developed throughout the years. Hodgson & Rosing (1992) express demand in a hybrid model as a combination of passing flows and aggregated consumer nodes and examine the trade-off between fulfilling both demand types. For example, convenience stores will be visited by people who are on their way from one location to another (represented as flows) as well as those who will make a trip from home for the sole purpose of visiting the store (represented as nodes). A more practical application to vehicle inspection stations (Hodgson et al., 1996) attaches higher value to facilities that are located close to the origin of flows, such that maximum protection against hazardous drivers is achieved.

Although the FCLP solves the cannibalisation problem, it still lacks a few elements that would make it applicable to the case of refueling stations. Primarily, it does not yet take restrictions of the driving range into account. Instead, it regards a flow as covered if there is at least one facility on its path from origin to destination, which is more interesting for facilities such as ATMs or convenience stores. Averbakh & Berman (1996) do incorporate a benefit of encountering multiple facilities on a customers pre-planned tour, but do not yet impose it as a prerequisite for coverage or service.

Kuby and Lim (2005) constructed a model that applies the concept of flow capturing to refueling stations. Their Flow Refueling Location Problem (FRLP) makes two major adjustments to the Hodgson's approach. Firstly, a flow is not covered if the next refueling station cannot be reached within the vehicle's driving range. Multiple facilities may thus be required to cover one flow. Secondly, vehicles along a certain flow perform a round-trip between their origin and destination. It is therefore important that the driving range is sufficient to survive the round-trip from the first facility on the path to the origin and from the last facility on the path to the destination. Altogether, Kuby and Lim's version of the FRLP boils down to a set covering approach that checks all potential facility combinations and finds the one that maximises flow coverage.

Since then, numerous additions and extensions of the FRLP have been explored, all of which aim to formulate the problem as realistically as possible. This includes the incorporation of stochasticity in the driving range by De Vries & Duijzer (2017). As opposed to previous work, De Vries and Duijzer explicitely model the driving range as a parameter in the deterministic framework. By adopting a new formulation that does so, they can easily make adjustments that make the driving range stochastic. A more elaborate explanation of their approach can be found in Section 3. F. Wu & Sioshansi (2017) dive into the notion of uncertainty in the demand side of the problem and incorporate stochasticity in the vehicle flow volumes.

An interesting extension has been the Deviation Flow Refueling Location Problem by Kim & Kuby (2012). They relaxed the assumption that drivers will always stay on their shortest path from origin to destination. Instead, drivers may deviate from their pre-determined path to get refueling services up to a maximum tolerated distance. Especially in the early stages of rolling out a recharging network for AFVs, they may not be a facility on every route yet. Allowing for deviation takes away the problem that these routes are not considered covered, even though a recharging station may be a few minutes off the shortest path.

Another assumption of the original FRLP is that the capacity of facilities is unlimited. This is not too big of an issue in networks where the flow volumes are relatively small, but especially in densely-populated areas and/or during peak hours, capacity constraints might congest the refueling network. Moreover, recharging electric and hybrid to full battery power can take up to several hours (Yilmaz & Krein, 2013), which increases the risk of congestion. Therefore, Upchurch et al. (2009) introduced the capacitated FRLP. This model stipulates that each facility can cover up to certain number of vehicles. They do not only add a capacity constraint, they also change the structure of the facilities. The yes/no decision of opening a facility is replaced by one or more interchangeable modular units of capacity that can be installed at a location. This adjustment seems very plausible in the context of AFVs, as many users own single charging units at home or at work. Hosseini et al. (2017) combine this approach with the aforementioned possibility of deviation.

The analysis of Jung et al. (2014) goes a step further when it comes to the capacitated problem. Instead of simply considering a maximum number of vehicles that can be covered by a facility in total, they adopt a simulation-optimisation framework that minimises the average queueing delay at facilities. The advantage of this approach is that vehicles now have the option to wait at a charging station that is full at arrival until other vehicles leave again. Besides that, it makes the occupancy of the charging station time-dependent, which is reasonable to assume as traffic is heavier during certain hours of the day. Both aspects are disregarded by the capacitated FRLP, which determines beforehand whether a flow can be covered by a set of facilities or not.

Miralinaghi et al. (2017) take a two-fold approach to capacity. First, they assign a parameter to each facility that is defined as the maximum capacity of that facility. Second, a set of decision variables is introduced that determines the level of operation as a proportion of the maximum capacity. A flexible implementation of capacity opens the door to cost savings, since facilities can reduce operation costs by running facilities below capacity during quiet hours.

3 Methodology

This section elaborates upon the deterministic and stochastic models introduced by De Vries and Duijzer (2017). Furthermore, a more efficient formulation by Boujelben & Gicquel (2019) for the deterministic FRLP is discussed. Based on the latter, a capacitated model is created that makes use of the interchangable modular unit representation of capacity inspired by Upchurch et al. (2009). This extension will be treated in Section 4.

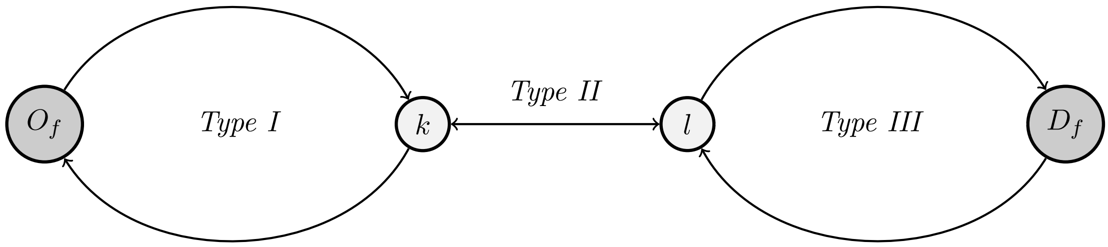
Several important assumptions are made about the problem that apply to all models that will be examined in this section. First of all, the battery or fuel consumption of vehicles is assumed to be proportional to the traveled distance. This simplification alleviates the need for in-depth technical knowledge about battery performance and deterioration over time. Furthermore, charging stations are accessible from both directions of travel. Although it is often the case that a refueling station along a highway is only accessible in one direction, facilities are typically located at both sides of the road. Finally, the charging stations are assumed to be uncapacitated.

3.1 Deterministic Flow Refueling Location Problem

The cornerstone of our analysis regarding refueling facilities is the FRLP. It considers a set of locations L, which consists of origins O, destinations D and potential refueling locations K. A binary decision variable x_k is introduced which equals 1 if a facility is opened at location $k \in K$, and 0 otherwise. Each distinct origin-destination pair is called a flow, with F denoting the set of all flows. A flow f has an origin O_f, destination D_f, and a number of refueling facilities K_f on its way. It is assumed that drivers along flow f travel infinitely often from O_f to D_f and back with volume v_f.

The formulation adopted by De Vries and Duijzer (2017) models the flows as cycles. These can be divided into three types of subcycles: (I) the segment from the opened facility $k \in K_f$ that is closest to the origin, to O_f and back, (II) the segment between two consecutively opened facilities $k, l \in K_f$ along the flow in the direction of the destination and vice versa, and (III) the segment from the opened facility $l \in K_f$ that is closest to the destination, to D_f and back. Figure 1 gives an illustration of the cycle segment types.

Figure 1: *Example of a flow f from origin O_f to destination D_f that runs through recharging facilities k and l. This flow contains a cycle segment of type I ($k \longrightarrow O_f \longrightarrow k$), two cycle segments of type II ($k \longrightarrow l$) and ($l \longrightarrow k$) and a cycle segment of type III ($l \longrightarrow D_f \longrightarrow l$).*



To indicate whether two locations k and l form a cycle segment on flow f, a binary decision variable i_{klf} is introduced that equals 1 if this is the case, and 0 if not. Let t_{kl} be the length of the cycle segment formed by locations $k \in L_f$ and $l \in L_f$ on flow f. It is important to note that this is not necessarily equal to the distance of the shortest path between k and l. For example, if we consider the cycle segment between O_f and k in Figure 1, $t_{O_f k}$ represents the length of segment $k \longrightarrow O_f \longrightarrow k$ and thus equals twice the distance of the shortest path between these two nodes. The definition of the entire set of parameters t_{kl} is defined as follows:

- $k \in K_f, l \in K_f$: t_{kl} equals the distance of the shortest path between k and l, i.e. the length of cycle segment $k \longrightarrow l$.
- $k = O_f, l \in K_f$: $t_{O_f l}$ equals *twice* the distance of the shortest path between k and l, i.e. the length of cycle segment $l \longrightarrow O_f \longrightarrow l$.
- $k \in K_f, l = D_f$: t_{kD_f} equals *twice* the distance of the shortest path between k and l, i.e. the length of cycle segment $k \longrightarrow D_f \longrightarrow k$.
- $k = O_f, l = D_f$: $t_{O_f D_f}$ equals M, a large number that exceeds the driving range.

It is important to count the entire length of the cycle segment, because the full segment will be crossed without encountering a new facility. Therefore, to consider a flow 'covered', the driving range R of the vehicles must be larger than the length of each of the segments of the flow as defined above. Otherwise, the vehicles on the flow run out of fuel before reaching the next refueling facility. A binary decision variable y_f is used to indicate whether flow f is indeed covered.

Even though the actual distance of the shortest path between origin and destination may be smaller than the driving range, a flow will never be covered if there are no recharging stations along the way. This is caused by the assumption that vehicles on a flow make an infinite number of trips between origin and destination. To avoid that the mathematical programming model considers a flow covered without any opened facilities along the way, $t_{O_f D_f}$ is set to a large number M.

Table 1 gives an overview of all the notation used:

Table 1: *Notation of sets, parameters and decision variables in the FRLP*

	Notation	Description
Sets	L	set of locations
	L_f	set of locations along flow f
	L^a_{kf}	set of locations along flow f **after** location k
	L^b_{kf}	set of locations along flow f **before** location k
	K	set of potential recharging locations
	K_f	set of potential recharging locations along flow f
	O	set of origins
	D	set of destinations
Parameters	p	number of new facilities
	R	driving range
	M	large number
	O_f	origin of flow f
	D_f	destination of flow f
	v_f	vehicle volume on flow f
	t_{kl}	length of the shortest path between location k and location l
Decision variables	x_k	1 if a new facility is installed at location k, 0 if not (binary)
	y_f	1 if flow f is covered by the recharging stations, 0 if not (binary)
	i_{klf}	1 if locations k and l form a cycle segment of flow f, 0 if not (binary)

The resulting maximisation problem is formulated as follows:

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

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

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

$$\sum_{l \in L_f^a} i_{klf} = x_k \qquad \forall f \in F, k \in K_f \tag{4}$$

$$\sum_{l \in L_f^a} i_{klf} = 1 \qquad \forall f \in F, k \in O_f \tag{5}$$

$$\sum_{k \in L_f^b} i_{klf} = x_l \qquad \forall f \in F, l \in K_f \tag{6}$$

$$\sum_{k \in L_f^b} i_{klf} = 1 \qquad \forall f \in F, l \in D_f \tag{7}$$

$$i_{klf} \in [0, 1] \qquad \forall f \in F, k \in L_f, l \in L_{kf}^a \tag{8}$$

$$x_k, y_f \in \mathbb{B} \qquad \forall f \in F, k \in K \tag{9}$$

Objective function (1) maximises the flow volume covered. Constraint (2) ensures that only p refueling facilities will be opened. The driving range constraint is captured by (3), where M represents a large number. Furthermore, constraints (4) to (8) stipulate that i_{klf} is set to 1 if facilities $k, l \in K_f$ form a cycle segment of flow f according to the presented definition of a cycle segment (De Vries et al., 2014). Finally, constraints (9) protect the binary nature of the decision variables.

3.2 Stochastic models

The previous model assumes that the driving range R is fixed for every vehicle. However, there are many factors that affect the driving range in practice, such as the weather conditions, traffic congestion, driving speed and tire wear (X. Wu et al., 2015). To allow the driving range to be different under these circumstances, De Vries and Duijzer (2017) propose two formulations that model it as a function $R(\boldsymbol{w})$ of random variables $\boldsymbol{w}$. For the sake of simplicity, it is assumed by the researchers that the driving range $R(\boldsymbol{w})$ is the same across all cycle segments along a flow f, given a realisation of $\boldsymbol{w}$.

The following additional notation is introduced:

Table 2: Additional notation of parameters and decision variables in the EFRLP and CCFLRP

	Notation	Description
Parameters	α	threshold probability of running out of fuel on a round-trip from origin to destination and back
	g_{kl}	probability of running out of fuel on the shortest path between location k and location l
Decision variables	z_f	probability of not running out of fuel on a round-trip from O_f to D_f and back

3.2.1 Expected Flow Refueling Location Problem (EFRLP)

The first of two stochastic models optimises the expected vehicle flow that is covered. It does so by introducing the variable z_f, which is defined as the probability that vehicles along flow f will not run out of fuel. This probability depends on the distribution of the driving range of the vehicles and the length of the segments that have to be travelled. Parameter g_{kl} is equal to the probability that the driving range is smaller than t_{kl}. Given a certain realisation of the random variables $\boldsymbol{w}$, the driving range $R(\boldsymbol{w})$ is the same on each cycle segment. This implies that if the driving range is large enough to cover the longest cycle segment of a flow, it is definitely large enough to cover all other segments. The probability of running out of fuel on the entire flow therefore equals the probability of running out of fuel on the longest segment within the flow. In the mathematical model, this is ensured by introducing Constraints (12). For every location $k \in \{O_f \cup K_f\}$ on flow f, we pick the probability g_{kl} of the cycle segment that starts at k by summing over all possible cycle segments ($\sum_{l \in L_f^a} i_{klf} g_{kl}$). We defined g_{kl} as the probability of running out of fuel on the shortest path from location k to l, so the reverse probability $1 - g_{kl}$ is the probability of finishing the trip from k to l. Hence, the probability z_f of finishing the entire round-trip from O_f to D_f on flow f is always less than or equal to each of the probabilities $1 - g_{kl}$ of the cycle segments of f.

All in all, the model turns out as follows:

$$
\begin{aligned}
\max \quad & \sum_{f \in F} v_f z_f && && (10) \\
\text{s.t.} \quad & \sum_{k \in K} x_k = p && && (11) \\
& (4) - (8) && && \\
& \quad z_f \leq 1 - \sum_{l \in L_f^a} i_{klf} g_{kl} && \forall f \in F, k \in \{O_f \cup K_f\} && (12) \\
& \quad z_f \geq 0 && \forall f \in F && (13) \\
& \quad x_k \in \mathbb{B} && \forall k \in K && (14)
\end{aligned}
$$

The objective function (10) maximises the expected flow covered by multiplying the volume of a flow by the probability that it will not run out of fuel, for every flow $f \in F$. Constraint (11) serves the same purpose as in the standard FRLP model, just as constraints (4) - (8). Moreover, constraints (13) stipulate that the probability of not running out of fuel is non-negative. Note that constraint (12) already puts an upper bound of 1 on z_f.

3.2.2 Chance Constrained Flow Refueling Location Problem (CCFRLP)

An alternative way of including stochasticity is to use another definition of a covered flow. In the CCFRLP, a flow is covered if the probability of running out of fuel on the entire round-trip from origin to destination is lower than a certain threshold α. The probability of running out of fuel on a single cycle segment from k to l is still defined by parameter g_{kl}.

Again, it is sufficient to look at the largest cycle segment in order to check whether a flow has a low enough probability of running out of fuel to consider it covered. Constraints (17) require a flow to be covered (i.e. $y_f = 1$) if and only if the probability that the driving range is insufficient to cover a cycle segment, is always smaller than or equal to the threshold α. The left-hand side of these constraints takes on the value of this probability for the cycle segment of flow f that starts at k by summing over all potentially active segments ($\sum_{l \in L_f} i_{klf} g_{kl}$). Because at most one cycle segment of flow f starts at location k, the left-hand side will equal g_{kl} if and only if k and l form a cycle segment. This probability has to be smaller than α to make it possible for y_f to take on the value of 1.

Apart from Constraints (17), the formulation of the CCFRLP is very similar to the one used for the FRLP.

$$\max \quad \sum_{f\in F} v_f y_f \tag{15}$$

$$\text{s.t.} \quad \sum_{k\in K} x_k = p \tag{16}$$

$$\sum_{l\in L_f} i_{klf} g_{kl} \leq \alpha + (1 - y_f) \qquad \forall f \in F, k \in \{O_f \cup K_f\} \tag{17}$$

$$(4) - (8)$$

$$x_k, y_f \in \mathbb{B} \qquad \forall k \in K, f \in F \tag{18}$$

3.2.3 Value of the Stochastic Solution (VSS)

Although incorporating stochasticity adds a degree of realism to the refueling location problem, it is not clear-cut what the benefit is in terms of coverage. The Value of the Stochastic Solution (VSS) is introduced to serve this purpose (Birge, 1982). It is defined as the difference between the optimal objective value of the EFRLP and expected value of the solution of the deterministic FRLP. The VSS can thus be regarded as the coverage that is lost by naively using the deterministic solution over the stochastic solution of the EFRLP.

The expected value of the deterministic solution can be obtained by first solving the deterministic FRLP. The result is an optimal solution $\{x_1^\star, x_2^\star, \ldots, x_{|K|}^\star\}$ for the facility allocation. Then, the values of the $\boldsymbol{x}$-variables in the stochastic EFRLP model are being fixed to $\{x_1^\star, x_2^\star, \ldots, x_{|K|}^\star\}$, whereafter we solve it for $\boldsymbol{z}$ and $\boldsymbol{i}$. The resulting objective value is the expected value of the deterministic solution. This value will always be lower than or equal to the objective value of the stochastic solution, as we restrict optimisation with respect to $\boldsymbol{x}$. The difference between these two objective values is the VSS.

3.3 FRLP formulation by Boujelben and Gicquel (2019)

The recently developed model by Boujelben & Gicquel (2019) shifts away from the subcycle method as discussed in Sections 3.1 and 3.2. Instead, it takes a different approach that focuses on the locations themselves. For a vehicle on flow f to arrive at a certain location $l \in L_f$, it must be able to reach that location from the preceding recharging point $k \in L_{lf}^b$. In other words, every location $l \in L_f$ has to be assigned a preceding facility location $k \in L_{lf}^b$ that enables the vehicle to make it up l without running out of fuel. Only if this holds for all locations, flow f can be considered covered. To serve this purpose, a new set of decision variables w_f^{kl} makes its appearance. This variable is equal to 1 if a recharging station is installed at location k that enables the vehicles on flow f to reach at least location l without running out of fuel. In case location $k = O_f$, $w_f^{O_f l}$ equals 1 if a recharging station is installed at l that enables the vehicles on flow f to make the round-trip from l to O_f without running out of fuel.

Apart from the new decision variables w_f^{kl}, the notation used in this model is the same as the one described in Table 1. The corresponding formulation (Boujelben & Gicquel, 2019) is the following:

$$
\begin{aligned}
\max \quad & \sum_{f \in F} v_f y_f & & (19)\\
\text{s.t.} \quad & \sum_{k \in K} x_k = p & & (20)\\
& \sum_{k \in L_{lf}^b} w_f^{kl} = y_f & \forall f \in F, l \in \{K_f \cup D_f\} & \quad (21)\\
& w_f^{kl} \leq x_k & \forall f \in F, k \in K_f, l \in L_{kf}^a & \quad (22)\\
& \sum_{k \in L_{lf}^b} t_{kl} w_f^{kl} \leq R & \forall f \in F, l \in \{K_f \cup D_f\} & \quad (23)\\
& w_f^{kl} \in \mathbb{B} & \forall f \in F, l \in \{K_f \cup D_f\}, k \in L_{lf}^b & \quad (24)\\
& x_k, y_f \in \mathbb{B} & \forall f \in F, k \in K & \quad (25)
\end{aligned}
$$

The objective function (19) still aims to maximise vehicle flow coverage, whereas Constraint (20) sets the number of installed recharging facilities to p. Constraints (21) ensure that a flow can only be covered if every location is assigned to a preceding recharging station. For every flow $f \in F$ and location $l \in \{K_f \cup D_f\}$, it sums over the assignment variables of the preceding locations $k \in L_{lf}^b$. If none of these locations is assigned to l, the flow will not be covered (i.e. $y_f = 0$). One of the assignment variables has to be equal to 1 in order to set y_f to 1. Because of the binary nature of y_f, only a single location can be assigned to l.

To ascertain that only opened facility sites can be assigned to other locations, Constraints (22) are invoked. This is done by limiting the assignment variables to be less than or equal to the facility installment variables, such that w_f^{kl} can only be equal to 1 if x_k equals 1 as well. Furthermore, Constraints (23) requires that the assignment variable w_f^{kl} can only be 1 if the driving range R does not exceed the driving distance t_{kl} of the segment formed by locations k and l. The definition of the distance is the same as in Section 3.1. Finally, Constraints (24) and (25) stipulate that all decision variables are binary.

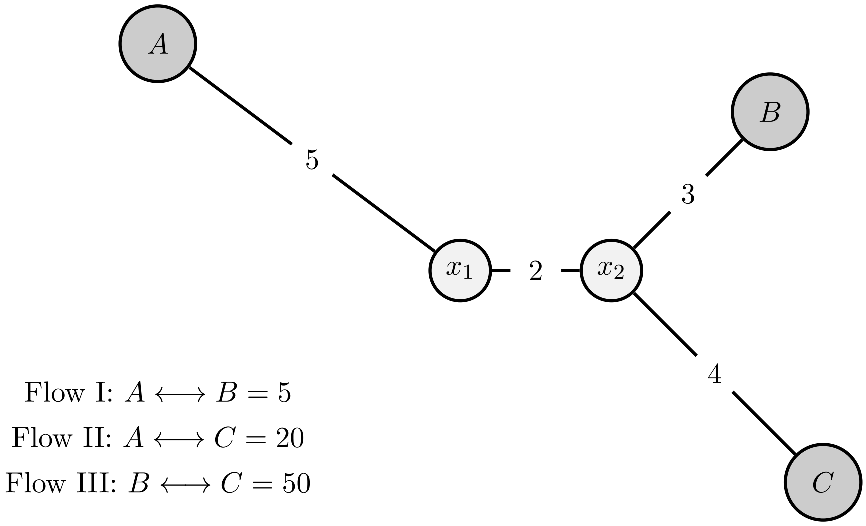
4 Extension

The models discussed in Section 3 all rely on the assumption of unlimited facility capacity. No matter how large a vehicle flow is, it will be covered as long as consecutive recharging stations can be reached within the vehicle's driving range. When vehicle volumes are relatively small and maximum capacity is barely attained, assuming uncapacitated facilities in the model is unlikely to affect the solution radically. However, with the forecasted growth of AFVs in the upcoming decades, recharging facilities run a higher risk of congestion. In that case, uncapacitated models will still deem vehicles covered, despite the facilities on their path having reached maximum capacity. To avoid suboptimal outcomes correspondingly, a capacitated model is desired.

Modelling capacity can be done in several ways. In our model, we are inspired by the approach of Upchurch et al. (2009), who model capacity as interchangeable modular units. These units can be thought of as single gas or petrol pumps at a filling station or single charging points for electrical vehicles. Apart from the clear interpretation of this approach, it allows for much more flexibility in constructing a recharging network. De Vries & Duijzer (2017) already showed that synergy effects exist when the number of facilities is increased, as multiple facilities are often needed to fully cover a flow. In this line of reasoning, many small units of capacity would be preferred over fewer large units. Furthermore, the size of facilities can be varied to more accurately align with the amount of capacity demanded at certain locations or on certain flows.

We can demonstrate the potential benefits of the modular unit approach by means of a simple example. Let us consider the network in Figure 2 with three O-D nodes (A, B and C) and two potential recharging station sites (x_1 and x_2). The vehicle flow volumes and path lengths are shown in the Figure, the driving range of vehicles is set to 10 and total capacity will equal 100.

Figure 2: *Example of a network with 3 O-D nodes (A, B and C) and two potential facility locations (x_1 and x_2). The values on the edges denote the length of the corresponding edge. The left-bottom corner shows the vehicle volumes on the flows.*



In order for flow I to be covered, x_1 must contain a facility with a capacity of at least 5. If only x_2 is opened, the length of the cycle segment $x_2 \longrightarrow A \longrightarrow x_2$ will exceed the driving range ($14 \geq 10$) and flow I will not be covered. Moreover, to cover flow II, both x_1 and x_2 must be opened and have a capacity of at least 20. Finally, x_2 is the only facility on flow III and must hence be installed with minimum capacity of 50 to cover it.

Now, let us consider the case that we have one unit of capacity 100 to allocate. Obviously, it would be installed at x_2 such that the largest flow (III) with volume 50 can be covered. Placing it at x_1 only yields a coverage level of 5, with flow I being the only covered flow. Synergy effects already arise when we split up the total capacity of 100 into two modular units of capacity 50. By allocating one unit to both x_1 and x_2, flow I and III can both be covered with a resulting total coverage level of 55. Flow II cannot be covered yet, since the maximum capacity of 50 at x_2 is reached by covering flow III through this location. However, most of the capacity at x_1 is redundant as it is only used to cover a flow (I) of volume 5. An even more efficient distribution of capacity can be obtained through splitting up the modular units once more into four units of capacity 25. Allocating one unit to x_1 and three units to x_2 yields full coverage of all flows in the network. Facility x_1 with capacity 25 serves flow I (volume of 5) and II (volume of 20), whilst facility x_2 with capacity 75 serves all flows (5 + 20 + 50).

4.1 Comparison of FRLP formulations

When introducing capacity constraints into our FRLP framework, the question arises which FRLP formulation would be the best candidate to extend. The reason why we pick the FRLP of Boujelben & Gicquel (2019) instead of the model of De Vries & Duijzer (2017) lies in the definition of the cycle segments. If a certain facility is opened, all flows that pass this facility will stop here to recharge due to Constraints (4) and (6). This may cause flows to be uncovered due to the capacity constraints of the facilities, even though they could be covered if there was the opportunity to skip facilities that are full. The advantage of the assignment variable in Boujelben and Gicquel's approach is that facilities are assigned for each flow seperately and can be set to 0 despite the location containing one or more recharging units. Note that in both models specifications, a flow can only be covered if the recharging stations on its path have enough capacity for the flow's entire vehicle volume. If not, flows cannot be split up and will hence be fully uncovered.

Moreover, Boujelben & Gicquel (2019) have shown in their research that their formulation has significantly smaller running times than De Vries and Duijzer. This is desirable, as the incorporation of capacity constraints and potentially many small units of capacity will make the problem more complex and therefore computationally more challenging to solve.

4.2 Unit-Capacitated Flow Refueling Location Problem (UCFRLP)

Given the formulation of the FRLP in Section 3.3, including capacitated facilities is relatively straightforward. First of all, the number of facilities to be installed, p, is replaced by the number of interchangeable modular units that have allocated, u. Each unit has a capacity of c, such that the total capacity of the network equals cu. There is no limit on the amount of locations that has to be utilised or the maximum number of units per location. To serve this purpose, decision variable x_k will be integer-valued and non-negative from this moment onward.

The number of vehicles that recharges at a facility location $k \in K$ is measured by taking the sum over all flow volumes of flows that make use of k as a recharging station. A new decision variable s_f^k is introduced, which is equal to 1 if location k is used by flow f to recharge, and 0 if not. To guarantee that s_f^k takes on the correct values, Constraints (29) stipulate that it is equal to 1 if assignment variable w_f^{kl} equals 1 as well. This will be the case if location k is assigned to serve any other location l on flow f.

The result is the following Unit-Capacitated Flow Refueling Location Problem (UCFRLP):

$$\max \quad \sum_{f \in F} v_f y_f \tag{26}$$

$$\text{s.t.} \quad \sum_{k \in K} x_k = u \tag{27}$$

$$\sum_{f \in F} v_f s_f^k \leq c x_k \qquad \forall k \in K \tag{28}$$

$$s_f^k \geq w_f^{kl} \qquad \forall f \in F, k \in K_f, l \in L_{kf}^a \tag{29}$$

$$(21) - (24)$$

$$s_f^k \in \mathbb{B} \qquad \forall f \in F, k \in K \tag{30}$$

$$x_k \in \mathbb{Z}_{\geq 0} \qquad \forall k \in K \tag{31}$$

$$y_f \in \mathbb{B} \qquad \forall f \in F \tag{32}$$

The formulation of the UCFRLP has many similarities with the FRLP of (Boujelben & Gicquel, 2019). The objective function remains unchanged and Constraints (21) - (24) are recycled as well. Constraint (27) ensures that exactly u units will be installed. The capacities are enforced by Constraints (28) for every location $k \in K$. The left-hand side sums over all vehicle volumes that recharge at location k ($\sum_{f \in F} v_f s_f^k$). This constitutes the total outgoing flow at k, which cannot exceed the combined capacity of all installed units at this location (i.e. cx_l). Moreover, the new decision variables take on their desired values through Constraints (29) as explained earlier this section. Lastly, Constraints (30), (31) and (32) define the decision variables.

5 Results

5.1 Problem instances

All problem instances have been solved by means of CPLEX. The programming classes used are explained in Appendix B. The results are obtained by using randomly generated networks as input for the models (De Vries & Duijzer, 2017). A total of X potential refueling locations are randomly drawn from a uniform distribution on a $[0, 1000]^2$ grid. Then, a minimum spanning tree is created amongst all locations by means of Kruskal's algorithm. Out of the X locations, Y are replicated to be our Origin-Destination (O-D) nodes. This gives us a total of Y(Y - 1)/2 O-D pairs, of which the shortest path within each pair is determined using Dijkstra's algorithm.

After generating the network, the vehicle volume of each flow v_f is calculated with Equation (33). In this formula, e_f^O and e_f^D denote the random numbers drawn for the standard uniform distribution that are assigned to the origin and destination node of flow f, respectively. Furthermore, T_f is the distance of the shortest path between these nodes. The indicator function at the end of the equation stipulates that the vehicle volume is set to 0 if this distance is smaller than 100. After the volumes have been computed, they are normalised such that they add up to 10^6.

$$v_f = \frac{e_f^O e_f^D}{T_f} * 1_{T_f \geq 100} \tag{33}$$

Finally, α is set to 0.05 and the driving range follows a Gamma distribution. The shape and scale parameter are equal to 50 and 5, respectively, such that the expected value equals 250. This will be the number used for the fixed driving range in the deterministic model.

5.2 Comparison of the deterministic and stochastic models

This section dives into the solutions of the EFRLP, CCFRLP and deterministic FRLP (or DFRLP). All models have been solved for three problem instances (80/40, 60/30 and 40/20), where the first number denotes the amount of potential recharging locations and the second number the amount of O-D nodes. This is repeated 9 times for different values of number of installed facilities p (1, 2, 3, 4, 5, 10, 15, 20 and 25).

Table 3 depicts the average outcomes for the optimality gap and Value of the Stochastic Solution (VSS) over all values of p for every problem instance. The optimality gap is defined as the percentage of coverage decrease when the solutions of a model are evaluated in terms of expected coverage and chance constrained coverage. The optimality gap of the EFRLP solution in terms of expected coverage and of the CCFRLP solution in terms of chance constrained coverage are naturally equal to zero. The VSS yields the difference between the EFRLP and DFRLP solution's coverage in terms of expected flow volume. More elaborate results can be found in Appendix A.

Table 3: *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 (p = 1, 2, 3, 4, 5, 10, 15, 20, 25).*

	Opt. gap in terms of expected flow volume covered (%)			Opt. gap in terms of chance constr. flow volume covered (%)			
Instance	EFRLP Solution	CCFRLP Solution	DFRLP Solution	EFRLP Solution	CCFRLP Solution	DFRLP Solution	VSS
80/40	0.00	15.14	6.92	28.26	0.00	62.77	3.21E+04
60/30	0.00	12.68	7.94	20.19	0.00	46.86	4.20E+04
40/20	0.00	6.55	6.07	11.33	0.00	30.95	4.33E+04

The first thing that stands out in Table 3 is the difference between the optimality gaps in terms of expected flow volume covered and in terms of chance constrained flow volume covered. The solution of the CCFRLP yields much smaller optimality gaps in terms of expected coverage than the solution of the EFRLP in terms of chance constrained coverage. This can be explained by looking at the coverage definitions. Chance constrained coverage only counts flows with at least 95% (in our case of $\alpha = 0.05$) coverage, whereas expected coverage also considers flows with much lower coverage. Counting flows with low coverage affects the optimal allocation of recharging stations, but does not lead to coverage when measured by the chance constrained model. Alternatively, the covered flows in the CCFRLP solution will contribute to coverage in the expected flow model too.

Furthermore, the deterministic model performs much better in terms of expected coverage as opposed to chance constrained coverage. We know that in the deterministic solution, the maximum distance of an active cycle segment of a covered flow is at most equal to the expected value of the driving range. In the stochastic models, this solution will therefore always have a substantial coverage. Although this will always contribute to expected coverage, only the extremely well covered flows will add to chance constrained coverage. This explains the large differences in optimality gaps for the DFRLP solution in Table 3.

Finally, the Value of the Stochastic Solution shows that using a stochastic driving range yields an increase in coverage of roughly 32.100, 42.000 and 43.300 for the 80/40, 60/30 and 40/20 instances respectively. If we consider that the total flow volume in the network equal

1.000.000, these numbers can be translated to 3.21%, 4.20% and 4.33% of the total flow volume that is covered under the EFRLP solution, but not under the DFRLP according to the expected coverage definition.

5.3 Synergy effects

Similar to De Vries & Duijzer (2017), we investigate the possible existence of synergy effects in our network. This requires a closer look into the increase in coverage when the number of facilities increases. Figure 3 depicts the coverage levels in terms of expected flow volume covered of the DFRLP, EFRLP and CCFRLP solutions for 9 different values of p (1, 2, 3, 4, 5, 10, 15, 20 and 25).

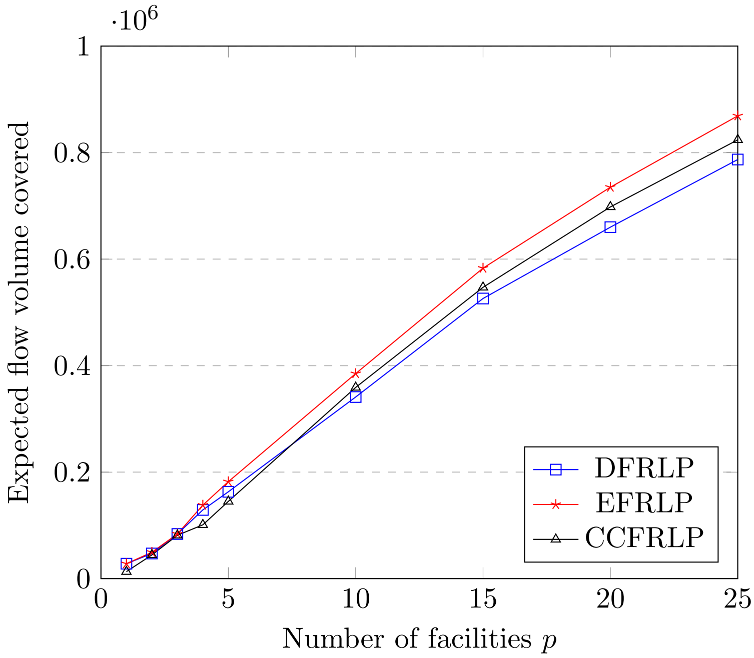


Figure 3: *Expected flow volume covered of the DFRLP, EFRLP and CCFRLP solutions for the 80/40 instance*

Let us take a look at the coverage levels of the EFRLP solution. Increasing the number of facilities from 1 to 2 yields a coverage increase of 21428, whereas adding another facility causes 34326 extra vehicles to be covered. A fourth facility would be even more efficient, as this will yield another 53842 covered vehicles. This shows that synergy effects exist in terms of expected flow volume covered.

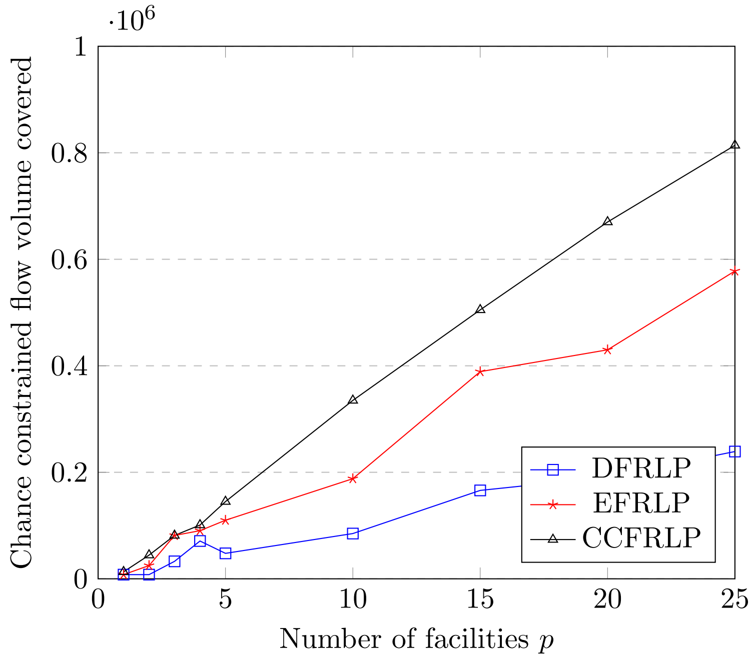


Figure 4: *Chance constrained flow volume covered of the DFRLP, EFRLP and CCFRLP solutions for the 80/40 instance*

Figure 4 is similarly constructed as Figure 3, but then for the chance constrained flow volume covered. We can derive that synergy effects are present as well when this definition of coverage is used. For example, the coverage of the CCFRLP solution increases with 32104 when p is incremented from 1 to 2, and with 37260 when another facility is added.

A reason that synergy effects arise is that often multiple facilities are needed to cover a flow. Especially when the number of facilities is low, adding one more recharging station can make a lot of difference for flows for which a single facility does not suffice to cover the entire round-trip from origin to destination. When deciding on a number of facilities to install, it is therefore worthwhile to consider more than one facility.

5.4 Capacitated model

For the capacitated model, we want to investigate whether a more efficient allocation of modular capacity units can be obtained if more yet smaller units are used, without increasing the combined capacity of the units. In other words, the total capacity cu will be kept constant, whilst varying the number of units u. It must be noted that a decrease in coverage is impossible when u is doubled, as a solution can always be reconstructed with double the amount of units. This can be done by assigning two half-sized units to a facility location for every unit that was allocated to it in the previous solution.

The simple example in Section 4 already gives a rationale for examining synergy effects in a capacitated context. In order to do this properly, we start with 1 unit of capacity and split it up into twice the number of units of half the capacity until we have 32 units. Given certain values for c and u, the UCFRLP is solved to optimality for the 80/40 instance. This procedure is repeated for a total capacity of 10.000, 20.000, 50.000, 100.000 and 200.000. The results are depicted in Figure 5.

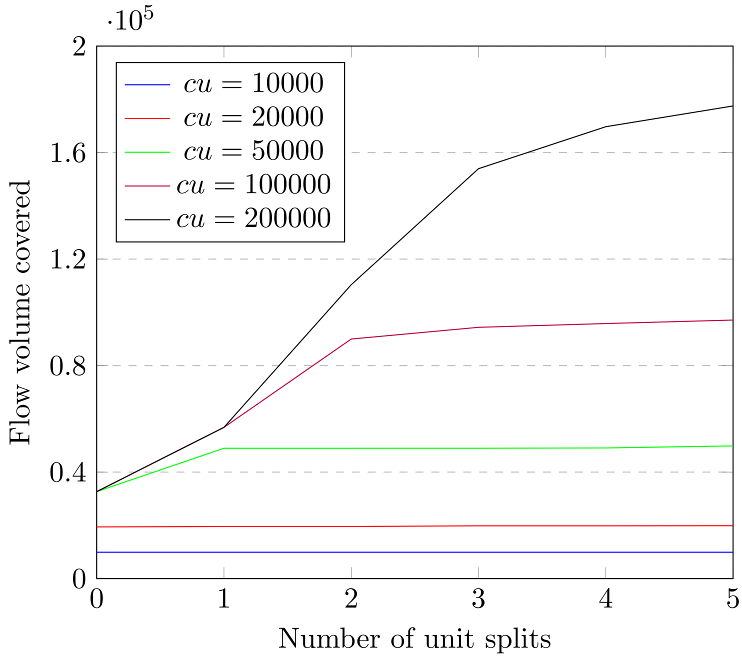


Figure 5: *Flow volume covered of the UCFRLP for the 80/40 instance. The unit splits correspond to the total number of units in the following way: number of units* $= 2^{\textit{number of unit splits}}$. *Driving range* $R = 250$.

Table 4 provides a more detailed overview of the coverage increase per unit split and the number of installed facilities per solution.

Table 4: *Coverage gained by splitting up the modular units compared to the situation with half the number of units with double capacity. The number of facility locations is denoted inbetween brackets. Driving range* $R = 250$.

		Number of units					
		1	2	4	8	16	32
Total capacity	10.000	-	0% (1)	0% (1)	0% (1)	0% (1)	0% (1)
	20.000	-	0.77% (2)	0% (2)	1.34% (2)	0% (2)	0.20% (2)
	50.000	-	49,86% (2)	0% (2)	0,02% (2)	0,23% (4)	1,55% (3)
	100.000	-	73,92% (2)	58,38% (4)	4,89% (5)	1,50% (4)	1,37% (6)
	200.000	-	73,92% (2)	94,29% (4)	39,48% (8)	10,24% (11)	4,59% (13)

The main observation from Figure 5 and Table 4 is that the efficiency gains are larger when total capacity is larger as well. Furthermore, the benefit of splitting up seems to decrease in the number of unit splits. An increase in coverage can have two major reasons. Firstly, coverage can increase due to redundant capacity. Especially when total capacity is large and the number of units is small, most of the capacity cannot be used as only a limited amount of flow volume can be covered from one location. As a result, the coverage level of the 50.000, 100.000 and 200.000 instances is equal to the DFRLP coverage for one facility, whereas the coverage of the 100.000 and 200.000 instances is equal to the DFRLP coverage for two facilities, too. A larger number of units allows rebalancing capacity among opened facility locations. An example of this phenomenon is the increase from 8 to 16 units with a total capacity of 100.000. Instead of 5 facility locations with 8 units, only 4 locations are opened with 16 units due to the possibility of more flexible assigning capacity. Secondly, a wider spread of the facility network contributes to extra coverage. Flows that could not be served with only a few facilities can be covered when more units are installed along their pre-planned path. This is clearly demonstrated for the instance of 200.000 total capacity, where the number of used locations increases with every split.

In general, the instances with more total capacity benefit more from splitting up, as this allows them to fully utilise their redundant capacity. The smaller instances already approach the full potential of their capacity with a few units. A possible cause of this is the presence of multiple flows that can be covered with only one facility. Consequently, the efficiency gain of splitting up units is capped.

5.4.1 Sensitivity analysis on the driving range

The structure of the network might have an effect on the magnitude of synergy effects. Larger networks with longer flows will typically require more facilities per flow in order to cover it. Instead of increasing the size of the network, the same effect can be obtained by decreasing the driving range. We perform the same analysis as in Section 5.4 with half the original driving range, resulting in $R = 125$.

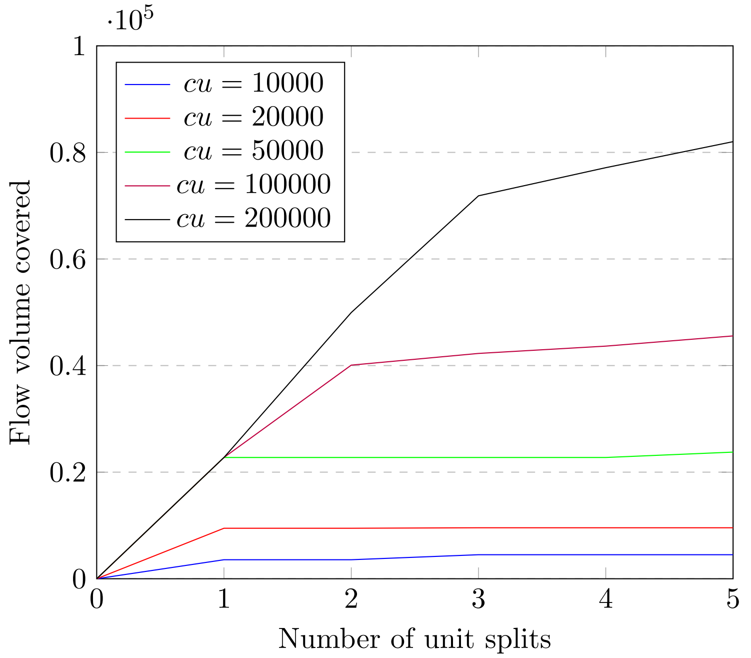


Figure 6: *Flow volume covered of the UCFRLP for the 80/40 instance. The unit splits correspond to the total number of units in the following way: number of units* $= 2^{\text{number of unit splits}}$. *Driving range* $R = 125$.

Table 5 provides a more detailed overview of the coverage increase per unit split and the number of installed facilities per solution.

Table 5: *Coverage gained by splitting up the modular units compared to the situation with half the number of units with double capacity. The number of facility locations is denoted inbetween brackets. Driving range* $R = 125$.

		Number of units					
		1	2	4	8	16	32
Total capacity	10.000	-	-	0% (2)	26.36% (4)	0% (4)	0% (4)
	20.000	-	-	0% (2)	0.95% (4)	0% (4)	0% (4)
	50.000	-	-	0% (2)	0% (2)	0% (2)	4.42% (4)
	100.000	-	-	76.13% (4)	5.48% (5)	3.24% (6)	4.39% (7)
	200.000	-	-	119,48% (4)	43,87% (7)	7,34% (11)	6,33% (11)

It can be derived from Figure 6 and Table 5 that most of the efficiency gains are larger compared to the case with twice the driving range. Furthermore, for most problem instances the number of opened facilities has increased. This can be explained by the fact that there is both more excess capacity and need for more facilities due to a decreasing number of flows that can be covered with only 1 or 2 facilities.

5.4.2 Quality of the uncapacitated solution

In our final part of the analysis, we dive into the quality of the 'naive' uncapacitated solution. In a similar fashion as the VSS, we set the values of the unit installment variables ($\boldsymbol{x}$) to the optimal facility allocation of the DFRLP. This entails that at each facility that is opened in the DFRLP, one unit is installed at that location in the UCFRLP. The resulting coverage can be compared with the coverage of UCFRLP solution, as is done in Table 6.

Table 6: *Average optimality gap and covered lost of the uncapacitated DFRLP solution in terms of the UCFRLP over 5 total capacities (cu* $= 10.000, 20.000, 50.000, 100.000, 200.000$*).*

		R = 250		R = 125	
		Optimality gap (%)	Lost coverage	Optimality gap (%)	Lost coverage
Number of units	1	0.22%	56	0%	0
	2	9,12%	3503	4,69%	763
	4	16,97%	9463	26,85%	6757
	8	38,46%	25155	46,38%	14004
	16	43,24%	29776	58,67%	18496
	32	51,88%	36752	67,42%	22303

The uncapacitated solution clearly reduces in quality as the number of units increases. This can be explained by the capacity of each unit, which decreases in the number of units. The smaller the capacity of each facilities, the more large flows cannot be covered anymore, which are the most interesting in terms of coverage for the uncapacitated model. Hence, if many units have to be allocated, the UCFRLP is highly recommended over the DFRLP. For smaller problem instances though, the uncapacitated model performs relatively well.

6 Conclusion

In this paper, we have dived deeper into several instances of the Flow Refueling Location Problem. Although deterministic formulations of this model provide a good starting point for further analysis into location-allocation problems for recharging facilities, they yield a number of unrealistic assumptions. One of these assumptions, the fixed driving range of vehicles, is tackled by De Vries & Duijzer (2017). They introduce two models, the EFRLP and the CCFRLP, that incorporate the stochasticity of the driving range. The EFRLP aims to maximise expected coverage given the uncertainty in the driving range, whereas the CCFRLP maximise the number of vehicles that have a probability of at least $1 - \alpha$ of finishing their trip without running out of fuel.

Analysis of the stochastic models shows that decision-makers benefit from using the optimal solutions of these models over the deterministic solution. The Value of the Stochastic Solution indicates that expected coverage increases substantially with the use of the stochastic solution. Besides that, the CCFRLP solution performs relatively well in terms of expected coverage, whereas the EFRLP solution faces higher optimality gaps in terms of chance constrained coverage. These results point in favour of using the CCFRLP to allocate recharging facilities, although the decision should mostly be based on the desired definition of coverage. Vehicles that fully rely on alternative fuel do not want to risk running out of fuel, which makes the chance constrained definition more applicable. Expected coverage might be more interesting in case of hybrid vehicles. High coverage is desired here, but not a necessity as vehicles can switch to more traditional power sources when running out of their alternative power source.

One of the disadvantages of the EFRLP and CCFRLP is the assumption of uncapacitated facilities. Although a network with small vehicle volumes relative to the facilities' capacity might not benefit a lot from accounting for capacity, future large-scale use will face issues with congested facilities if capacity is not modelled. To serve this purpose, we introduce the Unit-Capacitated Flow Refueling Location Problem, based on the FRLP formulation of Boujelben & Gicquel (2019). Besides the fact that their formulation is computationally more efficient than the formulation of De Vries and Duijzer, their use of assignment variables instead of cycle segment variables is more appropriate for a clear implementation and interpretation of capacity constraints.

The notion of capacity that is used in the UCFRLP is one of interchangeable modular unit as explored by Upchurch et al. (2009). This representation allows us to flexibly allocate capacity over opened facilities, which opens the door to reallocating excess capacity in small units and creating a more widespread network that covers flow that require multiple facilities on their path.

Analysing the EFRLP and CCFRLP shows that such synergy effects do indeed exist, and our investigation of the UCFRLP confirms this for the capacitated model. Efficiency gains appear to be even larger when the driving range is smaller relative to the size of the network. Finally, the uncapacitated solution turns out to perform rather poorly when applied to the capacitated case. When modular units are few and have high capacity, the loss in coverage is limited, but when the full benefit is reaped from the unit representation by modelling many units with low capacity, coverage can drop over 50% when the uncapacitated solution is naively adopted.

One of the disadvantages of our approach of modelling capacities is that every vehicle in a flow is obliged to visit the same facilities. This implies that flows cannot be split up in smaller subflows that can visit different facilities if there does not exist one set of facilities that yields enough capacity to cover the entire flow volume. If this can be realised, coverage will increase and our current unit allocation might be suboptimal. It would be interesting for further research to dive deeper into the possibility of assigning different sets of facilities to the vehicles in one flow. This could be combined with the idea of deviation from pre-planned shortest paths, similar to the approach of Kim & Kuby (2012).

Furthermore, our model does not take into account that vehicle volume within a flow is time-varying. It is unlikely that a facility requires the same level of capacity throughout the day. At night, recharging stations will be much quieter than during traditional peak-hours during the day. Accurate approximations of recharging demand and traffic behaviour could be a valuable addition when multi-period modelling is examined and will add a degree of realism to the problem. Finally, in the same line of reasoning, it is assumed that the capacity of a facility can never be exceeded. In reality however, a vehicle can enter a fully occupied recharging station, wait for a few minutes and then hop into the spot that opened up. Modelling this will probably require a simulation approach that incorporates queueing theory into the problem.

References

Averbakh, I., & Berman, O. (1996). Location flow-capturing units on a network with multi-counting and diminishing returns to scale. *European Journal of Operational Research*, *91*(3), 495-506.

Birge, J. R. (1982). The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming*, *24*(1), 314-325.

Boujelben, M. K., & Gicquel, C. (2019). Efficient solution approaches for locating electric vehicle fast charging stations under driving range uncertainty. *Computers and Operations Research*, *109*, 288-299.

Chapman, L. (2007). Transport and climate change: A review. *Journal of Transport Geography*, *15*, 354-367.

Choi, H., Nakagawa, D., Matsunaka, R., Oba, T., & Yoon, J. (2013). Research on the causal relationship between urban density, travel behaviours, and transportation energy consumption by economic level. *International Journal of Urban Sciences*, *17*(3), 362-384.

Church, R., & Velle, C. R. (1974). The maximal covering location problem. *Papers in regional science*, *32*, 101-118.

De Vries, H., & Duijzer, E. (2017). Incorporating driving range variability in network design for refueling facilities. *Omega*, *69*, 102–114.

De Vries, H., Van de Klundert, J. J., & Wagelmans, A. P. M. (2014). The roadside healthcare facility location problem. *Econometric Institute Research Papers*(EI 2014-09).

Hodgson, M. J. (1990). A flow-caputing location-allocation model. *Geographical Analysis*, *22*, 270-279.

Hodgson, M. J., & Rosing, K. E. (1992). A network location-allocation model trading off flow capturing and p-median objectives. *Annals of Operations Research*, *40*(1), 247-260.

Hodgson, M. J., Rosing, K. E., & Zhang, J. (1996). Locating vehicle inspection stations to protect a transportation network. *Geographical Analysis*, *28*(4), 299-314.

Hosseini, M., MirHassani, S. A., & Hooshmand, F. (2017). Deviation-flow refueling location problem with capacitated facilities: Model and algorithm. *Transportation Research Part D: Transport and Environment*, *54*, 269-281.

Jung, J., Chow, J., Jayakrishnan, R., & Park, J. (2014). Stochastic dynamic itinerary interception refueling location problem with queue delay for electric taxi charging stations. *Transportation Research Part C*, *40*, 123-142.

Kim, J., & Kuby, M. (2012). The deviation-flow refueling location model for optimizing a network of refueling stations. *International Journal of Hydrogen Engergy*, *37*, 5406-5420.

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

Lim, S., & Kuby, M. (2010). Heuristic algorithms for alternative-fuel stations using the flow-refueling location model. *European Journal of Operational Research*, *204*(1), 51-61.

McIlroy, R. C., & Stanton, N. A. (2017). What do people know about eco-driving? *Ergonomics*, *60*(6), 754-769.

Melendez, M. (2006). Transitioning to a hydrogen future: Learning from the alternative fuels experience. *National Renewable Energy Laboratory*.

Miralinaghi, M., Keskin, B. B., Lou, Y., & Roshandeh, A. M. (2017). Capacitated refueling station location problem with traffic deviations over multiple time periods. *Networks and Spatial Economics*, *17*(1), 129-151.

Upchurch, C., Kuby, M., & Lim, S. (2009). A model for location of capacitated alternative-fuel stations. *Geographical Analysis*, *41*, 85-106.

Wu, F., & Sioshansi, R. (2017). A stochastic flow-capturing model to optimize the location of fast-charging stations with uncertain electric vehicle flows. *Transportation Research Part D*, *53*, 354-376.

Wu, X., Freese, D., Cabrera, A., & Kitch, W. (2015). Electric vehicles' energy consumption measurement and estimation. *Transportation Research Part D: Transport and Environment*, *34*, 52-67.

Yilmaz, M., & Krein, P. (2013). Review of battery charger topologies, charging power levels, and infrastructure for plug-in electric and hybrid vehicles. *IEEE Transactions on Power Electronics*, *28*(5), 2151-2169.

A Results

Table 7: *Expected volume covered and chance constrained volume covered for the EFRLP, CCFRLP and DFRLP solution for the 40/20 instance*

p	Optimal objective value and opt. gap in terms of expected flow volume covered (%)					Optimal objective value and opt. gap in terms of chance constrained flow volume covered (%)					VSS
	EFRLP Solution	CCFRLP Solution		DFRLP Solution		EFRLP Solution		CCFRLP Solution	DFRLP Solution		
1	9,55E+04	8,06E+04	-15,64%	9,55E+04	0,00%	6,45E+04	-11,72%	7,31E+04	6,45E+04	-11,72%	0,00E+00
2	1,88E+05	1,67E+05	-11,37%	1,88E+05	0,00%	1,36E+05	-18,52%	1,67E+05	1,36E+05	-18,52%	0,00E+00
3	2,39E+05	2,16E+05	-9,69%	2,28E+05	-4,62%	1,84E+05	-15,52%	2,18E+05	1,36E+05	-37,46%	1,10E+04
4	3,00E+05	2,75E+05	-8,34%	2,76E+05	-7,80%	2,38E+05	-11,93%	2,70E+05	1,87E+05	-30,83%	2,34E+04
5	3,76E+05	3,54E+05	-5,71%	3,76E+05	0,00%	3,26E+05	-8,68%	3,57E+05	3,26E+05	-8,68%	0,00E+00
10	5,72E+05	5,62E+05	-1,70%	5,13E+05	-10,23%	5,15E+05	-8,56%	5,63E+05	3,29E+05	-41,54%	5,85E+04
15	7,21E+05	6,90E+05	-4,24%	6,81E+05	-5,57%	5,50E+05	-18,25%	6,72E+05	3,46E+05	-48,53%	4,01E+04
20	9,56E+05	9,55E+05	-0,15%	8,28E+05	-13,43%	8,66E+05	-8,20%	9,43E+05	5,33E+05	-43,47%	1,28E+05
25	9,90E+05	9,69E+05	-2,10%	8,62E+05	-12,95%	9,79E+05	-0,58%	9,84E+05	6,12E+05	-37,82%	1,28E+05

Table 8: *Expected volume covered and chance constrained volume covered for the EFRLP, CCFRLP and DFRLP solution for the 60/30 instance*

p	Optimal objective value and opt. gap in terms of expected flow volume covered (%)					Optimal objective value and opt. gap in terms of chance constrained flow volume covered (%)					VSS
	EFRLP Solution	CCFRLP Solution		DFRLP Solution		EFRLP Solution		CCFRLP Solution	DFRLP Solution		
1	5,19E+04	5,19E+04	0,00%	5,19E+04	0,00%	2,62E+04	0,00%	2,62E+04	2,62E+04	0,00%	1,00E-02
2	1,00E+05	7,46E+04	-25,59%	1,00E+05	0,00%	3,41E+04	-54,81%	7,55E+04	3,41E+04	-54,81%	0,00E+00
3	1,25E+05	8,68E+04	-30,37%	1,23E+05	-1,17%	8,74E+04	-0,48%	8,78E+04	3,41E+04	-61,12%	1,46E+03
4	1,67E+05	1,42E+05	-14,50%	1,60E+05	-3,81%	1,30E+05	-8,79%	1,42E+05	9,07E+04	-36,12%	6,36E+03
5	2,12E+05	1,95E+05	-8,36%	1,89E+05	-10,90%	1,65E+05	-13,17%	1,89E+05	1,33E+05	-29,64%	2,32E+04
10	4,12E+05	4,12E+05	0,00%	3,72E+05	-9,62%	3,76E+05	0,00%	3,76E+05	2,09E+05	-44,41%	3,96E+04
15	5,43E+05	4,79E+05	-11,76%	4,42E+05	-18,67%	3,94E+05	-17,04%	4,75E+05	1,81E+05	-61,91%	1,01E+05
20	7,13E+05	6,37E+05	-10,74%	5,73E+05	-19,62%	3,26E+05	-47,68%	6,24E+05	1,57E+05	-74,82%	1,40E+05
25	8,64E+05	7,53E+05	-12,79%	7,98E+05	-7,64%	4,48E+05	-39,72%	7,43E+05	3,05E+05	-58,90%	6,60E+04

Table 9: *Expected volume covered and chance constrained volume covered for the EFRLP, CCFRLP and DFRLP solution for the 80/40 instance*

p	Optimal objective value and opt. gap in terms of expected flow volume covered (%)					Optimal objective value and opt. gap in terms of chance constrained flow volume covered (%)					VSS
	EFRLP Solution	CCFRLP Solution		DFRLP Solution		EFRLP Solution		CCFRLP Solution	DFRLP Solution		
1	2,80E+04	1,31E+04	-53,24%	2,80E+04	0,00%	7,86E+03	-40,17%	1,31E+04	7,86E+03	-40,17%	1,00E-02
2	4,95E+04	4,45E+04	-9,97%	4,72E+04	-4,57%	2,49E+04	-45,05%	4,52E+04	7,86E+03	-82,63%	2,26E+03
3	8,38E+04	8,14E+04	-2,88%	8,38E+04	0,00%	8,14E+04	-1,35%	8,25E+04	3,28E+04	-60,19%	0,00E+00
4	1,38E+05	1,01E+05	-26,59%	1,29E+05	-6,49%	9,04E+04	-11,88%	1,03E+05	7,10E+04	-30,81%	8,93E+03
5	1,82E+05	1,45E+05	-20,15%	1,63E+05	-10,32%	1,10E+05	-24,32%	1,45E+05	4,78E+04	-67,16%	1,87E+04
10	3,85E+05	3,59E+05	-6,87%	3,41E+05	-11,45%	1,88E+05	-43,77%	3,35E+05	8,50E+04	-74,63%	4,41E+04
15	5,83E+05	5,47E+05	-6,22%	5,26E+05	-9,74%	3,89E+05	-22,94%	5,05E+05	1,66E+05	-67,09%	5,68E+04
20	7,35E+05	6,98E+05	-5,13%	6,60E+05	-10,27%	4,30E+05	-35,86%	6,70E+05	1,91E+05	-71,57%	7,55E+04
25	8,69E+05	8,24E+05	-5,17%	7,87E+05	-9,47%	5,78E+05	-29,00%	8,14E+05	2,39E+05	-70,66%	8,23E+04

B Programming Code

The following classes have been written in Java to perform the analyses as described in this paper:

- **Dijkstra.java**: Performs Dijkstra's algorithm to find the shortest path from one node in a graph to all others.
- **Edge.java**: Models an edge between two nodes in a graph.
- **Flow.java**: Models a flow instance as a list of locations on said flow and a certain volume.
- **Graph.java**: Models a graph with locations and edges.
- **Location.java**: Models a location as either an origin, destination or potential facility location.
- **Main.java**: Solves all models introduced in this paper for numerous problem instances.
- **ModelCCFRLP.java**: Models the Chance Constrained Flow Refueling Location Problem by means of CPLEX, according to the formulation of De Vries & Duijzer (2017).
- **ModelEFRLP.java**: Models the Expected Flow Refueling Location Problem by means of CPLEX, according to the formulation of De Vries & Duijzer (2017).
- **ModelFRLP.java**: Models the deterministic Flow Refueling Location Problem by means of CPLEX, according to the formulation of De Vries & Duijzer (2017).
- **ModelFRLP2.java**: Models the deterministic Flow Refueling Location Problem by means of CPLEX, according to the formulation of Boujelben & Gicquel (2019).
- **ModelUCFRLP2.java**: Models the Unit-Capacitated Flow Refueling Location Problem by means of CPLEX.