Robust Uncapacitated Hub Location

Carlos Armando Zetina[a], Ivan Contreras[a], Jean-François Cordeau[b,*], Ehsan Nikbakhsh[c]

[a]Concordia University and CIRRELT, Montréal, Canada H3G 1M8
[b]HEC Montréal and CIRRELT, Montréal, Canada H3T 2A7
[c]Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran

Abstract

In this paper we present robust counterparts for uncapacitated hub location problems in which the level of conservatism can be controlled by means of a budget of uncertainty. We study three particular cases for which the parameters are subject to interval uncertainty: demand, transportation cost, and both simultaneously. We present mixed integer programming formulations for each of these cases and a branch-and-cut algorithm to solve the latter. We present computational results to evaluate the performance of the proposed formulations when solved with a general purpose solver and study the structure of the solutions to each of the robust counterparts. We also compare the performance between solutions obtained from a commensurable stochastic model and those from our robust counterparts in both risk neutral and worst-case settings.

Keywords: hub location, robust discrete optimization, budget of uncertainty

1. Introduction

In many transportation, telecommunications and computer networks, direct routing of commodities between a large number of origin-destination (O/D) pairs is not possible due to economic and technological constraints. Instead, hub-and-spoke networks are commonly used to connect O/D pairs so as to efficiently route flows by using a small number of links. The key feature of these networks is the use of

*Corresponding author.
Email addresses: c_zetina@encs.concordia.ca (Carlos Armando Zetina), icontrer@encs.concordia.ca (Ivan Contreras), jean-francois.cordeau@hec.ca (Jean-François Cordeau), nikbakhsh@modares.ac.ir (Ehsan Nikbakhsh)

centralized units, known as hub facilities, for consolidation, sorting and transshipment purposes. *Hub location problems* (HLPs) consider the design of hub-and-spoke networks by locating a set of hub facilities, activating a set of inter-hub links, and routing a predetermined set of commodities through the network while optimizing a cost-based (or service-based) objective.

Applications of HLPs in the design of transportation and distribution systems are abundant. These include air freight and passenger travel, postal delivery, express package delivery, trucking, and rapid transit systems. Since the seminal work of O'Kelly (1986), hub location has evolved into a rich research area. Early works focused mostly on a first generation of HLPs which are analogue to fundamental discrete facility location problems, while considering a set of assumptions (hubs fully interconnected, no direct connections, constant discount factor, all commodities must be routed, etc.) that allow to simplify network design decisions (see, Campbell and O'Kelly, 2012; Contreras, 2015, for a discussion). Recent works have studied more complex models that relax some of these assumptions and incorporate additional features of real-life applications. For instance, *hub arc location problems* (HALPs) (Campbell et al., 2005) extend HLPs by relaxing full interconnection of hub nodes and incorporating hub arc selection decisions. *Hub network design problems with profits* (Alibeyg et al., 2016) further extend HALPs by integrating within the decision-making process additional network design decisions on the nodes and commodities that have to be served. Other models consider: the design of multimodal networks (Alumur et al., 2012a; Serper and Alumur, 2016), competition and collaboration (O'Kelly et al., 2015b; Mahmutogullari and Kara, 2016), capacitated networks (Correia et al., 2010; Contreras et al., 2011b), flow dependent discounted costs (O'Kelly et al., 2015a; Tanash et al., 2016), and the design of particular network topologies (Martins de Sá et al., 2015; Contreras et al., 2016), among other things. We refer the reader to Campbell and O'Kelly (2012), Zanjirani Farahani et al. (2013), and Contreras (2015) for recent surveys on hub location.

In most HLPs considered in the literature, the input parameters are assumed to be known and deterministic. In practice, however, this assumption is unrealistic. Long-term strategic decisions such as the location of hub facilities have to be made under high uncertainty on future conditions for relevant parameters (i.e., costs, demands and distances) that have a direct impact on the performance of hub networks. In some cases, some probabilistic information is known for these parameters and can be used to minimize the total expected cost by using stochastic programming techniques. However, in other cases, no information about their probability distributions is known except for the specification of intervals containing the uncertain values and thus, one must rely on robust optimization techniques to design hub networks which are robust in the sense that they can perform well even in the worst-case scenarios that may arise.

In this paper we show how discrete robust optimization techniques can be used in hub location to incorporate both independently and jointly demand and transportation costs uncertainties when the only available information is an interval of uncertainty. In particular, we study several robust counterparts for one of the most fundamental problems in hub location research, the *uncapacitated hub location problem with multiple assignments* (UHLP). In this problem, a predetermined set of commodities has to be routed via a set of hubs. It is assumed that open hubs are fully interconnected with more effective pathways, which allow a flow-independent discount factor to be applied to the transportation costs between hub nodes. The number of hubs to locate is not known in advance, but a setup cost for each hub facility is considered. It is also assumed that the incoming and outgoing flows at hubs as well as the flow routed through each link of the network are unbounded. Commodities having the same origin but different destinations can be routed through different sets of hubs, i.e. a multiple assignment strategy is allowed. Demand between O/D pairs and transportation costs are assumed to be known and deterministic. The objective is to minimize the sum of the hub setup costs and of demand transportation costs over the solution network. To the best of our knowledge, the most efficient formulations for the UHLP are those of Hamacher et al. (2004), Marín et al. (2006), and Contreras and Fernández (2014), whereas the best exact algorithm is the Benders decomposition of Contreras et al. (2011a).

The main contributions of this paper are the following. We introduce three different robust counterparts of the UHLP. The first is the *robust uncapacitated hub location problem with uncertain demands* (UHLP-D) in which demands between O/D pairs are considered to be uncertain values lying in a known interval. The second is the *robust uncapacitated hub location problem with uncertain transportation costs* (UHLP-TC) in which the transportation costs for all links of the network are uncertain values lying in a known interval and independent for each link. The third is the *robust uncapacitated hub location problem with uncertain demands and transportation costs* (UHLP-DTC) where uncertainty exists in both demands between O/D pairs and transportation costs for all links. This problem considers that the uncertainties of both classes of parameters are independent from each other. In these robust counterparts of the UHLP, the objective is to minimize the sum of the hub setup costs and of demand transportation costs in the worst-case scenario that may arise for the uncertain parameters. In the spirit of Bertsimas and Sim (2003), we use a budget of uncertainty to allow decision-makers to control the desired level of conservatism in an independent way for both demand and transportation costs.

For each of the proposed robust models, we present mathematical programming formulations which are non-linear due to the min-max nature of the objective functions. For the case of UHLP-D and UHLP-TC, we use a dual transformation to reformulate them as compact *mixed integer linear programs* (MIP) having a polynomial number of variables and constraints. However, for the case of UHLP-DTC this transformation cannot be used due to the interaction of demand and transportation costs parameters in the objective function. We show how the UHLP-DTC can be modeled as an MIP with a polynomial number of variables but an exponential number of constraints. As a result, we develop a simple branch-and-cut algorithm to handle this formulation. We perform extensive computational experiments on several sets of benchmark instances to assess the computational performance of the proposed MIP formulations when solved with a general purpose solver and our branch-and-cut algorithm. We study the effects of the intervals of uncertainty and of the budgets of uncertainty in the structure of optimal solution networks. In addition, we compare the performance between solution networks obtained from a deterministic model, a commensurable stochastic model and those from our robust counterparts in both risk neutral and worst-case settings.

The remainder of the paper is organized as follows. Section 2 provides a literature review on hub location problems dealing with uncertainty. In Section 3 we introduce the three considered robust counterparts of the UHLP. Section 4 describes the computational experiments we have run. The results produced by each model are presented and analyzed. Conclusions follow in Section 5.

2. Literature Review

Discrete facility location problems under uncertainty have been widely studied in the literature. Louveaux (1993), Snyder (2006), and more recently Correia and da Gama (2015) provide comprehensive reviews on modeling approaches for stochastic and robust facility location. However, much less work has been done to study how different uncertainty aspects can be taken into account when designing hub networks. One of the first is the paper by Marianov and Serra (2003) for an application in the airline industry. The authors model hubs as M/D/c queuing systems to derive an analytic expression for the probability of a number of customers in the system. This expression is then represented in the model as a constraint that limits the number of airplanes in the system. Sim et al. (2009) consider the p-hub center problem with travel times following a normal distribution. The authors use chance constraints to incorporate service level guarantees. Their model takes into account the uncertainty in travel times when designing the network so that the maximum travel time through the network is minimized.

Yang (2009) proposes a two-stage stochastic model for air freight hub location under a finite set of possible demands. The locational decisions are treated in the first stage while the routing is the second stage. Data from the air freight market in Taiwan and China is used to test the proposed model. Contreras et al. (2011b) show how the UHLP can be stated as a two-stage integer stochastic program with recourse in the presence of uncertainty in demands and transportation costs. Three different variants are introduced. The first assumes demand between each O/D pair to be stochastic. The second considers that uncertainty in transportation costs is given by a single parameter equally influencing all links of the network. The third focuses on the more general case in which the uncertainty of transportation costs is independent for each link of the network. It is shown that the first two variants are equivalent to their associated expected value problem in which uncertain demand and transportation costs are replaced by their expectation. However, this situation does not hold for the third case. The authors present a sample average approximation method based on Monte Carlo simulation to obtain an estimate on the stochastic solution. Alumur et al. (2012b) study the uncapacitated hub location problem both with single and multiple assignments under uncertain setup costs for the hubs and demands. The first class of models deals with uncertainty in the setup costs assuming there is no known probability distribution for there random parameters. The authors propose the use of a min-max regret model where the objective deals with the minimization of the worst-case regret over a finite set of scenarios. The second class focuses on uncertainty in demand and uses a two-stage stochastic program with recourse. However, as shown in Contreras et al. (2011b), these problems are actually equivalent to their associated expected value problem. The third class considers uncertainty in both setup costs and demands and are models as two-stage min-max regret programs with recourse.

Ghaffari-Nasab et al. (2015) present robust capacitated hub location problems with both multiple and single assignments in which demand uncertainty is modeled with an interval of uncertainty. The authors consider the uncertainty in the capacity constraints and use a budget of uncertainty for each of them. However, they do not consider demand uncertainty in the objective function and use the nominal demand value instead. Habibzadeh Boukani et al. (2016) study the same capacitated hub location problems with multiple and single assignments but now the uncertainties relate to the setup costs and capacities. The authors present min-max regret models involving five scenarios for each uncertain parameter. Shahabi and Unnikrishnan (2014) present robust counterparts for uncapacitated hub location problems with both multiple and single assignments with demand uncertainty. Demand is assumed to be affinely dependent on a known mean and a number of independent random variables, i.e. ellipsoidal demand uncertainty. The authors propose mixed integer conic quadratic programming formulations and a linear relaxation strategy. Given the inherent difficulty for solving this class of mathematical programs, instances with up to 25 nodes are solved with a general purpose solver. Merakli and Yaman (2016) study robust uncapacitated p-hub median problems with multiple assignments under polyhedral demand uncertainty. The authors use a hose uncertainty model and a hybrid model to characterize demand uncertainty. The former considers that the only available information is an upper limit on the total flow adjacent to each node, while the latter incorporates lower and upper bounds on each O/D flow. The authors present MIP formulations and a Benders algorithm to solve these problems for instances with up to 150 nodes. We would like to highlight that, compared to our proposed models, none of these papers dealing with robust optimization in hub location considers uncertainty in transportation costs and its interaction with demand uncertainty.

Demand uncertainty has also been studied when designing hub networks from a congestion perspective. Elhedhli and Hu (2005) study a hub location model that considers hub congestion costs, caused by delays at hub facilities, as an exponential function of the hub flow. Elhedhli and Wu (2010) propose a different approach in which hubs behave as a queue with exponential service rate determined by its capacity. The congestion cost is modeled using a Kleinrock average delay function. de Camargo and Miranda (2012) extend the previous models by considering two different perspectives: a network owner perspective in which the the goal is to design a network with the least congestion cost, and a user perspective where the goal is the minimization of the maximum congestion effect. Other works have considered reliability issues. An et al. (2015) and Azizi et al. (2016) present models in which disruptions at hub facilities are taken into account when designing hub networks. The proposed models mitigate the resulting hub unavailability by using backup hubs and alternative routes for demands. Finally, Sun and Zheng (2016) study a probability model devised to identify promising hub locations for liner shipping networks.

3. Robust Uncapacitated Hub Location Problems

Before presenting the considered robust models, we formally define their deterministic counterpart, the UHLP. Let $G = (N, A)$ be a complete directed graph where N is the set of nodes and A the set of arcs. Without loss of generality, let N represent the set of potential hub locations, and K the set of commodities each with an origin, destination in N and demand, denoted by the triplet $(o(k), d(k), W_k)$. For each node $i \in N$, f_i is the fixed setup cost for locating a hub. For each $(i, j) \in A$

the distances, or transportation costs $d_{ij} \geq 0$ are assumed to be symmetric and satisfy the triangle inequality. The UHLP consists of locating a set of hub facilities and of determining the routing of commodities demand through the hubs, with the objective of minimizing the total setup and transportation cost.

In the case of UHLP, given that open hubs are fully interconnected at no cost, distances satisfy the triangle inequality, and direct connections between non-hub nodes are not allowed, every O/D path will contain at least one and at most two hubs, i.e. $P_{ak} = (o(k), a_1, a_2, d(k))$, where $a = (a_1, a_2) \in A$ is a *hub arc* and a_1, a_2 is the ordered pair of hubs to which $o(k)$ and $d(k)$ are assigned, respectively. The unit transportation cost for routing commodity k along path P_{ak} is given by $C_{ak} = \chi d_{d(k)a_1} + \alpha d_{a_1 a_2} + \delta d_{a_2 d(k)}$, where χ, α and δ represent the flow-independent collection, transfer and distribution costs along the path. To reflect economies of scale between hub nodes, we assume that $\alpha < \chi$ and $\alpha < \delta$.

For each $i \in N$, we define binary location variables z_i equal to 1 if and only if a hub is located at i. For each $k \in K$ and $a \in A$, we also introduce continuous routing variables x_{ak} equal to the fraction of commodity demand W_k routed via first hub a_1 and second hub a_2. The UHLP can be stated as follows (Hamacher et al., 2004):

$$\text{minimize} \quad \sum_{i \in N} f_i z_i + \sum_{k \in K} \sum_{a \in A} W_k C_{ak} x_{ak} \tag{1}$$

$$\text{subject to} \quad \sum_{a \in A} x_{ak} = 1 \qquad \forall k \in K \tag{2}$$

$$\sum_{a \in A: i \in a} x_{ak} \leq z_i \qquad \forall k \in K, i \in N \tag{3}$$

$$z_i \in \{0, 1\} \qquad \forall i \in N \tag{4}$$

$$x_{ak} \geq 0 \qquad \forall a \in A, k \in K. \tag{5}$$

The first term of the objective is the total setup cost of the hubs and the second term is the total transportation cost. Constraints (2) ensure that each commodity demand is fully routed. Constraints (3) prohibit commodities from being routed via a non-hub node, whereas (4)-(5) are the standard integrality and non-negativity constraints. This formulation contains $|N| + |K||A|$ variables and $|K| + |N||K|$ constraints. Given that there are no capacity constraints on the hubs or links of the network, there always exist an optimal solution of (1)-(5) in which all x_{ak} are integer. That is, each k is fully routed on a single path. As we will later see in Section 4, this property does not necessarily hold for some of the studied robust counterparts.

We next present three robust variants of UHLP that incorporate uncertainty on demands W_k and transportation costs C_{ak}. We note that an interesting characteristic of formulation (1)-(5) is that the uncertain parameters only appear in the objective function and not in the constraints defining the feasible region. As a consequence, when using (1)-(5) as a basis to model the robust counterparts the uncertain parameters will not affect the feasibility of the problem. This is not necessarily true when using other existing MIP formulations for the UHLP as the demand and transportation costs do appear in the constraint matrix (Contreras, 2015).

In a similar fashion to the approach used in Bertsimas and Sim (2003), we use a budget of uncertainty to allow decision-makers to control the desired level of conservatism. However, instead of defining a budget for the maximum number of variables for which the objective coefficient is allowed to differ from its nominal value, as in Bertsimas and Sim (2003), we use two budgets of uncertainty to control the maximum number of commodity demands and transportation costs, respectively, whose value is allowed to differ from its nominal value. This is an important modeling feature due to the fact that objective coefficients of x_{ak} variables are defined as functions of several uncertain parameters, i.e. $W_k C_{ak} = W_k(\chi d_{d(k)a_1} + \alpha d_{a_1 a_2} + \delta d_{a_2 d(k)})$. As we will show, this makes the proper modeling of the robust counterparts more involved, as compared to the approach of Bertsimas and Sim (2003), given that each uncertain parameter appears in several objective coefficients associated with different routing variables X_{ak}.

3.1. Case A: Uncertain Demands

We consider the UHLP-D in which demand is subject to interval uncertainty. For each commodity $k \in K$, let $W_k \in [W_k^L, W_k^L + W_k^\Delta]$ be the interval of uncertainty of W_k where W_k^L is its nominal value and $W_k^\Delta \geq 0$ its deviation. Let h_W denote the uncertainty budget on the maximum number of demand parameters W_k whose value is allowed to differ from its nominal value. The UHLP-D can be stated as:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in N} f_i z_i + \sum_{k \in K} \sum_{a \in A} W_k^L C_{ak} x_{ak} + \max_{S_W \subseteq K : |S_W| \leq h_W} \{ \sum_{k \in S_W} \sum_{a \in A} W_k^\Delta C_{ak} x_{ak} \} \\ \text{subject to} \quad & (x, z) \in \Theta, \end{aligned}$$

where $\Theta = \{(x, z) : (2) - (5) \text{ are satisfied}\}$. Given a solution $(x, z) \in \Theta$, the goal of the inner maximization of the objective function is to select a subset of commodities $S_W \subseteq K$ such that their demand perturbations W_k^Δ maximize the total cost. Reformulating this inner problem as a mathematical program by introducing the binary variable $u_k \in \{0, 1\}$ for each $k \in K$, we obtain

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i \in N} f_i z_i + \sum_{k \in K} \sum_{a \in A} W_k^L C_{ak} x_{ak} + \max\{\sum_{k \in K} \sum_{a \in A} W_k^\Delta C_{ak} x_{ak} u_k\} \\
& \sum_{k \in K} u_k \leq h_W && (6) \\
\text{subject to} \quad & 0 \leq u_k \leq 1 \qquad \forall k \in K && (7) \\
& (x, z) \in \Theta.
\end{aligned}
$$

Note that the integrality conditions of u_k variables have been relaxed given that the constraint matrix of (6) is totally unimodular. Let (μ, λ) denote the vector of dual multipliers of constraints (6)-(7) of appropriate dimension. Dualizing the inner maximization problem we obtain the following MIP formulation for the UHLP-D:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i \in N} f_i z_i + \sum_{k \in K} \sum_{a \in A} W_k^L C_{ak} x_{ak} + \sum_{k \in K} \lambda_k + h_W \mu \\
\text{subject to} \quad & \lambda_k + \mu \geq \sum_{a \in A} C_{ak} W_k^\Delta x_{ak} \qquad \forall k \in K && (8) \\
& \lambda_k, \mu \geq 0 \qquad \forall k \in K \\
& (x, z) \in \Theta.
\end{aligned}
$$

The above formulation contains $|K|+1$ additional continuous variables and $2|K|+1$ additional constraints compared to the deterministic UHLP.

3.2. Case B: Uncertain Transportation Costs

We now consider the UHLP-TC in which transportation costs are subject to interval uncertainty. For each arc $(i, j) \in A$, let $d_{ij} \in [d_{ij}^L, d_{ij}^L + d_{ij}^\Delta]$ be the interval of uncertainty of d_{ij}, where d_{ij}^L is its nominal value and $d_{ij}^\Delta \geq 0$ its deviation. Observe that now each coefficient of the routing variables x_{ak} involves up to three uncertain parameters, i.e. $C_{ak} = \chi d_{d(k)a_1} + \alpha d_{a_1 a_2} + \delta d_{a_2 d(k)}$. To simplify the exposition of the model, the transportation cost associated to each path P_{ak} is stated as

$$
C_{ak} = \sum_{(i,j) \in P_{ak}} \kappa_{ak}^{ij} d_{ij} = \sum_{(i,j) \in A} \kappa_{ak}^{ij} d_{ij},
$$

where

$$
\kappa_{ak}^{ij} = \begin{cases} \chi & \text{if } (i, j) = (o(k), a_1) \in P_{ak}, \\ \alpha & \text{if } (i, j) = (a_1, a_2) \in P_{ak}, \\ \delta & \text{if } (i, j) = (a_2, d(k)) \in P_{ak}, \\ 0 & \text{otherwise,} \end{cases}
$$

for each $a \in A$, $k \in K$ and $(i,j) \in A$. In addition, for each $a \in A$ and $k \in K$, we define the nominal transportation cost as $C^L_{ak} = \sum_{(i,j)\in P_{ak}} \kappa^{ij}_{ak} d^L_{ij}$. Let h_d denote the uncertainty budget of the maximum number of transportation cost coefficients whose values are allowed to differ from their nominal value. The UHLP-TC can be stated as:

$$\text{minimize} \sum_{i\in N} f_i z_i + \sum_{k\in K}\sum_{a\in A} W_k C^L_{ak} x_{ak} + \max_{S_d\subseteq A:|S_d|\le h_d} \{\sum_{k\in K}\sum_{a\in A}\sum_{(i,j)\in P_{ak}\cap S_d} d^\Delta_{ij}\kappa^{ij}_{ak} W_k x_{ak}\}$$

subject to $(x,z) \in \Theta$.

Given a solution $(x,z) \in \Theta$, the goal of the inner maximization is to select the subset of arcs $S_d \subseteq A$ whose transportation costs perturbations d^Δ_{ij} maximize the total cost. Reformulating this inner problem as a mathematical program by introducing the binary variable $w_{ij} \in \{0,1\}$ for each $(i,j) \in A$, we obtain

$$\begin{aligned}
\text{minimize} \quad & \sum_{i\in N} f_i z_i + \sum_{k\in K}\sum_{a\in A} W_k C^L_{ak} x_{ak} + \max\{\sum_{k\in K}\sum_{a\in A}\sum_{(i,j)\in P_{ak}\cap S_d} d^\Delta_{ij}\kappa^{ij}_{ak} W_k x_{ak} w_{ij}\} \\
\text{subject to} \quad & \sum_{(i,j)\in A} w_{ij} \le h_d && (9) \\
& 0 \le w_{ij} \le 1 \qquad \forall (i,j) \in A && (10) \\
& (x,z) \in \Theta.
\end{aligned}$$

Similarly to UHLP-D, the integrality conditions of w_{ij} variables can be relaxed given that the constraint matrix of (9) is totally unimodular. Let (μ, λ) denote the vector of dual multipliers of constraints (9)-(10) of appropriate dimension. Dualizing the inner maximization problem we obtain the following MIP formulation for the UHLP-TC:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i\in N} f_i z_i + \sum_{k\in K}\sum_{a\in A} W_k C^L_{ak} x_{ak} + \sum_{(i,j)\in A} \lambda_{ij} + h_d \mu \\
\text{subject to} \quad & \lambda_{ij} + \mu \ge \sum_{k\in K}\sum_{a\in A:(i,j)\in P_{ak}} d^\Delta_{ij}\kappa^{ij}_{ak} W_k x_{ak} \qquad \forall (i,j) \in A && (11) \\
& \lambda_{ij}, \mu \ge 0 \qquad \forall (i,j) \in A \\
& (x,z) \in \Theta.
\end{aligned}$$

The above formulation contains $|A|+1$ additional continuous variables and $2|A|+1$ additional constraints compared to the deterministic UHLP.

3.3. Case C: Uncertain Demand and Transportation Costs

We now focus on the more general variant UHLP-DTC in which both demand and transportation costs are subject to interval uncertainty. Similar to previous models, let $W_k \in [W_k^L, W_k^L + W_k^\Delta]$ and $d_{ij} \in [d_{ij}^L, d_{ij}^L + d_{ij}^\Delta]$ denote the interval of uncertainty for commodity demands and transportation costs, respectively, and h_d and h_W their respective uncertainty budgets. To simplify the exposition fo the proposed model, we define

$$F(z,x) = \sum_{i \in N} f_i z_i + \sum_{k \in K} \sum_{a \in A} W_k^L C_{ak}^L x_{ak},$$

as the nominal cost function, i.e. set-up cost and nominal routing cost, and

$$\begin{aligned} R(z,x,S_W,S_d) \quad = \quad & \sum_{k \in S_W} \sum_{a \in A} W_k^\Delta C_{ak}^L x_{ak} + \sum_{k \in K} \sum_{a \in A} \sum_{(i,j) \in P_{ak} \cap S_d} \kappa_{ak}^{ij} W_k^L d_{ij}^\Delta x_{ak} \\ & + \sum_{k \in S_W} \sum_{a \in A} \sum_{(i,j) \in P_{ak} \cap S_d} \kappa_{ak}^{ij} W_k^\Delta d_{ij}^\Delta x_{ak}, \end{aligned}$$

the uncertain routing cost function. Given that the coefficients of the routing variables x_{ak} contain the multiplication of the demand W_k and some transportation costs d_{ij}, three different scenarios may arise and those correspond to each term of $R(z,x,S_W,S_d)$. The first term represents an additional cost caused by an increase only in demands W_k, whereas the second term relates to an additional cost due to an increase only in transportation costs d_{ij}. The third term corresponds to the situation in which there is an additional cost caused by a simultaneous increase of both demand and transportation costs. The UHLP-CDT can be stated as:

$$\begin{aligned} \text{minimize} \quad & F(z,x) + \max_{S_d \subseteq A, |S_d| \leq h_d, S_W \subseteq K, |S_W| \leq h_W} R(z,x,S_W,S_d) \\ \text{subject to} \quad & (x,z) \in \Theta. \end{aligned}$$

For a given solution $(\bar{x}, \bar{z}) \in \Theta$, the goal of the inner maximization of the objective function is to select a subset of commodities $S_W \subseteq K$ and a subset of arcs $S_d \subseteq A$, such that their perturbations maximize the total cost. This inner program, denoted as SP, can be reformulated as the following mathematical program:

$$\begin{aligned} (SP) \quad \text{maximize} \quad & \sum_{k \in K} \sum_{a \in A} W_k^\Delta C_{ak}^L \bar{x}_{ak} u_k + \sum_{k \in K} \sum_{a \in A} \sum_{(i,j) \in A_{ak}} \kappa_{ak}^{ij} W_k^L d_{ij}^\Delta \bar{x}_{ak} w_{ij} \\ & + \sum_{k \in K} \sum_{a \in A} \sum_{(i,j) \in A_{ak}} \kappa_{ak}^{ij} W_k^\Delta d_{ij}^\Delta \bar{x}_{ak} u_k w_{ij} \\ \text{subject to} \quad & \sum_{k \in K} u_k \leq h_k \end{aligned}$$

$$\sum_{(i,j)\in A} w_{ij} \le h_d$$
$$u_k \in \{0,1\} \qquad \forall k \in K,$$
$$w_{ij} \in \{0,1\} \qquad \forall (i,j) \in A.$$

SP is a nonlinear integer program which can be linearized using standard techniques. However, it can also be stated as the following combinatorial optimization problem. Given a bipartite graph $B = (K \cup A, E)$ where each node in K and A has a profit a_k and b_{ij}, respectively, and each edge in E has a profit c_{ijk}, select node subsets $S_1 \subseteq K$ and $S_2 \subseteq A$ whose cardinality do not exceed $|S_1| \le h_W$ and $|S_2| \le h_d$, so as to maximize the total profit, given by the sum of the profits of the nodes in S_1 and S_2 and of the edges $e \in S_1 \times S_2 \subseteq E$. SP is thus a generalization of the *heaviest k-subgraph problem* (Bruglieri et al., 2006).

Proposition 1. *The SP problem is $\mathcal{NP}$-hard.*

Proof The result follows by reduction from the heaviest k-subgraph problem, which is known to be $\mathcal{NP}$-complete even for the case of bipartite graphs (see Corneil and Perl, 1984). ■

The above result implies that we cannot use a dual transformation as in UHLP-D and RHLP-TC to obtain a MIP formulation for UHLP-DTC. However, we can still reformulate it as an MIP by introducing an additional continuous variable y that keeps track of the maximum transportation cost associated with different sets S_W and S_d. The UHLP-DTC can thus be formulated as the following MIP:

$$\begin{aligned} \text{minimize} \quad & F(z,x) + y \\ \text{subject to} \quad & y \ge R(z,x,S_W,S_d) \quad \forall S_d \subseteq A, \forall S_W \subseteq K, |S_d| \le h_d, |S_W| \le h_W && (12) \\ & y \ge 0 && (13) \\ & (x,z) \in \Theta. \end{aligned}$$

Observe that the above formulation contains a polynomial number of variables - it only contains one extra continuous variable with respect to the deterministic UHLP. However, it requires an exponential number of constraints (12) whose separation problem is equivalent to the solution of SP, which is $\mathcal{NP}$-hard. Therefore, we resort to a cutting-plane algorithm to handle this formulation with a general purpose solver. Details are given in the next section.

4. Computational Experiments

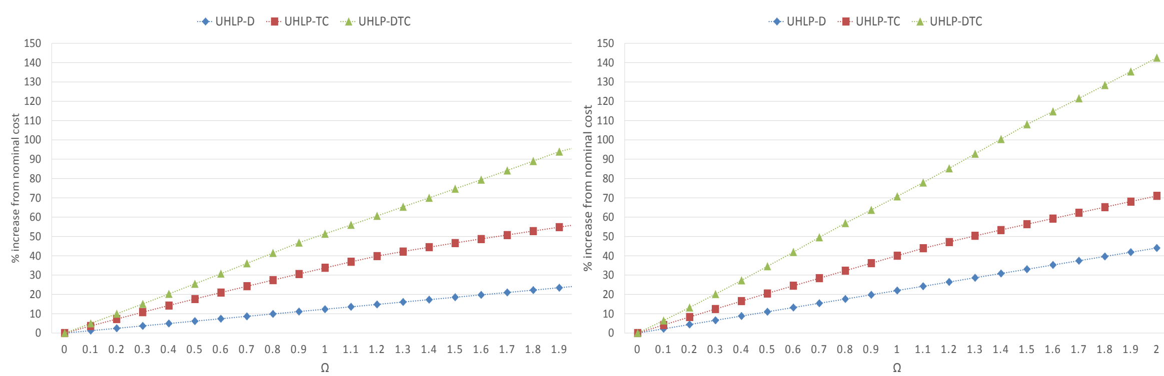
In this section we describe the extensive computational experiments we have run in order to analyze the performance of the three studied robust counterparts of the UHLP. This section is structured as follows. We first describe the computational environment and the sets of benchmark instances we have used in our experiments. In Section 4.1 we give insights on the impact of the level of uncertainty (worst estimation error of the nominal value) in optimal solutions of the robust counterparts, whereas in Section 4.2 we study the effect of the budgets of uncertainty h_d and h_W on the solution networks and objective values. Section 4.3 provides numerical results to assess the usefulness of the proposed MIP formulations when solved with a general purpose solver and to compare the computational difficulty of the three robust counterparts. Finally, in Section 4.4 we compare the performance of optimal solution networks of the deterministic model, the robust counterparts and a commensurable stochastic model in both worst-case and risk-neutral settings. This provides some indication of the virtues and limitations of each approach when dealing with demand and transportation costs uncertainty.

All experiments were run on HP servers managed by Calcul Québec and Compute Canada with an Intel Xeon X5650 Westmere processor at 2.67 GHz and 24 GB of RAM under Linux environment. All MIP formulations were coded in C and solved using the callback library of CPLEX 12.6.0. We use a traditional (deterministic) branch-and-bound solution algorithm with all CPLEX parameters set to their default values with the exception that the number of threads was set to only one.

We have used three different sets of benchmark instances found in the literature to construct a testbed for our experiments. The first one is the well-known CAB data set of the US Civil Aeronautics Board. These instances were obtained from the website *(http://www.researchgate.net/publication/269396247_cab100_mok)*. The data in the CAB set refers to 100 cities in the US. It provides Euclidean distances between cities, d_{ij}, and the values of the service demand W_k between each pair of cities, where $o(k) \neq d(k)$. Since CAB instances do not provide the setup costs for opening facilities, we use the setup costs f_i generated by de Camargo et al. (2008). We selected the instances with $|N| = 25$. The second set is the classical AP (Australian Post) set of instances, which is the most commonly used in the hub location literature. They were obtained from the OR library *(http://mscmga.ms.ic.ac.uk/jeb/orlib/phubinfo.html)*. Transportation costs are proportional to the Euclidian distances between 200 postal districts in the metro Sydney area and the values W_k represent postal flows between pairs of districts. We selected instances with $|N| = 10$, 20, 25, 40, and 50. Finally, the third set is the *Set I* instances introduced in Contreras et al. (2011b) for the stochastic counterparts of the UHLP. We used all instances within this set with $|N| = 10$, 20, 30, 40, and 50. For each of the considered instances in the entire testbed, the nominal values W_k^L and d_{ij}^L were set to the associated deterministic values provided in each dataset, and the collection and distribution factors $\chi = \delta = 1$. To control the size of the generated intervals of uncertainty for W_K and d_{ij}, we define an additional parameter Ω as the maximum possible deviation of the value of each parameter. For each $k \in K$, we independently set W_k^Δ by considering $W_k^\Delta \sim U\left[0, \Omega W_k^L\right]$, and for each $(i,j) \in A$ we independently set d_{ij}^Δ by considering $d_{ij}^\Delta \sim U\left[0, \Omega d_{ij}^L\right]$.

4.1. The Impact of the Interval of Uncertainty on Optimal Solutions

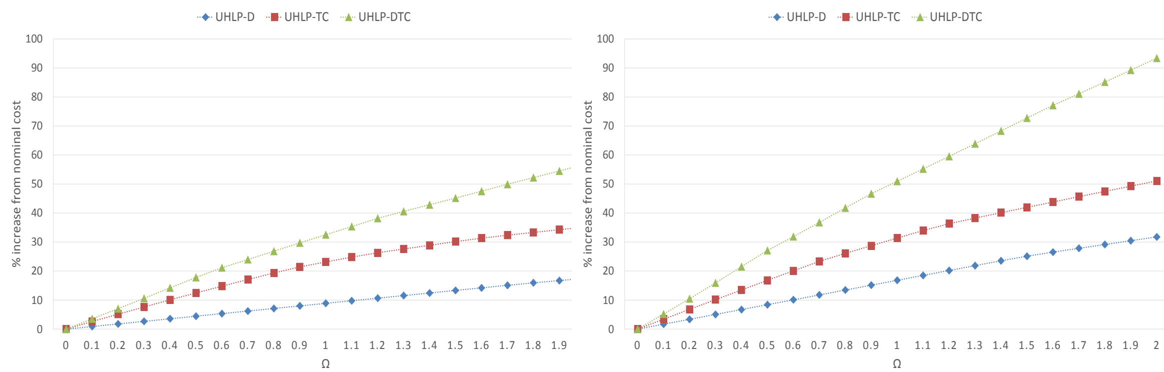
For the first set of experiments we use the CAB 25-node instances and the AP and *Set I* 20-node instances with a transfer discount factor $\alpha = 0.2$. We consider a 5% and 15% uncertainty budget for both the number of commodities and arcs. For each value, we generate instances with $\Omega \in \{0.1, 0.2, 0.3, ..., 2.0\}$ for the three robust counterparts. We note that smaller values of Ω correspond, on average, to narrower intervals whereas larger values to wider intervals. Figures 1, 2 and 3 illustrate for the CAB, AP and *Set I* instances, respectively, the increase in the optimal solution values of each robust counterpart as the intervals of uncertainty increase in size (i.e. Ω increases). In particular, they give the percentage increase of the optimal solution value with the respect to the optimal value of the deterministic (or nominal) UHLP.



(a) 5% budget

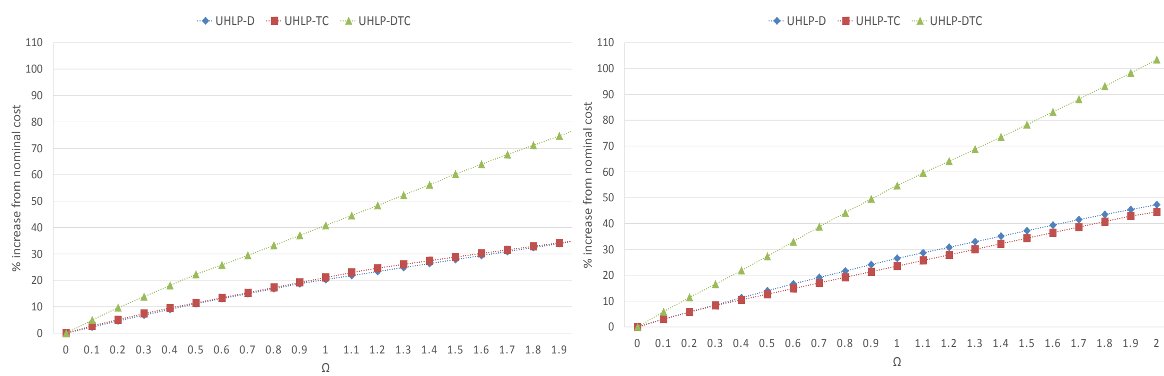
(b) 15% budget

Figure 1: Impact of interval of uncertainty for CAB 25-node instances with $\alpha = 0.2$



(a) 5% budget (b) 15% budget

Figure 2: Impact of interval of uncertainty for AP 20-node instances with $\alpha = 0.2$



(a) 5% budget (b) 15% budget

Figure 3: Impact of interval of uncertainty for *Set I* 20-node instances with $\alpha = 0.2$

For the considered instances, increasing Ω implies a higher optimal value for each of the robust counterparts. This is to be expected as larger values of Ω are associated with the possibility of wider intervals of uncertainty. It is interesting to note how the change in relative cost seems almost constant through the range of considered values of Ω. Moreover, observe that the behaviour of the robust counterparts varies depending on the dataset used. In general, the relative increase of the CAB instances is higher than that of both the AP and *Set I* instances. Of the latter two, the AP instances are the least sensitive to an increase in Ω. Also, the relative increase of the transportation cost and of the demand have very similar behaviour in the

Set I instances while for the other datasets, the robust counterpart of uncertain transportation costs is clearly more sensitive to an increased value of Ω.

Another interesting characteristic is the interaction of the robust counterparts. At a 5% budget, UHLP-D is the least sensitive to the increase of Ω followed by UHLP-TC and the most sensitive being UHLP-DTC. At a 15% budget, the gap in behaviour between UHLP-D and UHLP-TC becomes smaller. In fact for the *Set I* instances, the sensitivity relation is inverted, UHLP-D appears to be slightly more sensitive to Ω than UHLP-TC. This comes from the effect the uncertainty budget has on the robust counterparts. This is further explored in Section 4.2.

Tables 1, 2 and 3 show the open hubs at optimal solutions for varying values of Ω. We note that $\Omega = 0$ actually corresponds to the optimal hub configuration of the deterministic UHLP. Given that the set of open hubs for these counterparts changes for different levels of uncertainty Ω, each row on these tables corresponds to a set of values of Ω for which the optimal solution is the same.

Ω	UHLP-D	UHLP-TC	UHLP-DTC
$[0.0, 0.8]$	13 19 25	13 19 25	13 19 25
$[0.9, 1.2]$	13 19 25	13 19 25	5 13 19 25
$[1.3, 1.9]$	13 19 25	5 13 19 25	5 13 19 25
2.0	13 19 25	5 13 19 25	5 12 13 25

(a) Impact of uncertainty at 5% budget

Ω	UHLP-D	UHLP-TC	UHLP-DTC
$[0.0, 0.7]$	13 19 25	13 19 25	13 19 25
$[0.8, 1.1]$	13 19 25	13 19 25	5 13 19 25
$[1.2, 1.4]$	13 19 25	5 13 19 25	5 13 19 25
$[1.5, 2.0]$	13 19 25	5 13 19 25	5 12 13 25

(b) Impact of uncertainty at 15% budget

Table 1: Open hubs at optimal solutions for CAB 25-node instances

Ω	UHLP-D	UHLP-TC	UHLP-DTC
$[0.0, 0.5]$	7 14	7 14	7 14
$[0.6, 0.8]$	7 14	7 14	1 7 14
$[0.9, 1.1]$	7 14	1 7 14	1 7 14
1.2	7 14	6 8 14	1 7 14
$[1.3, 1.5]$	7 14	6 8 14	1 7 8 14
1.6	7 14	7 8 14	1 7 8 14
1.7	7 14	8 11 14	1 7 8 14
$[1.8, 2.0]$	6 14	8 11 14	1 2 8 11 14

(a) Impact of uncertainty at 5% budget

Ω	UHLP-D	UHLP-TC	UHLP-DTC
$[0.0, 0.4]$	7 14	7 14	7 14
$[0.5, 0.7]$	7 14	7 14	1 7 14
$[0.8, 0.9]$	7 14	1 7 14	2 7 14
$[1.0, 1.1]$	7 14	1 7 14	2 7 8 14
$[1.2, 1.4]$	7 14	1 7 8 14	2 7 8 14
1.5	6 14	1 7 8 14	2 7 8 14
$[1.6, 2.0]$	2 11 14	1 7 8 14	2 7 8 9 14

(b) Impact of uncertainty at 15% budget

Table 2: Open hubs at optimal solutions for AP 20-node instances

Ω	UHLP-D	UHLP-TC	UHLP-DTC
0.0	2 20	2 20	2 20
0.1	2 20	2 4 20	2 4 20
0.2	2 20	2 4 20	2 4 13
0.3	2 4 20	2 4 20	2 4 13
0.4	2 4 20	2 4 13	2 4 13
0.5	2 4 20	2 4 13	2 4 6 11
$[0.6, 0.9]$	2 4 13	2 4 13	2 4 6 11
$[1.0, 1.1]$	2 4 6 11	2 4 13	2 4 6 11
$[1.2, 1.5]$	2 4 6 11	2 4 6 11	2 4 6 11
$[1.6, 1.7]$	2 4 6 11	2 4 6 11	2 4 6 9 11
$[1.8, 2.0]$	2 4 6 11	2 4 6 11	2 4 6 9 11 20

(a) Impact of uncertainty at 5% budget

Ω	UHLP-D	UHLP-TC	UHLP-DTC
0.0	2 20	2 20	2 20
0.1	2 20	2 4 20	2 4 20
0.2	2 4 20	2 4 20	2 4 13
$[0.3, 0.5]$	2 4 20	2 4 13	2 4 13
$[0.6, 0.7]$	2 4 13	2 4 13	2 4 13
$[0.8, 0.9]$	2 4 13	2 4 13	2 4 6 20
1	2 4 6 11	2 4 13	2 4 6 9 20
$[1.1, 1.7]$	2 4 6 11	2 4 13	2 4 6 9 11 20
$[1.8, 1.9]$	2 4 6 11 20	2 4 13	2 4 6 9 11 20
2	2 4 6 11 20	2 4 6 9 20	2 4 6 9 11 20

(b) Impact of uncertainty at 15% budget

Table 3: Open hubs at optimal solutions for *Set I* 20-node instances

The results of Tables 1, 2 and 3 show that in general the optimal hub configuration of UHLP-DTC changes from that of the deterministic solution at a lower level of uncertainty than the other two robust models. For example, the CAB instance with a fixed 15% uncertainty budget changes at $\Omega = 0.8$ while for the *Set I* instance, it changes as early as $\Omega = 0.1$. With respect to the other two robust models, UHLP-TC is more susceptible to changing its optimal hub configuration than UHLP-D as seen in the tables. Another noteworthy characteristic is the difference in behaviour between datasets. Note that the CAB instance has a less varied array of optimal hub configurations for the different values of Ω despite the elevated increase in relative cost observed in Figure 1. However, in practice, this may be partially explained by the existence of very important O/D nodes that contain a high percentage of the total amount of commodity demands being routed on the network, thereby making these nodes promising hub candidates despite a high level of uncertainty.

4.2. Impact of Budget of Uncertainty on Optimal Solutions

Our modeling approach allows decision-makers to select the level of conservatism of the robust counterparts by means of the uncertainty budgets h_d and h_W. For some problems, it is unrealistic to assume that all values of the parameters will change and require protection against this possibility. With this in mind, we study the effect that different uncertainty budgets have on the robust counterparts using the same set of instances as in the previous section with $\Omega = 1.0$ and $\alpha = 0.2$. We test the models at budget values of $h_d \in \{\lfloor 0.05|A| \rfloor, \lfloor 0.1|A| \rfloor, \lfloor 0.15|A| \rfloor, ...|A|\}$ and $h_W \in \{\lfloor 0.05|K| \rfloor, \lfloor 0.1|K| \rfloor, \lfloor 0.15|K| \rfloor, ...|K|\}$. As before, we analyze the effects on UHLP-D, UHLP-TC, and UHLP-DTC with respect to the relative increase of the optimal solution value and the set of open hubs.

Figures 4, 4a and 4b give for the CAB, AP and *Set I* instances, respectively, the increase in the optimal solution values of each robust counterpart as the budget of uncertainty increases. In particular, they give the percentage increase of the optimal solution value with respect to the optimal value of the deterministic UHLP.

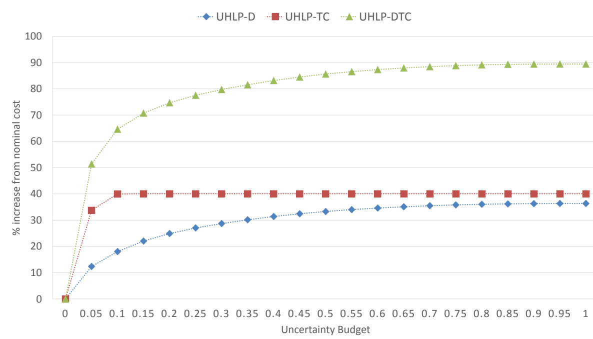
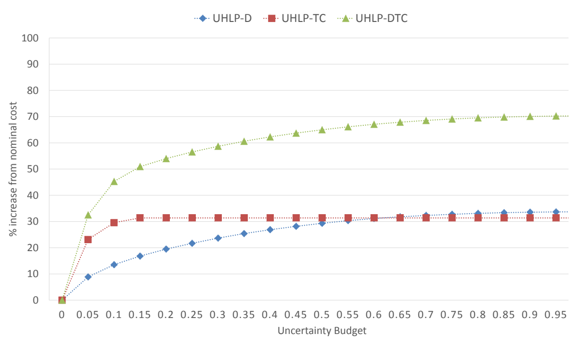
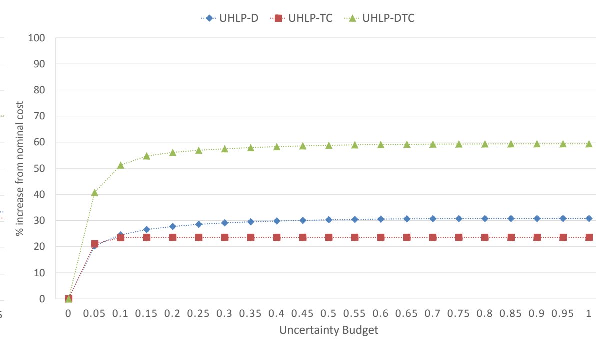


Figure 4: Effect of budget of uncertainty for CAB 25-node instance



(a) AP 20-node instance



(b) *Set I* 20-node instance

Figure 5: Effect of uncertainty budget

We note that while increasing the budget of uncertainty for UHLP-D leads to a greater optimal value for all budgets, for UHLP-TC there is a point after which the effect of increasing the budget becomes null and the objective value remains the same. In fact, we note from the figures that this crucial point tends to be at small values of the budget of uncertainty h_d. This can be partially explained by the fact that hub-and-spoke networks only use a small percentage of the total arcs of the underlying network in an optimal solution, thereby nullifying the additional protection acquired by using a higher budget. When taking into account both demand and transportation cost uncertainty in UHLP-DTC, we note a somehow similar behaviour. As the uncertainty budget becomes larger, the rate of increase in the objective function value decreases, although it does not completely become null before reaching a budget of 100%. Another interesting observation is the interaction between these models. In particular, that between UHLP-D and UHLP-TC. Note that at budgets close to 0, the optimal value of UHLP-TC is greater than that of UHLP-D. However, as the budgets increase past the critical point at which the optimal value of UHLP-TC becomes stable, the optimal value of UHLP-D draws closer to that of UHLP-TC as seen in Figure 4 for the CAB instance. In the AP and *Set I* instances, Figures 4a and 4b show that the optimal value of UHLP-D surpasses that of UHLP-TC.

As for the optimal hub configurations, in the case of the CAB 25-node instances these remained unchanged for all budgets strictly greater than 0. These are $\{13, 19, 25\}$ for UHLP-D and UHLP-TC which coincide with the deterministic case, and $\{5, 13, 19, 25\}$ for the UHLP-DTC. This reaffirms the stability (or robustness) of the hub configurations for this particular dataset as mentioned earlier. In Table 4, we present the sets of open hubs for each of the robust counterparts for the AP and *Set I* 20-node instances. Similarly to Tables 1, 2 and 3, each row corresponds to a set of values of the uncertainty budget for which the optimal solution is the same.

Budget	UHLP-D	UHLP-TC	UHLP-DTC
0.00	7 14	7 14	7 14
0.05	7 14	1 7 14	1 7 14
$[0.1, 0.55]$	7 14	1 7 14	2 7 8 14
$[0.6, 1.0]$	6 11 14	1 7 14	2 7 8 14

(a) AP 20-node instance

Budget	UHLP-D	UHLP-TC	UHLP-DTC
0.00	2 20	2 20	2 20
$[0.05, 0.10]$	2 4 6 11	2 4 13	2 4 6 11
$[0.15, 1.00]$	2 4 6 11	2 4 13	2 4 6 9 20

(b) *Set I* 20-node instance

Table 4: Optimal hub configurations for different budgets of uncertainty

We see that the budget of uncertainty does not lead to as wide a variety of hub configurations as that seen in the analysis of Ω. This supports the hypothesis that the impact caused by an increase in uncertainty budgets for UHLPs becomes negligible after a relatively small value, in particular for the robust counterpart of uncertain transportation costs. It is also interesting to see how early the optimal set of open hubs of the deterministic solution is no longer optimal for the UHLP-TC and the UHLP-DTC for both the AP and *Set I* instances. For these cases, the set of open hubs begins to differ from that of the deterministic solution from as early as a 5% uncertainty budget. For the UHLP-TC, these robust hub configurations remain unchanged for increased uncertainty budgets while for UHLP-DTC only an additional hub configuration is obtained at 10 and 15% for the AP and *Set I* instances, respectively. Finally, we remark the fact that each of the robust counterparts converges to a different optimal solution when assigned a 100% uncertainty budget, thereby showing the uniqueness of each model.

A somehow unexpected result, from our view-point, is that for the cases of UHLP-TC and UHLP-DTC we observed that at optimal solutions, a relatively small percentage of commodity demands were routed via more than one O/D path. This is indeed surprising given that these problems do not consider any capacity constraints. The maximum % of commodities routed via more than one path is 7.2 for the CAB instances and 6.5 for both AP and *Set I* instances. From a practical prespective, it is reasonable for a commodity to take multiple paths since it would allow for alternative routes for its routing in the event that one of the arcs of the cheapest route takes its worst-case value. On the other hand, if a commodity must be routed only through one path, then the fractional routing variables $\bar{x}_{ak}$ may be interpreted as the probability of sending commodity k through hub arc a. This phenomenon occurs only for uncertainty budgets of up to 15% of the number of arcs after which all commodities are routed through a unique path.

4.3. Computational Performance of MIP Formulations

For this set of experiments we use the AP and Set I instances with up to 50 nodes as well as the CAB 25-node instances. We consider $\alpha \in \{0.2, 0.5, 0.8\}$, and $\Omega \in \{0.5, 1\}$. Table 5 summarizes the results obtained by CPLEX when solving the MIP formulations of UHLP-D and UHLP-TC. The first two columns give the information associated with the instance size $|N|$ and interval of uncertainty Ω. The columns under the heading *Solved* give the number of optimal solutions found within 345,600 seconds (four days) of CPU time. The columns under the heading %LP provide the average optimality gap relative to the linear programming relaxation bound of the MIP formulations of each model. The columns under the heading *Time (sec)* give the average CPU time in seconds needed to obtain the optimal solution of each group if instances. The columns under the heading *Nodes* provide the number of explored nodes in the branch and bound.

From Table 5 we note a significant difference in solution times between these models, in particular for the larger instances. For example, the average CPU time for 50-node instances with $\Omega = 0.5$ is under five minutes for UHLP-D while for UHLP-TC it is over 19 hours. This behaviour can be partially explained by the fact that the additional *robust* constraints (11) used in the MIP formulation of UHLP-TC are much more dense than the analogue *robust* constraints (8) of UHLP-D. Another important difference that contributes to this increased difficulty is that the LP gap for UHLP-D is significantly tighter than that of UHLP-TC, which leads to the need

		UHLP-D				UHLP-TC			
$\lvert N \rvert$	Ω	Solved	LP%	Time (secs)	Nodes	Solved	LP%	Time (secs)	Nodes
10	0.5	6/6	0.06	0.07	0.00	6/6	1.11	0.59	4.67
	1	6/6	0.13	0.08	0.00	6/6	2.06	0.93	5.50
20	0.5	6/6	0.00	0.59	0.00	6/6	1.41	24.55	7.83
	1	6/6	0.00	0.67	0.00	6/6	4.86	102.47	44.33
25	0.5	6/6	0.00	1.52	0.00	6/6	1.62	143.93	10.33
	1	6/6	0.00	1.80	0.00	6/6	5.78	1,017.75	86.50
30	0.5	3/3	0.00	7.37	0.00	3/3	0.93	871.52	19.33
	1	3/3	0.00	10.14	0.00	3/3	3.72	3,996.24	47.33
40	0.5	6/6	0.01	52.47	0.50	6/6	1.70	5,254.59	21.00
	1	6/6	0.06	82.72	1.33	6/6	4.69	32,597.16	73.50
50	0.5	6/6	0.00	289.46	0.00	6/6	2.02	69,444.89	35.50
	1	6/6	0.01	474.61	0.00	1/6	7.04	185,957.97	107.00

Table 5: Summary results of UHLP-D and UHLP-TC

for further exploration in the enumeration tree for the latter as seen in Table 5. Another interesting observation is the increase in computation time required when solving instances with $\Omega = 1$ when compared to $\Omega = 0.5$. This characteristic is seen for both models and shows how instances with larger uncertainty sets are more difficult to solve. Finally, we point out that 5 of the 50 node instances were not solved within the time limit for the UHLP-TC model while for the UHLP-D model 20 minutes was the longest computation time required.

In order to solve the more challenging UHLP-DTC using the MIP formulation given in Section 3.3, we need to handle the huge number of *robust* constraints (12) in an efficient way. For this reason, we develop a simple branch-and-cut algorithm based on this MIP. The idea is to solve its LP relaxation with a cutting-plane algorithm by initially relaxing all of constraints (12) and then iteratively adding only the ones violated by the current LP solution. When no more violated inequalities (12) exist, within a threshold value $\epsilon = 1\text{E}{-4}$, or when the improvement in the objective function is below 0.5% we resort to CPLEX for solving the resulting formulation by enumeration, using a call-back function for generating additional violated inequalities at some nodes of the enumeration tree. In particular, at most one cut is added at a fractional solution obtained at nodes where its depth in the tree is divisible by three. The separation problem of constraints (12) is the following: given a fractional solution $(\bar{z}, \bar{x}, \bar{y})$ of the linear relaxation, determine whether there exists one or more inequalities in the set (12) that are not satisfied by $(\bar{z}, \bar{x}, \bar{y})$. Note that separation problem is equivalent to the solution of the SP for a given solution $(\bar{z}, \bar{x}, \bar{y})$. In our implementation, we solve the linearized version of SP with CPLEX to find the most violated inequality, if any.

The detailed results of the branch-and-cut algorithm for the UHLP-DTC are given in Table 6. The columns under the headings *Solved*, *%LP*, *Nodes*, and *Time (sec)*, have the same interpretation as in the previous table. The columns under the headings *Int cuts* and *Frac cuts* provide the average number of cuts added at integer and fractional solutions, respectively. The last column under the heading *% Sep time* gives the average percentage of the total time spent solving the separation problem SP.

$\|N\|$	Ω	Solved	%LP	Int cuts	Frac cuts	Nodes	Time (sec)	% Sep time
10	0.5	6/6	0.89	97.83	7.17	5.83	4.20	42.54
	1	6/6	1.93	231.83	10.83	17.17	10.61	38.37
20	0.5	6/6	1.68	58.17	14.33	22.33	164.45	10.93
	1	6/6	4.80	437.33	53.83	197.33	1,997.09	8.11
25	0.5	6/6	1.67	65.50	16.67	32.50	848.30	7.63
	1	6/6	6.21	611.67	151.50	621.50	37,273.18	3.51
30	0.5	3/3	0.74	134.33	24.33	38.67	5,660.83	2.97
	1	2/3	1.40	153.00	37.00	73.00	12,268.55	1.92
40	0.5	6/6	1.46	117.17	25.17	54.00	35,489.97	1.45
	1	2/6	1.73	223.00	50.00	157.00	218,011.02	0.53
50	0.5	4/6	N/A	147.75	28.25	55.25	247,633.79	0.82
	1	1/6	N/A	222.00	22.00	3.00	322,896.84	0.80

Table 6: Summary results of UHLP-DTC

From Table 6, we observe that UHLP-DTC requires the longest solution time of all three robust counterparts. Our algorithm is capable of consistently solving instances with up to 30 nodes. After that, it is capable of solving 10 out of 12 instances with 40 and 50 nodes with $\Omega = 0.5$ but is only capable of solving 3 out of 12 instances with $\Omega = 1$ in four days of CPU time. The long solution times come as a result of the need to add a relative large number of dense cuts which in turn, considerably increases the time to solve the LP relaxations, and to explore more nodes in the enumeration tree to prove optimality. Note that although many cuts are added, very little time is spent solving the separation problem, showing that despite the theoretical NP-hardness of SP, CPLEX is still able to solve it efficiently. Finally, we point out the relatively small %LP gaps of this formulation.

4.4. A Comparison of Deterministic, Stochastic and Robust Solutions

In this final set of computational experiments, we compare the solutions obtained from the deterministic UHLP, a two-stage stochastic variant of the UHLP presented in Contreras et al. (2011b), and our proposed robust counterpart UHLP-TC. To evaluate these solutions as fairly as possible we define $W_k(\xi)$ and $d_{ij}(\omega)$ as random variables following uniform distributions over the intervals $W_k(\xi) \in [W_k^L - W_k^\Delta, W_k^L + W_k^\Delta]$ and $d_{ij}(\omega) \in [d_{ij}^L - d_{ij}^\Delta, d_{ij}^L + d_{ij}^\Delta]$, respectively.

As mentioned in Section 2, Contreras et al. (2011b) show that when considering demand uncertainty, the two-stage stochastic model of UHLP is actually equivalent to its associated expected value problem in which the uncertain demand is replaced by its expectation. However, when considering both demand and independent transportation cost uncertainty this situation does not hold. For the sake of completeness, we next present the stochastic model with demand and transportation cost uncertainty.

We assume that the random parameters $d_{ij}(\omega)$ and $W_k(\xi)$ are independent and use the notation $C_{ak}(\omega) = \chi d_{o(k)i}(\omega) + \alpha d_{ij}(\omega) + \delta d_{jd(k)}(\omega)$ where $a = (i, j)$. We define z_i as first stage variables and $x_{ak}(\omega, \xi)$ as second stage variables that adjust depending on the realization of the uncertainty parameter. The *uncapacitated hub location problem with stochastic demands and transportation costs* (UHLP-SDTC) can be stated as follows:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i \in N} f_i z_i + E_{\omega,\xi}[\sum_{k \in K} \sum_{a \in A} C_{ak}(\omega) W_k(\xi) x_{ak}(\omega, \xi)] && (14) \\
\text{subject to} \quad & \sum_{a \in A} x_{ak}(\omega, \xi) = 1 \qquad \forall k \in K, \omega \in \Omega, \xi \in \Xi && (15) \\
& \sum_{a \in A: i \in a} x_{ak}(\omega, \xi) \leq z_i \qquad \forall k \in K, i \in N, \omega \in \Omega, \xi \in \Xi && (16) \\
& z_i \in \{0, 1\} \qquad \forall i \in H && (17) \\
& 0 \leq x_{ak}(\omega, \xi) \leq 1 \qquad \forall a \in A, k \in K, \omega \in \Omega, \xi \in \Xi. && (18)
\end{aligned}
$$

Since both parameters are independently distributed then $E_{\omega,\xi} = E_\omega E_\xi$. Also, note that for a first-stage vector z and a fixed realization of distances ω, the optimal route $x_{ak}(\omega, \xi)$ is the same independent of the realization of ξ, i.e. $x_{ak}(\omega, \xi) =$

$x_{ak}(z,\omega)$. Therefore, given (z,ω) we have

$$E_\omega E_{\xi|\omega}\left[\sum_{k,a} C_{ak}(\omega)W_k(\xi)x_{ak}(\omega,\xi)\right] = E_\omega E_{\xi|\omega}\left[\sum_{k,a} C_{ak}(\omega)W_k(\xi)x_{ak}(z,\omega)\right] \quad (19)$$

$$= E_\omega\left[\sum_{k,a} E_\xi\left[C_{ak}(\omega)W_k(\xi)x_{ak}(z,\omega)\right]\right] \quad (20)$$

$$= E_\omega\left[\sum_{k,a} C_{ak}(\omega)W_k^L x_{ak}(z,\omega)\right], \quad (21)$$

where (19) comes from the independence of $x_{ak}(\omega,\xi)$ of the realization of ξ, (20) comes from the independence of ω and ξ, and (21) is from our assumptions on the uncertainty distribution of the $W_k(\xi)$. As a consequence, UHLP-SDTC reduces to the uncapacitated hub location problem with stochastic independent transportation costs presented in Contreras et al. (2011b). Hence, we use the sample average approximation algorithm proposed in Contreras et al. (2011b) to obtain ϵ-optimal solutions for the stochastic counterparts. In our experiments, we take 20 samples of size 1,000 to estimate a lower bound and a larger sample size of 100,000 to estimate an upper bound.

Our comparison methodology is comprised of two steps. The first is to note when the solutions obtained from the stochastic, robust and deterministic models coincide with respect to the open hubs. The second is to study instances for which these differ and compare each proposed solution in both a risk-neutral and worst-case scenario. This analysis is done with the AP and *Set I* 20-node instances as well as the CAB 25-node instance for values of $\Omega \in \{0.5, 1.0, 1.5, 2.0\}$ and $\alpha \in \{0.2, 0.5, 0.8\}$. To obtain the corresponding robust solutions, we solve UHLP-TC at budget values of $\{0.1, 0.2, 0.3, ..., 1\}$ and consider all hub configurations obtained for these. Tables 7, 8a and 8b show the optimal hub configurations obtained by each approach for the CAB, AP and *Set I* instances, respectively.

Ω	α	Det	Stoch	Robust
0.5	0.2, 0.5, 0.8	13 19 25	13 19 25	13 19 25
1.0	0.2, 0.5, 0.8	13 19 25	13 19 25	13 19 25
1.5	0.2, 0.5, 0.8	13 19 25	13 19 25	5 13 19 25
2.0	0.2, 0.5, 0.8	13 19 25	5 19 25	5 13 19 25

Table 7: Optimal hubs for each approach with CAB instances

Ω	α	Det	Stoch	Robust
0.5	0.2, 0.5 0.8	7 14	7 14	7 14
1	0.2	7 14	7 14	1 7 14
1	0.5, 0.8	7 14	7 14	7 14
1.5	0.2	7 14	7 14	1 7 8 14
1.5	0.5	7 14	7 14	1 8 11 14 1 7 14
1.5	0.8	7 14	7 14	6 8 14
2	0.2	7 14	7 11 14	1 7 8 14
2	0.5	7 14	7 11 14	1 8 11 14 1 7 8 14
2	0.8	7 14	7 11 14	1 8 11 14 1 7 8 14 7 8 14

(a) Optimal hubs for AP instances

Ω	α	Det	Stoch	Robust
0.5	0.2, 0.5	2 20	2 20	2 4 13
0.5	0.8	2 20	2 20	2 20
1	0.2, 0.5	2 20	2 20	2 4 13
1	0.8	2 20	2 20	2 6 11 2 13
1.5	0.2	2 20	2 11 20	2 4 13
1.5	0.5	2 20	2 11 20	2 4 6 11 2 4 13
1.5	0.8	2 20	2 11 20	2 6 11 19 2 4 6
2	0.2	2 20	2 11 20	2 4 6 9 20
2	0.5	2 20	2 11 20	2 4 6 11 2 4 13
2	0.8	2 20	2 11 20	2 6 11 19 2 4 6 11 2 13 19

(b) Optimal hubs for *Set I* instances

Table 8: Optimal hub configurations for *Set I* and AP instances

From Table 7 we note that for the CAB instances, the open hubs of all three approaches coincide for all instances with $\Omega \in \{0.5, 1.0\}$. This shows the stability of the open hubs for this dataset requiring an increase of 150% of the nominal cost for the hub configuration to change. Even at a 150% level of uncertainty, the stochastic approach still keeps the same hub configuration as the deterministic case while the robust counterparts open only one additional hub. Table 8a shows that for the case of the AP instances, for $\Omega \in \{0.5, 1.0\}$ and $\alpha \in \{0.5, 0.8\}$, all three approaches give the same optimal hub configuration. At all other values, the robust solutions differ from those obtained from the other approaches while the stochastic approach opens the same hubs as the deterministic model for all values except for $\Omega = 2.0$. For the *Set I* instances, Table 8b shows that all three approaches give the same hub configuration only when $\Omega = 0.5$ and $\alpha = 0.8$. The robust approach differs from the other approaches for all other tested values while the stochastic approach obtains different hub configurations when $\Omega \in \{1.5, 2.0\}$.

From each dataset, we arbitrarily select a value of Ω and α that yield different hub configurations for each approach to measure their performance in a worst-case and risk-neutral setting. To evaluate in a worst-case framework, we fix the hub variables z_i from UHLP-TC to the hub configurations obtained from each approach and solve the remaining model for budgets of $\{0.1, 0.2, 0.3, ..., 1.0\}$, while in the risk-neutral setting we use Monte Carlo simulation to evaluate the average performance of all hub configurations.

A summary of the performance of optimal hub configurations obtained from each approach (deterministic, stochastic, and robust) in a worst-case setting for the CAB, AP and *Set I* instances is given in Figures 6, 6a and 6b, respectively. In particular, these figures give the percentage increase in the optimal solution value with respect to the deterministic (nominal) cost for different values of the budget of uncertainty. Table 9 shows a similar analysis for a risk-neutral setting.

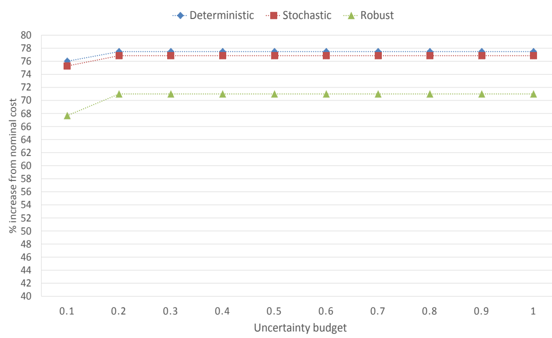
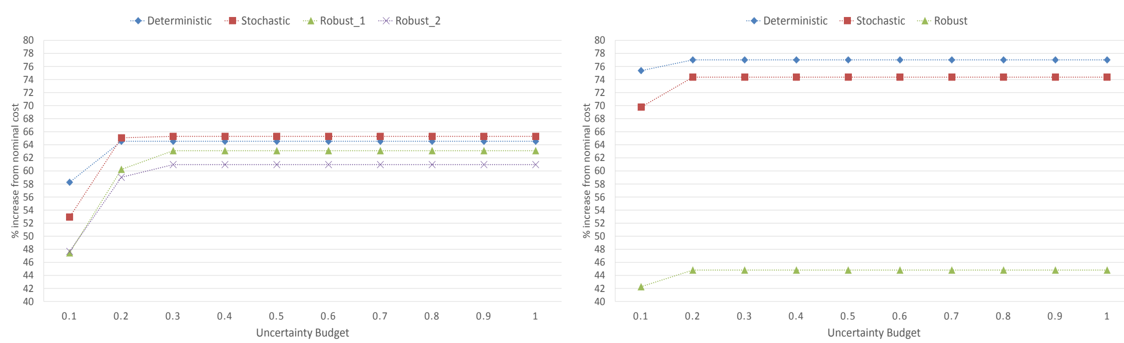


Figure 6: Solution performance in worst-case setting for CAB instance with $\Omega = 2$ and $\alpha = 0.2$



(a) AP instance with $\Omega = 2$ and $\alpha = 0.2$ (b) *Set I* instance with $\Omega = 2$ and $\alpha = 0.5$

Figure 7: Solution performance in worst-case setting

Instance	Approach	Open Hubs	Exp. Cost	σ	%Increase
	Stochastic	5 19 25	1,421,798.65	824.57	
CAB	Deterministic	13 19 25	1,424,828.17	866.89	0.21
	Robust	5 13 19 25	1,433,177.30	758.16	0.80
	Stochastic	7 11 14	199,390.29	95.38	
AP	Deterministic	7 14	200,442.11	106.96	0.53
	Robust1	1 8 11 14	214,087.97	78.69	7.37
	Robust2	1 7 8 14	216,080.27	74.48	8.37
	Stochastic	2 11 20	40,576.19	31.96	
Set I	Deterministic	2 20	41,946.05	38.78	3.38
	Robust	2 4 6 9 20	42,832.30	17.07	5.56

Table 9: Solution performance in a risk-neutral setting

For the case of the *Set I* instance, note that in a worst-case scenario, the relative increase in optimal value is almost 30% greater for the hub configuration from the stochastic model than that from the robust model. Also, note that the hub configuration of the deterministic model has the worst performance of the three in this setting. However, observe that in the case of the risk-neutral setting the average performance of the hub configuration of the robust model is only 5.56% worse than that of the stochastic model while that of the deterministic model is 3.38% worse. This shows how the hub configuration of our robust approach significantly outperforms those obtained from the stochastic and deterministic approaches in a worst-case setting and is only slightly worse in a risk-neutral setting. A somehow similar behaviour is noted for the CAB instance as seen in Figure 6 and Table 9, although in a risk-neutral setting both deterministic and robust models have an increase that does not exceed 1%. For the AP instance the hub configuration of the stochastic approach is the worst performing in a worst-case setting as seen in Figure 6a while the hub configurations of the robust approach are on average under 8.4% more costly than the stochastic solution in a risk-neutral setting as seen in Table 9.

5. Conclusion

In this paper we introduced robust counterparts of the well-known uncapacitated hub location problem with multiple assignments for the cases considering uncertain demands, transportation costs, and both simultaneously. We considered an interval of uncertainty for demands and transportation costs and used a budget of uncertainty to control the level of conservatism in solution networks. We presented mixed integer linear programming formulations for each of the considered robust counterparts. We performed extensive computational experiments to evaluate the performance of our formulations and to study the impact of the intervals of uncertainty and of the budgets of uncertainty in the optimal hub configurations. In addition, we compared the performance between solutions obtained from deterministic, stochastic and robust models in both worst-case and risk-neutral settings.

The experiments showed that the interval of uncertainty has a considerable effect on the optimal configuration and cost of hub networks. In particular, UHLP-DTC and UHLP-D seem to be the most and the least susceptible models to uncertainty, respectively. Also, the optimal hub configuration of UHLP-DTC changes from that of the deterministic solution at a lower level of uncertainty than the other two robust models. The computational results also showed that an increase in the budget of uncertainty does not lead to a large variety of hub configurations as the ones observed when increasing the intervals of uncertainty. Among the three considered sets of instances, CAB instances demonstrated the most robustness in terms of changes in optimal hub network configurations. Finally, the comparison of robust UHLP-TC to deterministic and stochastic UHLPs showed that the hub networks of the proposed robust model significantly outperformed those obtained from the stochastic and deterministic models in a worst-case setting and was only slightly worse in a risk-neutral setting.

Acknowledgments

This research was partially funded by the Canadian Natural Sciences and Engineering Research Council [Grants 418609-2012 and 04959-2014]. This support is gratefully acknowledged. Computations were made on the supercomputers managed by Calcul Québec and Compute Canada. The operation of these supercomputers is funded by the Canada Foundation for Innovation (CFI), the ministère de l'Économie, de la science et de l'innovation du Québec (MESI) and the Fonds de recherche du Québec - Nature et technologies (FRQ-NT).

References

Alibeyg, A., Contreras, I., Fernández, E., 2016. Hub network design problems with profits. Transportation Research Part E: Logistics and Transportation Review 96, 45–59.

Alumur, S. A., Kara, B. Y., Karasan, O. E., 2012a. Multimodal hub location and hub network design. Omega 40 (6), 927 – 939.

Alumur, S. A., Nickel, S., da Gama, F. S., 2012b. Hub location under uncertainty. Transportation Research Part B: Methodological 46 (4), 529 – 543.

An, Y., Zhang, Y., Zeng, B., 2015. The reliable hub-and-spoke design problem: Models and algorithms. Transportation Research Part B: Methodological 77, 103 – 122.

Azizi, N., Chauhan, S., Salhi, S., Vidyarthi, N., 2016. The impact of hub failure in hub-and-spoke networks: Mathematical formulations and solution techniques. Computers and Operations Research 65, 174 – 188.

Bertsimas, D., Sim, M., 2003. Robust discrete optimization and network flows. Mathematical Programming 98 (1-3), 49–71.

Bruglieri, M., Ehrgott, M., Hamacher, H. W., Maffioli, F., 2006. An annotated bibliography of combinatorial optimization problems with fixed cardinality constraints. Discrete Applied Mathematics 154 (9), 1344 – 1357.

Campbell, J. F., Ernst, A., Krishnamoorthy, M., 2005. Hub arc location problems part I: Introduction and results. Management Science 51 (10), 1540–1555.

Campbell, J. F., O'Kelly, M. E., 2012. Twenty-five years of hub location research. Transportation Science 46 (2), 153–169.

Contreras, I., 2015. Hub location problems. In: Laporte, G., Saldanha da Gama, F., Nickel, S. (Eds.), Location Science. Springer, pp. 311–344.

Contreras, I., Cordeau, J.-F., Laporte, G., 2011a. Benders decomposition for large-scale uncapacitated hub location. Operations Research 59 (6), 1477–1490.

Contreras, I., Cordeau, J.-F., Laporte, G., 2011b. Stochastic uncapacitated hub location. European Journal of Operational Research 212 (3), 518 – 528.

Contreras, I., Díaz, J. A., Fernández, E., 2011b. Branch and price for large-scale capacitated hub location problems with single assignment. INFORMS Journal on Computing 23 (1), 41–55.

Contreras, I., Fernández, E., 2014. Hub location as the minimization of a supermodular set function. Operations Research 62 (3), 557–570.

Contreras, I., Tanash, M., Vidyarthi, N., 2016. Exact and heuristic approaches for the cycle hub location problem. Annals of Operations ResearchDOI 10.1007/s10479-015-2091-2.

Corneil, D., Perl, Y., 1984. Clustering and domination in perfect graphs. Discrete Applied Mathematics 9 (1), 27 – 39.

Correia, I., da Gama, F. S., 2015. Facility location under uncertainty. In: Laporte, G., Nickel, S., Saldanha da Gama, F. (Eds.), Location Science. Springer, pp. 177–203.

Correia, I., Nickel, S., da Gama, F. S., 2010. Single-assignment hub location problems with multiple capacity levels. Transportation Research Part B: Methodological 44 (8–9), 1047 – 1066.

de Camargo, R., Miranda, G., 2012. Single allocation hub location problem under congestion: Network owner and user perspectives. Expert Systems with Applications 39 (3), 3385 – 3391.

de Camargo, R. S., Miranda, G., Luna, H. P., 2008. Benders decomposition for the uncapacitated multiple allocation hub location problem. Computers and Operations Research 35 (4), 1047–1064.

Elhedhli, S., Hu, F. X., 2005. Hub-and-spoke network design with congestion. Computers and Operations Research 32 (6), 1615 – 1632.

Elhedhli, S., Wu, H., 2010. A lagrangean heuristic for hub-and-spoke system design with capacity selection and congestion. INFORMS J. on Computing 22 (2), 282–296.

Ghaffari-Nasab, N., Ghazanfari, M., Teimoury, E., 2015. Robust optimization approach to the design of hub-and-spoke networks. The International Journal of Advanced Manufacturing Technology 76, 1091–1110.

Habibzadeh Boukani, F., Farhang Moghaddam, B., Pishvaee, M., 2016. Robust optimization approach to capacitated single and multiple allocation hub location problems. Computational and Applied Mathematics 35, 45–60.

Hamacher, H. W., Labbé, M., Nickel, S., Sonneborn, T., 2004. Adapting polyhedral properties from facility to hub location problems. Discrete Applied Mathematics 145 (1), 104 – 116.

Louveaux, F., 1993. Stochastic location analysis. Location Science 1, 127 – 154.

Mahmutogullari, A. I., Kara, B. Y., 2016. Hub location under competition. European Journal of Operational Research 250 (1), 214 – 225.

Marianov, V., Serra, D., 2003. Location models for airline hubs behaving as m/ d / c queues. Computers & Operations Research 30 (7), 983 – 1003.

Marín, A., Canovas, L., Landete, M., 2006. New formulations for the uncapacitated multiple allocation hub location problem. European Journal of Operational Research 172, 274–292.

Martins de Sá, E., Contreras, I., Cordeau, J., 2015. Exact and heuristic algorithms for the design of hub networks with multiple lines. European Journal of Operational Research 246, 186–198.

Merakli, M., Yaman, H., 2016. Robust intermodal hub location under polyhedral demand uncertainty. Transportation Research Part B: Methodological 86, 66 – 85.

O'Kelly, M. E., 1986. The location of interacting hub facilities. Transportation Science 20 (2), 92–106.

O'Kelly, M. E., Campbell, J. F., Camargo, R. S., Miranda, G., 2015a. Multiple allocation hub location model with fixed arc costs. Geographical Analysis 47 (1), 73—96.

O'Kelly, M. E., Luna, H. P. L., De Camargo, R. S., de Miranda Jr, G., 2015b. Hub location problems with price sensitive demands. Networks and Spatial Economics 15 (4), 917–945.

Serper, E. Z., Alumur, S. A., 2016. The design of capacitated intermodal hub networks with different vehicle types. Transportation Research Part B: Methodological 86, 51 – 65.

Shahabi, M., Unnikrishnan, A., 2014. Robust hub network design problem. Transportation Research Part E: Logistics and Transportation Review 70, 356 – 373.

Sim, T., Lowe, T. J., Thomas, B. W., 2009. The stochastic -hub center problem with service-level constraints. Computers & Operations Research 36 (12), 3166 – 3177, new developments on hub location.

Snyder, L., 2006. Facility location under uncertainty: A review. IIE Transactions 38, 537 – 554.

Sun, Z., Zheng, J., 2016. Finding potential hub locations for liner shipping. Transportation Research Part B: Methodological, –.

Tanash, M., Contreras, I., Vidyarthi, N., 2016. An exact algorithm for the modular hub location problem. submitted.

Yang, T.-H., 2009. Stochastic air freight hub location and flight routes planning. Applied Mathematical Modelling 33 (12), 4424 – 4430.

Zanjirani Farahani, R., Hekmatfar, M., Arabani, A. B., Nikbakhsh, E., 2013. Hub location problems: A review of models, classification, solution techniques, and applications. Computers & Industrial Engineering 64 (4), 1096–1109.