ISE Industrial and Systems Engineering

A distributionally robust optimization approach for a stochastic mobile facility routing and scheduling problem

Karmel S. Shehadeh[1]

[1]Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA

ISE Technical Report 20T-19

Lehigh University

A distributionally robust optimization approach for a stochastic mobile facility routing and scheduling problem

Karmel S. Shehadeh[a,*]

[a]*Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA, USA*

Abstract

A mobile facility (MF) is a facility capable of moving from one place to another, providing real-time services to customers when it is stationary. In this paper, we study an MF routing and scheduling problem in which probability distributions of the time-dependent customers demands for MF services are unknown. We first construct an ambiguity set of all possible demand distributions that share identical support, mean, and dispersion in each period. Then, we define a *distributionally robust mobile facility routing and scheduling* (DMFRS) problem that seeks optimal routing and scheduling decisions for a fleet of MFs to minimize the fixed operating costs and worst-case expected transportation and unmet demand costs generated during a planning horizon. We take the worst-case expectation over all distributions residing within the ambiguity set. To solve our DMFRS model, we propose a decomposition-based algorithm and derive lower bound and two-families of symmetry breaking inequalities to strengthen the master problem and speed up convergence. We conduct extensive computational experiments that demonstrate (1) how our DRO approach has a superior computational and operational performance as compared to the SP approach, (2) efficiency of the symmetry breaking and lower bound inequalities, (3) value of considering the distributional ambiguity of random demand, and (4) the trade-off between cost, number of MFs, and MF capacity.

Keywords: (R) Facilities planning and design, mobile facility, demand uncertainty, routing and scheduling, distributionally robust optimization;, symmetry-breaking

---

1. Introduction

A mobile facility (MF) is a facility capable of moving from one place to another, providing real-time service to customers in the vicinity of its location when it is stationary (Halper and Raghavan, 2011). In this paper, we study a plain mobile facility (MF) routing and scheduling problem with stochastic demand first introduced by Lei et al. (2014). Specifically, we seek to find the routes and time schedule for a fleet of MFs in a given service region over a specified planning horizon.

*Corresponding author.
Email address: kshehadeh@lehigh.edu, kas720@lehigh.edu (Karmel S. Shehadeh)

Preprint submitted to Optimization Online (under journal review) *January 22, 2021*

Customers demand for MF service are time-dependent and random. The probability distribution of the demand in each period is unknown, and only the mean and range of the demand are known. The quality of the MFs routing and scheduling decisions is a function of the fixed operation costs, cost of assigning demands to the MFs (e.g., transportation, shipping), and cost of unsatisfied demand (e.g., outsourcing.).

The concept of MF routing and scheduling is very different than conventional static facility location (FL) problem and vehicle routing (VR) problems. In static FL problems, we usually consider opening facilities at fixed locations. Conventional VR problems aims at handling the movement of items between facilities (e.g., depots) and customers. MF is a "*facility-like-vehicle*" that behaves in a way similar to traditional facilities when they are stationary except that they can move from one place to another if necessary (Lei et al., 2014). Thus, the most evident advantage of MF over fixed facilities is their flexibility in moving to accommodate the change in the demand over time and location (Halper and Raghavan, 2011; Lei et al., 2014, 2016).

MFs are used in many applications ranging from cellular services to humanitarian relief logistics. For example, light trucks with portable cellular stations can provide cellular service in areas where existing cellular network of base stations temporarily fails (Halper and Raghavan, 2011). In humanitarian relief logistics, MFs give relief organizations the ability to provide aid to populations dispersed in remote regions and dense urban areas. Mobile clinics, for example, can travel to the heart of communities, both rural and urban, and provide prevention and health care services (Ahmadi-Javid et al., 2017; Alexy and Elnitsky, 1996). Recently, mobile health clinics played a significant role in providing drive-through COVID-19 testing sites or triage locations in parking lots near emergency departments or at regional areas in response to the COVID-19 pandemic (Attipoe-Dorcoo et al., 2020).

MF operators often seek a tactical plan, including the MF fleet's routes and time schedules, that reduces their fixed operating costs and maximizes demand satisfaction. The MF routing and scheduling problem (MFRSP) is a challenging optimization problem for two primary reasons. First, customers demand in each period and at each location is random and hard to predict in advance. Second, even in a perfect world in which we know with certainty the amount of demand in each period, the deterministic MFRSP is NP-hard as it is reduced to the classic FL problem (Halper and Raghavan, 2011; Lei et al., 2014). Thus, the incorporation of demand variability makes the problem more challenging.

To model uncertainty, Lei et al. (2014) propose the first a *priori* two-stage stochastic optimization model (SP) for MFRSP, which seeks optimal routing and scheduling decisions to minimize the total expected system-wide cost, where the expectation is taken with respect to a known probability distribution of customers' demand. Although attractive, the applicability of the SP approach is limited to the case in which we know the distribution of the demand. In practice, it is unlikely that decision-makers can estimate the actual probability distributions of random demand accurately, especially with limited data during the planning process before observing customer demand (Basciftci et al., 2019; Lei et al., 2016; Liu et al., 2019). If we solve an SP model with a data sample from a biased distribution, then the resulting (biased) optimal decisions may have a disappointing performance under the true distribution (Esfahani and Kuhn, 2018). Additionally, SP approaches suffer from the "*curse of dimensionality*," and thus are often intractable.

Lei et al. (2016) proposed the first two-stage robust optimization (RO) approach for an MF fleet sizing and routing problem. Lei et al. (2016) assume that demand resides in a polyhedral uncertainty set based on the maximum "positive" deviation from the mean. The objective is to minimize a fixed cost of establishing the MF fleet (first stage), and a penalty cost for unmet demand (second stage). Optimization is performed with respect to the worst-case scenario (i.e., mean+maximum positive deviation) in the uncertainty set, which, as pointed out by (Chen et al., 2020; Delage and Saif, 2018; Thiele, 2010), can inevitably lead to overly conservative and suboptimal decisions for other more-likely demand's scenarios.

In this paper, we aim to address the lack of distributional information (i.e., distributional ambiguity) in MFRSP using distributionally robust optimization (DRO). In DRO, we assume that the distribution of random parameters resides in a so-called ambiguity set, and optimization is based on the worst-case distribution within this set. The ambiguity set is a family of distributions characterized by some known properties of random parameters (Esfahani and Kuhn, 2018; Rahimian and Mehrotra, 2019). DRO aims to unify SP and RO while overcoming their drawbacks and has recently gained significant attention for the following reasons. First, DRO alleviates the unrealistic assumption of the decision-maker's complete knowledge of random parameter distributions. Second, DRO models are often more computationally tractable than their SP and RO counterparts for most real-world problems (Delage and Saif, 2018; Rahimian and Mehrotra, 2019; Shehadeh and Padman, 2020; Wang et al., 2019; Wiesemann et al., 2014). Third, one can use easy approximate distributional information (e.g., mean values and range) to construct the ambiguity sets and build DRO models that mimic reality and less conservative than RO. Finally, as Kuhn et al. (2019) pointed out, DRO faithfully anticipates the possibility of "*black swans*", i.e., mitigate the disappointing consequences (black swans) associated with implementing the SP or sample-based solutions under the true distribution or data that differ from all training samples.

1.1. Contributions

In this paper, we study a *distributionally robust mobile facility routing and scheduling* (DMFRS) problem. Specifically, given a set of MFs, we aim to find: (1) the number of MFs to use within T, (2) a routing plan for the selected MFs, i.e., routes that specify the movement of each MF, (3) a timetable of the MFs routing plan (i.e., time schedule), and (4) assignment plan from facility to customers. Decisions (1)–(3) are first-stage decisions that we make before realizing the demand. The assignment decisions (4) represent the recourse actions in response to the first-stage decisions and demand realizations. The objective is to minimize first-stage fixed cost (i.e., cost of using MFs and traveling inconvenience cost) and the worst-case expectation of transportation and unsatisfied demand costs over a planning horizon. We take the worst-case expectations over a family of demand distributions that have the same support, mean, and dispersion (MAD) in each period. As in Wang et al. (2019) AND Wang et al. (2020), we use MAD as the dispersion or variability measure because, as we will show later in Section 3.3, it enables a linear and computationally attractive reformulation. We summarize our main contributions as follows.

- To the best of our knowledge, and according to our literature review in Section 2, our paper is the first to addresses the distributional ambiguity of the time-based demand in this specific MF routing and scheduling problem. In contrast to Lei et al. (2016)'s RO model, we additionally incorporate the MF traveling inconvenience cost in the first-stage objective and the random transportation cost in the second-stage objective. Additionally, instead of focusing on demand's worst-case scenario (mean+maximum positive deviation) as in Lei et al. (2016), we optimize the system performance over all demand distributions residing within the ambiguity set.
- We propose a computationally efficient decomposition-based algorithm to solve DMFRS. We derive valid lower bound inequalities to strengthen the master problem and improve convergence. Lei et al. (2016) also applied the same art of cutting plane-based decomposition algorithms. Nevertheless, our master and sub- problems and lower bound inequalities have a different structure than those of Lei et al. (2016) due to the differences in the decision variables and objectives. Our decomposition approach is also different than that of Shehadeh and Sanci (2020). Shehadeh and Sanci (2020)'s DRO approach address a "static" FL problem with time–independent and bimodal demand. They aim to select a subset of static facilities to open to minimize the opening facility's fixed cost and the worst-cost expected transportation and unmet penalty costs, where the worst-case expectation is taken with respect to all bimodal distributions residing within a scenario-wise ambiguity set. In contrast, we address an MF routing and scheduling problem with time–dependent and unimodal demand. Thus, our optimization problem, formulation, ambiguity set, and consequently, the master and sub-problems and bounding inequalities are different (see Section 2 for a detailed comparison).
- We derive two–families of symmetry breaking inequalities, which break symmetries in the solution space of the first-stage routing and scheduling decisions. These inequalities are novel in the sense that they are independent of the method of modeling uncertainty and thus valid for and can improve the solvability of any SP, DRO, RO, and deterministic formulation that uses the same routing and scheduling decisions in the first stage. Lei et al. (2016), Lei et al. (2014), and other pioneering work did not address the issue of symmetry in deterministic and stochastic MFRSP. Thus, our inequalities are the first of their kind.

- To analyze a DRO approach for MFRSP, we conduct computational experiments that demonstrate (1) how our DRO approach has a superior computational and operational performance as compared to the SP approach, (2) efficiency of the symmetry breaking and lower bound inequalities, (3) value of considering the distributional ambiguity of random demand, and (4) the trade-off between cost, number of MFs, and MF capacity.

1.2. Structure of the paper

The remainder of the paper is structured as follows. In Section 2, we review the relevant literature. In Section 3, we formally define DMFRS and its reformulation. In Section 4, we present our DMFRS–decomposition algorithm and strategies to improve convergence. In Section 5, we present our computational results. Finally, we draw conclusions and discuss future directions in Section 6.

2. Relevant Literature

There is limited literature on mobile facilities as compared to stationary facilities. However, as pointed out by Lei et al. (2014), MFRSP share some features with several well-studied problems, including dynamic Facility Location Problem (DFLP), Vehicle Routing Problem (VRP), and the Covering Tour Problem (CTP). In this review, we briefly discuss the similarities and differences between MFRSP and these problems.

Given that we consider making decisions over a given planning period, then MFRSP is somewhat similar to DFLP, which seeks to locate/re-locate facilities over a planning horizon. To mitigate the impact of demand fluctuation along the planning period, decision-makers may open new facilities and close or relocate existing facilities at a relocation cost (Albareda-Sambola et al. (2009); Antunes et al. (2009); Contreras et al. (2011); Drezner and Wesolowsky (1991); Jena et al. (2017); Manzini and Gebennini (2008); Owen and Daskin (1998); Van Roy and Erlenkotter (1982); Wesolowsky and Truscott (1975)). Most DFLPs assume that the relocation time is relatively short as compared to the planning horizon. In contrast, MFRSP takes into account the relocation time of MFs. In addition, each MF needs to follow a specific route during the entire planning horizon, which is not a requirement in the DFLP.

In CTP, one seeks to select a subset of nodes to visit that can cover other nodes within a particular coverage (Current et al. (1985); Flores-Garza et al. (2017); Gendreau et al. (1997); Hachicha et al. (2000); Tricoire et al. (2012)). In contrast to MFRSP, CTP does not consider the variations of demand over the planning horizon and assumes that the amount of demand to be met by vehicles is not related to the length of time the MF spending at the stop.

The VRP is one of the most extensively studied problems in operations research that has numerous applications and variants (Subramanyam et al., 2020). Both MFRSP and VRP consider the routing decisions of vehicles. As described in Lei et al. (2014), MFRSP is different than VRP in the following ways. First, in MFRSP, we can meet customer demand by a nearby MF (e.g., cellular stations). In VRP, vehicles visit customers to meet their demands. Second, the amount of demand that an MF can serve at each location depends on the duration of the MF stay at each location, which is a decision variable. In contrast, VRPs often assume a fixed service time. Finally, most of the VRPs require that each customer be visited exactly once in each route. In contrast, the MFRSP allows zero or multiple visits to any customer in each route. For comprehensive surveys of stochastic VRP, we refer to Bertsimas and Simchi-Levi (1996); Cordeau et al. (2007); Oyola et al. (2017); Ritzinger et al. (2016), andToth and Vigo (2014).

Halper and Raghavan (2011) introduced the concept of MF and proposed a continuous-time formulation to model the maximum covering mobile facility routing problem for multiple facilities under deterministic settings. To solve their model, Halper and Raghavan (2011) proposes several computationally effective heuristic algorithms. In DMFRS, we consider uncertainty of demand distribution toward minimizing the average cost of the entire system.

To avoid the challenges of dynamic and re-optimization approaches, Lei et al. (2014) propose the first a *priori* two-stage stochastic optimization model (SP) for the plain MFRSP. Lei et al. (2014)'s SP seeks optimal first-stage routing and scheduling decisions to minimize the total expected system-wide cost, where the expectation is taken with respect to a known probability distribution of the demand. A priori optimization has a managerial advantage since it guarantees the regularity of service, which is beneficial for both customer and service provider. That is, a prior plane allow the customers to know when and where to obtain service and enable MF service providers to be familiar with routes and better manage their time schedule during the day. Although attractive, the applicability of the SP approach is limited to the case in which the distribution of the demand is fully known. In addition, SP approaches suffer from the "*curse of dimensionality*," and thus they are often intractable.

Robust optimization (RO) is another approach for modeling uncertainty that assumes a complete ignorance about the distributions of uncertain parameters. Instead, RO assumes that uncertain parameters reside in an uncertainty set of possible outcomes with some structure (Bertsimas and Sim, 2004; Ben-Tal et al., 2015; Soyster, 1973). In RO, optimization is based on the worst-case scenario occurring within the uncertainty set. Notably, Lei et al. (2016) proposed the first and only two-stage RO approach for MF feet sizing and routing problem with demand uncertainty. Lei et al. (2016)'s model aims to find the MF fleet's size and routing decisions that minimize fixed cost of establishing the MF fleet (first stage), and a penalty cost for unmet demands (second stage). Optimization in Lei et al. (2016)'s RO model is based on the worst-case scenario of the demand (i.e., mean+ maximum positive deviation) occurring within a polyhedral uncertainty set. By focusing the optimization on the worst-case scenario, RO inevitably lead to overly conservative and suboptimal decisions for other more-likely scenarios and poor expected performance(Chen et al., 2020; Delage and Saif, 2018; Thiele, 2010).

Distributionally robust optimization (DRO) is another alternative approach for modeling uncertain problem data when their distributions are hard to characterize and so are subject to uncertainty (Rahimian and Mehrotra, 2019). In DRO, we assume that the distribution of random parameters resides in a so-called ambiguity set, and optimization is based on the worst-case distribution within this set. The ambiguity set is a family of all possible distributions of random parameters characterized by some known properties of random parameters (Esfahani and Kuhn, 2018). DRO is becoming an attractive approach to model the lack of distributional information of uncertain data in stochastic optimization problems because: (1) it alleviates the unrealistic assumption of the decision-makers' complete knowledge of the probability distribution governing the uncertain parameters, (2) it is usually more computationally tractable than its SP and RO counterparts (Delage and Saif, 2018; Rahimian and Mehrotra, 2019), and (3) one can use easy to approximate minimal distributional information (e.g., mean and range) to construct the ambiguity set and then build DRO models that better approximate reality and less conservative than RO models. Moreover, as Kuhn et al. (2019) pointed out, DRO faithfully anticipates the possibility of "*black swans*", i.e., mitigate the disappointing consequences (black swans) associated with implementing the SP or sample-based solutions under the true distribution or data that differ from all training samples. We refer to Rahimian and Mehrotra (2019) for a comprehensive survey of DRO literature.

Despite the potential advantages, there are no DRO approaches for the specific MF routing and scheduling problem that we study in this paper. Therefore, our paper is the first to propose a DRO approach for DMFRS. Our paper, however, builds on and uses similar techniques in recent DRO static FL literature (see, e.g., Basciftci et al. (2019); Luo and Mehrotra (2018); Saif and Delage (2020); Shehadeh and Sanci (2020); Shehadeh and Tucker (2020); Wang et al. (2020); Wu et al. (2015) and references therein). In particular, our decomposition-based algorithm for DMFRS is based on the same theory and art of cutting plane-based algorithms employed in Shehadeh and Sanci (2020). We also use the fact that we know the lower bound of the demand to derive valid lower bounding inequality. However, our master and sub- problems and bounding inequalities have a different structure than those of Shehadeh and Sanci (2020) due to the differences in decisions variables, objectives, and ambiguity set. Shehadeh and Sanci (2020) address a static FL problem with time-independent and bimodal demand (i.e., display two spatially distinct and ambiguous distributions). To model bimodality, they construct a scenario-wise ambiguity set and develop a DRO model that seeks to find the number of static facilities to open that minimize the fixed cost of opening facilities and the worst-case expectation of transportation and unmet demand cost.

In contrast to Shehadeh and Sanci (2020), we aim to determine MF routing and scheduling decisions and assume that demand distribution for MF services is unimodal, unknown, and time-dependent. Our ambiguity set is based on the mean, support, and mean absolute deviations of the demand in each period. Our objective is to optimize the system performance over a planning horizon. Thus, DMFRS is a different optimization problem than the one considered in Shehadeh and Sanci (2020).

Paper	Modeling Uncertainty	Optimization Approach	Decisions		Objectives		Symmetry breaking
			1st-stage	2nd-stage	1st-stage	2nd-stage	
Lei et al. (2016)	polyhedral uncertainty set based on maximum "positive" deviation from mean value	RO: optimization is based on worst-case scenario, i.e., mean+maximum positive deviation in each time interval	(1) size of MF fleet (2) flow of MFs over arcs in each interval	(1) amount of customer demand covered in each time interval (2) unmet demand in each interval	(1)fixed cost of establishing MF fleet	cost of unmet demand	
Our paper	ambiguity set: a family of all possible demand probability distributions with the same mean, MAD, and support	DRO: optimization is based on the worst-case distribution residing in ambiguity set, i.e., demand's distribution is a decision variable. The ambiguity set consists of all distributions that share the same mean, MAD, and support	(1) # of MFs to "activate" (2) routes and time schedule of active MFs	(1) amount of customer demand served by MFs in each time period (2) unmet demand in each period	(1) fixed cost of establishing MF fleet (2) traveling inconvenience cost	(1) cost of unmet demand (2) transportation or shipping cost	✓

Notation: RO is robust optimization, DRO is distributionally robust optimization, MF is mobile facility

Figure 1: Theoretical Comparison between Lei et al. (2016) and our work.

In contrast to RO model of Lei et al. (2016), we additionally incorporate the MF traveling inconvenience cost in the first-stage objective and the random transportation cost in the second-stage objective. Additionally, we optimize the system performance over all demand distributions residing within the ambiguity set. Our master and sub- problems and lower bound inequalities have a different structure than those of Lei et al. (2016) due to the differences in the decision variables and objectives. Finally, we propose two–families of symmetry-breaking inequalities, which break symmetries in the solution space of the first-stage routing and scheduling decisions. These inequalities are independent of the method of modeling uncertainty and thus can improve the solvability of any SP, DRO, RO, or deterministic formulation that uses the same routing and scheduling decisions in the first stage. Lei et al. (2014), Lei et al. (2016), and pioneering references therein did not address the issue of symmetry in deterministic MFRSP. In Figure 1, we present a theoretical comparison between Lei et al. (2016) and our work.

3. DMFRS Formulation and Analysis

In this section, we formally define DMFRS (Section3.1) and formulate it as a two-stage DRO model (Section 3.2). Then, we use duality theory, standard DRO reformulation techniques, and DMFRS properties to derive an equivalent solvable reformulation (Section 3.3).

3.1. ***Problem Statement***

As in Lei et al. (2014), we consider a fleet of M mobile facilities and define DMFRS on a directed network $G(V, E)$ with node set $V := \{v_1, \dots, v_n\}$ and edge set $E := \{e_1, \dots, e_m\}$. The sets $I \subseteq V$ and $J \subseteq V$ are the set of all customers points and the subset of nodes where MFs can be located, respectively. The distance matrix $D = (d_{i,j})$ is defined on E and satisfies the triangle inequality, where $d_{i,j}$ is a deterministic and time-invariant distance between any pair of nodes i and j.

We consider a planning horizon of T identical time periods, and we assume that the length of each period $t \in T$ is sufficiently short such that, without loss of generality, problem parameters are static (Halper and Raghavan, 2011; Lei et al., 2014, 2016). The demand level, $W_{i,t}$, of each customer i in each time period t is random and time-dependent. The probability distribution of the demand is unknown, and only the mean μ and range $[\underline{W}, \overline{W}]$ of W are known.

We consider the following basic features of the DMFRS as in Lei et al. (2014): (1) each MF has all the necessary service equipment and can move from one place to another, (2) all MFs are homogeneous, providing the same service, and traveling at the same speed, (3) we explicitly account for the travel time of the MF in the model, and service time are only incurred when the MF is not in motion, (4) the travel time $T_{j,j'}$ from location j to j' is an integer multiplier of a single time period (Lei et al. (2014, 2016), and (5) the amount of demand to be served is proportional to the duration of the service time at the location serving the demand.

We consider a cost f for using an MF, which represents the expenses associated with purchasing or renting an MF, staffing cost, equipment, etc. Each MF has a capacity limit C, which represents the amount of demand that an MF can serve in a single time unit. Due to the random fluctuations of demand and limited capacity, there is a possibility that the MF fleet fail to satisfy customers' demand fully. We consider a penalty cost γ for each unit of unmet demand. This penalty cost can represent, the opportunity cost for the loss of demands or expense for outsourcing the excess demands to other companies, among others (Basciftci et al., 2019; Lei et al., 2016).

Given that an MF cannot provide service when in motion, it is not desirable to keep it moving for a long time to avoid losing potential benefits. On the other hand, it is not desirable to keep the MF stationary all the time, as this may also lead to losing the potential benefits of making a strategic move to locations with higher demands. Thus, the trade-off of the problem includes the decision to move or keep the MF stationery. We consider a traveling inconvenience cost α to discourage unnecessary moving in cases where moving would neither improve or degrade the total performance. As in Lei et al. (2014), we assume that α is much lower than other costs such that its impact over the major trade-off is negligible.

We assume that the service quality a customer receives from an MF is inversely proportional to the distance between the two to account for the "access cost" (this assumption is common in practice and in the literature, see, e.g., Ahmadi-Javid et al. (2017); Reilly (1931); Drezner (2014); Lei et al. (2016); Berman et al. (2003); Lei et al. (2014)). Accordingly, we consider a demand assignment cost that is linearly proportional to the distance between the customer point and the location of an MF, i.e., $\beta d_{i,j}$, where $\beta \geq 0$ represents the assignment cost factor per demand unit and per distance unit.

Given a set of MF, M, we seek to find: (1) the number of MFs to use within T, (2) a routing plan for the selected MFs, i.e., routes that specify the movement of each MF, (3) a timetable of the MFs routing plan (i.e., time schedule), and (4) assignment plan from facility to customers for different realization of demand patterns. Decisions (1)–(3) are first-stage decisions that we make before realizing the demand patterns. The assignment decisions (4) represent the recourse actions in response to the first-stage decisions and the realizations of demand patterns (herein, from the

Table 1: Notation.

Indices	
m	index if MF, $m = 1, \ldots, M$
i	index of customer location, $i = 1, \ldots, I$
j	index of MF location, $j = 1, ..., J$
Parameters and sets	
M	number, or set, of MFs
J	number, or, set of locations
f	fixed operating cost
$d_{i,j}$	distance between any pair of nodes i and j
C	the amount of demand that can be served by an MF in a single time unit
$W_{i,t}$	demand at customer site i for each period t
$\underline{W}_{i,t}/\overline{W}_{i,t}$	lower/upper bound of demand at customer location i for each period t
γ	penalty of not satisfying demand
First-stage decision variables	
y_m	$\begin{cases} 1, & \text{if MF } m \text{ is permitted to use,} \\ 0, & \text{otherwise.} \end{cases}$
x^t_{jm}	$\begin{cases} 1, & \text{if MF } m \text{ stays at location } j \text{ at period } t, \\ 0, & \text{otherwise.} \end{cases}$
Second-stage decision variables	
$z^t_{i,j,m}$	amount of demand of customer point i being served by MF m located at j in period t
u_t	total amount of unmet demands by MFs in time period t

worst-case distribution of the demand). The quality of the MF routing and scheduling decisions are function of (1) first-stage fixed operation cost, and (2) the worst-case expectation of the recourse assignment cost and unsatisfied demand penalty cost.

Notation: For $a, b \in \mathbb{Z}$, we define $[a] := \{1, 2, \ldots, a\}$ and $[a, b]_{\mathbb{Z}} := \{c \in \mathbb{Z} : a \leq c \leq b\}$. The abbreviations "w.l.o.g." and "w.l.o.o." respectively represent "without loss of generality" and "without loss of optimality." For notation brevity, we use $(I,\ J,\ T)$ to denote both the number and set of (customer locations, MF locations, time periods). Table 1 summarizes other notation.

3.2. ***DMFRS Formulation***

We assume that demand distribution $\mathbb{P}$ is unknown but belongs to an ambiguity set of possible distributions, which incorporates demand's mean values, Mean Absolute Deviations (MADs), and support set. Here, we use MAD as the dispersion or variability measure because, as we will show later in Section 3.3, it enables a linear reformulation and hence is computationally attractive. Specifically, we let $\mu_{i,t} = \mathbb{E}_{\mathbb{P}}[W_{i,t}]$ and $\eta_{i,t} = \mathbb{E}_{\mathbb{P}}(|W_{i,t} - \mu_{i,t}|)$ respectively represent the mean value and MAD of the demand $W_{i,t}$ at node i in period t, for all $i \in [I]$ and $t \in [T]$. We let $\underline{W}_{i,t}$ and $\overline{W}_{i,t}$ represent the known lower bound and upper bound of $W_{i,t}$, for all $i \in [I]$ and $t \in [T]$.

Mathematically, we define the following support $\mathcal{S}$ of $\boldsymbol{W}$:

$$\mathcal{S} := \Big\{ W \geq 0 : \ \underline{W}_{i,t} \leq W_{i,t} \leq \overline{W}_{i,t}, \ \forall i \in [I], \ \forall t \in [T]. \ \Big\} \tag{1}$$

The probability distribution $\mathbb{P}$ of $\boldsymbol{W}$ belongs to the following ambiguity set

$$\mathcal{F}(\mathcal{S}, \mu, \eta) := \left\{ \mathbb{P} \in \mathcal{P}(\mathcal{S}) : \begin{array}{l} \int_S d\mathbb{P} = 1 \\ \mathbb{E}_{\mathbb{P}}[W_{i,t}] = \mu_{i,t}, \ \forall i \in [I], \ t \in [T] \\ \mathbb{E}_{\mathbb{P}}(|W_{i,t} - \mu_{i,t}|) \leq \eta_{i,t}, \ \forall i \in [I], \ t \in [T] \end{array} \right\} \tag{2}$$

Where $\mathcal{P}(S)$ in $\mathcal{F}(\mathcal{S}, \mu)$ represents the set of probability distributions supported on $\mathcal{S}$ and share the same mean value μ and dispersion measure η. For all $m \in [M]$, let binary variable $y_m = 1$ if MF m is permitted to use, and zero otherwise. For all $j \in [J]$, $m \in [M]$, and $t \in [T]$, let binary variable $x^t_{j,m} = 1$ if MF m stays at location j at period t. The feasible region $\mathcal{X}$ of variables x and y is defined in (3). As defined in Lei et al. (2014), region $\mathcal{X}$ represent: (1) the requirement that an MF can only be in service when it is stationary, (2) MF m at location j in period t can only be available at location $j' \neq j$ after a certain period of time depending on the time it takes to travel from location j to j', $T_{j,j'}$, i.e., $x^{t'}_{j,m} = 0$ for all $j' \neq j$ and $t' \in \{t, \ldots, \min\{t + T_{j,j'}, T\}\}$, and (3) MF m has to be in an active condition before providing service.

$$\mathcal{X} = \left\{ x, y : \begin{array}{l} x^t_{j,m} + x^{t'}_{j',m} \leq y_m, \quad \forall t, m, j, \ j \neq j', \ t' \in \{t, \ldots, \min\{t + T_{j,j'}, T\}\} \\ x^t_{j,m} \in \{0,1\}, \ y_m \in \{0,1\}, \ \forall j, m, t \end{array} \right\} \tag{3}$$

For all (i, j, m, t), let variable $z^t_{i,j,m}$ represents the amount of demand of customer point i being served by MF m located at j in period t. Let variable u_t represents total amount of unmet demands by MFs in time period t. Let parameter f represent a fixed operating cost needs to be paid for using an MF. Similar to Lei et al. (2014), we consider a demand assignment cost β, which is linearly proportional to the distance between the customer point and the location of an MF, and a penalty of not satisfying demand γ. Finally, we impose a capacity limit C on the amount of demand that can be served by an MF in a single time unit. Table 1 summarizes these notation. Using this notation and ambiguity set $\mathcal{F}(S, \mu, \eta)$, we formulate the DMFRS problem as:

$$\min_{(y,x) \in \mathcal{X}} \left\{ \sum_{m \in M} f y_m - \sum_{t \in T} \sum_{j \in J} \sum_{m \in M} \alpha x^t_{j,m} + \sup_{\mathbb{P} \in \mathcal{F}(S,\mu,\eta)} \mathbb{E}_{\mathbb{P}}[Q(y, x, W)] \right\} \tag{4a}$$

where for a feasible $(y, x) \in \mathcal{X}$ and a realization of uncertain demand $\xi := W$

$$Q(y, x, \xi) := \min_{z,u} \Big(\sum_{j \in J} \sum_{i \in I} \sum_{m \in M} \sum_{t \in T} \beta d_{i,j} z^t_{i,j,m} + \gamma \sum_{t \in T} \sum_{i \in I} u_{i,t} \Big) \tag{5a}$$

$$\text{s.t.} \quad \sum_{j \in J} \sum_{m \in M} z^t_{i,j,m} + u_{i,t} = W_{i,t}, \qquad \forall i \in [I], \ t \in [T] \tag{5b}$$

$$\sum_{i \in I} z^t_{i,j,m} \leq C x^t_{j,m} \qquad \forall j \in [J], \ m \in [M], \ t \in [T] \tag{5c}$$

$$u_t \geq 0, \ z^t_{i,j,m} \geq 0, \qquad \forall i \in [I], \ j \in [J], m \in [M], \ t \in [T] \tag{5d}$$

Formulation DMFRS in (4) searches for first stage decisions $(x, y) \in \mathcal{X}$ that minimize the first stage cost and the maximum expectation of the total second stage cost: (1) cost of assigning demands to the MFs, and (2) the penalty costs for unmet demands, over a family of distributions residing in the ambiguity set $\mathcal{F}(S, \mu, \eta)$. Constraints (5b) account for the amount of demand from each customer in each time period that is satisfied and the amount of demand that is not satisfied. Constraints (5c) respect the capacity of MF m. Finally, constraints (5d) specify feasible ranges of the decision variables. Our recourse formulation (5a)–(5d) is based on Lei et al. (2014)'s SP model (the only SP model for MFRSP).

3.3. ***Reformulation***

Our DMFRS model in (4) incorporates the decision of the probability distribution function, which is generally intractable and not solvable in the presented form. In this section, we use duality theory and exploit the characteristics of our recourse formulation and the ambiguity set, in deriving a solvable and tractable reformulation of our model. To do so, as in Wang et al. (2019), we define an epigraphical random variable $k_{i,t}$ for the MAD term $|W_{i,t} - \mu_{i,t}|$, for all $i \in [I]$ and $t \in [T]$, and rewrite the support set (1) and ambiguity set (2) as follows

$$\mathcal{S} := \left\{ (\boldsymbol{W}, \boldsymbol{k}) \geq 0 \;\middle|\; \begin{array}{l} W_{i,t} \in [\underline{W}_{i,t}, \overline{W}_{i,t}], \ \forall i \in [I], \ \forall t \in [T] \\ k_{i,t} \geq W_{i,t} - \mu_{i,t}, \ \forall i \in [I], \ \forall t \in [T] \\ k_{i,t} \geq \mu_{i,t} - W_{i,t}, \ \forall i \in [I], \ \forall t \in [T] \end{array} \right\} \text{ and} \tag{6}$$

$$\mathcal{F}(\mathcal{S}, \mu, \eta) := \left\{ \mathbb{P} \in \mathcal{P}(\mathcal{S}) \;\middle|\; \begin{array}{l} \int_S d\mathbb{P} = 1 \\ \mathbb{E}_{\mathbb{P}}[W_{i,t}] = \mu_{i,t}, \ \forall i \in [I], \ t \in [T] \\ \mathbb{E}_{\mathbb{P}}[k_{i,t}] \leq \eta_{i,t}, \ \forall i \in [I], \ t \in [T] \end{array} \right\} \tag{7}$$

We first consider the inner maximization problem $\sup\limits_{\mathbb{P} \in \mathcal{F}(S,\mu,\eta)} \mathbb{E}_{\mathbb{P}}[Q(y, x, W)]$ for a fixed $(y, x) \in \mathcal{X}$, where $\mathbb{P}$ is the decision variable. For a fixed $(y, x) \in \mathcal{X}$. we can explicitly write $\sup\limits_{\mathbb{P} \in \mathcal{F}(S,\mu,\eta)} \mathbb{E}_{\mathbb{P}}[Q(y, x, W)]$ as the following functional optimization problem.

$$\max_{\mathbb{P} \geq 0} \int_{\mathcal{S}} Q(y, x, W) \ d\mathbb{P} \tag{8a}$$

$$\text{s.t.} \int_{\mathcal{S}} W_{i,t} \ d\mathbb{P} = \mu_{i,t} \qquad \forall i \in [I], \ t \in [T] \tag{8b}$$

$$\int_{\mathcal{S}} k_{i,t} \ d\mathbb{P} \leq \eta_{i,t} \qquad \forall i \in [I], \ t \in [T] \tag{8c}$$

$$\int_{\mathcal{S}} d\mathbb{P} = 1 \tag{8d}$$

Where the probability distribution of W is the decision variable, i.e., we are choosing the distribution that maximizes $Q(y, x, W)$. As we show in the proof of Proposition 1 in Appendix A, the stochastic optimization problem in (8) is equivalent to problem (9).

Proposition 1. *For any fixed $(y,x) \in \mathcal{X}$, problem (8) is equivalent to*

$$\min_{\rho,\psi \geq 0} \left\{ \sum_{t\in T}\sum_{i\in I}(\mu_{i,t}\rho_{i,t} + \eta_{i,t}\psi_{i,t}) + \max_{(\boldsymbol{W},\boldsymbol{k})\in\mathcal{S}} \left\{ Q(y,x,W) + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t}) \right\} \right\} \tag{9}$$

Note that $Q(x,y,W)$ is a minimization problem, and thus in (9) we have an inner max-min problem, which is not suitable to use in standard solution methods. For a given solution (x,y) and realized value of W, $Q(x,y,W)$ is a linear program. We formulate $Q(x,y,W)$ defined in (5) in its dual form as

$$Q(y,x,W) = \max_{\lambda,v} \sum_{t\in T}\sum_{i\in I}\lambda_{i,t}W_{i,t} + \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} Cx^t_{j,m}v^t_{j,m} \tag{10a}$$

$$\text{s.t.} \quad \lambda_{i,t} + v^t_{j,m} \leq \beta d_{i,j}, \qquad \forall i \in [I], j \in [J], m \in [M], t \in [T] \tag{10b}$$

$$\lambda_{i,t} \leq \gamma, \qquad \forall i \in [I],\ t \in [T] \tag{10c}$$

$$v^t_{j,m} \leq 0, \qquad \forall j \in [J],\ t \in [T] \tag{10d}$$

where λ and v are the dual variables associated with constraints (5b) and (5c), respectively. It is easy to see that, w.l.o.o., $\lambda_{i,t} \geq 0$ for all $i \in [I]$ due to constraints (10c) and the objective of maximizing a positive and bounded variable times λ. Note that the objective of $Q(y,x,W)$ in (10a) is separable with respect to each time period t. In addition, region $\Omega := \{(10b) - (10d)\}$ is feasible and bounded and thus the optimal solution (λ^*, v^*) is an extreme point of Ω. Given that $W \in [\underline{W}_{i,t}, \overline{W}_{i,t}]$ by definition (see (6)), in view of dual formulation (10), we can rewrite $\max_{(\boldsymbol{W},\boldsymbol{k})\in\mathcal{S}} \left\{ Q(y,x,W) + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t}) \right\}$ in (9) as

$$\max_{\lambda,v,W,k} \left\{ \sum_{t\in T}\sum_{i\in I}\lambda_{i,t}W_{i,t} + \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} Cx^t_{j,m}v^t_{j,m} + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t}) \right\} \tag{11a}$$

$$\text{s.t.} \ (10b) - (10d), \quad W_{i,t} \in [\underline{W}_{i,t}, \overline{W}_{i,t}],\ \forall i \in [I],\ \forall t \in [T] \tag{11b}$$

$$k_{i,t} \geq W_{i,t} - \mu_{i,t},\ k_{i,t} \geq \mu_{i,t} - W_{i,t},\ \forall i \in [I],\ \forall t \in [T] \tag{11c}$$

Note that the objective function in (11) contains the interaction term $\lambda_{i,t}W_{i,t}$. To linearize formulation (11), we define $\pi_{i,t} = \lambda_{i,t}W_{i,t}$ for all $i \in [I]$ and $t \in [T]$. Also, we introduce the following McCormick inequalities for variables $\pi_{i,t}$:

$$\pi_{i,t} \geq \lambda_{i,t}\underline{W}_{i,t}, \qquad \pi_{i,t} \leq \lambda_{i,t}\overline{W}_{i,t}, \quad \forall i \in [I],\ \forall t \in [T] \tag{12a}$$

$$\pi_{i,t} \geq \gamma W_{i,t} + \overline{W}_{i,t}(\lambda_{i,t} - \gamma) \qquad \pi_{i,t} \leq \gamma W_{i,t} + \underline{W}_{i,t}(\lambda_{i,t} - \gamma),\ \forall i \in [I],\ \forall t \in [T] \tag{12b}$$

Thus, for a fixed $(y,x) \in \mathcal{X}$, problem (11) is equivalent to the following mixed-integer linear program

$$\max_{\lambda,v,W,\pi,k} \left\{ \sum_{t\in T}\sum_{i\in I} \pi_{i,t} + \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} Cx^t_{j,m}v^t_{j,m} + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t}) \right\} \tag{13a}$$

$$\text{s.t.} \quad (10\text{b}) - (10\text{d}),\ (12\text{a}) - (12\text{b}) \tag{13b}$$

$$W_{i,t} \in [\underline{W}_{i,t}, \overline{W}_{i,t}],\ k_{i,t} \geq W_{i,t} - \mu_{i,t},\ k_{i,t} \geq \mu_{i,t} - W_{i,t},\ \forall i \in [I],\ \forall t \in [T] \tag{13c}$$

Combining the inner problem in the form of (13) with the outer minimization problems in (9) and DMRS model in(4), we derive an equivalent reformulation of DMFRS as

$$\min \left\{ \sum_{m\in M} fy_m - \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} \alpha x^t_{j,m} + \sum_{t\in T}\sum_{i\in I} \Big[\mu_{i,t}\rho_{i,t} + \eta_{i,t}\psi_{i,t}\Big] + \delta \right\} \tag{14a}$$

$$\text{s. t.} \quad (y, x) \in \mathcal{X},\ \psi \geq 0 \tag{14b}$$

$$\delta \geq h(x, W) \tag{14c}$$

where $h(x, W) := \max_{\lambda,v,W,\pi,k} \Big\{ \sum_{t\in T}\sum_{i\in I} \pi_{i,t} + \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} Cx^t_{j,m}v^t_{j,m} + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t}) :$ $(10\text{b}) - (10\text{d}), , (12\text{a}) - (12\text{b}), (13\text{c}) \Big\}$.

Proposition 2. *For any feasible values of variables x and y, $h(x, W) < \infty$. Furthermore, $h(x, W)$ is a convex piecewise linear function in x (We refer to Appendix B for a detailed proof.)*

4. Solution Method

In this section, we present a decomposition-based algorithm to solve the two-stage DMFRS formulation in (14), and strategies to improve the solveability of the formulation. In Section 4.1, we present our DMFRS–Decomposition Algorithm. Then, in Section 4.2, we identify valid inequalities to improve the convergence of the algorithm. Finally, in Section 4.3, we derive two–families of symmetry–breaking inequalities to reduce symmetry in the solution space of first-stage routing and scheduling decisions of DMFRS.

*4.1. **DMFRS–Decomposition Algorithm***

Proposition 2 suggests that constraint (14c) describes the epigraph of a convex and piecewise linear function of decision variables in formulation (14). Therefore, given the two-stage characteristics of DMFRS in (14), it is natural to attempt to solve problem (14) (or equivalently, the DMFRS problem in (4)) via a separation-based decomposition algorithm. Algorithm 1 presents DMFRS–decomposition algorithm. Algorithm 1 is finite because we identify a new piece of the function $\max_{\lambda,v,W,\pi,k} \{ \sum_{t\in T}\sum_{i\in I} \pi_{i,t} + \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} Cx^t_{j,m}v^t_{j,m} + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t}\}$ each time when the set $\{L(x, \delta) \geq 0\}$ is augmented in step 4, and the function has a finite number of pieces according to Proposition 2.

Algorithm 1: DMFRS–decomposition algorithm.

1. Input. Feasible region $\mathcal{X}$; set of cuts $\{L(x,\delta) \geq 0\} = \emptyset$; $LB = -\infty$ and $UB = \infty$.

2. Master Problem. Solve the following master problem

$$Z = \min \left\{ \sum_{m \in M} f y_m - \sum_{t \in T} \sum_{j \in J} \sum_{m \in M} \alpha x^t_{j,m} + \sum_{t \in T} \sum_{i \in I} \Big[\mu_{i,t} \rho_{i,t} + \eta_{i,t} \psi_{i,t} \Big] + \delta \right\} \tag{15a}$$

$$\text{s. t.} \qquad (x,y) \in \mathcal{X}, \quad \psi_{i,t} \geq 0, \qquad \forall i \in [I], \ \forall t \in [T] \tag{15b}$$

$$\{L(x,\delta) \geq 0\} \tag{15c}$$

and record an optimal solution (x^*, ρ^*, ψ^*) and optimal value Z^*. Set $LB = Z^*$.

3. Sub-problem. With (x, ρ, ψ) fixed to (x^*, ρ^*, ψ^*), solve the following problem

$$h(x,W) = \max_{\lambda, v, W, \pi, k} \left\{ \sum_{t \in T} \sum_{i \in I} \pi_{i,t} + \sum_{t \in T} \sum_{j \in J} \sum_{m \in M} C x^t_{j,m} v^t_{j,m} + \sum_{t \in T} \sum_{i \in I} -(W_{i,t} \rho_{i,t} + k_{i,t} \psi_{i,t}) \right\} \tag{16a}$$

$$\text{s.t.} \quad (10\text{b}) - (10\text{d}),, (12\text{a}) - (12\text{b}), (13\text{c}) \tag{16b}$$

record optimal solution $(\pi^*, \lambda^*, W^*, v^*, k^*)$ and $h(x,W)^*$. Set
$UB = \min\{UB, \ h(x,W)^* + (LB - \delta^*)\}$

4. if $\delta^* \geq \sum_{t \in T} \sum_{i \in I} \pi^*_{i,t} + \sum_{t \in T} \sum_{j \in J} \sum_{m \in M} C x^{t*}_{j,m} v^{t*}_{j,m} + \sum_{t \in T} \sum_{i \in I} -(W^*_{i,t} \rho^*_{i,t} + k^*_{i,t} \psi^*_{i,t})$ **then**

stop and return x^* and y^* as the optimal solution to DMFRS formulation (15)

else add the cut $\delta \geq \sum_{t \in T} \sum_{i \in I} \pi^*_{i,t} + \sum_{t \in T} \sum_{j \in J} \sum_{m \in M} C x_{j,m} v^{t*}_{j,m} + \sum_{t \in T} \sum_{i \in I} -(W^*_{i,t} \rho_{i,t} + k^*_{i,t} \psi_{i,t})$ to the set of cuts $\{L(x,\delta) \geq 0\}$ and go to step 2.

end if

It is worth mentioning that Algorithm 1 is based on the same classical theory and art of cutting-plane and decomposition approaches, and particularly the same high-level steps of Shehadeh and Sanci (2020)'s decomposition algorithm. However, Shehadeh and Sanci (2020)'s decomposition algorithm is for a DRO model that address a static facility location problem with bimodal demand. Therefore, given the differences in the variables, constraints, objectives, and ambiguity set of DMFRS and those of Shehadeh and Sanci (2020) (see Section 2), our master and sub-problems, and the set of cuts, are entirety different than those of Shehadeh and Sanci (2020).

4.2. ***Multiple Optimality cuts and Lower Bound Inequalities***

In this section, we aim to incorporate more second-stage information into the first stage without adding optimality cuts into the master problem by exploiting the structural properties of the recourse problem. We first observe that once the first-stage solutions and the random demand are known, the second-stage problem can be decomposed into independent sub-problems with respect to time periods. Accordingly, we can construct cuts for each sub-problem in step 4. Let δ_t represent the optimality cut for each period t, we replace δ in (15a) with $\sum_t \delta_t$, where

$$\delta_t \geq \sum_{i\in I} \pi^*_{i,t} + \sum_{j\in J}\sum_{m\in M} Cx_{j,m} v^{t*}_{j,m} + \sum_{i\in I} -(W^*_{i,t}\rho_{i,t} + k^*_{i,t}\psi_{i,t}) \qquad \forall t \in [T] \tag{17}$$

The original single cut is the summation of multiple cuts of the form, i.e., $\delta = \sum_{t\in T} \delta_t$. Hence, in each iteration, we incorporate more or at least an equal amount of information into the master problem using (17) as compared with the original single cut approach. In this manner, the optimality cuts become more specific, which may result in better lower bounds and, therefore, a faster convergence. Inspired by the SP work of Lei et al. (2014) and using the fact that we know the minimum value of the demand, in Proposition 3, we identify valid inequalities for each time period to tighten the lower bound of the master problem (see Appendix C for a proof).

Proposition 3. *Inequalities* (18) *are valid lower bound inequalities for DMFRS.*

$$\delta_t \geq \sum_{i\in I} \min\{\gamma, \min_{j\in J}\{\beta d_{i,j}\}\}\underline{W}_{i,t}, \qquad \forall t \in [T] \tag{18}$$

4.3. ***Symmetry-Breaking Inequalities***

In this section, we derive two–families of symmetry–breaking inequalities to reduce symmetry in the solution space of first-stage decisions of DMFRS. First, recall that the MFs are homogeneous. As such, solutions $y = [1, 1, 0]^\top$, $y = [0, 1, 1]^\top$, and $y = [1, 0, 1]^\top$ are equivalent (i.e., yield the same objective) in the sense that they all permit 2 out of 3 MFs to be used in the planning period. To avoid wasting time exploring such alternative and equivalent solutions, we assume that MFs are numbered sequentially and add the following inequalities to the master problem in (15).

$$y_{m+1} \leq y_m \qquad \forall m < M. \tag{19}$$

enforcing arbitrary ordering or activation of MFs. Second, recall that in the first period, $t = 1$, we decide the initial locations of the MFs. Therefore, it doesn't matter which MF to assign to location j. For example, suppose that we have 3 locations, and MFs 1 and 2 are active. Then, feasible solutions ($x^1_{1,1} = 1$, $x^1_{3,2} = 1$) and ($x^1_{1,2} = 1$, $x^1_{3,1} = 1$) yield the same objectives. To avoid exploring such equivalent solutions, we define a dummy location $J + 1$ and add the following inequalities to the master problem in (15).

$$x^1_{j,m} - \sum_{j'>j}^{J+1} x^1_{j',m+1} \leq 0, \qquad \forall m < M, \forall j \in [J] \tag{20a}$$

$$x^1_{J+1,m} = 1 - \sum_{j=1}^{J} x^1_{j,m}, \qquad \forall m \in [M]. \tag{20b}$$

Inequalities (19)–(20) are independent of the method of modeling uncertainty, and so they are valid for the SP, RO, and deterministic formulations, i.e., these inequalities are valid for any formulation that uses the same sets of routing and scheduling decisions.

Table 2: DMFRS instances. Notation: I is # of customers, J is # of locations, and T is # of periods.

Inst	$\boldsymbol{I}$	$\boldsymbol{J}$	$\boldsymbol{T}$
1	10	10	10
2	10	10	20
3	15	15	10
4	15	15	20
5	20	20	10
6	20	20	20
7	25	25	10
8	25	25	20
9	30	30	10
10	30	30	20
11	40	40	10

5. Computational Experiments

The primary objective of our computational study is to quantify the benefit of using a DRO approach for the plain MFRSP as compared to the SP approach (see Appendix D for the formulation). The SP minimizes the first stage operation costs plus the expected cost of recourse activity via the the sample average approximation (SAA) approach (see, e.g., Kim et al. (2015); Kleywegt et al. (2002) for a detailed discussion on SAA). Section 5.1 presents the details of data generation and experimental design. In Section 5.3, we demonstrate efficiency of the lower bound and symmetry breaking inequalities (18)–(20). In Section 5.4, we compare the out-of-sample performance of the optimal solutions of the DR and SP. We close by analyzing the sensitivity of the DRO model to different parameter settings in Section 5.5.

*5.1. **Experimental Setup***

Due to lack of data, we construct 11 DMFRS instances, in part based on the same parameters settings and assumption made in Lei et al. (2014) and Lei et al. (2016), who address MF facility location problems. We summarize our test instances in Table 2. Each of the 11 DMFRS instances is characterized by the number of customers locations I, number of candidate locations J, and the number of periods T. Instances 1–4 are from Lei et al. (2014) and instances 5–10 are from Lei et al. (2016). Instance 11 is a hypothetical instance.

For each DMFRS instance, we generate a total of I vertices as uniformly distributed random numbers on a 100 by 100 plane and compute the distance between nodes in Euclidean sense as in Lei et al. (2014). We follow the same procedures in the DR scheduling and facility location literature (e.g., Jiang et al. (2017); Shehadeh and Sanci (2020); Shehadeh and Tucker (2020); Mak et al. (2014)) to generate random parameters as follows. For most instances, we randomly generate the mean values $\mu_{i,t}$ of the stochastic demand of each customer i in period t from a uniform distribution $U[\underline{W}, \overline{W}] = [20, 60]$ and the standard deviation $\sigma_{i,t} = 0.5\mu_{i,t}$. We set MAD $\eta_{i,t} = \max(\mu_{i,t} - \underline{W}_{i,t}, \overline{W}_{i,t} - \mu_{i,t})$, for all $i \in [I]$ and $t \in [T]$. We conduct sensitivity analysis for these parameters in Section 5.5.

We sample N realizations $W^n_{i,t}, \ldots, W^n_{I,T}$, for $n = 1, \ldots, N$, by following lognormal (LogN) distributions with the generated means $\mu_{i,t}$ and standard deviations $\sigma_{i,t}$. Kamath and Pakkala

(2002) results suggest that the LogN is a suitable distribution for modeling stochastic demands in the economic context, and Lei et al. (2014) use LogN distribution to model W. We round each random parameter to the nearest integer. We optimize the SP by using all of the N data points, and the DR model with the corresponding mean, lower bound $\underline{W}$, and upper bound $\overline{W}$. Lei et al. (2014) suggest that $N = 200$ is an adequate sample-size to obtain near-optimal solutions using the SP of MFRSP. Accordingly, we report results using $N = 200$. We set the travel time and number of MFs as in Lei et al. (2014).

We assume that all of the cost parameters are calculated in terms of present monetary value. Specifically, as in Lei et al. (2014), for each instance we randomly generate (1) the fixed cost from a uniform distribution $U[a, b]$ with $a = 1000$ and $b = 1500$, (2) the assignment cost factor per unit distance per unit demand β from $U[0.0001a, 0.0001b]$, and (3) the penalty cost per unit demand γ form $U[0.01a, 0.01b]$. Finally, we set the traveling inconvenience cost factor to $0.0001a$, and unless stated otherwise, we use a capacity parameter $C = 100$.

We implemented the SP model, DRO model, and DMFRS–decomposition algorithm using AMPL2016 Programming language calling CPLEX V12.6.2 as a solver with default settings. We run all experiments on MacBook Pro with Intel Core i7 processor, 2.6 GHz CPU, and 16 GB (2667 MHz DDR4) of memory. We imposed a solver time limit of 1 hour. To ensure fairness, the parameters of instances tested among the SP and DRO approaches are all the same.

5.2. ***CPU Time***

In this section, we first compare solution times of SP and DR models. In addition to the base-case settings $(W \in [\underline{W}, \overline{W}], C) = ([20, 60], 100)$, we consider $([50, 100], 100)$. For each of the 11 DMFRS instances and parameter settings, we randomly generate 5 instances as described in Section 5.1 for a total of 110 SAAs and DR instances. We solve each instance using the SAA formulation and our DMFRS–decomposition algorithm.

First, we present details of solving the 110 DMFRS instances using our DMFRS-decomposition algorithm. Table 3 presents the minimum (Min), average (Avg), and maximum (Max) CPU time and number of iterations taken to solve these instances via the DMFRS–decomposition algorithm. Results marked with $*$ are obtained with a relaxed tolerance level (relative gap between lower and upper bound) of ϵ =2% in Algorithm 1 (i.e., we obtain near-optimal solutions).

From Table 3, we first observe that the computational effort (i.e., solution time per iteration) tend to increase as the scale of the instances increases. The average solution times of DMFRS instances 1–7 ranges from 1 seconds to 10 minutes. The average solution times of larger instances, instances 8–10, ranges from 8 to 23 minutes. Note that when solving instances 8–10 and instance 11 with $\epsilon = 0$, the optimality gap remain at 2% even after several hours. This could indicate that CPLEX finds a good integer solution early, but examine many additional nodes to prove optimality. In reality, practitioners seek to find near-optimal solutions that can guarantee good

Table 3: Computational details of solving DMFRS model. Results marked with $*$ are obtained with $\epsilon = 2\%$.

Inst	I	T	$W \in [20, 60]$, $C = 100$						$W \in [50, 100]$, $C = 100$					
			CPU time			iteration			CPU time			iteration		
			Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
1	10	10	1	4	6	3	12	16	2	2	3	3	4	4
2	10	20	4	10	13	5	13	18	3	5	7	3	4	5
3	15	10	20	30	51	9	15	24	9	14	19	4	5	7
4	15	20	29	35	44	8	40	122	26	42	55	4	5	7
5	20	10	39	74	147	9	26	69	5	55	82	4	5	6
6	20	20	147	352	872	11	30	75	61	85	134	3	4	4
7	25	10	267	624	933	22	47	60	58	140	249	4	6	9
8	25	20	802	1253	1790	12	23	40	512	1024	1776	4	5	7
9*	30	10	277	1259	1802	12	27	33	237	456	737	5	6	8
10*	30	20	133	1354	1775	2	13	38	500	1246	1631	1	4	6
11*	40	10	924	1612	1812	11	17	22	500	1353	2830	1	2	3

Table 4: Comparison of solution times (in seconds) using the SAA and DMFRS formulations

Inst	I	T	$W \in [20, 60]$, $C = 100$						$W \in [50, 100]$, $C = 100$					
			SP			DMFRS			SP			DMFRS		
			Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
1	10	10	122	156	203	1	4	6	157	266	355	2	2	3
2	10	20	838	918	1134	4	10	13	487	781	1875	3	5	7
3	15	10	629	1723	3604	20	30	51	2188	2618	3605	9	14	19

operational performance within a reasonable time. Thus, it may make sense to relax ϵ to 2% for large instances.

In contrast, using the SP–SAA formulation, we were able to solve 40 out of the 110 SAAs instances to optimality within 1 hour, namely, all of the SAAs that correspond to instances 1-4, while the remaining instances terminated with a large optimality gap ($\geq 40\%$). For brevity and illustrative purposes, in Table 4, we compare the Min, Avg, and Max SP and DRO solution times (in seconds) of instances 1–3. From Table 4, we observe that solution times tend to increase as the scale of instances increases. The SP takes a substantially longer time to solve each instance. Solution times of the SP ranges from 122 to 3600 seconds seconds, and those of the DRO model ranges from 1 to 51 seconds. Finally, it is worth mentioning that, using an *enhanced multicut L-shaped* (E-LS) method to solve their SP, Lei et al. (2014) are able to solve instance 1–4. The average solution time of instance 4 using E-LS is 3000 seconds obtained at 5% optimality gap.

5.3. ***Efficiency of Inequalities*** (18)–(20)

In this Section, we investigate the importance of symmetry breaking inequalities (19)–(20) and lower bound inequalities (18). Given the challenges of solving DMFRS instances without these inequalities, we use Instance 1 and Instance 2 under $W \in [20, 60]$ in this experiment for illustrative

Table 5: Average solution times (seconds) with and without (w/o) symmetry-breaking (SB) inequalities (19)-(20). $W \in [20, 60]$

			SP		DRO	
Ins	I	T	With SB Ineq.	w/o SB Ineq.	With SB Ineq.	w/o SB Ineq.
1	10	10	156	2491	6	280
2	10	20	918	3868	10	1407

purposes and brevity.

First, we investigate the computational advantage of adding symmetry-breaking inequalities (19)-(20) to the SP and DRO model. To do so, we separately solve the SP and DRO model with and without (w/o) these symmetry-breaking inequalities (SB Ineq.). Table 5 compare solution times of instances 1-2 with and w/o SB Ineq. Not surprisingly, inequalities (19)-(20) substantially improve SP and DRO's solvability. These results demonstrate the importance of breaking the symmetry in the routing and scheduling decisions and the effectiveness of our SB Ineq (19)-(20).

Next, we analyze the impact of including the lower bounding (LB) inequalities (18) in the master problem of the DMFRS–decomposition algorithm. We first note that the DRO model is very challenging to solve, and thus the algorithm takes hundreds of iteration and a longer time until convergence without these LB inequalities. Therefore, for illustrative purposes, in Figure 2 and 3, we respectively compare the LB and gap (i.e., relative difference between the upper and low bounds on the objective value) values from the first 20 iterations of the algorithm with and w/o LB inequalities (18).

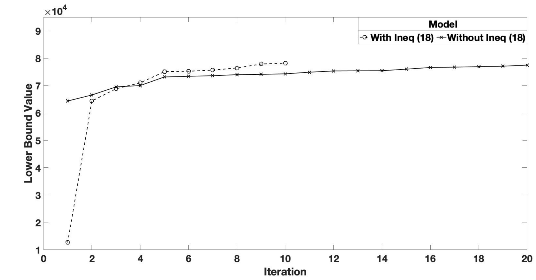
It is easy to see from Figure 2 and 3 that both the lower bound and gap values converge faster when we introduce inequalities (18) into the master problem. The differences in the lower bound and gap values are significant, especially in the first few iterations of the algorithm. Moreover, because of the better bonding effect, the algorithm converges to the optimal solution in fewer iterations and shorter solution times. For example, with these inequalities, the algorithm takes 6 seconds and 10 iterations to converge to the optimal solution of DMFRS instance 1 as compared to an hour and $\sim$ 800 iterations without these inequalities.

The results in this section demonstrate the importance and efficiency of inequalities (18)–(20).

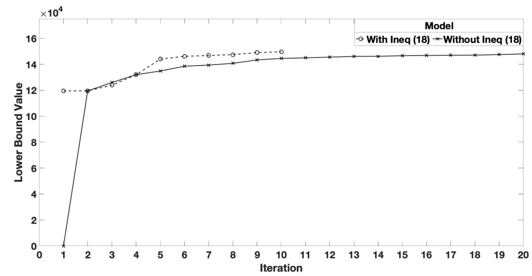
5.4. *DRO and SP solutions quality*

In this section, we compare how the DRO and SP models perform when the underlying uncertainty distributions are perfectly specified and misspecified. We also compare the optimal solutions performances with those of a moment-based DRO model based on mean values and range of the demand. We denote the moment based and our DRO model (which is based on mean, range, and MADs) as DRO-M and DRO-MAD, respectively.

Given that the SP cannot solve terminate with a large gap for most of the DMFRS instances, for fair comparison and brevity, we use instance 2 in this experiment (recall that the SP can solve this

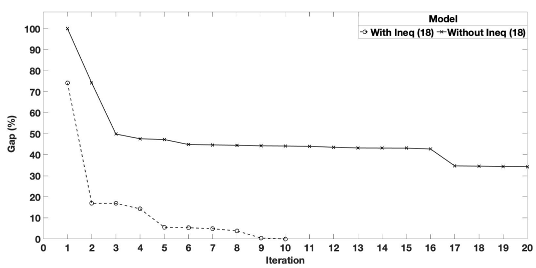


(a) Instance 1

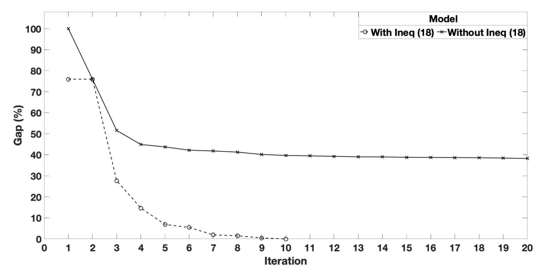


(b) Instance 2

Figure 2: Comparisons of lower bound value with and without inequalities (18).



(a) Instance 1



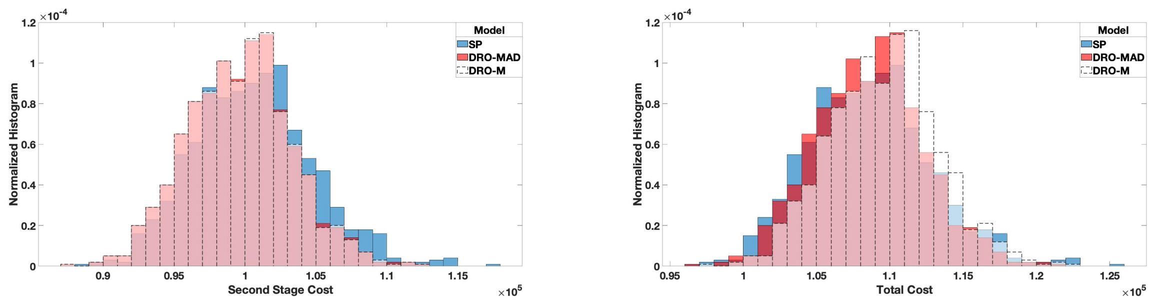
(b) Instance 2

Figure 3: Comparisons of gap values with and without inequalities (18).

instance to optimality). We first note that, in their optimal solutions, the DRO-M, DRO-MAD, and SP models activate 10, 9, and 8 MFs, respectively, i.e., 10, 9, 8 MFs are respectively chosen to serve customers during the planning horizon. Next, we analyze how these optimal solutions perform via in-sample and out-of-sample simulation. Specifically, first, we fix the optimal first-stage decisions (y, x) yielded by each model in the SP model. Then, we solve the second-stage recourse problem of the SP using the following two sets of $N' = 10,000$ out-of-sample data of $W_{i,t}^n$, for all $i \in [I], t \in [T]$, and $n \in [N']$, to compute the corresponding second-stage transportation and unmet demand costs.

Set 1. *Perfect distributional information*: we use the same parameter settings and distributions that we use for generating the N in-sample data points to generate the N' data points. This is to simulate the performance under perfect distributional information, i.e., the true distribution is the same as the one used in optimizing the SP (Logn).

Set 2. *Misspecified distributional information*: we vary the distribution type of random parameters to generate the $N' = 10,000$ samples as follows. First, we simulate the optimal solutions using N′ samples from a Normal distribution with the same mean values and standard deviations as in the in-sample data (i.e., we vary the distribution type). Second, we follow the same out-of-sample simulation testing procedure described in Wang et al. (2020) and employed in Shehadeh and Tucker (2020). That is, we perturb the distribution of the demand by a



(a) Second stage cost

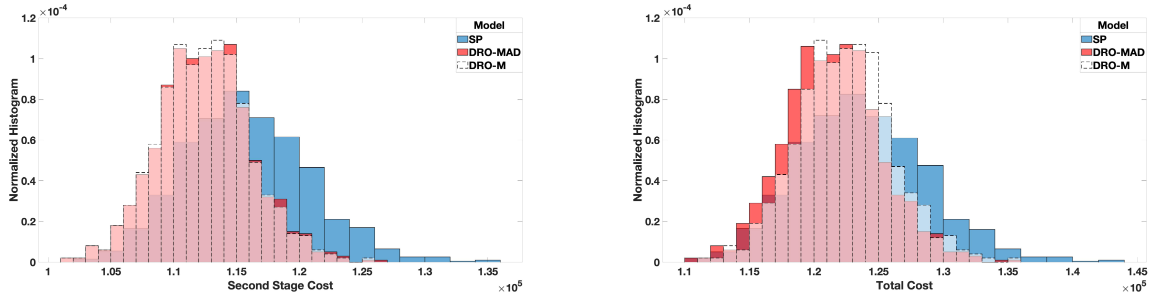
(b) Total cost

Figure 4: Comparison of out-of-sample simulation results under perfect distributional information

parameter Δ and use a parameterized uniform distribution $U[(1-\Delta)W, (1+\Delta)W\ ]$, where a higher value of Δ corresponds to a higher variation level. We apply $\Delta \in \{0, 0.25, 0.5\}$. A zero value of Δ indicates that we only vary the demand distribution, i.e., we do not perturb the demand's distribution and simulate the optimal solutions under a uniform distribution defined on the orange of the demand. We generate $N' = 10,000$ samples from these uniform distribution to evaluate the performance of the optimal solutions obtained from the DRO-M, DRO-MAD, and SP models. This is to simulate the models' optimal decisions when the true distributions are different from the in-sample distribution.

Figure 4 present the normalized histogram of out-of-sample total cost and second-stage cost with perfect distributional information (under Set 1). From this figure, we observe that there is no clear winner. That is, when the decision-maker has perfect knowledge about the demand's distribution, the three models yield similar performance. However, the SP has a slightly lower total cost because it chooses less number of MF's to operate during the planning horizon (8 versus 10 and 9 MFs). The DRO-M model yields a slightly higher average total cost and upper quantiles than the DRO-MAD and SP because it activates more MF's and thus yields higher fixed operating costs. It is not surprising that the SP's total cost is slightly lower than that of the DRO model since we assume perfect information of the exact demand distribution in this simulation.

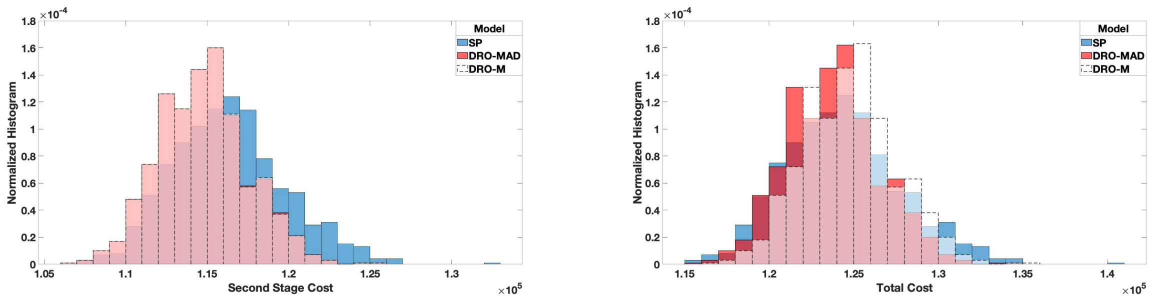
Figures 5–8 present the results from misspecified distributions. These figures show the normalized histogram of out-of-sample total cost and second-stage cost for each choice of variation, Δ. It is quite apparent that the DRO solutions consistently outperform the SP solutions under all levels of variation, $\Delta \in \{0, 0.25, 0.5\}$, and across the criteria of mean and quantiles of total and second stage costs. In particular, the DRO models has substantially lower unmet demand costs than the SP. For example, when $\Delta = 0$, the average unmet demand of SP, DRO-M, and DRO-MAD models are 3035 and 3, and 0, respectively. By activating a fewer number of MF's the DRO-MAD (9 versus 10 MFs) has a lower total cost than the DRO-M (which is more conservative). However, when $\Delta = 0.5$ (i.e., under high variations in the demand), the DRO-M yields a lower second-stage



(a) Second stage cost

(b) Total cost

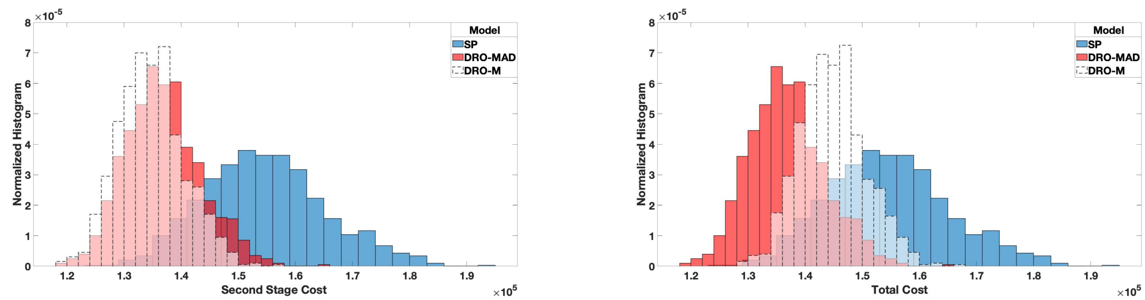
Figure 5: Comparison of out-of-sample simulation results under missspecified distribution, Normal



(a) Second stage cost

(b) Total cost

Figure 6: Comparison of out-of-sample simulation results under missspecified distribution, $\Delta = 0$



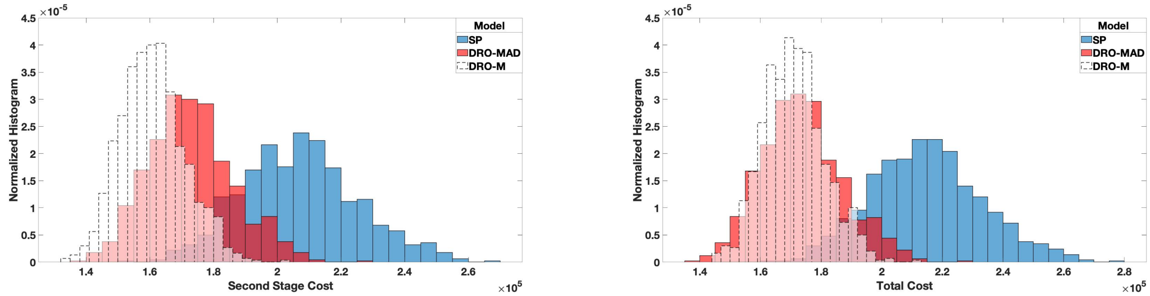
(a) Second stage cost

(b) Total cost

Figure 7: Comparison of out-of-sample simulation results under missspecified distribution, $\Delta = 0.25$

cost than DRO-MAD (see Figure 8).

Finally, we note that the DRO models performances are more stable in attaining the lowest standard deviations (i.e., variations) in the second-stage and total costs. The DRO model's superior performance, which focuses on hedging against distributional ambiguity, reflects the value of modeling the distributional ambiguity of random parameters in DMFRS. Moreover, these results show the advantages of using the computationally efficient DRO approach compared to the challenging SP approach.



(a) Second stage cost

(b) Total cost

Figure 8: Comparison of out-of-sample simulation results under missspecified distribution, $\Delta = 0.5$

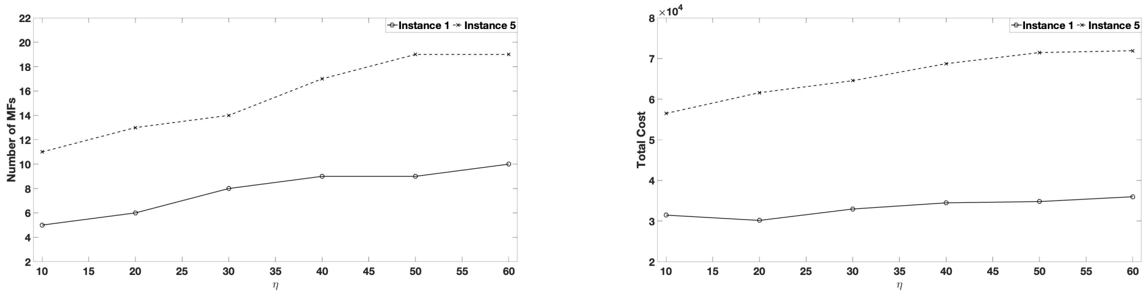
5.5. ***Sensitivity Analysis***

In this section, we study the sensitivity of DR model to different parameter settings. Given that we observe similar results for all of the 11 DMFRS instances, for presentation brevity and illustrative purposes, we present results for instance 1 (I, J, T)=(10, 10, 10) and instance 5 (I, J, T)=(20, 20,10) as examples of small and relatively average DMFRS instances.

First, we analyze the optimal number of active/scheduled MFs under different values of η (MAD or dispersion measure). We fix all parameters as described in Section 5.1 and solve our DRO model with $\eta \in \{10, 20, 30, 40, 50, 60\}$. Figure 9 presents the optimal number of active MFs and the associated total cost under these values of η. It is quite apparent that the number of active MFs, and thus the associated fixed cost increase as η increases. These results make sense and suggest that our DRO model tends to mitigate the increased variability of demand by scheduling a larger number of MFs.

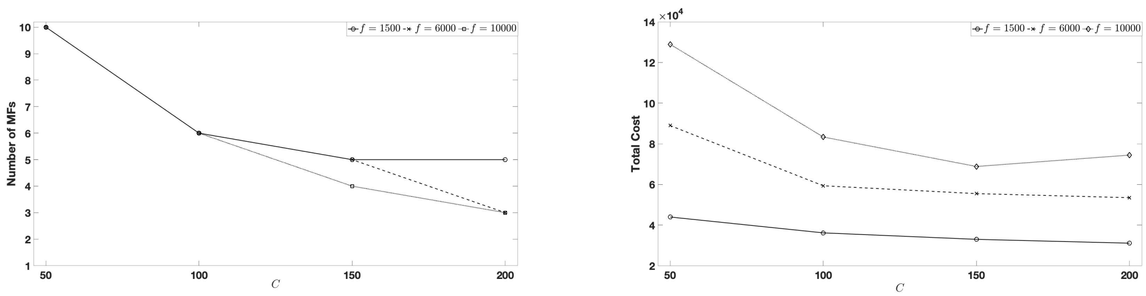
Second, we study the effect of the fixed cost, f, and MF capacity, C. We fix all parameters as described in Section 5.1 and solve our DRO model with $C \in \{50, 100, 150, 200\}$ and $f = 1,500$ (low), $f = 6,000$ (average), and $f = 10,000$ (high). From Figures 10–11, we observe that irrespective of f, the optimal number of scheduled MFs decreases as the capacity of the MFs increases, which makes sense because, with higher capacity, each MF can serve a larger amount of customer demand in each period. The total cost also decreases as C increases, which makes sense because, with higher capacity, we schedule fewer MFs (i.e., which reduces the fixed cost), and each of the scheduled MFs can serve a larger amount of demand. Finally, we note that as f increases, the number of scheduled MFs decreases irrespective of the capacity.

Third, we fix $M = 20$ and solve the DRO model with unmet demand penalty $\gamma \in \{0.10\gamma_o, 0.25\gamma_o, 0.35\gamma_0, 0.50\gamma_o\}$ (where γ_o is the base case penalty in Section 5.1) and $f \in \{1,500, 6,000, 10,000\}$. For brevity, in Figure 12, we present the number of MFs and unmet demand cost for instance 5. It is not surprising that as γ increases (i.e., satisfying customer demand becomes more important), the number of scheduled MFs increases (Figure 12a), and accordingly, a larger amount of customer demand is satisfied, i.e., unmet demand cost decrease (Figure 12b).



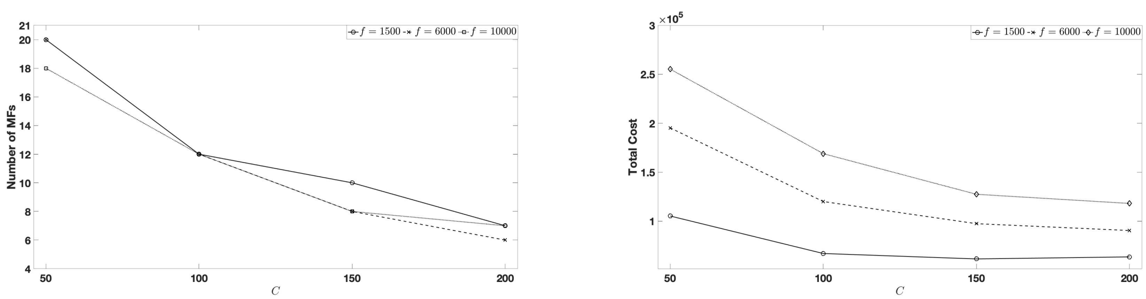
(a) Number of Active MFs (b) Total cost

Figure 9: Comparison of the results for different values of η



(a) Number of MFs (b) Total cost

Figure 10: Comparison of the results for different values of f and C. Instance 1



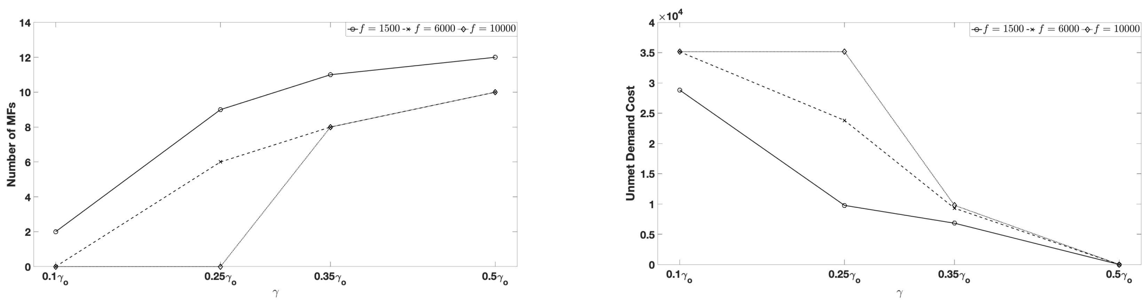
(a) Number of MFs (b) Total cost

Figure 11: Comparison of the results for different values of f and C. Instance 5

Our experiments in this section provide an example of how decision-makers can use our computationally efficient DRO approach to generate DMFRS solutions under different parameter settings.

6. Conclusion

In this paper, we study a DMFRS problem, which seeks optimal routing and scheduling decisions for a fleet of MFs to minimize the fixed operating cost and worst-case expected cost generated dur-



(a) Number of MFs

(b) Unmet demand cost

Figure 12: Comparison of the results for different values of γ and f

ing the planning horizon. We take the worst-case expectation over an ambiguity set characterized by the known mean, MAD, and random demand range. Using duality theory, we derive an equivalent mixed-integer non-linear programming reformulation of DMFRS. We linearize and propose a decomposition algorithm to solve the reformulation. We also derive lower bound and two-families of symmetry breaking inequalities to strengthen the master problem and speed up convergence.

To justify the value of a DRO approach for MFRSP, we conduct extensive numerical experiments. Our results demonstrate that our approach is computationally efficient and able to solve instances with up to 40 customers and 10 times periods in a reasonable time compared to the SP approach, which can only solve instances with 15 customers and 10 periods at a significantly larger time than our DRO approach. Considering the scale of the problem in the sense of a static facility location problem, our DRO approach can solve instances of $30 \times 20 = 600$ customers (instance 10), which are relatively large for some practical applications. Our results also show the optimal DRO solutions' superior performance under different probability distributions of random demand compared to the optimal SP solutions. Additionally, our results show that our lower bound and symmetry breaking inequality can tighten the lower bound during the solution process and lead to faster convergence of the decomposition algorithm. Finally, we demonstrate how decision-makers can use our computationally efficient approach to study the trade-offs between cost, number of MFs, and MF's capacity. More broadly, our paper draws attention to the need to consider the distributional ambiguity of the demand in strategic real-world stochastic optimization problems.

We suggest the following areas for future research. First, our results are based on assumptions and parameter settings made in prior studies and assume that we know the capacity and the size of the MF fleet. We aim to extend our model to optimize the capacity and size of the MF fleet. Second, we assume that the distribution of the demand is unimodal. We want to extend our approach by incorporating multi-modal probability distributions and higher moments of the demand in a data-driven DR approach based on emergent ambiguity set such as distance-based and scenario-wise ambiguity sets. Third, we aim to incorporate the uncertainty of MF travel time in a DR model. Finally, it would be theoretically interesting to investigate exact methods to improve a

DRO approach's computational performance for stochastic MFRS problems.

Acknowledgment

We want to thank all colleagues who have contributed significantly to the related literature. Dr. Karmel S. Shehadeh dedicates her effort in this paper to every little dreamer in the whole world who has a dream so big and so exciting. Believe in your dreams and do whatever it takes to achieve them–the best is yet to come for you

Appendix A. Proof of Proposition 1

Letting $\rho_{i,t}, \psi_{i,t}$ and θ be the dual variables associated with constraints (8b), (8c), (8d), respectively, we present problem (8)) in its dual form:

$$\min_{\rho,\theta,\psi\geq 0} \sum_{t\in T}\sum_{i\in I}(\mu_{i,t}\rho_{i,t} + \eta_{i,t}\psi_{i,t}) + \theta \tag{A.1a}$$

$$\text{s.t.} \sum_{t\in T}\sum_{i\in I}(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t}) + \theta \geq Q(y,x,W) \qquad \forall(\boldsymbol{W},\boldsymbol{k}) \in \mathcal{S} \tag{A.1b}$$

where ρ and θ are unrestricted in sign, $\psi \geq 0$, and constraint (A.1b) is associated with the primal variable $\mathbb{P}$. Under the standard assumptions that $\mu_{i,t}$ lies in the interior of the set $\{\int_{\mathcal{S}} W_{i,t}\ d\mathbb{Q} : \mathbb{Q}$ is a probability distribution over $S\}$, strong duality hold between (8) and (A.1) (Bertsimas and Popescu (2005); Jiang et al. (2017); Mak et al. (2014); Shehadeh et al. (2019)). Note that for fixed (ρ,θ,ψ), constraint (A.1b) is equivalent to $\theta \geq \max_{W\in\mathcal{S}}\{Q(y,x,W) + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t})\}$. Since we are minimizing θ in (A.1), the dual formulation of (8) is equivalent to:

$$\min_{\rho,\psi\geq 0} \left\{\sum_{t\in T}\sum_{i\in I}(\mu_{i,t}\rho_{i,t} + \eta_{i,t}\psi_{i,t}) + \max_{W\in\mathcal{S}}\left\{Q(y,x,W) + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t})\right\}\right\}$$

Appendix B. Proof of Proposition 2

First, note that the feasible regions $P := \{(10b) - (10d), , (12a) - (12b), (13c)\}$ and $\mathcal{S}$ are independent of x and bounded. In addition, DMFRS has a complete recourse (i.e., feasible for any feasible $(x,y) \in \mathcal{X}$) Therefore, $\max_{\lambda,v,W,\pi,k}\Big[\sum_{t\in T}\sum_{i\in I}\pi_{i,t} + \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} Cx^t_{j,m}v^t_{j,m} + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t})\Big] < \infty$. Second, for any fixed π, v, W, k, $\Big[\sum_{t\in T}\sum_{i\in I}\pi_{i,t} + \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} Cx^t_{j,m}v^t_{j,m} + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t})\Big]$ is a linear function of x. It follow that $\max_{\lambda,v,W,\pi,k}\Big[\sum_{t\in T}\sum_{i\in I}\pi_{i,t} + \sum_{t\in T}\sum_{j\in J}\sum_{m\in M} Cx^t_{j,m}v^t_{j,m} + \sum_{t\in T}\sum_{i\in I} -(W_{i,t}\rho_{i,t} + k_{i,t}\psi_{i,t})\Big]$ is the maximum of a linear functions of x, and hence convex and piecewise linear. Finally, it is easy to see that each linear piece of this function is is associated with one distinct extreme point of polyhedra P and S. Given that each of these polyhedra has a finite number of extreme points, the number of pieces of this function is finite. This completes the proof.

Appendix C. Proof of Proposition 3

Recall from the definition of the ambiguity set that the lowest demand of each customer i in period t equals to the integer parameter $\underline{W}_{i,t}$. Now, if we treat the MFs as uncapacitated facilities, then we can fully satisfy $\underline{W}_{i,t}$ at the lowest assignment cost from the nearest location $j \in J'$, where $J' := \{j : x^t_{j,m} = 1\}$. Note that $J' \subseteq J$. Thus, the lowest assignment cost, δ_t, must be at least equal to or larger than $\sum_{i \in I} \min_{j \in J}\{\beta d_{i,j}\}\underline{W}_{i,t}$. If $\gamma < \min_{j \in J}\{\beta d_{i,j}\}$ then δ_t must be at least equal to or larger than $\sum_{i \in I} \gamma \underline{W}_{i,t}$. Accordingly, $\delta_t \geq \sum_{i \in I} \min\{\gamma, \min_{j \in J}\{\beta d_{i,j}\}\}\underline{W}_{i,t}$ is a valid lower bound $\delta_t,\ \forall t \in [T]$. This complete the proof.

Appendix D. Stochastic Mixed-Integer Linear Program

$$\min \Big[\sum_{m \in M} f y_m - \sum_{t \in T} \sum_{j \in J} \sum_{m \in M} \alpha x^t_{j,m} + \frac{1}{N} \sum_{n=1}^{N} \Big(\sum_{j \in J} \sum_{i \in I} \sum_{m \in M} \sum_{t \in T} \beta d_{i,j} z^{t,n}_{i,j,m} + \gamma \sum_{t \in T} \sum_{i \in I} u^n_{i,t} \Big) \Big] \tag{D.1a}$$

$$\text{s.t.} \quad (x, y) \in \mathcal{X} \tag{D.1b}$$

$$\sum_{j \in J} \sum_{m \in M} z^{t,n}_{i,j,m} + u^n_{i,t} = W^n_{i,t}, \qquad \forall i \in [I],\ t \in [T],\ n \in [N] \tag{D.1c}$$

$$\sum_{i \in I} z^{t,n}_{i,j,m} \leq C x^t_{j,m} \qquad \forall j \in [J],\ m \in [M],\ t \in [T],\ n \in [N] \tag{D.1d}$$

$$u^n_{i,t} \geq 0,\ z^{t,n}_{i,j,m} \geq 0, \qquad \forall i \in [I],\ j \in [J], m \in [M],\ t \in [T,\ n \in [N] \tag{D.1e}$$

References

Ahmadi-Javid, A., Seyedi, P., Syam, S. S., 2017. A survey of healthcare facility location. Computers & Operations Research 79, 223–263.

Albareda-Sambola, M., Fernández, E., Hinojosa, Y., Puerto, J., 2009. The multi-period incremental service facility location problem. Computers & Operations Research 36 (5), 1356–1375.

Alexy, B. B., Elnitsky, C. A., 1996. Community outreach: rural mobile health unit. JONA: The Journal of Nursing Administration 26 (12), 38–42.

Antunes, A., Berman, O., Bigotte, J., Krass, D., 2009. A location model for urban hierarchy planning with population dynamics. Environment and Planning A 41 (4), 996–1016.

Attipoe-Dorcoo, S., Delgado, R., Gupta, A., Bennet, J., Oriol, N. E., Jain, S. H., 2020. Mobile health clinic model in the covid-19 pandemic: lessons learned and opportunities for policy changes and innovation. International Journal for Equity in Health 19 (1), 1–5.

Basciftci, B., Ahmed, S., Shen, S., 2019. Distributionally robust facility location problem under decision-dependent stochastic demand. arXiv preprint arXiv:1912.05577.

Ben-Tal, A., Den Hertog, D., Vial, J.-P., 2015. Deriving robust counterparts of nonlinear uncertain inequalities. Mathematical programming 149 (1-2), 265–299.

Berman, O., Drezner, Z., Wesolowsky, G. O., 2003. Locating service facilities whose reliability is distance dependent. Computers & Operations Research 30 (11), 1683–1695.

Bertsimas, D., Popescu, I., 2005. Optimal inequalities in probability theory: A convex optimization approach. SIAM Journal on Optimization 15 (3), 780–804.

Bertsimas, D., Sim, M., 2004. The price of robustness. Operations research 52 (1), 35–53.

Bertsimas, D. J., Simchi-Levi, D., 1996. A new generation of vehicle routing research: robust algorithms, addressing uncertainty. Operations research 44 (2), 286–304.

Chen, Z., Sim, M., Xiong, P., 2020. Robust stochastic optimization made easy with rsome. Management Science.

Contreras, I., Cordeau, J.-F., Laporte, G., 2011. The dynamic uncapacitated hub location problem. Transportation Science 45 (1), 18–32.

Cordeau, J.-F., Laporte, G., Savelsbergh, M. W., Vigo, D., 2007. Vehicle routing. Handbooks in operations research and management science 14, 367–428.

Current, J. R., Velle, C. R., Cohon, J. L., 1985. The maximum covering/shortest path problem: A multiobjective network design and routing formulation. European Journal of Operational Research 21 (2), 189–199.

Delage, E., Saif, A., 2018. The Value of Randomized Solutions in Mixed-Integer Distributionally Robust Optimization Problems. GERAD HEC Montréal.

Drezner, T., 2014. A review of competitive facility location in the plane. Logistics Research 7 (1), 114.

Drezner, Z., Wesolowsky, G., 1991. Facility location when demand is time dependent. Naval Research Logistics (NRL) 38 (5), 763–777.

Esfahani, P. M., Kuhn, D., 2018. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. Mathematical Programming 171 (1-2), 115–166.

Flores-Garza, D. A., Salazar-Aguilar, M. A., Ngueveu, S. U., Laporte, G., 2017. The multi-vehicle cumulative covering tour problem. Annals of Operations Research 258 (2), 761–780.

Gendreau, M., Laporte, G., Semet, F., 1997. The covering tour problem. Operations Research 45 (4), 568–576.

Hachicha, M., Hodgson, M. J., Laporte, G., Semet, F., 2000. Heuristics for the multi-vehicle covering tour problem. Computers & Operations Research 27 (1), 29–42.

Halper, R., Raghavan, S., 2011. The mobile facility routing problem. Transportation Science 45 (3), 413–434.

Jena, S. D., Cordeau, J.-F., Gendron, B., 2017. Lagrangian heuristics for large-scale dynamic facility location with generalized modular capacities. INFORMS Journal on Computing 29 (3), 388–404.

Jiang, R., Shen, S., Zhang, Y., 2017. Integer programming approaches for appointment scheduling with random no-shows and service durations. Operations Research 65 (6), 1638–1656.

Kamath, K. R., Pakkala, T., 2002. A bayesian approach to a dynamic inventory model under an unknown demand distribution. Computers & Operations Research 29 (4), 403–422.

Kim, S., Pasupathy, R., Henderson, S. G., 2015. A guide to sample average approximation. In: Handbook of simulation optimization. Springer, pp. 207–243.

Kleywegt, A. J., Shapiro, A., Homem-de Mello, T., 2002. The sample average approximation method for stochastic discrete optimization. SIAM Journal on Optimization 12 (2), 479–502.

Kuhn, D., Esfahani, P. M., Nguyen, V. A., Shafieezadeh-Abadeh, S., 2019. Wasserstein distributionally robust optimization: Theory and applications in machine learning. In: Operations Research & Management Science in the Age of Analytics. INFORMS, pp. 130–166.

Lei, C., Lin, W.-H., Miao, L., 2014. A multicut l-shaped based algorithm to solve a stochastic programming model for the mobile facility routing and scheduling problem. European Journal of operational research 238 (3), 699–710.

Lei, C., Lin, W.-H., Miao, L., 2016. A two-stage robust optimization approach for the mobile facility fleet sizing and routing problem under uncertainty. Computers & Operations Research 67, 75–89.

Liu, K., Li, Q., Zhang, Z.-H., 2019. Distributionally robust optimization of an emergency medical service station location and sizing problem with joint chance constraints. Transportation research part B: methodological 119, 79–101.

Luo, F., Mehrotra, S., 2018. Distributionally robust optimization with decision dependent ambiguity sets. arXiv preprint arXiv:1806.09215.

Mak, H.-Y., Rong, Y., Zhang, J., 2014. Appointment scheduling with limited distributional information. Management Science 61 (2), 316–334.

Manzini, R., Gebennini, E., 2008. Optimization models for the dynamic facility location and allocation problem. International Journal of Production Research 46 (8), 2061–2086.

Owen, S. H., Daskin, M. S., 1998. Strategic facility location: A review. European journal of operational research 111 (3), 423–447.

Oyola, J., Arntzen, H., Woodruff, D. L., 2017. The stochastic vehicle routing problem, a literature review, part ii: solution methods. EURO Journal on Transportation and Logistics 6 (4), 349–388.

Rahimian, H., Mehrotra, S., 2019. Distributionally robust optimization: A review. arXiv preprint arXiv:1908.05659.

Reilly, W. J., 1931. The law of retail gravitation. WJ Reilly.

Ritzinger, U., Puchinger, J., Hartl, R. F., 2016. A survey on dynamic and stochastic vehicle routing problems. International Journal of Production Research 54 (1), 215–231.

Saif, A., Delage, E., 2020. Data-driven distributionally robust capacitated facility location problem. European Journal of Operational Research.

Shehadeh, K. S., Cohn, A. E., Jiang, R., 2019. A distributionally robust optimization approach for outpatient colonoscopy scheduling. European Journal of Operational Research.

Shehadeh, K. S., Padman, R., 2020. A distributionally robust optimization approach for stochastic elective surgery scheduling with limited intensive care unit capacity. European Journal of Operational Research.

Shehadeh, K. S., Sanci, E., 2020. Distributionally robust facility location with bimodal random demand. Optimization Online E-Print.

Shehadeh, K. S., Tucker, E. L., 2020. A distributionally robust optimization approach for location and inventory prepositioning of disaster relief supplies. arXiv preprint arXiv:2012.05387.

Soyster, A. L., 1973. Convex programming with set-inclusive constraints and applications to inexact linear programming. Operations research 21 (5), 1154–1157.

Subramanyam, A., Repoussis, P. P., Gounaris, C. E., 2020. Robust optimization of a broad class of heterogeneous vehicle routing problems under demand uncertainty. INFORMS Journal on Computing.

Thiele, A., 2010. A note on issues of over-conservatism in robust optimization with cost uncertainty. Optimization 59 (7), 1033–1040.

Toth, P., Vigo, D., 2014. Vehicle routing: problems, methods, and applications. SIAM.

Tricoire, F., Graf, A., Gutjahr, W. J., 2012. The bi-objective stochastic covering tour problem. Computers & operations research 39 (7), 1582–1592.

Van Roy, T. J., Erlenkotter, D., 1982. A dual-based procedure for dynamic facility location. Management Science 28 (10), 1091–1105.

Wang, S., Chen, Z., Liu, T., 2020. Distributionally robust hub location. Transportation Science 54 (5), 1189–1210.

Wang, Y., Zhang, Y., Tang, J., 2019. A distributionally robust optimization approach for surgery block allocation. European Journal of Operational Research 273 (2), 740–753.

Wesolowsky, G. O., Truscott, W. G., 1975. The multiperiod location-allocation problem with relocation of facilities. Management Science 22 (1), 57–65.

Wiesemann, W., Kuhn, D., Sim, M., 2014. Distributionally robust convex optimization. Operations Research 62 (6), 1358–1376.

Wu, C., Du, D., Xu, D., 2015. An approximation algorithm for the two-stage distributionally robust facility location problem. In: Advances in Global Optimization. Springer, pp. 99–107.