Flex Serv Manuf J (2018) 30:123–152
https://doi.org/10.1007/s10696-017-9286-6

A stochastic approach for designing two-tiered emergency medical service systems

Rania Boujemaa[1] · Aida Jebali[2] · Sondes Hammami[1,3] · Angel Ruiz[4] · Hanen Bouchriha[1]

Published online: 22 March 2017

Abstract Emergency medical services (EMS) systems provide out-of-hospital acute medical care and transportation to the appropriate health care provider to patients with illnesses and injuries. The objective of EMS systems is to satisfy demand requests by providing timely first care medical assistance to patients at the incident scene. This paper aims at designing a robust two-tiered EMS system while accounting for the inherent uncertainty of the demand. A two-stage stochastic programming location-allocation model is proposed to simultaneously determine the location of ambulance stations, the number and the type of ambulances to be deployed, and the demand areas served by each station. This problem is then solved efficiently using the sampling average approximation algorithm. Computational experiments highlight the performance of the proposed solution approach and its practical applicability.

Keywords Stochastic programming · Emergency medical services (EMS) · Location-allocation model · Sample average approximation (SAA)

✉ Aida Jebali
aida_jebali@yahoo.fr

[1] Laboratoire de Recherche Analyse, Conception et Commande des Systèmes, Ecole Nationale d'Ingénieurs de Tunis, Université de Tunis El Manar, 1002 Tunis, Tunisia

[2] Department of Industrial Engineering and Engineering Management, University of Sharjah, PO Box 27272, Sharjah, United Arab Emirates

[3] Ecole Nationale d'Ingénieurs de Carthage, Université de Carthage, 2035 Tunis, Tunisia

[4] Département d'opérations et systèmes de décision, and CIRRELT, Faculté des sciences de l'administration, Université Laval, Quebec G1K 7P4, Canada

1 Introduction

Emergency medical services (EMS, for short) are a critical component of any healthcare system. Over the last four decades, the design of EMS systems has received the scientific community's growing attention due to its crucial role in saving lives. The implementation of an efficient EMS system can reduce human suffering and economic losses resulting from disabilities due to injuries and sudden illnesses by providing the fastest and the highest quality of healthcare services available in a pre-hospital setting. The effectiveness and efficiency of an EMS system are gauged through the response time, i.e. the time between an emergency call and the ambulance's arrival at the incident scene. There is a common consensus that out-of-hospital care must be provided in a timely manner, mainly for life-threatening emergencies.

Even though only a few studies have discussed the relationship between the response time of EMS and the lives saved (Blanchard et al. 2012), the current findings underline the importance of reducing the response time to increase the likelihood of patient survival. For instance, O'Keeffe et al. (2010) claimed that the response time is the most important predictive factor for patient survival. They demonstrated that the estimated effect of a 1 min reduction in response time improved the odds of survival by 24%. Moreover, Gonzales et al. (2009) established that an increased EMS response time is associated with higher mortality rates. Furthermore, it has been shown that using first responders (usually fire and rescue services) to reduce response time increases survival rates (Sund et al. 2011).

The response time achieved by an EMS system depends on several factors. Some of them are defined by the system designer, like the number and the locations of EMS stations (hereafter, EMS and ambulance stations will be used interchangeably), the number and type of ambulances assigned to each station and the ambulance dispatch policy (the assignment of an available ambulance to a call), but others are random (such as traffic conditions, the number of ambulances available when the call is received, etc.). A wide stream of research has been dedicated to study the design and the deployment of EMS systems, mostly seeking to achieve two objectives: provide an acceptable response time and contain the costs related to the system's operation (Beraldi et al. 2004). Some papers investigate EMS facility location problems at the strategic level, aiming at ensuring the best demand coverage to the population within a given geographical area. Other works consider the tactical decision level that aims to define the type and the number of resources to allocate to each EMS station. But the majority of the literature is dedicated to the operational level of EMS systems management. The latter addresses the ambulance dispatching and relocation problem, which aims at deciding which ambulance to assign to each demand request and, where to locate available ambulances at the short run. Markedly, the design and management of an EMS system bear challenging problems, especially when subjected to economic constraints and different uncertainty sources.

In this paper, the emphasis is placed on the EMS system's design, hereafter, referred to as the ambulance location-allocation problem. This involves a strategic-tactical decision that encompasses the selection of ambulance base stations to open and the number of ambulances to house at each of these stations. But unlike most of the works addressing such strategic-tactical decision level problem, uncertainty on the demand is considered explicitly. Three decisions are tackled jointly: (1) where to site ambulance stations (2), which demand zones will be served by each ambulance station and vice versa, and (3) the type and the number of ambulances to assign to each station. Two types of ambulances are considered: (1) basic life support (BLS) ambulances equipped with basic equipment and (2), advanced life support (ALS) ambulances able to perform life-saving procedures in addition to all the procedures performed by BLS ambulances. That is, ALS ambulances can respond to a call requesting a BLS care level. Henceforth, the paper addresses the design of two-tiered ‘‘successively inclusive’’ EMS systems while accounting for demand uncertainty. The problem is formulated as a two-stage stochastic programming model and solved using the sample average approximation (SAA) algorithm. The proposed model aims at ensuring an efficient and cost-effective coverage of the demand within a threshold response time while accounting for demand uncertainty.

The main contribution of this paper stems from the novelty of the proposed model. It is worth noting that the literature related to EMS system design underlines the paucity of papers that handle the design of two-tiered EMS systems under demand uncertainty. This work is designed to start filling this gap.

The remainder of this paper is organized as follows: Sect. 2 presents the relevant literature related to our study. Section 3 comprises the problem's description and mathematical model's formulation. The SAA algorithm is detailed in Sect. 4. Section 5 is devoted to the presentation and discussion of the experimentation results. Finally, concluding remarks and directions for future research are given in Sect. 6.

2 Literature review

Various modelling approaches have been devoted to the study of the ambulance location-allocation problem. The proposed models are classified within the wide body of the literature dedicated to capacitated facility location problems (CFLP). It is, however, worth noting that extensive literature reviews on facility location models (Current et al. 2001; Snyder 2006) reveal that facility location models are application specific and that a generic and basic model that can be adapted to all potential applications does not exist (Current et al. 2001). For instance, capacity constraints in ambulance location-allocation models should prevent the assignment of an incoming call to a busy ambulance. This feature related to ambulance congestion conspicuously distinguishes the ambulance location-allocation model from the other facility location models. Therefore, we will focus on the literature stream devoted to the design of EMS systems. Furthermore, we do not aim to review all developed approaches, but rather consider the approaches that are most relevant to our research. The focus will be henceforth placed on papers that either tackle ambulance location-allocation problems while accounting for demand uncertainty, or address the design of two-tiered EMS systems.

As far as uncertainty is concerned, the earliest papers addressed the randomness of the availability of emergency vehicles. The proposed models considered the probability of each ambulance being unavailable to respond to a call, referred to as a busy fraction (Daskin 1983; ReVelle and Hogan 1989). The shortcoming of these models is related to the estimation of emergency vehicles' busy fractions that are a priori unknown as they depend on ambulance location's plan and the demand. Therefore, EMS systems design has attracted significant attention in the recent years to deal with demand uncertainty in order to incorporate the randomness of EMS systems in an explicit way. Some other works considered the randomness of ambulances travel time (Ingolfsson et al. 2008; Erkut et al. 2008).

Three main approaches have been proposed for ambulance location-allocation under demand uncertainty: (1) stochastic programming models, (2) robust programming models (Zhang and Jiang 2014; Lam et al. 2016) and (3), queuing models (Larson 1974, 1975; Iannoni et al. 2011; Geroliminis et al. 2011). Obviously, stochastic programming models for ambulance location-allocation are those that are particularly tied to our work. In this kind of model, contrarily to robust optimization, it is assumed that the probability distributions governing the data are known or can be estimated.

To design a reliable EMS system, Ball and Lin (1993), Beraldi et al. (2004), Beraldi and Bruni (2009) and Noyan (2010) developed chance-constrained stochastic programming models where the main uncertainty was assumed to be due to the stochastic call arrival process. The reliability is represented by the EMS's capability to guarantee a target service level while ensuring that demand coverage is kept above a specified value of probability for all demand areas. EMS's reliability is enforced in the proposed models by the chance constraints. In Ball and Lin (1993), the authors developed a reliability model where system failure is entailed by a vehicle's unavailability to respond to a request within acceptable time. Based on a bound on the probability of system failure, the authors transformed their initial model into a 0–1 integer programming optimization model. However, as it was pointed out by Erkut et al. (2008), the probability of having an available vehicle within a standard time is seldom used in practice. Instead of focusing on the randomness in the availability of vehicles, Beraldi et al. (2004), Beraldi and Bruni (2009) and Noyan (2010) focused on the demand satisfaction's randomness that allows for directly determining the service level, i.e., the fraction of calls covered within a response time below a predetermined threshold, which is a common performance measure used by EMS managers. In Beraldi et al. (2004), the authors provided a deterministic equivalent formulation of the chance constraints using the so-called p-efficient points of a joint probability distribution function. This formulation is based on the assumption that the demand is independent. The latter can be relaxed when a scenario-based formulation is devised (Beraldi and Bruni 2009; Noyan 2010). However, it is worth noting that the demand independence assumption is well justified in normal operating conditions. A correlation among demand points can only be established in the case of large-scale emergency situations (Beraldi et al. 2004). Beraldi and Bruni (2009) introduced chance constraints in the traditional two-stage stochastic programming framework. The facilities' location and the definition of the corresponding capacities present the first-stage strategic decisions. Once uncertainty has been resolved, tactical decisions concerning the allocation of demand points to facilities are taken while considering non-splittable demand, i.e., each demand point must be served by exactly one ambulance station under each scenario. In order to account for the cost-reliability trade-off, the authors introduce chance constraints. Henceforth, the decision makers can assess EMS system costs for different reliability levels. In Noyan (2010), the author proposed two models that account for target service level by including risk measures on random unmet demand. The first model incorporates the integrated chance constraints (ICCs) and the second one includes ICCs and a stochastic dominance constraint to account for the largest acceptable expected unmet demand. He modelled the random demands using the scenario approach. With the addition of ICCs constraints, the unmet demand is capped to a predefined nonnegative risk aversion parameter that represents the largest acceptable expected unmet demand value. It is worth noting that in the two-stage stochastic programming model, the ICCs also serve to restrict the risk to the solvability of the second-stage problem. In order to address the complexity inherent to the two-stage formulation, the author introduced the ICCs in the single-stage formulation, where it is assumed that the assignment of vehicles to demand nodes is scenario-independent. However, this single-stage formulation incurs a significant increase of the total system's cost compared to its counterpart two-stage formulation. More recently, Zhang and Li (2015) devised a novel stochastic model with chance constraints to design EMS systems by considering the randomness in the maximum number of concurrent demands occurring at a demand site over a day. The original model is transformed into a conic quadratic mixed-integer program by approximating the chance constraints as a second-order cone constraints. The obtained model is then solved by a commercial solver for problem instances of practical sizes. Zhang and Jiang (2014) proposed a bi-objective robust optimization model to design a cost-responsiveness efficient EMS system while considering the maximum number of concurrent demands occurring at a demand site over a day. This kind of formulation, contrarily to the stochastic programming approach, is used when there is little probability information on the uncertainty of the demand. The latter is accounted for in the robust optimization model through the definition of ellipsoidal uncertainty sets. In Nickel et al. (2015), the authors developed a scenario-indexed model to locate and size ambulances based on stochastic demand. The objective was to minimize the total cost associated with the EMS facilities. Nevertheless, the objective function does not consider the minimization of the expected costs of ambulance deployment. A pre-specified coverage level is enforced by ensuring that the expected number of ambulances allocated to a demand point is greater than or equal to the product of the considered service level factor and the expected demand. A sampling approach is proposed to solve the problem. Through numerical experiments conducted on small-sized problem instances, the authors highlighted the relevance of using a stochastic approach to design EMS systems. Lam et al. (2016) proposed a two-stage stochastic programming model to find an ambulance deployment that minimizes the overall shortfall in demand coverage. They reformulated their problem as a robust mathematical program by replacing the chance constraints with a set of deterministic constraints based on the Poisson arrival rates and Markov inequality and solved it using a standard solver. Van Essen et al. (2014) presented a two-stage stochastic program for the joint strategic and tactical ambulance planning. At the first-stage, the model determines the location of ambulance bases and the number of ambulances to assign to each opened ambulance base. In a second-stage, the model considers all the potential demand scenarios and specifies which ambulance to dispatch to which emergency call. As an ambulance is busy for about 1 h when responding to an emergency, the model incorporates scenarios based on the number of hourly incoming emergency calls. The demand scenarios are generated from a stochastic Poisson process. The authors sought to minimize the weighted combination of the number of located ambulances and the number of bases. The fraction of demand coverage, which does not have to be the same for each region (differs for urban and rural areas), is accounted for through the constraints. The problem is then solved by adopting a two-stepped heuristic approach: the first step solves the strategic level (location of ambulance bases' decision) and the second step considers the tactical level (number of ambulances' decision). The authors considered different types of demand resulting in different coverage requirements but covered by one type of vehicle.

As it can be noticed, all the aforementioned papers assumed that the ambulances were equally equipped. The papers that considered different types of ambulances developed deterministic and hypercube models. The first models were the tandem equipment allocation model (TEAM) and the facility location equipment emplacement technique (FLEET) model proposed in Schilling et al. (1979). These two models extended the maximal covering location problem (MCLP) presented in Church and ReVelle (1974) by enforcing the hierarchy between the two types of vehicles considered. A demand node was covered only if it was equipped with the two types of vehicles with the prescribed standards. The TEAM supposed that specific equipment could only be located in tandem with basic equipment. Conversely to the TEAM model, the FLEET model considered that the two-vehicle types were free to be located either in tandem or individually. These models were deterministic and were applied to locate vehicles of the Fire Protection System in the Baltimore City. ReVelle and Marianov (1991) and Marianov and ReVelle (1992) extended the model's framework to the fire protection system while considering the busy fraction. Their model sought to distribute standard and specific vehicles in order to maximize the population served by both types of vehicles within specified time standards, either with independent availabilities or with a joint availability. Mandell (1998) proposed a probabilistic covering-type model for two-tiered EMS systems in which it was not necessary to restrict the sites of ALS units to those sites where a BLS unit is located. A call's response was adequate if: (1) an ALS ambulance arrived within time t_B or (2) a BLS ambulance arrived in t_B time and an ALS ambulance arrived within time t_A ($t_B < t_A$). Hence, the model did not consider multiple types of calls and assumed the same response time standard t_B for every call. The model took server availability into account through a two-dimensional queuing model, thereby avoiding the need to assume the independence of server busy probabilities. Maximizing the expected coverage was used as the objective function. More recently, McLay (2009) extended the maximum expected coverage location problem (MEXCLP) introduced by Daskin (1983) to optimally locate servers for public service applications to include two types of medical units (the ALS non-transport Quick Response Vehicles (QRV) and the BLS) and multiple customer types (based on case acuity, an emergency call is classified of Priority 1, 2 or 3). A Hypercube queuing model is developed to estimate the busy probabilities of each vehicle type and the fraction of time where calls of a given type are assigned to their preferred server type. The model supposed that each type of demand arrived according to a Poisson process. The results of the hypercube model were then used to input an integer programming model (referred to as MEXCLP2) that determined the number of ALS QRV and BLS vehicles to locate at each facility. The objective was to maximize the expected number of Priority 1 calls (life-threatening calls) satisfied within a given response time threshold. However, such approaches require defining the possible combinations of medical units that might be dispatched and solving the problem for each of these combinations. Given that the number of combinations significantly increases in real-life problem instances, these approaches become of little practical use because of their computational intractability.

The literature discussed above shows that there are some recent papers that model the ambulance location-allocation problem as a two-stage stochastic programming model with recourse. These papers either reformulate the problem as a deterministic robust model (Lam et al. 2016) or solve it using heuristics (Van Essen et al. 2014). In all cases, limited discussion on the quality and the characteristics of resulting solutions is provided.

Moreover, the models that used stochastic programming to deal with the ambulance location-allocation problem do not consider the two types of commonly used ambulances in EMS systems. The papers that address this realistic feature develop deterministic models that can use input from queuing models to take the dependencies between ambulances (servers) into account.

In this paper, we propose to investigate the ambulance location-allocation problem with two types of vehicles under demand uncertainty. To directly deal with this uncertainty, we propose a cost-based two-stage stochastic program with recourse where the demand is assumed to be a random variable with known probability distribution. This model is solved using the SAA algorithm that allows for computing lower and upper bounds for problem solutions and providing the corresponding optimality gaps. To the best of our knowledge, this is the first paper that informs on the quality of the generated solutions. Hence, the proposed solution approach is a valuable contribution in itself. A detailed problem formulation and solution approach is detailed in the following two sections.

3 The ambulance location-allocation model

3.1 Problem description

In this paper, we address the design of a two-tiered EMS system under demand uncertainty. The EMS system is equipped with two types of ambulances (ALS and BLS units). Life-threatening calls, such as those involving a cardiac arrest require an ALS care level. Conversely, non-life-threatening calls require a BLS care level. An

ALS can be used to serve a non-life-threatening call. However, a BLS is under-equipped for a life-threatening call. The target response time for a life-threatening call is obviously shorter than that of a non-life-threatening call.

A two-stage stochastic programming model with recourse is proposed to locate ambulance stations, make decisions on number of vehicles of each type to be housed at each opened station, and the allocation of ambulances to demand points. Based on the time horizon of impact, the location of ambulance stations and the definition of their capacities are first-stage decisions. These decisions cannot be changed on the short run. Conversely, the allocation of ambulances to demand zones is decided based on demand realization, and thus constitutes a second-stage decision. The number of life-threatening calls and non-life-threatening calls coming from a demand point within a defined time interval is assumed to follow a Poisson distribution (Ingolfsson et al. 2008). An ambulance is busy for a certain amount of time when responding to an emergency call. Similarly to most studies in this area, namely Beraldi et al. (2004), Beraldi and Bruni (2009), Noyan (2010), Van Essen et al. (2014) and Zhang and Jiang (2014), we assume that 1 h is a reasonable time requirement for a service trip. Based on this assumption, the proposed model formulates the ambulance location-allocation problem over a horizon of 1 h and does not explicitly consider time. In this formulation, each ambulance can be assigned to at most one emergency call over 1 h (Van Essen et al. 2014).

The uncertainty related to the number of emergency calls coming from each demand point over 1 h is denoted by a scenario ξ. A scenario defines the vector of outcomes of two independent random variables: the number of life-threatening calls and the number of non-life-threatening calls arriving to each demand point over 1 h. The set of scenarios Ξ is supposed to be finite and $P(\xi)$, the probability distribution of scenario $\xi \in \Xi$ is discrete.

The objective is to minimize the location-allocation costs incurred by the system's infrastructure (the fixed cost of setting up the elected ambulance stations and the fixed per-unit capacity cost related to the cost of ambulances) and the expected transportation and penalty cost. The latter accounts for unsatisfied demand. It is therefore tied to the reliability of the EMS system to provide adequate demand coverage.

We should point out that, as we consider 1 h horizon, the strategic cost incurred by the opening and the per-unit capacity cost needs to be related to the incumbent time. Thus, the depreciations presented in the first-stage cost are converted to hourly costs. Moreover, the tactical cost is composed of the expected transportation and penalty costs entailed by the assignment of ambulances and the unsatisfied demands within 1 h time.

At this point, it is worth noting that only the transportation costs between the ambulance station and the location of the incident scene (the demand point) are accounted for. This can be justified by the fact that the location of ambulance stations does not depend on the decision regarding the hospital to which the patient will be transported and admitted. Indeed, this decision is made at a later stage by involving more accurate information on the patient case, hospital network and resource availability. Additionally, given that the transportation cost is proportional to the distance and the time between ambulance station and demand zone, its minimization tends to favour the design of an EMS system that tries to reduce the response time. As it can be noticed, even though the proposed model is not intended to tackle the ambulance operational planning level, the consideration of transportation costs entails an ambulance assignment to emergency calls in adherence with the closest ambulance dispatch rule.

Additionally, it is worth mentioning that, in cases where the ambulance service trip time differs from 1 h, it suffices to consider the Poisson distribution that reflects the number of life-threatening calls and non-life-threatening calls coming within this time frame. Similarly, the opening and the per-unit capacity costs should be tied to the considered horizon.

We consider a finite set of demand points and a set of potential stations where ambulances may be located. A demand point can only be covered by an ambulance station if the average travelling time between the demand point and the ambulance station is within a threshold value. Given that life-threatening calls may occur at any demand point, this threshold value is set to the target response time for this type of call.

The two-stage stochastic programming EMS model is described in the following paragraph.

3.2 A two-stage stochastic programming model

Let us introduce the following notation for our two-stage stochastic program with recourse, referred to as (SPEMS). We begin by introducing the parameters of the model (Table 1).

The following decision variables are used in the model formulation, with the two latter variables pertaining to the second-stage model (Table 2).

The proposed model assumes that each demand point can be served by more than one ambulance station (splittable demand), and that the ambulances housed at a given ambulance station are not dedicated to serve a specific demand point. Given that the variable X_{ijkl}^{ξ} also represents the number of ambulances of type k located at site j that serve calls of type l occurring at demand point i under scenario ξ, the total number of vehicles of type k allocated to the ambulance station j is determined by:

$$\max_{\xi=1..|\Xi|} \sum_{i\in I} \sum_{l\in L} X_{ijkl}^{\xi}.$$

At this point, we can present the complete formulation of the two-stage stochastic program with recourse (SPEMS):

$$Min \sum_{j\in J} f_j Y_j + \sum_{j\in J} \sum_{k\in K} Z_{jk} V_k + \sum_{\xi\in\Xi} p^{\xi} Q(Z, Y, \xi) \tag{1}$$

Subject to:

$$\sum_{j\in J} Z_{jk} \leq P_k \quad \forall k \in K \tag{2}$$

Table 1 Indices and parameters

Notation	Description
I	Set of demand points
i	Demand points index, $i = 1\ldots\lvert I\rvert$
J	Set of potential locations
j	Index of potential locations for an ambulance station, $j = 1\ldots\lvert J\rvert$
L	Set of demand types (life-threatening calls and non-life-threatening calls)
l	Demand types index, $l = 1\ldots\lvert L\rvert$
K	Set of ambulance types (ALS and BLS)
k	Ambulance types index, $k = 1\ldots\lvert K\rvert$
Ξ	Set of scenarios
ξ	Scenarios index, $\xi = 1\ldots\lvert\Xi\rvert$
Δ_{kl}	1 If ambulance of type k can cover demand of type l; 0 otherwise
δ_{ijk}	Travelling time between ambulance station j and demand zone i by ambulance of type k
R_l	Response time threshold for a call of type l
N_{ijkl}	1 If a call of type l occurring at demand point i can be covered by an ambulance of type k located at station j within response time threshold R_l (i.e. $N_{ijkl} = 1$ if $\delta_{ijk}\Delta_{kl} \leq R_l$); 0 otherwise
P_k	Maximum number of ambulances of type k that can be allocated to the EMS system
d_{il}^{ξ}	Number of calls of type l arriving from demand point i under scenario ξ
f_j	Fixed cost of opening ambulance station j
V_k	Fixed cost of ambulance of type k
W_{ijk}	Travelling cost between ambulance station j and demand point i using ambulance of type k
E_l	Penalty cost for not satisfying a call of type l
p^{ξ}	Probability of scenario ξ
M	A sufficiently large number

Table 2 Decision variables

Notation	Description
Y_j	1 If ambulance station j is open; 0 otherwise
Z_{jk}	Number of ambulances of type k to be housed at station j
X_{ijkl}^{ξ}	Number of calls of type l occurring at demand point i that are covered by ambulances of type k located at site j under scenario ξ
U_{il}^{ξ}	Number of unsatisfied calls of type l coming from demand point i under scenario ξ

$$Y_j \in \{0, 1\} \quad \forall j \in J \tag{3}$$

$$Z_{jk} \in \mathbb{Z}^{+} \quad \forall j \in J, \forall k \in K \tag{4}$$

With:

$$Q(Z,Y,\xi) = Min \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \sum_{l \in L} W_{ijk} \Delta_{kl} X^{\xi}_{ijkl} + \sum_{i \in I} \sum_{l \in L} E_l U^{\xi}_{il} \tag{5}$$

$$\sum_{i \in I} \sum_{k \in K} \sum_{l \in L} X^{\xi}_{ijkl} \leq MY_j \quad \forall j \in J,\ \forall \xi \in \Xi \tag{6}$$

$$d^{\xi}_{il} \leq U^{\xi}_{il} + \sum_{j \in J} \sum_{k \in K} N_{ijkl} X^{\xi}_{ijkl} \quad \begin{array}{c} \forall i \in I, \forall l \in L, \\ \forall \xi \in \Xi \end{array} \tag{7}$$

$$Z_{jk} \geq \sum_{i \in I} \sum_{l \in L} X^{\xi}_{ijkl} \quad \begin{array}{c} \forall j \in J,\ \forall k \in K, \\ \forall \xi \in \Xi \end{array} \tag{8}$$

$$U^{\xi}_{il}, X^{\xi}_{ijkl} \in \mathbb{Z}^{+} \quad \forall i \in I,\ \forall j \in J,\ \forall k \in K,\ \forall l \in L,\ \forall \xi \in \Xi \tag{9}$$

The objective function (1) minimizes the sum of the first-stage costs and the expected second-stage costs. The first-stage costs are composed of the fixed cost of opening ambulance stations and the per-unit capacity cost. The second-stage costs are composed of the expected travelling cost between the demand points and the ambulance stations and, the penalty cost incurred by unsatisfied demand. Constraints (2) respect the maximum number of ambulances that can be allocated to the EMS system. This number can translate to the maximum budget that can be allocated to ambulance acquisition or the number of available crews. Constraints (3) and (4) express the domain of the first-stage decision variables. Constraints (6) indicate that calls occurring at a demand point can only be served by an open station. Constraints (7) determine the number of unsatisfied demands. Note that the combination of constraints (7) with the objective function ensures that X^{ξ}_{ijkl} takes the value 0 if a call of type l occurring at demand point i cannot be served by an ambulance of type k located at station j, while respecting the considered target response time (i.e. $N_{ijkl} = 0$). Moreover, this combination forces X^{ξ}_{ijkl} to be equal to the number of calls of type l occurring at demand point i that are served by an ambulance of type k housed at station j. Additionally, the term in the objective function pertaining to the penalty cost forces the decision variable U^{ξ}_{il} to be equal to the number of calls of type l occurring at demand point i that are not served under scenario ξ. Constraints (8) guarantee that the total number of ambulances allocated to demand points under a given scenario is less or equal to the total number of ambulances available in the EMS system. Constraints (9) represent the integrality constraints of the second-stage decision variables.

The majority of the papers that formulate the ambulance location-allocation problem use stochastic programming recourse to chance constraints in order to maintain the risk of not satisfying the demand under a prescribed threshold. Nevertheless, except in some special cases, chance-constrained models are computationally intractable (Luedtke and Ahmed 2008). Even moderate-sized problems are very difficult to solve using off-the-shelf solvers (Ahmed and Shapiro 2008). As an alternative, in the current model, the risk of not satisfying the demand is restricted by enforcing a penalty cost on the unmet demand. The use of a cost-based modelling approach allows for solving the ambulance location-allocation problem at hand within a reasonable computing time while maintaining a high service level.

One can also see from the literature review presented above that there is diversity in terms of protocols used for the deployment of ambulances in a two-tiered EMS system. These protocols stem from the EMS system considered in each of these studies. Indeed, the deployment of ambulances in a two-tiered EMS system depends on the level of ambulance equipment as well as the skills of the crew assigned to each type of ambulances. In this paper, the ambulance deployment protocol is inspired from the Tunisian EMS system. However, the proposed model can be slightly modified and/or extended to adapt to other ambulance deployment protocols. For example, let us consider that for a life-threatening call ($l = 2$), both an ALS ($k = 2$) and a BLS ($k = 1$) ambulances are required. This protocol is particularly used when the ALS is a non-transport quick response vehicle (McLay 2009). A non-life-threatening call ($l = 1$) cannot be served by an ALS. It is covered if a BLS arrives to the patient within a time inferior or equal to the response time standard (R_1). For life-threatening calls an ALS or a BLS should arrive to the patient within the specified response time standard (R_2), with a preference for the ALS vehicle arriving first at the incident scene. In all cases, the BLS and the ALS must arrive to the scene within or under a given time threshold R_{21} and R_{22}, respectively. Without loss of generality, let us define the set $R = \{R_2, R_1, R_{22}, R_{21}\}$ of time thresholds arranged in ascending order given that $R_2 < R_1 < R_{22} \leq R_{21}$. Let us denote by $R(r)$ the rth element ($r = 1\ldots4$) of the set R.

Let us replace the parameter N_{ijkl} by N_{ijkr} that takes the value 1 if demand point i can be covered by an ambulance of type k located at station j within time threshold r; 0 otherwise. In order to accommodate the above-mentioned protocol, constraints (7) will be replaced by constraints (7.1), (7.2), (7.3) and (7.4):

$$d_{il}^{\xi} \leq U_{il}^{\xi} + \sum_{j \in J} \sum_{k \in K} N_{ijkr} X_{ijkl}^{\xi} \quad \forall i \in I,\ \forall l \in L,\ r/R(r) = R_l,\ \xi \in \Xi \tag{7.1}$$

$$X_{ij21}^{\xi} = 0 \quad \forall i \in I,\ \forall j \in J,\ \forall \xi \in \Xi \tag{7.2}$$

$$\sum_{j \in J} N_{ij11} X_{ij12}^{\xi} \leq \sum_{j \in J} N_{ij23} X_{ij22}^{\xi} \quad \forall i \in I,\ \forall \xi \in \Xi \tag{7.3}$$

$$\sum_{j \in J} N_{ij21} X_{ij22}^{\xi} \leq \sum_{j \in J} N_{ij14} X_{ij12}^{\xi} \quad \forall i \in I,\ \forall \xi \in \Xi \tag{7.4}$$

Similarly to constraints (7), constraints (7.1) determine the unsatisfied demand. The combination of constraints (7.1), (7.2) and the objective function ensures that an adequate response to a non-life-threatening call if a BLS arrives to the patient within response time threshold R_1. The combination of constraints (7.1), (7.3), (7.4) and the objective function ensures that a life-threatening call ($l = 2$) is served if: (1) a BLS ($k = 1$) arrives within R_2 ($r = 1$) and an ALS ($k = 2$) arrives within R_{22}

($r = 3$) or (2) an ALS ($k = 2$) arrives within R_2 ($r = 1$) and a BLS ($k = 1$) arrives within R_{21} ($r = 4$). More specifically, constraints (7.3) and the objective function ensure that all life-threatening calls coming from a demand point and served by a BLS within R_2 are also served by an ALS within R_{22}. In the same way, constraints (7.4) and the objective function ensure that all the life-threatening calls coming from a demand point and served by an ALS within R_2 are also served by a BLS within R_{21}.

The proposed changes aim to demonstrate how the proposed model can accommodate other possible ambulance deployment protocols that can be used in two-tiered EMS systems. Nevertheless, the rest of the paper will be devoted to investigating the design of a two-tiered EMS system in accordance with the ambulance deployment paradigm adopted in Tunisia.

4 Solution approach

A typical problem instance in a real-life case entails a large number of scenarios. Thus, solving the proposed (SPEMS) would be computationally intractable. To overcome this challenge, we propose to use the SAA, which allows for finding a good solution while considering a modest number of scenarios (Ruszczynski and Shapiro 2004). Random samples with S ($S < |\Xi|$) scenarios (or realizations) of the uncertain parameters are generated using Monte Carlo simulation technique and integrated in the model. The expected value of the recourse costs (travelling and penalty costs) is approximated by the average of these scenarios. The following mathematical model describes the SAA problem of the (SPEMS) with a sample size S. The index s will be used hereafter to denote a scenario included in a sample of size S ($s = 1\ldots S$).

As it can be noticed, the integrality constraints on second-stage variables U^s_{il} and X^s_{ijkl} are relaxed. Indeed, given the following conditions: (1) the matrix defining the feasible region of U^s_{il} and X^s_{ijkl} is totally unimodular, (2) Y_j, M and d^s_{il} are integer numbers; the linear relation of U^s_{il} and X^s_{ijkl} will have integer solutions (Wolsey 1998).

$$\begin{aligned} Min \sum_{j\in J} f_j Y_j + \sum_{j\in J}\sum_{k\in K} Z_{jk} V_k \\ + \frac{1}{S}\left[\sum_{s\in S}\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\sum_{l\in L} W_{ijk}\Delta_{kl}X^s_{ijkl} + \sum_{s\in S}\sum_{i\in I}\sum_{l\in L} E_l U^s_{il}\right] \end{aligned} \tag{10}$$

Subject to:

$$\sum_{j\in J} Z_{jk} \le P_k \quad \forall k \in K \tag{11}$$

$$Y_j \in \{0,1\} \quad \forall j \in J \tag{12}$$

$$Z_{jk}\ \text{integer} \quad \forall j \in J,\ \forall k \in K \tag{13}$$

$$\sum_{i\in I}\sum_{k\in K}\sum_{l\in L} X^{s}_{ijkl} \le MY_j \quad \forall j \in J,\ \forall s \in S \tag{14}$$

$$d^{s}_{il} \le U^{s}_{il} + \sum_{j\in J}\sum_{k\in K} N_{ijkl}X^{s}_{ijkl} \quad \forall i \in I,\ \forall l \in L,\ \forall s \in S \tag{15}$$

$$Z_{jk} \ge \sum_{i\in I}\sum_{l\in L} X^{s}_{ijkl} \quad \forall j \in J,\ \forall k \in K,\ \forall s \in S \tag{16}$$

$$U^{s}_{il}, X^{s}_{ijkl} \ge 0 \quad \forall i \in I,\ \forall j \in J,\ \forall k \in K,\ \forall l \in L,\ \forall s \in S \tag{17}$$

The optimal solutions of the SAA problem converge with probability one to an optimal solution of the original problem as the sample size S increases (Kleywegt et al. 2001). Nevertheless, solving the SAA problem with a large sample size S would incur an excessive computational burden. Thus, choosing the sample size requires a trade-off between the quality of an optimal solution of the SAA problem and the computational time needed to obtain it. To overcome this shortcoming, Kleywegt et al. (2001) proposed to solve the SAA problem repeatedly by generating M independent samples, each of a reasonable size S. Indeed, solving the considered M SAA problems, each with a sample of size S, and retaining the best solution among the M obtained ones can be more efficient than increasing the sample size, S. Henceforth, an optimality gap of an obtained SAA solution can be estimated and used to select the best solution. The details of the general SAA algorithm can be found in Kleywegt et al. (2001). The algorithm is implemented in this paper with some modifications to fit the features of (SPEMS).

The procedure for the implemented SAA algorithm is described below:

Step 1 For $m = 1\ldots M$

Step 1.1 Generate a sample of size S

Step 1.2 Solve the associated SAA problem (i.e. model 10–17) and record its optimal objective value $\hat{\vartheta}^{m}_{S}$ and the optimal first-stage solution $\hat{\upsilon}^{m}_{S}$

Step 1.3 Generate a sample of size S'. Typically S' is chosen to be quite larger than S ($S' > S$) and independent of the samples used in the SAA problems

Step 1.4 Estimate the true objective value $\hat{g}_{S'}\left(\hat{\upsilon}^{m}_{S}\right)$ of the SAA optimal first-stage solution $\hat{\upsilon}^{m}_{S}$ and its variance $\sigma^{2}_{\hat{g}_{S'}\left(\hat{\upsilon}^{m}_{S}\right)}$ by (18)

$$\sigma^{2}_{\hat{g}_{S'}\left(\hat{\upsilon}^{m}_{S}\right)} = \frac{1}{(S'-1)S'}\sum_{s=1}^{S'}\left[\hat{g}_{S}\left(\hat{\upsilon}^{m}_{S}\right) - \hat{g}_{S'}\left(\hat{\upsilon}^{m}_{S}\right)\right]^{2} \tag{18}$$

Step 2 Compute $\bar{\vartheta}_S^M$ and its variance $\sigma^2_{\bar{\vartheta}_S^M}$ over the M replications by (19) and (20)

$$\bar{\vartheta}_S^M = \frac{1}{M}\sum_{m=1}^{M}\hat{\vartheta}_S^m \tag{19}$$

$$\sigma^2_{\bar{\vartheta}_S^M} = \frac{1}{(M-1)M}\sum_{m=1}^{M}\left[\hat{\vartheta}_S^m - \bar{\vartheta}_S^M\right]^2 \tag{20}$$

Step 3 For each solution $\hat{\upsilon}_S^m$, $m = 1\ldots M$, compute the optimality gap $\hat{g}_{S'}\left(\hat{\upsilon}_S^m\right) - \bar{\vartheta}_S^M$ and a corresponding estimate of variance $\sigma^2_{gap} = \sigma^2_{\hat{g}_{S'}\left(\hat{\upsilon}_S^m\right)} + \sigma^2_{\bar{\vartheta}_S^M}$. We should point out that $\bar{\vartheta}_S^M$ and $\hat{g}_{S'}\left(\hat{\upsilon}_S^m\right)$, respectively, provide a statistical lower and upper bounds on the optimal objective function of the original problem (Norkin et al. 1998; Mak et al. 1999).

The confidence interval for the optimality gap at a given solution is calculated as:

$$\left(\hat{g}_{S'}\left(\hat{\upsilon}_S^m\right) - \bar{\vartheta}_S^M\right) + Z_\alpha \sigma_{gap} \tag{21}$$

With $Z_\alpha = \Phi^{-1}(1-\alpha)$, where Φ is the cumulative distribution function of the standard normal distribution.

After completing step 3, we have to inspect the values of the optimality gap and its variance. If these values are too large, one must repeat the procedure with larger values of S and M. Readers can refer to (Kleywegt et al. 2001) for a more detailed description of the SAA algorithm.

5 Numerical experiments

The numerical experiments were carried out on a dual Intel Xeon X5650 processor 2.66 GHz and 72 GB DDR3 ECC Reg Memory RAM. The models were implemented in MS Visual C++ and linked with ILOG CPLEX 12.3 optimization library.

The test problems considered in the experimentation come from a real-life case study arising in the Northern Region of Tunisia (referred to as SAMU 01). In Tunisia, the calls received by the EMS system are classified as follows: (1) calls that do not require an ambulance (code 1); (2) calls that require transportation service without any emergency (code 2); (3) calls associated with a non-life-threatening incident (code 3) and (4) calls involving danger to human life (code 4). Markedly, the degree of priority associated with each of these calls is different. The emergency calls requesting an ambulance and a medical team are those of codes 3 and 4. Obviously, the highest priority is given to code 4 calls. Calls of codes 1 and 2 do not need an ambulance from the EMS system and are oriented to the appropriate services. To satisfy calls of codes 3 and 4, a team composed of one or two nurses and an emergency physician, and an adequately equipped ambulance are assigned to serve the patient. Two types of ambulances could be used: (1) BLS ambulances and (2) ALS ambulances. Only the crew of the ALS ambulance type includes an emergency physician. This type of ambulance is intended to serve patients of code 4 as it can provide cardiac and medical monitoring and ensure the patient's treatment during transport to the hospital. ALS ambulances can also cover demand of code 3. However, the BLS ambulance type can only cover the demand of code 3. The response time threshold is set to 20 min for emergency calls of code 3 and 15 min for emergency calls of code 4.

5.1 Data description

In this section, we will present the necessary details for our model: data related to potential sites for ambulance stations, demand zones and vehicles. The Northern Region of Tunisia is divided into seven governorates, where each of them is further divided into a number of delegations. Henceforth, the demand points represent the different delegations, while the potential sites for ambulance stations, are proposed by the SAMU 01 manager, and correspond to the region's hospitals. Therefore, 31 potential ambulance stations' sites and 78 demand points are considered. SAMU 01 historical data is used to determine the average hourly demand of codes 3 and 4 associated with each demand point. It is worth noting here that, in the Northern Region of Tunisia, the demand of code 4 contributes up to 90% of the total demand. The number of incoming calls of codes 3 and 4 for each demand point are independently drawn from a Poisson distribution. It is important to note here that the hourly demand's variations over the course of the day and from one season to another can be directly accounted for in the model by generating demand realizations for different hourly time periods of the year.

The maximum number of ALS and BLS units that can be acquired by the EMS system is set to 80. The per-unit capacity cost is set to 7.412 (resp. 6.156) Tunisian Dinar (TND) per hour for ALS (resp. BLS). Note that, in the considered base case, the ALS cost is 20% higher than the BLS cost. The hourly capacity of each ambulance type is set to one. In fact, as it has been mentioned earlier, an ambulance can only cover one demand within 1 h.

The travelling time between each demand point and potential ambulance site is determined based on the distance between the two locations and the average speed of the ambulance. The ALS ambulance is faster than BLS (the speed of BLS is assumed to be approximately equal to 0.85 * the speed of ALS). The transportation cost depends on the distance and is equal to 1.2 TND/km for ALS and 0.8 TND/km for BLS (this cost is based on fuel and maintenance costs of each ambulance type).

As far as penalty costs are concerned, they are given relatively large values (in comparison to other costs) in order to prevent the violation of demand coverage constraints as much as possible. Thus, the penalty for an unsatisfied code 3 call (resp. code 4) is set to 300 TND (resp. 500 TND). Moreover, as it can be noticed, penalty cost values are used to enforce a higher priority for demand of code 4 over that of code 3.

5.2 Design of the EMS system

In this section we describe the obtained results. We choose to solve it using 10 replications ($M = 10$). For the SAA problems, we use sample sizes of $S = 5$, 10, 20, 50, 100 and 200 scenarios. We estimate the "true objective value" of each SAA optimal solution by simulation with $S' = 1000$.

Table 3 reports the average computing time (Avg. CPU) in seconds, the produced Lower Bound (LB) and Upper Bound (UB) in Tunisian Dinar (TND) and their respective standard deviation (Std.). As expected, it can be seen that, when the sample size S increases, the LB increases whereas the UB decreases, leading to a tighter optimality gap. The CPU increases with higher sample sizes. However, it remains reasonable for a sample size $S = 200$. In addition, Table 3 shows reasonable standard deviation of the bounds. Therefore, we can conclude that 10 replications are enough to obtain a reasonable confidence interval of the bounds. If the standard deviations are too large, then the value of M must be augmented (Kleywegt et al. 2001).

Table 4, presents the estimator for the optimality gap, as well as its 95% confidence interval for the different sample sizes. The estimator for the optimality gap is the difference between UB and LB. These results show the convergence of the SAA solutions. As it can be seen from Fig. 1, the optimality gap diminishes as the sample size increases, meaning that better solutions are obtained for larger sample sizes. For the problem at hand, a near-optimal solution of 0.5% optimality gap is obtained with sample size $S = 200$. The results for the optimality gap indicate that the solutions produced by the SAA algorithm scheme are good enough to be used in a practical application.

Figure 2 depicts the total cost (TC), the first-stage cost and the expected second-stage cost as a function of the sample size, showing the trade-off between the cost's components. Indeed, when the sample size increases, it can be seen that first-stage cost increases whereas the expected second-stage cost decreases. This behavior is explained by the nature of the simulation scheme: when larger samples are used, more robust solutions are produced but they require the use of a larger number of facilities and vehicles, thus increasing the first-stage cost. Consequently, the penalty cost is reduced by the higher reliability of the EMS system. This result points out the value of a well-dimensioned infrastructure to achieve better robustness.

Table 3 Statistical lower and upper bounds for $M = 10$ and $S' = 1000$

Sample size S	Avg. CPU (s)	Lower bound		Upper bound	
		LB	Std. (%)	UB	Std. (%)
5	1	497	3.00	1417	2.00
10	2	538	1.80	822	2.00
20	7	568	1.50	769	2.00
50	19	627	1.50	684	1.00
100	38	650	1.00	673	0.96
200	76	660	0.40	663	0.85

Table 4 Estimated optimality gap and variance

Sample size S	Estimated optimality gap			95% confidence interval			
	Gap	%	Std. (Gap)	Min	%	Max	%
5	920	>100.0	37	847	>100.0	992	>100.0
10	284	53.0	23	239	44.0	329	61.0
20	201	35.0	21	160	28.0	242	43.0
50	57	9.0	18	22	3.0	92	15.0
100	23	4.0	13	−2	−0.3	48	7.0
200	3	0.5	9	−15	−2.0	21	3.0

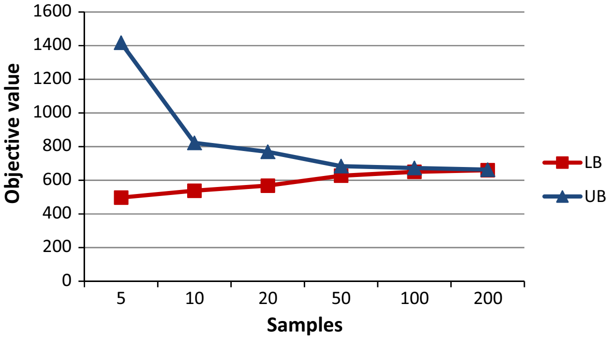

Fig. 1 Convergence of the objective value

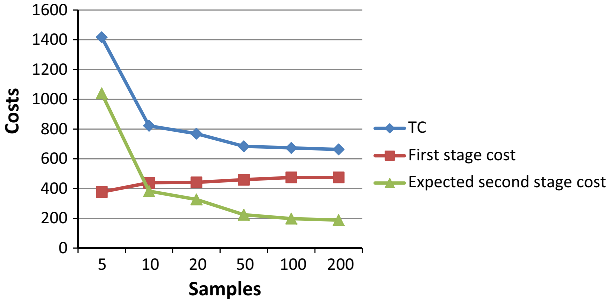

Fig. 2 EMS system costs associated with the different SAA solutions

In order to evaluate the added value of dealing with stochastic representation of the demand instead of its average value, other simulation experiments were performed. The deterministic version was obtained by replacing the random parameters by their means and then solving the resulting problem. Table 5 reports the costs produced by the deterministic version (EVP) and the stochastic one (SAA), confirming that the SAA solution outperforms the one obtained for the EVP in terms

Table 5 EMS system costs associated with the SAA and EVP solutions and its performance

	Penalty cost (TND)	Opening cost (TND)	Capacity cost (TND)	Travelling cost (TND)	Demand covered (%)
EVP	1267	148	146	122	81.0
SAA	43	157	317	146	99.3

of reliability. As it can be noticed, the penalty cost of EVP solution is much larger than the one incurred by the SAA solution. Subsequently, the percentage of covered demand is higher when a stochastic approach is adopted for the design of the EMS system. The latter favours system reliability at the expense of higher first-stage costs.

5.3 Sensitivity analysis

A sensitivity analysis is conducted in order to investigate the behaviour of the proposed two-tiered EMS system. A sensitivity analysis is a procedure that allows for determining how sensitive the optimal solution is to changes in data values that could stem from inappropriate predictions. Some input data are based on estimates from data analysis and expert opinions, thus, the real values are probably higher or lower than these estimates. In the sensitivity analysis, the following parameters are particularly considered: the penalty cost, the demand and the ambulance service trip time. Moreover, some experiments are conducted in order to investigate the effect of the response time threshold on the configuration of the EMS system. Additional tests are performed in order to see how the EMS configuration is impacted by the demand distribution among code 3 and code 4, and ambulances per-unit capacity costs. Table 6 reports the characteristics of the different tests carried out with an attempt to investigate the variation of the aforementioned parameters. For each test, the following information is provided: the statistical lower bound (LB) in (TND), the upper bound (UB) in (TND), the average upper bound (AUG) in (TND), the average first-stage cost (AFS) in (TND), the average second-stage cost for solutions without penalty costs (ACS) in (TND), the average penalty cost (AP) in (TND), the demand covered (DC) in percentages, the number of opened stations (NS), the number of ambulances of type ALS (NALS) and BLS (NBLS) and the location of ambulance stations proposed by the obtained solutions of the tests (N°C), as reported in Table 7. All SAA problems are solved with $M = 10$, $S = 200$ and $S' = 1000$.

The results obtained are summarized in Table 8. The effects of varying the considered parameters on the configuration of the EMS system and its performance are illustrated in Figs. 3, 5 and 6.

Figure 3 delineates the effect of penalty costs on the EMS system's configuration and performance. The penalty cost is multiplied by a factor lying between 0.125 and 2. The results obtained for tests 2–6 are summarized in Table 8.

Table 6 Characteristics of the tests

Test	Characteristics
1	Base case
2	Base case penalty cost * 0.125
3	Base case penalty cost * 0.25
4	Base case penalty cost * 0.5
5	Base case penalty cost * 1.5
6	Base case penalty cost * 2
7	Base case demand * 1.4
8	Base case demand * 1.8
9	Base case demand * 0.8
10	Base case demand * 0.6
11	Response time threshold in minutes (15–10)
12	ALS cost = 1.4 * BLS cost
13	ALS cost = 1.8 * BLS cost
14	Demand of code 3 (50%)-Demand of code 4 (50%); ALS cost = 1.2 * BLS cost
15	Demand of code 3 (50%)-Demand of code 4 (50%); ALS cost = 1.4 * BLS cost
16	Demand of code 3 (50%)-Demand of code 4 (50%); ALS cost = 1.8 * BLS cost
17	Demand of code 3 (90%)-Demand of code 4 (10%); ALS cost = 1.2 * BLS cost
18	Demand of code 3 (90%)-Demand of code 4 (10%); ALS cost = 1.4 * BLS cost
19	Demand of code 3 (90%)-Demand of code 4 (10%); ALS cost = 1.8 * BLS cost
20	Ambulance service trip time 45 min
21	Ambulance service trip time 75 min

Table 7 Location of ambulance stations proposed by the solutions of the tests

N°C	Open stations
1	2, 3, 4, 5, 6, 15, 16, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
2	2, 3, 4, 5, 6, 16, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
3	2, 5, 6, 16, 22, 23, 25, 27, 28, 30, 31
4	2, 3, 4, 5, 6, 14, 15, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
5	2, 3, 5, 6, 16, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
6	2,3,5,16, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
7	2, 3, 4, 5, 6, 15, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
8	2,3,5,16, 22, 23, 24, 25,27, 28, 29, 30, 31

As it can be noticed, when the penalty cost increases, on one hand the EMS system cost increases and, on the other hand, the percentage of covered demand increases. Thus, an increase in the penalty cost favours the design of a more reliable

Table 8 Results obtained for the different tests with $S = 200$, $M = 10$ and $S' = 1000$

Test	LB	UB	AUG	AFS	ACS	AP	DC (%)	NS	NALS	NBLS	N°C
1	660	663	678	475	147	43	99.3	17	42	1	1
2	522	523	557	250	121	152	82.0	11	20	0	3
3	589	591	605	385	151	58	97.0	16	33	0	2
4	629	630	637	427	152	58	99.0	16	39	0	2
5	677	687	711	486	152	73	99.4	17	43	1	1
6	679	690	837	494	152	191	99.7	17	46	1	1
7	795	808	831	544	212	75	99.3	17	50	1	1
8	916	934	1021	606	274	141	99.5	17	60	2	1
9	578	592	603	412	135	45	99.1	15	37	0	5
10	527	578	535	352	98	75	98.2	15	28	1	5
11	720	731	739	508	150	73	98.7	18	46	0	4
12	705	717	788	499	145	72	98.9	17	39	1	1
13	809	811	830	584	150	70	99.0	17	38	1	1
14	601	611	625	401	168	41	99.3	14	32	6	6
15	637	654	668	455	150	48	99.2	15	31	8	5
16	712	726	797	519	143	64	98.9	16	29	8	7
17	523	533	543	355	126	52	98.9	14	16	19	8
18	550	554	603	378	127	49	99.0	13	15	21	8
19	578	584	680	392	129	62	99.0	13	13	21	8
20	459	464	529	316	120	27	99.5	16	37	0	2
21	877	894	985	667	184	43	99.5	17	50	1	1

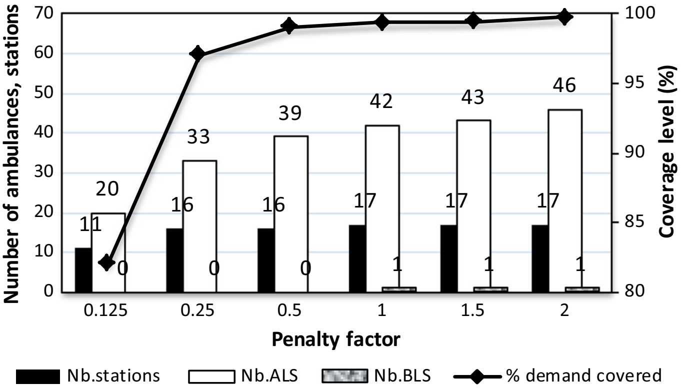

Fig. 3 The effect of varying penalty costs on the configuration of the EMS system and its performance

EMS system. However, since the coverage is already very good, managers should wonder if this improvement is worth the additional cost. Furthermore, the results show that a penalty cost increase has a little impact on the number of BLS and opened stations. Conversely, the number of ALS required for the EMS system clearly increases when the penalty cost increases. This explains the EMS system cost's increase and demand coverage's improvement. Additionally, in Fig. 4, one can see that an increase in the penalty cost incurs an increase in the average optimality gap. Henceforth, with an increase in the penalty cost, the quality of the SAA solution can be improved by considering in the SAA algorithm a larger sample size (S) and/or larger number of replications (M).

Emergency demand values are derived from the considered real-life case study. These values might increase in the future due to an increase in the number of inhabitants or/and due to an aging population. Moreover, the demand might decrease because of a decrease in the number of inhabitants or the implementation of more effective and adequately coordinated preventive care. Henceforth, we conducted a sensitivity analysis to test the effect of an increase/decrease of the demand on the optimal configuration. The results for test 7, 8, 9 and 10 are reported in Table 8. They show the optimal configuration of the EMS system when the demand is increased by 40 and 80% and then decreased by 20 and 40%, respectively. It should be noted that an increase in the demand value does not affect the decisions pertaining to the opened ambulance stations. However, it impacts the number of ambulances required by the EMS system. Conversely, with a decrease in the demand, the number of opened ambulance stations is reduced. However, an increase or a decrease in the demand clearly impacts the number of ambulances required by the EMS system. As it can be seen in Fig. 5, the number of ambulances required increases along with the demand. Moreover, as it can be observed in Fig. 5, an increase of the demand entails an increase in the percentage of covered demand. This means that the increase in unsatisfied demand is lesser than the increase in the demand. We should point out that, by an increase of 80% on the current demand, the EMS system should be equipped with 60 ALS and 2 BLS. The proposed configuration allows for a coverage of 99.5% of the demand within the required response time threshold.

Test 11 was performed to assess the effect of the response time threshold on the EMS configuration, while shortening the required response time from 20 to 15 min (i.e. 20 and 15 min are the response time threshold to cover demand of codes 3 and 4, respectively) to 15–10 min. The results related to test 11 in Table 8 show that a decrease in the response time threshold incurs an increase in the EMS system's cost. With the considered decrease in response time threshold, the number of opened

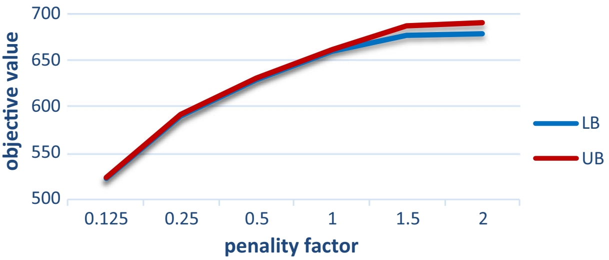

Fig. 4 The effect of varying penalty costs on LB and UB

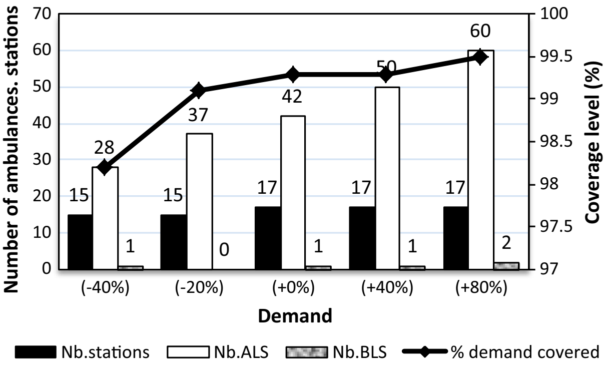

Fig. 5 The effect of varying demand on the configuration of the EMS system and its performance

ambulance stations is increased from 17 to 18 and the EMS system should acquire a larger number of ambulances. This partially explains the EMS system cost's increase. Indeed, there is an increase in the penalty cost even though more resources are allocated to the EMS system. This highlights that a lower response time threshold makes it more challenging to improve the reliability of the EMS system (Fig. 6).

The results obtained for the base case point out that the number of BLS type ambulances is much lesser than the number of ALS type ambulances, even when the demand is increased by 80%. This can be explained by the flexibility of the ALS that can respond to code 3 and 4 calls, but more importantly by the fact that 90% of the calls are of code 4. In this case, it would be preferable to opt for an all-ALS EMS system that, contrarily to its two-tiered counterpart, does not require the implementation of a sophisticated triage system that can effectively identify the type of ambulance to dispatch to an incoming call (Stout et al. 2000). These results spark the interest for investigating how the EMS configuration changes based on (1) the distribution of the demand and (2) the per-unit ALS cost. In the base case, the per-unit cost of ALS is 20% higher than that of a BLS. However, the per-unit ALS capacity cost becomes higher if the ALS is equipped with more cutting-edge equipment and devices. Tests 12–19 are performed in order to investigate the effects

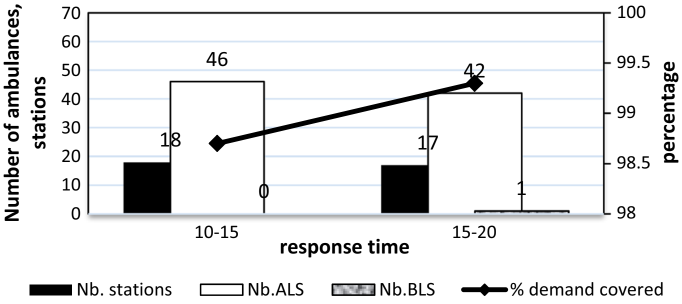

Fig. 6 The effect of varying response time on the configuration of the EMS system and its performance

of demand distribution and the level of ALS equipment/technology (that translates into per-unit ALS cost) on the EMS configuration.

As it can be observed from Fig. 7, the number of BLS ambulances required by the EMS system augments when the percentage of the code 3 demand increases. Moreover, the number of ambulance stations decreases when the percentage of the code 3 demand increases. This can be explained by the fact that the response time

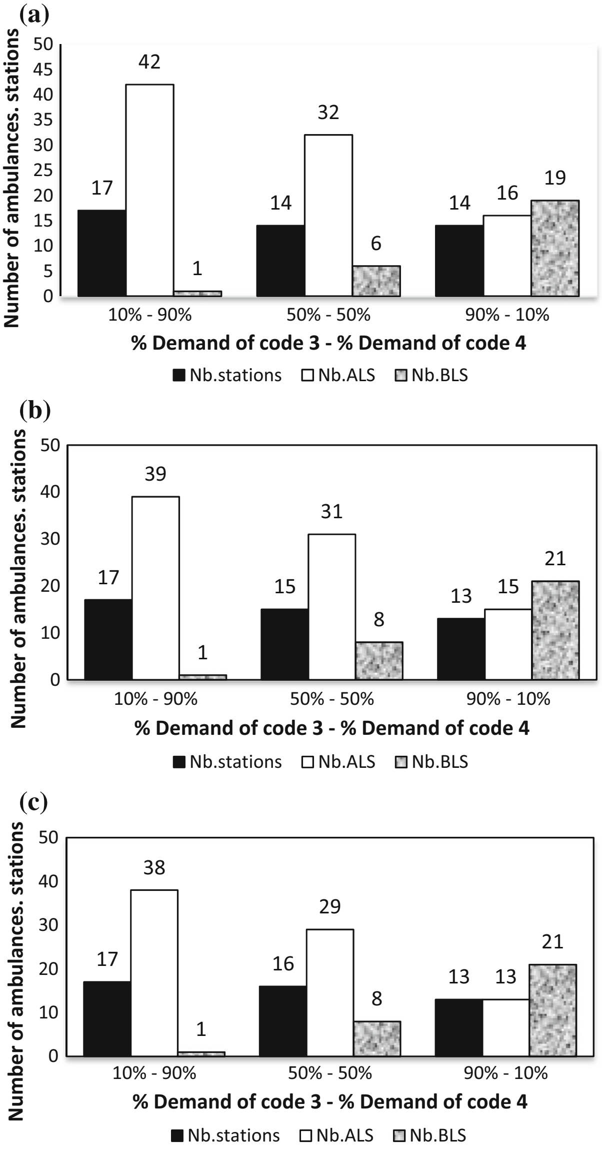

Fig. 7 The effect of demand distribution and ALS cost on the configuration of the EMS system. **a** Cost ALS = 1.2 * Cost BLS, **b** Cost ALS = 1.4 * Cost BLS, **c** Cost ALS = 1.8 * Cost BLS

threshold for a code 3 call is higher than the one associated with a code 4 call. The increase of the per-unit ALS capacity cost is generally followed by a slight decrease in the number of ALS units counter a slight increase in the number of BLS units. Hence, the EMS configuration is barely affected by the ALS per-unit capacity cost. In particular, these results stress the interest of considering a two-tiered EMS system when the code 3 demand is relatively important (such as, the cases with a code 3 demand that represents 50 and 90% of the total demand).

According to the ambulance operator, 1 h is a reasonable ambulance service trip time. Nevertheless, it remains interesting to know how the ambulance service trip time affects the EMS configuration. Tests 20 and 21 have been conducted in this perspective while assuming ambulance service trip times of 45 and 75 min, respectively. The data used in these tests has been adjusted as follows: (1) the demand is generated while considering the average number of incoming calls over 45 min (resp. 75 min), and (2) the opening and the per-unit capacity costs are those associated with a horizon of 45 min (resp. 75 min). Remarkably, one can see from Fig. 8 that the number of ambulances increases with the service trip time. Additionally, the number of opened ambulance stations is lesser when the service trip time is 45 min.

5.4 Evaluation of the EMS configuration

The proposed stochastic model and solution approach are aimed at designing a reliable two-tiered EMS system. The model does not tackle the dynamics of EMS systems operations, as this would require the addition of the time dimension through the division of the horizon of 1 h into time segments of 1–3 min (Naoum-Sawaya and Elhedhli 2013) in order to account for the arrival time of emergency calls. Markedly, this integration would result in a computationally intractable model that cannot be used in practice. In order to account for system dynamics, simulation experiments are carried out to evaluate the performance of the proposed EMS configuration and confirm the quality of the obtained SAA solution. The simulation is also used to evaluate the performance of the deterministic solution (EVP solution) and compare it to the one generated by the SAA algorithm. This comparison provides another evaluation on the benefit of a stochastic model over its computational efforts.

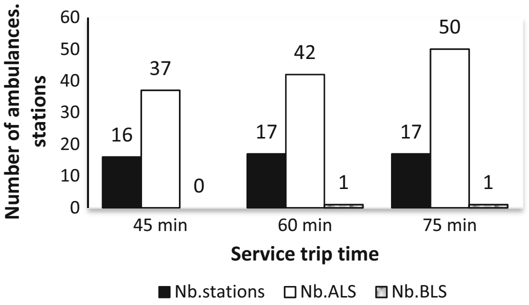

Fig. 8 The effect of ambulance service trip time on the configuration of the EMS system

In particular, the simulation allows the evaluation of the service level of the EMS system proposed by these two solutions, which is a performance measure commonly used by EMS managers. The service level certainly depends on the EMS system's configuration, but also on the time and location of the incoming calls, the ambulance service trip time and ambulance dispatching decisions. The latter pertains to the determination of the ambulance to assign to a received emergency call. Typically, the objective is to minimize the response times for all emergency calls received in the course of a given day (Schmid 2012; Aboueljinane et al. 2012). Various dispatching rules are proposed in the EMS simulation literature. However, most works use the "closest available ambulance dispatch" rule (Kergosien et al. 2015; van Buuren et al. 2012; Aboueljinane et al. 2012; Maxwell et al. 2009; Ingolfsson et al. 2003; Fitzsimmons 1971) and its variants, namely the "closest available vehicle with preemption" rule (Savas 1969; Lubicz and Mielczarek 1987), the "closest base" rule (Iskander 1989), the "lower response vehicle" rule (Silva and Pinto 2010), the "nearest available vehicle conditioned by call priorities" rule (Aringhieri et al. 2007) and the "regionalized response" rule (Swoveland et al. 1973; Su and Shih 2003). In addition to these simple rules, some papers devise more sophisticated algorithms to address the dynamic ambulance relocation and/or dispatching problem (Andersson and Värbrand 2007; Schmid 2012; Jagtenberg et al. 2015). For a more detailed survey on ambulance dispatching rules, we refer the reader to the literature reviews presented in Aboueljinane et al. (2012) and Bélanger et al. (2015).

Even though the proposed stochastic model focuses on the strategic-tactical ambulance planning problem, it indirectly tackles ambulance dispatching by favouring the closest ambulance dispatch rule through the consideration of transportation costs. In addition to the closest available ambulance dispatch rule, we consider that the emergency calls are served on a first-come first-served basis without preemption. In the considered two-tiered EMS system, a higher priority is given to code 4 calls over those of code 3. Hence, code 4 calls are the first to be assigned to the available ALS ambulances. Moreover, when a code 4 call occurs, the closest ALS ambulance is dispatched to the emergency scene. When a code 3 call occurs, the closest BLS ambulance is dispatched unless all BLS ambulances are busy; in this case the closest ALS is dispatched.

The simulation is coded in MS Visual C++. Arbitrarily, the run length is set to 24 consecutive hours. This allows to determine an estimate of the average daily service level. Thirty replications are performed for statistical evaluations.

From the obtained simulation results reported in Table 9 and those presented in Fig. 9, one can see that the EMS configuration proposed by the SAA solution

Table 9 Simulation results

	Service level (%)			
	Min	Avg.	Max	Std.
EVP	54	58	62	2.11
SAA	87	92	100	4.13

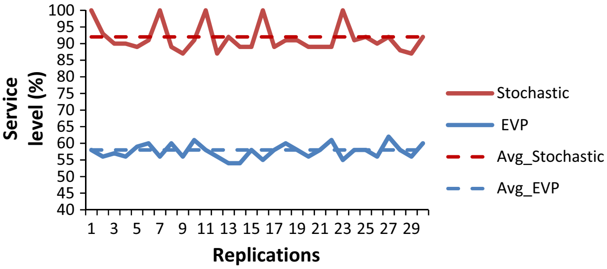

Fig. 9 EMS system service level

outperforms the one given by the EVP solution in terms of service level. More importantly, the results show that an average service level of 92% can be achieved for the SAA EMS system. At this point, it is worth noting that the targeted average service level for the EMS system under study is 90%, and the obtained EMS configuration is deemed satisfactory by the EMS system planners. Furthermore, it might be useful to recall here that the simulation is used to provide a service level estimate based on simple ambulance dispatch rules and that the dynamic relocation is not considered. Henceforth, this service level can be improved if more sophisticated algorithms are developed for dynamic ambulance dispatching and relocation.

6 Conclusion

In this paper, we proposed a two-stage stochastic programming model for the design of a two-tiered EMS system under demand uncertainty. The objective was to find the configuration of the EMS system that minimizes the total cost composed of the ambulance station opening cost, per-unit capacity cost, transportation cost and penalty cost associated with demand unsatisfaction. A sample average approximation algorithm was then proposed to solve the considered problem. Numerical experiments were carried out to assess the performance of the proposed stochastic approach. Experimental results based on a real-life case study demonstrated the convergence of the proposed algorithm and its usefulness in practice. Moreover, they pointed out the relevance of using a stochastic approach instead of a deterministic one to improve the robustness of the EMS system. Finally, a simulation study was conducted to evaluate the performance of the EMS system while accounting for its dynamics. The simulation showed that the SAA solution outperformed the deterministic one. The current work addressed the design of the EMS system. Future research will offer a more in-depth investigation of the dynamic ambulance dispatching and relocation problem encountered at the operational decision level. A comprehensive assessment of the EMS system will be conducted in order to gauge the significance of the design decisions and various ambulance dispatching and relocation strategies on system performance.

Acknowledgements Funding was provided by Ministry of Higher Education and Scientific Research (TN).

References

Aboueljinane L, Jemai Z, Evren S (2012) Reducing ambulance response time using simulation: the case of val-de-marne department emergency medical service. In: Proceedings of the winter simulation conference, Berlin, Germany

Ahmed S, Shapiro A (2008) Solving chance-constrained stochastic programs via sampling and integer programming. Tutorials in operations research. In: Chen Z-L, Raghavan S (eds) INFORMS, pp 261–269

Andersson T, Värbrand P (2007) Decision support tools for ambulance dispatch and relocation. J Oper Res Soc 58:195–201. doi:10.1057/palgrave.jors.2602174

Aringhieri R, Carello G, Morale D (2007) Ambulance location through optimization and simulation: the case of Milano urban area. In: Proceedings of the annual conference of the operations research society optimization and decision sciences, Milan, Italy

Ball O, Lin LF (1993) A reliability model applied to emergency service vehicle location. Oper Res 41:18–36. doi:10.1287/opre.41.1.18

Bélanger V, Ruiz A, Soriano P (2015) Recent advances in emergency medical services management. Available via CIRRELT. https://www.cirrelt.ca/DocumentsTravail/CIRRELT-2015-28. Accessed July 2015

Beraldi P, Bruni ME (2009) A probabilistic model applied to emergency service vehicle location. Eur J Oper Res 196:323–331. doi:10.1016/j.ejor.2008.02.027

Beraldi P, Bruni ME, Conforti D (2004) Designing robust emergency medical service via stochastic programming. Eur J Oper 158:183–193. doi:10.1016/S0377-2217(03)00351-5

Blanchard IE, Doig CJ, Hagel BE, Anton AR, Zygun DA, Kortbeek JB, Powell DG, Williamson TS, Fick GH, Innes GD (2012) Emergency medical services response time and mortality in an urban setting. Prehosp Emerg Care 16:142–511. doi:10.3109/10903127.2011.614046

Church R, ReVelle C (1974) The maximal covering location problem. Pap Reg Sci Assoc 32:101–118. doi:10.1111/j.1435-5597.1974.tb00902.x

Current J, Daskin M, Schilling D (2001) Discrete network location models. In: Drezner Z, Hamacher HW (eds) facility location: applications and theory. Springer, Berlin, pp 83–120

Daskin MS (1983) A maximum expected location model: formulation, properties and heuristic solution. Transp Sci 7:48–70. doi:10.1287/trsc.17.1.48

Erkut E, Ingolfsson A, Erdogan G (2008) Ambulance location for maximum survival. Nav Res Logist 55:42–58. doi:10.1002/nav.20267

Fitzsimmons JA (1971) An emergency medical system simulation model. In: Proceedings of the winter simulation conference, ACM, New York, USA, pp 18–25

Geroliminis N, Kepaptsoglou K, Karlaftis MG (2011) Ahybrid hypercube-genetic algorithm approach for deploying many emergency response mobile units in an urban network. Eur J Oper Res 210:287–300. doi:10.1016/j.ejor.2010.08.031

Gonzales RP, Cummings GR, Phelan HA, Muleker MS, Rodning CB (2009) Does increased emergency medical services prehospital time affect patient mortality in rural motor vehicle crashes? A statewide analysis. Am J Surg 197:30–34. doi:10.1016/j.amjsurg.2007.11.018

Iannoni AP, Morabito R, Saydam C (2011) Optimizing large-scale emergency medical system operations on highways using the hypercube queuing model. Socioecon Plan Sci 45:105–117. doi:10.1016/j.seps.2010.11.001

Ingolfsson A, Erkut E, Budge S (2003) Simulation of single start station for Edmonton EMS. J Oper Res Soc 54:736–746. doi:10.1057/palgrave.jors.2601574

Ingolfsson A, Budge S, Erkut E (2008) Optimal ambulance location with random delays and travel times. Health Care Manag Sci 11:262–274. doi:10.1007/s10729-007-9048-1

Iskander WH (1989) Simulation modeling for emergency medical service systems. In: Proceedings of the winter simulation conference, Washington, USA, pp 1107–1111

Jagtenberg CJ, Bhulai S, van der Mei RD (2015) An efficient heuristic for real-time ambulance redeployment. Oper Res Health Care 4:27–35. doi:10.1016/j.orhc.2015.01.001

Kergosien Y, Bélanger V, Soriano P, Gendreau M, Ruiz A (2015) A generic and flexible simulation-based analysis tool for EMS management. Int J Prod Res 53:7299–7316. doi:10.1080/00207543.2015.1037405

Kleywegt AJ, Shapiro A, Homem-de-Mello T (2001) The sample average approximation method for stochastic discrete optimization. SIAM J Optim 12:479–502. doi:10.1137/S1052623499363220

Lam SSW, Ng YS, Lakshmanan MR, Ng YY, Marcus EHO (2016) Ambulance deployment under demand uncertainty. J Adv Manag Sci 4:187–194. doi:10.12720/joams.4.3.187-194

Larson RC (1974) A hypercube queueing model for facility location and redistricting in urban emergency service. Comput Oper Res 1:67–95. doi:10.1016/0305-0548(74)90076-8

Larson RC (1975) Approximating performance of urban emergency service systems. Oper Res 23:845–868. doi:10.1287/opre.23.5.845

Lubicz M, Mielczarek B (1987) Simulation modelling of emergency medical services. Eur J Oper Res 29:178–185. doi:10.1016/0377-2217(87)90107-X

Luedtke J, Ahmed S (2008) A Sample approximation approach for optimization with probabilistic constraints. SIAM J Optim 19:674–699. doi:10.1137/070702928

Mak WK, Morton DP, Wood RK (1999) Monte Carlo bounding techniques for determining solution quality in stochastic programs. Oper Res Lett 24:47–56. doi:10.1016/S0167-6377(98)00054-6

Mandell M (1998) Covering models for two-tiered emergency medical services systems. Locat Sci 6:355–368. doi:10.1016/S0966-8349(98)00058-8

Marianov V, ReVelle CS (1992) The capacitated standard response fire protection siting problem: deterministic and probabilistic models. Ann Oper Res 40:303–322. doi:10.1007/BF02060484

Maxwell MS, Henderson SG, Topaloglu H (2009) Ambulance redeployment: an approximate dynamic programming approach. In: Proceedings of the winter simulation conference, Piscataway NJ, USA, pp 1850–1860

McLay LA (2009) A maximum expected covering location model with two types of servers. IIE Trans 41:730–741. doi:10.1080/07408170802702138

Naoum-Sawaya J, Elhedhli S (2013) A stochastic optimization model for real-time ambulance redeployment. Comput Oper Res 40:1972–1978. doi:10.1016/j.cor.2013.02.006

Nickel S, Reuter-Oppermann M, Saldanha-da-Gama F (2015) Ambulance location under stochastic demand: a sampling approach. Oper Res Health Care 8:24–32. doi:10.1016/j.orhc.2015.06.006

Norkin VI, Pflug GC, Ruszczynski A (1998) A branch and bound method for stochastic global optimization. Math Program 83:425–450. doi:10.1007/BF02680569

Noyan N (2010) Alternate risk measures for emergency medical service system design. Ann Oper Res 181:559–589. doi:10.1007/s10479-010-0787-x

O'Keeffe C, Nicholl J, Turnerl J, Goodacre S (2010) Role of ambulance response times in the survival of patients with out-of-hospital cardiac arrest. Emerg Med J 28:703–706. doi:10.1136/emj.2009.086363

ReVelle C, Hogan K (1989) The maximum reliability location problemand α-reliable P-Center problems: derivatives of the probabilistic location set covering problem. Ann Oper Res 18:155–174. doi:10.1007/BF02097801

ReVelle C, Marianov V (1991) A probabilistic FLEET model with individual reliability requirements. Eur J Oper Res 53:93–105. doi:10.1016/0377-2217(91)90095-D

Ruszczynski A, Shapiro A (2004) Stochastic programming. Handbooks in Operations research and management science. Elsevier, Amsterdam

Savas ES (1969) Simulation and cost-effectiveness analysis of New York's emergency ambulance service. Manag Sci 15:608–627. doi:10.1287/mnsc.15.12.B608

Schilling DA, Elzinga DJ, Cohon J, Church RL, ReVelle CS (1979) The TEAM/FLEET models for simultaneous facility and equipment sitting. Transp Sci 13:163–175. doi:10.1287/trsc.13.2.163

Schmid V (2012) Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. Eur J Oper Res 219:611–621. doi:10.1016/j.ejor.2011.10.043

Silva PMS, Pinto LR (2010) Emergency medical systems analysis by simulation and optimization. In: Proceedings of the winter simulation conference, Baltimore, Maryland, pp 1850–1860

Snyder LV (2006) Facility location under uncertainty: a review. IIE Trans 38:537–554. doi:10.1080/07408170500216480

Stout J, Pepe PE, Mosesso VN (2000) All-advanced life support vs tiered-response ambulance systems. Prehosp Emerg Care 4:1–6. doi:10.1016/S1090-3127(00)70065-7

Su S, Shih C-L (2003) Modeling an emergency medical services system using computer simulation. Int J Med Inform 72:57–72. doi:10.1016/j.ijmedinf.2003.08.003

Sund B, Svensson L, Rosenqvist M, Hollenberg J (2011) Favourable cost-benefit in an early defibrillation programme using dual dispatch of ambulance and fire services in out-of-hospital cardiac arrest. Eur J Health Econ 13:811–818. doi:10.1007/s10198-011-0338-7

Swoveland C, Uyeno D, Vertinsky I, Vickson R (1973) A simulation-based methodology for optimization of ambulance service policies. Socio-Econ Plan Sci 7:697–703. doi:10.1016/0038-0121(73)90033-5

van Buuren M, van der Mei R, Aardal K, Post H (2012) Evaluating dynamic dispatch strategies for emergency medical services: TIFAR simulation tool. In: Proceedings of the winter simulation conference, Berlin, Germany

Van Essen JT, Hurink JL, Nickel S, Reuter M (2014) Models for ambulance planning on the strategic and the tactical level. Beta publishing. http://doc.utwente.nl/87377. Accessed 2013

Wolsey LA (1998) Integer programming. Wiley-Interscience series in Discrete Mathematics and Optimization, New York

Zhang Z-H, Jiang H (2014) A robust counterpart approach to the bi-objective medical service design problem. Appl Math Model 38:1033–1040. doi:10.1016/j.apm.2013.07.028

Zhang Z-H, Li K (2015) A novel probabilistic formulation for locating and sizing emergency medical service stations. Ann Oper Res 229:813–835. doi:10.1007/s10479-014-1758-4

Rania Boujemaa is currently a Ph.D. candidate in the department of industrial engineering at the National Engineering School of Tunis (ENIT). Her research focuses on healthcare management and more specifically on the design and the planning of medical emergency service system. In 2012, before undertaking her doctoral studies, she obtained a Master degree in Automatic Control and Production engineering from the Higher School of Sciences and Techniques of Tunis.

Aida Jebali is currently assistant professor in the Industrial Engineering and Engineering Management Department at University of Sharjah (UAE). Prior to joining the University of Sharjah, she was research scientist at Masdar Institute of Science and Technology (UAE), assistant professor of Operations Management at Prince Sultan University (KSA) and assistant professor of Industrial Engineering at National Engineering School of Tunis (Tunisia). She holds a Ph.D. in Industrial Engineering from Grenoble Institute of Technology (France). Her research is focused on supply chain management, maritime logistics, healthcare management, operating room planning and scheduling, and medical emergency planning.

Sondes Hammami is assistant professor of Industrial Engineering at National Engineering School of Carthage (Tunisia). She holds a Ph.D. in Industrial Engineering from Grenoble Institute of Technology (France). Her research is focused on supply chain management, healthcare management, operating room planning and scheduling, and medical emergency planning.

Angel Ruiz is professor in the department of Operations and Decision Systems at Laval University. He holds a Ph.D. from the Université de Technologie de Compiègne (France). His current research interests are in emergency logistics, healthcare management, and medical emergency planning. He is member of the Interuniversity Research Center on Enterprise Networks, Logistics and Transportation (CIRRELT) and head of the CIRRELT's Laboratory on Healthcare Networks.

Hanen Bouchriha is a Professor in industrial Engineering at National Engineering School of Tunis (ENIT). She obtained her HDR in Industrial Engineering from ENIT in April 2011. She holds a Ph.D. in Industrial Engineering from Grenoble Institute of Technology (France) in October 2002. Prior to joining ENIT, she was a Postdoctoral researcher at For@ac Research Consortium, Laval University (2002–2004). She served as the Head of industrial engineering department, ENIT (September 2012–July 2014) and a Member of the scientific board of ENIT (2008–2014). Her main research interests lie in developing optimization tools for Supply Chain configuration and planning with many application areas: Pulp and Paper supply chain, Textile supply chain, recycling waterway sediments networks and hospital networks.