

# A SIMULATION-BASED OPTIMIZATION FRAMEWORK FOR BALANCING EFFICIENCY AND INTRA-GROUP EQUITY IN HEALTHCARE RESOURCE ALLOCATION

AHLAM BAHARI<sup>1\*</sup>, LINA ABOUELJINANE<sup>2</sup>, MARIA LEBBAR<sup>1,2</sup>

<sup>1</sup>MOAD6 Research Team (EMI), Mohammed V University of Rabat, Morocco

<sup>2</sup>Higher National School of Rabat (ENSMR), Rabat, Country

E-mail: \*bahariahlam@gmail.com

## ABSTRACT

**Purpose:** Efficient resource allocation is fundamental to mitigating Emergency Department (ED) congestion. However, global optimization often masks significant service disparities across patient groups. This study investigates the complex "Efficiency-Equity-Cost" trilemma by developing a dual-phase optimization framework designed to minimize boarding times and budgetary expenditures while explicitly enforcing intra-group equity across different patient severity levels.

**Methodology:** A Discrete Event Simulation (DES) model was developed to capture the stochastic dynamics of patient flow across an ED and three specialized Inpatient Units (IUs). The study employs a Simulation-Based Optimization (SBO) approach using the  $\epsilon$ -constraint method to identify Pareto-optimal trade-offs between budget and efficiency. In the second phase, an intra-group equity constraint was integrated to stabilize service performance within each Emergency Severity Index (ESI) category.

**Findings:** Results indicate that systemic boarding reduction is primarily achieved by expanding capacity in downstream IUs rather than the ED itself, identifying these units as the primary structural bottlenecks. While efficiency-only optimization yielded significant performance gains, the introduction of equity constraints revealed a critical feasibility ceiling. However, a "sweet spot" was identified at a budget of 150 units, where substantial improvements in equity are achieved with negligible impacts on global system efficiency.

**Originality/Value:** This research contributes a robust decision-support tool that quantifies the "Price of Fairness" in healthcare. By identifying the specific budget thresholds where equity and efficiency converge, this work provides hospital administrators with a theoretically grounded and practically viable strategy for ethically aligned resource allocation in high-pressure stochastic environments.

**Keywords:** *Emergency Department, Inpatient Unit, Simulation-Based Optimization, Multi-Objective Approach, Boarding Time, Equity.*

## 1. INTRODUCTION

The emergency department (ED) is a complex and dynamic organization that must provide urgent quality care services for heterogeneous patients [1]. This establishment faces too many challenges related to the uncertainties of demands, as well as service times which vary according to the availability of the resources [2]. The increasing demand for services, coupled with tight budgets, contributes significantly to overcrowding in the ED. These challenges result in long waiting times, limited resources, and inefficiencies in patient flow, making it difficult to provide timely and efficient care.

While decision-making in healthcare covers diverse areas, from staff scheduling [3], [4], [5] to resource optimization [6], [7], [8], [9], the

literature often overlooks the critical impact of boarding time and equity in integrated flows. Boarding, the practice of retaining admitted patients in the ED due to Inpatient Unit (IU) saturation [10], [11], consumes up to 40% of staff time [1] and is directly correlated with increased mortality rates [12]. Furthermore, achieving efficiency must not come at the cost of equity, ensuring that high-acuity patients are not disadvantaged by resource constraints.

Despite the critical nature of these issues, current multi-objective models in ED management frequently focus on aggregate performance indicators (e.g., average waiting time), failing to address the complex stochastic interactions between the ED and downstream units (IU) under strict budgetary and equity constraints. There is a lack of simulation-based frameworks capable of

simultaneously minimizing boarding times and maintaining service equity without exceeding operational budgets.

To address this gap, this study proposes a Multi-objective Simulation-Based Optimization (SBO) framework. The primary objective is to evaluate the trade-offs between boarding reduction, equity, and budget using the  $\epsilon$ -constraint method. This research is significant as it introduces a novel equity constraint focusing on service variability for patients within the same Emergency Severity Index (ESI), providing healthcare managers with a robust tool to identify critical resource thresholds.

The key contributions of this work are:

1. **Integrated Stochastic Modeling:** Capturing interactions between the ED, Observation Unit (OU), and Inpatient Units (IU) to identify downstream capacity as the primary lever for boarding reduction.
2. **Bi-objective Optimization:** Proposing a formulation that minimizes boarding time and budget via an  $\epsilon$ -constraint method integrated into an SBO engine.
3. **Novel Equity Constraint:** Introducing a constraint aimed at stabilizing service variability for patients of the same acuity level, moving beyond simple global averages.
4. **Sensitivity Analysis:** Identifying critical resource "tipping points" below which equity-driven strategies become operationally infeasible.

The remainder of this paper is structured as follows: Section 2 presents a systematic literature review of multi-objective approaches in healthcare, highlighting current gaps in boarding and equity modeling. Section 3 details the problem definition and the development of the Discrete Event Simulation (DES) model used to capture system stochasticity. Section 4 constitutes the core of our contribution, detailing the bi-objective mathematical formulations and providing a structured analysis of the experimental results, including sensitivity analyses. Finally, Section 5 concludes the paper by summarizing key findings, discussing managerial implications, and outlining future research directions.

## 2. LITERATURE REVIEW

The application of Simulation-Based Optimization (SBO) in healthcare has seen significant growth, particularly for resource dimensioning and strategic planning [13]. As illustrated in Table 1, which provides an overview of key references, target objectives, and function types,

the literature is rich but presents three major gaps that this study aims to bridge.

First, the scope of flow modeling. Many studies focus exclusively on internal ED processes, such as triage or bedside staffing [14], [15], [16], while treating IU as an infinite sink. This simplification ignores the boarding phenomenon, the bottleneck occurring at the interface of ED and IU, which is a primary driver of systemic congestion [17], [18], [19], [20]. Our model overcomes this by integrating the stochastic interactions between ED, OU, and IU capacity to identify downstream levers for boarding reduction.

Second, the treatment of conflicting objectives. As shown in Table 1, most existing healthcare models rely on weighted-sum objective functions [21], [22]. While computationally simple, these methods require arbitrary weight assignments that can bias results and fail to capture the true Pareto front. In contrast, our work employs the  $\epsilon$ -constraint method [23], [24]. This approach provides a mathematically rigorous trade-off analysis between boarding time and operational budget without the subjectivity of weighting, allowing for a clearer exploration of non-dominated solutions.

Third, the definition of equity. Equity is often absent from SBO models or reduced to global averages of waiting times [20], [25]. Such aggregate measures mask the high variability experienced by patients of the same acuity level. We propose a novel equity constraint that specifically targets the reduction of service variability (measured by standard deviation) within Emergency Severity Index (ESI) groups, ensuring a fairer distribution of care compared to traditional expected-value formulations [6].

The integration of these elements into a cohesive decision-support framework is detailed in Figure 1, which outlines our SBO process from data collection to sensitivity analysis.

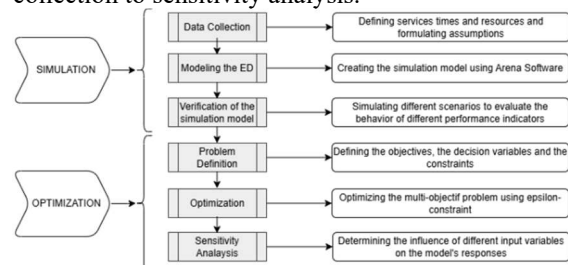


Figure 1: Overview of the Simulation-Based Optimization Process

The synthesis of these research gaps leads to the development of our proposed integrated approach. Figure 1 illustrates the global architecture of the Simulation-Based Optimization (SBO)

process followed in this study. This process is divided into two major phases:

1. **The Simulation phase**, which ensures the robust modeling of stochastic ED-IU interactions using Arena software.
2. **The Optimization phase**, which applies the  $\epsilon$ -constraint method to resolve the conflicts between boarding time, budget, and equity.

This structured workflow ensures that the sensitivity analysis (final step) is grounded in a verified and validated representation of the healthcare system, as detailed in the following section.

### 3. SYSTEM DESCRIPTION AND SIMULATION MODEL

This section delineates the architectural framework of the ED and the subsequent development of its Discrete-Event Simulation (DES) model. The system is conceptualized as a complex stochastic network where patient flow is governed by urgency-based priority rules and resource availability.

#### 3.1 Stochastic Patient Flow and Triage Logic

The ED environment is characterized by a non-stationary arrival process. Patients enter the system either via ambulance or as walk-ins. The hourly arrival intensity  $\lambda(t)$  is empirically derived from historical data, as illustrated in Figure 2.

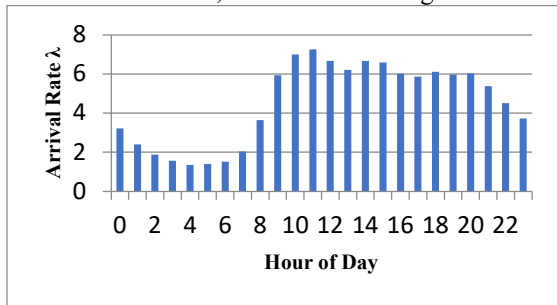


Figure 2: Mean hourly arrival rate distribution  $\lambda$  based on empirical data[26]

Upon arrival, patients undergo a formal triage process to determine their clinical priority using the Emergency Severity Index (ESI). The ESI scale ranges from 1 (immediate life-threatening) to 5 (non-urgent). As shown in Figure 3, the majority of walk-in patients are classified as ESI 4 (47%), while high-acuity cases (ESI 1 and 2) represent a smaller

but more resource-intensive fraction of the total volume.

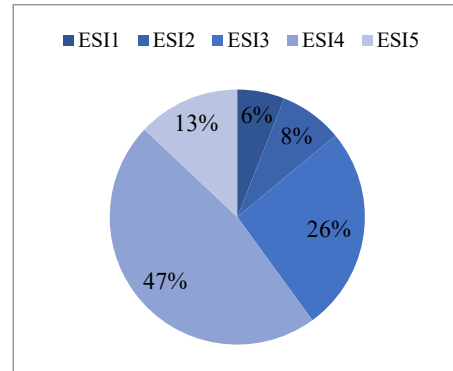


Figure 3: The distribution of patients arriving on their own according to the Emergency Severity Index

#### 3.2 Simulation Framework

The simulation model was developed using Arena (Rockwell Automation) to replicate the logical workflow depicted in Figure 4.

□ **Entry & Stabilization:** Initial processing (Registration and Triage) and immediate installation for ESI 1 cases.

□ **Clinical Assessment:** Resource-constrained examinations (Physicians and Nurses) where service times follow specific probability distributions (e.g., Beta, Gamma, and Triangular) as detailed in Table 2.

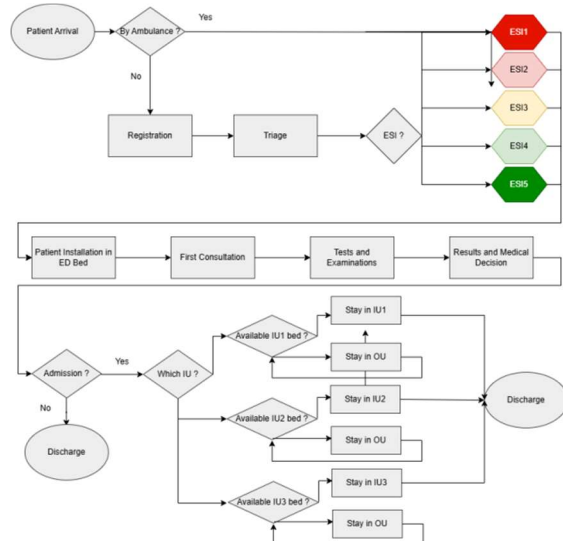


Figure 4: Flowchart representation of the emergency department simulation model

Table 2: Activities, Resources and Time Distributions

Activity	Resources	Service time distribution
Registration	Receptionist	Uniform (5, 10) (min)
Triage	Triage Nurse	Normal (7, 1.5) (min)
Patient installation	Nurse ED Bed	Uniform (3, 5) (min)
First consultation	Physician	20 + 10 * BETA(1.02, 6.08) (min)
Tests	Nurse	ESI 1 14+GAM M (1.32,3.53) (min)
		ESI 2 10+GAM M (3.11,1.42) (min)
		ESI 3 8+GAMM (3.11,1.42) (min)
		ESI 4 6+GAMM (3.11,1.42) (min)
		ESI 5 6+EXPO (1) (min)
Physician check	Physician	TRIA (8, 10, 12) (min)
Length of stay in IU1	IU1 Bed	TRIA (0.7, 1, 1.3) (days)
Length of stay in IU2	IU2 Bed	TRIA (0.8, 1, 1.4) (days)
Length of stay in IU3	IU3 Bed	TRIA (0.8, 1.2, 1.8) (days)

□Decision & Disposition: Following the medical decision, patients are either discharged or admitted to an Inpatient Unit (IU).

□Observation Unit (OU) & Bed Management: A critical feature of this model is the interaction between the OU and IUs. If a target IU bed is unavailable, the patient remains in the OU, creating a "bed-blocking" effect that impacts overall system throughput.

### 3.3 Input Data and Resource Configuration

The reliability of the simulation is ensured by the use of fitted stochastic distributions for each activity. Table 2 summarizes the service time distributions, including the specific parameters for tests and examinations based on ESI levels.

The system's initial capacity is defined in Table 3. By adopting a priority-based queuing discipline, the model reflects the real-world prioritization of more critical cases (ESI 1 & 2), ensuring they receive faster responsiveness while monitoring congestion levels across the facility.

Table 3: Resources' Initial Capacities

Resources	Initial capacities
Receptionists	2
Triage Nurses	3
Nurses	3
Physicians	4
ED Beds	30

OU Beds	8
IU1 Beds	20
IU2 Beds	20
IU3 Beds	25

### 3.4 Verification and Validation of the Simulation Model

To ensure the reliability of the proposed model, a rigorous Verification and Validation (V&V) process was conducted. To ensure the reliability of the proposed model, a rigorous Verification and Validation (V&V) process was conducted. To ensure the statistical significance of the results, the simulation protocol followed a rigorous three-step procedure. First, a warm-up period of 10 days was implemented to ensure the system reached a steady state, eliminating initialization bias before data collection. Second, the number of replications ( $n = 15$ ) was calculated to achieve a half-width of less than 5% of the mean for the primary KPI (Boarding Time) at a 95% confidence level. Third, each simulation run was synchronized with the optimization algorithm via an automated data exchange interface, ensuring that every candidate resource configuration was evaluated over the same stochastic arrival profile to maintain experimental consistency.

#### 3.4.1. Verification

The verification phase confirmed that the Arena implementation accurately reflected the conceptual logic. Two primary techniques were employed:

Face Validity: The model's internal logic was audited step-by-step. Using Arena's built-in debugging tools, we confirmed the absence of bottlenecks or logical errors during execution.

Sensitivity Analysis: Following [27], we tested the model's responsiveness by varying the patient arrival rate  $\lambda$ . Using the Process Analyzer (PAN), we simulated eight scenarios, increasing demand from 0% (baseline) to 35% in 5% increments. This analysis confirmed that the model's performance indicators responded realistically to increased pressure.

#### 3.4.2. Validation Metrics

The following Key Performance Indicators (KPIs) were selected to validate the model and evaluate ED efficiency across the simulated scenarios:

- Door-to-Doctor Time (DTDT): The cumulative interval from arrival to the first physician contact, including registration, triage, and bed

assignment. It serves as a primary metric for triage efficiency.

- Length of Stay (LOS): The total duration from patient entry to either discharge or admission to an Inpatient Unit (IU). LOS is the global benchmark for ED throughput.
- Boarding Time (BT): The critical waiting period between the medical decision to admit and the actual transfer to an IU bed. This metric specifically highlights the "bed-blocking" effect between the ED and IUs.
- Resource Utilization: The occupancy rates of physical assets (beds). This allows for identifying capacity constraints during peak demand periods.

The box plot in Figure 5 illustrates the variation in BT across the eight simulated demand scenarios. A clear upward trend is observed, validating the model's logical responsiveness to arrival rate fluctuations. For demand increases up to 10% (Scenarios 1 to 3), BT exhibits a sharp, nearly linear increase with significant dispersion (as shown by the larger whiskers), indicating high system sensitivity to initial growth.

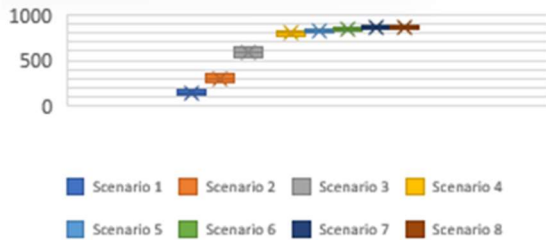


Figure 5: Impact of increased demand on patient boarding time (BT)

However, for demand increases exceeding 15% (Scenarios 4 to 8), the BT progression slows down as it approaches a plateau, accompanied by reduced variability. This pattern suggests system saturation, where the ED reaches its maximum operational capacity. At this stage, the congestion becomes constant across replications, highlighting the critical risks associated with limited bed capacity and the non-linear impact of overcrowding on patient safety.

Beyond the specific analysis of boarding delays, the overall stability and responsiveness of the ED were evaluated across multiple performance dimensions. While the BT highlights the critical bottleneck at the admission interface, other indicators such as DTDT, LOS, and resource occupancy provide a holistic view of the system's degradation under pressure. To maintain a concise yet rigorous validation, the mean values for these secondary KPIs are synthesized in Table 4 and Table

5. This comparative approach allows for a direct correlation between arrival rate fluctuations and the physical saturation of hospital assets.

Table 4: Evolution of Time-based KPIs

Scenario	Demand	DTDT (min)	LOS (min)
Scenario 1	Baseline (0%)	100	200
Scenario 4	+ 15%	200	350
Scenario 8	+ 35%	> 5000	> 3500

Table 5: Evolution of Bed Utilization Rates

Scenario	Demand	ED Bed Util.	OU Bed Util.	IU1, IU2, and IU3 Util (Avg)
Scenario 1	Baseline (0%)	20%	10%	70%
Scenario 4	+ 15%	40%	45%	90%
Scenario 8	+ 35%	100%	100%	100%

To provide a comprehensive overview of the system's performance, Table 4 and Table 5 summarize the evolution of time-based KPIs and bed utilization rates, respectively. The data reveals a direct correlation between resource occupancy and patient flow degradation. As bed utilization reaches 100% in the Observation Unit (OU) and Inpatient Units (IU), both Door-to-Doctor Time (DTDT) and Length of Stay (LOS) exhibit an exponential increase, confirming that physical capacity is the primary driver of ED congestion.

Table 5 shows that IU1, IU2, and IU3 occupancy rates reach near-saturation (90-100%) as early as Scenario 4. This confirms that the Inpatient Units act as the primary bottleneck of the entire patient flow.

Once the OU and IU beds are fully occupied (Scenario 8), the Boarding Time (BT) and Door-to-Doctor Time (DTDT) explode, exceeding 5000 minutes. This phenomenon occurs because patients cannot be moved out of the ED beds, effectively "freezing" the intake capacity at the front door.

The comparison between the two tables demonstrates that the system remains stable as long as bed utilization is below 85%. Beyond this threshold, every 5% increase in demand leads to a disproportionate increase in Length of Stay (LOS).

This validation phase confirms that the simulation model accurately replicates the real-world "crowding" behavior of the emergency department. The identified bottlenecks, specifically the physical bed capacities in OU and IU, provide a solid justification for the Simulation-Based

Optimization (SBO) developed in the following section.

#### 4. SIMULATION-BASED OPTIMIZATION (SBO) FRAMEWORK

After presenting and validating the simulation model in Section 3, this section introduces the simulation-based optimization framework used to improve system performance. We first reformulate the problem by defining the objective functions, the decision variables, and the constraints. In the first phase, the optimization of Model 1 focuses on minimizing both the total boarding time and the budget allocated to material resources. In the second phase, we extend the model into Model 2 by incorporating an equity constraint to ensure a fair distribution of boarding time across patients. The optimization was carried out using OptQuest, integrated within Arena Software, and complemented by the  $\epsilon$ -constraint method to explore trade-offs between conflicting objectives.

##### 4.1 Model 1: Bi-objective Optimization without Equity Constraint

###### 4.1.1 Problem Formulation and Assumptions

As a first stage, our basic model (Model 1) aims to minimize both boarding time and the budget for material resources. For this purpose, Table 6 presents the notations used for the following models.

Table 6: Simulation Based-Optimization of Model 1 Notations

Notation	Description
Set of Index	
$J$	The set of material resources $j=1$ to 5 ('1'=ED Bed, '2'=OU Bed, '3'=IU1 Bed, '4'=IU2 Bed, '5'=IU3 Bed)
Parameters	
$LB_j$	The lower bound of additional material resources $j=1$ to 5
$UB_j$	The upper bound of additional material resources $j=1$ to 5
$CM_j$	The annual costs of purchasing and maintaining material resources $j=1$ to 5
$R$	The number of DES replications
$w$	The randomness of the simulation model
Decision variables	
$X=(A_j, j=1$ to 5)	Integer variables corresponding to the additional number of material resources
KPIs in objective function	
$BT(X, w)$	Boarding time resulting from the vector of decision variables for a specific replication with a realization of random variables in the simulation expressed by the vector $w$

Notation	Description
$B(X)$	Budget for additional material resources resulting from the vector of decision variables

The models make the following assumptions:

- Stochastic Arrivals: Heterogeneous patient arrivals follow time-dependent rates as illustrated in Figure 2.
- Patient Categorization: Patients are modeled as a single entity type in Arena, and their acuity levels are assigned according to the ESI. The distribution of patient acuity across ESI levels is presented in Figure 3.
- Service Logic: The ED components (e.g., registration, triage, consultation) are modeled with specific service durations.
- OU Representation: The Observation Unit (OU) is represented solely by the length of stay, as patients do not receive active care there but merely wait for an available IU bed.
- Resource Scope: Only material resources are considered in the OU and IUs.

To identify the optimal configuration of resources, we define the following bi-objective mathematical model based on the notations in Table 6.

$$\min z_1 = E_R(BT(X, w)) \quad (1.1)$$

$$\min z_2 = B(X) = \sum_{j=1}^5 CM_j * A_j \quad (1.2)$$

Subject to:

$$LB_j \leq A_j \leq UB_j \quad j=1, 2, 3, 4, 5 \quad (1.3)$$

$$A_j \text{ integer } j=1, 2, 3, 4, 5 \quad (1.4)$$

In our model 1, we aim to minimize  $E_R(BT(X, w))$  the expected value boarding time  $BT$  computed across the  $R$  replications (1.1) and to minimize the total budget  $B$  for additional material resources (1.2). The expression  $B(X)$  represents a deterministic function of the input decision variables  $X$ , meaning it doesn't involve any randomness. In contrast,  $BT(X, w)$  depends not only on the decisions  $X$ , but also on the stochastic behavior introduced through  $w$ . Thus, Taking the expectation  $E_R(BT(X, w))$  allows us to evaluate the average performance of the system under the uncertainty modeled by  $w$ . In addition, the capacity constraints are expressed by the inequality (1.3), where the number of additional resources is limited

by both lower and upper bounds. The type of variables is expressed by (1.4).

To address our bi-objective problem, the  $\epsilon$ -constraint is employed. This method provides solutions for convex and non-convex optimization problems ensuring a comprehensive identification of the entire Pareto Front [28], [29], [30]. The epsilon-constraint method consists of optimizing one objective while the other is treated as a constraint bounded by a predefined threshold ( $\epsilon$ ). In our case, the primary objective is to minimize the total boarding time, while the budget for material resources acquisition is constrained by an upper limit  $\epsilon_m$ . The index  $m$  presents the  $\epsilon$ -constraint iteration number.

The epsilon-constraint method is conducted by varying the upper bound  $\epsilon_m$  of the secondary objective  $B$  across  $M$  iterations in (1.5), where each  $\epsilon_m$  corresponds to a different budget level. The resulting mathematical formulation is as follows:

$$\min z_1 = E_R(BT(X, w)) \quad (1.1)$$

Subject to:

$$LB_j \leq A_j \leq UB_j \quad j= 1, 2, 3, 4, 5 \quad (1.3)$$

$$A_j \text{ integer } j= 1, 2, 3, 4, 5 \quad (1.4)$$

$$B(X) \leq \epsilon_m \quad m \text{ denotes the } \epsilon\text{-constraint iteration index, } m \in \{1, 2, \dots, M\} \quad (1.5)$$

The index  $m$  in (1.5) denotes the iteration number in the epsilon-constraint procedure, with  $m = 1$  corresponding to the minimum budget and increasing progressively, varying  $\epsilon_m$  from 10 to 250. To solve this simulation-based optimization models, we use the OptQuest optimizer. OptQuest is a tool for optimization in simulation software like ARENA that enables the expression of the objective function and constraints using a combination of built-in functions and variables specific to the simulation model. OptQuest enables the search for optimal decision variables ( $X$ ) by applying metaheuristic algorithms, such as Tabu Search and Genetic Algorithms, to navigate the complex stochastic space defined by the simulation [31].

#### 4.1.2 Model 1 Optimization Results

The multi-objective simulation-based optimization for Model 1 was executed using the OptQuest engine within the Arena environment. The primary objective was to minimize the expected BT  $z_1 = E_R(BT(X, w))$  while simultaneously

minimizing the total budget  $z_2$  required for additional material resources.

We ran OptQuest for 800 simulations and 15 replications for each, doing this for each value of  $\epsilon_m$ . Table 7 summarizes the various parameters used in OptQuest.

Table 7: OptQuest settings for Model 1

<b>Controls</b>	-	<b>Number of additional beds <math>A_j</math> in each unit: ED, OU, IU1, IU2, and IU3</b>
<b>Responses</b>	-	Boarding Time <i>BT</i>
<b>Constraints</b>	-	Additional bed capacity constraint
	-	Budget constraint
<b>Objective</b>	-	Minimize Boarding Time <i>BT</i>
<b>Options</b>	-	Number of simulations : 800
	-	Number of replications per simulation: 15
	-	Tolerance for the equality of two solutions: 0

Figure 5 illustrates the optimization progress for a specific budget threshold  $\epsilon_5 = 90$ , showing how the heuristic search algorithms (such as Tabu Search and Genetic Algorithms) successfully identified configurations that minimized boarding time within the set constraints.

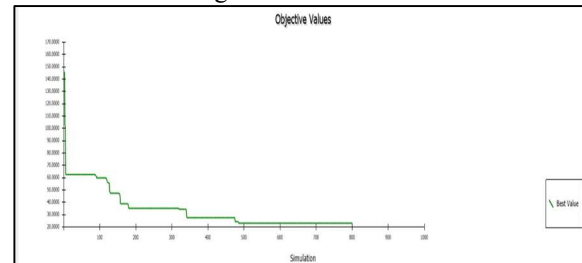


Figure 6: The results obtained for  $\epsilon_5 = 90$  using OptQuest

By running OptQuest for the different  $\epsilon$ -constraint iteration indexes  $m$ , namely different budget levels ranging from 10 to 250 units, we obtained the following Front Pareto shown in Figure 6. In this regard, Figure 6 shows the trade-off between the two objective functions  $z_1$  and  $z_2$ .

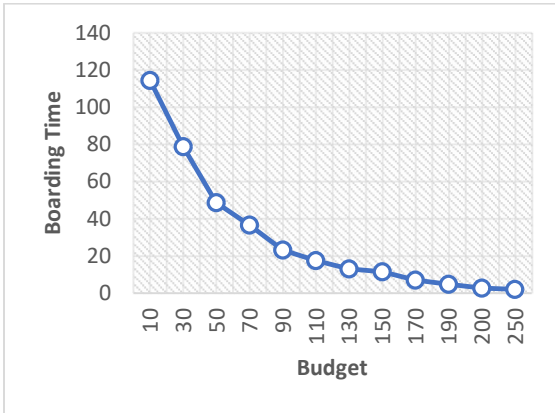


Figure 7: Pareto Front illustrating the trade-off between total budget  $z_2$  and expected boarding time  $z_1$  for Model 1

The proposed SBO framework demonstrates high efficiency in identifying the critical 'Knee Point' within the budget range of 90 to 130 units. In this specific situation, the model achieves the best trade-off by prioritizing IU2 and IU3, which were identified as the primary physical bottlenecks. Conversely, the method shows diminishing returns beyond a budget of 230, where further investment yields less than 1% improvement in boarding time, indicating that the system has reached its theoretical capacity limit.

The results in Figure 7 demonstrate a clear non-linear, convex relationship between financial investment and operational efficiency.

The analysis reveals a significant diminishing returns effect as the budget increases:

- High-Impact Zone: Increasing the budget from 10 to 30 results in a sharp reduction in boarding time from 114.37 to 78.77, a gain of over 35 units.
- Saturation Zone: Conversely, increasing the budget from 230 to 250 only yields a marginal improvement, with boarding time decreasing from 2.65 to 2.06, a gain of less than 0.6 units.

The convexity of the Pareto curve offers critical insights for strategic resource allocation:

- Optimal Investment Range: A budget between 90 and 130 achieves a substantial reduction in boarding time (from 23.15 to 13.06) while maintaining moderate costs.
- Feasibility vs. Performance: Aiming for the absolute minimum boarding time at a budget of 250 requires a disproportionately high investment that may not be feasible or justifiable in resource-constrained environments.

Ultimately, this analysis highlights the critical trade-off between cost and operational performance in hospital units. By understanding the shape of this curve, decision-makers can avoid over-investment and identify the "knee point" where system efficiency is maximized for the least relative cost.

The distribution of additional beds across the Pareto front provides critical insights into the system's structural constraints. As synthesized in Table 8, the OptQuest engine consistently maintained the number of additional beds for the Emergency Department  $A_1$  and the Observation Unit  $A_5$  at zero across all budget levels.

Table 8: Values of decision variables across budget levels

(B ; BT)	A1 (ED Bed)	A2 (IU1 Bed)	A3 (IU2 Bed)	A4 (IU3 Bed)	A5 (OU Bed)
(10 ; 114.37)	0	0	1	0	0
(30 ; 78.77)	0	1	2	0	0
(50 ; 48.71)	0	1	2	2	0
(70 ; 36.66)	0	2	3	2	0
(90 ; 23.15)	0	2	4	3	0
(110 ; 17.49)	0	3	4	4	0
(130 ; 13.06)	0	4	5	4	0
(150 ; 11.58)	0	3	6	5	0
(170 ; 7.01)	0	5	6	6	0
(190 ; 4.80)	0	4	7	8	0
(200 ; 2.65)	0	6	8	7	0
(250 ; 2.06)	0	6	9	8	0

Instead, capacity increases were exclusively directed toward the Inpatient Units (IU1, IU2, and IU3), with IU2 receiving the highest priority (up to +9 beds at  $\epsilon_{13} = 250$ ). This allocation pattern confirms that downstream congestion in specialized units is the primary driver of the "boarding" phenomenon. By prioritizing IU capacity, the model effectively "pulls" patients out of the ED, thereby reducing the systemic blockage.

While Model 1 identifies cost-efficient solutions, a deeper analysis reveals significant disparities in service performance across different patient groups. Figures 7 to 11 illustrate the standard deviation of boarding times for ESI levels 1 through 5 across various budget thresholds.

The results show a clear downward trend: as the budget increases from  $B = 90$  to  $B = 170$  the variability of boarding times within and between ESI groups significantly decreases. For instance, in ESI 1 (Figure 7), the standard deviation narrows as

investment grows, indicating a more predictable and uniform service. The most equitable outcomes, characterized by the lowest variability, are exclusively observed at the highest budget levels.

This highlights a critical operational trade-off: achieving fairness in patient care requires a greater investment in resources. In Model 1, equity is merely a byproduct of high spending rather than a primary objective.

To explicitly address this limitation, the following section introduces Model 2. This model integrates an equity constraint directly into the optimization framework, allowing for a systematic exploration of solutions that balance operational efficiency and cost while ensuring a fair distribution of performance across all patient acuity levels.

Standard Deviation ES11 by Scenario

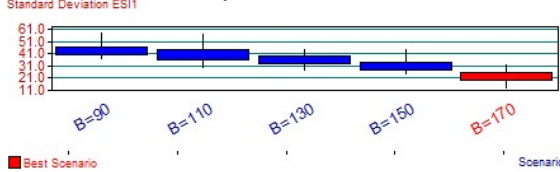


Figure 8: Boxplots of standard deviations of boarding times across ESI 1 patients for each budget level

Standard Deviation ES12 by Scenario

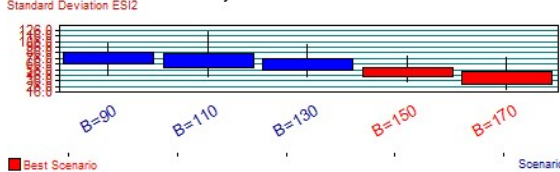


Figure 9: Boxplots of standard deviations of boarding times across ESI 2 patients for each budget level

Standard Deviation ES13 by Scenario

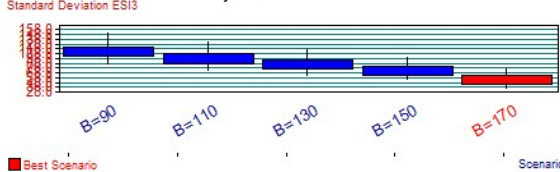


Figure 10: Boxplots of standard deviations of boarding times across ESI 3 patients for each budget level

Standard Deviation ES14 by Scenario

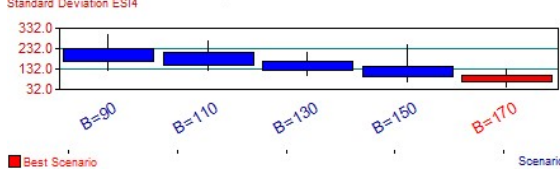


Figure 11: Boxplots of standard deviations of boarding times across ESI 4 patients for each budget level

Standard Deviation ES15 by Scenario

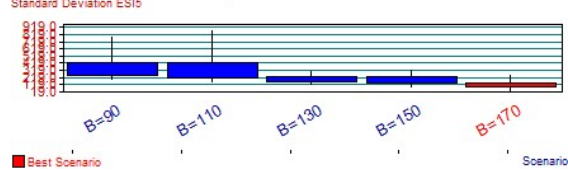


Figure 12: Boxplots of standard deviations of boarding times across ESI 5 patients for each budget level

This highlights a critical operational trade-off: achieving fairness in patient care requires a greater investment in resources. In Model 1, equity is merely a byproduct of high spending rather than a primary objective.

To explicitly address this limitation, the following section introduces Model 2. This model integrates an equity constraint directly into the optimization framework, allowing for a systematic exploration of solutions that balance operational efficiency and cost while ensuring a fair distribution of performance across all patient acuity levels.

## 4.2 Model 2: Bi-objective Optimization with Equity Constraint

### 4.2.1 Problem Formulation and Assumptions

While the patient population under study has been segmented into clinically homogeneous groups, ensuring intra-group equity remains a critical component of the resource allocation model.

By explicitly introducing an intra-group equity constraint, the model promotes a fair distribution of resources or outcomes among individuals with similar needs [32], [33], and supports the integration of equity as a formal, quantitative objective within operational decision-making. For this purpose, we implemented a second model, Model 2, where we considered an additional constraint to optimize the equity between patient for each level of ESI. Thus, we considered the standard deviation of each ESI. The notations of this model are summarized in Table 10.

Table 10: Simulation Based-Optimization Model 2 with Equity Constraint Notations

Notation	Description
Set of Index	
$J$	The set of material resources $j=1$ to 5 ('1'=ED Bed, '2'=OU Bed, '3'=IU1 Bed, '4'=IU2 Bed, '5'=IU3 Bed)
$K$	The set of Emergency Severity Indexes $k=1$ to 5
Parameters	
$LB_j$	The lower bound of additional material resources $j=1$ to 5

Notation	Description
$UB_j$	The upper bound of additional material resources $j=1$ to 5
$CM_j$	The annual costs of purchasing and maintaining material resources $j=1$ to 5
$UBSD_k$	The upper bound of standard deviations of boarding time for each $ESI_k$ $k=1$ to 5
$R$	The number of DES replications
$w$	The randomness of the simulation model
Decision variables	
$X=(A_j, j=1$ to 5)	Integer variables corresponding to the additional number of material resources
KPIs in objective function	
$BT(X, w)$	Boarding time resulting from the vector of decision variables for a specific replication with a realization of random variables in the simulation expressed by the vector $w$
$B(X)$	Budget for additional material resources resulting from the vector of decision variables
$SD_k(X, w)$	The standard deviations of boarding time for each $ESI_k$ $k=1$ to 5 resulting from the vector of decision variables for a specific replication with a realization of random variables in the simulation expressed by the vector $w$

The model 2 with equity constraint is presented as follows:

$$\min z_1 = E_R(BT(X, w)) \tag{2.1}$$

$$\min z_2 = B(X) = \sum_{j=1}^5 CM_j * A_j \tag{2.2}$$

Subject to:

$$LB_j \leq A_j \leq UB_j, j=1, 2, 3, 4, 5 \tag{2.3}$$

$$E_R(SD_k(x, w)) \leq UBSD_k, k=1 \text{ to } 5 \tag{2.4}$$

$$A_j \text{ integer } j=1, 2, 3, 4, 5 \tag{2.5}$$

Model 2 consists of minimizing the same two objectives  $z_1$  and  $z_2$ , namely BT and B. Model 2 adds to Model 1 the constraint (2.4) that presents the expected value of standard deviations  $SD_k$  of boarding times delayed by patients with different  $ESI_k$  which is computed across the  $R$  replications. The standard deviation indicates a more equitable distribution among the patients of the same level of acuity, namely ESI.

To investigate the impact of the added constraint, five mid-range solutions of the Pareto front were selected and used to re-execute the model. The selected  $B$  values (90, 110, 130, 150, and 170) were subsequently re-applied in the model, where the  $BT$  variable was reassessed with the equity

constraint explicitly imposed. Accordingly, Model 2 is formulated as follows:

$$\min z_1 = E_R(BT(X, w)) \tag{2.1}$$

Subject to:

$$LB_j \leq A_j \leq UB_j, j=1, 2, 3, 4, 5 \tag{2.3}$$

$$E_R(SD_k(X, w)) \leq UBSD_k, k=1 \text{ to } 5 \tag{2.4}$$

$$A_j \text{ integer } j=1, 2, 3, 4, 5 \tag{2.5}$$

$$B(X) = \sum_{j=1}^5 CM_j * A_j \leq \epsilon_m, m \text{ denotes the } \epsilon\text{-constraint iteration index, } \epsilon_m \in \{90, 110, 130, 150, 170\} \tag{2.6}$$

The iterative process expressed by (2.6) allowed us to analyze the sensitivity of the boarding time  $BT$  to increasing levels of intra-group equity enforcement for the selected  $B$  values (90, 110, 130, 150, and 170). Starting from the initial standard deviation values obtained from the unconstrained model, Model 1, we progressively tightened it by decreasing the allowed  $UBSD_k$ .

The objective was to evaluate how the system's  $BT$  would be affected by the addition of the intra-group equity constraint. This approach allows us to assess the impact of equity on the temporal performance of compromise solutions in terms of the budget-boarding time trade-off.

#### 4.2.2 Model 2 Optimization Results

Model 2 is solved using the epsilon constraint method implemented in OptQuest. At each step, we re-ran the model and recorded the corresponding changes in  $BT$ . The objective was to observe the trade-off between fairness and efficiency, i.e., how enforcing a more equitable solution impacts the overall system performance in terms of time, particularly  $BT$ .

Table 11 shows the setting introduced in OptQuest.

Table 11: OptQuest settings for Model 2

Controls	Number of additional beds $A_j$ in each unit: ED, OU, IU1, IU2, and IU3
Responses	Boarding Time BT Additional bed capacity constraint
Constraints	Budget constraint BT standard deviation constraint for each ESI
Objective Options	Minimize Boarding Time BT Number of simulations: 800 Number of replications per simulation: 15 Tolerance for equality of two solutions: 0

Model 2 integrates an intra-group equity constraint  $UBSD_k$ . We performed a sensitivity analysis by tightening  $UBSD_k$  in 5% increments.

- The Feasibility Ceiling: For budgets B=90 and B=110, solutions become infeasible as soon as the equity constraint requires a reduction of 10% or more.
- Critical Finding: At a 20% equity reduction, the system becomes entirely infeasible across all budget levels, highlighting that absolute fairness is structurally impossible within current resource limits.

To assess the "Price of Fairness," we utilized the PAN to compare Model 1 and Model 2. The Price of Fairness (Figures 13 and 14): At lower budget levels, specifically B=90 and B=110, a slight upward shift in the median boarding time is observed when the equity constraint ( $UBSD_k = 5\%$ ) is applied.



Figure 13: Boxplot Showing Overlapping Boarding Times Distributions for B=90

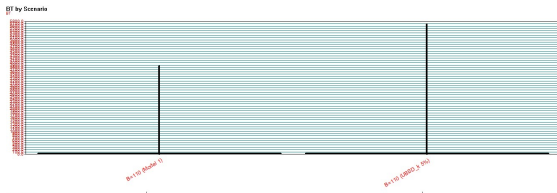


Figure 14: Boxplot Showing Overlapping Boarding Times Distributions for B=110

- Efficiency-Equity Convergence (Figures 15 and 16): For higher budgets, such as B=150 and B=170, the distributions between Model 1 and Model 2 show a high degree of overlap. Interestingly, at B=150, the differences in BT between the models for varied  $UBSD_k$  (5%, 10%, and 15%) are minimal and statistically insignificant.

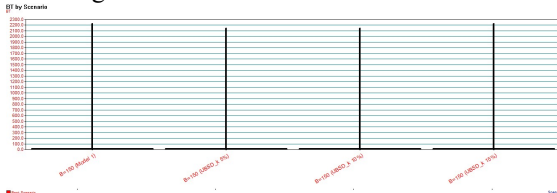


Figure 15: Boxplot Showing Overlapping Boarding Times Distributions for B=150



Figure 16: Boxplot Showing Overlapping Boarding Times Distributions for B=170

The analysis of resource allocation decisions generated by OptQuest reveals meaningful insights into system dynamics under varying budget levels. Table 12, Table 13, and Table 14 display the decision variables configurations corresponding to each budget level considered.

Table 12: Values of decision variables across budget levels for the model with equity constraint (5%)

(B ; BT)	A1 (ED Bed)	A2 (IU1 Bed)	A3 (IU2 Bed)	A4 (IU3 Bed)	A5 (OU Bed)
(90 ; 24.32)	0	2	3	4	0
(110 ; 16.98)	0	3	5	3	0
(130 ; 13.06)	0	4	5	4	0
(150 ; 8.36)	0	3	7	5	0
(170 ; 6.30)	0	4	5	8	0

Table 13: Values of decision variables across budget levels for the model with equity constraint (10%)

(B ; BT)	A1 (ED Bed)	A2 (IU1 Bed)	A3 (IU2 Bed)	A4 (IU3 Bed)	A5 (OU Bed)
(90 ; INFEASIBLE)					
(110 ; INFEASIBLE)					
(130 ; 13.06)	0	4	5	4	0
(150 ; 8.36)	0	3	7	5	0
(170 ; INFEASIBLE)					

Table 14: Values of decision variables across budget levels for the model with equity constraint (15%)

(B ; BT)	A1 (ED Bed)	A2 (IU1 Bed)	A3 (IU2 Bed)	A4 (IU3 Bed)	A5 (OU Bed)
(90 ; INFEASIBLE)					
(110 ; INFEASIBLE)					
(130 ; INFEASIBLE)					
(150 ; 8.62)	0	4	5	6	0
(170 ; INFEASIBLE)					

The shift from Model 1 to Model 2 also influences the physical distribution of beds. While the general strategy remains focused on downstream

units, the specific configurations change to meet equity goals:

- **Infeasibility Zones:** Tables 13 and 14 highlight that for budgets of 90 and 110, requiring a 10% or 15% reduction in variability leads to infeasible results. The system simply lacks the capacity to satisfy both the flow requirements and the strict equity bounds.
- **Priority Units:** At the viable spot of  $B=150$ , the model maintains a robust presence in IU2 (7 beds) and IU3 (5 to 6 beds) to stabilize the standard deviation across all acuity levels.

A key finding from the sensitivity analysis is the structural infeasibility observed in Tables 13 and 14 for lower budget levels ( $B=90$  and  $B=110$ ). This represents a critical systemic limitation: when resources are scarce, the model cannot mathematically satisfy the 10% or 15% equity improvement threshold without violating flow constraints. This 'Risk of Rigidity' suggests that equity policies must be flexible and budget-dependent to avoid operational deadlock.

The comparative analysis of Model 1 and Model 2 offers two critical takeaways for hospital practitioners:

- **Equity is Budget-Dependent:** Achieving fairness is expensive at low budgets but becomes nearly free in terms of performance loss at higher budget levels (above 150 units).
- **The Risks of Rigidity:** The infeasibilities recorded in Table 15 warn against enforcing equity as a "hard constraint" in resource-limited environments. If resources are too scarce, a strict equity policy may inadvertently lead to systemic failure rather than better care.
- **Optimal Strategy:** We recommend a budget level of 150 as it provides the necessary flexibility to reduce intra-group variability by up to 15% while keeping boarding times at a minimum.

## 5. CONCLUSION AND FUTURE RESEARCH

This research investigated the critical trade-offs between boarding time reduction, budgetary constraints, and service equity within a multi-ESI (Emergency Severity Index) healthcare environment. By integrating a severity-based queuing model with a secondary intra-group equity framework, this study moves beyond traditional efficiency-only metrics to address the ethical dimensions of patient flow management.

The application of the Simulation-Based Optimization (SBO) approach, validated through a

Projective Adaptive Neural Network (PAN), provided several key insights. Notably, the model demonstrated that systemic boarding time reduction is primarily achieved by expanding capacity in downstream Inpatient Units (IU2 and IU3), rather than the Emergency Department itself. Furthermore, while the equity-constrained model successfully reduced intra-group variability, it revealed a significant "feasibility ceiling." Budget levels below 130 units proved insufficient to satisfy strict equity improvements, often leading to infeasible solutions due to the lack of resource "slack." However, a pivotal finding is the identification of a "sweet spot" at a budget of 150 units, where a 15% improvement in fairness can be achieved with statistically insignificant impacts on global efficiency.

From a managerial perspective, these results warn practitioners against enforcing equity as a "hard" constraint in resource-limited settings, as it may paradoxically lead to systemic failure. Instead, administrators should aim for identified budget thresholds where fairness and efficiency converge. While minimizing average wait times remains a valid goal, ignoring intra-group variability can lead to extreme service disparities that undermine patient satisfaction and ethical alignment.

Despite its contributions, this study opens several avenues for further investigation. Future work should explore the relaxation of equity bounds into "soft constraints" [34] or equity-weighted objective functions to maintain operational viability in low-budget scenarios. Additionally, incorporating real-time staff workload, bed utilization rates, and nurse-to-patient ratios would enhance the model's fidelity. Additionally, exploring machine learning-based arrival forecasting could further enhance the proactive allocation of resources during peak demand periods. Methodologically, the integration of AI-driven metamodels [35] with real historical data could further improve predictive accuracy. Finally, exploring behavioral factors, such as patient preferences and canceled demands, would refine the model's applicability to increasingly complex and dynamic healthcare markets.

## REFERENCES

- [1] G. Savioli *et al.*, "Emergency Department Overcrowding: Understanding the Factors to Find Corresponding Solutions," *J. Pers. Med.*, vol. 12, no. 2, Art. no. 2, Feb. 2022, doi: 10.3390/jpm12020279.
- [2] "Health - OECD." Accessed: Feb. 23, 2024. [Online]. Available: <https://www.oecd.org/health/>

- [3] O. EL-Rifai, T. Garaix, V. Augusto, and X. Xie, "A stochastic optimization model for shift scheduling in emergency departments," *Health Care Manag. Sci.*, vol. 18, no. 3, pp. 289–302, Sep. 2015, doi: 10.1007/s10729-014-9300-4.
- [4] H. R. Feili, "Improving The Health Care Systems Performance By Simulation Optimization," *J. Math. Comput. Sci.*, vol. 07, pp. 73–79, Jul. 2013, doi: 10.22436/jmcs.07.01.08.
- [5] F. Zaerpour, M. Bijvank, H. Ouyang, and Z. Sun, "Scheduling of Physicians with Time-Varying Productivity Levels in Emergency Departments," *Prod. Oper. Manag.*, vol. 31, no. 2, pp. 645–667, Feb. 2022, doi: 10.1111/poms.13571.
- [6] M. A. Ahmed and T. M. Alkhamis, "Simulation optimization for an emergency department healthcare unit in Kuwait," *Eur. J. Oper. Res.*, vol. 198, no. 3, pp. 936–942, Nov. 2009, doi: 10.1016/j.ejor.2008.10.025.
- [7] E. Cabrera, M. Taboada, M. L. Iglesias, F. Epelde, and E. Luque, "Optimization of Healthcare Emergency Departments by Agent-Based Simulation," *Procedia Comput. Sci.*, vol. 4, pp. 1880–1889, 2011, doi: 10.1016/j.procs.2011.04.204.
- [8] D. Daldoul, I. Nouaouri, H. Bouchriha, and H. Allaoui, "Optimization on human and material resources in Emergency Department," in *2015 International Conference on Industrial Engineering and Systems Management (IESM)*, Seville, Spain: IEEE, Oct. 2015, pp. 633–638. doi: 10.1109/IESM.2015.7380224.
- [9] F. Zeinali, M. Mahootchi, and M. M. Sepehri, "Resource planning in the emergency departments: A simulation-based metamodeling approach," *Simul. Model. Pract. Theory*, vol. 53, pp. 123–138, Apr. 2015, doi: 10.1016/j.simpat.2015.02.002.
- [10] M. Martin, R. Champion, L. Kinsman, and K. Masman, "Mapping patient flow in a regional Australian emergency department: A model driven approach," *Int. Emerg. Nurs.*, vol. 19, no. 2, pp. 75–85, Apr. 2011, doi: 10.1016/j.ienj.2010.03.003.
- [11] P. Shi, M. C. Chou, J. G. Dai, D. Ding, and J. Sim, "Models and Insights for Hospital Inpatient Operations: Time-Dependent ED Boarding Time," *Manag. Sci.*, vol. 62, no. 1, pp. 1–28, Jan. 2016, doi: 10.1287/mnsc.2014.2112.
- [12] L. He, S. C. Madathil, A. Oberoi, G. Servis, and M. T. Khasawneh, "A systematic review of research design and modeling techniques in inpatient bed management," *Comput. Ind. Eng.*, vol. 127, pp. 451–466, 2019, doi: <https://doi.org/10.1016/j.cie.2018.10.033>.
- [13] L. Wang and E. Demeulemeester, "Simulation optimization in healthcare resource planning: A literature review," *IIEE Trans.*, vol. 55, no. 10, pp. 985–1007, Oct. 2023, doi: 10.1080/24725854.2022.2147606.
- [14] C. X. Qin and X. Feng, "Study on Stress Response Method of Nurses in Department of Emergency Medicine," *Adv. Mater. Res.*, vol. 433–440, pp. 1962–1966, 2012, doi: 10.4028/www.scientific.net/AMR.433-440.1962.
- [15] D. Sinreich, O. Jabali, and N. P. Dellaert, "Reducing emergency department waiting times by adjusting work shifts considering patient visits to multiple care providers," *IIE Trans.*, vol. 44, no. 3, pp. 163–180, Mar. 2012, doi: 10.1080/0740817X.2011.609875.
- [16] D. Sinreich and O. Jabali, "Staggered work shifts: a way to downsize and restructure an emergency department workforce yet maintain current operational performance," *Health Care Manag. Sci.*, vol. 10, no. 3, pp. 293–308, Jun. 2007, doi: 10.1007/s10729-007-9021-z.
- [17] S. Lucidi, M. Maurici, L. Paulon, F. Rinaldi, and M. Roma, "A Simulation-Based Multiobjective Optimization Approach for Health Care Service Management," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 4, pp. 1480–1491, Oct. 2016, doi: 10.1109/TASE.2016.2574950.
- [18] M. E. H. Petering, O. T. Aydas, K. Kuzu, and A. Ross, "Simulation analysis of hospital intensive care unit reimbursement policies from the triple bottom line perspective," *J. Simul.*, vol. 9, no. 2, pp. 86–98, May 2015, doi: 10.1057/jos.2014.24.
- [19] X. Wang, X. Gong, N. Geng, Z. Jiang, and L. Zhou, "Metamodel-based simulation optimisation for bed allocation," *Int. J. Prod. Res.*, vol. 58, no. 20, pp. 6315–6335, Oct. 2020, doi: 10.1080/00207543.2019.1677962.
- [20] L. Zhou, N. Geng, Z. Jiang, and X. Wang, "Multi-objective capacity allocation of hospital wards combining revenue and equity," *Omega*, vol. 81, pp. 220–233, 2018, doi: <https://doi.org/10.1016/j.omega.2017.11.005>.
- [21] M. Bazirha, A. Kadrani, and R. Benmansour, "Stochastic home health care routing and scheduling problem with multiple synchronized services," *Ann. Oper. Res.*, vol. 320, no. 2, pp. 573–601, Jan. 2023, doi: 10.1007/s10479-021-04222-w.
- [22] K. J. Klassen and R. Yoogalingam, "Appointment scheduling in multi-stage outpatient clinics," *Health Care Manag. Sci.*,

- vol. 22, no. 2, pp. 229–244, Jun. 2019, doi: 10.1007/s10729-018-9434-x.
- [23] “Multi-objective Optimization,” in *Multi-objective Management in Freight Logistics*, London: Springer London, 2008, pp. 11–36. doi: 10.1007/978-1-84800-382-8\_2.
- [24] K. Miettinen, “Introduction to Multiobjective Optimization: Noninteractive Approaches,” in *Multiobjective Optimization*, vol. 5252, J. Branke, K. Deb, K. Miettinen, and R. Słowiński, Eds., in Lecture Notes in Computer Science, vol. 5252, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 1–26. doi: 10.1007/978-3-540-88908-3\_1.
- [25] A. Saremi, P. Jula, T. ElMekawy, and G. G. Wang, “Bi-criteria appointment scheduling of patients with heterogeneous service sequences,” *Expert Syst. Appl.*, vol. 42, no. 8, pp. 4029–4041, May 2015, doi: 10.1016/j.eswa.2015.01.013.
- [26] K. Ghanes *et al.*, “Simulation-based optimization of staffing levels in an emergency department,” *SIMULATION*, vol. 91, no. 10, pp. 942–953, Oct. 2015, doi: 10.1177/0037549715606808.
- [27] R. G. Sargent, “Verification and validation of simulation models,” *J. Simul.*, vol. 7, no. 1, pp. 12–24, Feb. 2013, doi: 10.1057/jos.2012.20.
- [28] G. Mavrotas and K. Florios, “An improved version of the augmented  $\epsilon$ -constraint method (AUGMECON2) for finding the exact pareto set in multi-objective integer programming problems,” *Appl. Math. Comput.*, vol. 219, no. 18, pp. 9652–9669, May 2013, doi: 10.1016/j.amc.2013.03.002.
- [29] S. Makhoul, A.-L. Olteanu, and M. Sevaux, “A multiobjective  $\epsilon$ -constraint based approach for the robust master surgical schedule under multiple uncertainties,” *Eur. J. Oper. Res.*, Aug. 2024, doi: 10.1016/j.ejor.2024.08.022.
- [30] J. Liu, J. Feng, S. Yang, H. Zhang, and S. Liu, “Dynamic  $\epsilon$ -multilevel hierarchy constraint optimization with adaptive boundary constraint handling technology,” *Appl. Soft Comput.*, vol. 152, p. 111172, Feb. 2024, doi: 10.1016/j.asoc.2023.111172.
- [31] N. Chemkomnerd, W. Pannakkong, T. Tanantong, van-N. Huynh, and J. Karnjana, “A Simulation-Based Multi-Objective Optimization Framework to Enhance Patient Satisfaction: A Case Study of Ophthalmology Department Management,” *IEEE Access*, vol. 12, pp. 93197–93220, 2024, doi: 10.1109/ACCESS.2024.3423394.
- [32] I. P. Mahmood, H. Zhu, and A. Zaheer, “Centralization of intragroup equity ties and performance of business group affiliates,” *Strateg. Manag. J.*, vol. 38, no. 5, pp. 1082–1100, 2017, doi: 10.1002/smj.2542.
- [33] D. A. Strauss, “Must Like Cases Be Treated Alike?,” May 01, 2002, *Social Science Research Network, Rochester, NY*: 312180. doi: 10.2139/ssrn.312180.
- [34] K.-C. Chen, S.-Y. Lin, and M.-M. Yu, “Exploring the efficiency of hospital and pharmacy utilizations in Taiwan: An application of dynamic network data envelopment analysis,” *Socioecon. Plann. Sci.*, vol. 84, p. 101424, Dec. 2022, doi: 10.1016/j.seps.2022.101424.
- [35] F.-Z. Sahlaoui, L. Aboueljinane, and M. Lebbar, “A review on simulation-based metamodeling in emergency healthcare: methodology, applications, and future challenges,” *SIMULATION*, vol. 99, no. 10, pp. 989–1009, Oct. 2023, doi: 10.1177/00375497231183722.
- [36] M. Lamiri, X. Xie, A. Dolgui, and F. Grimaud, “A stochastic model for operating room planning with elective and emergency demand for surgery,” *Eur. J. Oper. Res.*, vol. 185, no. 3, pp. 1026–1037, Mar. 2008, doi: 10.1016/j.ejor.2006.02.057.
- [37] K. J. Klassen and R. Yoogalingam, “Appointment system design with interruptions and physician lateness,” *Int. J. Oper. Prod. Manag.*, vol. 33, no. 4, pp. 394–414, Mar. 2013, doi: 10.1108/01443571311307253.
- [38] Y.-Y. Feng, I.-C. Wu, and T.-L. Chen, “Stochastic resource allocation in emergency departments with a multi-objective simulation optimization algorithm,” *Health Care Manag. Sci.*, vol. 20, no. 1, pp. 55–75, Mar. 2017, doi: 10.1007/s10729-015-9335-1.
- [39] M. Lamiri, F. Grimaud, and X. Xie, “Optimization methods for a stochastic surgery planning problem,” *Int. J. Prod. Econ.*, vol. 120, no. 2, pp. 400–410, Aug. 2009, doi: 10.1016/j.ijpe.2008.11.021.
- [40] K. J. Klassen and R. Yoogalingam, “Strategies for Appointment Policy Design with Patient Unpunctuality,” *Decis. Sci.*, vol. 45, no. 5, pp. 881–911, Oct. 2014, doi: 10.1111/deci.12091.
- [41] E. K. Lee *et al.*, “Transforming Hospital Emergency Department Workflow and Patient Care,” *Interfaces*, vol. 45, no. 1, pp. 58–82, Feb. 2015, doi: 10.1287/inte.2014.0788.

- [42] F. Mallor, C. Azcárate, and J. Barado, "Control problems and management policies in health systems: application to intensive care units," *Flex. Serv. Manuf. J.*, vol. 28, no. 1–2, pp. 62–89, Jun. 2016, doi: 10.1007/s10696-014-9209-8.
- [43] W. Abo-Hamad and A. Arisha, "Simulation-based framework to improve patient experience in an emergency department," *Eur. J. Oper. Res.*, vol. 224, no. 1, pp. 154–166, Jan. 2013, doi: 10.1016/j.ejor.2012.07.028.
- [44] R.-C. Lin, M. Y. Sir, and K. S. Pasupathy, "Multi-objective simulation optimization using data envelopment analysis and genetic algorithm: Specific application to determining optimal resource levels in surgical services," *Omega*, vol. 41, no. 5, pp. 881–892, Oct. 2013, doi: 10.1016/j.omega.2012.11.003.
- [45] S. C. Tsai, Y. Yeh, and C. Y. Kuo, "Efficient optimization algorithms for surgical scheduling under uncertainty," *Eur. J. Oper. Res.*, vol. 293, no. 2, pp. 579–593, Sep. 2021, doi: 10.1016/j.ejor.2020.12.048.
- [46] W. Vancroonenburg, P. De Causmaecker, and G. Vanden Berghe, "Chance-constrained admission scheduling of elective surgical patients in a dynamic, uncertain setting," *Oper. Res. Health Care*, vol. 22, p. 100196, Sep. 2019, doi: 10.1016/j.orhc.2019.100196.
- [47] P. Landa, R. Aringhieri, P. Soriano, E. Tànfani, and A. Testi, "A hybrid optimization algorithm for surgeries scheduling," *Oper. Res. Health Care*, vol. 8, pp. 103–114, Mar. 2016, doi: 10.1016/j.orhc.2016.01.001.
- [48] T.-L. Chen and C.-C. Wang, "Multi-objective simulation optimization for medical capacity allocation in emergency department," *J. Simul.*, vol. 10, no. 1, pp. 50–68, Feb. 2016, doi: 10.1057/jos.2014.39.
- [49] Z. Aghaabdellahian and M. Bijari, "Bed management considering bed-blocking and elective patient admissions using simulation optimisation," *Int. J. Simul. Process Model.*, vol. 15, no. 3, p. 278, 2020, doi: 10.1504/IJSPM.2020.107330.
- [50] M. Cildoz, A. Ibarra, and F. Mallor, "Accumulating priority queues versus pure priority queues for managing patients in emergency departments," *Oper. Res. Health Care*, vol. 23, p. 100224, Dec. 2019, doi: 10.1016/j.orhc.2019.100224.
- [51] X. Fan, J. Tang, C. Yan, H. Guo, and Z. Cao, "Outpatient appointment scheduling problem considering patient selection behavior: data modeling and simulation optimization," *J. Comb. Optim.*, vol. 42, no. 4, pp. 677–699, Nov. 2021, doi: 10.1007/s10878-019-00487-x.
- [52] A. Goienetxea Uriarte, E. Ruiz Zúñiga, M. Urenda Moris, and A. H. C. Ng, "How can decision makers be supported in the improvement of an emergency department? A simulation, optimization and data mining approach," *Oper. Res. Health Care*, vol. 15, pp. 102–122, Dec. 2017, doi: 10.1016/j.orhc.2017.10.003.
- [53] A. Saremi, P. Jula, T. ElMekkawy, and G. G. Wang, "Appointment scheduling of outpatient surgical services in a multistage operating room department," *Int. J. Prod. Econ.*, vol. 141, no. 2, pp. 646–658, Feb. 2013, doi: 10.1016/j.ijpe.2012.10.004.
- [54] Y. Xiao and R. Yoogalingam, "Reserved capacity policies for operating room scheduling," *Oper. Manag. Res.*, vol. 14, no. 1–2, pp. 107–122, Jun. 2021, doi: 10.1007/s12063-020-00172-x.
- [55] Z. Zhang and X. Xie, "Simulation-based optimization for surgery appointment scheduling of multiple operating rooms," *IIE Trans.*, vol. 47, no. 9, pp. 998–1012, Sep. 2015, doi: 10.1080/0740817X.2014.999900.
- [56] J. Song, Y. Qiu, and Z. Liu, "Integrating Optimal Simulation Budget Allocation and Genetic Algorithm to Find the Approximate Pareto Patient Flow Distribution," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 1, pp. 149–159, Jan. 2016, doi: 10.1109/TASE.2015.2424975.
- [57] C. J. Jagtenberg and A. J. Mason, "Improving fairness in ambulance planning by time sharing," *Eur. J. Oper. Res.*, vol. 280, no. 3, pp. 1095–1107, Feb. 2020, doi: 10.1016/j.ejor.2019.08.003.
- [58] H. Mousavi, S. A. Darestani, and P. Azimi, "An artificial neural network based mathematical model for a stochastic health care facility location problem," *Health Care Manag. Sci.*, vol. 24, no. 3, pp. 499–514, Sep. 2021, doi: 10.1007/s10729-020-09533-1.
- [59] R. D. Baker, "Improving appointment scheduling for medical screening," *IMA J. Manag. Math.*, vol. 13, no. 4, pp. 225–243, Oct. 2002, doi: 10.1093/imaman/13.4.225.
- [60] R. Braune, W. J. Gutjahr, and P. Vogl, "Stochastic radiotherapy appointment scheduling," *Cent. Eur. J. Oper. Res.*, vol. 30, no. 4, pp. 1239–1277, Dec. 2022, doi: 10.1007/s10100-021-00762-5.
- [61] M. Deceuninck, S. De Vuyst, and D. Fiems, "An efficient control variate method for appointment scheduling with patient unpunctuality," *Simul.*

- Model. Pract. Theory*, vol. 90, pp. 116–129, Jan. 2019, doi: 10.1016/j.simpat.2018.11.001.
- [62] A. K. Dogru and S. H. Melouk, “Adaptive appointment scheduling for patient-centered medical homes,” *Omega*, vol. 85, pp. 166–181, Jun. 2019, doi: 10.1016/j.omega.2018.06.009.
- [63] W. Fan, L. J. Hong, and X. Zhang, “Distributionally Robust Selection of the Best,” *Manag. Sci.*, vol. 66, no. 1, pp. 190–208, Jan. 2020, doi: 10.1287/mnsc.2018.3213.

Table 1 : Overview of References, Target Objectives, and Objective Function Types

Ref	Objectives						Approach		
	Waiting times	Patients related costs	Overtime costs	Profit	Resources Utilization	Others	Weighted	Pareto based	Stage-based
[36]		X	X				X		
[37]	X				X		X		
[38]	X		X		X			X	
[39]		X	X				X		
[40]	X				X		X		
[22]		X			X		X		
[41]	X				X		X		
[17]			X	X	X			X	
[42]	X	X						X	
[43]				X		X	X		
[44]						X		X	
[25]	X					X		X	
[45]			X			X			X
[46]	X				X	X			X
[47]		X			X				X
[48]	X				X			X	
[49]					X	X		X	
[50]	X					X	X		
[51]					X	X		X	
[44]	X				X	X	X		
[38]	X				X			X	
[52]						X		X	
[53]	X					X	X		
[54]						X	X		
[55]			X		X		X		
[56]	X				X	X		X	
[57]						X		X	

Ref	Objectives						Approach		
	Waiting times	Patients related costs	Overtime costs	Profit	Resources Utilization	Others	Weighted	Pareto based	Stage-based
[58]	X	X	X					X	
[59]	X				X		X		
[60]	X				X		X		
[61]						X	X		
[62]	X	X			X		X		
[63]						X	X		

Table 15 : Expected Standard Deviations for Each Upper Bound, ESI, and Budget

		Model 1	Model 2 with 5% decrease in initial $UBSD_k$		Model 2 with 10% decrease in initial $UBSD_k$		Model 2 with 15% decrease in initial $UBSD_k$	
		$E_R(SD_k(X, w))$	$UBSD_k$	$E_R(SD_k(X, w))$	$UBSD_k$	$E_R(SD_k(X, w))$	$UBSD_k$	$E_R(SD_k(X, w))$
<b>B =90</b>	<b>ESI 1</b>	43.284	41.1198	40.119				
	<b>ESI 2</b>	76.971	73.12245	72.96				
	<b>ESI 3</b>	112.148	106.5406	103.96				
	<b>ESI 4</b>	199.968	189.9696	186.358				
	<b>ESI 5</b>	342.839	325.69705	303.193				
<b>B =110</b>	<b>ESI 1</b>	40.044	38.0418	36.348				
	<b>ESI 2</b>	71.951	68.35345	62.491				
	<b>ESI 3</b>	98.23	93.3185	87.711				
	<b>ESI 4</b>	182.353	173.23535	150.033				
	<b>ESI 5</b>	322.142	306.0349	206.47				
<b>B =130</b>	<b>ESI 1</b>	36.066	34.2627	31.701	32.4594	31.701		
	<b>ESI 2</b>	65.38	62.111	56.756	58.842	56.756		
	<b>ESI 3</b>	85.289	81.02455	74.252	76.7601	74.252		
	<b>ESI 4</b>	146.616	139.2852	118.25	131.9544	118.25		
	<b>ESI 5</b>	201.285	191.22075	165.572	181.1565	165.572		
<b>B =150</b>	<b>ESI 1</b>	30.659	29.12605	24.984	27.5931	24.984	26.06015	24.275
	<b>ESI 2</b>	50.893	48.34835	44.412	45.8037	44.412	43.25905	40.476
	<b>ESI 3</b>	72.426	68.8047	55.523	65.1834	55.523	61.5621	55.818
	<b>ESI 4</b>	118.96	113.012	94.88	107.064	94.88	101.116	96.288
	<b>ESI 5</b>	191.034	181.4823	125.447	171.9306	125.447	162.3789	114.762
<b>B =170</b>	<b>ESI 1</b>	21.979	20.88005	20.774				
	<b>ESI 2</b>	41.556	39.4782	29.916				
	<b>ESI 3</b>	53.033	50.38135	50.299				
	<b>ESI 4</b>	84.627	80.39565	69.931				
	<b>ESI 5</b>	113.338	107.6711	93.235				