RUTCOR
RESEARCH
REPORT

ALTERNATE RISK MEASURES FOR EMERGENCY MEDICAL SERVICE SYSTEM DESIGN

Nilay Noyan[a]

RRR 13-2009, JULY, 2009

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/∼rrr

[a]Manufacturing Systems/Industrial Engineering Program, Faculty of Engineering and Natural Sciences, Sabancı University, Orhanli, Tuzla, 34956 Istanbul, Turkey. E-mail: nnoyan@sabanciuniv.edu

Rutcor Research Report
RRR 13-2009, July, 2009

Alternate Risk Measures for Emergency Medical Service System Design

Nilay Noyan

Abstract. The stochastic nature of emergency service requests and the unavailability of emergency vehicles when requested to serve demands are critical issues in constructing valid models representing real life emergency medical service (EMS) systems. We consider an EMS system design problem with stochastic demand and locate the emergency response facilities and vehicles in order to ensure target levels of coverage, which are quantified using risk measures on random unmet demand. The target service levels for each demand site and also for the entire service area are specified. In order to increase the possibility of representing a wider range of risk preferences we develop two stochastic optimization models involving alternate risk measures. Our first model includes integrated chance constraints (ICCs), whereas the second one incorporates ICCs and a second order dominance constraint. We propose solution methods for our stochastic optimization problems and present extensive numerical results demonstrating their computational effectiveness.

Keywords: *Stochastic programming, Random demand, Integrated chance constraints, Stochastic dominance constraints, Emergency facilities, Ambulance allocation.*

Introduction

Determining the optimal location of emergency vehicles is a significant problem in designing EMS systems and has received considerable attention in the literature (Brotcorne et al., 2003; Marianov and ReVelle,1995). A key point in effective emergency response is the prompt availability of emergency vehicles at response facilities. The service area of an EMS system is often modeled by defining a network consisting of a set of geographical nodes, where each node represents a source of requests for response (emergency) vehicles and/or a candidate facility site.

We consider an EMS system design problem of locating the response facilities (ambulance stations) and determining the number of vehicles (ambulances) to allocate to each facility. In real life situations, the future emergency service requests from demand sites are not known with certainty. Furthermore, due to the various constraints, such as budget or capacity, pre-allocated emergency vehicles may not be sufficient to cover all demands in the service area within an acceptable time. Since unmet demands in emergency situations may result in loss of life, it is critical to design systems that guarantee reasonable levels of coverage for potential users. Hence, demand uncertainty must be addressed effectively in designing EMS systems.

Stochastic programming is one of the fundamental approaches that can be used to model decision problems in the presence of uncertainty. Here, we model the uncertainty in demand for the emergency vehicles. In particular, we represent the uncertain demand by random variables and use a scenario approach. We describe two new stochastic programming formulations that determine the optimal location and allocation decisions minimizing the total cost while meeting the target service levels. The level of service is measured by keeping the total unmet demand values below some prescribed target values, guaranteeing a high level of coverage. We specify an individual target service level for each demand node and a target service level for the entire geographical area (system-wide coverage). Defining only a system-wide target service level may result in a lower level of coverage at one demand site and a higher level of coverage at another site, hence leading to inequitable solutions. Providing equal access to users is an important issue in EMS system design (Felder and Brinkman, 2002). However, there is no agrement on what is fair in an EMS system and how to measure the equity. For instance, Felder and Brinkman (2002) discuss the conflicts between the equal access policy guaranteeing a nationwide uniform response time and the policy aiming to provide equal per-capita resources in EMS systems. Felder and Brinkman (2002) suggest that every person should not be given equal access to emergency service irrespective of locations. For further discussion on equity in locating public facilities, we refer to Marsh and Schilling (1994). In this study, the way we define the individual target service levels may be regarded as an alternative approach to model the coverage equity. We consider an equal access policy which guarantees a systemwide uniform response time and also guarantee equal service at each demand site based on the proportion of the unmet demand. This implies that our models consider the level of demand as a criterion and allocate more vehicles to higher-demand areas than lower-demand areas, but in terms of proportionality different areas would get equal service. To the best of our knowledge, there does not exist an EMS design model like ours considering simultaneously individual target service levels and a system-wide service level.

Stochastic programming formulations using probabilistic constraints are widely applied in stochastic EMS design models. These models usually consider the probability of having an available vehicle within a standard acceptable distance as the performance measure. However, Erkut et al. (2008b) argue in detail that this probabilistic performance measure is not consistent with the performance measures used by most EMS operators in practice. A common performance measure is the fraction of calls covered whose response time is below a specified threshold. Discussions by Erkut et al. (2008b) indicate the significance of the models based on expected coverage performance measures. Following this line of thought, we propose to use alternate risk measures based on expected unmet demand; and therefore, our proposed risk measures may potentially be better aligned with the performance measures employed by EMS operators in practice. The major contribution of our study is the use of integrated chance constraints and stochastic dominance constraints as alternatives to probabilistic constraints in EMS system design problems.

Probabilistic constraints are commonly used; however, it is well known that they pose great computational difficulties. Therefore, in general we can only solve small to moderate size problems involving probabilistic constraints. In this study, we show that switching to alternate risk constraints and developing corresponding solution methods we obtain computationally tractable models for a larger set of scenarios compared to the existing literature. Handling a larger set of scenarios is significant in modeling uncertainties of real life. In our first stochastic optimization model, the target service levels are defined using integrated chance constraints. Note that probabilistic constraints measure the probabilities of violating the coverage constraints, irrespective of how violated the constraints are. In other words, the probabilistic constraints do not take the magnitude of the unmet demand into account. As an alternative, ICCs are based on the magnitude of violation in coverage constraints. ICCs were introduced by Klein Haneveld (1986) and have only been used in finance applications so far. The use of ICCs in EMS design is novel. In our second model, the individual target service level for each demand node is defined using an ICC on the random unmet demand, whereas the system-wide service level is defined using a second order stochastic dominance (SSD) constraint introduced by Dentcheva and Ruszczyński (2003). The stochastic dominance relation allows us to obtain location and allocation decisions for which the random total unmet demand dominates a benchmark (reference) random total unmet demand. Such a reference outcome may be defined based on an EMS standard or on a potential/candidate solution. In either case, there is a reference decision vector and the model involving stochastic dominance constraints constructs a decision vector which is better than the reference with respect to the distribution of the total unmet demand. We propose to define the reference outcome based on a common EMS performance standard which imposes a lower bound on the fraction of calls whose response time is below a threshold. By proposing two models we increase the possibility of representing a wider range of risk preferences. In many applications, where the distribution of a random outcome is of significant interest and it is possible to define a reasonable reference random outcome, we recommend the decision makers to use the model involving stochastic dominance constraints.

Literature review is presented in Section 1. In Section 2, we describe the underlying deterministic problem for the proposed stochastic programming problems. In Section 3, we first introduce a mixed 0-1 linear programming (LP) formulation for the stochastic EMS system design model with ICCs and then develop an associated computationally effective alternate formulation. In Section 4, a stochastic optimization problem involving ICCs and an SSD constraint is presented. We also describe effective and practical methods to solve this problem. We present numerical results in Section 5 to demonstrate the computational effectiveness of the developed solution methods and illustrate how input parameters and risk measures affect optimal location and allocation solutions. In Section 5, we also discuss the computational study performed to compare the proposed models with a closely related existing model. Finally, in Section 6 we conclude and discuss further research directions.

1 Literature review

The problem of determining the optimal locations of emergency vehicles has been quite popular in the MS/OR literature. For extensive reviews on emergency vehicle location and allocation models we refer to Brotcorne et al. (2003), Marianov and ReVelle (1995), and Goldberg (2004). EMS design problems are closely related to facility location problems. The readers should consult the book by Daskin (1995) and the references therein for a detailed discussion on facility location theory. The models we propose in this study are stochastic versions of the capacitated facility location problem. Stochastic facility location models have received significant attention in the last decades and there is a huge literature on such models. Our aim is not to provide an extensive review; but to briefly discuss some selected relevant papers on modeling the uncertainty in location problems.

Comprehensive reviews on facility location under uncertainty can be found in Berman and Krass (2001), Louveaux (1993), Owen and Daskin (1998), Current et al. (2002), Snyder (2006), and Snyder and Daskin (2006). The first probabilistic location models were proposed by Chapman and White (1974). Their formulations are extensions of the set covering and p-center problems, and they account for both vehicle availability and stochastic travel times. Based on a similar approach to Chapman and White, Daskin (1983) develops the maximum expected covering location model (MECLM) to account for the potential unavailability of ambulances. It is one of the first probabilistic models for locating emergency vehicles; it maximizes the expected demand coverage for a given number of facilities to be located on the network. MECLM is built on the assumption that each ambulance has the same probability of being unavailable to answer a call. As a generalization of the maximum covering model, ReVelle and Hogan (1989) propose chance constrained stochastic models which maximize the demand covered with a given probability value. There is also a rich literature on emergency vehicle location models focusing on randomness in response times and the demand. For example, Ingolfsson et al. (2007) model the uncertain delay and travel times and the ambulance availability. A recent study by Erkut et al. (2008) incorporates a survival function, which returns the probability of survival for a patient associated with the response time of the emergency vehicle dispatched to serve that patient, into existing coverage models. After this brief review of stochastic location models in general, we next focus on stochastic versions of the classical capacitated fixed charge facility location problem (CFLP) which are particularly related to our study.

Research focusing on stochastic CFLP includes Louveaux (1986), Ball and Lin (1993), Beraldi et al. (2004) and Beraldi and Bruni (2009). Louveaux (1986) presents a stochastic version of the CFLP in which the expected utility of profit is maximized while considering a penalty for unmet demand. Ball and Lin (1993), Beraldi et al. (2004) and Beraldi and Bruni (2009) assume that the main uncertainty is due to the stochastic call arrival process, and they propose stochastic programming formulations under probabilistic constraints. Ball and Lin (1993) incorporate a probabilistic constraint for each demand site to ensure that the probability of unavailability of a vehicle to serve a request from the demand point within an acceptable time is less than a certain value. On the other hand, Beraldi et al. (2004) and Beraldi and Bruni (2009) incorporate probabilistic constraints to ensure that all requests are served with a prescribed high probability. Since Beraldi et al. (2004) and Beraldi and Bruni (2009) directly focus on the randomness in demand rather than the randomness in the availability of vehicles, these studies are more closely related to our study than Ball and Lin (1993) and we would like to discuss our contribution relative to these studies.

Beraldi and Bruni (2009) introduce a two-stage stochastic programming problem, where the second stage decision variables are associated with scenarios to represent the assignment of vehicles to demand nodes under each scenario. Here, we do not model how an emergency operating system assigns an available vehicle to an emergency call and only focus on locating the response facilities and determining the number of vehicles to allocate to each facility. These decisions can be viewed as first stage decisions and similar to Beraldi et al. (2004) we formulate single-stage stochastic programming problems. Moreover, Beraldi and Bruni (2009) do not allow splittable demand, i.e., the demand at each node must be served by exactly one facility under each scenario. Therefore, it seems that our paper has more commonalities with Beraldi et al. (2004) and we performed a computational study to compare our results to those that would be obtained by using probabilistic constraints as in Beraldi et al. (2004).

Similar to previous studies like Daskin (1983) and Revelle and Hogan (1989), Beraldi et al. (2004) assume that the service providers and demand points are independent. We model the random demand using the scenario approach and relax the assumption that the service providers and demand points are independent. For a review of scenario planning models, see Owen and Daskin (1998). Beraldi et al. (2004) assume that the demand distributions at each node are given and the demand nodes are independent. Under the independence assumption they reformulate the problem involving probabilistic constraints and solve it using the CPLEX solver. Without the independence assumption it is hard to solve such a problem and the proposed reformulation by Beraldi et al. (2004) is not valid for a given set of scenarios. Therefore, a different reformulation is required to incorporate the probabilistic constraints when a scenario approach is used to model risk. These discussions support the potential contribution of the proposed risk constraints alternate to probabilistic constraints.

In general, one cannot claim that one risk measure is better than others. Depending on the decision maker's risk preference, either probabilistic constraint, or ICCs, or stochastic dominance constraints may be employed. However, our motivation to propose risk constraints alternate to probabilistic constraints in EMS system design models is based on the arguments by Erkut et. al. (2008b). Apart from this, the computational difficulties inherent in probabilistic constraints further motivate our research. Integrated chance constraints can be considered as relaxations of probabilistic constraints. Therefore, ICCs can be used to obtain convex approximations of the generally non-convex feasible sets defined by probabilistic constraints. (For details see Klein Haneveld and Van Der Vlerk, 2006.) Thus, the risk measures we propose to model different types of risk preference, lead to computationally tractable models and allow us to handle a larger set of scenarios for improving the validity of the models.

2 The Underlying Deterministic Model

In this section we present our assumptions and notation, and describe the underlying deterministic model for the stochastic programming models that are presented in the following sections. Then, we discuss how to incorporate stochastic demand into the model.

We say that a candidate facility can cover a demand node if the distance between them is less than or equal to an acceptable value, which is known as the coverage distance threshold and can be determined according to a response time standard.

Inputs

I: finite set of demand sites (nodes);

J: finite set of candidate facility sites, where response facilities can be located;

d_{ij}: traveling distance between demand node i and a candidate facility node j, $i \in I$, $j \in J$;

D_c: coverage distance threshold;

$M_j = \{i \in I \mid d_{ij} \leq D_c\}$: set of demand nodes that can be covered by a facility located at node j, $j \in J$;

$N_i = \{j \in J \mid d_{ij} \leq D_c\}$: set of all candidate facility nodes that are within acceptable distance of node i, $i \in I$;

f_j: (hourly) fixed cost of opening a facility at node j, $j \in J$;

a: (hourly) cost of purchasing and maintaining an emergency vehicle;

β: cost per unit distance per unit demand;

c_{ij}: cost of shipping a unit of demand from a facility at node j to node i (notice that $c_{ij} = \beta d_{ij}$), $i \in I$, $j \in J$;

U_j: maximum number of vehicles that can be assigned to a facility located at node j, $j \in J$;

h_i: number of service requests (demand) generated at node i, $i \in I$, during a specified amount of time;

Decision variables

$$y_j = \begin{cases} 1 & \text{if a facility is located at node } j \in J, \\ 0 & \text{otherwise.} \end{cases}$$

$$x_{ij} = \text{ Number of vehicles located at facility site } j \in J \text{ due to demand at node } i \in I.$$

We remark that we do not explicitly model how an available response vehicle is assigned to an emergency call, i.e., we do not consider any particular dispatching rule. The allocation variable x_{ij} is interpreted as the number of vehicles located at facility site j due to the demand at node i. However, *these vehicles are not dedicated to node i* in real life dispatching problems. In other words, $\sum_{i \in M_j} x_{ij}$ only represents the total number of vehicles allocated to node j and these vehicles are not reserved to serve specific demand nodes. Depending on the realized demand values, they would be dispatched to serve any node that are within acceptable distance to facility j. When the demand is stochastic, it may happen that less than x_{ij} vehicles are needed to cover demand at node i, then those unused vehicles could be dispatched to serve requests from other demand nodes. Thus, for the determined facility locations and the number of vehicles allocated to each facility, existing methods may be utilized to find practical solutions for dispatching and reallocating vehicles. As an alternative approach one can define the decision variable x_{ij} associated with each scenario and determine how to assign the vehicles to demand nodes. Beraldi and Bruni (2009) have recently presented such a model. Here we do not focus on assignment decisions and we define x_{ij} variables only to represent the inter-facility resource allocation. Our models provide us with determined x_{ij} values that are large enough to guarantee the satisfaction of demand with prescribed service levels, regardless of the realization of (actual value taken by) demand.

We consider service requests at each demand site during a certain amount of time. The length of this time period is chosen as a reasonable time required for a service trip, which we define as the total time required for an emergency vehicle to return to the original location before responding to another service request. Similar to other studies in this area, such as Beraldi et al. (2004) and Beraldi and Bruni (2009) we consider hourly demand.

For our stochastic programming problems, we have the following *underlying deterministic* *problem*:

$$\min \quad \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} f_j y_j + \sum_{i\in I}\sum_{j\in J} a x_{ij} \tag{1}$$

$$\text{subject to} \quad \sum_{j\in N_i} x_{ij} \geq h_i, \quad \forall i \in I, \tag{2}$$

$$\sum_{i\in M_j} x_{ij} \leq U_j y_j, \quad \forall j \in J, \tag{3}$$

$$x_{ij} \in \mathbb{Z}_+, \forall i \in I,\ j \in J, \tag{4}$$

$$y_j \in \{0,1\}, \forall j \in J. \tag{5}$$

This is similar to the well known mathematical programming formulation of the capacitated fixed charge facility location problem with splittable demands. The objective function (1) minimizes the sum of the variable transportation costs, the fixed setup cost for opening the facilities and the total cost of purchasing and maintaining vehicles. Different than the traditional CFLP, we also incorporate the total ambulance cost which is obtained by multiplying the total number of ambulances by the unit cost a. In ambulance service, 24-hour availability is required and therefore, it has relatively high fixed costs associated with equipment and staffing. However, the variable costs based on distances are significant in allocating vehicles since the closest available vehicle is usually dispatched to an emergency call. Each demand node has a total request of h_i vehicles, and coverage constraints (2) ensure that all of the demand at each node i must be served by one or more open facilities. Constraints (3) are capacity constraints, which guarantee that the amount of vehicles allocated to each facility site $j \in J$ is less than or equal to the maximum number of vehicles that can be assigned to facility node j. Constraints (4) and (5) are the integrality and nonnegativity constraints.

In practice, as we determine the values of the allocation vector $\mathbf{x}$ and the location vector $\mathbf{y}$, the actual values of $h_i,\ i \in I$, are not known; they will become known in the future. In this paper, we consider models where the demand parameters, $h_i,\ i \in I$, are not constants but random variables. This implies that coverage constraints (2) are stochastic. We assume that we are given a discrete set of scenarios, a set of realizations of joint service requests at demand nodes, and their associated probabilities. Let S denote the finite set of (global) scenarios; p_s denote the probability associated with scenario $s,\ s \in S$ and $h_i(s)$ denote realization of demand at node i under scenario s, $i \in I, s \in S$. It is worthwhile to point out that using scenarios allows the demand values to be dependent. In the following Sections 3 and 4, we develop mathematical programming formulations, which incorporate risk measures to model the uncertainty in demand using the scenario approach.

3 An integrated chance constrained EMS system design model

In Section 3.1 we discuss the integrated chance constraints and then in Section 3.2 we introduce *the integrated chance constrained EMS system design* problem. Finally, in Section 3.3 we develop an equivalent alternate formulation, which leads to a significant reduction in number of variables and provides us with an efficient solution method.

3.1 Integrated chance constraints

Integrated chance constraints (ICCs) introduced by Klein Haneveld (1986) as alternate to probabilistic constraints. Let $e_i : \mathbb{Z}_+{}^{|I||J|} \times S \to \mathbb{R},\ i \in I$, be the outcome mappings. For a given allocation vector $\mathbf{x} \in \mathbb{Z}_+{}^{|I|\ |J|}$ let us define the mapping $e_{(\mathbf{x},i)} : S \to \mathbb{R}$ by $e_{(\mathbf{x},i)}(s) = e_i(\mathbf{x}, s)$ for all $i \in I,\ s \in S$. Also let $[\eta]_+ = \max(0, \eta)$ and $[\eta]_- = \max(0, -\eta)$, $\eta \in \mathbb{R}$.

We denote the random unmet demand at node $i \in I$ by $e_{(\mathbf{x},i)}$ and the random total unmet demand (for the network) by $\xi_{\mathbf{x}}$, where

$$e_{(\mathbf{x},i)}(s) = \left[h_i(s) - \sum_{j \in N_i} x_{ij} \right]_+, \quad \forall s \in S \tag{6}$$

and

$$\xi_{(\mathbf{x})}(s) = \sum_{i \in I} e_{(\mathbf{x},i)}(s) = \sum_{i \in I} \left[h_i(s) - \sum_{j \in N_i} x_{ij} \right]_+, \quad \forall s \in S. \tag{7}$$

We model the risk, which can be broadly defined as the effect of variability of random service requests, by specifying constraints on random unmet demands, $e_{(\mathbf{x},i)},\ i \in I$, and the random total unmet demand $\xi_{\mathbf{x}}$. The underlying idea of the models we consider is to allow infeasibilities in the stochastic constraints (2), but specify restricting constraints on the amount of their violations. In connection with the stochastic constraints there are several measures of violation that can be incorporated into an optimization model as constraints. One way of measuring such violations is via probabilistic constraints, also called chance constraints. Here are several probabilistic constraints that can be considered for the EMS system design problem:

$$\begin{aligned} &P[e_{(\mathbf{x},i)} \le \eta_i^{(1)}] \ge 1 - \gamma_i^{(1)},\ i \in I \quad \text{(individual probabilistic constraints)},\\ &P[e_{(\mathbf{x},i)} \le \eta_i^{(2)},\ i \in I] \ge 1 - \gamma^{(2)} \quad \text{(a joint probabilistic constraint)},\\ &P[\xi_{\mathbf{x}} \le \eta^{(3)}] \ge 1 - \gamma^{(3)} \quad \text{(individual probabilistic constraint)}, \end{aligned} \tag{8}$$

where $\eta_i^{(1)}$, $\eta^{(2)}$ and $\eta^{(3)}$ are some fixed target values and $\gamma_i^{(1)}$, $\gamma^{(2)}$, $\gamma^{(3)} \in (0, 1)$ are the maximum allowed probabilities of violating the stochastic constraints. Models involving probabilistic constraints were introduced by Charnes (1958), Miller (1965), and Prékopa (1970). Prékopa (1995) discusses in detail the probabilistic optimization theory and the associated numerical techniques. Beraldi et al. (2004) consider joint probabilistic constraints, where the target values are 0 (in order to measure the probability of covering all the requests at all demand nodes or a set of demand nodes). Notice that probabilistic constraints measure the probabilities of violating the stochastic constraints, and they do not take the magnitude of the unmet demand into account if there is excess demand. Whereas, ICCs are based on the amounts of violations in connection with stochastic constraints.

Using the definition of Klein Haneveld (1986) we have the ICCs on random unmet demands as follows:

$$\mathbb{E}([h_i - \sum_{j\in N_i} x_{ij}]_+) \leq q_i,\ i \in I, \tag{9}$$

where $\mathbb{E}$ stands for the expected value operator and $q_i, i \in I$, are nonnegative risk aversion parameters representing the largest acceptable expected unmet demand values. The constraints of type (9) guarantee that for all demand nodes the average magnitude of unmet demand is less than or equal to the maximum acceptable risk aversion parameters. For example, one can set $q_i = \alpha_i E[h_i]$, where α_i is another type of risk aversion parameter and this would mean that at most a fraction α_i of the expected demand be unmet. Here as proposed by Klein Haneveld and Van Der Vlerk (2006), we construct alternative individual ICCs by choosing the risk parameters dependent on the distributions of random unmet demand instead of specifying the maximum acceptable risk parameters as fixed numbers $q_i, i \in I$. These alternative ICCs are

$$\mathbb{E}\left([h_i - \sum_{j\in N_i} x_{ij}]_+\right) \leq \alpha_i \mathbb{E}\left(\mid h_i - \sum_{j\in N_i} x_{ij} \mid\right),\ i \in I, \tag{10}$$

where $\alpha_i \in [0, 1/2]\,,\ i \in I$, is a risk aversion parameter associated with demand site i, specified by decision makers according to their risk preferences. Let us denote the expected value of the random demand h_i by $\bar{h}_i$. We note that when the risk parameters $\alpha_i = 1/2, i \in I$, the ICCs (14) take the form of the coverage constraints, where the random variables are replaced by their expected values:

$$\sum_{j\in N_i} x_{ij} \geq \bar{h}_i, \quad \forall i \in I.$$

Therefore, it is only meaningful to consider $\alpha_i \in [0, 1/2]\,,\ i \in I$. Otherwise, we would obtain solutions for which the unmet demand values would be even higher than those associated with the solutions constructed by using a naive approach based on the expected demand values.

Since $[a]_+ + [a]_- = |a|$ and $[a]_- = [a]_+ - a$ for $a \in \mathbb{R}$, the alternative ICCs (10) are equivalently represented by

$$(1 - 2\alpha_i)\mathbb{E}\left(\left[h_i - \sum_{j\in N_i} x_{ij}\right]_+\right) \leq \alpha_i \left(\sum_{j\in N_i} x_{ij} - \bar{h}_i\right),\ i \in I. \tag{11}$$

Similarly, the ICC on the random total unmet demand defined in (7) is given by

$$(1 - 2\delta) \sum_{i\in I} \mathbb{E}\left(\left[h_i - \sum_{j\in N_i} x_{ij}\right]_+\right) \leq \delta \left(\sum_{i\in I} \sum_{j\in N_i} x_{ij} - \sum_{i\in I} \bar{h}_i\right), \tag{12}$$

where $\delta \in [0, 1/2]$ is a risk aversion parameter associated with the total unmet demand.

These constraints would allow the decision makers to evaluate different location and allocation decisions based on the tradeoff between the quality of service and costs by varying the risk parameters.

3.2 An integrated chance constrained EMS system design problem

Replacing coverage constraints (2) in the *underlying deterministic problem* by ICCs (11) and (12) leads to the following stochastic programming problem:

$$
\begin{aligned}
\min \quad & \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} f_j y_j + \sum_{i\in I}\sum_{j\in J} a x_{ij} && (13)\\
\text{subject to} \quad & (1-2\alpha_i)\mathbb{E}\left(\left[h_i - \sum_{j\in N_i} x_{ij}\right]_+\right) \le \alpha_i \left(\sum_{j\in N_i} x_{ij} - \bar{h}_i\right),\ i\in I, && (14)\\
& (1-2\delta)\sum_{i\in I}\mathbb{E}\left(\left[h_i - \sum_{j\in N_i} x_{ij}\right]_+\right) \le \delta\left(\sum_{i\in I}\sum_{j\in N_i} x_{ij} - \sum_{i\in I}\bar{h}_i\right), && (15)\\
& (\mathbf{x},\mathbf{y}) \in Q, && (16)
\end{aligned}
$$

where $\alpha_i \in [0,1/2]$, $i \in I$, and δ are prescribed risk aversion parameters, and $Q = \{(\mathbf{x} \in {\mathbb{Z}_+}^{|I|\times|J|}, \mathbf{y} \in \{0,1\}^{|J|}) : \sum_{i\in M_j} x_{ij} \le U_j y_j,\ \forall\ j \in J\}$. We refer to this problem as *the integrated chance constrained EMS system design* problem (ICCsP).

We set $\alpha_i, i \in I$, values to be equal for providing fair service to each demand site in terms of the proportion of the unmet demand. For the rest of the paper we let $\alpha_i = \alpha,\ i \in I$. The ICCs (14) defined for each demand node are referred as to local constraints and the constraint (15) defined for the system-wide service level is referred to as a global constraint. We chose the risk parameters, $\delta < \alpha < 0.5$, so that both types of constraints drive the system. This model is significant since it allows us to control simultaneously the target levels for each demand node and the entire network.

The optimization problem (13)-(16) can be represented by an mixed integer programming (MIP) formulation by creating an $|S| \times |I|$ matrix of new variables representing the excess demand values. Recall that the excess value of demand realization $h_i(s),\ s \in S$, with respect to the total number of vehicles that are allocated due to the demand at node $i \in I$ is represented by $e_{(\mathbf{x},i)}(s)$ defined in (6).

Then, we obtain the following deterministic equivalent formulation of ICCsP:

$$\min \quad \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} f_j y_j + \sum_{i\in I}\sum_{j\in J} a x_{ij} \tag{17}$$

$$\text{subject to} \quad (1-2\alpha)\sum_{s\in S} p_s \hat{e}_{(\mathbf{x},i)}(s) \le \alpha\left(\sum_{j\in N_i} x_{ij} - \bar{h}_i\right),\ \forall\, i\in I, \tag{18}$$

$$(1-2\delta)\sum_{i\in I}\sum_{s\in S} p_s \hat{e}_{(\mathbf{x},i)}(s) \le \delta\left(\sum_{i\in I}\sum_{j\in N_i} x_{ij} - \sum_{i\in I}\bar{h}_i\right), \tag{19}$$

$$\hat{e}_{(\mathbf{x},i)}(s) \ge h_i(s) - \sum_{j\in N_i} x_{ij},\ s\in S,\ i\in I, \tag{20}$$

$$\hat{e}_{(\mathbf{x},i)}(s) \ge 0,\ s\in S,\ i\in I, \tag{21}$$

$$(\mathbf{x},\mathbf{y}) \in Q, \tag{22}$$

where decision variables $\hat{e}_{(\mathbf{x},i)}(s),\ s\in S,\ i\in I$, represent the excess values of demand realizations. We refer to this problem as DirectICCsP. The following observation is required to argue that we obtain the optimal solution even if we use $\hat{e}_{(\mathbf{x},i)},\ i\in I$, variables instead of $\mathbf{e}_{(\mathbf{x},i)},\ i\in I$.

Observation 1. *For every feasible solution* $(\mathbf{x},\mathbf{y},\hat{\mathbf{e}}_{(\mathbf{x},i)},\ i\in I)$ *of* (18)-(22) *we have* $\hat{\mathbf{e}}_{(\mathbf{x},i)} \ge \mathbf{e}_{(\mathbf{x},i)},\ i\in I$, *and the pair* $(\mathbf{x},\mathbf{y},\mathbf{e}_{(\mathbf{x},i)},\ i\in I)$ *is also feasible for* (18)-(22).

One may solve this MIP formulation directly using a standard mixed integer programming solver such as CPLEX. However, in case of large instances it would be difficult for a MIP solver to provide an optimal solution. In the following section, we describe an alternative equivalent formulation for DirectICCsP in order to reduce the number of variables and develop an efficient method to solve our original problem ICCsP.

3.3 An alternate formulation based on local demand distributions

A scenario represents a realization of joint service requests at demand nodes. Notice that the same demand realization for a node can be observed under multiple scenarios, and therefore, the number of different demand realizations for each node would be significantly smaller than the number of global scenarios, $|S|$. For a given set of scenarios we can easily find the different demand realizations and the associated probabilities for each node. Basically, we decompose the global scenarios into local scenarios, and we denote the set of different demand realizations for node i by $S_i,\ i\in I$. Let $p_i(m)$ denote the probability of the demand realization, which is equal to $m,\ m\in S_i$, then $p_i(m) = \sum_{s\in S}\{p_s\ :\ h_i(s)=m\}$. Then, the expected value of the random unmet demand at node $i\in I$ is rewritten as follows

$$\mathbb{E}\left(\left[h_i - \sum_{j\in N_i} x_{ij}\right]_+\right) = \sum_{s\in S} p_s\left(\left[h_i(s) - \sum_{j\in N_i} x_{ij}\right]_+\right) = \sum_{m\in S_i} p_i(m)\left(\left[m - \sum_{j\in N_i} x_{ij}\right]_+\right).$$

Let

$$\tau_{(\mathbf{x},i)}(m) = [m - \sum_{j\in N_i} x_{ij}]_+, \ i \in I, \ m \in S_i.$$

Then, the alternate formulation of problem ICCsP is

$$\min \quad \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} f_j y_j + \sum_{i\in I}\sum_{j\in J} a x_{ij} \tag{23}$$

$$\text{subject to} \quad (1-2\alpha)\sum_{m\in S_i} p_i(m)\hat{\tau}_{(\mathbf{x},i)}(m) \le \alpha\left(\sum_{j\in N_i} x_{ij} - \bar{h}_i\right), \ \forall\, i \in I, \tag{24}$$

$$(1-2\delta)\sum_{i\in I}\sum_{m\in S_i} p_i(m)\hat{\tau}_{(\mathbf{x},i)}(m) \le \delta\left(\sum_{i\in I}\sum_{j\in N_i} x_{ij} - \sum_{i\in I}\bar{h}_i\right), \tag{25}$$

$$\hat{\tau}_{(\mathbf{x},i)}(m) \ge m - \sum_{j\in N_i} x_{ij}, \ i \in I, \ m \in S_i, \tag{26}$$

$$\hat{\tau}_{(\mathbf{x},i)}(m) \ge 0, \ i \in I, \ m \in S_i, \tag{27}$$

$$(\mathbf{x}, \mathbf{y}) \in Q, \tag{28}$$

where decision variables $\hat{\tau}_{(\mathbf{x},i)}(m), \ i \in I, \ m \in S_i$, are introduced to represent the excess values of demand realizations. We refer to this problem as *the alternative integrated chance constrained EMS system design* problem (AlterICCsP).

This alternate formulation creates only $\sum_{i\in I} |S_i|$ new variables instead of $|S||I|$. Thus, it leads to a significant reduction in number of variables and provides us with an efficient solution method. For example, in our computational studies for a problem instance where $|I| = |J| = 400$ and $|S| = 50,000$ we have $\max_{i\in I} |S_i| = 9$ and $\sum_{i\in I} |S_i| = 2471$, whereas $|S||I| = 400 * 50,000 = 20,000,000$.

In the next section, we describe another model involving a different type of constraint on the system-wide service level, which imposes a lower bound on the fraction of calls whose response time is below a threshold. This is a common EMS performance criterion (see Erkut et al., 2008).

4 The SSD-based EMS system design model

We consider the same underlying deterministic model and the scenario approach presented in Section 2. In Section 4.1 we discuss the second order stochastic dominance constraints. Then, in Section 4.2 we describe *the SSD-based EMS system design* problem and develop an equivalent alternate formulation. Since it is hard to solve the proposed problem with SSD constraints, we develop a heuristic procedure in Section 4.3. Finally, in Section 4.4 we describe how to utilize an EMS standard to generate the reference distribution of the total random unmet demand, which is required to apply the stochastic dominance based approach.

4.1 Second order stochastic dominance constraints

In many applications where the distribution of a random outcome is of significant interest, a single ICC may not be sufficient to model our preferences. We may introduce many individual ICCs, which is closely related to the concept of second order stochastic dominance. The relation of stochastic dominance is one of the fundamental concepts of statistics and decision theory (Mann and Whitney, 1947; Lehmann, 1955). The concept of stochastic dominance introduces a preorder in the space of real random variables [1]. It has been widely used in economics and finance (Hadar and Russell, 1969; Hanoch and Levy,1969; also see Levy, 1992 and references therein). We refer to Müller and Stoyan (2002) for a modern perspective on stochastic dominance relations.

Comparing uncertain outcomes is one of the fundamental interests of decision theory. Suppose that smaller values of random outcomes are preferred to larger ones. In the stochastic dominance based approach, random variables are compared by a point-wise comparison of some performance functions constructed from their distribution functions. Let $F_X(\eta) = P[X \leq \eta]$ denote the distribution function of a random variable X. The performance function we consider here is given by an area below the distribution function F_X:

$$\mathbb{E}([X - \eta]_+), \ \eta \in \mathbb{R}.$$

This function is well defined for all random variables X with finite expected value.

Definition 1. *For two integrable random variables X and Y, X* dominates *Y in the second order, which we denote by $X \succeq_{(2)} Y$, if*

$$\mathbb{E}([X - \eta]_+) \leq \mathbb{E}([Y - \eta]_+) \quad \textit{for all} \quad \eta \in \mathbb{R}. \tag{29}$$

This definition featuring the expected excess values of the random variables is intuitive; preferring smaller realizations of a random outcome implies preferring smaller expected excess values with respect to some threshold values.

Suppose that Y has a discrete distribution with realizations $y_k, \ k = 1, \dots, D$, then the inequalities (29) are equivalent to (see Dentcheva and Ruszczyński 2003):

$$\mathbb{E}([X - y_k]_+) \leq \mathbb{E}([Y - y_k]_+), \quad k = 1, \dots, D. \tag{30}$$

It is well-known (see, e.g., Müller and Stoyan, 2002) that the second order stochastic dominance relation $X \succeq_{(2)} Y$ can be equivalently expressed as follows:

$$\mathbb{E}[u(X)] \leq \mathbb{E}[u(Y)]$$

holds for all non-decreasing convex functions $u : \mathbb{R} \to \mathbb{R}$ for which the above expectations are finite; these u functions are interpreted as the risk-averse disutility functions. Due to this correspondence with the utility theory and in particular due to the correspondence with the risk-averse preferences, the second-order dominance relation has been receiving significant attention.

[1] A preorder is a relation that is reflexive and transitive, but not necessarily antisymmetric and each preorder induces an equivalence relation between elements.

Dominance relations can be involved in stochastic optimization problems as constraints, allowing us to obtain solutions for which the random outcomes of interest dominate some benchmark (reference) random outcomes. Recently, there has been significant interest in such stochastic optimization models. They have been introduced and analyzed by Dentcheva and Ruszczyński (2003 and 2006).

The expected excess values are closely related to the Conditional Value at Risk (CVaR) defined by Rockafeller and Uryasev (2000); the SSD relation is equivalent to a continuum of CVaR constraints (see Dentcheva and Ruszczyński, 2006). It has been quite popular in different fields to develop models involving CVaR; as a relevant example, it has been used to model stochastic facility location decisions (Chen et al., 2006). We would like to emphasize that Chen et al. (2006) model the risk by some modifications of the objective function; in particular minimizing the CVaR for a single confidence level (corresponds to a single threshold). However, we specify constraints on risk associated with multiple thresholds while minimizing the total cost.

In this paper we propose to introduce an SSD constraint into EMS system design models. In particular, we introduce a new EMS system design optimization model involving local ICCs on random unmet demands and a single SSD constraint on the random total unmet demand. The SSD constraint allows us to construct location and allocation decisions for which the random total unmet demand dominates a reference random total unmet demand. In this study, we define the reference outcome based on a common EMS performance standard which is to respond to $\rho * 100\%$ of all calls within r_1 minutes. Therefore, our model with the SSD constraint constructs solutions consistent with this common EMS performance standard. By proposing the second stochastic optimization model, we intend to provide a useful analytical tool for the EMS decision makers who are more interested in the distribution of the random outcome and would like to obtain a decision vector dominating a potential one under consideration.

4.2 Integer linear optimization problem with second order constraints

Consider our problem of designing an EMS system in which the decision vector $\mathbf{x}$ affects random unmet demands, $e_{(\mathbf{x},i)}, i \in I$, defined by (6). In order to find the best decision vector $\mathbf{x}$, we compare the corresponding random unmet demands and prefer the decisions with smaller unmet demand values. Thus, the problem of choosing between decisions becomes the problem of choosing between random variables representing unmet demands according to a preference relation. Here, we specify a preference relation among random variables based on ICCs (14) and the second order stochastic relation. Let Y be some random variable (*benchmark outcome*). We develop the following stochastic optimization model involving ICCs and an SSD constraint:

$$\min \quad \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} f_j y_j + \sum_{i\in I}\sum_{j\in J} a x_{ij}$$

$$\text{subject to} \quad (1-2\alpha)\mathbb{E}\left(\left[h_i - \sum_{j\in N_i} x_{ij}\right]_+\right) \le \alpha\left(\sum_{j\in N_i} x_{ij} - \bar{h}_i\right),\ i \in I, \tag{31}$$

$$\xi_{(\mathbf{x})} \succeq_{(2)} Y, \tag{32}$$

$$(\mathbf{x},\mathbf{y}) \in Q.$$

We refer to this problem as *the SSD-based EMS system design* problem (SSDP).

Suppose that the reference random outcome Y has a discrete distribution with realizations $y_1 < y_2 < \ldots < y_D$. Then, by the use of (30) the SSD relation (32) is equivalent to

$$\mathbb{E}([\xi_{(\mathbf{x})} - y_k]_+) \le \mathbb{E}([Y - y_k]_+), \quad k = 1,\ldots,D. \tag{33}$$

We would like to point out that the set of inequalities (33) can be viewed as finitely many ICCs of type (9). In the discrete case, the SSD relation (32) can be represented by finitely many linear inequalities. In order to solve problem SSDP efficiently we propose to use the set of linear inequalities introduced by Luedtke (2007), where large realizations of an uncertain outcome are preferred. We adapt this set of inequalities to the case of reverse preference since smaller total unmet demand values are preferred to larger ones.

Theorem 1. ***(A representation of the*** **SSD** ***relation.)*** *Suppose that X is a random variable and the realization of X under scenario s is denoted by z_s, $s \in S$, and Y has a discrete distribution with realizations y_k, $k = 1,\ldots,D$, and the associated probabilities are denoted by v_k, $k = 1,\ldots,D$. Then, $X \succeq_{(2)} Y$ if and only if there exists $\pi \in \mathbb{R}_+^{|S|D}$ such that*

$$z_s \le \sum_{k=1}^{D} y_k \pi_{sk}, \quad s \in S \tag{34}$$

$$\sum_{k=1}^{D} \pi_{sk} = 1, \quad s \in S \tag{35}$$

$$\sum_{s\in S}\sum_{j=k}^{D} (y_j - y_k)\pi_{sj} \le \sum_{j=k}^{D} v_j(y_j - y_k), \quad k \in \{1,\ldots,D\}. \tag{36}$$

In our study z_s is the realization of the total unmet demand under scenario s, i.e., $z_s = \xi_{(\mathbf{x})}(s) = \sum_{i\in I} e_{(\mathbf{x},i)}(s)$. Here, we cannot utilize the local demand distributions as we did for ICCsP, since the realization of the total unmet demand under each scenario s, z_s, is required to represent the SSD constraint, and thus, the global scenarios cannot be decomposed. Then, using the variables representing the excess demand values we obtain a mixed-integer LP formulation of SSDP referred to as AlterSSDP:

$$
\begin{aligned}
\min \quad & \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} f_j y_j + \sum_{i\in I}\sum_{j\in J} a x_{ij} \\
\text{subject to} \quad & (1-2\alpha)\sum_{s\in S} p_s \hat{e}_{(\mathbf{x},i)}(s) \le \alpha\left(\sum_{j\in N_i} x_{ij} - \overline{h_i}\right),\ i\in I, && (37)\\
& (20)-(21), && (38)\\
& \sum_{i\in I} \hat{e}_{(\mathbf{x},i)}(s) \le \sum_{k=1}^{D} y_k \pi_{sk}, \quad s\in S, && (39)\\
& (35)-(36) && (40)\\
& \pi_{sk} \ge 0, \quad s\in S,\ k\in\{1,\ldots,D\}, && (41)\\
& (\mathbf{x},\mathbf{y}) \in Q. && (42)
\end{aligned}
$$

This formulation creates $|S| * D$ new variables and $O(S + D)$ constraints in order to represent the SSD relation (32) by linear inequalities. To improve the computational effectiveness we propose to rewrite ICCs (37) using a binary search algorithm. Notice that ICCs (31) are local constraints, each of which independently imposes a lower bound on the total number of vehicles to be allocated in order to ensure the individual target service levels. It is easy to see that these constraints depend monotonically on the number of vehicles allocated to cover demand at each node, i.e., if $\sum_{j\in N_i} x_{ij} = k$ satisfies the associated ICC, then for any number of vehicles greater than k the same constraint is satisfied. Due to this special structure of the local constraints we implement a binary search algorithm for each demand node in order to find the minimum number of vehicles needed to reach the individual target service levels. We denote these lower bounds on the number of vehicles allocated due to demand at each node by $LB_i^{\text{ICC}},\ i\in I$.

Finally, the proposed formulation for our original problem SSDP, which we refer to as AlterSSDPWithBS, is given by

$$
\begin{aligned}
\min \quad & \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} f_j y_j + \sum_{i\in I}\sum_{j\in J} a x_{ij} \\
\text{subject to} \quad & \sum_{j\in N_i} x_{ij} \ge LB_i^{\text{ICC}},\ i\in I, \\
& (38)-(42).
\end{aligned}
$$

4.3 Heuristic Algorithm

In our numerical study, we could solve moderately large problem instances using the proposed formulation AlterSSDPWithBS. However, solving AlterSSDPWithBS requires substantially more effort as the number of potential sites and the number of scenarios increase due to the variables introduced to represent the SSD relation. Hence, we propose a mathematical programming based heuristic procedure, HSSDP, that yields a feasible solution with a small optimality gap within a reasonable computation time. Intuitively, we have a two step problem. First, we need to find an estimate for the optimal number of vehicles required to satisfy the target service levels, which is accomplished by solving the LP relaxation of AlterSSDPWithBS. Then, given the vehicle requirements we solve the deterministic problem CFLP in order to determine the facility locations and the vehicle allocation at each facility.

Algorithm 1 Heuristic Algorithm HSSDP for SSDP.

Initialize the iteration number, $t = 0$.
Solve the LP relaxation of AlterSSDPWithBS to obtain the optimal solution $(\mathbf{x}^{LP}, \mathbf{y}^{LP})$.
Let h_i^t, the total number of vehicles allocated due to node i, be equal to $\lfloor \sum_{j \in N_i} x_{ij}^{LP} \rfloor$.
if the allocation vector h^t does not satisfy SSD constraints (33) **then**
  Find the set $I^* = \{i \in I : (\lceil \sum_{j \in N_i} x_{ij}^{LP} \rceil - \lfloor \sum_{j \in N_i} x_{i,j}^{LP} \rfloor) > 0\}$.
end if
while the allocation vector h^t does not satisfy SSD constraints (33) **do** {Iteratively increase the number of vehicles allocated}
  Let $t := t + 1$. Find the smallest index $i^* \in I^*$ such that $i^* = \arg\max_{i \in I^*} \{\lceil \sum_{j \in N_i} x_{ij}^{LP} \rceil - \lfloor \sum_{j \in N_i} x_{ij}^{LP} \rfloor\}$.
  Update the allocation vector h^t by setting $h_{i^*}^t = \lceil \sum_{j \in N_{i^*}} x_{i^* j}^{LP} \rceil$.
  Let $I^* := I^* \setminus i^*$.
end while
Solve the *the underlying deterministic problem* defined by (1)-(5), where $h_i = h_i^t, \quad i \in I$.

This finite procedure provides us with a feasible solution of SSDP. It is easy to see that $\sum_{j \in N_i} x_i = \lceil \sum_{j \in N_i} x_i^{LP} \rceil$ would satisfy the SSD constraints. However, such a direct round up approach would lead to an excessive allocation of vehicles. Thus, instead we propose to increase the number of vehicles to be allocated iteratively until the feasibility of SSD constraints is achieved. This iterative procedure turns out to be effective to obtain feasible solutions with reasonable small optimality gaps (see Section 5.3).

4.4 How to generate the distributions of reference outcomes

As discussed before, applying the stochastic dominance based approach requires that the reference distribution of the total random unmet demand is available in advance. To this end, we need to specify target values $y_k, k \in \{1, \ldots, D\}$, as realizations of the reference total unmet demand along with the associated probabilities $v_k, k \in \{1, \ldots, D\}$.

A widely applied EMS standard is to respond to $\rho * 100\%$ of all calls within r_1 minutes (generally, $\rho = 0.9$ and $r_1 = 8$). In order to construct solutions in line with this standard we construct the empirical distribution of the total demand, denoted by TD, given the set of scenarios and their associated probabilities. Then, we set the reference random unmet demand to $Y = (1 - \rho)$ TD. For instance, suppose that a total demand realization is 50 with probability 0.001. For $\rho = 0.9$, we have $50 * 0.1 = 5$ as a reference target value for the total unmet demand with the probability of 0.001. The parameter $r_1 = 8$ is taken into consideration while defining the set of demand nodes that can be covered by a facility located at node $j \in J$ and the set of all candidate facility nodes that are within acceptable distance of node $i \in I$. For illustrative examples, please see Figures 1(a) and 1(b).

The parameter values $\rho = 0.9$ and $r_1 = 8$ are common in North America, but not necessarily in elsewhere in the world. However, the method described above can be applied for any parameter values.

5 Computational Results

In the following section we give some details on generating the problem instances. Then, in Section 5.2 we provide results to demonstrate the computational effectiveness of the proposed alternate formulations of ICCsP and SSDP. Section 5.3 presents numerical results illustrating the computational effectiveness of the heuristic developed for SSDP. In Sections 5.4 and 5.5 we present numerical results to analyze how the optimal location and allocation solutions change with respect to the input parameters and different risk preferences represented by ICCs and the SSD constraint. Finally, we discuss the computational study performed to compare the proposed models to the most relevant existing model (Beraldi et al., 2004).

5.1 Generation of problem instances

In order to test the computational performance of our solution methods, we considered several problem instances of different sizes. A total of 171 randomly generated problem instances are used to obtain the numerical results presented in this section. For a specified number of nodes ($N = |J| = |I|$) in the network, where each node is a demand point as well as a facility candidate, problem instances were randomly generated as follows:

- ⋄ We randomly generate the set of demand points in the $[0, 30]^2$ square according to a continuous uniform distribution as proposed by Gendreau et al. (1997). We set $d_{ij},\ i \in I,\ j \in J$, values to be the Euclidean distance between these points.
- ⋄ As Gendreau et al. (1997) we assume the side of the square region to be 30 km and the ambulance speed to be 40 km/h. Then, the coverage distance threshold D_c is $40*8/60 \approx 5.33$ km, when the response time standard is chosen to be 8 minutes.
- ⋄ The (annual) fixed facility cost vector $365 * 24 * \mathbf{f}$ is sampled from the uniform distribution on the interval $[1000, 4000]$. The (annual) cost of purchasing and maintaining an emergency vehicle $(365 * 24 * a)$ is 100. The cost per unit distance per unit demand (β) is 0.001 or 0.01.

⋄ Demand realizations at each node $i \in I$, $h_i(s)$, $s \in S$, are generated from a Poisson distribution with arrival rate parameter λ_i, which is sampled from the uniform distribution on the interval $[0.1, 0.8]$. As an alternative, in order to allow higher demand density in the city center, we divided the $[0, 30]^2$ square in 9 equal square zones. The arrival rate parameter for the nodes in the city center is sampled from the uniform distribution on the interval $[0.8, 1]$. For the corner zones we use the the interval $[0.1, 0.2]$, whereas, for the remaining four zones we use the interval $[0.3, 0.4]$.

⋄ The maximum number of vehicles that can be allocated to each facility $j \in J$, U_j, is sampled from the uniform distribution on the interval $[2, 4]$ or $[6, 10]$.

⋄ Scenario probabilities p_s, $s \in S$, are set to be equal or sampled from the uniform distribution on the interval $[0.2, 0.7]$ and then normalized.

⋄ The risk aversion parameter, α for individual ICCs are chosen to be 0.2, whereas the risk aversion parameter, δ, for the global ICC is chosen to be 0.04 or 0.02.

⋄ The proportion of demand covered within $r_1 = 8$ minutes (ρ) is 0.9 unless stated otherwise. As discussed in Section 4.4, the ρ parameter is used to generate reference random total unmet demand.

We would like to point out that generating the scenarios is not our main concern here. Existing methods can be applied to generate alternate scenarios, or if available, real historical data may be employed. Also note that the demand varying depending on the time of day or the day of the week, can be incorporated into the model by generating demand realizations for different hourly time periods of a week.

All problems were solved using the AMPL modeling language (Fourer et al., 1993) running on ILOG CPLEX 10.2 (as a reference, please see ILOG AMPL CPLEX System Version 10.0 User's Guide, 2006). The binary search algorithms were implemented in MATLAB R2006. The numerical experiments were performed on a 64-bit, 2 quad-core CPU HP workstation running on Linux. In our computational study, we terminate CPLEX when the prescribed CPU time limit ($t = 7200$ seconds) is reached.

In order to give an idea about the cost structure in the generated data sets, we present in Table 1 the average optimal objective function values decomposed into three types of cost components for 23 test problems presented in Table 2. We assume that the cost associated with the amount of labor needed to operate facilities is also incorporated into the fixed costs $f_i, i \in I$, associated with the facility locations. As labor is typically the dominating cost component (Goldberg, 2004), the cost parameters are chosen accordingly in order to give the largest priority to the fixed costs of facilities in our model.

Variable Cost $\sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij}$	Fixed Cost $\sum_{j\in J} f_j y_j$	Vehicle Cost $\sum_{i\in I}\sum_{j\in J} a x_{ij}$	Total Cost
6,880.57	57,096.65	36,752.17	100,729.39

Table 1: Average Optimal Cost Values Associated with Table 2.

5.2 On alternate formulations

Here, we present some results indicating how substantial the computational improvement is after decomposing the global scenarios into the local ones. Table 2 shows the CPU times that are in favor of AlterICCsP even for the small problem instances.

Problem Instances	$\sum_{i\in I} S_i$	DirectICCsP CPU (sec.)	AlterICCsP CPU (sec.)	Relative Reduction in CPU (%)
N=100 S=|500|	430	125.96	1.24	99.02%
N=150 S=|500|	652	852.15	7.68	99.10%
N=200 S=|500|	884	1780.02	23.78	98.66%
N=300 S=|500|	1326	1692.14	25.17	98.51%
N=400 S=|500|	1784	5924.99	94.08	98.41%
N=100 S=|1000|	459	2928.00	3.94	99.87%
N=150 S=|1000|	704	1134.38	4.34	99.62%
N=200 S=|1000|	941	7210.96*	19.34	>99.73%
N=300 S=|1000|	1383	7219.09*	20.58	>99.71%
N=400 S=|1000|	1904	7235.69*	44.10	>99.39%
N=100 S=|3000|	507	7210.17*	6.14	>99.91%
N=150 S=|3000|	772	7218.36*	30.82	>99.57%
N=200 S=|3000|	1011	7230.04*	70.83	>99.02%
N=300 S=|3000|	1556	7617.45*	109.67	>98.56%
N=400 S=|3000|	2073	7295.23*	184.85	>97.47%
N=300 S=|10000|	1671	N/A	34.29	N/A
N=400 S=|10000|	2230	N/A	426.12	N/A
N=200 S=|30000|	1205	N/A	29.25	N/A
N=300 S=|30000|	1841	N/A	47.91	N/A
N=400 S=|30000|	2409	N/A	67.11	N/A
N=200 S=|50000|	1241	N/A	18.36	N/A
N=300 S=|50000|	1906	N/A	40.23	N/A
N=400 S=|50000|	2471	N/A	170.89	N/A

Table 2: Performance of AlterICCsP versus DirectICCsP
* : time limit with integer solution.
N/A: No solution is available since CPLEX terminated due to solver error (ran out of memory).

It can be seen from Table 3 that when the formulation AlterSSDP is used, CPLEX could not provide optimal solutions within the time limit of 7200 seconds even for small instances of SSDP. However, these instances could be solved to optimality by using formulation AlterSSDPWithBS.

The proposed formulation AlterSSDPWithBS utilizing binary search algorithms substantially outperforms the formulation AlterSSDP and leads to great reductions in CPU time to solve SSDP. We remark that the CPU time to obtain the local sets $S_i,\ i \in I$, and the associated probabilities from a given set of global scenarios (S), and to calculate the lower bounds $LB_i^{\text{ICCs}},\ i \in I$, were negligible and therefore, we did not report them. For example, for a large problem instance with $N = 400$ and $|S| = 50000$, we obtain the sets $S_i,\ i \in I$, and the associated probabilities using MATLAB R2006 in 6.02 CPU seconds. For the same problem instance it took the binary search algorithms 0.06 CPU seconds in total to calculate the lower bounds, $LB_i^{\text{ICCs}}, i \in I$.

Problem Instances	AlterSSDP CPU (sec.)	AlterSSDPWithBS CPU (sec.)	Relative Reduction in CPU (%)
N=50 S=|300|	76.06	11.60	84.75%
N=75 S=|300|	118.14	14.42	87.80%
N=100 S=|300|	2129.35	82.21	96.14%
N=150 S=|300|	7201.81*	137.52	>98.09%
N=200 S=|300|	7202.95*	1043.47	>85.51%
N=50 S=|500|	43.07	5.33	87.62%
N=75 S=|500|	754.54	84.85	88.75%
N=100 S=|500|	7201.29*	1213.52	>83.15%
N=150 S=|500|	3692.95	442.86	88.01%

Table 3: Performance of AlterSSDPWithBS versus AlterSSDP

* : time limit with integer solution.

5.3 On the heuristic HSSDP

Results presented in this section illustrate the computational effectiveness of the heuristic HSSDP. Since it is hard to solve AlterSSDPWithBS for large problem instances, we solve smaller problem instances in order to calculate the relative optimality gaps, as defined below. We use the best known lower bound on the objective value found by the branch-and-bound algorithm of CPLEX, as most of the problem instances cannot be solved for optimality within the prescribed time limit. Let Obf_t^{LP} denote the best lower bound on the objective function value that is provided by the CPLEX solver, when the prescribed time limit t is reached. Obf^* denotes the objective function value obtained by HSSDP. The feasible solution obtained by the described heuristic algorithm gives an upper bound on the objective value. Then, we define the relative optimality gap associated with the heuristic HSSDP as follows:

$$\text{Relative optimality gap (ROG)} = \frac{\text{Obf}^* - \text{Obf}_t^{LP}}{\text{Obf}_t^{LP}}.$$

To take the randomness in data generation into account, 5 test problem instances were generated for each selected combination of parameters. Table 4 reports the CPU times and relative optimality gaps averaged over 5 randomly generated instances. This table shows that heuristic HSSDP is quite effective for generating feasible solutions with reasonably small relative optimality gaps. Numerical results presented in Sections 5.2 and 5.3 indicate that the proposed solution methods

Problem Instances	SSDP CPU (sec.)	HSSDP CPU (sec.)	Relative Reduction in CPU (%)	ROG
N=100 S=|300|	55.27	8.35	84.90%	2.28%
N=150 S=|300|	378.45	18.87	95.01%	1.49%
N=200 S=|300|	3839.03*	30.31	>99.21%	0.75%
N=300 S=|300|	7219.26*	87.69	>98.79%	0.74%
N=400 S=|300|	7209.33*	223.47	>96.90%	0.76%
N=100 S=|500|	2008.75*	18.13	>99.10%	3.46%
N=150 S=|500|	4844.03*	50.44	>98.96%	1.72%
N=200 S=|500|	7209.28*	75.31	>98.96%	1.23%
N=300 S=|500|	7216.89*	233.30	>96.77%	0.84%
N=400 S=|500|	7219.07*	469.80	>93.49%	0.56%
N=100 S=|1000|	5685.66*	54.41	>99.04%	2.91%
N=150 S=|1000|	6866.01*	155.04	>97.74%	2.43%
N=200 S=|1000|	7207.28*	304.96	>95.77%	1.89%
N=300 S=|1000|	7215.96*	889.93	>87.67%	1.09%
N=400 S=|1000|	7235.78*	1875.20	>74.08%	0.98%

Table 4: Performance of Heuristic HSSDP

* : time limit with integer solution (for at least one instance out of 5).

are quite sufficient to solve ICCsP and SSDP for large problem instances.

5.4 Solution sensitivity to input parameters

We solve ICCsP for a particular problem instance with $N = 200$ and $|S| = 500$ in order to check the sensitivity of the model to changes in input parameters. We refer to the originally generated test problem instance as the *base case*. The value(s) of only a certain type of parameter is (are) changed while everything else is kept fixed. The entries in the first column of Table 5 state which parameter values have been changed in the data of the *base case*. As expected the upper bounds on the number of vehicles that can be allocated to facilities have a significant effect on the number of facilities to be opened. It can also be seen from Table 5 that the risk parameter δ significantly effects the number of vehicles to be allocated. We will elaborate on the effects of the risk parameters more in the next section.

The fixed cost parameters of the *base case* are turned out to be large enough to construct solutions with the main objective of minimizing the number of facilities to be opened. That is why when we increase the values of parameters $f_i, i \in I$, the optimal solution does not really change. The results also show that when we decrease the values of the fixed costs $f_i, i \in I$, the number of vehicles allocated stays almost constant due to the risk constraints, whereas the number of facilities to be opened increases significantly. When the unit variable cost β is set to be a high value the model may result in more facilities to reduce the total cost. However, we prefer small values for the unit variable cost in order to give lower priority to the total variable cost than the other cost components. Due to the risk constraints imposing lower bounds on the number of vehicles, the value of the cost parameter $a > 0$ is not significant on the total number of vehicles allocated and the number of facilities to be opened.

Using a worst case scenario approach, a decision maker might construct a solution based on the highest demand realizations to avoid any unmet demand situations, i.e., demand would be satisfied under all scenarios. Notice that this approach is equivalent to enforcing a joint probabilistic constraint (8), where the probability level $\gamma^{(2)} = 1$. Instead of the number of vehicles, we report the ratio of the total number of vehicles allocated to the total number vehicles that would be required based on the worst case scenario approach. We will refer to this ratio as the "ratio of total number of vehicles".

	Total Cost	Total # of Facilities	Ratio of Total # of Vehicles	Total Expected Unmet Demand
Base Case	79,488.66	33.4	0.4233	8.39
Response time standard: 7 mins.	81,098.91	34	0.4239	8.41
Upper Bounds $U_i,\ i \in I$, Divided by 2	146,279.19	70.4	0.4233	8.39
Upper Bounds $U_i,\ i \in I$, Multiplied by 2	60,204.50	18.5	0.4241	8.42
Risk Parameter $\delta = 0.01$	110,141.35	45.4	0.5731	2.99
Risk Parameter $\delta = 0.1$	74,968.00	31.8	0.3979	10.08
Fixed Costs, $f_i,\ i \in I$, Divided by 20	35,635.77	40.00	0.4240	8.40
Fixed Costs, $f_i,\ i \in I$, Multiplied by 100	4,498,674.53	33.20	0.4242	8.41
Unit Purchase Cost, a, Divided by 1000	51,113.48	33.40	0.4251	8.45
Unit distance cost, β, Multiplied by 100	429,547.36	67.80	0.4380	8.82

Table 5: Results for Different Input Parameters (averaged over 5 test problems)

5.5 On risk constraints

For comparison purposes, we find optimal decisions by solving a problem based on a naive approach which uses expected demand values. We refer to the *underlying deterministic problem* (1)-(5), where h_i set to $\lceil \bar{h_i} \rceil,\ \forall\ i \in I$, as the *base problem*. The stochastic models, which consider the variability in demand, provide solutions with more flexibility to avoid unmet demand situations. Therefore, it is not surprising that the solutions of our models are better in terms of the amount of unmet demand than the ones found by the base problem (see Table 6).

As it can be seen from Table 6 the optimal numbers of vehicles obtained by ICCsP and SSDP are between the values obtained by the base problem and the worst case scenario approach (ratio of total # of vehicles < 1). Thus, risk measures deal with the problem of placing too much weight on extreme scenarios.

The level of risk modeled by ICCsP and SSDP depend on risk parameters and the specified distribution of the reference random outcome. Either of these two types of models may provide us with more risk averse solutions according to the risk parameter δ and the reference distribution. In

Problem Instances	Total Number of Facilities			Ratio of Total # of Vehicles			Total Expected Unmet Demand		
	Base Prob.	ICCsP	SSDP	Base Prob.	ICCsP	SSDP	Base Prob.	ICCsP	SSDP
N=100 S=|3000|	16	18	23	0.25	0.35	0.47	9.19	4.23	1.64
N=200 S=|3000|	24	33	41	0.25	0.34	0.42	19.39	8.28	4.50
N=300 S=|3000|	34	50	58	0.24	0.34	0.40	32.18	12.80	8.80
N=100 S=|5000|	15	17	24	0.23	0.33	0.44	9.80	4.30	1.57
N=150 S=|5000|	18	25	32	0.22	0.29	0.41	14.82	6.40	3.02
N=200 S=|5000|	23	32	40	0.23	0.32	0.40	18.93	8.27	4.35

Table 6: Total Number of Facilities, Ratio of Total # of Vehicles and Expected Unmet Demand by the Base Problem, ICCsP and SSDP

order to compare the level of risk aversion of the optimal solutions we can check the distribution of total unmet demand when exactly the optimal number of vehicles are assigned to each demand node, and the distribution of reference random total unmet demand. For example, for a problem instance where $N = 200$ and $|S| = 3000$, the optimal total number of vehicles required according to ICCsP is smaller than the amount required according to SSDP when $\delta = 0.04$ (see Table 7). For this instance, by setting the risk parameter δ to 0.04 and 0.01, and keeping everything else fixed, we obtain two different Figures 1(a) and 1(b). These figures show the reference cumulative distribution function and cumulative distribution functions of the random total unmet demand associated with the optimal solutions of ICCsP, SSDP and the base problem. We note that when $\delta = 0.04$ the total unmet demand for the optimal solution constructed by ICCsP is dominated by the reference random total unmet demand in the second order. Therefore, SSDP which requires the second order dominance with respect to the reference random outcome leads to a more risk averse solution than the one obtained by ICCsP. On the other hand, when $\delta = 0.01$ the reference random total unmet demand is dominated by the total unmet demand for the optimal solution obtained by ICCsP in the second order. Therefore, in this case the ICCsP results in a more risk averse solution; the total numbers of vehicles assigned to cover demand is higher (comparing ratios: $0.46 > 0.41$). In conclusion, we cannot claim that ICCsP provides more risk averse solutions than SSDP for all values of the risk parameter δ. It can also be seen from the results in Table 8 corresponding to a problem instance, where $N = 200$ and $|S| = 500$, we cannot claim that SSDP constructs more risk averse solutions than ICCsP for all reference random outcomes (for all values of the parameter ρ).

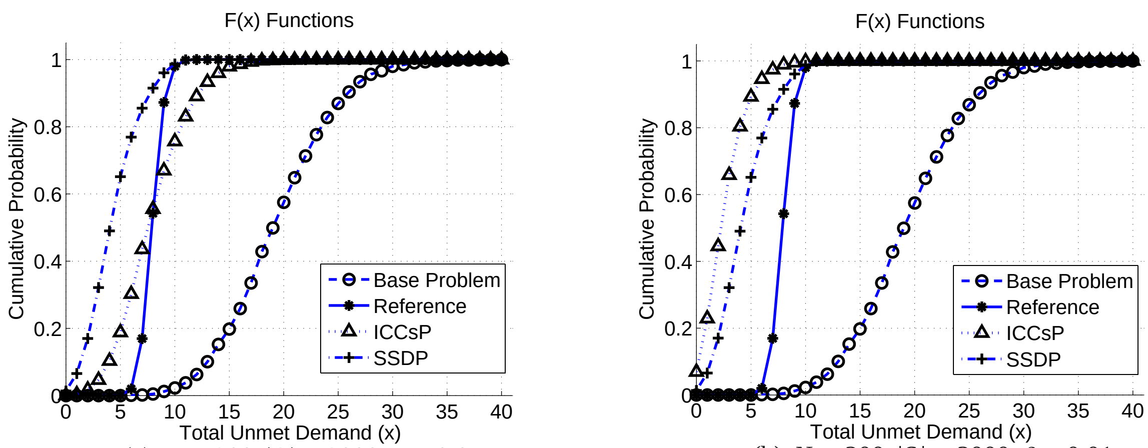
Risk parameter	For Reference Benchmark	Ratio of Total # of Vehicles		Total # of Facilities	
δ	ρ	ICCsP	SSDP	ICCsP	SSDP
0.04	0.9	0.34	**0.41**	31	38
0.01	0.9	**0.46**	0.41	43	38

Table 7: For an instance with $N = 200$ and S=|3000|.

Figure 2 illustrates that our models obtain more risk averse optimal solutions when the level of risk aversion gets higher. Thus, the total cost is nonincreasing while the risk aversion parameters α and δ are increasing, i.e., while the level of risk aversion is getting smaller. Figure 2 also

Risk parameter	For Reference Benchmark	Ratio of Total # of Vehicles		Total # of Facilities	
δ	ρ	ICCsP	SSDP	ICCsP	SSDP
0.04	0.85	**0.42**	0.41	34	33
0.04	0.9	0.42	**0.48**	34	39

Table 8: For an instance with $N = 200$ and S=|500|.



(a) $N = 200$, $|S| = 3000$, $\delta = 0.04$

(b) $N = 200$, $|S| = 3000$, $\delta = 0.01$

Figure 1: Reference Distribution Function and Distribution Functions of the Total Unmet Demand associated with the optimal solutions of ICCsP, SSDP and the Base Problem.

demonstrates the interaction between the local and global constraints. It is easy to see that for large enough δ values the global ICC constraint would be implied by ICCs (14) and it would become redundant. In this case increasing the value of δ would not affect the optimal solutions, since a certain level of service is ensured by the local constraints. Similarly, for large enough α values the global ICC constraint would be more significant and the optimal solutions are not affected by further increasing α values. Figure 2 is consistent with these observations; the total cost, the number of facilities and vehicles become constant after some large enough δ and α values. In particular, these large enough values are $\delta = 0.052$ and $\alpha = 0.17$, respectively. These values can be justified by solving ICCsP after dropping the global ICC. For the optimal solution of the modified ICCsP, which involves only the local constraints with $\alpha = 0.2$ and $\alpha = 0.17$, we calculate the associated value of the global risk parameter δ for which the global constraint (15) is (externally) satisfied as 0.052 and 0.0399, respectively. That is why we observe that results stay constant around $\delta = 0.05$ and $\alpha = 0.17$ in Figure 2. In order to provide a more clear understanding of the interaction between the local and global constraints, we also present figures for the optimal solutions obtained by ICCsP after dropping the global ICC. As expected, Figure 3 shows that the optimal solutions keep changing when the risk parameter α is modified since there is no effect of the global risk parameter δ.

Tables 6 and 9 provide results for several problem instances to show how the optimal solutions change with respect to different risk preferences. We also generate reference distributions to represent more risk averse preferences by increasing ρ, the proportion of demand that must be covered within $r_1 = 8$ minutes. It is not surprising that for larger values of ρ the model SSDP constructs more conservative solutions with higher total cost values.

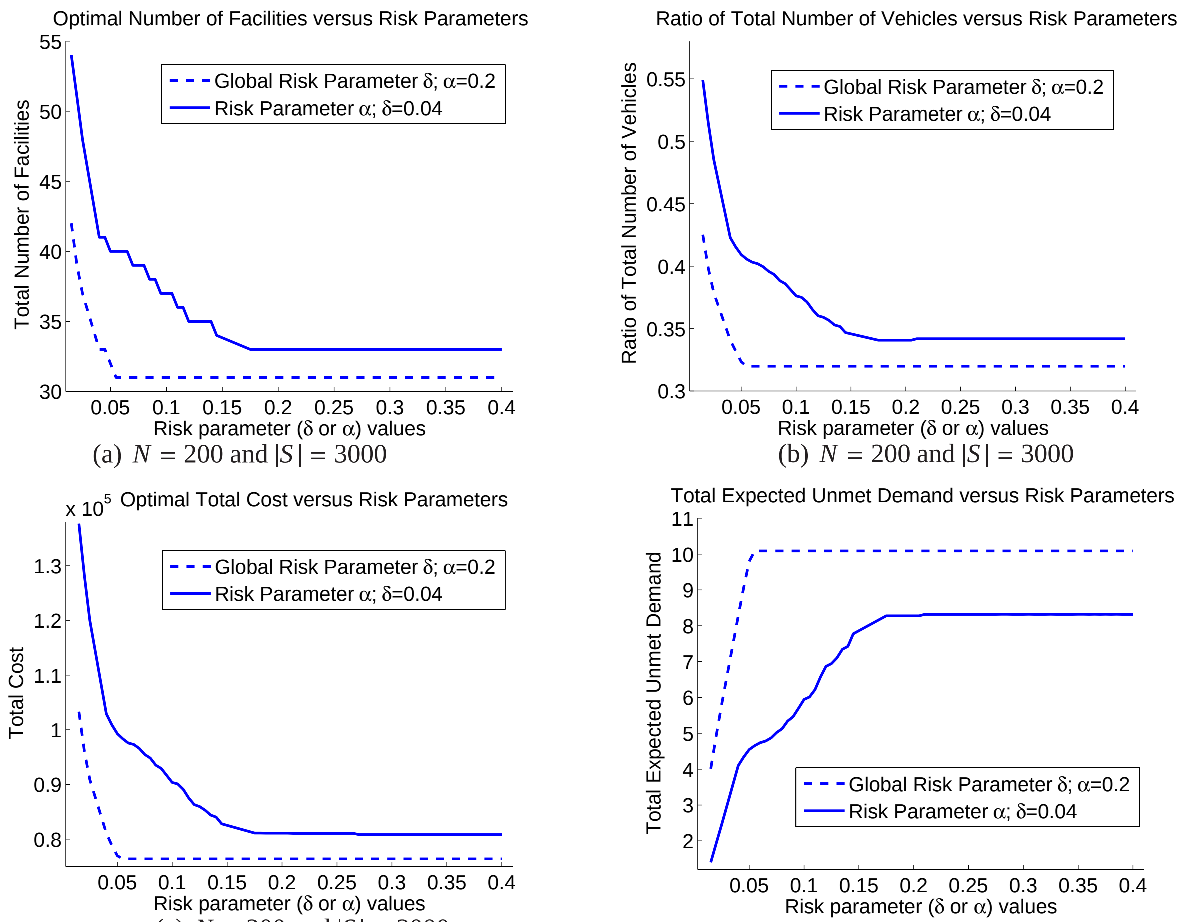
Problem Instances	For Reference Benchmark ρ	Total # of Facilities	Ratio of Total # of Vehicles	Total Cost	Total Expected Unmet Demand
N=200 S=|500|	0.8	32	0.39	73,590.76	10.29
N=200 S=|500|	0.85	33	0.41	77,758.72	9.19
N=200 S=|500|	0.9	39	0.48	91,198.71	6.24
N=200 S=|500|	0.95	50	0.62	120,686.61	2.37
N=200 S=|3000|	0.85	32	0.34	82,632.97	8.43
N=200 S=|3000|	0.9	38	0.41	100,510.27	4.80
N=200 S=|3000|	0.95	53	0.55	142,386.77	1.61

Table 9: Results by SSDP for Different Reference Random Total Unmet Demand

5.6 Comparison with an existing model

We perform a computational study to compare our models to the one proposed by Beraldi et al. (2004). In general, Beraldi et al. (2004) consider the following probabilistic constraints:

$$P\left(\sum_{j\in N_i} x_{ij} \geq h_i,\ \forall i \in G(k)\right) \geq 1 - \epsilon_k, \quad k = 1, \ldots, K, \tag{43}$$



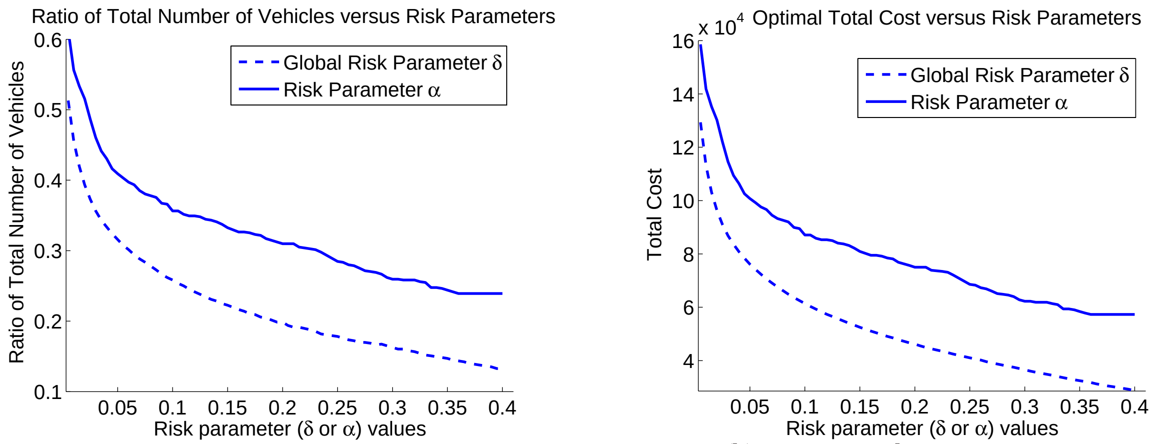
(a) $N = 200$ and $|S| = 3000$

(b) $N = 200$ and $|S| = 3000$

(c) $N = 200$ and $|S| = 3000$

(d) $N = 200$ and $|S| = 3000$

Figure 2: Results obtained solving ICCsP for different values of risk aversion parameters.



(a) $N = 200$ and $|S| = 3000$ (b) $N = 200$ and $|S| = 3000$

Figure 3: Results obtained solving ICCsP after dropping the global ICC for different values of risk aversion parameters.

where K denotes the number of sub-area indexed by k, $G(k)$ denotes the components of random demand vector associated with the group k and $1 - \epsilon_k$ is the prescribed probability level for the group k. They consider three cases described below as quoted from their paper:

1. " a global reliability system for the entire geographical territory (Jo);
2. an individual reliability system for each demand point (In);
3. an individual reliability system for sub-area (Jo-In)."

Notice that for the first case, $K = 1$ and for the second case K is equal to the number of demand nodes (each group consists of an individual demand node). In the last case we have divided the demand points in 5 (Jo-In 5) or 10 (Jo-In 10) sub-areas.

As mentioned before, the reformulation proposed by Beraldi et al. (2004) is not valid for a given set of scenarios. Therefore, in order to solve the problems which involve probabilistic constraints instead of our risk constraints, we need to consider another reformulation. A common approach introduces binary variables ζ_s associated with each scenario $s \in S$. Then, for a group $k \in K$ we reformulate (43) as

$$\sum_{j \in N_i} x_{ij} \geq h_i(s)(1 - \zeta_s), \qquad \forall i \in G(k),\ s \in S, \tag{44}$$

$$\sum_{s \in S} p_s \zeta_s \leq \epsilon_k, \tag{45}$$

$$\zeta \in \{0, 1\}^{|S|}. \tag{46}$$

When the binary variable ζ_s takes value of 0, it is guaranteed by (44) that all the inequalities $\sum_{j \in N_i} x_{ij} \geq h_i(s)$, $i \in G(k)$ hold. Constraints (44) and (45) require that demand at each node will be satisfied for a set of scenarios whose aggregate probability is at least equal to the enforced probability level $1 - \epsilon_k$. Since Luedtke et al. (2009) has been the most recent study on efficiently solving problems with joint probabilistic constraints, we have decided to implement their proposed formulation, which is a stronger formulation of (44)-(46). (For details please see Luedtke et al., 2009). In Tables 10 and 11 we present results obtained for three cases considered by Beraldi et al. (2004) and also the results obtained by our models for several risk parameters. We also report the minimum of the resulting probability of satisfying demand at each node.

Problems	Total # of Facilities	Total # of Vehicles	Total Cost	Total Expected Unmet Demand	Minimum of all local probabilities
Base Prob.	32	102	85508.51	5.86	0.71
Worst Case Approach	72	229	222779.59	0.00	1.00
Jo	63	207	195487.23	0.22	0.99
Jo-In(5)	55	180	162799.40	0.62	0.90
Jo-In(10)	49	161	142576.17	1.12	0.90
In	33	106	92416.44	4.35	0.90
ICCsP, $\alpha = 0.01$	61	199	181955.12	0.37	0.97
ICCsP, $\alpha = 0.05$	45	148	130488.77	1.55	0.95
ICCsP, $\alpha = 0.2$	41	137	116091.63	2.28	0.92
SSDP, $\alpha = 0.01$	50	170	154080.23	0.85	0.97
SSDP, $\alpha = 0.2$	45	147	127929.99	2.01	0.89

Table 10: Comparison results for a problem instance where $N = 100$ and $|S| = 100$

Problems	Total # of Facilities	Total # of Vehicles	Total Cost	Total Expected Unmet Demand	Minimum of all local probabilities
Base Prob.	29	102	78183.33	6.15	0.73
Worst Case Approach	*	268	*	*	*
Jo	76	242	223909.80	0.13	0.98
Jo-In(5)	56	189	156299.87	0.59	0.90
Jo-In(10)	48	160	127193.54	1.16	0.90
In	32	111	84612.15	3.78	0.91
ICCsP, $\alpha = 0.01$	59	195	165053.65	0.50	0.98
ICCsP, $\alpha = 0.05$	45	149	118316.89	1.46	0.94
ICCsP, $\alpha = 0.2$	39	114	103143.52	2.15	0.93
SSDP, $\alpha = 0.01$	49	166	135967.37	0.94	0.98
SSDP, $\alpha = 0.2$	45	152	120072.41	1.65	0.92

Table 11: Comparison results for a problem instance where $N = 100$ and $|S| = 300$
*:Not feasible; the total demand can not be satisfied due to the upper bounds on the number of vehicles that can be allocated to each facility.

As seen from Tables 10 and 11 similar results can be obtained using probabilistic constraints or the proposed alternate risk measures by varying the associated risk parameters. Obviously, considering just individual probabilities leads to high unmet demand values. On the other hand, enforcing a single joint probabilistic constraint is too conservative. A rational decision maker would prefer an approach somewhere between these two extremes. Such a decision maker can enforce joint probabilistic constraints for sub-areas as proposed by Beraldi et al. (2004) or can use ICCs or a stochastic dominance constraint depending on his/her risk preferences. As the performed computational study shows our proposed models and the associated solution methods enable the decision makers to solve relatively large problem instances.

6 Conclusion and Further Work

In situations where decisions are taken repeatedly under similar conditions, one can justify the optimization of the expected value by the Law of Large Numbers. However, the expectation is not a good optimality criterion in general: the standard stochastic programming problem maximizing expected coverage may provide solutions that result in poor performance for certain realizations of the random data. Therefore, risk measures should be incorporated into facility location and allocation problems in order to consider the inherent variability in the system and the decision makers' risk aversion. In this paper, we have shown that new stochastic programming models specifying alternate constraints on risk can be developed for the problem of designing an EMS system. We present a novel approach to EMS design problems by modeling risk through integrated chance and stochastic dominance constraints. The decision makers can evaluate different location and allocation decisions with respect to the quality of service and costs by varying the risk parameters. We presented numerical results to illustrate how the location and allocation solutions change with respect to different risk preferences. One of the main disadvantages of the scenario approach is the requirement to limit the number of scenarios for computational reasons. We developed alternative formulations and a heuristic for the proposed problems and by performing an extensive computational study we showed that we can overcome the main drawback of the scenario approach and solve relatively large problem instances.

In practice, the total demand varies depending on the time of the day or the day of the week. Our models can be used to solve the problem of determining the facility locations and the number of vehicles allocated to each facility. Then, given the optimal number of facilities and the allocation of the emergency vehicles, modified versions of our models, where the objective function represents the total hourly cost, can be solved 168 times for each hour of a week. The optimal allocation decisions for each time period would provide us with the number of vehicles required for each hourly time period of a week according to the specified risk preferences. Then, existing methods for reallocating emergency vehicles and multistage crew scheduling may be applied to improve the EMS system design.

The future research can focus on developing similar stochastic models which consider other features of an EMS system, such as different types of vehicles to serve different types of service requests. For example, EMS systems typically work with two types of service providers having different capabilities: basic life support units and advanced life support units. Finally, the dispatching rules that are related to the allocation of vehicles to specific nodes can also be modeled by using the proposed risk measures in the two-stage stochastic programming framework.

References

[1] Ball, M. O. and F. L. Lin, 1993, A reliability model applied to emergency service vehicle location, Operations Research 41, 18–36.

[2] Beraldi, P., M. E. Bruni, and D. Conforti, 2004, Designing robust emergency medical service via stochastic programming, European Journal of Operational Research 196, 323–331.

[3] Beraldi, P., and M. E. Bruni, 2009, A probabilistic model applied to emergency service vehicle location, European Journal of Operational Research 158 (1), 183–193.

[4] Berman, O. and D. Krass, 2001, Facility location problems with stochastic demands and congestion, in: Z. Drezner, H.W. Hamacher (Eds.), Facility Location: Applications and Theory, Springer-Verlag, New York, 331–373.

[5] Brotcorne L., G. Laporte, and F. Semet, 2003, Ambulance location and relocation models, European Journal of Operational Research 147, 451–63.

[6] Chapman, S., and J. White, 1974, Probabilistic Formulations of Emergency Service Facilities Location Problems, paper presented at the 1974 ORSA/TIMS Conference, San Juan, Puerto Rico.

[7] Charnes, A., W. W. Cooper, and G. H. Symonds, 1958, Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil, Management Science 4, 235–263.

[8] Chen, G., M. Daskin , M. Shen and S.Uryasev, 2006, The alpha-Reliable Mean-excess Regret Model For Stochastic Facility Location Modeling, Naval Research Logistics 53 (7), 617–626.

[9] Current, J., M. S. Daskin, and D. Schilling, 2002, Discrete network location models. In Z. Drezner and H. W. Hamacher, editors, Facility Location: Applications and Theory, chapter 3. Springer-Verlag, New York, 83–120.

[10] Daskin, M. S., 1983, A maximum expected covering location model: Formulation, properties and heuristic solution, Transportation Science 17 (1), 48–70.

[11] Daskin, M. S., 1995, Network and Discrete Location: Models, Algorithms and Applications, John Wiley and Sons, Inc., New York.

[12] Dentcheva, D., and A. Ruszczyński, 2003, Optimization with stochastic dominance constraints, SIAM Journal on Optimization 14, 548–566.

[13] Dentcheva, D., and A. Ruszczyński, 2006, Portfolio Optimization with First Order Stochastic Dominance Constraints, Journal of Banking and Finance 30(2), 433–451.

[14] Erkut, E., A. Ingolfsson, and G. Erdogan, 2008, Ambulance Location for Maximum Survival, Naval Research Logistics Quarterly 55 (1), 42–58.

[15] Erkut, E, A. Ingolfsson, S. Budge, 2008b, Maximum availability/reliability models for selecting ambulance station and vehicle locations: a critique. Working paper, available from http://www.business.ualberta.ca/aingolfsson/publications.htm.

[16] Felder, S., and H. Brinkmann, 2002, Spatial Allocation of Emergency Medical Services: Minimizing the Death Rate or Providing Equal Access?, Regional Science and Urban Economics, 32(1), 27–45.

[17] Fourer, R., D. M. Gay and B. W. Kernighan, 1993, AMPL: A Modelling Language for Mathematical Programming, The Scientific Press.

[18] Gendreau, M., G. LaPorte and F. Semet, 1997, Solving an ambulance location model by tabu search. Location Science 5 (2) 75-88.

[19] Goldberg, J.B., 2004, Operations Research Models for the Deployment of Emergency Services Vehicles, EMS Management Journal, 1(1), 20–39.

[20] Hadar, J., and W. Russell, 1969, Rules for ordering uncertain prospects, Amer. Econom. Rev. 59, 25–34.

[21] Hanoch, G., and H. Levy, 1969, The efficiency analysis of choices involving risk, Rev. Econom. Stud. 36, 335–346.

[22] ILOG CPLEX, 2006, ILOG AMPL CPLEX System Version 10.0 User's Guide, ILOG CPLEX Division.

[23] Ingolfsson, A., S. Budge., E. Erkut, 2007, Optimal ambulance location with random delays and travel times, Health Care Management Science, DOI: 10.1007/s10729-007-9048-1.

[24] Klein Haneveld, W. K., 1986, Duality in Stochastic Linear and Dynamic Programming, Lecture Notes in Economics and Mathematical Systems 274, Springer-Verlag, New York.

[25] Klein Haneveld, W. K. and M. H. Van Der Vlerk, 2006, Integrated Chance Constraints: Reduced Forms and an Algorithm, Computational Management Science Vol.3, Number 4, 245–269.

[26] Lehmann, E., 1955, Ordered families of distributions, Annals of Mathematical Statistics 26, 399–419.

[27] Levy, H., 1992, Stochastic dominance and expected utility: survey and analysis, Management Science 38, 555–593.

[28] Louveaux, F.V., 1986, Discrete stochastic location models, Annals of Operation Research 6, 23-34.

[29] Louveaux F.V., 1993, Stochastic location analysis, Location Science 1(2), 127–154.

[30] Luedtke, J., 2008, New Formulations for Optimization Under Stochastic Dominance Constraints, accepted for publication in SIAM Journal on Optimization.

[31] Luedtke, J., S. Ahmed, G. Nemhauser, 2009, An Integer Programming Approach for Linear Programs with Probabilistic Constraints, to appear in Mathematical Programming.

[32] Mann, H. B., and D. R. Whitney, 1947, On a test of whether one of two random variables is stochastically larger than the other, Annals of Mathematical Statistics 18, 50–60.

[33] Marianov, V. and C. ReVelle, 1995, Siting Emergency Services. Facility Location: A Survey of Applications and Methods, ed. Z. Drezner, Springer.

[34] Marianov, V. and C. ReVelle, 1996, The queueing maximal availability location problem: a model for the siting of emergency vehicles, 1996, European Journal of Operational Research 93, 110-120.

[35] Marsh, M., and D., Schilling, 1994, Equity Measurement in Facility Location Analysis: A Review and Framework, European Journal of Operational Research 74 (1), 1–17.

[36] Miller, L. B. and H. Wagner, 1965, Chance-constrained programming with joint constraints, Operations Research 13, 930–945.

[37] Müller A., and D. Stoyan, 2002, Comparison Methods for Stochastic Models and Risks, John Wiley & Sons, Chichester.

[38] Owen, S. H. and M. S. Daskin, 1998, Strategic facility location: A review. European Journal of Operational Research 111(3), 423–447.

[39] Prékopa, A., 1970, On probabilistic constrained programming, in Proceedings of the Princeton Symposium on Mathematical Programming, Princeton University Press, Princeton, NJ, 113–138.

[40] Prékopa, A., 1995, Stochastic Programming, Kluwer Academic, Dordrecht, Boston.

[41] ReVelle C., and K. Hogan, 1989, The maximum availability location problem, Transportation Science 23, pp. 192-200.

[42] Rockafellar, R. T. and S. Uryasev, 2000, Optimization of Conditional Value-at-Risk, Journal of Risk 2, 21–41.

[43] Snyder, L.V., 2006, Facility location under uncertainty: a review. IIE Transactions 38(7),537–54.

[44] Snyder, L. V. and M. S. Daskin, 2006, Stochastic p-Robust Location Problems, IIE Transactions, 38 (11), 971–985.