

Robust Alternative Fuel Refueling Station Location Problem with Routing under Decision-Dependent Flow Uncertainty

Özlem Mahmutoğulları*, Hande Yaman

ORSTAT, Faculty of Economics and Business, KU Leuven, 3000 Leuven, Belgium

Abstract

The refueling station location problem with routing (RSLP-R) is defined as a maximal coverage problem that locates alternative fuel refueling stations (AFSs) on a road network to maximize the refueled alternative fuel vehicle flows by considering the limited range of vehicles and the willingness of drivers to deviate from their paths for refueling. In this study, we introduce the robust counterpart of RSLP-R using a decision-dependent polyhedral uncertainty set. We model the flow uncertainty set using a hybrid model that comprises a hose model and individual flow bounds. To take into account the fact that vehicle flows are affected by AFS deployment decisions in their neighborhoods, we incorporate the decision-dependency notion into the flow uncertainty set. We propose two linear mixed integer programming formulations and a Benders reformulation. Our computational experiments on instances based on the road network of Belgium confirm the effectiveness of the reformulation in solving larger instances. We also report the results of experiments to assess the value of incorporating uncertainty and decision-dependency into the problem.

Keywords: Location, Robust optimization, Decision-dependent uncertainty, Benders reformulation, Alternative fuel vehicles

1 Introduction

Transportation is heavily dependent on fossil fuels, especially petroleum-based products. This strong dependency has two main drawbacks. Firstly, consuming fossil fuels results in greenhouse gas emissions that cause environmental problems such as air pollution, climate change, and global warming. The transport sector accounts for a significant portion of total worldwide greenhouse gas emissions. The amount of emissions from transportation constituted 27% of total Europe greenhouse gas emissions in 2017 and this amount increased in 2018 and 2019 (European Environment Agency (2019)). Secondly, fossil fuels have limited reserves that may be depleted in the near future. By 2050, only 14% of oil and 18% of gas reserves will remain if current fossil fuel energy consumption rates are maintained (Martins et al. (2019)). Using alternative fuels is one of the solutions to deal with the problems caused by fossil fuel consumption. In recent years, there has been a substantial increase in the promotion of vehicles that need alternative fuels to break the transportation sector's reliance on fossil fuels. These cars commonly use biodiesel, electricity, ethanol, hydrogen, (renewable) natural gas, and propane as

*Corresponding Author

e-mail: ozlem.mahmutoullar@kuleuven.be, hande.yaman@kuleuven.be

energy sources (Alternative Fuels Data Center (2021)). According to European Alternative Fuels Observatory (EAFO) (2020) data, while the overall number of alternative fuel vehicles (AFVs) in use in 2018 is 2.76 percent of total passenger cars in use, this percentage is 5.92 in 2020 for EU countries, the UK, EFTA countries, and Turkey. These numbers show that the total number of AFVs on the roads has more than doubled in two years. The AFVs considered in this survey are commonly used vehicles like battery electric (BEV), compressed natural gas (CNG), hydrogen (H₂), liquefied natural gas (LNG), liquefied petroleum gas (LPG), and plug-in hybrid electric (PHEV) vehicles.

The lack of alternative fuel station (AFS) infrastructure and the rather limited range of AFVs are two significant obstacles that are slowing down the introduction of AFVs and, as a result, the wide adoption of these vehicles by drivers (Belgium National Policy Framework (2016)). In this regard, the efficient deployment of the AFS infrastructure has recently started to be studied in the literature. The refueling station location problem (RSLP) for AFVs was first introduced and a maximal covering location model was presented by Kuby & Lim (2005). The problem aims to maximize the total amount of AFV flows that can be refueled by locating a predetermined number of AFSs on a network. The vehicle flows are defined on the shortest paths between their origins and destinations, and refueling a flow may require more than one station because of the limited range of the vehicles. Kim & Kuby (2012) extended the problem and considered the willingness of drivers to deviate from their shortest paths to refuel their vehicles. The drivers can tolerate traversing other paths, rather than their shortest paths, between their origin and destination (O-D) pairs if the total driving distance is within a certain bound, namely tolerance. Yıldız et al. (2016) generalized the problem by including non-simple path deviations as well, i.e., drivers are allowed to make cycles to refuel their vehicles. Therefore, a flow can be refueled if a combination of AFSs on at least one of the paths, between its origin and destination, satisfies the range and tolerance limits. Since the locations of stations and the routes of drivers are determined simultaneously in this problem, the authors referred to this problem as the refueling station location problem with routing (RSLP-R).

The majority of the RSLP studies in the literature assume that the decision maker has full knowledge of AFV flows before locating AFSs. However, it is likely to observe uncertainties in the flows because the rollout of AFVs and the development of the AFS network are still at their initial stages. In this respect, employing deterministic flows may result in an inefficient deployment of the AFS infrastructure. Furthermore, because fostering the use of AFVs is in its early stages, there is not enough historical data to predict a probability distribution for the flows. Consequently, robust optimization can be an advantageous tool in designing AFS networks.

The relationship between the lack of AFS network design and the low level adoption of AFVs is regarded as a “chicken-egg-problem” in the literature. The insufficient number of AFSs causes a poor incentive for drivers to use AFVs and vice versa. This interplay between the number of AFSs and the number of AFVs can be seen in the data provided by EAFO (2020) for Belgium. According to the Belgium data, by 2020, the numbers of different types of AFVs in use are 64 H₂ vehicles, 18,229 CNG vehicles, 33,703 BEVs, and 74,988 PHEVs, and there are 2 hydrogen, 168 natural gas, and 8,482 electricity stations (most of these electricity stations are convenient for short-distance intra-city trips). To analyze how facility location decisions affect the demands at customer sites, Jorge & Correia (2013) made a study for the carsharing business and Erlenkotter (1977) made a study to deploy warehouses. Both works ended up with the result that opening a facility increases customer demand in the neighborhood. In the aspects of these remarks, it is thus important to consider that the availability of AFSs in the neighborhood affects the proliferation of AFVs during the development

of infrastructure.

In this study, we incorporate robustness and decision-dependency into the RSLP-R. The contributions of the study are as follows:

- We introduce the robust RSLP-R under decision-dependent polyhedral vehicle flow uncertainty. To the best of our knowledge, this is the first study that considers the effects of location decisions on the uncertainty sets of the flow-based demands.
- We derive mathematical programming formulations and propose a Benders reformulation and a branch-and-cut algorithm for the Benders reformulation.
- We generate data sets based on the road network and population distribution of Belgium for the computational experiments.
- We perform the following computational experiments: We first compare the performances of the proposed mathematical models and the Benders reformulation. Then, we investigate the changes in station locations and total covered flows when the optimal solutions of the deterministic, robust and decision-dependent robust problems are employed. We also analyze the changes under different parameter settings. We observe that recognizing the uncertainty in flows and the decision-dependency of uncertain flow realizations can lead to significant gains in the total AFV flows covered.

The rest of the study is organized as follows. In Section 2, we review the related studies in the literature. In Section 3, we define the deterministic version of the RSLP-R, introduce the decision-dependent robust counterpart of the problem, and present formulations. We also derive a Benders reformulation and describe a branch-and-cut algorithm to solve it. We report our computational results in Section 4. We discuss possible extensions and conclude our study in Section 5.

2 Literature Review

In the context of location-allocation problems, Hakimi (1965) first introduces p -median and p -center problems by extending the absolute median and absolute center problems described by Hakimi (1964). These problems are based on locating p facilities on a network to minimize the sum of demand weighted distances from customer nodes to facilities (p -median) and the maximum of total demand weighted distances between a customer and facilities (p -center). Toregas et al. (1971) deal with a facility location problem where each facility has a maximum service distance. They propose a set-covering model to minimize the total number of facilities required to meet the total demand. Church & ReVelle (1974) take into account the resource constraint as well as the maximum service distance. The resource constraint may arise where there is an insufficient number of facilities to be located because of the high capital cost of building a facility. They propose a maximal covering model to locate p facilities maximizing the total demand covered. As a result, there are studies on the set-covering and maximal covering versions of the RSLP. Although the major part of the RSLP literature contains the maximal covering models, there are a few papers (Wang & Lin (2009); Wang & Wang (2010)) that consider the set-covering notion.

Bapna et al. (2002) present the first RSLP formulation based on maximal covering and p -center problems to locate AFSs. Nicholas et al. (2004) derive a p -median model that minimizes the total average driving time from demand nodes to the stations. In both works, demands occur at the nodes

of the road network and the population of the nodes determines the magnitude of their demands. Unlike node-based studies, Kuby & Lim (2005) make a flow-based (path-based) analysis and assume the demands as traffic flows on the shortest paths of their O-D pairs. They extend the flow-capturing location model (FCLM) proposed by Hodgson (1990) and develop the flow-refueling location model (FRLM), which is used interchangeably with RSLP in the literature, for AFS infrastructure. In FCLM, a flow is regarded as covered if it passes at least one facility along its path. In FRLM, due to the limited range of vehicles, one station on their paths does not always guarantee that the flows are refueled. Kuby & Lim (2005) consider a traffic flow as refueled if a vehicle departing from its origin can reach its destination and return to the origin without running out of fuel. The authors propose a two-stage solution method to locate a given number of AFSs and develop a maximal covering model that requires identifying all combinations of candidate AFS locations capable of serving each vehicle flow to make a round trip on their paths beforehand. Upchurch & Kuby (2010) compare flow-based and node-based formulations which are the FRLM and p -median model, respectively. They investigate how well these formulations perform in each other's objectives and conclude that the performance of the stations located by the FRLM in the objective of the p -median model is better than the performance of the stations located by the p -median in the FRLM's objective.

The FRLM requires a pre-generation of all candidate station location combinations for refueling the flow of each path before using a mixed integer programming model. Generating an enormous number of feasible combinations is impractical and can be computationally burdensome for large-sized networks. In this regard, Lim & Kuby (2010) propose heuristic methods. They adapt greedy and genetic algorithms to solve the FRLM. Capar & Kuby (2012) reformulate the FRLM in such a way that the pre-generation stage is no longer required. They used the logic that the flow is refuelable if each node on the path is reachable without running out of fuel. Capar et al. (2013) present a new reformulation based on covering all arcs, rather than nodes, that comprise each path. Another reformulation is proposed by MirHassani & Ebrazi (2013). They use the idea of network transformation, i.e., a given path is refuelable if all its intermediate nodes have AFSs. Tran et al. (2018) employ the model proposed by Capar et al. (2013) and propose a heuristic algorithm that outperforms the heuristic algorithms developed by Lim & Kuby (2010). The algorithm is based on exploiting promising candidate station locations during the solution process, considering that the optimal station locations obtained from the LP relaxation of the model can perform well for the original problem.

In the FRLM, drivers are assumed to use a fixed path, generally the shortest path, between their O-D pairs and the AFSs are located along this path to refuel the vehicle flow. However, in real-world situations, drivers are likely to deviate from their shortest paths to refuel their vehicles if the provision of refueling stations is sparse. Kim & Kuby (2012) accommodate the willingness of drivers to deviate from their shortest paths into the FRLM and introduce a new version of the problem, which is the deviation-FRLM. The authors develop a mixed integer linear programming model that takes all paths and eligible combinations of refueling stations that can make a path between O-D pairs refuelable as inputs and assigns the flows to the station combinations. For this model, computational efficiency is a great concern because considering several paths instead of a fixed path for an O-D pair makes the problem more complex. In this respect, Kim & Kuby (2013) propose a heuristic algorithm. Yıldız et al. (2016) extend the problem where drivers can make cycles not only at either end of their paths, as stated in Kim & Kuby (2013), but also within their paths to refuel their vehicles. The authors refer to the problem as the RSLP-R and devise a mathematical model and a branch-and-price algorithm

where the explicit pre-generation of each route between O-D pairs is not required. Finally, Arslan et al. (2019) propose a branch-and-cut algorithm for a natural formulation of the RSLP-R, and improve the solution times and the solvable problem sizes substantially.

Most studies assume that the capacity of AFSs is not a concern because of the small amount of flows on the roads at the adoption stage of AFVs. There are a few studies that consider a restriction on the amount of flows that can be refueled at each station. Upchurch et al. (2009) present the first work that incorporates capacity constraints into the FRLM. Hosseini & MirHassani (2017) also study the capacitated FRLM and propose a mathematical model by using the concept of network transformation described in MirHassani & Ebrazi (2013). Based on the formulation proposed by Capar et al. (2013) for the problem with uncapacitated stations, Hosseini et al. (2017) introduce the capacitated problem where the drivers may deviate from their pre-determined paths while Zhang et al. (2017) address the problem of locating capacitated stations considering a multi-period planning horizon. Zhang et al. (2017) also draw attention to the fact that the state of the AFS infrastructure, i.e., the percentage flow coverage of each path, in the previous period will affect the AFV flows in the future periods. Although, like Hosseini et al. (2017) and Zhang et al. (2017), Rose et al. (2020) build their proposed model upon the model in Capar et al. (2013), they define the objective with respect to the set-covering notion. The authors assume that there is a limit on the number of stations located at each node as well as a limit on the capacity of stations.

A number of extensions of the conventional deterministic FRLM, apart from the studies mentioned before, have recently been studied in the literature. Kweon et al. (2017) use the idea, in Kuby & Lim (2007), of employing additional candidate AFS sites along the arcs rather than regarding only the vertices and apply it to the case in which there are multiple paths for the O-D pairs. Kang & Recker (2015) incorporate the drivers' out-of-home activities scheduling and routing decisions into AFS location decisions. The authors propose a location-routing model based on the set-covering problem for the AFS location decisions and the household activity pattern problem for the routing decisions. Hwang et al. (2015) investigate the AFS location decisions on roads divided into opposite directions by barriers, where single-access stations can only serve vehicles on one side of the road and dual-access stations can serve vehicles on both sides. The problem in Hwang et al. (2015) is extended by Hwang et al. (2017) and Hwang et al. (2020) to the cases with different vehicle ranges and driver tolerances, respectively. Ventura et al. (2017) also base their proposed model upon the work of Hwang et al. (2015) but consider two conflicting objectives, which are minimizing the cost of constructing stations and maximizing the total trip distance that can be traveled. Honma & Kuby (2019) aim at minimizing the total deviations that drivers must make to refuel their vehicles where there is a single station that can refuel each path-based demand.

The uncertainty in AFV flow volumes is often incorporated into the decision-making process for designing an AFS infrastructure using robust or stochastic optimization. Mak et al. (2013) present a study that focuses on the demand uncertainty in the RSLP. They propose a distributionally robust optimization model for the problem that aims to minimize the total cost of locating electrical battery swapping stations. Hosseini & MirHassani (2015) also incorporate uncertainty into vehicle flows. They propose a stochastic optimization model based on the formulation introduced by MirHassani & Ebrazi (2013). Another work that handles vehicle flow uncertainty is by Miralinaghi et al. (2017). The authors consider multiple time periods in which AFV flows fluctuate and present a robust optimization model by using an uncertainty budget notion (Bertsimas & Sim (2004)). Most stochastic optimization works propose two-stage stochastic programming formulations to solve the problem. Hosseini & MirHassani

(2015) determine the locations of permanent AFSs in the first stage and portable AFSs after the demands are realized. Other works focus specifically on EVs and thus aim at locating the recharging stations. Faridimehr et al. (2018) consider the demand uncertainty depending on factors such as time, driver behaviors, and remaining battery charges. The authors propose a model to locate semi-rapid chargers for intra-city trips where drivers can only charge their cars once. MirHassani et al. (2020) extend this work to the case for which there are two types of chargers and the drivers are allowed to charge their vehicles two times. Wu & Sioshansi (2017) address the problem of locating recharging stations for inter-city trips while assuming that the battery can only be recharged once per trip, considering both the vehicle range and the driver tolerances. Instead of defining multiple deviation paths for each trip, the authors employ a capturing circle for each station, i.e., the stations are eligible to charge the vehicles within a pre-determined radius, to include the tolerances without making the problem more complex. Yıldız et al. (2019) take into account the multiple deviation paths and the limited capacity of the recharging stations. A multi-stage stochastic programming model is proposed by Kadri et al. (2020) to locate the charging stations over a multi-period horizon. Some studies regard the uncertainty in other parameters rather than in vehicle flows. For example, Xie et al. (2018) incorporate the uncertainty into the waiting times of the drivers to find an available charging station, Riemann et al. (2015) deal with the uncertainty in drivers' route selection choices, and de Vries & Duijzer (2017) and Kchaou-Boujelben & Gicquel (2020) consider the uncertainty in vehicle ranges. Hosseini et al. (2021) simultaneously take into account the uncertainties in vehicle ranges and flow volumes.

Depending on whether the decisions to be made play a role in the uncertain information resolution, the uncertainty can be classified as exogenous or endogenous (Jonsbråten (1998)). If the realization of uncertain parameters is independent of decisions, the uncertainty is exogenous to the decision-making process. The parameters in a problem with endogenous uncertainty, on the other hand, are affected by the decisions. This effect, according to Goel & Grossmann (2006) can be observed in two ways: the decisions can change the time when the uncertainty is revealed or favor the possibility of some parameter realizations. In the former type of endogenous uncertainty framework, i.e., the uncertain information is sequentially discovered by the decisions, there are studies on gas field investment and operational planning with uncertain gas reserves (Goel & Grossmann (2004)), resource allocation for projects with uncertain returns (Solak et al. (2010)), vehicle routing with uncertain customer demands (Hooshmand Khaligh & MirHassani (2016)), and operating room scheduling with uncertain operation durations (Hooshmand et al. (2018)). Novel formulations and solution methods for generic planning problems are proposed by Vayanos et al. (2011), Goel & Grossmann (2005), Tarhan et al. (2013) and Basciftci et al. (2019). The problems with the latter type of endogenous uncertainty, i.e., the decisions taken can alter the known probability distribution or uncertainty set of parameters, are investigated using stochastic and robust optimization approaches. Stochastic programming is applied by Shao et al. (2006) to user equilibrium traffic assignment in which travel times depend on the route choice decisions and by Basciftci et al. (2020) to power system management in which generator failures depend on their degradation levels caused by the dispatch amount and maintenance scheduling decisions. Employing robust programming for the decision-dependent uncertainty in which the decisions account for the uncertainty sets, which is also the focus of this paper, has recently received attention in the literature. The works within this framework address uncertainties in a software-partitioning problem (Spacey et al. (2012)), a knapsack problem (Poss (2013)), a scheduling problem (Vujanic et al. (2016)), and a newsvendor problem (Hu et al. (2019)). Several studies (Royset & Wets (2017); Noyan et al. (2018))

represent the uncertainty set as a set of probability distributions where the partial information about the distribution of parameters is known, i.e., distributionally robust optimization. In the context of literature on facility location problems, in terms of node-based demands, Basciftci et al. (2021) and Luo (2020) study the impact of facility decisions on demand uncertainty and model decision-dependent distributionally robust formulations. Basciftci et al. (2021) integrate the decision-dependency into the demand uncertainty set, formulate the problem using the decision-dependent uncertainty set and perform polyhedral analysis on the formulation. Their results underline the computational efficiency of their proposed approach and highlight the advantage of considering the impact of location decisions on the demand uncertainty to obtain a well-planned network. Unlike Basciftci et al. (2021), Luo (2020) decouples the decision-dependency of demand from the demand uncertainty set and asserts that the decoupling makes the formulation more data-driven and computationally efficient.

3 Problem Definition and Mathematical Formulations

3.1 The deterministic setting

In this section, we first define the deterministic problem and then present the model proposed by Arslan et al. (2019) for the sake of completeness.

The RSLP-R is defined on a directed network $G = (N, A)$ with node set $N = \{1, \dots, n\}$ and arc set $A \subseteq \{(i, j) : i, j \in N, i \neq j\}$. Set Q represents the set of demands. An AFV demand $q \in Q$ is defined as a five-tuple $\langle o_q, d_q, f_q, r_q, w_q \rangle$ where o_q is the origin node, d_q is the destination node, f_q is the vehicle flow volume, r_q is the range of vehicle, and w_q is the total deviation tolerated by drivers. The deviation can be defined as a fixed length or as a percentage of the shortest path length. We represent the driver tolerances in percentages. We use the round trip notion proposed by Kuby & Lim (2005) to consider a vehicle as refueled. In this regard, it is assumed that an AFV departs from its origin with a half-full tank and arrives at its destination with at least a half-full tank if there is no AFS at these nodes. If the origin node has an AFS, the AFV departs from the node with a full tank. If there is an AFS at the destination node, the AFV can arrive at the node with an empty tank.

Consider flow q where $o_q = 1$ and $d_q = 9$ on the example road network illustrated in Figure 1. There are 5 paths from node 1 to node 9: 1-2-7-9 with a length of 20, 1-2-6-9 with a length of 14, 1-2-5-9 with a length of 17, 1-3-4-5-9 with a length of 20, and 1-8-5-9 with a length of 15. The shortest path length is 14. If the driver is willing to drive at most a length of 21, namely, 50% driver tolerance, and the range of the vehicle is 40, a single facility at the origin or destination node makes each path feasible to complete the round trip. Alternatively, a station at any node on one of these paths makes the corresponding path feasible. If the range is 12, one station is not enough, so at least two stations are required. In this case, path 1-2-7-9 is feasible if and only if two stations are located at nodes 2 and 7. If the range of the vehicle is 8, path 1-2-7-9 is no longer feasible regardless of the tolerance. Let the total distance tolerated by the driver is equal to the shortest path length, in other words, the driver is not willing to deviate from the shortest path to refuel the vehicle. In this case, if the range is 5, the flow cannot be covered since the only path that satisfies the range limitation is path 1-3-4-5-9, which does not satisfy the tolerance.

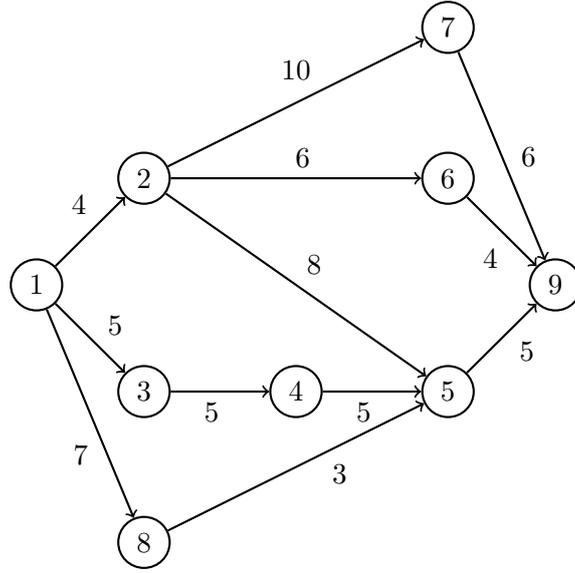


Figure 1: Example road network

While locating AFSs, we need to consider range and tolerance limitations to refuel flows. As seen in the previous example, there may be nodes in the road network that cannot satisfy these limitations for some flows. As a result, for these flows, network G can be reduced. This network transformation notion is first proposed by MirHassani & Ebrazi (2013) where the stations can be located only on a fixed path between an O-D pair. Arslan et al. (2019) adapt this notion to the case where each path between an O-D pair can be selected to locate stations. We use the network transformation presented by Arslan et al. (2019). Let δ_{ij} be the shortest path distance between nodes $i \in N$ and $j \in N$. Graph $G_q = (N_q, A_q)$ is called the transformed graph of flow q with $N_q = \{s_q, t_q\} \cup \{i \in N : \delta_{o_q i} + \delta_{i d_q} \leq (1 + w_q)\delta_{o_q d_q}\}$ where s_q and t_q are dummy nodes and $A_q = A_q^1 \cup A_q^2 \cup A_q^3$ where

$$\begin{aligned} A_q^1 &= \{(s_q, i) : \delta_{o_q i} \leq r_q/2, i \in N_q \setminus \{s_q, t_q\}\}, \\ A_q^2 &= \{(i, t_q) : \delta_{i d_q} \leq r_q/2, i \in N_q \setminus \{s_q, t_q\}\}, \\ A_q^3 &= \{(i, j) : \delta_{ij} \leq r_q, i, j \in N_q \setminus \{s_q, t_q\}, i \neq j\}. \end{aligned}$$

Arcs $(s_q, i) \in A_q^1$, $(i, t_q) \in A_q^2$ and $(i, j) \in A_q^3$ have lengths $\delta_{o_q i}$, $\delta_{i d_q}$ and δ_{ij} , respectively. It is assumed that an AFV departs from s_q with a half-full tank. Traversing arc (s_q, o_q) represents that the AFV is refueled at the origin node. The AFV should arrive at t_q with at least a half-full tank, and traversing arc (d_q, t_q) represents that the AFV is refueled at the destination node. Each node $i \in N_q \setminus \{s_q, t_q\}$ on a path from s_q and t_q is called an internal node, and traversing arc (i, j) represents that the AFV is refueled at node $j \in N_q \setminus \{s_q, t_q\}$. Note that, if all internal nodes of a path from s_q to t_q have an AFS, the round trip is guaranteed. Hence, a path on graph G_q is called feasible for demand q if it has an AFS at each of its internal nodes and has a length of at most $(1 + w_q)\delta_{o_q d_q}$.

Let P_q be the set of all feasible paths for a demand q . Arslan et al. (2019) define a q -node-cut to be subset $S \subseteq N_q \setminus \{s_q, t_q\}$ if removing S from the node set N_q disconnects every path in P_q . Let Γ_q be the set of all q -node-cuts. If no proper subset of set $S \in \Gamma_q$ is a q -node-cut, then set S is called a minimal q -node-cut for P_q .

As an example, the transformed graph of demand q with $o_q = 1$ and $d_q = 9$ is illustrated in Figure 2 where $r_q = 8$ and $w_q = 10\%$. For example, $\{2, 8\}$ and $\{6, 8\}$ are the minimal q -node-cuts for

this demand. Note that there remain paths after removing nodes 6 and 8, e.g., s-1-2-5-9-t, s-2-5-9-t. However, these remaining paths are not feasible and thus not in P_q because they have a length of 17, which is greater than total tolerated distance 15.4.

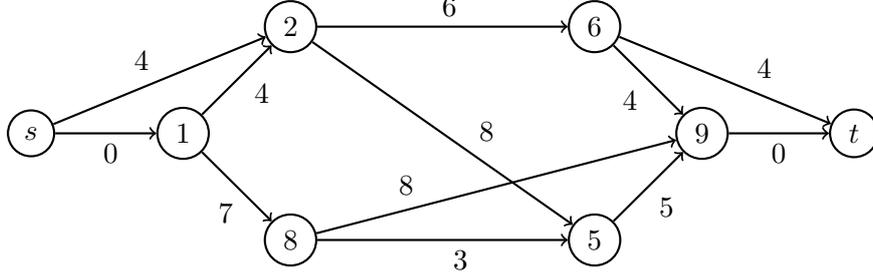


Figure 2: Transformed graph for demand q with $o_q = 1$ and $d_q = 9$ where $r_q = 8$ and $w_q = 10\%$

Let the location and cover variables be

$$x_i = \begin{cases} 1, & \text{if there is an AFS at node } i \in N \\ 0, & \text{otherwise;} \end{cases}$$

$$y_q = \begin{cases} 1, & \text{if a feasible path is constructed for } q \in Q \\ 0, & \text{otherwise.} \end{cases}$$

Arslan et al. (2019) formulate the deterministic version of the problem as follows:

$$\max \sum_{q \in Q} f_q y_q \quad (1a)$$

$$\text{s.t. } \sum_{i \in N} x_i \leq p \quad (1b)$$

$$y_q \leq \sum_{i \in S} x_i \quad \forall q \in Q, S \in \Gamma_q \quad (1c)$$

$$x_i \in \{0, 1\} \quad \forall i \in N \quad (1d)$$

$$y_q \in \{0, 1\} \quad \forall q \in Q. \quad (1e)$$

The objective function (1a) maximizes the total AFV flows to be covered. Constraint (1b) allows to locate at most p AFSs. To have a feasible path for demand q , there must be at least one AFS at one of the sites in each q -node-cut $S \in \Gamma_q$. If there is a feasible path, then demand q can be refueled. On the other hand, if there exists a q -node-cut $S \in \Gamma_q$ where no AFS has been located at any sites in S , i.e., $\sum_{i \in S} x_i = 0$, then G_q is not connected and there is no feasible path for demand q . Constraints (1c) ensure that, in this case, demand q cannot be refueled. Note that these constraints are exponential in number. To separate these constraints, we use the algorithm proposed by Arslan et al. (2019). We briefly explain this algorithm for the sake of completeness. For given $x = \hat{x}$ and $y = \hat{y}$, a support graph $G_q(\hat{N}_q)$ where $\hat{N}_q = \{s_q, t_q\} \cup \{i \in N_q : \hat{x}_i = 1\}$ is constructed for each q with $\hat{y}_q = 1$. If the corresponding graph is not connected or the length of the shortest path between nodes s_q and t_q is greater than $(1+w_q)\delta_{o_q d_q}$, then the inequality is violated for the q -node-cut $N_q \setminus \hat{N}_q$. As this q -node-cut may not be a minimal q -node-cut for path set P_q , if graph $G_q(\hat{N}_q \cup \{i\})$ for a node $i \in N_q \setminus \hat{N}_q$ is not

connected or the length of the shortest path between s_q and t_q in graph $G_q(\hat{N}_q \cup \{i\})$ is greater than $(1 + w_q)\delta_{o_q d_q}$, then node i is removed from the q -node cut. This process is repeated until a minimal q -node-cut is obtained.

3.2 The decision-dependent uncertainty set and the robust counterpart

We introduce our uncertainty set using the hybrid model, which is proposed by Altın et al. (2011) for the robust network loading problem and is also used by Meraklı & Yaman (2016) in the context of the robust uncapacitated hub location problem. The hybrid model comprises a hose model and an interval model. Duffield et al. (1999) and Fingerhut et al. (1997) introduce the hose model to model demand uncertainty in telecommunication networks. The model limits the total traffic associated with each node instead of estimating individual pairwise demands. The hose model has resource sharing flexibility, but it may lead to conservative location decisions because of considering unlikely worst-case flow realizations (Altın et al. (2011)), e.g., setting only one of the flows emanating from a node to the aggregate bound and the other flows to zero. Restricting the hose model with bounds of individual O-D pair flows alleviates the conservativeness of the hose model.

We define the hybrid uncertainty set of the vehicle flows under the impact of station location decisions as

$$\mathcal{F}(x) = \left\{ f \in \mathbb{R}_+^{|Q|} : \sum_{q \in Q: i=o_q \vee i=d_q} f_q \geq b_i(x) \quad \forall i \in N, \quad u_q(x) \geq f_q \geq l_q(x) \quad \forall q \in Q \right\} \quad (2)$$

where $b_i(x)$ is an aggregate bound for the incoming and outgoing flows of node i , $l_q(x)$ is a lower bound and $u_q(x)$ is an upper bound for flow q under decision x . The set $\mathcal{F}(x)$ contains all flows such that the total flow of the demands whose origin or destination is node $i \in N$ is at least as large as the aggregate bound $b_i(x)$ and the flow of each demand $q \in Q$ is within its upper and lower bounds, $u_q(x)$ and $l_q(x)$. Restricting the individual flow constraints $u_q(x) \geq f_q \geq l_q(x)$ for all $q \in Q$ with the aggregate flow constraints $\sum_{q \in Q: i=o_q \vee i=d_q} f_q \geq b_i(x)$ for all $i \in N$ helps to eliminate conservative solutions, i.e., if the uncertainty set is defined by only the box constraints, each individual flow is set to its lower bound. We assume that $\mathcal{F}(x)$ is nonempty for all feasible x . This assumption may seem restrictive, however, if there is no feasible flow for a given choice of AFS locations, then either the aggregate bounds $b_i(x)$'s and the upper bounds $u_q(x)$'s or the upper and lower bounds for some demands are not consistent and should be updated.

The robust RSLP-R under decision-dependent hybrid uncertainty aims at maximizing the total AFV flows refueled under the worst-case flow realization scenario. The problem can be formulated as

$$\max_{(x,y) \in \mathcal{D}} \min_{f \in \mathcal{F}(x)} \sum_{q \in Q} f_q y_q \quad (3)$$

where set \mathcal{D} is defined by the constraints of the deterministic problem. Note that formulation (3) is a max-min formulation. To obtain a monolithic formulation, since the inner problem is linear, feasible and bounded, we can use the dual transformation (see, e.g., Bertsimas & Sim (2003); Meraklı & Yaman (2016); Altın et al. (2011)). For given $(\hat{x}, \hat{y}) \in \mathcal{D}$, the inner problem is

$$\min \sum_{q \in Q} f_q \hat{y}_q$$

$$\begin{aligned}
\text{s.t. } \quad & \sum_{q \in Q: i=o_q \vee i=d_q} f_q \geq b_i(\hat{x}) \quad \forall i \in N & (\beta_i) \\
& l_q(\hat{x}) \leq f_q \quad \forall q \in Q & (\theta_q) \\
& f_q \leq u_q(\hat{x}) \quad \forall q \in Q & (\eta_q) \\
& f_q \geq 0 \quad \forall q \in Q.
\end{aligned}$$

Let β , θ and η be the dual variable vectors associated with the above constraints. Then, the dual of inner problem is

$$\begin{aligned}
\max \quad & \sum_{i \in N} \beta_i b_i(\hat{x}) + \sum_{q \in Q} (\theta_q l_q(\hat{x}) - \eta_q u_q(\hat{x})) \\
\text{s.t. } \quad & \beta_{o_q} + \beta_{d_q} + \theta_q - \eta_q \leq \hat{y}_q \quad \forall q \in Q \\
& \beta_i \geq 0 \quad \forall i \in N \\
& \theta_q, \eta_q \geq 0 \quad \forall q \in Q.
\end{aligned}$$

By strong duality, both problems have equal optimal values. Hence, problem (3) can be reformulated as follows:

$$\begin{aligned}
\max \quad & \sum_{i \in N} \beta_i b_i(x) + \sum_{q \in Q} (\theta_q l_q(x) - \eta_q u_q(x)) & (4a) \\
\text{s.t. } \quad & \beta_{o_q} + \beta_{d_q} + \theta_q - \eta_q \leq y_q \quad \forall q \in Q & (4b) \\
& (x, y) \in \mathcal{D} & (4c) \\
& \beta_i \geq 0 \quad \forall i \in N & (4d) \\
& \theta_q, \eta_q \geq 0 \quad \forall q \in Q. & (4e)
\end{aligned}$$

We suppose that, when a new station is opened, vehicle flows in the neighborhood increase because the drivers will be more willing to use AFVs if there are AFSs nearby. The impact of new stations can be incorporated into the flow parameters in different ways. Basciftci et al. (2021) and Yu & Shen (2020) study the decision-dependent distributionally robust problems. In both works, the authors consider the interplay between moment (mean and variance) information of distributions and location variables. Basciftci et al. (2021) interpret this relation using piecewise linear functions, whereas Yu & Shen (2020) assume that the parameters are affinely dependent on the decisions. We investigate the case where b_i , l_q and u_q are affine functions of x :

$$\begin{aligned}
b_i(x) &= b_{i0} + \sum_{j \in N} b_{ij} x_j \quad i \in N, \\
l_q(x) &= l_{q0} + \sum_{j \in N} l_{qj} x_j \quad q \in Q, \\
u_q(x) &= u_{q0} + \sum_{j \in N} u_{qj} x_j \quad q \in Q.
\end{aligned}$$

Then the objective function (4a) becomes:

$$\max \sum_{i \in N} \beta_i \left(b_{i0} + \sum_{j \in N} b_{ij} x_j \right) + \sum_{q \in Q} \left(\theta_q \left(l_{q0} + \sum_{j \in N} l_{qj} x_j \right) - \eta_q \left(u_{q0} + \sum_{j \in N} u_{qj} x_j \right) \right) \quad (5)$$

which can be rewritten as

$$\max \sum_{i \in N} b_{i0} \beta_i + \sum_{q \in Q} (l_{q0} \theta_q - u_{q0} \eta_q) + \sum_{j \in N} \left(\sum_{i \in N} b_{ij} \beta_i + \sum_{q \in Q} (l_{qj} \theta_q - u_{qj} \eta_q) \right) x_j. \quad (6)$$

Note that there are nonlinear terms in this objective function. To linearize the bilinear terms, we use McCormick envelopes (McCormick (1976)). If there is a bilinear term $\nu = vz$ where z is a binary variable, and v^u and v^l are the upper and lower bounds of variable v , McCormick inequalities

$$v - (1 - z)v^u \leq \nu, \quad (7a)$$

$$\nu \leq v - (1 - z)v^l, \quad (7b)$$

$$v^l z \leq \nu, \quad (7c)$$

$$\nu \leq v^u z \quad (7d)$$

can be used to represent ν . Linear formulation (7) guarantees that $\nu = v$ if $z = 1$ and $\nu = 0$ otherwise. We linearize the bilinear terms in objective function (6) in two different ways leading to an aggregated and a disaggregated formulation.

3.3 An aggregated formulation

First, we define a new decision variable $\phi_j = \left(\sum_{i \in N} b_{ij} \beta_i + \sum_{q \in Q} (l_{qj} \theta_q - u_{qj} \eta_q) \right) x_j$ for each $j \in N$. Since ϕ_j , $j \in N$ is an unrestricted variable (see Example 1), upper and lower bounds are taken as M_j^l and M_j^u for the McCormick inequalities where M_j^l and M_j^u are arbitrarily large numbers. The mixed integer linear programming model of the problem can be given as

(Decision-dependent Robust RSLP-R-Aggregated Model)

$$\max \sum_{i \in N} b_{i0} \beta_i + \sum_{q \in Q} (l_{q0} \theta_q - u_{q0} \eta_q) + \sum_{j \in N} \phi_j \quad (8a)$$

$$\text{s.t. } \beta_{o_q} + \beta_{d_q} + \theta_q - \eta_q \leq y_q \quad \forall q \in Q \quad (8b)$$

$$\phi_j \geq \sum_{i \in N} b_{ij} \beta_i + \sum_{q \in Q} (l_{qj} \theta_q - u_{qj} \eta_q) - (1 - x_j) M_j^u \quad \forall j \in N \quad (8c)$$

$$\phi_j \leq \sum_{i \in N} b_{ij} \beta_i + \sum_{q \in Q} (l_{qj} \theta_q - u_{qj} \eta_q) - (1 - x_j) M_j^l \quad \forall j \in N \quad (8d)$$

$$\phi_j \geq M_j^l x_j \quad \forall j \in N \quad (8e)$$

$$\phi_j \leq M_j^u x_j \quad \forall j \in N \quad (8f)$$

$$(x, y) \in \mathcal{D} \quad (8g)$$

$$\beta_i \geq 0 \quad \forall i \in N \quad (8h)$$

$$\theta_q, \eta_q \geq 0 \quad \forall q \in Q \quad (8i)$$

$$\phi_j \text{ is unrestricted} \quad \forall j \in N. \quad (8j)$$

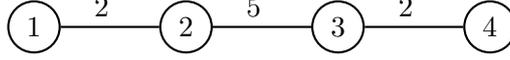


Figure 3: Example network

Example 1 Consider the network in Figure 3. The lengths of edges are as shown in the figure. The range of the vehicle is 4 and the drivers do not tolerate to deviate from their shortest paths. The O-D pairs are as follows: $(o_1, d_1) = (1, 2)$, $(o_2, d_2) = (1, 3)$, $(o_3, d_3) = (2, 3)$, $(o_4, d_4) = (2, 4)$ and $(o_5, d_5) = (3, 4)$. The aggregated flow parameters are: $b_{10} = b_{40} = 108$, $b_{20} = b_{30} = 126$ and $b_{ij} = 0$ for all $i, j \in N$, and the individual flow parameters are: $l_{10} = l_{50} = 10$, $l_{20} = l_{30} = l_{40} = 2$, $u_{10} = u_{50} = 90$, $u_{20} = u_{30} = u_{40} = 18$, $l_{21} = l_{22} = l_{23} = l_{42} = l_{43} = l_{44} = 5.4$, $u_{21} = u_{22} = u_{23} = u_{42} = u_{43} = u_{44} = 0.6$, $l_{24} = l_{41} = u_{24} = u_{41} = 0$ and $l_{qj} = u_{qj} = 0$ for all $j \in N$ and $q \in \{1, 3, 5\}$.

If model (8) is used to locate 2 AFSs for the given network, the optimal value is 178.8 with optimal solution $x_1 = x_4 = 1$, $y_1 = y_5 = 1$, $\beta_1 = \beta_4 = 1$, $\eta_2 = \eta_4 = 1$, $\phi_1 = \phi_4 = -0.6$ and other decision variables are 0. If we enforce all ϕ variables to be greater than or equal to 0, then the optimal value is 177.2. Hence, for this instance there is no optimal solution with nonnegative ϕ .

There are four constraints with big M values in model (8). The performance of the formulation can be improved by strengthening the big M values. Since ϕ is dependent on β, θ and η , stronger bounds for ϕ can be found if there exist upper bounds for variables β, θ and η .

Proposition 1 For an extreme point $\xi = (x, y, \beta, \theta, \eta, \phi)$ of the convex hull of the set defined by inequalities (8b)-(8j), the following hold:

$$\beta_i \leq 1 \quad \forall i \in N, \quad (9a)$$

$$\theta_q \leq 1 \quad \forall q \in Q, \quad (9b)$$

$$\eta_q \leq 2 \quad \forall q \in Q. \quad (9c)$$

Proof. We first prove that for each extreme point ξ of the convex hull, θ_q and η_q cannot be positive at the same time for all $q \in Q$. Assume the contrary that there exists an extreme point $\hat{\xi} = (\hat{x}, \hat{y}, \hat{\beta}, \hat{\theta}, \hat{\eta}, \hat{\phi})$ with $\hat{\theta}_{\tilde{q}} > 0$ and $\hat{\eta}_{\tilde{q}} > 0$ for a flow $\tilde{q} \in Q$. Let's define, where $\epsilon > 0$ is a very small number, a point ξ' with $\theta'_{\tilde{q}} = \hat{\theta}_{\tilde{q}} + \epsilon$, $\eta'_{\tilde{q}} = \hat{\eta}_{\tilde{q}} + \epsilon$ and $\phi'_j = \hat{\phi}_j + \epsilon(l_{\tilde{q}j} - u_{\tilde{q}j})$ for $j \in N$ with $\hat{x}_j = 1$, and a point ξ'' with $\theta''_{\tilde{q}} = \hat{\theta}_{\tilde{q}} - \epsilon$, $\eta''_{\tilde{q}} = \hat{\eta}_{\tilde{q}} - \epsilon$ and $\phi''_j = \hat{\phi}_j - \epsilon(l_{\tilde{q}j} - u_{\tilde{q}j})$ for $j \in N$ with $\hat{x}_j = 1$, and all other entries of ξ' and ξ'' are the same as the entries of $\hat{\xi}$. Note that points ξ' and ξ'' are also in the convex hull and $\hat{\xi}$ can be written as a convex combination of points ξ' and ξ'' , i.e., $\hat{\xi} = \frac{1}{2}\xi' + \frac{1}{2}\xi''$. Thus, $\hat{\xi}$ is not an extreme point of the convex hull. This contradicts with the assumption.

For $q \in Q$, if $\eta_q = 0$, then (9b) holds. If $\eta_q > 0$, then $\theta_q = 0$. So (9b) is satisfied.

Next, we show that for each extreme point ξ of the convex hull, β_i is less than or equal to 1 for all $i \in N$. To the contrary, assume that there exists an extreme point $\hat{\xi} = (\hat{x}, \hat{y}, \hat{\beta}, \hat{\theta}, \hat{\eta}, \hat{\phi})$ with $\hat{\beta}_{\tilde{i}} > 1$ for a node $\tilde{i} \in N$. Let \tilde{Q} be the set of flows such that \tilde{i} is their origin or destination. If $\hat{\beta}_{\tilde{i}} > 1$, then $\hat{\eta}_q > 0$ for all $q \in \tilde{Q}$. Consider the feasible points ξ' and ξ'' where $\beta'_i = \hat{\beta}_i + \epsilon$, $\eta'_q = \hat{\eta}_q + \epsilon$ for all $q \in \tilde{Q}$ and $\phi'_j = \hat{\phi}_j + \epsilon(b_{ij} - \sum_{q \in \tilde{Q}} u_{qj})$ for all $j \in N$ with $\hat{x}_j = 1$, and $\beta''_i = \hat{\beta}_i - \epsilon$, $\eta''_q = \hat{\eta}_q - \epsilon$ for all $q \in \tilde{Q}$ and $\phi''_j = \hat{\phi}_j - \epsilon(b_{ij} - \sum_{q \in \tilde{Q}} u_{qj})$ for all $j \in N$ with $\hat{x}_j = 1$. All other entries of ξ' and ξ'' are the same

as the entries of $\hat{\xi}$. Consequently, $\hat{\xi}$ cannot be an extreme point because $\hat{\xi} = \frac{1}{2}\xi' + \frac{1}{2}\xi''$ and ξ' and ξ'' are also in the convex hull. So (9a) holds.

Finally, since (9a) and (9b) hold, we can conclude that inequality (9c) holds for each extreme point. \square

For an extreme point, β_{o_q} and β_{d_q} can be equal to 1 simultaneously. In this case, η_q may take a value up to 2 as shown in Example 2.

Example 2 Consider the network in Figure 3 and the same individual flow parameters as in Example 1. Additionally, a new flow between nodes 1 and 4 is defined where $l_{60} = 2$, $u_{60} = 18$ and $l_{6j} = u_{6j} = 0$ for all $j \in N$. The aggregated flow parameters are $b_{i0} = 70$ for all $i \in N$, $b_{1j} = 7$ for all $j \in N$ and $b_{ij} = 0$ for all $i \in \{2, 3, 4\}$ and $j \in N$.

When the problem is solved for $p = 2$, the optimal value is 80.8 and the optimal solution is $x_1 = x_4 = 1$, $y_1 = y_5 = 1$, $\beta_1 = \beta_4 = 1$, $\phi_1 = \phi_4 = 6.4$, $\eta_2 = \eta_4 = 1$ and $\eta_6 = 2$. The optimal value decreases to 73.8 when we impose an upper bound of 1 to η variables.

By Proposition 1, we can update the big M values in model (8) as $M_j^u = \sum_{i \in N} b_{ij} + \sum_{q \in Q} l_{qj}$ and $M_j^l = -2 \sum_{q \in Q} u_{qj}$ for each $j \in N$.

3.4 A disaggregated formulation

Next, to linearize problem (4), we define three new vectors Φ^b , Φ^l and Φ^u where $\Phi_{ij}^b = x_j \beta_i$ for each $i, j \in N$, $\Phi_{qj}^l = x_j \theta_q$ and $\Phi_{qj}^u = x_j \eta_q$ for each $j \in N$ and $q \in Q$. Since the new bilinear terms are nonnegative, their lower bounds are 0 and thus McCormick inequalities (7b)-(7d) are used to linearize them. Inequalities (9a)-(9c) also hold for an extreme point of the convex hull of a polyhedron defined by constraints (4b)-(4e) and the McCormick inequalities of the new bilinear terms (it can be proven in the same way as Proposition 1). Accordingly, the upper bounds can be chosen as 1 for Φ_{ij}^b and Φ_{qj}^l , and 2 for Φ_{qj}^u . Since the problem is a maximization problem and Φ^b and Φ^l have nonnegative coefficients in the objective function, lower bounding constraints for the linearization of these terms can be omitted. Similarly, since Φ^u has a nonpositive objective coefficient, its upper bounding constraints can be omitted. Subsequently, problem (4) is rewritten as

(Decision-dependent Robust RSLP-R-Disaggregated Model)

$$\max \sum_{i \in N} b_{i0} \beta_i + \sum_{q \in Q} (l_{q0} \theta_q - u_{q0} \eta_q) + \sum_{j \in N} \left(\sum_{i \in N} b_{ij} \Phi_{ij}^b + \sum_{q \in Q} (l_{qj} \Phi_{qj}^l - u_{qj} \Phi_{qj}^u) \right) \quad (10a)$$

$$\text{s.t. } \beta_{o_q} + \beta_{d_q} + \theta_q - \eta_q \leq y_q \quad \forall q \in Q \quad (10b)$$

$$\Phi_{ij}^b \leq x_j \quad \forall i \in N, j \in N \quad (10c)$$

$$\Phi_{ij}^b \leq \beta_i \quad \forall i \in N, j \in N \quad (10d)$$

$$\Phi_{qj}^l \leq x_j \quad \forall q \in Q, j \in N \quad (10e)$$

$$\Phi_{qj}^l \leq \theta_q \quad \forall q \in Q, j \in N \quad (10f)$$

$$\Phi_{qj}^u \geq \eta_q + 2(x_j - 1) \quad \forall q \in Q, j \in N \quad (10g)$$

$$(x, y) \in \mathcal{D} \quad (10h)$$

$$\beta_i \geq 0 \quad \forall i \in N \quad (10i)$$

$$\theta_q, \eta_q \geq 0 \quad \forall q \in Q \quad (10j)$$

$$\Phi_{ij}^b \geq 0 \quad \forall i \in N, j \in N \quad (10k)$$

$$\Phi_{qj}^l, \Phi_{qj}^u \geq 0 \quad \forall j \in N, q \in Q. \quad (10l)$$

3.5 A Benders reformulation

As the problem size grows, we encounter difficulties in solving the aggregated and disaggregated models because of the big M constraints and the large number of variables, respectively. In this section, we propose a Benders reformulation based on the disaggregated model.

Benders decomposition (Benders (1962)) has been used to solve various problems related to facility location, transportation, vehicle routing, and network design (Rahmaniani et al. (2017)). Here we provide a Benders reformulation and implement the decomposition algorithm in a branch-and-cut framework. This reformulation has the advantage that we can characterize the optimality cuts and separate them by inspection.

If we fix $x = \hat{x}$, $y = \hat{y}$, $\beta = \hat{\beta}$, $\theta = \hat{\theta}$ and $\eta = \hat{\eta}$ in the disaggregated model, then we obtain the following Benders subproblem:

$$\max \sum_{j \in N} \left(\sum_{i \in N} b_{ij} \Phi_{ij}^b + \sum_{q \in Q} (l_{qj} \Phi_{qj}^l - u_{qj} \Phi_{qj}^u) \right) \quad (11a)$$

$$\text{s.t. } \Phi_{ij}^b \leq \hat{x}_j \quad \forall i \in N, j \in N \quad (11b)$$

$$\Phi_{ij}^b \leq \hat{\beta}_i \quad \forall i \in N, j \in N \quad (11c)$$

$$\Phi_{qj}^l \leq \hat{x}_j \quad \forall q \in Q, j \in N \quad (11d)$$

$$\Phi_{qj}^l \leq \hat{\theta}_q \quad \forall q \in Q, j \in N \quad (11e)$$

$$\Phi_{qj}^u \geq \hat{\eta}_q + 2(\hat{x}_j - 1) \quad \forall q \in Q, j \in N \quad (11f)$$

$$\Phi_{ij}^b \geq 0 \quad \forall i \in N, j \in N \quad (11g)$$

$$\Phi_{qj}^l, \Phi_{qj}^u \geq 0 \quad \forall j \in N, q \in Q. \quad (11h)$$

As \hat{x} , $\hat{\beta}$ and $\hat{\theta}$ are nonnegative and $\hat{\eta}_q \leq 2$ for all $q \in Q$, this problem is always feasible. Hence, we only need optimality cuts. The subproblem decomposes first for each $j \in N$ and then for each $i \in N$ and $q \in Q$ and can be solved by inspection: $\Phi_{ij}^b = \min\{\hat{x}_j, \hat{\beta}_i\}$ for all $i \in N$ and $j \in N$, $\Phi_{qj}^l = \min\{\hat{x}_j, \hat{\theta}_q\}$ and $\Phi_{qj}^u = \max\{0, \hat{\eta}_q + 2(\hat{x}_j - 1)\}$ for all $q \in Q$ and $j \in N$ is an optimal solution. Using this, we obtain the following Benders reformulation:

$$\max \sum_{i \in N} b_{i0} \beta_i + \sum_{q \in Q} (l_{q0} \theta_q - u_{q0} \eta_q) + \sum_{j \in N} \phi_j \quad (12a)$$

$$\text{s.t. } \phi_j \leq \sum_{i \in N_j} b_{ij} x_j + \sum_{i \in N \setminus N_j} b_{ij} \beta_i + \sum_{q \in Q_j^l} l_{qj} x_j + \sum_{q \in Q \setminus Q_j^l} l_{qj} \theta_q - \sum_{q \in Q_j^u} u_{qj} \eta_q \quad (12b)$$

$$\beta_{o_q} + \beta_{d_q} + \theta_q - \eta_q \leq y_q \quad \forall q \in Q \quad (12c)$$

$$(x, y) \in \mathcal{D} \quad (12d)$$

$$\beta_i \geq 0 \quad \forall i \in N \quad (12e)$$

$$\theta_q, \eta_q \geq 0 \quad \forall q \in Q \quad (12f)$$

We solve this formulation using a branch-and-cut algorithm. We separate the optimality cuts (12b) at

both fractional and integer solutions at the root node and only at integer solutions at the remaining nodes of the branch-and-cut tree. The separation is done by inspection: For a given vector $(\hat{x}, \hat{y}, \hat{\beta}, \hat{\eta}, \hat{\theta})$, we let $N_j = \{i \in N : \hat{x}_j \leq \hat{\beta}_i\}$, $Q_j^l = \{q \in Q : \hat{x}_j \leq \hat{\theta}_q\}$ and $Q_j^u = \{q \in Q : 0 \leq \hat{\eta}_q + 2(\hat{x}_j - 1)\}$. Note that this separation is exact and polynomial.

4 Computational Experiments

In this section, we present the details of the instances used in the computational experiments and analyze the computational results obtained. The computational results are given in two sections. First, we set a one-hour time limit and evaluate the performances of the branch-and-cut algorithms to solve the Benders reformulation and the aggregated and disaggregated formulations. Then, for the second set of experiments, we solve the instances to optimality without time limitation because some instances may take longer than an hour to solve. We compare the station location decisions obtained by solving the deterministic, robust (without decision-dependency, i.e., the aggregated and individual flow bounds are the same for all x) and decision-dependent robust problems. We assess the importance of considering only uncertainty, and uncertainty and decision-dependency simultaneously. In these experiments, we also examine the effect of different parameter settings on the results. Under all settings, we highlight the gain of incorporating uncertainty and decision-dependency into strategic-level decisions.

The experiments are performed on a 64-bit machine with AMD Ryzen 5 PRO 3500U 2.10 GHz and 16 GB of RAM. The algorithms are coded in Java using CPLEX 12.9 and implemented using the lazy and user constraint callback functions in CPLEX. For the settings, default strategies provided by CPLEX are used.

4.1 Details of the instances

We use four data sets to perform our computational experiments. The first data set is the 25-node network introduced by Simchi-Levi & Berman (1988) for the traveling salesman location problem. Figure 4 depicts the network where the node sizes are proportional to their weights. This network is commonly used in the RSLP literature (Hodgson (1990), Kuby & Lim (2005), Kim & Kuby (2012), Capar et al. (2013), MirHassani & Ebrazi (2013)) where each pair of nodes is regarded as an O-D pair and flow volumes are generated by using the node weights provided by Simchi-Levi & Berman (1988) and the gravity model proposed by Hodgson (1990). According to the gravity model, the flow volume of an O-D pair is directly proportional to the weights of its endpoints and inversely proportional to the shortest path length between its endpoints. We generated the other data sets based on the road network of Belgium (Figure 5). In these Belgium data sets, the nodes represent the municipalities, in total 581, and the edges (two-way arcs) represent the roads (only highways are considered) between the municipalities. The populations of municipalities are regarded as their weights. The total numbers of nodes, edges, and O-D pairs for each data set are given in Table 1. For the BE-1 and BE-2 networks, the most populous 83 municipalities whose populations are higher than 30,000, and for the BE-3 network, the most populous 115 municipalities whose populations are higher than 25,000, are used as the endpoints of the O-D pairs. The most populous 83 municipalities are taken as the set of nodes for the BE-1 network, whereas all municipalities are taken as the set of nodes for the BE-2 and BE-3 networks. For each data set, all nodes in the network are regarded as candidate refueling station sites.

Data Set	Nodes	Edges	O-D pairs
25-node	25	43	300
BE-1	83	236	3403
BE-2	581	911	3403
BE-3	581	911	6555

Table 1: Properties of the instances

Let \bar{f}_q and \bar{b}_i represent the nominal values of individual and aggregate flow bounds, respectively. We set \bar{b}_i to $\kappa \sum_{q \in Q: i=o_q \vee i=d_q} \bar{f}_q$, $\kappa \geq 0$. The nominal values of individual flows are calculated based on the gravity model. To consider the whole Belgian population, while computing the flow volumes, we add the population of a node that is not regarded as an origin or a destination to the nearest node that is the origin or the destination of an O-D pair. The populations are obtained from STATBEL (2020) and the road distances are taken from Google Maps (2021). The generated Belgian data sets are available at <https://github.com/OzlemMahmutogullari/BeData>.

We define robustness parameter ψ , $0 \leq \psi \leq 1$ to adjust the level of robustness. Moreover, we define decision-dependency parameters λ_{ij}^B and λ_{qj}^I to represent the effect of opening an AFS at node j on the estimated aggregate flow bound of node i and the interval bounds of flow q , respectively. We suppose that when a new station opens in the neighborhood, vehicle flows increase and individual flow deviations diminish. Accordingly, we let

$$b_{i0} = \bar{b}_i, l_{q0} = \bar{f}_q(1 - \psi), u_{q0} = \bar{f}_q(1 + \psi),$$

$$b_{ij} = \lambda_{ij}^B \bar{b}_i, l_{qj} = \bar{f}_q(1 + \psi) \lambda_{qj}^I, u_{qj} = \bar{f}_q(1 - \psi) \lambda_{qj}^I.$$

It is noteworthy that if we set all decision-dependency parameters to zero, then we only consider the uncertainty in the flow volumes and obtain the robust RSLP-R; if we set the robustness parameter to zero as well as the decision-dependency parameters, then we obtain the deterministic RSLP-R because, in this case, there is one possible realization for each flow.

We assume that the flow parameters are affine functions of the location decisions. The choices for decision-dependency parameters λ^B and λ^I determine the degree of a station's effect on vehicle flows. As presented by Basciftci et al. (2021), we consider that the stations affect the flows by the magnitudes of their proximities to the flows, i.e., the closer stations have more effect than the further ones, and choose the decision-dependency parameters using nonincreasing functions of distances between stations and flows. In this regard, the aggregate flow at node i is most affected by the nearest stations, while the flow q is most affected by the stations on the shortest path between its endpoints. The nonincreasing functions, $e^{-\delta_{ij}/2}$ and $e^{-(\delta_{o_q,j} + \delta_{j,d_q})/2}$, are used to assign values to the decision-dependency parameters λ^B and λ^I , respectively. By assumption, $\sum_{j \in N} \lambda_{qj}^I < 1$ for each q so that $l_q(x) < u_q(x)$ under all station deployment plan. Moreover, we assume that $\sum_{j \in N} \lambda_{ij}^B \leq 0.5$ for each i and select κ values during the experiments without violating the feasibility of hybrid uncertainty set. κ is taken as 1 and ψ is taken as 0.8 if they are not stated otherwise. The decision-dependency parameters are normalized for each node i and flow q dividing each λ_{ij}^B by $2 \sum_{j \in N} \lambda_{ij}^B$ and each λ_{qj}^I by $1.05 \sum_{j \in N} \lambda_{qj}^I$.

The ranges of the vehicles are selected based on the information provided by the report of the European Commission (2016) for the experiments. According to this report, the recommended distances between two consecutive AFSs are 60 km for EV, 195 km for CNG, 220 km for LPG, 290 km for LNG, and 295 km for H2 vehicles. The driver tolerances are expressed in percentages of the shortest path

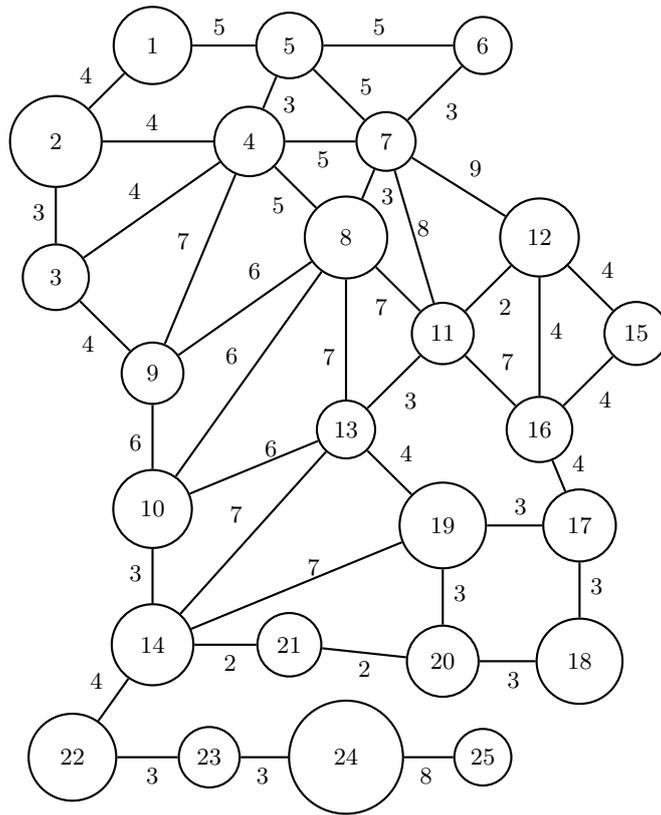


Figure 4: 25-node network

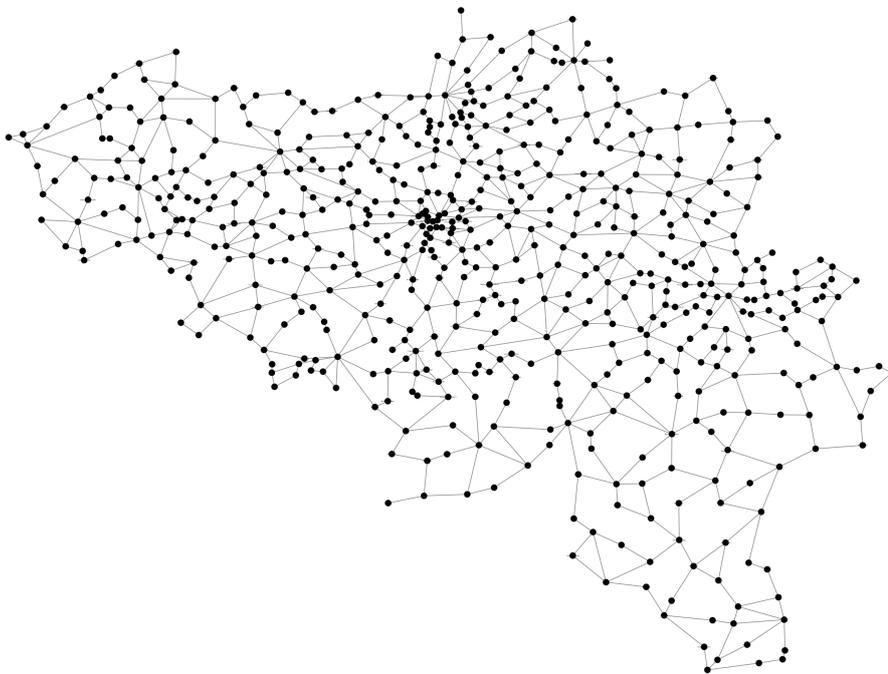


Figure 5: Belgian network

lengths. We assume that all drivers have the same tolerance percentage. However, the models can also tackle various tolerance percentages of the drivers who use the same endpoints.

4.2 Comparison of exact approaches

First, we compare the computational performance of the three proposed exact solution approaches, namely branch-and-cut algorithms for the two formulations and the Benders reformulation. For this experiment, we use the 25-node data set with a range of 8 and a tolerance of 10%. Table 2 shows the results for different numbers of stations. In the table, p is the given number of refueling stations, $\# Cuts$ is the total number of Benders cuts added, $\# Nodes$ is the number of nodes explored in the branch-and-cut tree, and CPU is the total time spent in seconds solving the instance set.

p	Benders			Aggregated		Disaggregated	
	# Cuts	# Nodes	CPU	# Nodes	CPU	# Nodes	CPU
1	225	17	0.88	26	2.56	10	7.78
2	725	43	1.14	91	2.01	67	15.29
3	800	57	1.20	104	2.42	34	10.98
4	225	96	0.72	182	2.82	55	13.37
5	1450	179	2.58	203	3.08	122	17.67
6	1025	236	2.21	394	4.39	153	23.33
7	975	381	3.70	229	3.79	183	32.21
8	1150	414	7.18	277	4.78	182	32.66
9	1125	509	6.57	443	5.50	335	39.85
10	975	655	7.60	240	4.32	323	36.06
11	1475	346	7.91	265	4.26	182	34.24
12	1000	466	5.04	249	3.53	168	30.73
13	2075	372	12.47	437	3.97	198	34.40
14	325	499	2.01	161	2.68	99	26.60
15	375	470	2.19	94	1.86	194	36.21
16	725	146	1.93	102	1.73	133	26.13
17	1600	59	2.83	83	1.83	57	15.95
18	1850	38	2.82	33	1.20	62	13.99
19	1475	18	1.87	26	1.06	28	11.62
20	950	13	0.91	16	1.11	38	10.46
21	300	7	0.32	10	0.92	18	5.65
22	550	0	0.28	0	0.16	0	1.04
23	325	0	0.14	0	0.07	0	0.67
24	225	0	0.11	0	0.07	0	0.75
25	150	0	0.07	0	0.10	0	0.78

Table 2: Results for the 25-node network with range 8 and tolerance 10%

As shown in Table 2, all proposed approaches can solve the data set to optimality in one minute for each p value. All methods can solve the problem instances in less time if the p value is small or large, rather than medium. The total number of nodes explored is the least for the disaggregated model, but the other methods outperform the disaggregated model in terms of CPU time. Since the comparison between the results of the Benders reformulation and the aggregated formulation is not clear for the 25-node network, we also investigate the performance of these approaches for larger-sized networks. We set the time limit to one hour. Accordingly, for each BE data set with varying range and tolerance values, the averages of results obtained for different numbers of stations ($p = \{1, \dots, 15\}$) are provided in Table 3. In the table, columns $\# Opt$ show the total number of instances that can be solved to optimality within the time limit, and columns $Opt. Gap$ represent the percentage gap between the best upper bound (BUB) and the best lower bound (BLB) found within the time limit for the instances that cannot be solved optimally within the time limit. This percentage gap is computed as $100 \times \frac{BUB - BLB}{BLB}$. Table 3 demonstrates that the Benders reformulation outperforms the aggregated formulation for all measures under each parameter choice. The algorithm can solve a greater number of instances in less time by exploring a smaller number of nodes in the branch-and-cut tree. When both approaches fail to reach optimality within the time limit, the Benders reformulation ends up with lower percentage gaps.

Data	r	w	Benders				Aggregated			
			# Opt (/15)	# Nodes	CPU	Opt. Gap	# Opt (/15)	# Nodes	CPU	Opt. Gap
BE-1	200	10%	15	837.67	762.40	0.00%	8	7366.40	1963.44	4.69%
		20%	15	1013.07	926.10	0.00%	11	6099.33	1758.51	1.42%
	250	10%	15	799.20	655.99	0.00%	10	8398.27	1643.64	3.15%
		20%	15	867.07	612.82	0.00%	12	6719.33	1409.42	0.89%
BE-2	200	10%	13	527.47	808.69	0.48%	11	3594.33	1734.50	2.68%
		20%	13	631.20	1028.21	0.48%	12	1876.20	1621.14	1.10%
	250	10%	14	843.07	572.17	0.00%	11	3460.40	1584.99	2.32%
		20%	13	570.93	1051.22	0.48%	12	1877.93	1444.63	1.01%
BE-3	60	10%	5	137.00	2519.86	15.14%	4	586.60	2879.92	39.83%
		20%	4	155.73	2739.59	19.20%	3	223.87	3233.47	56.57%
	100	10%	6	258.20	2515.22	5.95%	4	722.20	3121.89	30.10%
		20%	5	136.40	2884.03	9.59%	3	207.67	3151.91	46.16%

Table 3: Average results for larger data sets

4.3 The changes in total covered flows

Next, we analyze how the station location decisions change considering uncertainty and decision-dependency under different parameter settings. We present the results of the deterministic, robust, and decision-dependent robust problems, how well the deterministic solution performs under uncertainty and how well the robust solution performs under decision-dependent uncertainty. As stated before, the robust problem refers to the decision-dependent robust problem with decision-dependency parameters $\lambda^I = \mathbf{0}$ and $\lambda^B = \mathbf{0}$, and all instances for each problem are solved to optimality.

Tables 4 and 5 show the optimal values of the problems, the worst case total flows using the deterministic solutions under the flow uncertainty set and using the robust solutions under the decision-dependent flow uncertainty set, and the percentage increments in the total covered flows if the robust and the decision-dependent robust solutions are used. In the tables, z_{det} , z_r and z_{dd} represent the optimal values of the deterministic, robust, and decision-dependent robust problems, respectively, $z_{det,r}$ is the objective function value of the robust problem if the stations are selected according to the optimal solution of the deterministic problem and $z_{r,dd}$ is the objective function value of the decision-dependent problem where the station locations are as in the robust solution. The percentage increments in the flows that can be covered when uncertainty and decision-dependency are taken into account are given in column *Extra*. The value $\frac{z_r - z_{det,r}}{z_{det,r}} \times 100$ gives the percentage improvement in the total flow covered if the robust solution is used instead of the deterministic solution under uncertainty. The percentage of extra flow that can be covered in a decision-dependent environment, i.e., we consider the decision-dependency as well as the uncertainty, can be computed as $\frac{z_{dd} - z_{r,dd}}{z_{r,dd}} \times 100$.

If the optimal station locations of the deterministic problem are compared with those of the robust problem, it is seen that the locations change for 4 out of 15 instances in the 25-node network and for 9 out of 15 instances in the BE-2 network. If we compare the optimal station locations of the robust model with those of the decision-dependent robust model, we observe that there is a change in the locations for 12 out of 15 instances in the 25-node network and for 10 out of 15 instances in the BE-2 network. As examples, Figures 6 and 7 illustrate the optimal station locations for the instances in Tables 4 and 5 with $p = 5$ and $p = 7$, respectively. In the deterministic case, the stations are situated at cities 2, 14, 17, 20, and 23 as shown in Figure 6. The station in city 2 (red) is moved to city 16 (yellow), which is closer to the big cities, under the uncertain flow realizations. If we make the uncertainty set dependent on the location decisions, city 16 is replaced by city 18 (green), which is one of the largest cities and close to high volumes of vehicle flows. The deterministic model places 7 stations in Antwerp, Ghent, Liège, Brussels, Schaerbeek, Uccle, and Leuven in the BE-2

p	Deterministic	Robust			Decision-Dependent Robust		
	z_{det}	z_r	$z_{det,r}$	Extra (%)	z_{dd}	$z_{r,dd}$	Extra (%)
1	17.13	8.10	8.10	0.00	14.68	14.68	0.00
2	32.58	17.40	17.40	0.00	32.51	32.51	0.00
3	44.41	31.95	31.95	0.00	46.35	45.11	2.76
4	55.96	46.56	46.56	0.00	59.96	59.96	0.00
5	63.52	49.86	48.86	2.05	69.06	65.28	5.79
6	68.08	56.31	51.88	8.54	77.93	71.96	8.30
7	72.32	59.75	58.61	1.93	87.02	84.51	2.97
8	77.86	67.54	67.54	0.00	95.97	94.80	1.24
9	82.77	74.71	74.49	0.29	104.13	102.06	2.03
10	90.06	83.33	83.33	0.00	113.50	112.31	1.06
11	94.41	90.25	90.25	0.00	123.25	119.54	3.10
12	96.80	94.45	94.45	0.00	132.63	126.22	5.08
13	97.77	96.21	96.21	0.00	140.24	128.54	9.10
14	98.43	97.23	97.23	0.00	150.00	128.85	16.42
15	98.74	97.79	97.79	0.00	157.45	134.49	17.07

Table 4: Total refueled vehicle flow analysis for the 25-node network with range 8 and tolerance 10%

p	Deterministic	Robust			Decision-Dependent Robust		
	z_{det}	z_r	$z_{det,r}$	Extra (%)	z_{dd}	$z_{r,dd}$	Extra (%)
1	34.74	17.26	17.26	0.00	23.88	23.88	0.00
2	44.26	26.27	26.27	0.00	38.09	38.09	0.00
3	50.19	31.56	31.56	0.00	45.69	45.69	0.00
4	55.42	36.14	36.14	0.00	52.86	52.74	0.23
5	60.52	40.55	40.36	0.48	59.83	59.83	0.00
6	64.85	44.40	43.44	2.21	65.52	65.52	0.00
7	68.80	48.68	48.02	1.37	70.58	69.52	1.52
8	72.13	53.18	52.05	2.19	75.17	72.71	3.39
9	75.04	57.63	57.08	0.97	79.72	77.18	3.28
10	77.65	61.75	61.73	0.03	84.35	82.02	2.85
11	79.62	64.68	64.48	0.33	87.97	85.00	3.49
12	81.56	67.39	67.03	0.53	91.05	87.13	4.49
13	83.21	70.06	69.95	0.16	93.92	90.02	4.33
14	84.69	72.62	72.62	0.00	96.62	92.28	4.70
15	86.10	75.16	75.16	0.00	99.37	97.50	1.92

Table 5: Total refueled vehicle flow analysis for the BE-2 network with range 250 and tolerance 10%

network, as shown in Figure 7. Dilbeek (population 94,033) replaces Leuven (195,679) if we switch to the robust model, and Dilbeek is replaced by Charleroi (324,686) in response to considering the decision-dependency of flow uncertainty sets.

Although the deterministic station locations are robust for most instances of the small network, the extra flow that can be covered by employing the robust solution can be drastic for some instances, such as 8.54%. Moreover, when the decision-dependent uncertainty is considered, the optimal station locations are more likely to change, and the advantage of integrating decision-dependency in the flow uncertainty sets grows for larger p values. These findings underline the need of incorporating decision-dependent uncertainty into the strategic-level planning of locating stations, regardless of network size.

Table 6 presents the average percentage flow increments for the BE-1 and BE-2 data sets when uncertainty and decision-dependency are taken into account for a variety of range and tolerance combinations. The main difference between these two data sets is that the BE-2 network has a greater number of candidate locations than the BE-1 network, as if the stations can be located at some points on the edges, not only the endpoints, of the BE-1 network. In Table 6, $\# Ins$ represents the number of instances that have different optimal solutions for the corresponding problems, and $Mean$ is the

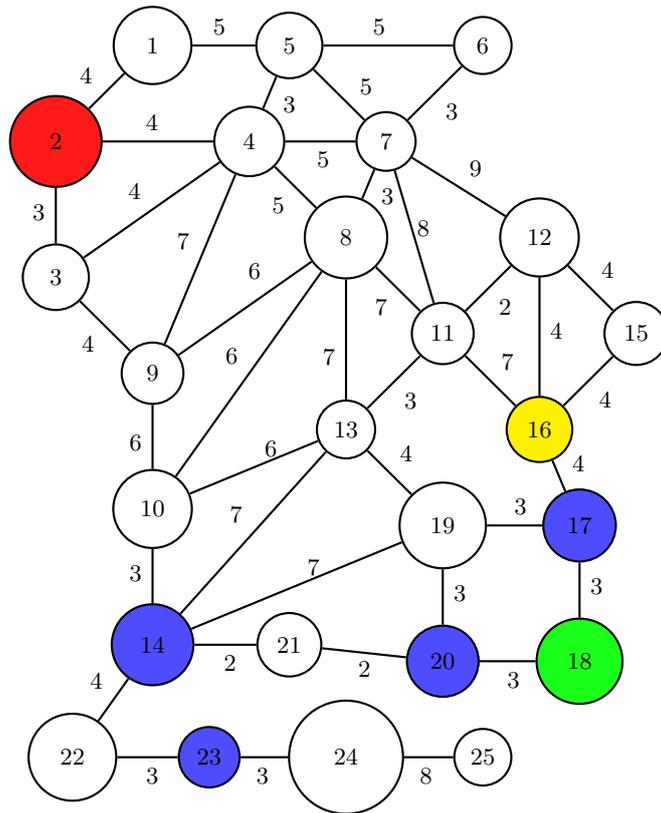


Figure 6: Optimal locations for the 25-node network with $p = 5$

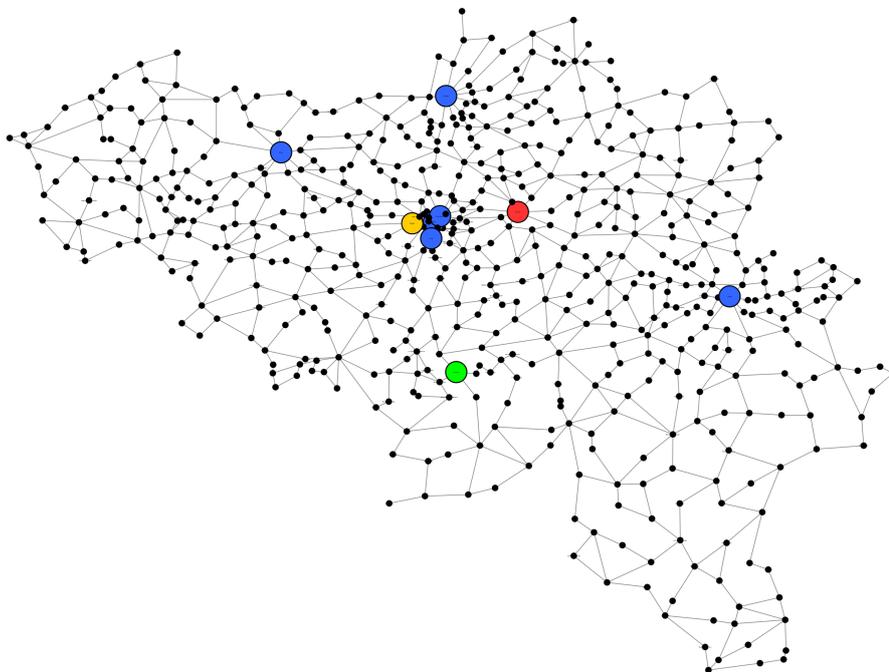


Figure 7: Optimal locations for the BE-2 network with $p = 7$

average of extra flows and Max is the maximum of extra flows if we apply the formulation that is suitable for the environment. $Mean$ and Max are given in percentages. For each range and tolerance

Data	r	w	Uncertainty			Decision-Dependency		
			# Ins (/45)	Mean (%)	Max (%)	# Ins (/15)	Mean (%)	Max (%)
BE-1	200	10%	19	1.84	5.37	10	4.69	10.45
		20%	16	0.59	4.69	10	4.95	10.55
	250	10%	18	1.02	7.56	10	5.38	11.07
		20%	16	0.51	2.04	11	5.09	11.00
	300	10%	16	0.87	3.27	11	4.84	11.29
		20%	14	0.72	2.34	11	5.11	11.04
BE-2	200	10%	16	0.56	2.63	8	3.07	4.33
		20%	10	1.01	2.54	9	3.16	5.61
	250	10%	15	0.59	2.21	10	3.02	4.70
		20%	11	0.37	1.45	10	2.86	4.61
	300	10%	17	0.53	2.70	10	3.32	5.54
		20%	10	0.51	2.09	10	2.79	4.73

Table 6: Average extra flows that can be covered considering the uncertainty and decision-dependency for the BE-1 and BE-2 networks under different parameter settings

setting, the problems are solved to locate $p = \{1, \dots, 45\}$ stations for the uncertainty and to locate $p = \{1, \dots, 15\}$ stations for the decision-dependency. We can deduce, from the table, that the station locations are more likely to change under uncertain flow volumes as the tolerance decreases for both data sets. The effects of station location changes under uncertainty are greater in the BE-1 network than in the BE-2 network. This argument is also true for the decision-dependent case. Additionally, under decision-dependent uncertainty, changes in the station locations are common in all instances and the percentage of extra flows can be up to 11.29%.

4.3.1 The effects of robustness parameter values

We examine how total covered flows change under different robustness parameter values. The results are reported in Tables 7 and 8 when considering uncertainty and decision-dependency, respectively. As expected, when we compare the station locations of the deterministic model with those of the robust model under hybrid flow uncertainty sets with robustness parameters of 0.5 and 0.8, we observe that the gain of considering uncertainty declines as the uncertainty level gets smaller. On the other hand, incorporating decision-dependency into uncertainty sets provides a significant benefit for each uncertainty level.

r	w	$\psi = 0.5$			$\psi = 0.8$		
		# Ins (/45)	Mean (%)	Max (%)	# Ins (/45)	Mean (%)	Max (%)
200	10%	15	1.03	2.35	19	1.84	5.37
	20%	9	0.34	1.93	16	0.59	4.69
	50%	9	1.08	3.18	10	2.86	10.67
250	10%	16	0.49	3.61	18	1.02	7.56
	20%	12	0.25	0.71	16	0.51	2.04
	50%	6	0.86	3.93	9	1.79	8.51
300	10%	14	0.34	1.61	16	0.87	3.27
	20%	10	0.36	1.21	14	0.72	2.34
	50%	7	0.76	4.18	8	1.99	8.72

Table 7: Average extra flows that can be covered considering the uncertainty for the BE-1 network with different ψ values

It is also observed that, in Table 7, employing the robust model under uncertain flows benefits a greater number of instances with different p values when the tolerance is reduced. However, the

r	w	$\psi = 0.5$			$\psi = 0.8$		
		# Ins (/15)	Mean (%)	Max (%)	# Ins (/15)	Mean (%)	Max (%)
200	10%	11	4.85	10.17	10	4.69	10.45
	20%	10	5.44	10.12	10	4.95	10.55
	50%	13	5.76	11.13	11	5.23	10.44
250	10%	11	5.22	10.63	10	5.38	11.07
	20%	11	5.52	10.70	11	5.09	11.00
	50%	10	7.37	11.15	13	5.50	10.46
300	10%	11	5.49	10.84	11	4.84	11.29
	20%	11	5.49	10.77	11	5.11	11.04
	50%	12	6.87	11.07	13	5.41	10.40

Table 8: Average extra flows that can be covered considering the decision-dependency for the BE-1 network with different ψ values

percentage increase in total covered flows on average increases when the tolerance is raised. As seen in Table 8, the advantage of recognizing the decision-dependency by using the decision-dependent robust station locations increases as the range and tolerance increase.

4.3.2 The effects of aggregated flow bound values

We further analyze the effects of κ values, i.e., the aggregated flow bound estimations, on the optimal station locations. In Tables 9 and 10, we present the performances of the deterministic station locations under the uncertain flow realizations and of the robust locations under the decision-dependent uncertain flows for $\kappa = \{0.8, 1.0, 1.2\}$. In the tables, it can be seen that the problem instances become more sensitive to the changes in the flows as the aggregate bounds of the nodes decrease. It is worth noting that, for the robust problem, this pattern is not followed when κ converges to 0. If $\kappa = 0$, the hybrid flow uncertainty set is defined only by the individual flow volumes, and the worst case realizations of all pairwise flows are equal to $(1 - \psi)$ proportion of their nominal values, i.e., their lower bounds. Consequently, when $\kappa = 0$, the deterministic and robust locations are the same. However, the gain of incorporating decision-dependency remains and continues to increase as κ decreases.

r	w	$\kappa = 0.8$			$\kappa = 1.0$			$\kappa = 1.2$		
		# Ins (/45)	Mean (%)	Max (%)	# Ins (/45)	Mean (%)	Max (%)	# Ins (/45)	Mean (%)	Max (%)
200	10%	28	1.74	4.98	19	1.84	5.37	13	1.20	4.98
	20%	23	1.39	5.02	16	0.59	4.69	6	0.69	2.09
	50%	17	1.60	5.54	10	2.86	10.67	7	1.05	3.28
250	10%	26	1.65	7.34	18	1.02	7.56	10	0.13	0.51
	20%	23	1.21	3.92	16	0.51	2.04	8	0.71	2.70
	50%	17	1.14	2.97	9	1.79	8.51	5	0.84	3.76
300	10%	23	1.84	6.14	16	0.87	3.27	5	0.27	0.76
	20%	19	1.76	5.23	14	0.72	2.34	8	0.57	3.21
	50%	16	1.14	3.16	8	1.99	8.72	3	1.23	3.64

Table 9: Average extra flows that can be covered considering the uncertainty for the BE-1 network with different κ values

In Table 9, although the average percentage changes in the objective function values become smaller as κ gets larger, we observe that recognizing uncertainty can still be quite valuable for some instances. For example, average percentage flow increments are generally around 1% if $\kappa = 1.2$, but 4.98% increment can be obtained for an instance. As seen in Table 10, employing decision-dependent uncertainty sets provides benefits for all instances under each parameter setting, and the gain can be up to 22.74% when $\kappa = 0.8$.

r	w	$\kappa = 0.8$			$\kappa = 1.0$			$\kappa = 1.2$		
		# Ins (/15)	Mean (%)	Max (%)	# Ins (/15)	Mean (%)	Max (%)	# Ins (/15)	Mean (%)	Max (%)
200	10%	12	13.59	20.51	10	4.69	10.45	13	2.99	6.51
	20%	12	12.05	21.13	10	4.95	10.55	13	3.08	6.36
	50%	13	12.62	21.08	11	5.23	10.44	12	1.74	3.96
250	10%	12	13.15	21.53	10	5.38	11.07	11	2.90	5.79
	20%	13	12.21	22.59	11	5.09	11.00	12	2.36	4.00
	50%	12	14.67	21.32	13	5.50	10.46	13	1.51	3.71
300	10%	12	14.31	21.84	11	4.84	11.29	12	3.01	5.37
	20%	12	13.88	22.74	11	5.11	11.04	12	2.29	3.72
	50%	12	15.58	21.17	13	5.41	10.40	11	1.69	3.17

Table 10: Average extra flows that can be covered considering the decision-dependency for the BE-1 network with different κ values

5 Conclusion

In this study, we introduced the robust RSLP-R under decision-dependent flow uncertainty using a hybrid uncertainty model that restricts the hose model by imposing lower and upper bounds on the pairwise flows. We proposed two mixed integer programming formulations as well as a Benders reformulation. We generated new benchmark instances based on the real Belgium road network. Our computational experiments showed that the Benders reformulation outperforms the other formulations for larger instances and that the gains obtained by recognizing the uncertainty and the decision-dependency of AFV flows can be significant.

The current study suggests some exciting future research directions: One such direction is to extend the problem by incorporating refueling station capacities and then time-dependent vehicle flows (rush hours, weekends, etc.). Another is to apply the idea of the Benders reformulation to model other problems with a decision-dependent hybrid uncertainty set.

Acknowledgments

This research is supported by the KU Leuven grant 3H180528.

References

- Alternative Fuels Data Center. (2021). *Alternative fuels and advance vehicles*. Retrieved from <https://afdc.energy.gov/fuels/>. Accessed July 16, 2021.
- Altın, A., Yaman, H., & Pınar, M. Ç. (2011). A hybrid polyhedral uncertainty model for the robust network loading problem. In *Performance models and risk management in communications systems* (pp. 157–172). Springer.
- Arslan, O., Karışan, O. E., Mahjoub, A. R., & Yaman, H. (2019). A branch-and-cut algorithm for the alternative fuel refueling station location problem with routing. *Transportation Science*, 53(4), 1107–1125.
- Bapna, R., Thakur, L. S., & Nair, S. K. (2002). Infrastructure development for conversion to environmentally friendly fuel. *European Journal of Operational Research*, 142(3), 480–496.
- Basciftci, B., Ahmed, S., & Gebraeel, N. (2020). Data-driven maintenance and operations scheduling in power systems under decision-dependent uncertainty. *IIEE transactions*, 52(6), 589–602.

- Basciftci, B., Ahmed, S., & Shen, S. (2021). Distributionally robust facility location problem under decision-dependent stochastic demand. *European Journal of Operational Research*, 292(2), 548–561.
- Basciftci, B., Shabbir, A., & Nagi, G. (2019). Adaptive two-stage stochastic programming with an application to capacity expansion planning. *arXiv preprint arXiv:1906.03513*.
- Belgium National Policy Framework. (2016). *Alternative fuels infrastructure*. Retrieved from <https://www.eafo.eu/sites/default/files/npf/1%20BELGIUM%20NPF.en.pdf>. Accessed July 18, 2021.
- Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1), 238–252.
- Bertsimas, D., & Sim, M. (2003). Robust discrete optimization and network flows. *Mathematical Programming*, 98(1), 49–71.
- Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, 52(1), 35–53.
- 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.
- Church, R., & ReVelle, C. (1974). The maximal covering location problem. In *Papers of the regional science association* (Vol. 32, pp. 101–118).
- de Vries, H., & Duijzer, E. (2017). Incorporating driving range variability in network design for refueling facilities. *Omega*, 69, 102–114.
- Duffield, N. G., Goyal, P., Greenberg, A., Mishra, P., Ramakrishnan, K. K., & van der Merive, J. E. (1999). A flexible model for resource management in virtual private networks. In *Proceedings of the conference on applications, technologies, architectures, and protocols for computer communication* (pp. 95–108).
- EAFO. (2020). *Belgium summary*. Retrieved from <https://www.eafo.eu/countries/belgium/1724/summary>. Accessed July 30, 2021.
- Erlenkotter, D. (1977). Facility location with price-sensitive demands: private, public, and quasi-public. *Management Science*, 24(4), 378–386.
- European Alternative Fuels Observatory (EAFO). (2020). *Total number of vehicles (2020)*. Retrieved from <https://www.eafo.eu/europe>. Accessed July 18, 2021.
- European Commission. (2016). *Clean transport - support to the member states for the implementation of the directive on the deployment of alternative fuels infrastructure good practice examples*. Retrieved from <https://ec.europa.eu/transport/sites/default/files/themes/urban/studies/doc/2016-01-alternative-fuels-implementation-good-practices.pdf>. Accessed July 30, 2021.

- European Environment Agency. (2019). *Greenhouse gas emissions from transport in Europe*. Retrieved from <https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12>. Accessed July 16, 2021.
- Faridimehr, S., Venkatachalam, S., & Chinnam, R. B. (2018). A stochastic programming approach for electric vehicle charging network design. *IEEE Transactions on Intelligent Transportation Systems*, *20*(5), 1870–1882.
- Fingerhut, J. A., Suri, S., & Turner, J. S. (1997). Designing least-cost nonblocking broadband networks. *Journal of Algorithms*, *24*(2), 287–309.
- Goel, V., & Grossmann, I. E. (2004). A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. *Computers & chemical engineering*, *28*(8), 1409–1429.
- Goel, V., & Grossmann, I. E. (2005). A lagrangean duality based branch and bound for solving linear stochastic programs with decision dependent uncertainty. In *Computer aided chemical engineering* (Vol. 20, pp. 55–60). Elsevier.
- Goel, V., & Grossmann, I. E. (2006). A class of stochastic programs with decision dependent uncertainty. *Mathematical programming*, *108*(2), 355–394.
- Google Maps. (2021). *Directions*. Retrieved from <https://maps.google.com>. Accessed February 01, 2021.
- Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, *12*(3), 450–459.
- Hakimi, S. L. (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research*, *13*(3), 462–475.
- Hodgson, M. J. (1990). A flow-capturing location-allocation model. *Geographical Analysis*, *22*(3), 270–279.
- Honma, Y., & Kuby, M. (2019). Node-based vs. path-based location models for urban hydrogen refueling stations: Comparing convenience and coverage abilities. *International journal of hydrogen energy*, *44*(29), 15246–15261.
- Hooshmand, F., MirHassani, S., & Akhavein, A. (2018). Adapting GA to solve a novel model for operating room scheduling problem with endogenous uncertainty. *Operations Research for Health Care*, *19*, 26–43.
- Hooshmand Khaligh, F., & MirHassani, S. (2016). A mathematical model for vehicle routing problem under endogenous uncertainty. *International Journal of Production Research*, *54*(2), 579–590.
- Hosseini, M., & MirHassani, S. (2015). Refueling-station location problem under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, *84*, 101–116.
- Hosseini, M., & MirHassani, S. A. (2017). A heuristic algorithm for optimal location of flow-refueling capacitated stations. *International Transactions in Operational Research*, *24*(6), 1377–1403.

- 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.
- Hosseini, M., Rahmani, A., & Hooshmand, F. (2021). A robust model for recharging station location problem. *Operational Research*, 1–44.
- Hu, J., Li, J., & Mehrotra, S. (2019). A data-driven functionally robust approach for simultaneous pricing and order quantity decisions with unknown demand function. *Operations Research*, *67*(6), 1564–1585.
- Hwang, S. W., Kweon, S. J., & Ventura, J. A. (2015). Infrastructure development for alternative fuel vehicles on a highway road system. *Transportation Research Part E: Logistics and Transportation Review*, *77*, 170–183.
- Hwang, S. W., Kweon, S. J., & Ventura, J. A. (2017). Locating alternative-fuel refueling stations on a multi-class vehicle transportation network. *European Journal of Operational Research*, *261*(3), 941–957.
- Hwang, S. W., Kweon, S. J., & Ventura, J. A. (2020). An alternative fuel refueling station location model considering detour traffic flows on a highway road system. *Journal of Advanced Transportation*, 2020.
- Jonsbråten, T. (1998). Optimization models for petroleum field exploitation [PhD thesis]. *Bergen, Norway: Norwegian School of Economics and Business Administration*.
- Jorge, D., & Correia, G. (2013). Carsharing systems demand estimation and defined operations: a literature review. *European Journal of Transport and Infrastructure Research*, *13*(3).
- Kadri, A. A., Perrouault, R., Boujelben, M. K., & Gicquel, C. (2020). A multi-stage stochastic integer programming approach for locating electric vehicle charging stations. *Computers & Operations Research*, *117*, 104888.
- Kang, J. E., & Recker, W. (2015). Strategic hydrogen refueling station locations with scheduling and routing considerations of individual vehicles. *Transportation Science*, *49*(4), 767–783.
- Kchaou-Boujelben, M., & Gicquel, C. (2020). Locating electric vehicle charging stations under uncertain battery energy status and power consumption. *Computers & Industrial Engineering*, *149*, 106752.
- 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.
- Kim, J.-G., & Kuby, M. (2013). A network transformation heuristic approach for the deviation flow refueling location model. *Computers & Operations Research*, *40*(4), 1122–1131.
- Kuby, M., & Lim, S. (2005). The flow-refueling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences*, *39*(2), 125–145.
- Kuby, M., & Lim, S. (2007). Location of alternative-fuel stations using the flow-refueling location model and dispersion of candidate sites on arcs. *Networks and Spatial Economics*, *7*(2), 129–152.

- Kweon, S. J., Hwang, S. W., & Ventura, J. A. (2017). A continuous deviation-flow location problem for an alternative-fuel refueling station on a tree-like transportation network. *Journal of Advanced Transportation*, 2017.
- Lim, S., & Kuby, M. (2010). Heuristic algorithms for siting alternative-fuel stations using the flow-refueling location model. *European Journal of Operational Research*, 204(1), 51–61.
- Luo, F. (2020). A distributionally-robust service center location problem with decision dependent demand induced from a maximum attraction principle. *arXiv preprint arXiv:2011.12514*.
- Mak, H.-Y., Rong, Y., & Shen, Z.-J. M. (2013). Infrastructure planning for electric vehicles with battery swapping. *Management Science*, 59(7), 1557–1575.
- Martins, F., Felgueiras, C., Smitkova, M., & Caetano, N. (2019). Analysis of fossil fuel energy consumption and environmental impacts in european countries. *Energies*, 12(6), 964.
- McCormick, G. P. (1976). Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems. *Mathematical Programming*, 10(1), 147–175.
- Meraklı, M., & Yaman, H. (2016). Robust intermodal hub location under polyhedral demand uncertainty. *Transportation Research Part B: Methodological*, 86, 66–85.
- Miralinaghi, M., Lou, Y., Keskin, B. B., Zarrinmehr, A., & Shabanpour, R. (2017). Refueling station location problem with traffic deviation considering route choice and demand uncertainty. *International Journal of Hydrogen Energy*, 42(5), 3335–3351.
- MirHassani, S., & Ebrazi, R. (2013). A flexible reformulation of the refueling station location problem. *Transportation Science*, 47(4), 617–628.
- MirHassani, S., Khaleghi, A., & Hooshmand, F. (2020). Two-stage stochastic programming model to locate capacitated EV-charging stations in urban areas under demand uncertainty. *EURO Journal on Transportation and Logistics*, 9(4), 100025.
- Nicholas, M. A., Handy, S. L., & Sperling, D. (2004). Using geographic information systems to evaluate siting and networks of hydrogen stations. *Transportation Research Record*, 1880(1), 126–134.
- Noyan, N., Rudolf, G., & Lejeune, M. (2018). Distributionally robust optimization with decision-dependent ambiguity set. *Optimization Online*.
- Poss, M. (2013). Robust combinatorial optimization with variable budgeted uncertainty. *4OR*, 11(1), 75–92.
- Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, 259(3), 801–817.
- Riemann, R., Wang, D. Z., & Busch, F. (2015). Optimal location of wireless charging facilities for electric vehicles: flow-capturing location model with stochastic user equilibrium. *Transportation Research Part C: Emerging Technologies*, 58, 1–12.
- Rose, P. K., Nugroho, R., Gnann, T., Plötz, P., Wietschel, M., & Reuter-Oppermann, M. (2020). Optimal development of alternative fuel station networks considering node capacity restrictions. *Transportation Research Part D: Transport and Environment*, 78, 102189.

- Royset, J. O., & Wets, R. J.-B. (2017). Variational theory for optimization under stochastic ambiguity. *SIAM Journal on Optimization*, 27(2), 1118–1149.
- Shao, H., Lam, W. H., & Tam, M. L. (2006). A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand. *Networks and Spatial Economics*, 6(3), 173–204.
- Simchi-Levi, D., & Berman, O. (1988). A heuristic algorithm for the traveling salesman location problem on networks. *Operations Research*, 36(3), 478–484.
- Solak, S., Clarke, J.-P. B., Johnson, E. L., & Barnes, E. R. (2010). Optimization of R&D project portfolios under endogenous uncertainty. *European Journal of Operational Research*, 207(1), 420–433.
- Spacey, S. A., Wiesemann, W., Kuhn, D., & Luk, W. (2012). Robust software partitioning with multiple instantiation. *INFORMS Journal on Computing*, 24(3), 500–515.
- STATBEL. (2020). *Population density by municipality*. Retrieved from <https://statbel.fgov.be/en/themes/population/population-density>. Accessed February 01, 2021.
- Tarhan, B., Grossmann, I. E., & Goel, V. (2013). Computational strategies for non-convex multistage MINLP models with decision-dependent uncertainty and gradual uncertainty resolution. *Annals of Operations Research*, 203(1), 141–166.
- Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). The location of emergency service facilities. *Operations Research*, 19(6), 1363–1373.
- Tran, T. H., Nagy, G., Nguyen, T. B. T., & Wassan, N. A. (2018). An efficient heuristic algorithm for the alternative-fuel station location problem. *European Journal of Operational Research*, 269(1), 159–170.
- Upchurch, C., & Kuby, M. (2010). Comparing the p-median and flow-refueling models for locating alternative-fuel stations. *Journal of Transport Geography*, 18(6), 750–758.
- Upchurch, C., Kuby, M., & Lim, S. (2009). A model for location of capacitated alternative-fuel stations. *Geographical Analysis*, 41(1), 85–106.
- Vayanos, P., Kuhn, D., & Rustem, B. (2011). Decision rules for information discovery in multi-stage stochastic programming. In *2011 50th IEEE conference on decision and control and european control conference* (pp. 7368–7373).
- Ventura, J. A., Kweon, S. J., Hwang, S. W., Tormay, M., & Li, C. (2017). Energy policy considerations in the design of an alternative-fuel refueling infrastructure to reduce GHG emissions on a transportation network. *Energy Policy*, 111, 427–439.
- Vujanic, R., Goulart, P., & Morari, M. (2016). Robust optimization of schedules affected by uncertain events. *Journal of Optimization Theory and Applications*, 171(3), 1033–1054.
- Wang, Y.-W., & Lin, C.-C. (2009). Locating road-vehicle refueling stations. *Transportation Research Part E: Logistics and Transportation Review*, 45(5), 821–829.

- Wang, Y.-W., & Wang, C.-R. (2010). Locating passenger vehicle refueling stations. *Transportation Research Part E: Logistics and Transportation Review*, 46(5), 791–801.
- 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: Transport and Environment*, 53, 354–376.
- Xie, F., Liu, C., Li, S., Lin, Z., & Huang, Y. (2018). Long-term strategic planning of inter-city fast charging infrastructure for battery electric vehicles. *Transportation Research Part E: Logistics and Transportation Review*, 109, 261–276.
- Yıldız, B., Arslan, O., & Karışan, O. E. (2016). A branch and price approach for routing and refueling station location model. *European Journal of Operational Research*, 248(3), 815–826.
- Yıldız, B., Olcaytu, E., & Şen, A. (2019). The urban recharging infrastructure design problem with stochastic demands and capacitated charging stations. *Transportation Research Part B: Methodological*, 119, 22–44.
- Yu, X., & Shen, S. (2020). Multistage distributionally robust mixed-integer programming with decision-dependent moment-based ambiguity sets. *Mathematical Programming*, 1–40.
- Zhang, A., Kang, J. E., & Kwon, C. (2017). Incorporating demand dynamics in multi-period capacitated fast-charging location planning for electric vehicles. *Transportation Research Part B: Methodological*, 103, 5–29.