European Journal of Operational Research 198 (2009) 936–942

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Simulation optimization for an emergency department healthcare unit in Kuwait

Mohamed A. Ahmed *, Talal M. Alkhamis

Department of Statistics and Operations Research, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait

ARTICLE INFO

Article history:
Received 8 February 2007
Accepted 27 October 2008
Available online 6 November 2008

Keywords:
Health services
Stochastic optimization
Simulation
Decision support system

ABSTRACT

This paper integrates simulation with optimization to design a decision support tool for the operation of an emergency department unit at a governmental hospital in Kuwait. The hospital provides a set of services for different categories of patients. We present a methodology that uses system simulation combined with optimization to determine the optimal number of doctors, lab technicians and nurses required to maximize patient throughput and to reduce patient time in the system subject to budget restrictions. The major objective of this decision supporting tool is to evaluate the impact of various staffing levels on service efficiency. Experimental results show that by using current hospital resources, the optimization simulation model generates optimal staffing allocation that would allow 28% increase in patient throughput and an average of 40% reduction in patients' waiting time.

1. Introduction

Busy healthcare systems constantly provide new challenges to their managers and decision-makers due to high demands for service, high costs, limited budget, and healthcare resources. As a result, decision-makers are continuously studying efficacy and efficiency of existing healthcare systems, and must be able to evaluate the outcomes of any changes they make to these systems. In this paper, we use simulation and optimization techniques to address the described management problems at a complex healthcare system. Health system managers must often maximize the utilization of their available resources while being constrained by specific budget limits. In order to do this, managers must implement highly efficient systems that minimize their costs of providing a specified level of care. In addition, due to humanistic and governmental obligations at public healthcare systems, these systems are required to provide quality care as quickly as possible. That is, the highest number of patients must be cared for adequately in a given time period to minimize waiting time and increase patient satisfaction. This may antagonize efficiency restrictions from a manager's standpoint, but is equally important in designing ideal healthcare systems.

This paper addresses the work of the authors with a simulation optimization model of an emergency department unit at a government hospital in Kuwait. The objective is to analyze patient flow throughout and evaluate the impact of alternative staffing distribution. Public healthcare in the state of Kuwait is maintained by a network of primary and secondary health centers, general and specialized hospitals, and research institutions. 76 primary healthcare centers are distributed throughout the country, each offering services which include general practitioner services, family medicine, obstetrics/gynecology, pediatric care, diabetes care, dentistry, preventive medical care, and pharmaceuticals. Six major hospitals cover secondary healthcare services in addition to providing primary health care services. Each of these hospitals is structured as a general hospital, a healthcare center, specialized clinics and dispensaries. Outpatient services and a 24-hour emergency service are provided by each general hospital. Kuwait also has several specialty hospitals, covering a range of specializations from cardiothoracic diseases to neurosurgery and oncology.

This paper uses optimization and simulation to evaluate alternative choices for decision-makers through "what-if" models. In a healthcare system model, such as in an emergency department, different categories of users may require multiple services through a common sequence. All patients within the same category undergo the same sequence of general services, but two patients from different categories may intersect at only some common services. For example, all patients who voluntarily walk into the emergency department must first see the receptionist as a required service for walk-in patients before proceeding to the examination room. On the other hand, all ambulance patients bypass service at the receptionist and are directed immediately to the examination room. The receptionist service here is not part of the care sequence for ambulance patients, while the examination room is a common service sequence for all categories. Therefore, each service center has its own costs depending on the service providers (number of doctors, lab technicians, nurses, etc.), the resources needed (number of beds, X-ray machines, etc.), and the patient demand for that center. Also, the amount of time patients must wait in between each

* Corresponding author. Tel.: +965 601 2051.
E-mail addresses: wakeel@kuc01.kuniv.edu.kw (M.A. Ahmed), alkhamis@kuc01.kuniv.edu.kw (T.M. Alkhamis).

0377-2217/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2008.10.025

service is a measure of the efficiency of the entire system. This paper tackles the problem of how to choose the configuration of the servers (number of servers of each type in each service) in order to optimize a measure of performance selected by a decision-maker within the constraints imposed by the system limitations. Using current hospital resources, the optimization simulation model presented in this paper generates optimal staffing allocation that would allow 28% increase in patient throughput and an average of 40% reduction in patients' waiting time.

Over the last decade, there have been fruitful efforts in developing simulation/optimization models for solving healthcare management problems. Côte (1999), Ferreira de Oliveira (1999), Swisher et al. (2001), Blasak et al. (2003), and Sinreich and Marmor (2005) use simulation models to reproduce the behaviour of a healthcare system in order to evaluate its performance and analyze the outcome of different scenarios. In these studies, the main objectives are to show decision-makers a realistic reproduction of the healthcare system at work. Optimization techniques have also been used as solution methods for healthcare management problems. Beaulieu et al. (2000) use a mathematical programming approach for scheduling doctors in the emergency room. Flessa (2000) uses a linear programming approach for the optimal allocation of healthcare resources in developing *countries*. Jacobson et al. (1999) use an integer programming model for vaccine procurement and delivery for childhood immunization. De Angelis et al. (2003) present a methodology that interactively uses system simulation, estimation of target function and optimization to calculate and validate the optimal configuration of servers in a transfusion center. Instead of using mathematical models that attempt to explicitly represent the functioning of the system, resulting in large linear and integer models with many variables and constraints, De Angelis et al. (2003) adopt a simpler model, where the complexity is captured by a non-linear function estimated from simulated data. Baesler and Sepúlveda (2001) present a multi-objective simulation optimization model for a cancer treatment center. Baesler et al. (2003) present a simulation model combined with a design of experiments for estimating maximum capacity in an emergency room.

As mentioned previously, simulation/optimization have been considered by several authors for the type of application dealt with in this paper. The method that we propose differs from others in the way it uses these techniques. Instead of dealing with an approximated mathematical model of the system, we solve the actual system by combining simulation with optimization. Our optimization model involves a complex stochastic objective function subject to a deterministic and stochastic set of constraints.

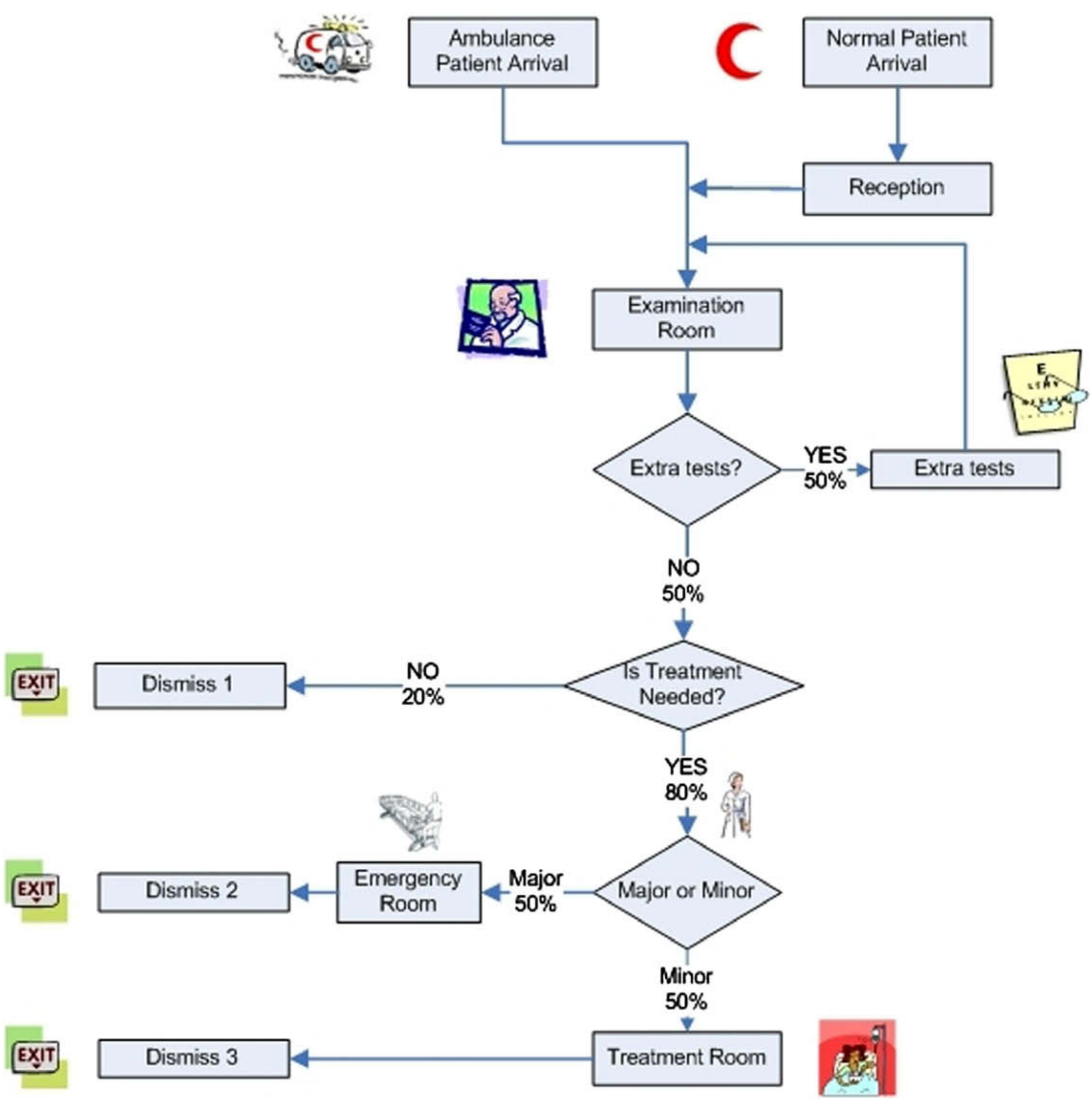

Fig. 1. Emergency department high-level process view.

Both the stochastic objective function and the stochastic set of constraints have no analytical form and can be evaluated through simulation only. We have employed special techniques that utilize new developments in the area of simulation optimization in order to search for the optimal solution as we will explain in Section 4. Moreover, the solution approach adopted results in a flexible tool that the managers can use to evaluate different configurations and different operating rules of the system.

The paper is organized as follows: Section 2 gives a brief description of the emergency department unit for the hospital. In Section 3, the basic elements of system simulation will be described. Problem formulation and system optimization will be discussed in Section 4. Section 5 presents some computational results. Section 6 includes some concluding remarks and future work.

2. System description

The emergency department (ED) is open 24 hours a day and receives an average of 145 patients daily. Besides its internal capacity, the emergency department shares resources with other hospital services, such as X-rays, scanners, clinical laboratories and pharmacy. Patients arriving at the emergency department follow a process as depicted in Fig. 1. The process begins when a patient arrives through the doors of the ED and ends when a patient is either released from the ED or admitted into the hospital for further treatment. The arriving patient goes through the receptionist who collects the patient's personal information and locates his/her file. Following this, the patient waits for availability of an examination room. The acuity of the patient's illness is assessed by a doctor in the examination room. Also in the examination room, doctors will decide if the patient needs further tests such as X-rays and clinical lab tests, performed by a patient care lab technician. Patients are classified as critical (category 1) and non-critical (categories 2 and 3) according to their conditions. After an assessment is performed by the doctor, the non-critical patients are further classified into two categories. Category 2 patients are asked to wait for a minor treatment, which is performed by a nurse in the treatment room. Category 3 patients receive their medication and are released from the hospital. Each critical patient is assigned to a bed in the emergency room where he/she receives complete treatment and stays under close observation. The treatment services in the emergency room are provided by a nurse and a doctor; the doctor is called from the examination room when needed. Finally, critical patients are either released or admitted into the hospital for further treatment. Patients who arrive at the hospital in an ambulance are considered critical patients (category 1) and are rushed immediately to the emergency room. It is observed that 88% of all patients are released from the emergency room, while the remaining 12% are admitted into the hospital for further treatment.

The emergency department has the following resources:

1. Receptionists (denoted by x_1),
2. Doctors (denoted by x_2),
3. Lab technicians (denoted by x_3),
4. Treatment room nurses (denoted by x_4),
5. Emergency room nurses (denoted by x_5).

Due to cost and layout considerations, hospital administrators have determined that the staffing level must not exceed three receptionists, four doctors, five lab technicians, six treatment room nurses and 12 emergency room nurses. The hospital would like to find the configuration of the above resources that maximize patient throughput (patient dismissed per unit time) subject to budget constraint and a constraint imposed on the average waiting time in the system for patients of category 1.

3. Model simulation

A comprehensive survey at the emergency department has been carried out in order to collect data on the arrival process, the service times at the examination room, the service times at the treatment room and the total turnaround time in the emergency department. After observing the process for three weeks and after collecting additional data from interviewed doctors, nurses and hospital personnel in charge of each of these activities, the results of these interviews were used to determine the best theoretical distribution to represent each stage of the process under study. The arrival process follows a non-homogenous Poisson process with rate $\lambda(t)$ ($\lambda(t)$ is the arrival rate of patients to the ED at time t).

Table 1 gives the distributions of the service times at each stage of the process. The overall process of the emergency department has been modeled by a discrete event simulation system. The system is described by

- The arrival process to the receptionist as a non-homogenous Poisson process with an estimate of $\lambda(t)$ given in Fig. 2,
- The arrival process to the examination room or X-ray room from outside the system as a Poisson process with rate 2 per hour,
- The routing probabilities of patients at each stage inside the system,
- The number of servers at each stage, defining the configuration of the system,
- The type of service time distributions and the value of their parameters.

4. The optimization model

The optimization problem considered in this paper aims to maximize the throughput (patient dismissed per unit time) subject to a budget constraint, a constraint imposed on the average waiting time in the system for patients of category 1, upper and lower

Table 1
Service time distributions at each stage of the process.

Stage	Distribution (minutes)
Reception	Uniform (5,10)
Lab tests	Triangular (10,20,30)
Examination room	Uniform (10,20)
Reexamination process	Uniform (7,12)
Treatment room	Uniform (20,30)
Emergency room	Uniform(60,120)

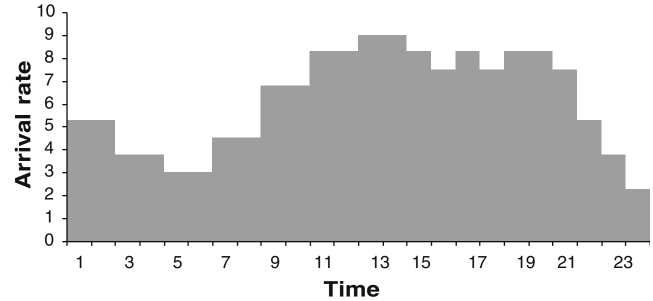

Fig. 2. Plot of the estimated rate function $\hat{\lambda}(t)$ in patients per hour for the arrival process.

bounds of the number of servers, and integrality conditions. Thus, the optimization problem can be represented mathematically as:

Problem A-1: (Maximize throughput)

$$\begin{aligned} \text{Max} \quad & f(x_1,x_2,x_3,x_4,x_5), && (1)\\ \text{s.t.} \quad & f_1(x_1,x_2,x_3,x_4,x_5) \leqslant B, && (2)\\ & f_2(x_1,x_2,x_3,x_4,x_5) \leqslant Q_1, && (3)\\ & L_i \leqslant x_i \leqslant U_i i = 1,\ldots,5, && (4)\\ & x_i \text{ integer } i = 1,\ldots,5, && (5) \end{aligned} \tag{A-1}$$

where f represents the throughput objective function value (number of dismissed patients from the system per hour), which is a function of the decision variables x_1, $x_2, \ldots,$ x_5. Function f_1 represents the budget constraint (deterministic) and function f_2 represents the constraint on the average waiting time in the system for patients of category 1 (stochastic). Function f and function f_2 are both stochastic functions that have no analytical form and can be evaluated only through simulation. Parameters B and Q_1 represent pre-specified requirement values for budget and average waiting time in the system for patients of category 1. Both parameters B and Q_1 must be specified before solving the problem and they represent management limitation imposed on the system. Due to cost and layout considerations, hospital administrators have determined that the staffing level must not exceed three receptionists, four doctors, five lab technicians, six treatment room nurses and 12 emergency room nurses. The hospital management would like to find the configuration of the above resources that maximizes patient throughput. Using complete enumeration, we would have to perform $3 \times 4 \times 5 \times 6 \times 12 = 4320$ experiments. If we use a sample size of, say, at least 50 replications per trial solution, then each experiment would take about 0.5 minute. This means that a complete enumeration of all possible solutions would take approximately 2160 minutes, or about 1.5 working days. This is obviously too long duration for finding a solution. In this paper, we develop a procedure based on simulation optimization to solve problem (A-1).

Simulation optimization is the practice of combining a simulation model with an optimization algorithm to obtain the optimal value of design parameters in order to maximize the performance of the simulated system. Problem (A-1) is a discrete stochastic optimization problem with two deterministic constraints (constraints (2) and (4)) and one stochastic constraint (constraint (3)), where the objective function is to be optimized over a non-empty, discrete, finite set of solution space. There is a considerable amount of literature in developing procedures for discrete simulation optimization. For more details on those procedures, see Yan and Mukai (1992), Andradottir (1995, 1996), Alrefaei and Andradottir (1999), Ahmed et al. (1998), Alkhamis et al. (1999), Ahmed and Alkhamis (2002), Alkhamis and Ahmed (2004), Kim and Nelson (2001), and Boesel et al. (2003). However, a limitation of those procedures lies in the fact that they are all suitable for optimizing unconstrained stochastic systems. In reality, we often face stochastic constraints that can only be analyzed through simulation. Discrete simulation optimization with stochastic constraints is currently a relatively underdeveloped area. However, a small amount of literature is available. See Butler et al. (2001), Andradottir et al. (2005), and Batur and Kim (2005).

To solve problem (A-1), we adopt a two-phase approach. Phase I finds a set, S, that contains only feasible or near-feasible solutions and then Phase II chooses the best among those solutions. In Phase I, we use a feasibility detection procedure which is composed of two steps. *Step 1* eliminates all solutions not satisfying deterministic constraint (2) and (4) of problem (A-1). *Step 2* refines the set obtained from *step 1* by eliminating solutions violating the stochastic constraint (3) using the feasibility detection procedure of Andradottir et al. (2005) where we have to specify lower and upper bound values for the parameter Q_1 (upper bound on the average waiting time in the system for patients of category 1), say (Q_{1L}, Q_{1U}) with $Q_{1L} < Q_{1U}$. The procedure of *step 2* defines the following regions for solution point i:

- $f_2(i) \leqslant Q_{1L}$: This is the acceptance region. Any solution point in this range is feasible and accepted.
- $Q_{1L} < f_2(i) < Q_{1U}$: This is the near acceptable region.
- $f_2(i) \geqslant Q_{1U}$: This is the rejection region. Any solution point in this range is infeasible and rejected.

As we mentioned above, at the end of phase I we will have a set that contains all acceptable and near acceptable systems. Accordingly, in phase II we will consider the following optimization problem:

$$\max_{i \in S} f(i), \tag{A-2}$$

where $f(i)$ represents a throughput function given solution point i, which represents the vector of decision variable values, x_1, $x_2 \ldots$, x_5, and S is the set that contains all acceptable and near acceptable points. In the next subsection, we present a simulation optimization approach that solves problem (A-2) and is based on work presented in Ahmed et al. (1998).

Phase II: Optimization algorithm for problem (A-2)

Our objective function, $f(i)$, includes a measure of performance that can only be evaluated through Monte Carlo simulation. Therefore, in this case, one usually uses the sample objective function, denoted by $L(i)$, as an estimate and requires $f(i) = E[L(i)]$, where $L(i) = f(i) + \tau_i$ and τ_i are independent random variables with mean zero and variance $\sigma_i^2 \tau_i$ can be considered as the noise that results from a simulation run). Thus, the optimization problem becomes

$$\text{Max}\{f(i) = E[L(i)] | i \in S\}$$

and our objective is to seek the global optimal solution $i^* \in S$, where

$$i^* = \{i \in S | f(i) = E[L(i) \geqslant E[L(j)] = f(j), \forall j \in S\}.$$

The optimization algorithm we propose to solve (A-2) uses a binary hypothesis test at each iteration. At iteration k, let i denote the current solution point. A new solution point j is selected randomly and uniformly from $S\backslash\{i\}$ (set S excluding point i). The objective is to select the best solution point, that is the point with the optimal objective function value. We might formulate this selection problem as a hypothesis testing procedure, where we want to discriminate among the following hypotheses:

$$H_0 : E[L(i)] > E[L(j)],$$
$$H_1 : E[L(i)] < E[L(j)].$$

We propose a sequential procedure for testing whether the new solution point, j, is better than the current solution point i. For each $i,j \in S$ generate sample pairs sequentially, say $\{L^1(i), L^1(j)\}$, $\{L^2(i), L^2(j)\}, \ldots$ and let χ_m denote the sign of $\{L^m(i), L^m(j)\}$ such that

$$\chi_m = \begin{cases} +1 & \text{if } L^m(i) < L^m(j), \\ 0 & \text{if } L^m(i) = L^m(j), \\ -1 & \text{if } L^m(i) > L^m(j). \end{cases}$$

Let $T_n(k) = \sum_{m=1}^{n} \chi_m$ denote our test statistic. At iteration k of the algorithm, we compare the current solution point, i, with a new solution point, j, by testing the above hypotheses using the test statistics $T_n(k)$ and decision rules as follows (assume that a_k is a given bound):

$$\begin{cases} \text{reject } H_0 & \text{if } T_n(k) \geqslant +a_k, \\ \text{accept } H_0 & \text{if } T_n(k) \leqslant -a_k, \\ \text{continue sampling} & \text{if } -a_k < T_n(k) < +a_k. \end{cases}$$

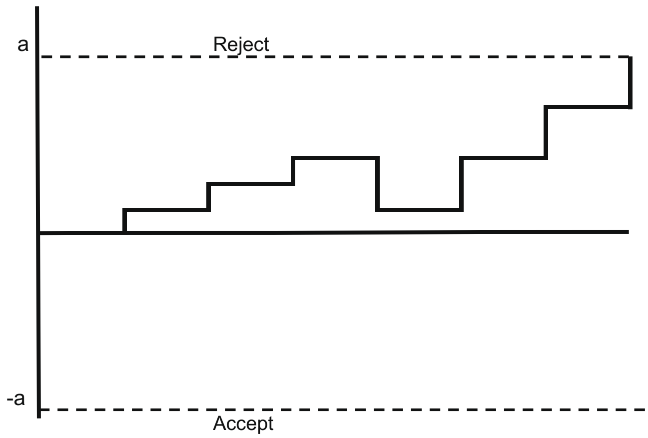

Fig. 3. A typical sample path of the random walk T_n.

The comparison procedure is conducted as follows: Run an experiment or a simulation to obtain samples from both i and j until the first time that $T_n\,(k)$ falls outside a given bound $\pm a_k$ If it falls out below $-a_k$ (i.e. i is better than j), the alternative point is rejected, whereas if it falls out above $+a_k$(i.e. j is better than i), the alternative point is accepted.

The process $\{T_n(k), n = 1, 2, \ldots\}, T_n(0) = 0$ is a random walk that moves a unit step in the positive direction with probability $p = P\{L(i) < L(j)\}$, moves a unit step in the negative direction with probability $q = P\{L(i) > L(j)\}$, does not move with probability $r = 1 - p - q$, and absorbs barriers at $\pm a_k$ (Fig. 3).

Define $\{I_k, k = 0, 1, \ldots\}$ to be the stochastic process where $I_k \in S$ is the current state of the search process at iteration k. The details of the algorithm for finding the optimal state $i \in S$ with maximum objective function value are as follows:

Step 0: Select a starting point I_0 randomly from S, set $k = 0$ and select initial boundary a_k.
Step 1: Given $I_k = i$, choose a new point J_k uniformly from $S\backslash\{i\}$.
Step 2: Calculate $T_n\,(k)$, continue sampling from I_k and J_k until $T_n\,(k)$ falls outside given bounds $\pm a_k$.
Step 3: Set $I_{k+1} = \begin{cases} J_k & \text{if } T_n(k) \geqslant +a_k, \\ I_k & \text{if } T_n(k) \leqslant -a_k. \end{cases}$
Step 4: Set k = k+1, update a_k. Go to *step 1.*

Remark 1. One of the following stopping criteria may be used for the above algorithm: Stop the search process after performing a predetermined number of iterations or when the current best configuration has not been changed for a predetermined number of iterations.

Remark 2. For convergence analysis of the above algorithm, the reader may refer to Alkhamis and Ahmed (2004).

5. Computational results

The first step in our computational procedure is to evaluate the current staffing level and the associated measures of performance. Our main objective in this step is to see whether the budget utilization is efficient or not. Due to the hospital request of information privacy, especially in budget data details, we have reprocessed the budget data and recoded them in terms of budget units (BU). The cost units for the staff are as follows: 0.4 BU for receptionist, 1.2 BU for doctor, 0.5 BU for lab technician, 0.3 BU for treatment and emergency nurses. The current staffing distribution at the emergency department is as follows: two receptionists, two doctors, three technicians, one treatment nurse and nine emergency nurses with total budget units of 7.7. We ran the system for 100 replications and obtained an average throughput of 4.9 patients per hour with 3.57 hours average waiting time in the system for patients of category 1. We applied our optimization approach using the current budget constraints of 7.7 BU and relaxing constraint (3) to determine the best staffing schedule. As the waiting time for different patient categories are very high, the question regarding staffing distribution efficiency has been raised by hospital managers. The hospital management would like to obtain more details on the effect of the new optimization model on the individual waiting time for all patient categories. Our new approach obtains a different staffing distribution as follows: one receptionist, three doctors, two technicians, two treatment nurses and seven emergency nurses with total budget units of 7.7. Table 2 presents the results of our optimization model along with the current situation. From Table 2, the improvement is obvious in both the overall system throughput and patients' waiting time for all categories. System throughput has improved from 4.9 patients per hour to 6.3 patients per hour (an improvement of 28%). Using the current staffing, the average waiting time in the system for category 1 (emergency patients), category 2 (less severe patients) and category 3 (normal patients) are 3.57 hours, 3.28 hours and 2.83 hours respectively. The average waiting time for category 1 patients has been reduced from 3.57 hours to 2.76 hours (an improvement of 22%) while average waiting time for category 2 patients has been reduced from 3.28 hours to 1.8 hours (an improvement of 45%) and average waiting time for category 3 patients has been reduced from 2.83 hours to 1.33 hours (an improvement of 53%).

We have conducted a number of experiments for problem (A-1) varying the budget parameter, B, with different patients' arrival rate. In Table 3, the optimization results of problem (A-1) are presented. We describe the results of the optimization with respect to the optimal value of f (system throughput) for different values of B (6.7, 8.7, 9.7 and 10.7 of BU) and different values of the arrival rate. For example, with a maximum daily budget of 6.7 and an arrival rate equal to 85% of the current arrival rate, $\lambda(t)$, the optimization model gives an average throughput of 5.03 patients per hour and staff distribution as follows: one receptionist, three doctors, one lab technician, two treatment nurses and five emergency room nurses. The average computer execution time is 5.5 minutes on Intel-based X86 Pentium 4, 2.8 GHz, 512 RAM.

Another optimization situation that is of interest to hospital management is to minimize staffing cost subject to average waiting time constraint for patients of category 1, average waiting time constraint for patients of category 2, integrality conditions and upper and lower bounds of the number of servers. In this case, our optimization model becomes:

Problem (B-1): (Minimize cost)

$$\begin{aligned} \text{Min} \quad & f_1(x_1, x_2, x_3, x_4, x_5), && (1') \\ \text{s.t.} \quad & f_2(x_1, x_2, x_3, x_4, x_5) \leqslant Q_1, && (2') \\ & f_3(x_1, x_2, x_3, x_4, x_5) \leqslant Q_2, && (3') \\ & L_i \leqslant x_i \leqslant \mathrm{U}_i i = 1, \ldots, 5, && (4') \\ & x_i \text{ integer } i = 1, \ldots, 5. && (5') \end{aligned} \qquad \text{(B-1)}$$

Table 2
Comparison between current and optimal staffing distribution.

Staff distribution	Measure of performance			Throughput
	Average waiting time in system			
	Category 1	Category 2	Category 3	
Current	3.56	3.28	2.83	4.9
New approach	2.76	1.8	1.33	6.3
Improvement (%)	22	45	53	28

Table 3
Optimization results for Problem A-1 for different values of budget and arrival rates.

Arrival rate	Available values of B in the budget constraint																							
	6.7						8.7						9.7						10.7					
	x_1^*	x_2^*	x_3^*	x_4^*	X_5^*	f^*	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	f^*	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	f^*	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	f^*
$\lambda(t)$	1	3	1	3	4	2.13	2	3	2	2	9	2.61	2	4	2	3	7	2.69	2	4	4	3	7	2.73
$.5\lambda(t)$	1	2	3	3	5	1.49	2	3	3	4	5	1.49	2	4	3	1	7	1.52	2	4	3	1	7	1.52
$.65\lambda(t)$	1	2	3	3	5	1.75	3	3	3	2	5	1.88	3	3	3	4	5	1.88	3	3	5	4	5	1.90
$0.75\lambda(t)$	2	2	2	2	5	1.87	2	3	2	2	8	2.06	1	4	2	4	7	2.09	4	4	3	2	6	2.15
$.85\lambda(t)$	1	3	1	2	5	2.01	2	3	4	2	6	2.30	1	3	4	2	6	2.30	3	4	2	6	6	2.37
$1.15\lambda(t)$	1	3	1	2	5	2.33	2	3	2	5	6	2.81	2	4	2	2	8	2.97	2	4	2	5	8	3.01
$1.3\lambda(t)$	1	3	2	2	3	2.26	2	3	3	3	6	2.93	2	4	3	2	6	3.12	3	4	3	2	8	3.38

Table 4
Optimization results for Problem (B-1) for different values of maximum average waiting time for patients of category 1 & 2.

Q_2	Q_1																							
	2 hours						2.33 hours						2.67 hours						3 hours					
	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	f_1^*	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	f_1^*	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	f_1^*	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	f_1^*
2 hours	3	4	4	2	9	11.3	2	5	2	3	9	11.4	2	5	1	2	9	10.6	1	3	4	1	9	9
2.33 hours	2	3	2	3	6	8.1	1	3	2	2	5	7.1	1	3	1	2	5	6.6	1	3	1	1	5	6.3
2.67 hours	2	3	2	2	7	8.1	1	2	1	1	4	4.8	1	2	1	2	3	4.8	1	2	1	1	2	4.2

The cost function, f_1, of problem (A-1) becomes the objective function for problem (B-1). Parameters Q_1 and Q_2 represent prespecified requirement values for the average waiting time for patients of categories 1 and 2, respectively. There is a major difference between problem (A-1) and problem (B-1). Problem (A-1) maximizes stochastic function subject to one stochastic constraint and one deterministic constraint, plus the constraints on the upper and lower values of the decision variables. Problem (B-1), on the other hand, minimizes deterministic function subject to two stochastic constraints, plus the constraints on the upper and lower values of the decision variables. Now the feasibility detection procedure of Andradottir et al. (2005), which was implemented for problem (A-1), needs to be extended to cover two stochastic constraints. Batur and Kim (2005) modified the procedure of Andradottir et al. (2005) to handle more than one stochastic constraint. We also utilize a two-phase procedure to solve problem (B-1). In phase I, we apply the feasibility detection approach as explained by Batur and Kim (2005) for more than one stochastic constraint to obtain the set, S, of all possible feasible solutions. If $|S| = 0$, no feasible solution exists. If $|S| = 1$, then this solution is the best feasible solution. If $|S| > 1$, then we will have the following optimization problem:

$$\min_{i \in S} \quad f_1(i). \tag{B-2}$$

Problem (A-2) is different than problem (B-2) in various aspects. Problem (A-2) is a stochastic optimization problem over the set S, while problem (B-2) is a deterministic optimization problem over the set S. Function f(i) in (A-2) is a stochastic function that represents system throughput while function f_1(i) in problem (B-2) is a deterministic function that represents system cost function. Problem (A-2) requires a special solution procedure as the one suggested in Section 4, while problem (B-2) is a normal integer programming problem, which can be solved by many computer packages. For problem (B-2), we have used the "solver optimization" add-on feature of the Excel software.

We have conducted a number of experiments for problem (B-1) varying the waiting time constraint parameters, Q_1 and Q_2. Table 4 presents the optimization results of problem (B-1) for different values of Q_1 (2, 2.33, 2.67, 3 hours) and Q_2 (2, 2.33, 2.67 hours). For example, with a maximum average waiting time for patients of category 1 of 3 hours and a maximum average waiting time for patients of category 2 of 2.33 hours, the optimization model gives a solution that costs 6.3 budget units and staff distribution as follows: one receptionist, three doctors, one lab technician, one treatment nurse and five emergency room nurses.

Despite the complex and technical elements of simulation and optimization that comprise the application in this paper, the discussed methodology is user-friendly and can be easily operated to produce results for analysis and interpretation by hospital decision-makers who have even minimal operations research experience. For example, if a hospital manager needed to keep the budget B or the waiting time for category 1 patients Q_1 or the waiting time for category 2 patients Q_2 below a certain level, this would impact the given values of the arrival rates as displayed in Tables 3. The manager can decide whether it is more beneficial to increase the budget or obligate patients to wait longer for treatment in cases of limited budget by evaluating B versus Q_1 or Q_2, using the given values. Therefore, not only does this methodology make the application useful in the Kuwaiti hospital case study, but it may be used more generally and in a diverse number of health care systems of varying complexity.

6. Conclusion and future work

This paper designs a decision support system for the operation of an emergency department unit at a governmental hospital in Kuwait by integrating simulation with optimization. We present a methodology that uses system simulation combined with optimization to determine the optimal number of staff members required to maximize patient throughput and to reduce patient time in the system subject to budget restrictions. The optimization simulation model presented in this paper provides optimal staffing allocation that would allow a 28% increase in patient throughput and an average of 40% reduction in patients' waiting time with the same resources. A decision support system is designed to help decision-makers at the hospital to either evaluate different situations of staffing distribution or optimize the system for optimal staffing distribution.

A plan is now under consideration as future work to construct a decision support system by establishing an interface linkage

Table 5
Input and output sections for the evaluation procedure for a given staff distribution.

Input section					
Staff	Reception	Doctors	Lab tec.	Nurse TR	Nurse ER
Number	2	2	3	1	9
Output section					
	Budget		Throughput		
	7.7		4.9		
Patient type	Expected number out (per hour)		Expected waiting time in system (hours)		
Category 1	2.2		3.58		
Category 2	1.9		3.29		
Category 3	0.8		2.85		
Queues					
Server	Expected length		Expected waiting time (minutes)		
Reception	0.13		1.26		
Doctors	13.3		99.45		
Laboratory	0.0		0.0		
Treat. Room	1.23		35.1		
Emrg. Room	0.0		0.0		
Efficiency					
Server	Expected number busy		Utilization		
Reception	0.8		0.4		
Doctors	1.8		0.9		
Laboratory	0.9		0.3		
Treat. Room	0.8		0.8		
Emrg. Room	3.4		0.4		

between Excel worksheet and SIMISCRIPT simulation software. The simulation program will take input data from Excel and perform the necessary simulation analysis. The evaluation section requires as an input the number of receptionists, doctors, lab technicians, treatment room nurses and emergency room nurses. Once these input data are entered through the interface input screen, the simulation program performs the necessary evaluation and produces the measure of performance for the current staffing distribution. The output section will include staff cost as well as system throughput and average waiting time in the system for patient categories 1, 2 and 3, along with their expected number out. The output will also provide detailed information about the queues formed in each of the services provided by the hospital. The expected length and the average waiting time for the queue in front of the reception, examination room, lab and X-ray room, treatment room and emergency room will be presented. Other information given in the evaluation output, which is of great importance to the hospital management, is staff utilization. This information will be presented in a table, which shows the expected number of busy staff of each type (receptionists, doctors, and nurses) as well as their utilization factor. As an example, for the current staffing distribution: two receptionists, two doctors, three technicians, one treatment nurse and nine emergency nurses with total budget units of 7.7, Table 5 presents a primary outline of the input and output section of the evaluation part of the planned decision support system.

For the optimization part, the decision-maker has to first select whether he/she would like to choose maximizing throughput or minimizing cost as his/her objective function. Then, he/she has to feed in the parameters for the constraint section, which include budget constraint, patient category 1 waiting time constraint, and staff lower and upper constraints. The information of the output section of the optimization part will be similar to the output section provided by the evaluation procedure. Both options provide decision-makers with plenty of information regarding the whole system.

Acknowledgements

This research was supported by the office of the Vice President for Scientific Research, Kuwait University, under project number SS01/08. The authors wish to thank the referees and the editor for their conscientious reading of this paper and their numerous suggestions for improvement which were extremely useful and helpful in modifying the manuscript.

References

Ahmed, M.A., Alkhamis, T.M., 2002. Simulation-based optimization using simulated annealing with ranking and selection. Computers and Operations Research 29, 387–402.

Ahmed, M.A., Alkhamis, T.M., Miller, D.R., 1998. Discrete search methods for optimizing stochastic systems. Computers and Industrial Engineering 34 (4), 703–716.

Alkhamis, T.M., Ahmed, M.A., 2004. Sequential stochastic comparison algorithm for simulation optimization. Engineering Optimization 36 (5), 513–524.

Alkhamis, T.M., Ahmed, M.A., Tuan, V.K., 1999. Simulated annealing for discrete optimization with estimation. European Journal of Operational research 116, 530–544.

Alrefaei, M.H., Andradottir, S., 1999. A simulated annealing algorithm with constant temperature for discrete stochastic optimization. Management Science 45 (5), 748–764.

Andradottir, S., 1995. A method for discrete stochastic optimization. Management Science 41 (12), 1946–1961.

Andradottir, S., 1996. Global search for discrete stochastic optimization. SIAM Journal on Optimization 6 (2), 513–530.

Andradottir, S., Goldsman, D., Kim, S.H., 2005. Fully sequential procedures for comparing constrained systems via simulation. In: Kuhl, M.E., Steiger, N.M., Armstrong, F.B., Joines, J.A. (Eds.), Proceedings of the 2005 Winter Simulation Conference. Institute of Electrical and Electronics Engineers, New Jersey.

Baesler, F.F., Sepúlveda, J.A., 2001. Multi-objective simulation optimization for a cancer treatment center. In: Proceedings of the 2001 Winter Simulation Conference, pp. 1405–1411.

Baesler, F.F., Jahnsen, H.E., DaCosta, M., 2003. Emergency departments I: the use of simulation and design of experiments for estimating maximum capacity in an emergency room. In: Proceedings of the 2003 Winter Simulation Conference, pp. 1903–1906.

Batur, D., Kim, S.H., 2005. Finding a set of feasible systems when the number of systems or constraints is large. Working Paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

Beaulieu, H., Ferland, J.A., Gendron, B., Michelon, P., 2000. A mathematical programming approach for scheduling physicians in the emergency room. Health Care Management Science 3 (3), 193–200.

Blasak, R.E., Armel, W.S., Starks, D.W., Hayduk, M.C., 2003. The use of simulation to evaluate hospital operations between the emergency department and a medical telemetry unit. In: Proceedings of the 2003 Winter Simulation Conference, pp. 1887–1893.

Boesel, J., Nelson, B.L., Kim, S.H., 2003. Using ranking and selection to "clean up" after simulation optimization. Operations Research 51, 814–825.

Butler, J., Morrice, D., Mullarkey, P., 2001. A multiple attribute utility theory approach to ranking and selection. Management Science 47, 800–816.

Côte, M.J., 1999. Patient flow and resource utilization in an outpatient clinic. Socio-economic Planning Sciences 33 (3), 231–245.

De Angelis, V., Felici, G., Impelluso, P., 2003. Integrating simulation and optimization in health care center management. European Journal of Operational Research 150, 101–114.

Ferreira de Oliveira, M.J. 1999. 3D visual simulation platform for the project of a new hospital facility. In: De Angelis, V., Ricciardi, N., Storchi, G.(Eds.), Monitoring, Evaluating, Planning Health Services. Proceedings to the 24th meeting of the ORAHS EURO-WG. World Scientific, Singapore, pp. 82–95.

Flessa, S., 2000. Where efficiency saves lives: A linear programme for the optimal allocation of health care resources in developing countries. Health Care Management Science 3 (3), 249–267.

Jacobson, S.H., Sewell, E.C., Deuson, E., Weniger, B.G., 1999. An integer programming model for vaccine procurement and delivery for childhood immunization: A pilot study. Health Care Management Science 2 (1), 1–9.

Kim, S.H., Nelson, B.L., 2001. A fully sequential procedure for indifference-zone selection in simulation. ACM TOMACS 11, 251–273.

Sinreich, D., Marmor, Y., 2005. Emergency department operations: The basis for developing a simulation tool. IIE Transactions 37, 233–245.

Swisher, J.R., Jacobson, S.H., Jun, J.B., Balci, O., 2001. Modeling and analyzing a physician clinic environment using discrete-event (visual) simulation. Computers and Operations Research 28 (2), 105–125.

Yan, D., Mukai, H., 1992. Stochastic discrete optimization. SIAM Journal on Control and Optimization 30, 594–612.