Decision Analytics Journal 3 (2022) 100059

Contents lists available at ScienceDirect

Decision Analytics Journal

journal homepage: www.elsevier.com/locate/dajour

An emergency medical services system design using mathematical modeling and simulation-based optimization approaches

Adel Hatami-Marbini [a], Nilofar Varzgani [b,*], Seyed Mojtaba Sajadi [c], Ahmad Kamali [d]

[a] Department of Management and Entrepreneurship, Leicester Castle Business School, De Montfort University, Hugh Aston Building, The Gateway, Leicester LE1 9BH, UK
[b] Department of Business Systems and Analytics, School of Business, La Salle University, Philadelphia, USA
[c] School of Strategy and Leadership, Faculty of Business and Law, Coventry University, Coventry, CV1 5DL, UK
[d] School of Industrial Engineering, Alborz Campus, University of Tehran, Tehran, Iran

ARTICLE INFO

Keywords:
Emergency medical services
Healthcare system design
Simulation-based optimization
Service operations and improvement

ABSTRACT

The healthcare system is no stranger to resource challenges in the face of unlimited demand to fulfill healthcare objectives of satisfying patients, maintaining service quality, and maximizing profit. Healthcare decision-makers are responsible for devising effective methods for allocating scarce resources fairly and in a way that maximizes benefits at the societal level. An emergency medical services (EMS) system plays a crucial role in stabilizing and transporting seriously injured patients to hospitals within healthcare systems. Several criteria affect the EMS function, such as call rate, traffic condition, setup, and operating costs. Therefore, the optimal design of EMS systems, including determining the location of emergency medical bases and allocating ambulances, helps improve service performance. This paper takes advantage of the mathematical modeling and simulation-based optimization approaches to identify the best location of emergency medical centers and assign the ambulances to the selected centers to maximize survival rate and minimize the total cost of the EMS system. A case study is presented to demonstrate the applicability and efficacy of the developed approach in this study.

1. Introduction and literature review

Over the last few years, the overall healthcare costs have shown a significant increase all across the globe. The complexity of the healthcare systems poses the challenge of managing scarce resources to meet growing patient needs. Therefore, it is essential to adopt the most efficient approaches for resource allocation to achieve this goal. In the area of Dublin, capital of the Republic of Ireland, the estimated cost of supplying the emergency ambulance services was over £28 million in 1996, when handling almost 300,000 patient journeys, in which there were nine emergency ambulance fleets with 272 emergency vehicles located at 89 ambulance stations [1]. In addition to securing pre-hospital emergency medical treatment and transportation of the utterly unwell and injured patients to the hospital, the ambulance fleets transport patients between hospitals for urgent or planned treatment [1]. Some researchers have studied minutely health care management problems to establish the extent to which the emergency ambulance services are provided efficiently and economically. The proposed models enable analyzing and understanding the insider problem, from theoretical as well as practical perspectives (e.g., see [2]). Emergency Medical Services (EMS) are an indispensable part of any health care system that aims at providing prehospital emergency medical care. EMS include the services of supervisors, managers, directors, administrators, and coordinators. When solving for today's greatest EMS challenges amidst rising demands, we need to take into consideration not only the complexity of the model but also the tremendous value of this infrastructure to the community. On the frontline of national disasters and health crises, EMS play a crucial role in preventing deaths and injuries. With the shift in community needs as a result of changing lifestyles and increased life expectancy, requests for EMS have been on the rise. The provision of high-quality EMS is expected to be consistently delivered at the national level. Healthcare managers constantly struggle with competing budget demands and staffing challenges, making delivery of adequate service levels even more difficult. This complexity of EMS systems calls for resource allocation approaches that are capable to optimize productivity and efficiency.

The primary duty of EMS is to offer urgent vital services and transport the patients to a specific hospital or clinic. Many studies have been carried out to seek the optimal locations of facilities, leading to performance improvement of the EMS system. In spite of a wide

* Corresponding author.
E-mail addresses: adel.hatamimarbini@dmu.ac.uk (A. Hatami-Marbini), varzgani@lasalle.edu (N. Varzgani), Mojtaba.Sajadi@coventry.ac.uk (S.M. Sajadi), ahmadekamali@ut.ac.ir (A. Kamali).

https://doi.org/10.1016/j.dajour.2022.100059
Received 11 March 2022; Received in revised form 6 April 2022; Accepted 25 April 2022
Available online 4 May 2022
2772-6622/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

variety of developed models, the key purpose of these models is to assign the emergency facilities optimally to services required, so that the desired outcomes are achieved. This includes studies that cover topics like the location set covering problem (LSCP) developed by Toregas et al. [3] and the maximal covering location problem (MCLP) proposed by Church and ReVelle [4]. Gendreau et al. [5] reconsidered the MCLP as a dynamic relocation strategy to seek new locations for the EMS centers to maximize the coverage level requested for ambulances in the Montreal region. Nevertheless, many related research projects have neglected the real conditions of EMS systems resulting in the research findings not being applicable [6,7]. Hence, to make the models more practical, the stochastic and dynamic nature of EMS needs to be considered when modeling the system. Galvao and Morabito [8] attempted to develope the stochastic models based on queueing theory, in which the ambulances play the role of servers. Ingolfsson et al. [9] extended a model to maximize the average coverage of requests by using the least number of ambulances allocated to patients when the time delay and travel time are assumed to be uncertain.

The existing models for EMS management include complex, dynamic, and contemporary challenges and it necessitates that the efficiency of the EMS system with distinctive properties is analyzed by simulation, which is based upon a computerized model [10]. McCormack and Coates [11] formulated a mathematical programming model for ambulance fleet allocation and base EMS location by the use of the simulation model and the genetic algorithm. Compared to other techniques, simulation has some advantages including the consideration of real characteristics and improvement of the model validity. Mathematical modeling and simulation methods have been originally used by Toro-Díaz et al. [12] to improve the logistics performance of EMS systems. In their paper, they briefly describe lean's application in the logistics and healthcare industries, and conceptually develope the lean-TOC (Theory of Constraint) approach.

With special emphasis on EMS, emergency station location, ambulance allocation, and facilities management are notable subjects for health care studies [7]. The main objective of facility location and resource allocation for EMS is to cover all regions of a city along with providing high service quality at a low cost. The number of requests for EMS depends on working days per week and working hours per day. This is due to the feasibility of improving the efficiency of the EMS system by optimally relocating the emergency stations and allotting ambulances to the stations. Brotcorne et al. [13] reviewed the state-of-the-art models for the emergency facility's location and service allocation over the last three decades.

Generally, location and allocation models in relation to EMS can be categorized into two major groups: deterministic models and stochastic models. The first group lays emphasis upon deterministic mathematical programming models to optimize the management of facilities in a static environment, and in the second group, scholars developed stochastics models with uncertain parameters under dynamic conditions to cover the largest possible demand and optimally allocate resources. One of the primary stochastic models was the maximum expected coverage locating problem (MEXCLP) proposed by Daskin [14]. Daskin's study aimed to meet demand as large as possible with the minimum number of facilities (vehicles). ReVelle and Hogan [15] proposed an optimization model to minimize the number of required ambulances in order to serve the maximum requests with sufficient reliability. Gendreau et al. [16] focused on the redeployment problem for a fleet of ambulances in the context of EMS. The authors attempted to develope a dynamic method to maximize the covered regions with the minimum response time via a parallel tabu search heuristic and simulation.

Given the fluctuating demand for ambulances, Rajagopalan et al. [17] developed a multiperiod model to identify the minimum number of ambulances and their locations over time in which coverage requirements are met with a predefined reliability level. The experimental findings of this study and a simulation model showed that tabu search gives fast and near-optimal solutions for maximizing coverage problems. A tabu search based heuristic has been developed by Toro-Díaz et al. [18] for large-scale EMS systems in which the model contains both strategic (location) and operative (dispatching) decisions to balance efficiency and fairness. Rajagopalan and Saydam [19] proposed a model for allocating a set of ambulances to minimize expected response distances using a heuristic search algorithm while fulfilling coverage requirements. Schmid and Doerner [20] thought of stochastic demand and multiple demand points that are assumed to be covered by more than one vehicle. Their multi-period model allows ambulances to be transported between the stations with the object of maximizing coverage at various points simultaneously over the planning horizon.

Incorporating dynamic and stochastic settings in the models for optimizing EMS systems are clearly valuable in improving the practical implementation of these systems by offering the system providers with high service quality in terms of response time and coverage level. These models are normally referred to as NP and the NP-hard complexity classes (see e.g., [21–23]). To simplify this complexity, the simulation approach has been used extensively in the literature. Simulation is a computerized approach to imitate the real-world operational processes and systems in a dynamic environment, and differs from the analytical approach, in which the system analysis is theoretical. Simulation has been widely used in healthcare applications. For instance, Sah et al. [24] exploited the combination of simulation and goal programming for total system improvement in an Indian hospital leading to the reduction in delays and bottlenecks through the hospital processes. In their research, they specified several criteria including wait time of patients and resources utilization to achieve the goal of analyzing and improving the system. Simulation helped them minimize the total cost of the hospital, subject to adequate allocation of hospital staff and beds.

Aboueljinane et al. [25] provided a survey of simulation models applied to EMS problems. They classified the decisions that affect the performance of EMS systems at the design and operational levels into three categories: long-term decisions, mid-term decisions, and short-term decisions. The *long-term decisions* include decisions about the identification of suitable skills, the number of human resources and the location of the central EMS station. For example, Harewood [26] proposed a multi-objective model to identify EMS locations and also the total ambulances required to meet the service requirement. The *mid-term decisions* consist of decisions around shift scheduling, the total EMS stations and needed resources for them, and scheduling resources like rescue teams and ambulances. For example, Goldberg et al. [27] developed a simulation model to schedule vehicles allocated to EMS centers. The *short-term decisions* are the decisions around dispatching rescue teams in order to improve their efficiency, choosing a suitable hospital or clinic for every patient, and determining redeployment strategies such as EMS relocation in order to offer better services. For example, Peleg and Pliskin [28] presented a redeployment strategy to reduce response time. They used the geographic information system (GIS) to determine the time between EMS locations and demand points and between demand points and hospital locations.

In the literature, the EMS performance is broadly assessed based upon three key factors: *timeliness*, *survival rate,* and *cost*. Hence, these factors have been considered in most studies for improving the system's performance. Wang et al. [29] developed an agent-based simulation of response to a disaster, to study assignment policies of victims to hospitals based on available geographic information systems and available response resources such as ambulances and hospital beds. Nogueira et al. [30] applied both optimization and simulation techniques to study the EMS of Belo Horizonte, Brazil. Their optimization model was aimed to identify the location of ambulance bases as well as allocating ambulances to those bases, and simulation was run to take account of the dynamic behavior of the system.

Van Barneveld et al. [31] studied the effect of ambulance relocations on the performance of ambulance service providers. They modeled the ambulance relocation from the current arrangement to the target arrangement by way of a linear bottleneck assignment problem. In their developed model, the performance of the ambulance service provider was measured by a general penalty function. Zaffar et al. [32] carried out a comparative study to contrast the performance of ambulance location models based upon four criteria; percentage of calls covered, survivability, average response time, and workload balance among the fleet. To this end, they used a simulation–optimization approach to compare the performance of three EMS location models including maximum coverage, minimum average response time, and maximum survivability. They eventually showed that the maximum survivability objective is more efficient than both response time and coverage criteria. Fritze et al. [33] proposed an integrated model of spatial information and integer programming for the EMS location problem. This model applies the MCLP to ensure all residents can cover by EMS at minimum cost. Andersson et al. [34] used the maximum expected performance location problem for heterogeneous regions (MEPLP-HR) developed by Leknes et al. [35] to analyze both strategic (locating ambulance stations), and tactical (allocating ambulances to the stations) decisions over multiple periods.

Unluyurt and Tuncer [36] employed four different mathematical models to determine EMS station locations along with maximizing coverage evaluated. Additionally, they evaluated the performance of EMS location models via discrete event simulation. Eventually, using Istanbul data they showed that a simulation-based evaluation methodology can give a fair framework to assess the effectiveness of models.

Aringhieri et al. [37] reviewed recent studies of the EMS systems to define ongoing challenges for future research avenues. Ahmadi-Javid et al. [38] surveyed articles relevant to a healthcare facility (HCF) location problems and classified the models into two categories: non-emergency facilities, and emergency facilities. Furthermore, they took different perspectives on HCF locations such as location and allocation models for developing health service (see e.g., [39]), ambulance location and relocation models (see e.g., [40]), location of HCFs from modeling aspects (see e.g., [41]), emergency response facility location (see e.g., [42]), methodological advancement in healthcare accessibility (see e.g., [43]), home healthcare logistics (see e.g., [44]), and an overview of planning and management of EMSs (see e.g., [6]). Liu et al. [45] used a robust optimization method for optimizing an EMS system. Their objective is to minimize the total cost of an EMS system based on the station construction, ambulances' location and allocation, and ambulances procurement and maintenance. Boujemaa et al. [46] proposed a multistage stochastic programming model for ambulance redeployment planning. They considered two types of ambulances and two sets of calls for requesting ambulances. Since their model possesses a high degree of complexity, the heuristic method was used for solving the developed model.

Recently, Bélanger et al. [47] provided a broad overview of studies relevant to vehicle location and relocation, as well as dispatching decisions in the context of ambulance fleet management. Firstly, they grouped the studies on static ambulance location models into three classes; (i) single coverage deterministic models (see e.g. [48]), (ii) multiple coverages deterministic models (see e.g. [49]), and (iii) probabilistic and stochastic models (see e.g. [50]), which aims to address ambulance location problems at the tactical level. Secondly, they reviewed the most recent approaches by classifying them into three classes; (i) stochastic and robust location–allocation models (see e.g., [51]), (ii) maximal survival models (see e.g., [52]), and (iii) equity models (see e.g., [53]). Thirdly, they examined multi-period relocation models and dynamic relocation models in ambulance location/relocation problems (see e.g. [54]). Finally, they reviewed research on dispatching decisions for allocating the vehicle to an emergency call.

Measuring the performance of health care systems is an essential process with some difficulties and challenges resulting in identifying weaknesses and inefficient sources. Data envelopment analysis (DEA) is a well-known non-parametric method for assessing the relative efficiencies of a group of decision-making units (DMUs). Farrell [55] originated many of the ideas and principles underlying DEA. After a long-term period, Charnes et al. [56] built on the provoking thought seminal work and introduced a powerful DEA methodology to assess the relative efficiencies of multi-input multi-output DMUs. Since the emergence of DEA, there has been a significant growth both in theoretical developments and applications [57–59]. DEA has also been used to assess different aspects of healthcare systems such as hospital efficiency [60], public policies efficiency [61,62], heart surgery efficiency [63] and health facilities efficiency [64–66]. Golabian et al. [67] conducted a study to obtain the best return strategy for ambulances to maximize the expected coverage concerning a predefined dispatch policy. They proposed a hypercube queuing model to maximize customers' coverage probability, in which locations of busy ambulances in each state are not known and approximated based on customer arrival rates.

There are some papers which extend the analysis to either predicting demand for ambulances or determining temporary emergency service center location decisions in disasters. Grekousis and Liu [68] put forth novel approach by attempting to predict the demand for ambulances in EMS using artificial intelligence. They proposed a three level model to first predict the future demand and then apply a location–allocation model to site ambulances prior to actual emergencies occurrence. They also used a case study based on data from Athens, Greece on the actual emergency events occurred to validate their model. Karatas and Yakıcı [69] proposed a multi-objective facility location analytics model for determining the number and locations of Temporary Emergency Service Centers (TESCs) for a regional natural gas distribution company in Turkey. While their work is important in natural disasters and other extreme events, it may not directly be applicable in meeting existing demand for EMS outside these events.

DEA developed by Charnes et al. [56] is an exceedingly endorsed and powerful method for measuring the performance of the public and private sectors [57,70]. Since the mid-1980s, there has been increasing interest in the application of DEA to healthcare problems. The literature review conducted by Hollingsworth et al. [71] examined 91 DEA-based studies on efficiency and productivity in healthcare systems from both theory and practice. Their review is divided into two groups of studies; the deployment of DEA to measure efficiency and productivity of hospitals such as Burgess and Wilson [72], and general healthcare such as Färe et al. [73]. Hollingsworth [74] built on the earlier survey and provided an overview of 188 related studies on non-parametric and parametric efficiency measurement in healthcare and health. Chilingerian and Sherman [75] reviewed the DEA literature focusing on efficiency measurement of health care providers such as general hospitals and academic medical centers, nursing homes, and physicians. They particularly discussed DEA models applied in health care application as well as listing inputs and outputs as the consequence of the extant research literature.

Although many research studies have been done to assess the EMS performance, the pertinent literature pays less attention to systems cost as a vital assessment factor. Considering real conditions help us to gain more practical and reliable solutions to EMS problems. Although some recent EMS studies such as Nickel et al. [76] and Boujemaa et al. [77] have tried to consider real conditions in their models, they have neglected some important conditions such as weather and traffic.

Fig. 1 displays the processes of patient rescue in the EMS system. This Figure defines all steps of the rescue process to make a valid model for the emergency location and allocating ambulances.

Referring to Fig. 1, response time, denoted by T_{R}, is defined as the time elapsed from a call received by the call center to the time that the ambulance arrives at the patient position location. Response time is a key factor and has a direct effect on the survival rate in such a way that the decrease in response time leads to a rise in the survival rate. That is, a quick EMS response is essential in improving survival rates and EMS performance. This study presents a simulation-based optimization approach based on the maximization of the survival rate and minimization of the total cost that is purposely developed for

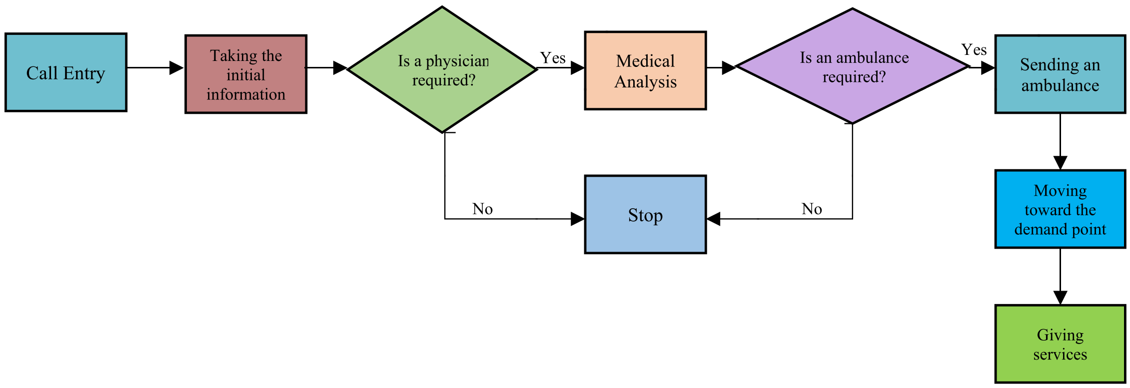

Fig. 1. Rescue processes for patients.

a project at the Emergency Management Centre in Isfahan, a central province in Iran to improve the service performance. In this respect, we build on the work by Knight et al. [78] and develop their work by adding cost minimization as an objective function thus enabling managers to improve the total performance of the EMS system. Contrary to Knight et al. [78], we also assume two different types of ambulances and establishment cost of EMS centers, as well as considering weather and traffic as stochastic conditions of EMS location and allocation. The results calculated from the optimization model is considered as the primary scenario to make our simulation model. Other scenarios build on the primary scenario by altering the variables such as the number and location of EMS. In addition, the stochastic conditions of emergency locations and their allocated ambulances are considered in the simulation model. To compare and prioritize the results obtained from simulation, the DEA method is applied on the basis of two predetermined objectives: survival rate, and total cost. In a nutshell, the theoretical and practical contributions of the framework developed in this study are threefold:

i A dual-objective optimization model is formulated to design an EMS system by maximizing survival rates and minimizing the total cost simultaneously. The mathematical model culminates in the optimal number and location of required EMS as well as allocating the optimal number of ambulances to the selected EMS locations.
ii A computer simulation analysis is applied to take account of the dynamic conditions of the EMS system in order to reach an appropriate response.
iii Finally, we draw on the dataset from the Emergency Management Centre in Isfahan to illustrate the applicability and efficacy of the proposed framework. The results show that the implementation of our proposed framework leads to improved survival rate.

The remainder of this paper is organized as follows; Section 2 describes the modeling methodology and its undelaying assumptions, Section 3 presents a case study from the city Isfahan, and Section 4 concludes this study with some remarks.

2. Preliminaries

In this section, the maximal expected survival location model for heterogeneous patients (MESLMHP) and DEA models are, in turn, reviewed and formulated [56,78].

2.1. *Maximal Expected Survival Location Model for Heterogeneous Patients (MESLMHP)*

The survival function plays a key role in locating the ambulance stations nearer to areas with high demand for EMS since the possibility for survival decreases exponentially with increasing response time. As a significant contribution, Erkut et al. [40] proposed EMS location models with the inclusion of a survival function in the existing coverage models. They showed that more traditional coverage models suffer from some weaknesses in practice and a survival-maximizing approach enables the decision-maker to obtain reliable and robust solutions. "A *survival function* is a monotonically decreasing function of the response time of an emergency medical service (EMS) vehicle to a patient that returns the probability of survival for the patient" ([40], p 42). Rather than designing the network location model, Knight et al. [78] developed the MESLMHP based upon Erkut et al. [40] to improve the performance of an extant EMS system. In this respect, MESLMHP aims to maximize the overall expected survival probability of different types of patients. Loosely speaking, several various types of survival probabilities can be defined based upon patient types. The patient types can be defined on the basis of time standard, for instance, type I patients are those that the ambulance reaches within 8 minutes, or by observing various patient conditions with respect to a relative survival curve.

Assume that there are m ambulance stations, and n demand locations. Patients are categorized into k different patient types on the basis of severity level and survival probability. Let λ_i^l indicate the demand of patient type l ($l = 1, 2, \ldots, k$) from demand location i ($i = 1, 2, \ldots, n$). The average utilization of ambulance type h at station j is denoted by π_j. The mathematical notations are provided in Table 1. In the view of π_j, the probability of ambulance availability is indicated by $(1 - \pi_j)$. The preference matrix $\rho = \left[\rho_{ij}\right]_{n\times m}$ is used to ensure that an emergency station is allocated to a demand point. We point out that if the jth emergency station has no available ambulance the $(j + 1)$th station will be selected. Furthermore, there can be two general survival functions where the first function is determined on the basis of severity and response time target, and the second one is based on considering different patient conditions with respect to survival probabilities. Notice that a combination of both functions can be observed in some circumstances. Fig. 2 displays three feasible survival functions. The red and green solid lines represent the survival functions that show that the survival probability decreases with time, and quick action of EMS is essential. Another function depicted by a blue solid line in Fig. 2 is a step function that is used for thinking of rigid targets such as defined service norms. Let $s_l : \Re \to [0, 1]_{\Re}$ be a survival function as for each patient type l and t_{ij} be the travel time between station j and demand node i. There might be l different survival functions, which are reliant on the number of patient types and the importance weights w_l are allocated to the lth patients. The MESLMHP model is given below:

$$\max \sum_{l=1}^{k} w_l \sum_{i=1}^{n} \lambda_i^l B_{i,\rho_{ij}}^l \tag{1}$$

Table 1
List of notations used in mathematical formulation of the MESLMHP.

Notations		Description
Indices	i	Index for demand locations $(i = 1, 2, \ldots, n)$
	l	Index for patient types $(l = 1, 2, \ldots, k)$
	j	Index for ambulance stations $(j = 1, 2, \ldots, m)$
	λ_i^l	Demand of patient type l from demand location i
Parameters	π_j	Average utilization of ambulance at station j
	Z	Total number of available ambulances
	ρ_{ij}	jth preferred station for the ith demand point
	t_{ij}	Travel time between station j and demand node i
	w_l	Relative importance of patient type l
	$s_l(t_{ij})$	Probability degree of survival associated to patient type l for time t_{ij}
Decision variables	x_j	Number of ambulances allocated to emergency station j

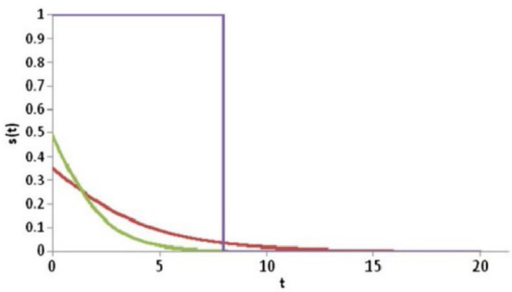

Fig. 2. Survival functions.
Source: Adopted from Knight et al. [78].

st.

$$\sum_{j=1}^{m} x_j = Z$$

$$x_j \in \mathbb{Z}^{0+}, \; j = 1, 2, \ldots, m$$

where $B^l_{i,\rho_{ij}} = \sum_{j=1}^{m} s_l(t_{i,\rho_{ij}}) \underbrace{(1 - \pi_{\rho_{ij}}^{x_{\rho_{ij}}}) \prod_{r=1}^{j-1} \pi_{\rho_{ir}}^{x_{\rho_{ir}}}}_{\varnothing}$ indicates the probability (expected degree) of survival of a patient of type l from demand node i where $\varnothing$ represents the probability of at least one ambulance located at j service node being free, while ambulances located at nodes 1 to -1 are busy. The equality constraint set of model (1) shows that Z ambulances are allocated to all EMS stations to maximize the weighted sum of all demand nodes in terms of each probability function. To find the actual utilization π_j, the iterative approach is developed on the basis of queuing theory where each EMS center can be modeled as an $M_j/M_j/x_j$ queue. The iterations are stopped after an allocation of ambulances matches the demand and the utilization.

The MESLMHP includes the following six phases [78]:

Phase 1: Estimate the mean service rate at station j (μ_j).

Phase 2: Set the average utilization of the jth ambulance station for the initial iteration $(\pi_j^{(0)})$.

Phase 3: Find x_j for all centers by solving the MESLMHP problem.

Phase 4: Compute the demand distribution $\Delta_j^{(k)}$ for the jth ambulance station.

Phase 5: Obtain π_j based on queuing model for $M_j/M_j/x_j$ which ensure the allocation calculated in Phase 3.

Phase 6: Iterate phases 3, 4 and 5 until convergence is achieved. Since this method does not guarantee convergence, the model runs for some pre-determined iterations or a given time. The steady state[1] graphs for cost or survival can be used to specify where convergence occurs.

[1] A steady state is a case where all state variables remain constant.

Our paper aims to consider the total costs of the EMS system installation, e.g., buildings and equipment costs. Therefore, our proposed model is a bi-objective optimization model to design an emergency medical system in which the first objective function aims to maximize the survival rate and the second one aims to minimize the total cost of the EMS system. Notably, the simulation method is used in this paper to seek the value of x_j, $(j = 1, 2, \ldots, m)$. In addition, the demand rate of each demand point is extracted from the data calls.

2.2. Data Envelopment Analysis (DEA)

DEA is a well-established approach for evaluating the relative efficiency of a group of functionally similar decision-making units (DMUs) (e.g., institutions, banks, hospitals, and hotels) that transforms multiple inputs (resources) into multiple outputs (goods and service) [56]. DEA builds on the "total weights flexibility" underpinning. The original DEA model proposes that the efficiency of a DMU is given as the maximum of a ratio of weighted outputs to weighted inputs, subject to the constraint that the same ratio for all the DMUs must be at most equal to one. Consider n DMUs, in which each DMU consumes m various inputs to produce s various outputs. Let $x_{ij}, i = 1, \ldots, m$ and $y_{rj}, r = 1, \ldots, s$ denote the input and output vectors of DMU$_j$, $j = 1, \ldots, n$, respectively. Charnes et al. [56] proposed the first DEA-CCR model under constant returns to scale (CRS) followed by the DEA-BCC model [79] that hypothesizes variable returns to scale (VRS). This paper's focus is on the following VRS model to measure the technical efficiency of the DMU$_o$:

$$\max \rho = \sum_{r=1}^{s} u_r y_{ro} + u_o \tag{2}$$

st.

$$\sum_{i=1}^{m} v_i x_{io} = 1,$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + u_o \le 0, \; j = 1, 2, \ldots, n,$$

$$v_i, u_r \ge 0, \quad i = 1, 2, \ldots, m; r = 1, 2, \ldots, s.$$

where u_r and v_i are the importance weights associated with rth output and ith input, respectively, and ρ represents the efficiency measure of DMU$_o$. Note that the free-in-sign variable, u_o, is associated with the convexity condition in the dual model leading to a convex hull. If the optimal objective value is equal to 1, $\rho^* = 1$, then the DMU$_o$ is called *efficient*. Otherwise, the DMU$_o$ is called *inefficient*.

3. Proposed model

In healthcare, fixed expenditures associated with buildings and equipment play a pivotal role in healthcare management and absorb more than 80 percent of the total cost. Although increasing the number of ambulances, stations, and crews on duty can improve the ability of EMS to respond to emergency calls, the system is not necessarily efficient because the rise in resources would result in long periods of inactivity for crews between calls and massive fixed cost. In this paper, the objective is to design an EMS system that minimizes the total cost (costs of preparation and construction of stations and purchasing cost of ambulance vehicles) while maximizing the overall expected survival probability of patients. Therefore, in reality, there are several types of ambulances in EMS systems. For instance, there are four types of ambulances in the US.[2] Types *i* and *iii* ambulances have a square patient compartment that is installed onto the chassis. The difference between types *i* and *iii* is their chassis whereby type *i* is attached to a truck similar to chassis while type *iii* is attached to a cut-a-way van chassis. Types *i* and *iii* ambulances are often equipped to be used in the locations such as airports, chemical plants, oil refineries as well as being ready for Advance Life Support. Type *ii* ambulances with a van type chassis are commonly used by hospitals, Health departments when patients require basic life support. Type *iv* is classified as mini ambulances and their compact design enables them to maneuver in the regions where other types of ambulances may not be possible, leading to a great decline in the overall response time.

Contrary to the MESLMHP, without loss of generality, it is assumed that two types of ambulances ($h = 1, 2$) operate in EMS, so-called *general*, and *advanced* ambulances. General ambulances are equipped with primary and emergency facilities while advanced ambulances have more facilities such as some devices for cardiac patients. According to the above-mentioned ambulance types, types *i* and *ii* are classified as *general ambulances* and type *iii* is classified as *advance ambulances*.

Let Z_h denote the maximum number of the hth ambulance type which are available to respond to any emergency. Here, c_1 and c_2 present the purchasing cost of *general* and *advanced* ambulance types, respectively. The fixed activation and construction cost of each station is assumed to be identical for all stations and denoted by C, and the maximum number of stations which can be in operation is denoted by D. We have two zero–one variables y_j and x_j^h in which y_j indicates whether or not station j is in operation and x_j^h shows the number of ambulances of types 1 and 2 allocated to emergency station j. Given the definitions of parameters and variables mentioned above, we propose the following bi-objective network location model to find the optimal number of necessary stations and allocated ambulances in terms of their types:

$$\min \sum_{j=1}^{m} C y_j + \sum_{j=1}^{m} \sum_{h=1}^{2} c_h x_j^h \tag{3a}$$

$$\max \sum_{h=1}^{2} \sum_{l=1}^{k} w_l \sum_{i=1}^{n} \lambda_i^l \sum_{j=1}^{m} s_l(t_{i,\rho_{ij}})(1 - \pi_{\rho_{ij}}^{x_{\rho_{ij}}^h}) \prod_{r=1}^{j-1} \pi_{\rho_{ir}}^{x_{\rho_{ir}}^h} \tag{3b}$$

st:.

$$\sum_{j=1}^{m} x_j^h \leq Z_h, \qquad h = 1, 2, \tag{3c}$$

$$\sum_{j=1}^{m} y_j \leq D, \tag{3d}$$

$$x_j^h \leq M y_j, \qquad h = 1, 2; j = 1, 2, \dots, m, \tag{3e}$$

$$x_j^h \geq y_j, \qquad h = 1, 2; j = 1, 2, \dots, m, \tag{3f}$$

$$x_j^h \in \mathbb{Z}^{0+}, \qquad h = 1, 2; j = 1, 2, \dots, m, \tag{3g}$$

$$y_j \in \{0, 1\}, \qquad j = 1, 2, \dots, m. \tag{3h}$$

The first objective function includes two components: the construction cost of EMS centers and the total cost of buying ambulances. Contrary to model (1), the second objective function (3–2) is generalized to two types of ambulances to maximize the overall expected survival probability of patients which is of the essence for designing an emergency medical services system. Notice that similar to model (1), ρ_{ij} denotes the jth preferred station for the ith demand point and $x_{\rho_{ij}}^h$ is the number of allocated ambulance type h to the preferred station. Constraints (3-3) ensure that the total allocated type h ambulances are at most Z_h. Constraint (3-4) guarantees that the total constructed stations are at most D. Constraints (3-5) and (3-6) help to define zero–one variables.

The demand distribution of each demand point is estimated by the use of the historical call data which are collected in the emergency system on a daily basis.

Markedly, the above model cannot be easily solved in the way that Knight et al. [78] proposed since solving two conflicting objective functions on the basis of queuing theory is quite complex and adds time. Due to the fact that the real-world EMS problems are dynamic, it is also highly desirable to analyze the problem over time while the existing optimization tools are not sufficiently appropriate. Therefore, we present a structured framework in this study to achieve our objectives.

To find an initial solution in model (3), the stochastic variables are first assumed to be represented by their expected values. Therefore, the stochastic model is transformed into a deterministic model. The solving methodology is based on a series of steps as follows:

1. Determining the potential locations for an emergency station based upon various factors such as population density and accessibility.
2. Defining the demand points. All people living in a city are potential demand points. For simplicity's sake, the gravity centers of areas in the city are considered as demand points.
3. Estimating traveling time between each station and demand points in the light of various important factors such as the traffic and weather conditions.
4. Estimating the costs of buying the ambulances and constructing the emergency stations.
5. Solving the proposed model (3) under the static condition to reach the primary solution. This solution is used as a primary input for the computer simulation model in Step 6.
6. Applying simulation to analyze the different scenarios which can help consider the real-world conditions in the mathematical model. A range of scenarios can be defined via changing (i) the number of emergency stations, (ii) the allocated ambulances and (iii) the probability of ambulances' availability.
7. Using DEA to assess and rank the results of the scenarios obtained from simulation in Step 6. The scenario with the superior score is the best arrangement (design) for the emergency medical services system.

4. Case study

In this section, we present an application to the province of Isfahan in Iran. Isfahan is located in southern central Iran, 1430 meters above sea level, with about 5 million population and covers an area of nearly 107,027 square km. The current state of the Isfahan emergency system shows that the costs associated with the emergency and service time for serving patients (transport, immediate treatment, and medical transmission) are relatively high. This issue has been discussed with experts and managers in several lengthy meetings in the form of brainstorming and the Delphi method. Isfahan health officials have concluded that the root cause of the problem is the improper location of emergency centers and ineffective ambulance allocation. Moreover, the Isfahan health authority has decided to rise the survival rate by increasing the number of ambulances.

The city of Isfahan is the provincial capital consisting of 15 areas (see Fig. 3). The most crowded areas of Isfahan are areas 1, 3, 5 and 6. On the basis of the ageing population, the average age of these areas is higher than others and historically more than 41 per cent of the

[2] http://metronixinc.com/site/ambulances.html

Table 2
Patient types.

Type	Description	Standard response time (Min)
I	Urgent and emergency help	9
II	Serious but not urgent	15
III	Non-life-threatening	20

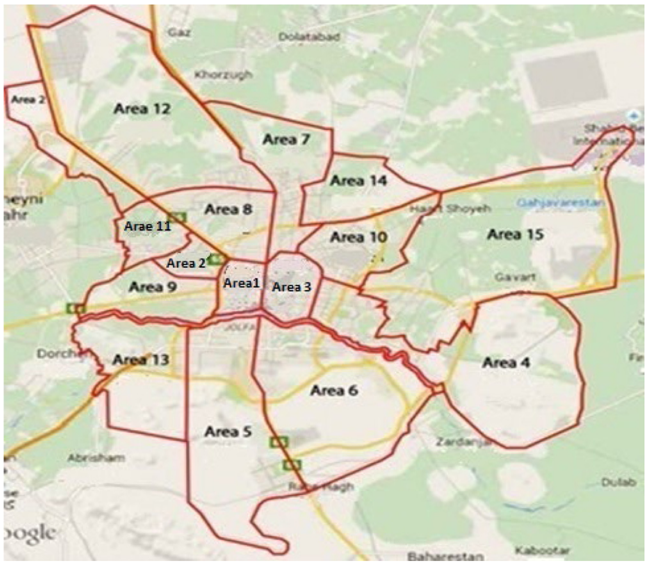

Fig. 3. Isfahan areas.

total emergency calls are received from these four areas of Isfahan. Therefore, their EMS coverage plays a vital role for the health care managers of the city. For this reason, we apply the proposed method in this paper to areas 1, 3, 5 and 6 of Isfahan. Each area is divided into 3 smaller sections, resultantly, there are 12 urban (demand) points. In addition, patients are classified into 3 types as shown in Table 2.

We point out that the standard response times in Table 2 are determined by the Ministry of Health, Treatment and Medical Education of Iran. Defining different response times reflect the heterogeneous groups of the population that the EMS system serves. As the demand rate and traffic conditions affect response time, we partition all hours of the day into four smaller intervals including dawn (0:00–5:59), morning (6:00–11:59), afternoon (12:00–17:59), and evening (18:00–23:59). The demand rate of the different demand points in terms of each patient type was collected for a 2-month period of time as summarized in Table 3.

The health care management team decided to establish 12 emergency stations at most to cover all sections in the four areas. In addition, in collaboration with several practitioners, a potential location for the emergency center in every section is determined, and the population gravity centers are used to form the coordinate of all demand points. Table 4 displays the travel time between emergency station j, $j \in \{A, B, C, D, E, F, G, H, I, J, K, L\}$ and demand point i, $i \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ in four pre-determined time intervals, which are obtained using GIS software and google maps.

A wide variety of survival functions can be used to show the probability of survival. There are two general survival functions consisting of exponential decay functions and step functions. In the exponential decay function, the survival probability decreases over a period of time while the step function is a piecewise constant function for the survival probability that jumps by a given amount at a finite number of points. In this paper, the standard response time is considered to define the step survival functions. Table 5 shows the survival functions and relative importance associated with each patient type.

After consultations with the health care management team of Isfahan, the maximum number of ambulances that the EMS procurement team can provide are 37 general ambulances and 12 advanced ambulances. The proposed model (3) is solved by the use of CPLEX solver in the general algebraic modeling system[3] (GAMS) to solve the Mixed Integer Linear Programming (MILP) and find a primary solution. The weighted sum method is used to solve the two-objective optimization problem.[4] The number of ambulances allocated to the emergency stations is reported in Table 6 and the respective total cost is $2,152,037. In addition, the survival rates of the patient types I, II and III are 0.125, 0.638 and 0.803, respectively. It should be noted that we do not consider the dynamic conditions and divided time when running the model. It means that the equal amount is considered for demand rate and travel time at all times.

It is of interest to study the design problem over time. The complexity in observing the dynamic conditions in the proposed model leads us to apply the simulation model and provide insight based on the results. Simulation is an advantageous method to assess and analyze the situation of current and newly designed systems [80,81].

While an optimization model aims to calculate the optimal values of decision-making variables, solving the model with multiple dynamic and uncertain parameters is often intricate and uneconomical. To solve and analyze these types of problems, simulation models can be combined into optimization models, which is referred to as a simulation-based optimization method. Simulation-based optimization is a structural method for determining the optimal parameters in the system and, in turn, the objective function is measured based on the simulation model [82]. The simulation based optimization method in a constant time enables us to evaluate the system's behavior at specific values of input variables. A simulation experiment is defined as a test or a series of tests in a way that significant variations are made to the input variables of the simulation based optimization.

Thereby, the results obtained from simulation experiments can help the decision-maker analyze changes in input variables (parameters) and select the best values for them [83]. Generally, a simulation based optimization model includes n input variables $(x_1, x_2, \ldots, x_n)$ and m output variables $(y_1, y_2, \ldots, y_m)$ with the aim of defining a set of optimal input variables which optimize the output variables. The simulation based optimization model is a suitable method for solving complicated and dynamic problems. As shown in Fig. 4, the output of a simulation model is used by an optimization strategy to find the optimal value for output variables along with getting feedback on how improvement can be made based on the [nearly] optimal solution. Obviously, the role of defining an appropriate optimization strategy is of great importance in this approach.

Let us construct the following six steps with the aim of analyzing the problem dynamically via simulation and reaching a more realistic and reliable solution:

1. The primary solution is calculated from model (3).
2. Assigning an emergency station to each demand point with the maximum 3 general and 1 advanced ambulance.
3. Assigning an emergency station to every two neighbor demand points with a maximum 6 general ambulances and 2 advanced ambulances.
4. Assigning a fixed emergency station to every two neighbor demand points with the maximum 6 general and 2 advanced ambulances as well as allocating a temporary emergency center (a mobile general ambulance located in crowded points temporarily) to each pair of demand points.

[3] https://www.gams.com/

[4] According to Toro-Díaz (2015), meta-heuristic methods can be used to solve large-scale problems.

Table 3
Demand rate of each patient type.

Time interval	Demand points	Patient types			Time interval	Demand points	Patient types		
		I	II	III			I	II	III
Dawn	1	52	61	60	Dawn	7	32	38	41
Morning		81	83	81	Morning		36	40	39
Afternoon		111	116	95	Afternoon		68	70	64
Evening		163	158	102	Evening		67	84	72
Dawn	2	42	46	40	Dawn	8	26	34	29
Morning		63	64	62	Morning		45	48	37
Afternoon		104	106	114	Afternoon		48	52	48
Evening		125	169	171	Evening		53	68	55
Dawn	3	76	82	62	Dawn	9	21	29	24
Morning		77	86	52	Morning		30	33	28
Afternoon		131	142	130	Afternoon		46	48	41
Evening		122	130	141	Evening		48	50	49
Dawn	4	65	66	68	Dawn	10	12	16	18
Morning		83	85	75	Morning		16	18	22
Afternoon		92	94	94	Afternoon		25	29	35
Evening		121	127	103	Evening		62	65	58
Dawn	5	33	36	28	Dawn	11	18	27	23
Morning		89	92	74	Morning		19	21	20
Afternoon		127	140	135	Afternoon		26	29	34
Evening		156	161	124	Evening		38	42	36
Dawn	6	47	54	46	Dawn	12	15	18	18
Morning		48	54	38	Morning		18	23	17
Afternoon		82	92	84	Afternoon		48	52	41
Evening		51	92	75	Evening		54	61	57

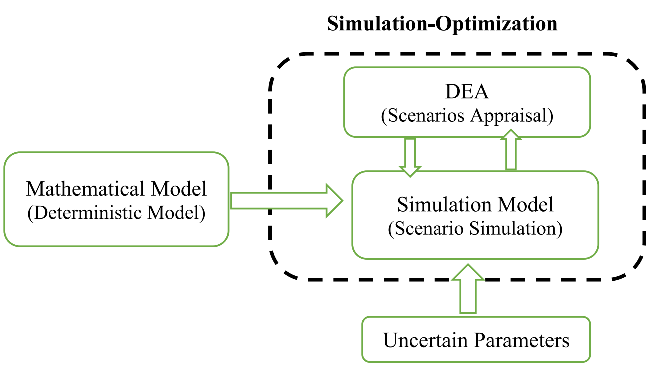

Fig. 4. A simulation–optimization model.

5. Allocating a fixed emergency station to every two neighbor demand points with the maximum 3 general ambulances and 1 advanced ambulance as well as allocating a temporary emergency center (a mobile general ambulance located in crowded points temporarily) to each pair of demand points.
6. Allocating an emergency center to demand points that receive the most call rate (80%) for patient type I with the maximum 3 general ambulances and 1 advanced ambulance. The remaining demand points are treated by Scenario 3.

Considering Fig. 1, the simulation runs the corresponding model for defined scenarios via Arena software. Notably, historical data collected from Isfahan Health Organization is fitted to a theoretical distribution of the proposed framework. The validation of the proposed model for the case study is carried out by tracking the entities ambulances and patients) and resources of the simulation model and their adaptability to the expected cases is inspected by Isfahan Health Organization experts. The subject matter experts therefore follow the entities to make sure that all the constraints in the proposed mathematical model have been covered by the simulation model. Furthermore, the outputs of simulation model for some certain situations and well-known conditions have been considered [84]. The simulation model includes probabilistic parameters. The historical data are used to estimate the probability distribution for each parameter as resulted in Table 7.

It should be noted that the demand rate of each patient type and ambulance response time is reported in Tables 3 and 4, respectively. Furthermore, the constant parameters are shown in Table 8.

As it is not straightforward to define the start and endpoints for the simulation model, a steady-state model is presumed. Therefore, the number of replications needs to be first determined. First, the model is run 30 times and the statistical description of the outputs is shown in Table 9.

The equation $n \cong z^2_{1-\frac{\alpha}{2}} \frac{s^2}{h^2}$ is used to identify the total number of replications, in which s is the standard deviation of the initial replications and h is a fixed number showing the half domain of confidence interval. Table 10 consequently shows the total number of replications of each scenario.

Having gone through Table 10, the results allow us to think of 110 imitations for the simulation model. The replication length in this model is 24 hours for each repetition. A warm-up period, T, is then applied to help determine the running time for simulation and ensure that the objective function values are measured in a steady state. Let us focus on Scenario 1 for example. As our model is sought to obtain the values of three survival rates and cost, fours Ts is considered for Scenario 1. Comparing the four trends in Fig. 5 shows that the longest convergence time is related to the survival rate (c). In other words, the survival rate for Scenario 1 reaches the steady-state status when $T = 8000$ h, Likewise, T can be estimated for the remaining scenarios.

Table 11 exhibits the survival rates and total cost of each scenario after simulating the model 110 times.

Generally, results show that the additional service station and ambulance lead to an increase in survival rates and total cost. Scenario 2 as the simplest scenario is not acceptable to some extent because it has the maximum cost compared with other scenarios and the survival rates of A and B are minimal. It means that increasing cost in Scenario 2 does not improve all the survival rates considerably. In addition, it might be difficult to make decisions on the basis of Scenarios 3 to 6 due to the conflicting results for the four decision variables as summarized in Fig. 6. That is to say, though the survival rate plays a key part in

Table 4
Travel time between demand points and service nodes.

Time interval	Demand points	Service nodes (minutes)											
		A	B	C	D	E	F	G	H	I	J	K	L
Dawn	1	3.1	12.4	13.3	4.1	3.6	14.5	15	3.2	2	14	15.1	13.8
Morning		3.7	12.6	13.3	4.1	3.7	14.8	15.1	3.4	2.4	14.2	15.4	14
Afternoon		3.9	13.5	13.8	4.9	4	15.4	15.6	3.7	2.5	14.5	16	14.6
Evening		4.2	14	14.1	5.3	5.1	15.8	16	14.1	3	15.3	16.2	15
Dawn	2	12.5	14.1	13.6	15	2.4	13.1	13	14.8	16.5	17.3	14.1	2
Morning		12.7	14.2	13.7	15	2.5	13.3	13.2	15	16.8	17.5	14.4	2.2
Afternoon		13	14.5	14	15.4	2.9	13.9	13.4	15.2	17.2	17.7	14.6	2.4
Evening		13.5	15	14.8	16.1	3.5	14.7	14	15.8	17.6	18.2	15.4	3.3
Dawn	3	17.5	15	3.7	16.8	4.1	12.2	13.8	14.9	6.4	8.2	3.4	2.7
Morning		17.8	15.4	4	17	4.5	12.5	14.2	15.4	7	8.5	3.8	13.4
Afternoon		18.2	16	4.5	17.5	5	13	14.9	16	7.7	9.1	4.2	14
Evening		18.5	16.6	5.2	18.1	5.7	13.6	15.4	16.7	8.2	9.8	4.9	14.6
Dawn	4	19.5	5.2	8.2	14.2	21.1	17.5	25.1	9.3	8.8	22.4	25.4	3.5
Morning		19.7	5.4	8.5	14.6	21.3	17.8	25.6	9.5	9	22.6	25.7	3.7
Afternoon		20	6.1	8.9	15	21.6	18.2	25.9	10.1	9.2	23.1	26.3	4.2
Evening		20.3	6.1	9.2	15.3	22.2	18.6	26	10.6	9.6	23.4	26.7	4.8
Dawn	5	9.9	28.1	18.1	16.8	26.2	8.5	4.2	16.2	18.3	21.1	18.4	3.2
Morning		10.2	28.4	18.3	16.9	26.2	8.6	4.3	16.2	18.5	21.3	18.6	3.5
Afternoon		10.6	28.5	18.8	17.5	26.5	9	4.6	16.5	18.9	21.6	19	3.7
Evening		11	28.7	19.2	17.9	26.9	9.4	4.6	17	19.2	21.8	19.4	4
Dawn	6	14.2	8.2	6.4	12.2	18.1	5.4	4.3	3.1	17.9	4.5	5	6.8
Morning		14.5	8.4	6.4	12.4	18.1	5.7	4.5	3.4	18.2	4.8	5.3	7
Afternoon		14.7	8.7	6.6	12.8	18.5	6	4.8	3.6	18.3	5	5.5	7.2
Evening		15.1	9	6.9	13	18.7	6.2	5	3.8	18.7	5.5	6	7.8
Dawn	7	3.2	4.1	5.6	10.3	11.4	5.6	17.8	14.3	3.1	2.2	5.7	16.3
Morning		3.3	4.3	5.9	10.5	11.6	5.8	18	14.5	3.3	2.5	6	16.5
Afternoon		3.5	4.5	6.2	10.7	11.8	6	18.2	14.7	3.5	2.7	6.3	16.8
Evening		3.7	4.8	6.5	11	12	6.3	18.5	15	3.8	3	6.7	17.2
Dawn	8	5.4	10	9.8	8.7	13.4	15.4	6.1	17.4	14.2	16.3	11.2	13
Morning		5.6	10.5	10.2	9	13.5	15.5	6.3	17.5	14.4	16.5	11.5	13.3
Afternoon		6	10.8	10.5	9.8	13.8	15.8	6.5	17.7	14.7	16.8	11.8	13.5
Evening		6.6	11.3	11	10.2	14.2	16.2	6.8	18.2	15	17.5	12.3	13.8
Dawn	9	15.8	14	10.2	8.9	13.5	16.8	3.2	4.1	5.3	4.8	18.3	6.8
Morning		16	14	10.5	9.1	13.7	17	3.5	4.3	5.5	5	18.5	7
Afternoon		16.2	14.5	10.8	9.4	14	17.3	3.8	4.5	5.8	5.3	18.7	7.2
Evening		16.5	14.7	11.2	9.6	14.2	17.5	4.2	5	6.3	5.8	19	7.8
Dawn	10	3.4	5.5	4.2	12.8	8.5	6.7	7.8	10.2	11.3	15.4	17.2	9.1
Morning		3.6	5.7	4.5	13	8.7	7	8	10.5	11.5	15.8	17.5	9.3
Afternoon		4	6	4.8	13.3	9	7.2	8.3	10.7	11.7	16	17.9	9.5
Evening		4.5	6.4	5.3	13.5	9.5	7.5	8.8	11.1	12	16.5	18.3	10
Dawn	11	12.2	10.1	6.8	4.2	3.3	14.8	15	12.1	3.4	2.1	5.8	14.8
Morning		12.5	10.3	7	4.5	3.5	15	15.2	12.3	3.6	2.2	6	15
Afternoon		12.8	10.5	7.3	4.8	6	15.2	15.5	12.5	4	2.5	6.3	15.3
Evening		13.2	10.8	7.5	5.2	6.5	15.5	16	12.8	4.5	3	6.8	15.8
Dawn	12	3.2	2	4.6	8.8	9.8	12.3	13.5	16.8	7.8	9.2	13.7	5.5
Morning		3.5	2.2	5	9	10	12.5	13.8	17	9	9.6	14	5.9
Afternoon		3.9	2.8	5.4	9.5	10.4	12.8	14.2	17.3	9.4	10	14.3	6.4
Evening		4.4	3.3	6	9.8	11	13	14.5	17.6	9.9	10.6	14.7	7

Table 5
Proposed survival functions and relative importance of each patient type.

Patient type	Step survival functions	Weights
I	$s_I(t) = \begin{cases} 1 & for\ 0 \le t \le 9 \\ 0 & for\ t > 9 \end{cases}$	4
II	$s_{II}(t) = \begin{cases} 1 & for\ 0 \le t \le 15 \\ 0 & for\ t > 15 \end{cases}$	2
III	$s_{III}(t) = \begin{cases} 1 & for\ 0 \le t \le 20 \\ 0 & for\ t > 20 \end{cases}$	1

an EMS system, resource limitation does not allow decision-makers to improve this factor by unreasonable and infeasible decisions.

Let us use the DEA model to compare these 6 scenarios and choose the best scenario. To do so, there is a need to define the inputs and outputs for each scenario. The inputs are (i) the number of selected EMS stations, (ii) the number of the general type of ambulances allocated to stations, and (iii) the number of advanced ambulances allocated to stations and the outputs are (i) total cost and (ii) survival rates of types I, II and III. The results obtained from the DEA model are presented in Table 12.

Table 6
Ambulances allocated to the emergency stations.

Service nodes	A	B	C	D	E	F	G	H	I	J	K	L
General type	4	5	3	5	4	3	3	2	2	2	2	2
Advanced type	1	2	1	1	1	0	2	1	1	0	1	1

On the basis of DEA analysis, scenarios 4 and 5 are superior followed by scenario 2. Although the survival rates of scenarios 4 and 5 are not the greatest, DEA compares the scenarios based on both inputs and outputs to find the best choices. Comparing scenario 4 with scenarios 1 and 2 shows that adding the portable station can help to increase the survival rate with incurring less cost. At a glance at the result in Table 12, it can be drawn that patient types I and II have a significant

Table 7
Parameters' distributions for the simulation model.

Receiving phone call	Time for getting initial information	Probability of physician requirement	Time for identifying the level of emergency care	Probability of ambulance requirement	Ambulance allocation time	Service time on the site
Poisson distribution with mean 860 calls per day	Triangular distribution with (1, 2, 3) minutes	Bernoulli distribution probability 75%	Exponential distribution with average 1 min	55% for the general type, 35% for the advanced type, and 10% does not need an ambulance	Exponential distribution with mean 1 min	Exponential distribution with mean 4 min

Table 8
Constant parameters.

Cost of the EMS construction ($)	Purchasing cost of ambulance vehicles ($)		Maximum ambulances		Maximum EMS
	General	Advanced	General	Advanced	
275,000	62,500	87,500	37	12	12

Table 9
Statistical output of initial replications.

Scenario	Average				Standard deviation				Minimum output				Maximum output			
	Survival rate			Cost	Survival rate			Cost	Survival rate			Cost	Survival rate			Cost
	I	II	III		I	II	III		I	II	III		I	II	III	
1	0.123	0.640	0.800	2,000,540	0.011	0.059	0.080	224,338	0.117	0.57	0.79	2,000,003	0.127	0.642	0.806	2,107,854
2	0.121	0.624	0.849	2,320,001	0.012	0.061	0.081	231,109	0.119	0.618	0.848	2,478,945	0.125	0.627	0.855	2,647,542
3	0.130	0.697	.0607	1,628,798	0.017	0.068	0.059	163,594	0.129	0.694	0.605	1,435,379	0.137	0.701	0.702	1,725,386
4	0.131	0.685	0.740	1,943,762	0.028	0.059	0.070	189,341	0.127	0.679	0.738	1,799,856	0.135	0.687	0.747	2,001,360
5	0.135	0.692	0.712	1,822,395	0.011	0.057	0.070	179,835	0.134	0.690	0.711	1,796,853	0.136	0.697	0.714	1,856,347
6	0.124	0.643	0.818	1,998,879	0.013	0.059	0.078	188,576	0.123	0.641	0.817	1.987,874	0.127	0.645	0.825	2,201,347

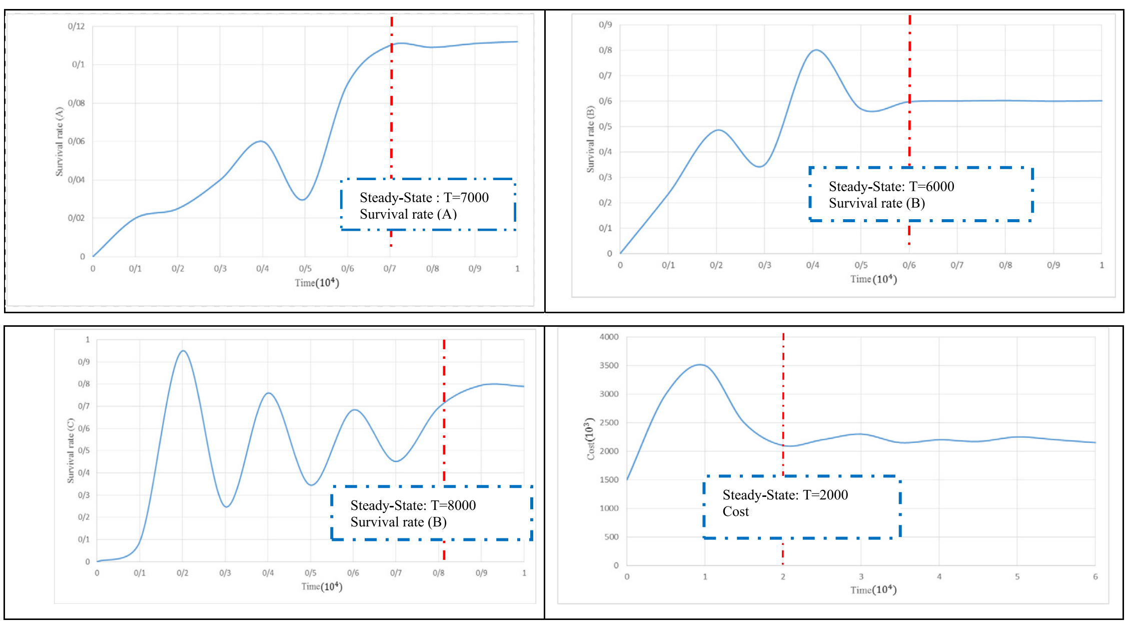

Fig. 5. Steady-State (Warm-up period).

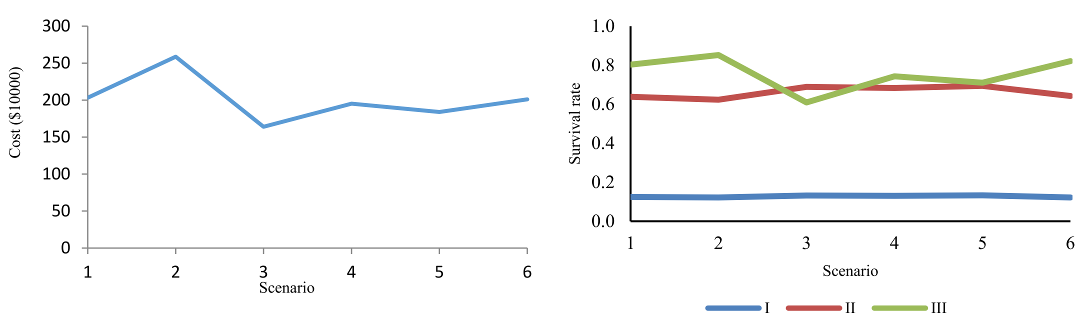

Fig. 6. Survival rates and total cost of each scenario.

Table 10
Required number of replications.

Scenario	Total required replications
1	106
2	112
3	101
4	107
5	103
6	110

Table 11
Results of simulation.

Scenario	Survival rate			Cost ($)	Estimation interval for cost ($)
	I	II	III		
1	0.125	0.638	0.803	2,031,665	2,031,665 ± 41923.22
2	0.122	0.623	0.852	2,587,221	2,587,221 ± 43188.55
3	0.132	0.698	0.609	1,639,999	1,639,999 ± 30571.67
4	0.130	0.683	0.743	1,953,359	1,953,359 ± 35383.14
5	0.133	0.694	0.710	1,839,025	1,839,025 ± 33606.71
6	0.122	0.642	0.821	2,011,628	2,011,628 ± 35240.18

Table 12
DEA's results for 6 scenarios.

Scenarios	DEA scores
1	0.72
2	0.94
3	0.86
4	1
5	1
6	0.88

Table 13
Survival rate of Isfahan municipality regions.

Areas of Isfahan	Survival rate		
	I	II	III
1,3,5 and 6 (scenario 4)	0.130	0.683	0.743
1,3, 5 and 6 (scenario 5)	0.133	0.694	0.710
2	0.121	0639	0.802
4	0.116	0.583	0.765
7	0.110	0.551	0.553
8	0.102	0.621	0.664
9	0.113	0.589	0.816
10	0.117	0.579	0.746
11	0.103	0.563	0.589
12	0.100	0.521	0.486
13	0.109	0.549	0.568
14	0.110	0.572	0.621
15	0.100	0.512	0.439

effect on choosing scenarios 4 and 5. Thus, building the portable station in the regions that have the most call of patient I can lead to the increase in the survival rate. Convincingly, scenario 3 is penultimate in the ranking order while it has the highest survival rate for types I and II and lower cost compared with scenarios 4 and 5.

According to the developed approach in this paper, scenarios 4 and 5 are the best configuration for the EMS system for areas 1, 3, 5 and 6 of Isfahan. Let us compare these findings with the existing survival rates in other areas (see Table 13). Clearly, the survival rates of types I and II for areas 1, 3, 5 and 6 are higher than other areas at the minimum cost and the survival rate of types III is acceptable.

5. Conclusion

This study integrates the simulation and optimization methods to find the best EMS configuration and ambulances allocation with the aim of increasing survival rate and decreasing cost across the EMS system. It is difficult to embed dynamic situations such as traffic conditions in mathematical models when analyzing EMS systems. Hence, this paper leverages the simulation method to solve the problem across different scenarios.

The simulation-based optimization model was implemented in four selected municipal regions of Isfahan to obtain a proper design for emergency center locations and ambulances allocation. In this regard, six scenarios were defined to simulate the model in a dynamic environment and measure the survival rate and the total cost of each scenario. In view of the survival rate and costs, DEA was then used to rank scenarios and select the best ones (scenarios 4 and 5). The chosen scenarios show that patient types I and II play a crucial role in increasing the survival rate and it is essential to be regarded in designing EMS facilities, which can help to improve the survival rate.

For future research, the proposed approach can be extended by considering some other key factors such as seasonal variations and weather affecting the travel time of ambulances. With the growing applications of machine learning techniques, this method could also be studied in combination with on-going research on identifying the location of expected emergency events through techniques such as artificial neural networks and decision models for location–allocation. Additionally, it could be a worthwhile study to apply the proposed method in this paper to similar problems in other emergency response systems such as urban firefighting facilities systems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Department of Health and Children (Ireland), The Emergency Ambulance Services, Report on Value for Money Examination, The Stationery Office, 1997.
[2] C. White, Health care spending growth: how different is the United States from the rest of the OECD, Health Affairs 26 (1) (2007) 154–161.
[3] C. Toregas, R. Swain, C. Revelle, L. Bergman, The location of emergency service facilities, Oper. Res. 19 (3) (1971) 1363–1373.
[4] R. Church, C. ReVelle, The maximal covering location problem, Pap. Reg. Sci. Assoc. 32 (1) (1974) 101–118.
[5] M. Gendreau, G. aporte, F. Semet, The maximal expected coverage relocation problem for emergency vehicles, J. Oper. Res. Soc. 57 (1) (2006) 22–28.
[6] X. Li, Z. Zhao, X. Zhu, T. Wyatt, Covering models and optimization techniques for emergency response facility location and planning: A review, Math. Methods Oper. Res. 74 (3) (2011) 281–310.
[7] J. Goldberg, Operations research models for the deployment of emergency services vehicles, J. EMS Manag. 1 (5) (2004) 20–39.
[8] R. Galvao, R. Morabito, Emergency service systems: The use of the hypercube queueing model in the solution of probabilistic location problems, Int. Trans. Oper. Res. 15 (5) (2008) 525–549.
[9] A. Ingolfsson, S. Budge, E. Erkut, Optimal ambulance location with random delays and travel times, Health Care Manag. Sci. 11 (3) (2008) 262–274.
[10] M.M. Gunal, M. Pidd, Discrete event simulation for the performance modelling in health care: A review of the literature, J. Simul. 4 (1) (2010) 42–51.
[11] R. McCormack, R. Coates, A simulation model to enable the optimization of ambulance fleet allocation and base station location for increased patient survival, European J. Oper. Res. 247 (5) (2015) 294–309.
[12] H. Toro-Díaz, M.E. Mayorga, S. Chanta, L.A. McLay, Joint location and dispatching decisions for emergency medical services, Comput. Ind. Eng. 64 (4) (2013) 917–928.
[13] L. Brotcorne, Laporte. G., F. Semet, Ambulance location and relocation models, European J. Oper. Res. 147 (4) (2003) 451–463.
[14] M. Daskin, A maximum expected covering location model: Formulation, properties and heuristic solution, Transp. Sci. 17 (2) (1983) 48–70.
[15] C. ReVelle, K. Hogan, The maximum availability location problem, Transp. Sci. 23 (3) (1989) 192–195.
[16] M. Gendreau, G. Laporte, F. Semet, A dynamic model and parallel Tabu search heuristic for real-time ambulance relocation, Parallel Comput. 27 (3) (2001) 1641–1653.
[17] H.K. Rajagopalan, C. Saydam, J. Xiao, A multiperiod set covering location model for dynamic redeployment of ambulances, Comput. Oper. Res. 35 (3) (2008) 814–826.

[18] H. Toro-Díaz, M.E. Mayorga, L.A. McLay, H.K. Rajagopalan, C. Saydam, Reducing disparities in large-scale emergency medical service systems, J. Oper. Res. Soc. 66 (7) (2015) 1169–1181.

[19] H. Rajagopalan, C. Saydam, A minimum expected response model: Formulation, heuristic solution, and application, Socio-Econ. Plan. Sci. 43 (3) (2009) 253–262.

[20] V. Schmid, K. Doerner, Ambulance location and relocation problems with time-dependent travel times, European J. Oper. Res. 90 (2) (2010) 580–595.

[21] C. Saydam, J. Repede, T. Burwell, Accurate estimation of expected coverage: a comparative study, Socio-Econ. Plan. Sci. 28 (2) (1994) 113–120.

[22] M. Gendreau, G. Laporte, F. Semet, Solving an ambulance location model by tabu search, Locat. Sci. 5 (2) (1997) 75–88.

[23] M. Gendreau, F. Guertin, J.Y. Potvin, E. Taillard, Parallel tabu search for real-time vehicle routing and dispatching, Transp. Sci. 33 (4) (1999) 381–390.

[24] B. Sah, R. Titiyal, S. Sonia, A goal programming and simulation based study for overall process improvement in an Indian hospital, Int. J. Serv. Oper. Manag. 27 (4) (2017) 439–456.

[25] L. Aboueljinane, E. Sahin, Z. Jemai, A review on simulation models applied to emergency medical service operations, Comput. Ind. Eng. 66 (3) (2013) 734–750.

[26] S. Harewood, Emergency ambulance deployment in Barbados, J. Oper. Res. Soc. 53 (2) (2002) 185–192.

[27] J. Goldberg, R. Dietrich, J. Chen, G. Mitwasi, T. Valenzuela, L. Criss, A simulation model for evaluating a set of emergency vehicle locations: development, validation, and usage, Socio-Econ. Plan. Sci. 24 (2) (1990) 125–141.

[28] k Peleg, J.S. Pliskin, A geographic information system simulation model of EMS: reducing ambulance response time, Am. J. Med. 22 (3) (2004) 164–170.

[29] Y. Wang, K.L. Luangkesom, L. Shuman, Modelling emergency medical response to a mass casualty incident using agent based simulation, Socio Econ. Plan. Sci. 46 (4) (2012) 281–290.

[30] L. Nogueira, L. Pinto, P. Silva, Reducing emergency medical service response time via the reallocation of ambulance bases, Health Care Manag. Sci. 58 (3) (2014) 511–551.

[31] T.C. Van Barneveld, S. Bhulai, R.D. Van der Mei, The effect of ambulance relocations on the performance of ambulance service providers, European J. Oper. Res. 252 (3) (2015) 257–269.

[32] M.A. Zaffar, H.K. Rajagopalan, C. Saydam, M. Mayorga, Coverage, survivability or response time: A comparative study of performance statistics used in ambulance location models via simulation–Optimization, Oper. Res. Health Care 11 (1) (2015) 1–12.

[33] R. Fritze, A. Graser, M. Sinnl, Combining spatial information and optimization for locating emergency medical service stations: A case study for lower Austria, Int. J. Med. Inform. 111 (2018) 24–36.

[34] H. Andersson, T.A. Granberg, M. Christiansen, E.S. Aartun, H. Leknes, Using optimization to provide decision support for strategic emergency medical service planning–three case studies, Int. J. Med. Inform. 133 (2020) 103975.

[35] H. Leknes, E.S. Aartun, H. Andersson, M. Christiansen, T.A. Granberg, Strategic ambulance location for heterogeneous regions, European J. Oper. Res. 260 (1) (2017) 122–133.

[36] T. Unluyurt, Y. Tuncer, Estimating the performance of emergency medical service location models via discrete event simulation, Comput. Ind. Eng. 102 (5) (2016) 467–475.

[37] R. Aringhieri, V. Knight, H. Smith, Or applied to health in a modern world, Health Syst. 5 (3) (2016) 163–165.

[38] A. Ahmadi-Javid, Z. Jalali, K.J. Klassen, Outpatient appointment systems in healthcare: A review of optimization studies, European J. Oper. Res. 258 (1) (2017) 3–34.

[39] M.D. Oliveira, G. Bevan, Modelling the redistribution of hospital supply to achieve equity taking account of patient's behaviour, Health Care Manag. Sci. 9 (1) (2006) 19–30.

[40] E. Erkut, A. Ingolfsson, G. Erdoğan, Ambulance location for maximum survival, Nav. Res. Logist. 55 (1) (2008) 42–58.

[41] S.S. Syam, .M.J. Côté, A location–allocation model for service providers with application to not-for-profit health care organizations, Omega 38 (3) (2010) 157–166.

[42] A.P. Iannoni, R. Morabito, C. Saydam, An optimization approach for ambulance location and the districting of the response segments on highways, European J. Oper. Res. 195 (2) (2009) 528–542.

[43] S.C. Brooks, J.H. Hsu, S.K. Tang, R. Jeyakumar, T.C. Chan, Determining risk for out of-hospital cardiac arrest by location type in a Canadian urban setting to guide future public access defibrillator placement, Ann. Emerg. Med. 61 (5) (2013) 530–5388.

[44] N. Bricon, F. Anceaux, N. Bennani, E. Dufresne, L. Watbled, A distributed coordination platform for home care: analysis, framework and prototype, Int. J. Med. Inform. 74 (10) (2005) 809-S25.

[45] K. Liu, Q. Li, Z. Zhang, Distributionally robust optimization of an emergency medical service station location and sizing problem with joint chance constraints, Transp. Res. B 119 (2019) 79–101.

[46] R. Boujemaa, A. Jebali, S. Hammami, A. Ruiz, Multi-period stochastic programming models for two-tiered emergency medical service system, Comput. Oper. Res. 123 (2020).

[47] V. Bélanger, A. Ruiz, P. Soriano, Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles, Eur. J. Oper. Res. 272 (1) (2018) 1–23.

[48] R. Galvao, C. ReVelle, A Lagrangean heuristic for the maximal covering location problem, European J. Oper. Res. 88 (1) (1996) 114–123.

[49] Y. Liu, M.R. Arash, Z. Li, K. Kepaptsoglou, H. Patel, X. Lu, Heuristic approach for optimizing emergency medical services in road safety within large urban networks, J. Transp. Eng. 140 (9) (2014) 14–43.

[50] L.A. McLay, A maximum expected covering location model with two types of servers, IIE Trans. 41 (8) (2009) 730–741.

[51] P. Beraldi, M.E. Bruni, D. Conforti, Designing robust emergency medical service via stochastic programming, European J. Oper. Res. 158 (1) (2004) 183–193.

[52] L.A. McLay, M.E. Mayorga, Evaluating emergency medical service performance measures, Health Care Manag. Sci. 13 (2) (2010) 124–136.

[53] S. Chanta, M.E. Mayorga, M.E. Kurz, L.A. Mclay, The minimum p-envy location problem: a new model for equitable distribution of emergency resources, IIE Trans. Healthc. Syst. Eng. 1 (2) (2011) 101–115.

[54] A. Basar, B. Catay, T. Uluyurt, A multi period double coverage approach for locating the emergency medical service stations in Istanbul, J. Oper. Res. Soc. 62 (4) (2011) 627–637.

[55] M.J. Farrell, The measurement of productive efficiency, J. R. Stat. Soc. 120 (1) (1957) 253–290.

[56] A. Charnes, W.W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European J. Oper. Res. 2 (6) (1978) 429–444.

[57] A. Emrouznejad, B.R. Parker, G. Tavares, Evaluation of research inefficiency and productivity: a survey and analysis of the first 30 years of scholarly literature in DEA, Socio-Econ. Plan Sci. 42 (3) (2008) 151–157.

[58] A. Hatami Marbini, A. Emrouznejad, M. Tavana, A taxonomy and review of the fuzzy data envelopment analysis literature, European J. Oper. Res. 214 (3) (2011) 457–472.

[59] W.D. Cook, L.M. Seiford, Data envelopment analysis (DEA) – Thirty years on, European J. Oper. Res. 19 (1) (2009) 1–17.

[60] A. Athanassopoulos, G. Chrysostomos, Assessing the technical and allocative efficiency of hospital operations in Greece and its resource allocation implications, European J. Oper. Res. 133 (2) (2001) 416–431.

[61] M.N. Coppola, Y.A. Ozcan, R. Bogacki, Evaluation of performance of dental providers on posterior restorations: does experience matter? A data envelopment analysis (DEA) approach, J. Med. Syst. 27 (5) (2003) 445–456.

[62] J.L. Miller, E.E. Adam, Slack and performance in health care delivery, Int. J. Qual. Reliab. Manag. (1996).

[63] J.A. Chilingerian, Evaluating physician efficiency in hospitals: A multivariate analysis of best practice, European J. Oper. Res. 80 (3) (1995) 548–574.

[64] B. Hollingsworth, The measurement of efficiency and productivity of health care delivery, Health Econ. 17 (10) (2008) 1107–1128.

[65] G.D. Ferrier, M.D. Rosko, V.G. Valdmanis, Analysis of uncompensated hospital care using a DEA model of output congestion, Health Care Manag. Sci. 9 (2) (2006) 181–188.

[66] p. Nayar, Y.A. Ozcan, Data envelopment analysis comparison of hospital efficiency and quality, J. Med. Syst. 32 (3) (2008) 193–199.

[67] H. Golabian, J. Arkat, R. Tavakkoli-Moghaddam, H. Faroughi, A multi-verse optimizer algorithm for ambulance repositioning in emergency medical service systems, J. Ambient Intell. Humaniz. Comput. 13 (1) (2022) 549–570.

[68] G. Grekousis, Y. Liu, Where will the next emergency event occur? Predicting ambulance demand in emergency medical services using artificial intelligence, Comput. Environ. Urban Syst. 76 (2019) 110–122.

[69] M. Karatas, E. Yakıcı, A multi-objective location analytics model for temporary emergency service center location decisions in disasters, Decis. Anal. J. 1 (2021) 100004.

[70] H. Eilat, B. Golany, A. Shtub, R & D project evaluation: An integrated DEA and balanced scorecard approach, Omega 36 (5) (2008) 895–912.

[71] B. Hollingsworth, P.J. Dawson, N. Maniadakis, Efficiency measurement of health care: a review of non-parametric methods and applications, Health Care Manag. Sci. 2 (3) (1999) 161–172.

[72] J.F. Burgess, P.W. Wilson, Decomposing hospital productivity changes 1985–1988 a nonparametric malmquist approach, J. Product. Anal. 6 (4) (1995) 343–363.

[73] R. Färe, S. Grosskopf, B. Lindgren, P. Roos, Productivity changes in Swedish pharmacies 1980–1989: A non-parametric malmquist approach, J. Product. Anal. 3 (2) (1992) 85–101.

[74] B. Hollingsworth, Non-parametric and parametric applications measuring efficiency in health care, Health Care Manag. Sci. 6 (4) (2003) 203–218.

[75] J.A. Chilingerian, H.D. Sherman, Health care applications, in: W.W. Cooper, L.M. Seiford, J. Zhu (Eds.), Handbook on Data Envelopment Analysis, in: International Series in Operations Research & Management Science, Vol. 71, Springer, Boston, MA, 2004.

[76] S. Nickel, M. Reuter-Oppermann, F. Saldanha-da Gama, Ambulance location under stochastic demand: A sampling approach, Oper. Res. Health Care 8 (1) (2016) 24–32.

[77] R. Boujemaa, A. Jebali, S. Hammami, A. Ruiz, H. Bouchriha, A stochastic approach for designing two-tiered emergency medical service systems, Flexible Serv. Manuf. J. 30 (1–2) (2018) 1–30.

[78] V. Knight, P. Harper, L. Smith, Ambulance allocation for maximal survival with heterogeneous outcome measure, Omega 40 (6) (2012) 918–926.

[79] R. Banker, A. Charnes, W. Cooper, Some models for the estimation of technical and scale inefficiencies in data envelopment analysis, Manage. Sci. 30 (9) (1984) 1078–1092.

[80] S. Robinson, Discrete-event simulation: from the pioneers to the present, what the next, J. Oper. Res. Soc. 56 (3) (2005) 619–629.

[81] H. Morohosi, T. Furuta, A simulation model to enable the optimization of ambulance fleet allocation and base station location for increased patient survival, European J. Oper. Res. 247 (5) (2013) 294–309.

[82] R. Aiassi, S.M. Sajadi, S.M. Hadji-Molana, A. Zamani-Babgohari, Designing a stochastic multi-objective simulation-based optimization model for sales and operations planning in built-to-order environment with uncertain distant outsourcing, Simul. Model. Pract. Theory 104 (2020) 102103.

[83] Y. Carson, A. Maria, Simulation optimization: methods and applications, in: Proceedings of the 29th Conference on Winter Simulation, 1997, pp. 118–126.

[84] A. Hatami-Marbini, S.M. Sajadi, H. Malekpour, Optimal control and simulation for production planning of network failure-prone manufacturing systems with perishable goods, Comput. Ind. Eng. 146 (2020) 106614.