January 2013

A Methodology for Scheduling Operating Rooms Under Uncertainty

Marbelly Paola Davila
University of South Florida, marbellydavila@yahoo.com

A Methodology for Scheduling Operating Rooms Under Uncertainty

by

Marbelly Paola Dávila

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Industrial and Management Systems Engineering
College of Engineering
University of South Florida

Major Professor: Grisselle Centeno, Ph.D.
Peter Fabri, Ph.D.
Katherine De Laurentis, Ph.D.
Kingsley Reeves, Ph.D.
Michael Weng, Ph.D.

Date of Approval:
April 9, 2013

Keywords: Simulation, Stochastic Procedure Duration, Data Mining, Support Vector Regression, Surgical Data

DEDICATION

To my beloved mother, *Maria Beatriz Dávila*, her support, encouragement, and constant love have sustained me throughout my life.

To *Jesus Alejandro* and *Maurizio Antonio*, my little angels, an unwavering source of love and motivation.

To my dear sisters, *Marggiory* and *Margondhy*, for their unconditional love and support.

ACKNOWLEDGMENTS

First and foremost, I want to thank my advisor Dr. Grisselle Centeno. It is an honor for me to be her first *female* Ph.D. student. Her support and encouragement during these years have carried me over my pursuits of the doctoral degree. I deeply admire her ability to balance research interests and personal pursuits.

Following my advisor, I want to specially recognize Dr. Peter Fabri's participation and guidance. He was instrumental in helping me not only to define the scope of the project but also to access real data for the model validation. I have grown so much, both professionally and personally, throughout our interactions.

I would also like to thank Dr. Katherine De Laurentis, Dr. Kingsley Reeves, Dr. Michael Weng, for serving in my dissertation committee and for supporting my research initiatives.

Thanks to Dr. Alfredo Santana Reynoso, who introduced me to the "simulation world" and collaborated with me on the implementation of the simulation model.

A special thanks to my mother, Maria Beatriz Dávila; words cannot express how grateful I am to her for all the sacrifices she has made on my behalf. Her prayers have sustained me thus far.

Lastly, I am in debt with Faisal (Dr. Zeidan), my mentor and best friend. Vida: "I deeply appreciate your unconditional help. You are always in my heart. Thanks for constantly being by my side even from afar. "

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

An operating room (OR) is considered to be one of the most costly functional areas within hospitals as well as its major profit center. It is known that managing an OR department is a challenging task, which requires the integration of many actors (e.g., patients, surgeons, nurses, technicians) who may have conflicting interests and priorities. Considering these aspects, this dissertation focuses on developing a simulation based methodology for scheduling operating rooms under uncertainty, which reflects the complexity, uncertainty and variability associated with surgery.

We split the process of scheduling ORs under uncertainty into two main components. First, we designed a research roadmap for modeling surgical procedure duration (from incision to wound closure) based on the surgery volume and time variability. Then, using a real surgical dataset we modeled the procedure duration using parametric and distribution-free predictive methods. We found that Support Vector Regression performs better that Generalized Linear Models increasing the prediction accuracy on unseen data by at least 5.5%.

Next, we developed a simulation based methodology for scheduling ORs through a case study. For that purpose, we initially built one day feasible schedules using the 60th, 70th, 80th, and 90th percentiles to allocate surgical procedures to ORs using four different allocation policies. We then used a discrete event simulation model to evaluate the robustness of these initial feasible schedules considering the stochastic duration of all the OR activities and the arrival of surgical emergency cases. We found that on average

elective waiting almost doubled the time for the emergency cases. In addition, we observed that there is not a clear effect of how being more conservative in scheduling within each scheduling policy impacts the elective waiting times. By contrast, there is a clear effect of how the scheduling policy and scheduling percentile impact the emergency waiting times. Thus, as we increase the percentile, the waiting times for emergency cases remarkably increases under half of the scheduling policies but reflects a lesser impact under scheduling the other half. OR utilization and OR overtime in a "virtual" eight operating room hospital fluctuate between 67% and 88% and 97 and 111 minutes respectively. Moreover, we noticed that both performance metrics depend not only on the levels of the scheduling policy and scheduling percentile but also are strongly affected by the increase of the emergency arrival rate.

Finally, we fit a multivariate-multiple-regression model using the output of the simulation model to assess the robustness of the model and the extent to which these results can be generalized to a single, aggregate hospital goal. Further research should include a true stochastic optimization model to integrate optimization techniques into simulation analysis.

1. INTRODUCTION

Over the last decade, healthcare providers in the United States (U.S.) have been facing the pressure of reconciling the increasing demand for delivering high quality services with the progressive reduction of government reimbursement. At the same time, the recent push for healthcare reform in the U.S brings new pressure into the healthcare system, as it is expected that by 2019 thirty-two million Americans will gain health insurance coverage under the Affordable Care Act [1]. In addition to this situation, other global factors such as the aging of the population and greater rates of obesity put an extra burden on healthcare systems [2]. As a result, healthcare providers in the U.S. are focusing on finding "new ways" to optimize the delivery of healthcare services.

A recent paper published by the Commonwealth Fund [3] stressed that U.S. spends far more on healthcare than any other industrialized country in the Organization for Economic for Cooperation and Development community (OECD). In 2009, U.S expenditures related to health accounted for 17% of the nation's Gross Domestic Product (GDP).

More recently, the Center for Medicare and Medicaid Services (CMS) reported that in 2010 the National Health Expenditures (NHE) in U.S. reached the amount of $2.6 trillion, which accounted for almost 18% of the Nation's Gross Domestic Product (GDP) and represented on a per capita basis the amount of $ 8,402 [4]. Furthermore, CMS estimates that healthcare expenditures in the U.S. by 2018 will reach $4.3 trillion; almost twice the amount spent in 2010. It is also projected by the CMS that by 2016 hospital care expenditures will reach $3.5 trillion (79% of the projected total NHE). The magnitude of these projections, has forced healthcare providers to rethinking the way in which they are delivering their services and to reengineer procedures to improve their productivity and reduce costly inefficiencies. Other strategies already implemented to cut operational costs include staff reduction and merging businesses with other healthcare providers [2].

Within a hospital, operating rooms (ORs) have been identified as one of its core financial component. It has been estimated that ORs are responsible for almost 30% of the total hospital expenditures [5] and account for more than 40% of its revenues [6]. Consequently, the operating room is considered to be one of the most costly functional areas within hospitals as well as its major profit center.

According to the National Center for Health Statistics (NCHS) in 2009 an estimated forty eight millions of inpatient surgical procedures were performed in the U.S.[7].

A common practice to schedule ORs is the use of deterministic surgical procedure durations (e.g., average of the last ten cases or surgeon's estimate), average patient flows and average needs [8]. Because of these practices, OR utilization typically falls short of the 80% target [9] and as a result has great financial implications for the hospitals.

It is known that planning and scheduling ORs is a challenging task since it requires the integration and interaction of many agents (e.g., patients, surgeons, nurses, technicians) under an uncertain environment and capacity constraints (e.g., availability of costly technologic equipment) [6, 10].

Both planning and scheduling activities in ORs deal with the allocation of available resources over time to perform a set of surgical procedures. In general, planning is described as "the process of reconciling demand and supply" [2] whereas scheduling deals with the assignment and sequence of tasks to servers.

One of the factors that strongly affects the planning and scheduling of a surgical facility is the presence of uncertainty and variability. Natural variability is omnipresent in healthcare processes and human domains (e.g., arrival of an emergency patient, uncertain duration of a surgical procedure) but artificial variability is the result of the design and implementation of inadequate planning and scheduling policies. In order to reduce the adverse impact of the artificial variability the uncertainty and stochastic nature of healthcare delivery processes must be considered.

1.1 Research Motivation

Despite the vast body of knowledge related to planning and scheduling of ORs [11], there are still issues that have been under addressed in the literature. Specifically, when dealing with the arrival of trauma/emergency patients, OR managers are in need of determining which scheduling policy to use to handle uncertainty. Therefore, there is a clear opportunity to seek for alternatives to address the following two questions: (1) How much OR capacity needs to be reserved to accommodate unplanned demand?; and (2) Is it possible to have an OR ready to accommodate an emergency case within a certain time after its arrival?.

In this dissertation, we present a novel methodology for scheduling ORs under uncertainty, integrating two aspects that are commonly considered independently in most of the literature. These include: (1) the impact of the arrival of trauma/ emergencies that needs to be treated as soon as possible (frequently within 20 minutes or less) and (2) the stochasticity governing the surgical procedure duration.

1.2 Organization of the Dissertation

The organization of this research is as follows: Chapter2 presents background information related to operating room's planning and scheduling. Chapter 3 summarizes the literature concerning ORs scheduling with regard to the operating room times and patient status. Based on the gaps found in the academic literature, the research objectives and methodology are formulated and presented in Chapter 4. Chapter 5 shows the analysis of an exploratory work conducted using real data from a Level-I Trauma Hospital. The predictive methods and model developed to analyze surgical case duration are presented in Chapter 6. Chapter 7 shows the methodology for scheduling ORs under uncertainty and presents a case study as a validation tool. Chapter 8 discusses the experimental results, and in Chapter 9 conclusions, contributions and, future research direction are presented.

2. THE SURGICAL CASE SCHEDULING PROCESS

Operations can be performed under an elective and non-elective scheme. An elective operation is one in which the date and place of the surgery can be planned in advance; whereas a non-elective operation usually is unanticipated and needs to be performed the same day Sometimes an operation is considered as an urgent case but not as an emergency case. Another classification used to categorize surgical patients is based on their status at the hospital. Inpatients are those patients admitted to the hospital either on or prior to the day of the surgery, they stay in the hospital for a recovery period. By contrast, outpatients are admitted and discharged from the hospital the same day of surgery.

In general, the term “scheduling” is defined as the allocation of resources to jobs or customers [12]. Within healthcare, a “surgical schedule” consists of a timetable for a particular time horizon (day, week, or month) that specifies the OR in which elective surgical procedures will be performed, the processing order, and their planned start times.

Commonly, the surgical case scheduling process is performed by an OR scheduler, who follows specific planning and scheduling rules during the booking process. Figure 1 depicts the general process used by the OR scheduler during the surgical case scheduling process. During the first stage, the scheduler builds a preliminary OR schedule allocating surgical procedures to ORs, fixing the date and providing surgeons with a tentative start time. During the second stage, the scheduler builds a final 1-day OR schedule that specifies the room number and the tentative starting time. Individual surgeons, groups, or surgical specialties could initiate this process as far as six months prior to the planned day of surgery.

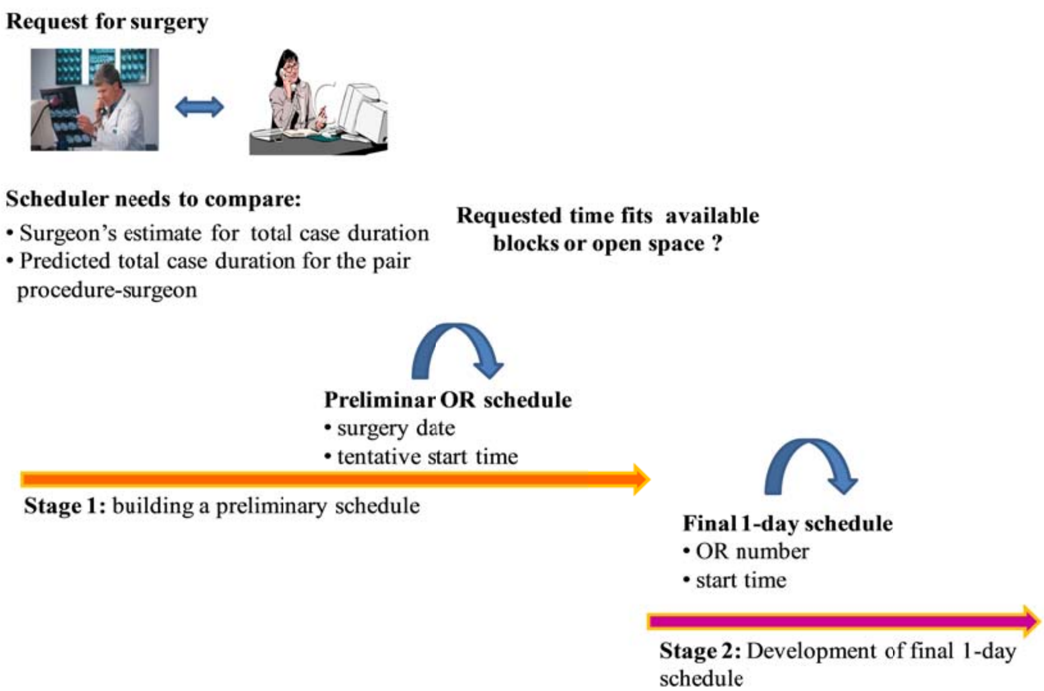


Figure 1 Surgical case scheduling process

2.1 Booking Systems

There are three well known booking schemes under which elective surgical procedures are assigned to ORs: block-booking, open-booking, and modified-block booking [6].

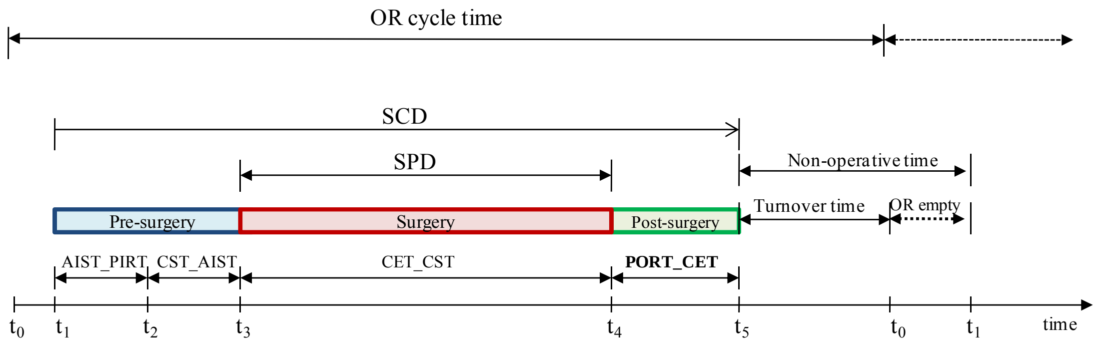
Under a block-booking scheme surgeons are allowed to schedule surgical procedures to their allotted block of time (pre-agreement with the OR department) if and only if the surgical procedure can be completed within the reserved block-time on the specific day of surgery. Otherwise, surgeons need to request an allowance for overbooking the surgical case.

In the open-book, surgeons compete for a space on a first come-first served basis (FCFS) until the pre-determined OR capacity is reached.

The modified-block booking scheme presents a more flexible system by allowing surgeons or surgical specialties to share or release their assigned blocks if they anticipate a conflict that may delay a surgery. This will open the opportunity to other procedures to be performed and will prevent penalties associated with late cancellations or no-shows. However, this must be done within a pre-established lead time prior to the day of surgery.

2.2 Operating Room Cycle Time Components

An operation is a procedure that involves the completion of several activities within an OR. In general, these activities are performed in a specific order and can be grouped in three interconnected phases or stages: pre-surgery, surgery and post-surgery. A typical surgery OR cycle time is depicted in Figure 2. The surgical procedure duration (time elapsed from incision to wound closure) is the amount of time during which the actual surgical procedure occurs. This corresponds to the defined Current Procedure Terminology (CPT) codes. Possible patient arrival delays (idle OR), times related to anesthesia induction, patient discharge and turnover intervals are also shown in Figure 2. The task of cleanup and preparing the OR for the next patient is carried out during the turnover time interval and is typically considered as a non-operative task.



Legend:

SCD: surgical case duration (patient in-patient out)
SPD: surgical procedure duration (incision-closure)
t_0 = Room ready time (RRT)
t_1 = Patient in room time (PIRT)
t_2 = Anesthesia induction start time (AIST)
t_3 = Case start time (CST)
t_4 = Case end time (CET)
t_5 = Patient out of room time (PORT)

Figure 2 A typical surgery OR cycle

Generally, OR schedules are assembled considering the expected values of surgical case duration. It is a common practice among hospitals to either use surgeons' estimates or departmental means during the process of allocating surgical procedures to ORs. A reliable OR schedule can only be assembled when accurate estimates about the time needed to perform an operation are available. Otherwise, operations that take significantly longer or shorter than predicted will increase the chance of having excessive OR overtime or higher rates of underutilization.

The problem of surgery planning and scheduling is complex due to the amount of sources of uncertainty. For instance, it is almost impossible to predict the exact duration of a surgical procedure or to determine in advance the arrival time of the next surgical emergency case. To overcome this, managers are likely to implement heuristic strategies or policies, such as reserving OR capacity and using dedicated ORs to absorb the effect of unpredictable events[13].

In addition to the arrival of emergencies (unplanned events), the uncertainty in the completion time of a surgical procedure strongly affects OR schedules. Operations completion times are stochastic by nature and depend on several factors and events that interconnect before and during the actual procedure.

2.3 Operating Room Performance Indicators

Efficient OR schedules are measured by their ability to execute elective operations as planned and, at the same time, being flexible to incorporate unplanned demand (emergency cases) [5].

Several competing criteria are involved in the process of evaluating the performance of the ORs schedules. Among the most common performance indicators are patient throughput, OR utilization, patient's waiting time, OR team waiting time, surgeon idle time, surgery cancellation rate, patient deferral, and satisfaction level of patients/surgeons/staff. This research focuses in the performance indicators that allow us to infer about the degree of robustness of OR schedules to absorb unplanned demand while incorporating surgery duration as a stochastic component.

2.4 Key Considerations in Scheduling Elective Surgical Cases

A common approach to estimate the duration of a surgical case is to use the average duration of the same procedure using historical data. However, since case duration in this research is treated a random variable, the variation around their expected value may cause delays in the start time of subsequent cases or even worst yield to the cancellation of delayed cases [14]. Furthermore, the arrival of unplanned surgical cases (emergency cases) may trigger the aforementioned scenario.

There are two approaches commonly used to reserve OR capacity when considering random arrivals of surgical emergencies [15, 16]: (1) use of dedicated OR sand, (2) use of planned slack time or slack capacity in elective ORs. In the first approach, when a surgical emergency case arrives it is assigned to the dedicated OR. If a second emergency arrives while the dedicated room is occupied with the first emergency, it will be assigned to the next available OR. In the second approach, for each elective case scheduled, an additional buffer based on a pre-established percentage by the hospital is added. That additional capacity is used to overcome any unanticipated emergency that may arrive [16].

In the United States, there is special type of hospital in which it is mandatory the use of dedicated ORs. After meeting specific criteria and passing a site review, these hospitals receive a certification by the Verification Review Committee of The American College of Surgeons to operate as trauma centers "level I". Trauma centers are defined as "hospitals that have resources and equipment needed to help care for severely injured patients" [17], they are classified in various categories, among which level I is the highest.

3. LITERATURE REVIEW

This chapter presents the state of the art in OR scheduling related to the various components of the surgery's cycle time (operative and non-operative) and the way that models handle the arrival of unplanned emergency cases. The body of knowledge associated with operative cycle time components, that is, surgical case duration and procedure duration, is summarized in Section 3.1. Section 3.2 presents an analysis of non-operative times or turnover times. Finally, a review on how mathematical models incorporate various preferences regarding the type of surgeries considered (elective vs. non-elective) is presented in Section 3.3. More extensive reviews are provided in [11, 18].

3.1 Prediction of Surgical Case Duration and Procedure Duration

Modeling procedure duration has been a topic of interest for operations management and the medical community. Better planning can be achieved when reliable predictions for the time required to complete elective operations are available. When an operation takes longer than predicted, subsequent operations are delayed or even postponed to a later day. When the actual time is shorter than predicted and planned, the operating rooms are unused and considered unproductive. Furthermore, in the absence of reliable predictions, the use of advanced planning techniques is futile. In some hospitals, the surgeons provide, based on their experience, an estimate on the time required to complete the procedure. Other hospitals use historical times (for example, the average of last five surgeries performed for a given procedure by a surgeon) to estimate procedures time [19].

To deal with the uncertainty related to the duration of procedures some researchers have used stochastic optimization models [20]. Alternatively, other authors incorporate a planned slack and model the Master Surgical Scheduling (MSS) as a mathematical program containing probabilistic constraints [21].

The existent literature can be classified based on the approach used to predict procedure time. These approaches incorporate one or more of the following information: patient characteristics, surgical team characteristics, CPT codes (five-digit number that represents a set of medical, surgical or diagnostic service) and the surgeon estimates. In general, three groups (based on the approached used) have been identified in the literature. The first one relies on linear regressions for estimating procedure times and identifying the crucial factors that affect variability in operations. The second group studies the fitness of known distributions, notably the normal distribution and the lognormal distribution, for the purpose of predicting surgical case duration (SCD). The third group attempts to incorporate a new line of approach based on database warehouse and knowledge discovery. Literature related to each of these approaches is discussed in the following subsections.

3.1.1 Linear Regression Models

Stepaniak et al. [22] performed an empirical analysis to quantify the effect of surgeon factors such as age, gender, experience, and team composition on total case durations (defined as the time from entry into until leaving the OR). Also, they studied if the combination of type of surgical procedure (main CPT code) and anesthesia type have any effect on case duration. The effect of the aforementioned factors was estimated for over 30 different types of operations in two hospitals, by means of one-way analysis of variance models for logarithmically transformed case durations. They concluded that total case duration depends on the type of operation (CPT-anesthesia combination) and surgeon specific factors. In particular, they found that for complex operations the surgeon learning curve and composition of the surgical team factors have a remarkable effect.

Eijkemans et al. [14] developed a model to predict surgical case duration based on factors including characteristics of the surgical team, type of surgery, patient characteristics and surgeon estimate of case duration. The outcome predicted was total-case-duration (patient in - patient out). They found that the factors associated with type of surgery and team composition had the largest predictive effect, whereas patient characteristics for some procedure or operations had a modest effect on surgical case duration. That is, operations were shorter for patients older than 60 years, and higher body mass index was associated with longer OR times.

Li et al. [23] developed a linear regression model to predict surgical case duration (from incision to wound closure) based on a single factor: the CPT code which is specified for each surgery. They developed a regression-based model using CPT codes as the unique explanatory variables when predicting surgical case durations. The model presented can incorporate up to eight CPT codes.

3.1.2 General Distributions Models

Strum, May and Vargas [24] used a large set of real data to tested how well the lognormal and normal distribution fit the surgical case duration and total case duration. They concluded that the use of the lognormal model outperformed the normal model when predicting surgical case duration for the factor CPT-anesthesia combination. In practice, they suggested that the selection of a model should be based on an examination of a normal probability plot (a possibly subjective procedure) in conjunction with a formal goodness-of-fit test (a more objective measure) to avoid unjustified model rejections. Finally, they strongly suggested that statistical tools such as regression and analysis of variance should be applied to the log transforms of the procedure times.

Stepaniak et al. [25] compared the fit of the normal distribution with 2-parameter and 3- parameter lognormal distributions for total case duration of a range of CPT-anesthesia combinations, including surgeon effects. Total case duration is defined as time elapsed from when the patient enters the OR until the patient leaves the OR. Only procedures or operation types with frequencies greater than 10cases, and one CPT-anesthesia combination segmented by surgeon, were considered. They found that the percentage of cases fitting the normal model and 2 or 3-parameter lognormal models are higher for surgical case duration (defining as the time used by the surgeon) than those for total case duration (defining as the time passing from entry into the operating room to leaving the OR). Also, they reported that the 3-parameter lognormal distribution provides the best result for the total case durations of one CPT-anesthesia combination segmented by surgeon, with a fit for almost 90% of the total cases.

3.1.3 Intelligent Based Models

Recent tendencies in the literature reflect a new line of work which attempts to incorporate the use of intelligent based models and data mining techniques such as rough sets, artificial neural networks and fuzzy inference systems to predict procedure times.

Combes et al. [26] presented a methodology for planning surgery in ORs based on data warehousing and knowledge discovery in database approaches. In the context of implementing a knowledge extraction process, they experimented with a series of prediction models based on data mining tools to forecast the total case duration. The models were based on variables related to patients' factors (administrative data, medical history, etc.) and also, to the procedure type (surgeon, type of anesthesia, etc.). Their models reported unsatisfactory results; the authors believe that their grouping of operations based on diagnoses rather than procedures types (CPT codes) in their model was the main reason for inaccuracy.

Devi P., Rao S. and Sangeetha S. [27] developed a model for forecasting total case duration considering only three different ophthalmologic operations. Among the variables considered were surgeon's experience, anesthesiologist's experience, and type of anesthesia, etc. The prediction of total case duration was done using three techniques:

Adaptative Neuro Fuzzy inference Systems, Artificial Neural Networks and Multiple Linear Regression Analysis.

3.2 Analysis of Non-Operative Times

Within an OR environment, the non-operative time is defined as the interval between two procedures [28]. When patients arrive on time (i.e. just after the OR is ready to receive the next patient), non-operative time corresponds to turnover times (TT). Turnover time is the interval of time devoted to tasks such as patient preparation for transportation to another healthcare unit (PACU or ICU), clean up and/or set up. In the literature, turnover time is also the time elapsed from patient-out-of-room to next patient-in-room. In that case, TTs could reflect any possible delay that may occur before the patient enters the OR (e.g., patient transportation delay).

Seim et al. [29] applied statistical process control as a tool for monitoring non-operative times. They applied this technique to assess the non-operative time performance between successive cases for same surgeon following himself in a experimentally controlled OR. The modified operating room implemented patient care pathway in order to improve the throughput by reducing non-operative time. The authors tested the efficiency of statistical process control when trying to detected reductions in non-operative time.

Marshall et al. [30] developed a group of standardized turnover charts for a specialized OR to balance the work among the OR staff and to guide the order in which each individual should perform specific turnover tasks. They also defined which type of turnover tasks can be started before the patient leaves the OR. A 45% reduction in non-operative time, as a result of the implementation of standardized turnover charts, was reported.

Krupta and Sathaye [28] presented a literature review on the effects of introducing operational improvements in the areas near to the OR to reduce turnover times. The methods used to reduce non-operative times were classified into five categories: process modifications, addition of staff, technology, facilities, and delay elimination.

Stepaniak et al. [31] studied the impact of scheduling similar consecutive cases on the turnover time, surgical case duration, and total case duration. They hypothesized that when a fixed OR team works on similar consecutive cases, turnover times will significantly decrease. Patients were assigned randomly to the study or control group for two types of operations or procedure types. For one of the operations, they found a considerably lower preparation time and a shorter case duration time in the study group compared to the control group. For the other operation being considered, only preparation time was significantly lower in the study group as compared to the control group. For both procedures there was a considerable decrease in the duration of turnover time.

3.3 Patient Status

Patient status refers to the type of patient who needs to undergo surgery. If the demand for health services can be planned and scheduled in advanced, it is considered an elective procedure. However, if the patient needs to undergo surgery immediately given an emergency, or during the same day (urgencies) then they are considered non-elective procedures.

The management of the non-elective patients is particularly difficult since their arrival is inherently uncertain and the speed of intervention is critical to the patient's potential for survivability and recovery. Such cases are not scheduled in advanced, but must be accommodated along with the cases that are scheduled on any given day.

Since there are several different environments in which surgical services are delivered, some hospitals reserve one or more operating rooms for emergencies; whereas in other, slack time is spread across multiple operating rooms to accommodate non-elective patients [15].

A recent comprehensive literature review [11] highlighted that the literature on elective patient planning and scheduling is rather vast compared to the non-elective counterpart. This study identified that the impact of planning and scheduling non-elective patients is hardly ever studied without the incorporation of elective patients.

Lamiri et al. [32] compared several optimization methods for elective surgery planning. The planing problem was considered as a stochastic optimization problem in order to minimize expected overtime costs and patient's costs. They assumed that surgical procedure times were known and deterministic.

Zhang et al. [33] developed a mixed integer programming model to allocate elective patients to operating rooms. A simulation model was used to assess the performance of the OR schedule. The methodology was illustrated through a case study. In their analysis the average elective waiting time was reduced through efficient allocation of OR time. In addition, they highlighted that the schedule templates generated by the optimization model, performs poorly under conditions of uncertainty.

Van Houdenhoven et al. [34] developed a mathematical model to investigate the association between OR utilization, the patient mix and overtime. Using statistical techniques, the association between the required reserved capacity and the acceptable risk of overtime and the variability of the case mix was established. The surgical case duration for all the procedures was known in advance. In addition, they assumed an identical average and standard deviation for the duration of all the cases per surgical department.

Bowers and Mould [35] used simulation to explore the balance between maximizing the utilization of ORs, minimizing overtime, and maintaining a reasonably quality of care. The simulation model was developed to examine a policy that included elective and non-elective patients. Their

Mehdi et al. [36] proposed and compared several optimization methods for the elective surgery planning problem when OR capacity is shared by elective and non-elective patients. The planning problem was formulated as a stochastic optimization problem. An optimal solution method combining the Monte Carlo sampling technique and mixed integer programming was presented. Some heuristic methods were also proposed for the planning problem. Optimal and heuristic methods were evaluated through numerical experiments.

Wullink et al. [15] showed, used a discrete event simulation model, to show that that reserving capacity for non-elective cases in ORs outperforms the policy of using dedicated ORs. They found that the amount of overtime and the OR utilization considerably improved when the reserved capacity was spread over multiple ORs.

Finally, Table 1 shows the most relevant contributions in the last decade to the prediction of procedure duration, analysis of non-operative times, and patient's status.

3.4 Limitations of Current Literature

The majority of the papers published assume that procedure times follow the lognormal distribution. There have been limited attempts to validate this assumption. Other researchers have proposed the use of linear models that consider one or multiple predictors such as surgeon's estimates or CPT codes to predict surgical case duration. However, to the best of our knowledge, there have been limited attempts to use data mining techniques or distribution-free models to create a general predictive methodology that could effectively extract information from multiple significant factors.

Despite the potential impact on the accuracy to execute the planned schedule, the uncertainty related to emergency case arrivals is generally ignored by most researchers. Moreover, when uncertainty is considered, most models assume that either the total operating room capacity is devoted to a single patient class (i.e. inpatient, outpatient, elective, emergency, etc.) or that the emergency cases will be exclusively allocated to dedicated operating rooms.

Finally, there have been limited attempts to create decision models to assist the multi-objective decision making faced by OR managers when evaluating the trade-offs between operational objectives and patient satisfaction.

These limitations inspired the research objectives for this dissertation as delineated in the following chapter.

Table 1 Contribution of the published work to the prediction of case duration and procedure duration, analysis of non-operative times and management of patient status

Ref	Title	ORT				PC/PE	SE	AE	ORTE	CPT		Strengths(SS) / Weakness(WK)
(Year)		OPT		NOT						S	A	
		TCD	SCD	TT	OT							
(2010)	Modeling and prediction of surgical procedure times	X					X	X	X			**SS:** This model can be applied to any surgical service within a hospital. The authors highlighted that predictions are more accurate when ANOVA is applied after the lognormal transformation of total case duration. **WK:** Requires detailed information about the work rate of the surgical team members
(2010)	Predicting the unpredictable	X				X	X	X	X	X		**SS:** The amount of detail of this model, using operation codes at the lowest level plus operations, team, and patient characteristics, allows for operational planning of operating rooms considering multiple variables
(2010)	Predicting surgical case durations using ill-conditioned CPT code matrix		X							X		**SS:** Allows prediction of surgical case duration based on combinations of multiple CPT codes. **WK:** Does not consider other important covariates together with CPT codes as explanatory variables.
(2000)	Modeling the Uncertainty of Surgical Procedure Times	X	X				X			X	X	**SS:** The prediction model developed help to legitimize the use of log transforms to normalize surgical procedure times before hypothesis testing using linear statistical models or other parametric statistical tests to investigate factors affecting the duration of operations **WK:** This study did not address the issue that when data is originally normal distributed the use of log transformation on it could conduct to a rejection of the null hypothesis that data is normal distributed
(2009)	Modeling Procedure and Surgical Times for Current Procedural Terminology-Anesthesia-Surgeon Combinations and Evaluation in Terms of Case-Duration Prediction and Operating Room Efficiency: A Multicenter Study	X					X	X		X	X	**SS:** One part of this model allows forecasting of total case duration of surgical cases with very few observations. The prediction model integrates previous data information with surgeons' prior guesses. Results were validated in a multicenter study.
(2008)	Using a KDD process to forecast the duration of surgery	X				X	X					**WK:** The inadequate selection of the model's variables affected the forecasting methodology and resulted in unsatisfactory results.
(2010)	Prediction of Surgery Times and Scheduling of Operation Theaters in Ophthalmology Department	X				X	X	X	X		X	**WK:** The validation of these models did not include testing on unseen data.
(2006)	Statistical Process Control as a Tool for Monitoring Non-operative Time			X	X		X					**WK:** difficult implementation due to costly redesign of facilities (operating room, PACU).

ORT= operating room times **OPT=operative times** **NOT=non-operative times** **SCD=surgical case duration**
PD=surgical procedure duration **TT=turnover time** **OT=other times (e.g., delays)** **PC=patient characteristics or status**
SE=Surgeon effect **AE=Anesthesiologist effect** **ORTE= OR team effect** **CPT=Current Procedure Terminology**
S=surgery **A=Anesthesia**

Table 1 Continued

Ref	Title	ORT				PC/PE	SE	AE	ORTE	CPT		Strengths(SS) / Weakness(WK)
(Year)		OPT		NOT						S	A	
		TCD	SCD	TT	OT							
(2006)	Using Lean Methods to improve turnover times			X	X				X			**SS:** Despite this approach requires a rigorous documentation of each process performed within an operating room (for each surgical service) and also, requires a continuous staff member training; it can be implemented in any surgical service.
(2008)	Reducing non-operative time: methods and impact on operating rooms economics			X	X							**WK:** The benefit of introducing the methods proposed (technology, facility redesign, process modification and additional staff) requires an additional resource investment, which could cause difficulty in its implementation
(2010)	Working with a fixed OR team on consecutive similar cases and the effect on case duration and turnover time	X		X			X		X			**SS:** The benefits showed by this approach could be implemented in low volume surgery facilities (ambulatory operations). **WK:** This approach imposes many restrictions to the operating room planning and scheduling, which could lead to schedule infeasibility.
(2007)	Closing emergency rooms improves efficiency	X		X		X						**WK:**(1)SPD for an emergency was based upon one lognormal distribution for all emergency procedures together
(2009)	Optimization methods for a stochastic surgery planning problem	X				X						**WK:** SPD were known and deterministic
(2006)	A Mixed Integer Programming Approach for Allocating Operating Room Capacity	X										**WK:** only one OR can be used to allocate emergencies
	A Norm Utilization for Scarce Hospital Resources: Evidence from Operating Rooms in a Dutch University Hospital	X				X						**WK:** SPD were known and deterministic
(2004)	Managing uncertainty in orthopedic trauma theatres	X										**SS:** through simulation determined the trade-off between OR utilization and overrunning **WK:** SPD were known and deterministic
(2008)	A stochastic model for operating room planning with elective and emergency demand for surgery	X		x		X						**SS:** stochastic emergency demand **WK:** SPD and TT were known and deterministic

ORT= operating room times **OPT=operative times** **NOT=non-operative times** **SCD=surgical case duration**
PD=surgical procedure duration **TT=turnover time** **OT=other times (e.g., delays)** **PC=patient characteristics or status**
SE=Surgeon effect **AE=Anesthesiologist effect** **ORTE= OR team effect** **CPT=Current Procedure Terminology**
S=surgery **A=Anesthesi**

4. RESEARCH OBJECTIVES AND METHODOLOGY

The global research objective of this work is to develop a methodology for scheduling operating rooms under uncertainty; which reflects the complexity, stochasticity, and variability that characterizes surgical facilities. This methodology is described in the next section. To address this challenge, five research goals as listed next were formulated.

- to identify the significant factors impacting surgical procedure times through the analysis and validation of real data (Chapter 5)
- to design a research roadmap that classifies procedures (operations) based on surgery volume and variability, (Chapter 6, section 6.1.1)
- to model surgical procedure times using parametric, and distribution-free predictive methods using real data, (Chapter 7)
- to develop a simulation-based methodology to schedule operating rooms under uncertainty, (Chapter 8) and
- to evaluate the robustness of the simulation model through statistical analysis of the simulation output.(Chapter 9)

4.1 Description of Research Methodology

An extensive review of the literature published until February 2012 was carried out in order to identify methods used for scheduling operating rooms that focus on two main topics: (1) prediction of the surgical case duration and (2) management of the arrival of emergency surgical cases. For that purpose, and also to facilitate the future implementation of the methods presented here, we had imitated as much as possible, the structure of the decisions frequently made in hospitals. Therefore, we split the process of scheduling ORs under uncertainty into two parts as shown in Figure 3. Part 1, refers to the process of modeling surgical case duration, whereas part 2 deals with scheduling multiple ORs under uncertainty. Specifically in part 1, we use an actual surgical dataset to design a model of a roadmap containing the steps to predict the surgical case duration, taking into account the volume and variability of the operations performed at a local hospital in Florida. In part 2, we initially built a one day feasible schedule using the 60^{th}, 70^{th}, 80^{th}, and 90^{th} percentiles of the cases for each procedure to identify the time of the procedure duration. For example, if 10 surgeries were recorded for total-knee-replacement, the durations of these 10 cases were ordered from minimum to maximum time. Then, the time for the procedure on the 60^{th} percentile position (which corresponded to the 6^{th} surgery in this example) was used as the average time for all total-knee-replacement procedures to be allocated in the schedule. This was repeated for the other levels of the percentile control factor.

Given the stochastic duration of all the OR activities as well as the arrival of surgical emergencies, we used a simulation model to assess the robustness of the initial feasible schedules under four scheduling strategies or policies: (1) Random (first come-first served); (2) Best Fit; (3) Best Fit and Shortest Processing Time; and (4) Modified Block Scheduling. Finally, we used a set of performance measures including: waiting time for emergency cases and elective patients, OR overtime, and overall efficiency to assess the robustness of each scheduling policy and to select the best strategy to achieve the hospital's goals.

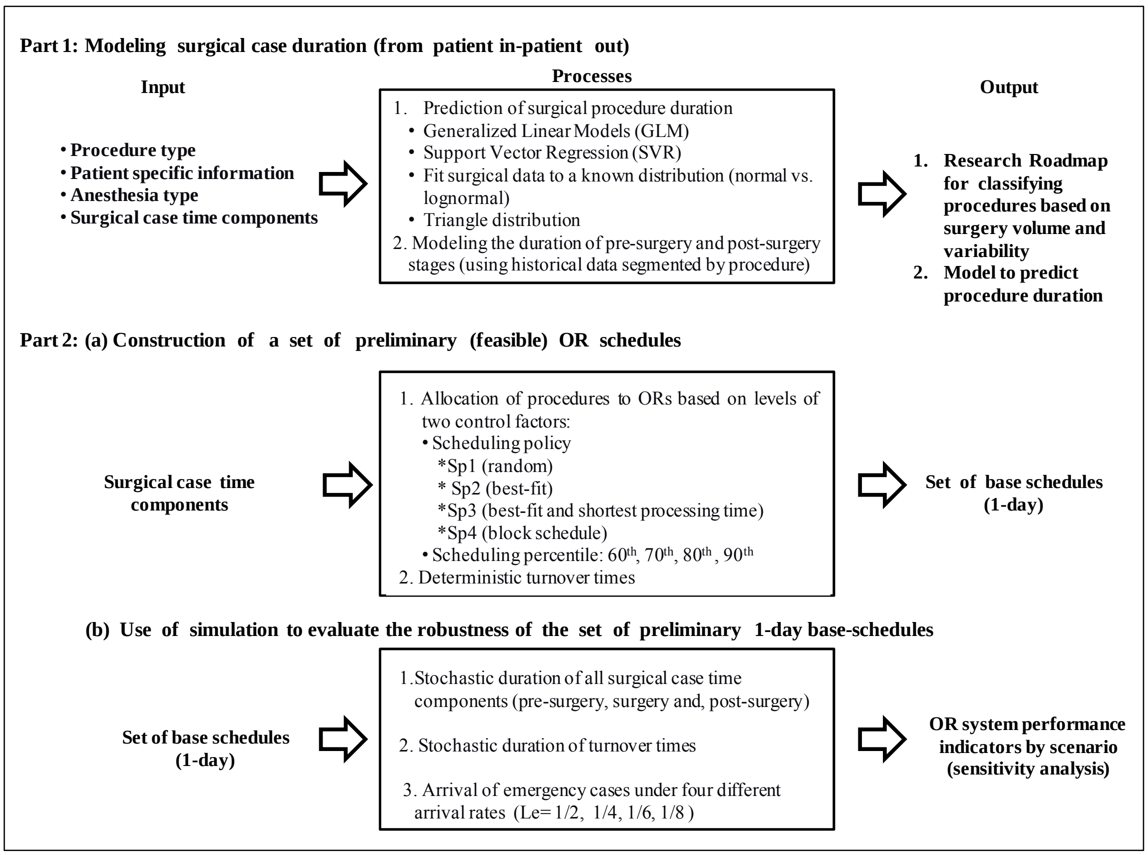


Figure 3 A methodology for scheduling ORs under uncertainty

5. EXPLORATORY DATA ANALYSIS

We received an IRB exemption (University of South Florida Institutional Review Board) which authorized us to work with completely de-identified surgical data belonging to a certified Level-I Trauma Hospital. The dataset consists of 43,679 surgical cases performed from 01/01/2008 to 12/31/2010. It involves 18 ORs, 17 surgical specialties or services, and 314 surgeons. Each surgical record as shown in Figure 4 is composed of 12 variables (attributes) related to the patient, anesthesia, surgical case information, and segmented operation times. Among the different surgical time intervals that can be obtained using the operation times shown in Figure 2, we focus our interest on two surgical time intervals: (1) the surgical case duration, and (2) the surgical procedure duration (incision-closure).

A description of the assumptions and methods used to assess the quality of the data received, and the preprocessing steps followed to clean the database are presented in the next section.

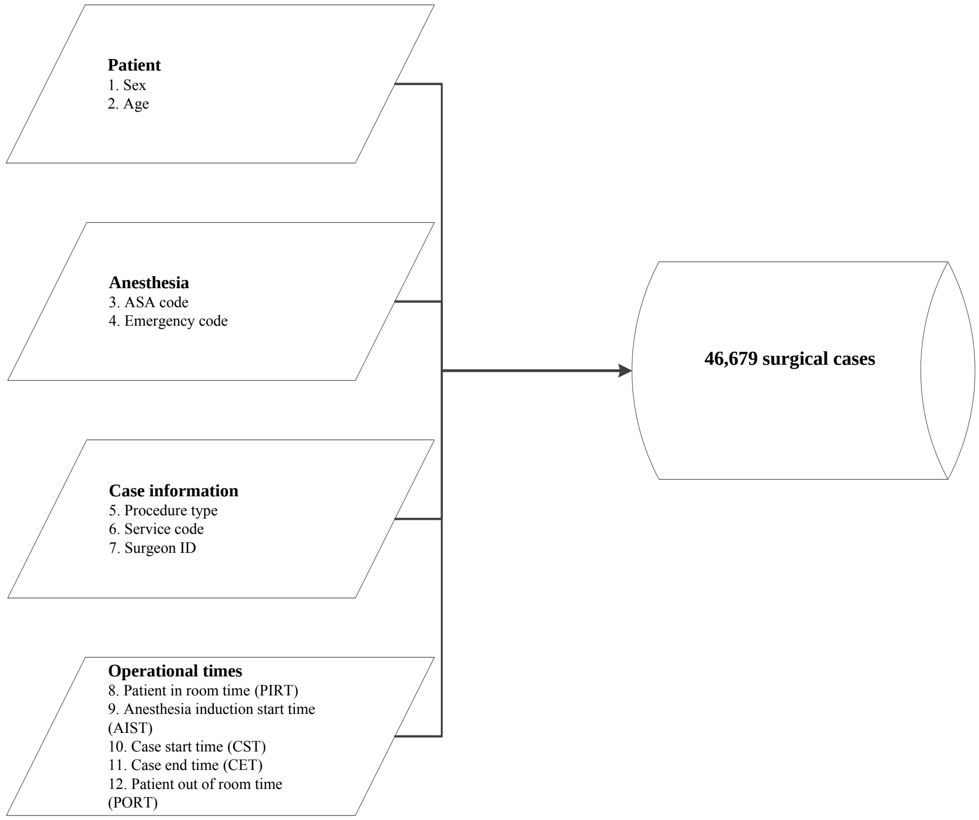


Figure 4 Variables contained in each surgical record in the database

5.1 Assessment of Data Quality

Through data mining methodologies the records received from the hospital have been reclassified and filtered. Thus, during the data quality assessment process each surgical record in the data set has been evaluated in order to plan the data cleansing and data enrichment strategies [37]. As a result of this process, data quality issues in the data set were identified. Specifically, surgical records were filtered and removed as a function of the following irregularities: (1) missing values, (2) inconsistent values, (3) duplicate data, and (4) time components with an anomalous duration (±3σ from the mean was considered a data entry error). Thus, surgical records in the data set that do not meet the inclusion criteria depicted in Figure 5have been discarded (about 29% of the total received) to assure reliable data for model development.

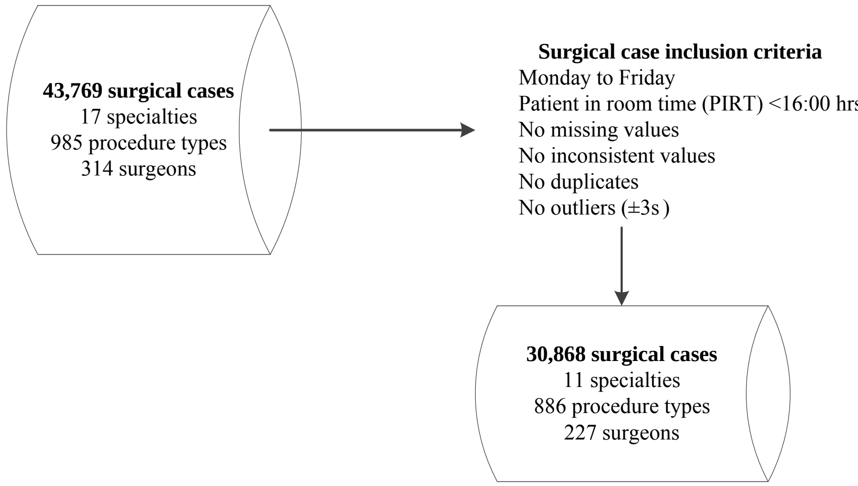


Figure 5 Data cleansing and filtering process

Table 2 contains the frequency distribution of the valid surgical records per surgical service after the cleaning process.

Table 2 Valid surgical records per surgical service in the database

Surgical service	Frequency	%	Cumulative frequency	Cumulative %
Burn/Plastic	2539	8.23	2539	8.23
General	6494	21.04	9033	29.26
Gynecology	700	2.27	9733	31.53
Hand	1316	4.26	11049	35.79
Neurology	4102	13.29	15151	49.08
Ophthalmology	359	1.16	15510	50.24
Oral	86	0.28	15596	50.52
Orthopedics	9823	31.82	25419	82.34
Otolaryngology	4097	13.27	29516	95.62
Urology	1292	4.19	30808	99.81
Vascular	60	0.19	30868	100

5.2 Analysis of Surgical Case Duration Time Components

Before operations can be scheduled their duration must be estimated. Thus, accurate estimates are required to efficiently assign procedures to ORs. For this study, in order to derive conclusions from actual data, we segmented the surgical case duration into four intervals or time components. These time components belong to one of the three surgery stages previously shown in Figure 2. Stage 1(pre-surgery stage) consists of two time components that represent together the time elapsed from when the patient enters the OR until the surgeon starts the actual operation. Thus, stage 1 involves activities associated with patient prepositioning and anesthesia induction. During stage 2, all the activities related with the surgical procedure are performed by the surgeon. Therefore, the duration of stage 2 is commonly known as the surgical procedure duration. Finally, stage 3 represents the interval which starts from the time the surgeon has completed the actual operation until the patient leaves the OR.

5.2.1 Summary Statistics

Prior to the modeling of the surgical case duration, summary descriptive statistics were measured. These quantities are useful to summarize the central tendency and also to quantify variability, detect extreme observations, and check for distribution assumptions [38]. Table 3 shows the summary statistics of each time component being considered. As several other studies have reported [24, 25, 39, 40] surgical times are positively skewed and have positive kurtosis, which indicates that the distributions of these time components have long tails to the right and are peaked in comparison with a normal distribution. Additionally, Table 3shows that all the time intervals have large coefficients of variation (CV), which means that the dataset can be considered to have high variance. In particular, this is significant for the time component related to the surgeon's stage.

Since the aforementioned statistics are not representative of robust measures of normality, we used the Kolmogorov-Smirnov test to assess whether the procedure duration could be considered normally or lognormally distributed after a log transformation. In addition, we have used some graphical tools such as histograms, normal probability plot and quantile-quantile plots, for detecting departures from the normal distribution.

Table 3 Summary statistics of each time component (n=30,868 cases)

Surgery stage	Time components	$\bar{x}$	$\tilde{x}$	σ	Var	CV	sk	kt
Pre-surgery	AIST_PIRT	12.5	11	5.9	35.51	47.62	1.5	3.8
Pre-surgery	CST_AIST	29.9	26	18.5	343.12	61.78	1.0	0.9
Surgery	CET_CST	111.5	89	90.8	8260.7	81.51	2.1	6.9
Post-surgery	PORT_CET	11.7	10	7.1	50.45	60.65	1.4	2.5

$\bar{x}$=mean, $\tilde{x}$=median, σ =standard deviation, Var=variance, CV=coefficient of variation, sk=skweness, kt=kurtosis

5.2.2 Correlation Matrix of Surgical Case Duration Time Components

In addition to the summary statistics, it is advisable to assess the correlation among the different time components that constitutes the surgical case duration before assuming that the time components can be modeled separately from each other. Figure 6 and Table 4 respectively show the values of the correlation matrix and the scatter plot matrix of all time components being considered. As can be seen in Table 4, only the anesthesia induction time component (CST_AIST) is moderately correlated (r^2=0.241) with the surgical procedure duration (CET_CST). Therefore, it can be assumed that the duration of the anesthesia induction interval is related to the complexity of the procedure being performed. As a consequence, we can assume that the time components are conditionally independent when the operation is fixed. Also, the scatter plot matrix shown in Figure 6 reinforces our previous finding that surgical time components are skewed to the right and have positive kurtosis.

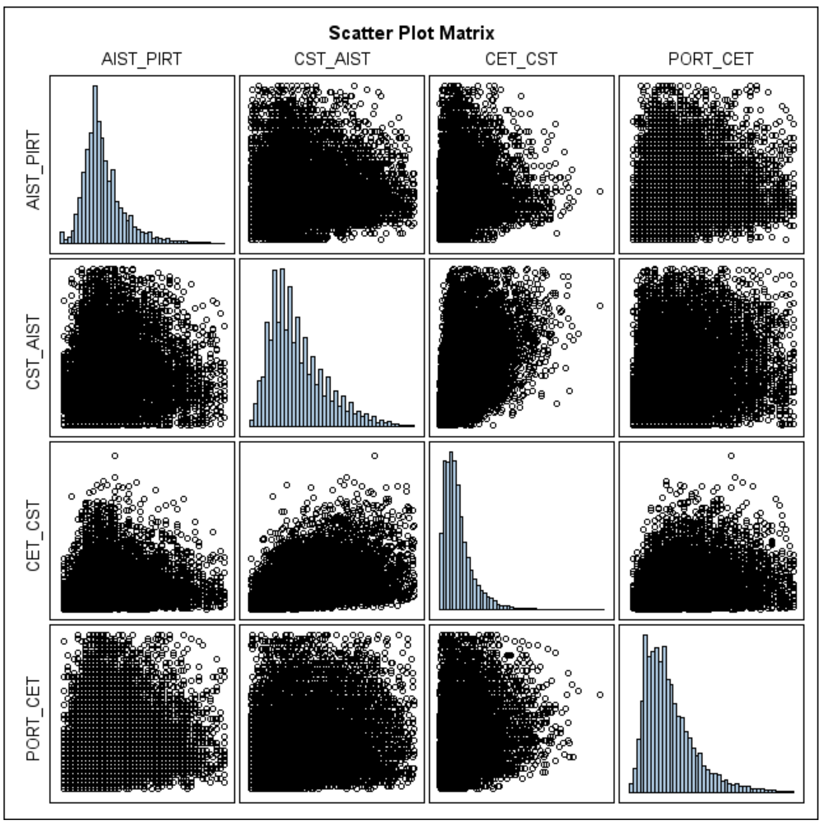


Figure 6 Scatter plot of surgical time components (n=30,868 surgical cases)

Table 4 Pearson correlation coefficients and coefficients of determination r^2 by time component (n=30,868 cases)

Time Component	AIST_PIRT	CST_AIST	CET_CST	PORT_CET
AIST_PIRT	1	0.130 (0.016)	0.133 (0.017)	0.139 (0.019)
CST_AIST	0.130	1	0.491 (0.241)	0.222 (0.049)
CET_CST	0.133	0.491	1	0.274 (0.074)
PORT_CET	0.139	0.222	0.274	1

5.3 Data Preprocessing

During this phase we evaluated which preprocessing steps should be applied to make the data more suitable for modeling purposes. Specifically, since our objective is to predict the duration of the surgical case, we have to convert in numerical variables the nominal and qualitative ordinal variables presented in Table 5.

Table 5 Set of variables contained in each surgical record in their original data format

Variable	Type of attribute	Label
Patient sex	Nominal	P_sex
Patient age	Numerical	P_age
ASA code	Ordinal	ASA_code
Emergency code	Binary	E_code
Service	Nominal	Service
Operation	Nominal	Proc_Id
Surgeon ID	Nominal	MS_id
Patient in room time	Numerical	PIRT
Anesthesia induction start time	Numerical	AIST
Case start time	Numerical	CST
Case end time	Numerical	CET
Patient out of room time	Numerical	PORT

Since most of the predictive models to be used require numerical attributes, we have created a new set of attributes that capture the information contained in the attributes operation type and surgeon ID, using the attribute construction approach suggested in [37]. The new attributes are built following steps shown in Table 6.

Table 6 Procedure for construct the new set of attributes (surgical case score and surgeon speed score)

Let N be the number of operations in the database, L be the number of surgeons in the database, and K be the number of surgical cases (surgical records)in the database

Step 1

For $i = 1, \ldots, N$

Compute the median of the surgical case duration of all the surgical cases by operation i

Call the resulting values SCD_i

Step2

Compute the overall median of all the surgical case durations in the database

Call this value $OSCD$

Step 3

For $i = 1, \ldots . L$

Compute the median of the surgical procedure durations of all the surgical cases performed by surgeon i and call the resulting vales SPD_i

Step 4

Compute the overall median of all the surgical procedure duration in the database

Call this value $OSPD$

Step 5

For $i = 1, \ldots . N$

Create a new variable (attribute) as follows

$$SCS_i = \frac{SCD_i}{OSCD}$$

Call the resulting variable the surgical case duration score

Step 6

For $i = 1, \ldots . L$

Create a new variable (attribute) as follows

$$SSS_i = \frac{SPD_i}{OSPD}$$

Call the resulting variable the surgeon speed score

Step 7

For $K = 1, \ldots, K$

Create two news columns and merge the values of SCS_i and SSS_i based on the corresponding values of the variables operation code and surgeon ID, respectively

The new attributes are discrete numerical variables and can be incorporated into predictive methods, such as generalized linear models and Support Vector Regression. Regarding the remaining nominal and ordinal variables (patient sex, service, and ASA code), we applied a recoding strategy to convert patient sex to a binary variable and ASA code into a discrete, ordered integer. Finally, the operational times were transformed into interval times or time components as described before. Table 7 contains the list of all the variables after applying the preprocessing steps.

Table 7 List of variables after the preprocessing step

Variable	Type of attribute	Label
Patient sex	Binary	P_sex
Patient age	Numerical	P_age
ASA code	Numerical	ASA_code
Emergency code	Binary	E_code
Service code	Numerical	Serv_code
Surgical case duration score	Numerical	SCS
Surgeon speed score	Numerical	SSS
Patient in room time	Numerical	PIRT
Anesthesia induction start time	Numerical	AIST
Case start time	Numerical	CST
Case end time	Numerical	CET
Patient out of room time	Numerical	PORT

6. MODELING SURGICAL CASE DURATION

Several significant studies about estimating the surgical case duration of an elective patient have been published since 1970 as reflected in [41]. In fact, during the last decade, numerous studies [24, 25, 39, 40, 42] have concluded that surgical case duration can be modeled using the lognormal distribution. Therefore, it is a widespread practice to assume that surgical times are lognormally distributed. However, not all previous studies support this conclusion [43]. As a consequence, we present in this section a general process to model and predict the surgical procedure duration using various predictive methods, such as general linear models and Support Vector Regression. Lastly, we have assessed and validated the performance of the models on unseen datasets.

6.1 Predictive Methods used to Model the Surgical Procedure Duration Based on Surgery Volume and Variability

As pointed out in [24], identifying and selecting the most suitable model to predict the surgical case duration and the surgical procedure duration is a crucial step in order to build reliable OR's schedules. Therefore, knowing the distribution of the data is essential for choosing the appropriate statistical method. In the following section, we focus our attention on the prediction of the surgical procedure duration (case start – case end) and consider that the remaining times components could be modeled in the same fashion. Thus, we also considered for the purpose of modeling the surgical procedure duration the use of Support Vector Regression as an alternative method, not requiring any assumption about the underlying probability distribution of the data being considered.

6.1.1 Classification of Surgical Procedures based on Surgery Volume and Variability

Ideally, the prediction of individual surgical procedure duration should be based on cases of the same operation type, performed by the same surgeon [44]. Actually, the majority of the surgical procedures are performed by a small number of experienced surgeons [45] and also represents, at most, 20% of the total procedure types in the database. As a result, it is difficult to obtain a variability assessment in case duration or procedure duration for surgical procedures within and among surgeons. To overcome this limitation, and to facilitate the use of the predictive models, the surgical data are split based on the surgery volume (frequency). As a result, two subsets are obtained as depicted in Figure 7. The first subset shows procedures with 30 or more cases (dataset 1), and the second shows procedures with less than 30 cases (dataset 2). Figure 7 also shows that almost 86% of all the surgical cases belong to the dataset 1, which only contains 20.4% of all the procedures types or operations saved in the database. This specific finding suggests that parametric analysis is a potential modeling technique to be applied only on dataset 1 (181 procedures). Therefore, the remaining surgical cases belonging to dataset 2 (705 procedures) requires modeling techniques that could handle less than 30 observations or surgical cases per operation.

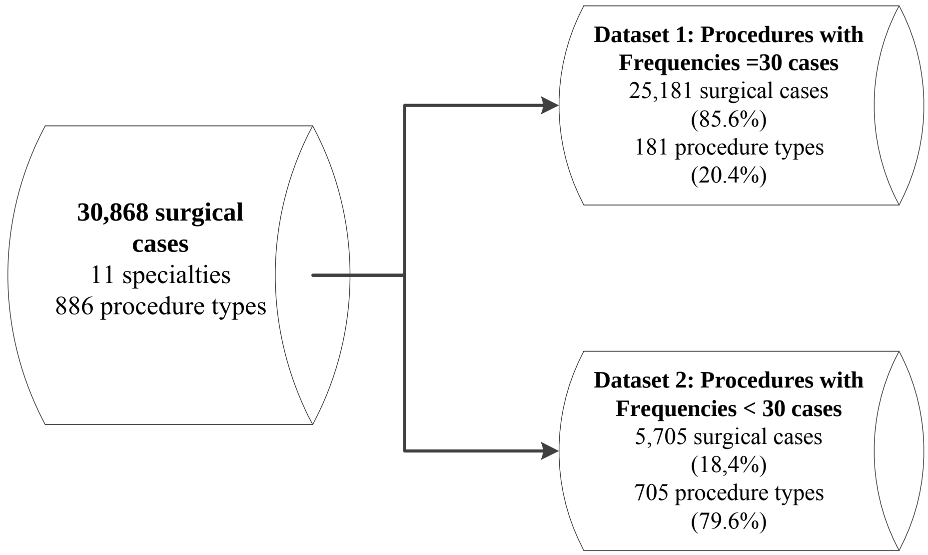


Figure 7 Split procedure to obtaining dataset 1 and dataset 2

Since we are interested in applying parametric techniques on dataset 1, a Univariate analysis on the surgical procedure duration segmented by operation type was initially conducted. After performing this analysis, it was noted that the distribution of surgical procedure duration by operation type was mainly skewed to the right.

Table 8 shows a cross-tabulation of the number of procedures versus the coefficient of variation. As pointed out by [46] the coefficient of variation (CV) appears to decrease as the number of surgical cases per operation increases; the p-value from the chi-square test of independence is 0.794. Table 9 shows a cross tabulation of the skewness values versus the number of surgical cases per procedures. Skewness appears to increase with smaller sample sizes; the p-value from the chi-square test for independence is 0.375.

Table 8 Cross-tabulation of the number of operations with n surgical cases versus the coefficient of variation (CV) of its values (number of procedures, row percentage)

Number of surgical cases n per procedure	$CV \leq 0.25$	$0.25 < CV \leq 0.35$	$0.35 < CV \leq 0.40$	$0.40 < CV \leq 0.50$	$CV > 0.5$	Total
$n \leq 50$	3(5.8)	6(11.6)	5(9.7)	14(26.9)	24(46.15)	52(28.8)
$50 < n \leq 100$	3(5.5)	6(10.9)	6(10.9)	12(21.8)	28(50.9)	55(30.4)
$150 < n \leq 250$	1(2.1)	9(18.4)	7(14.3)	7(14.3)	25(51.0)	49(27.0)
$250 < n$	2(8)	6(24)	2(8)	6(24)	9(36)	25(13.8)
Total	9(4.9)	27(14.9)	20(11.1)	39(21.6)	86(47.5)	181(100)

Table 9 Cross-tabulation of the number of procedures with n surgical cases versus the skewness (sk) of its values (number of procedures, row percentage)

Number of surgical cases n per procedure	$sk \leq 0.30$	$0.30 < sk \leq 0.75$	$0.75 < sk \leq 1.15$	$sk > 1.15$	Total
$n \leq 50$	4(7.6)	20(38.5)	16(30.8)	12(23.1)	52(28.7)
$50 < n \leq 100$	7(12.7)	14(25.5)	14(25.5)	20(36.4)	55(30.3)
$150 < n \leq 250$	4(8.2)	12(24.5)	12(24.5)	21(42.8)	49(27.1)
$250 < n$	0(0)	8(32)	9(36)	8(32)	25(13.9)
Total	15(8.3)	54(29.8)	51(28.2)	61(33.7)	181

Since there is not a clear cut-off to classify procedures based on its volume and measurements of central tendency and dispersion, we also used a theoretical normality test and probability plots to determine if it is possible to assume that procedures are normal or lognormal distributed as is suggested in the literature [24, 39, 40, 42, 47]. Table 10 displays the results of fitting procedure duration to the normal distribution. The procedure duration data fit the normal distribution for almost 36% of total number of operations. As reported in other empirical studies [24] it was noticed a decrease in the proportion of operations that fit the normal distribution as the sample size (number of surgical cases per operation) increased.

Table 10 Cross-tabulation of the number of procedures with (n) surgical cases versus the Kolmogorov-Smirnov statistic (KS) of its values (number of procedures, row percentage)

Number of surgical cases ($\boldsymbol{n}$) per procedure	$\boldsymbol{KS} < 0.5$	$\mathbf{0.05} \leq \boldsymbol{KS}$	**Total**
$n \leq 50$	9(38%)	31(62%)	40
$50 < n \leq 150$	6(66.7)	27(33.3)	33
$150 < n \leq 250$	74(80)	5(20)	79
$250 < n$	28(96)	1(4)	29
Total	117(64.7)	64(35.3)	181

Using the Univariate statistics of each operation, operations in dataset 1 were classified into four categories using the flowchart shown in Figure 8. They were classified according to volume and variability as follows: (a) high frequency, high variability; (b) high frequency, low variability; (c) low frequency, high variability; and (d) low frequency, low variability. For instance, a procedure or operation is classified as having high frequency and high variability if it has more than 250 surgical cases and at least one of the following indicators is present: Kolmogorov-Smirnov statistic less than 0.05, skewness value greater or equal to 0.75 or, coefficient of variation greater than 0.5. Table 11contains the classification of procedures belonging to dataset 1 based on its frequency and variability.

Table 11 Classification of procedures based on volume and variability (n or $\mathbf{n_p}$, row percentage)

Procedure category (or Group)	**Number of surgical cases($\boldsymbol{n}$)**	**Number of procedures ($\boldsymbol{n_p}$)**
High frequency, high variability (F_HV_H)	8,513(33.8)	19(10.5)
High frequency, low variability (F_HV_L)	2,899(11.5)	6(3.3)
Low frequency, high variability (F_LV_H)	11,326(44.9)	124(68.5)
Low frequency, low variability (F_LV_L)	2,443(9.8)	32(17.7)
Total	25,181	181

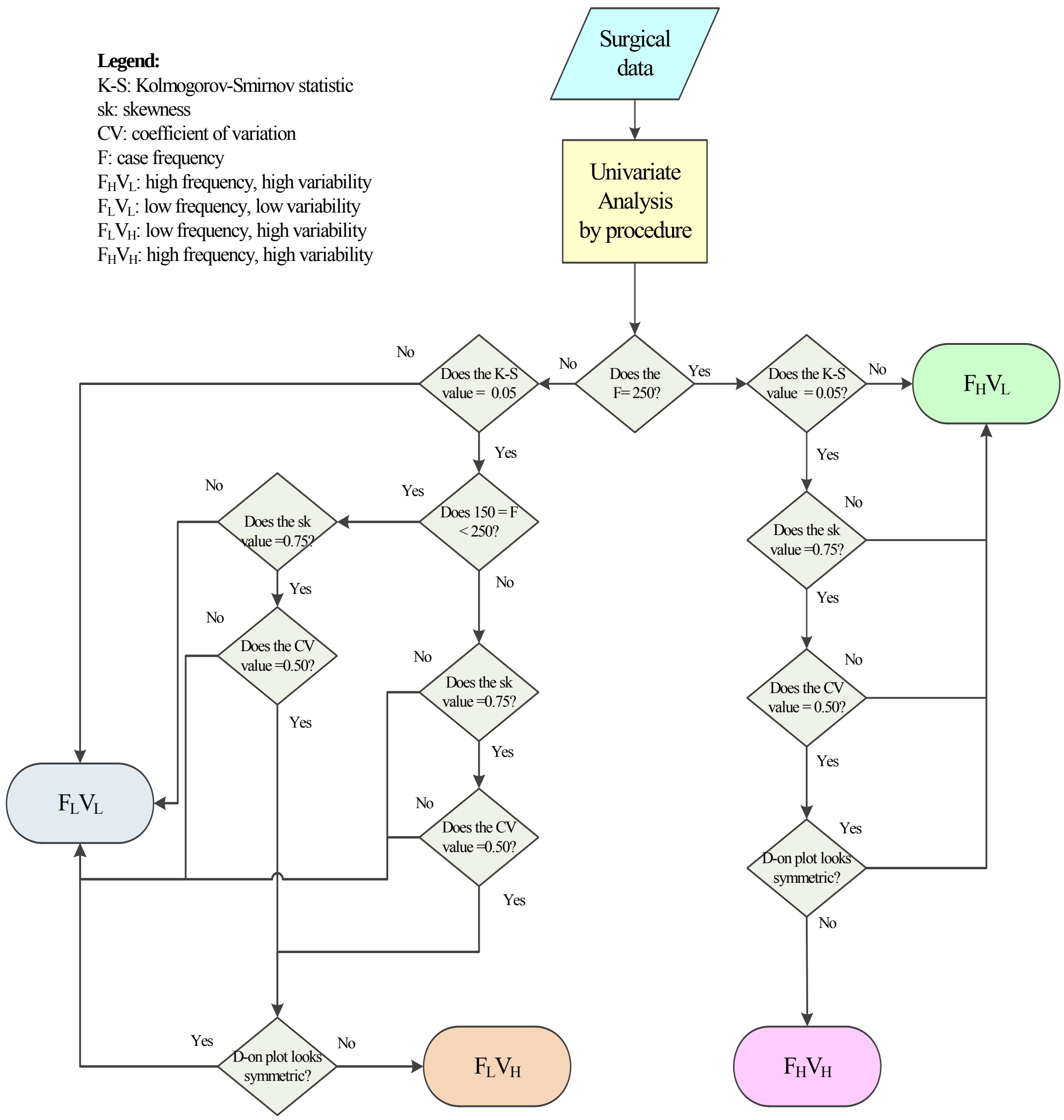


Figure 8 Roadmap to classify procedures based on surgery volume and variability

The information provided by the previous classification is used in the next sections to select the most suitable predictive model to deal with data that contains different levels of variability. For example, the least-squares estimator performs well when the conditions of data normality and homogeneity of variance hold. Nonetheless, in the presence of outliers the performance of least-squares regression can be extremely unsatisfactory [48]. Finally, we established a category of "infrequent cases" for all the operation types that had fewer than 30 instances in a 3-year period.

6.1.2 Splitting the Surgical Data for Purposes of Training and Testing the Models

According to the statistical literature [37] a typical protocol for validating empirical models is to split the original dataset into a training set and a test set. Suppose the n surgical cases in the original dataset are divided into the n_t cases in the training set and the n_s cases in the test set, where$n = n_t + n_s$. It is assumed that both the training data and the test datasets are representative samples of the underlying problem.

Another approach [37] is to split the original dataset into three disjoint sets: (a) a training set, (b) a test set, and (c) a validation set. Thus, the training set is used to build the model, and the test set is used to assess the performance on unseen data. Finally, once a model has been chosen, the validation set is used to assess the error rate of the model using an independent dataset.

Accordingly, we used three independent datasets obtained after splitting the data, using the following proportions: 60% for training, 20% for testing and comparing the model's performances on unseen data, and the remaining 20% for validating the chosen predictive model.

6.1.3 Fitting the Surgical Procedure Duration Data to a Known Distribution

Since surgical times are considered random variables that cannot be predicted with certainty, probability distributions must be used to describe their properties. In general, surgical times are assumed to be normal or lognormally distributed [24, 39, 40, 42, 46] because times are strictly positive. In the next section, the procedures belonging to dataset 1 (with 30 cases at least per procedure type) are analyzed to determine whether normal or lognormal models fit better. The triangular distribution is used when the number of surgical cases per procedure is less than 30, which is the case for procedures belonging to dataset 2.

6.1.3.1 Normal versus Lognormal Fitting

Using an approach similar to [46], we defined a stochastic variable X_i as the procedure duration of operation i in Equation (1).

$$X_i \sim f(\mu_i, \sigma_i) \tag{1}$$

where

X_i=Procedure duration of operation i

μ_i =Expected procedure duration of operation i

σ_i =Standard deviation of procedure duration of operation i

To test whether or not the procedure duration of operation i can be modeled using a normal or lognormal distribution we used the Kolmogorov-Smirnov goodness-of-fittest. Specifically, in Equation (2), we tested the null hypothesis that the model distribution fits the data. The segment durations by operation type were tested before and after applying the log transformation, since it is known that a lognormal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed [49]. Thus, to be considered "non-normal" the p-value for the transformed random variables should be less than a cutoff alpha value of 0.05.

$$H_0 = X_i \sim N \; or \; (log(X_i) \sim N \xrightarrow{yields} X_i \sim logN) \tag{2}$$

where

H_0=null hypothesis

X_i=procedure duration of operation i

For the normal and the lognormal models, dataset 1 was used to test the null hypothesis that these models fit the data. The μ and σ parameters were estimated for each procedure type in dataset 1. Then, Equation (3) was used when estimating the procedure duration for the next surgical case since that is the expectation of a normal or lognormal model with parameters μ and σ [50].

$$\widehat{Y_{Z_i}} = \exp(\hat{\mu}_i + \frac{\hat{\sigma}_i^2}{2}) \tag{3}$$

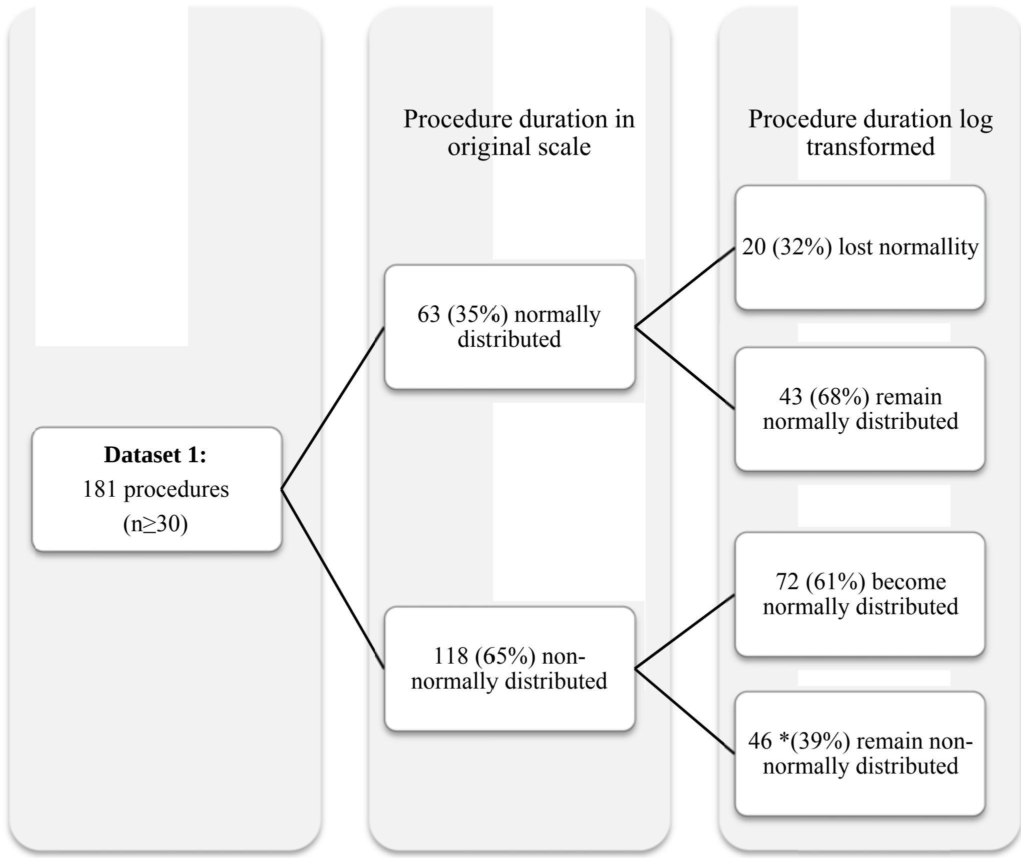
where

μ_i =Expected procedure duration of operation i using the training dataset

σ_i =Standard deviation of procedure duration of operation i using the training dataset

$\widehat{Y_{Z_i}}$ =Expected procedure duration of operation i for the next surgical case

Figure 9 displays the number (and percentage) of procedures contained in dataset 1 that were normally distributed in the original scale and those that became normally distributed after applying the log transformation.



*After residual plot inspection, only 10 of 46 pro cedures do not follow a normal or log normal distribution

Figure 9 Number and percentage of procedures best fit by normal or lognormal distribution

Specifically, Figure 9 shows that 35% of the procedures contained in dataset 1 were originally normally distributed and 65% were non-normally distributed. From those that were originally normally distributed, 32% lost their condition of normality after applying the log transformation.

As pointed out in [24], applying a log transformation to a dataset that follows a normal distribution could invalidate any statistical analysis. Therefore, in this specific case, the log transformation should be applied only when the data is not normally distributed in the original scale.

Furthermore, Figure 9 shows that the null hypothesis was rejected for 46 operation types (25% of the total 181 operations). As a result, we analyzed residual plots to verify the correctness of these rejections since it is known that the Kolmogorov-Smirnov statistic is sensitive to sample size [51]. Only 10 (5.5% of the total 181 operations) did not follow a normal or lognormal distribution.

For simplicity, a single distribution is chosen as the best fit for all the surgical times independently of the operation.

Figure 10 and Figure 11 show the histograms and cumulative density function (CDF) of the procedure duration(X) of 1,580 total-knee-replacement cases, before and after applying the log transformation. In Figure 10, we noted that $\log(X)$ of the transformed data, better resembles the shape of a normal distribution than the procedure duration in its original scale (X). Figure 11 displays the residual plots before and after the log transformation. Specifically, the residual plot of the $\log(X)$, looks fairly straight, when the few extreme outliers are ignored.

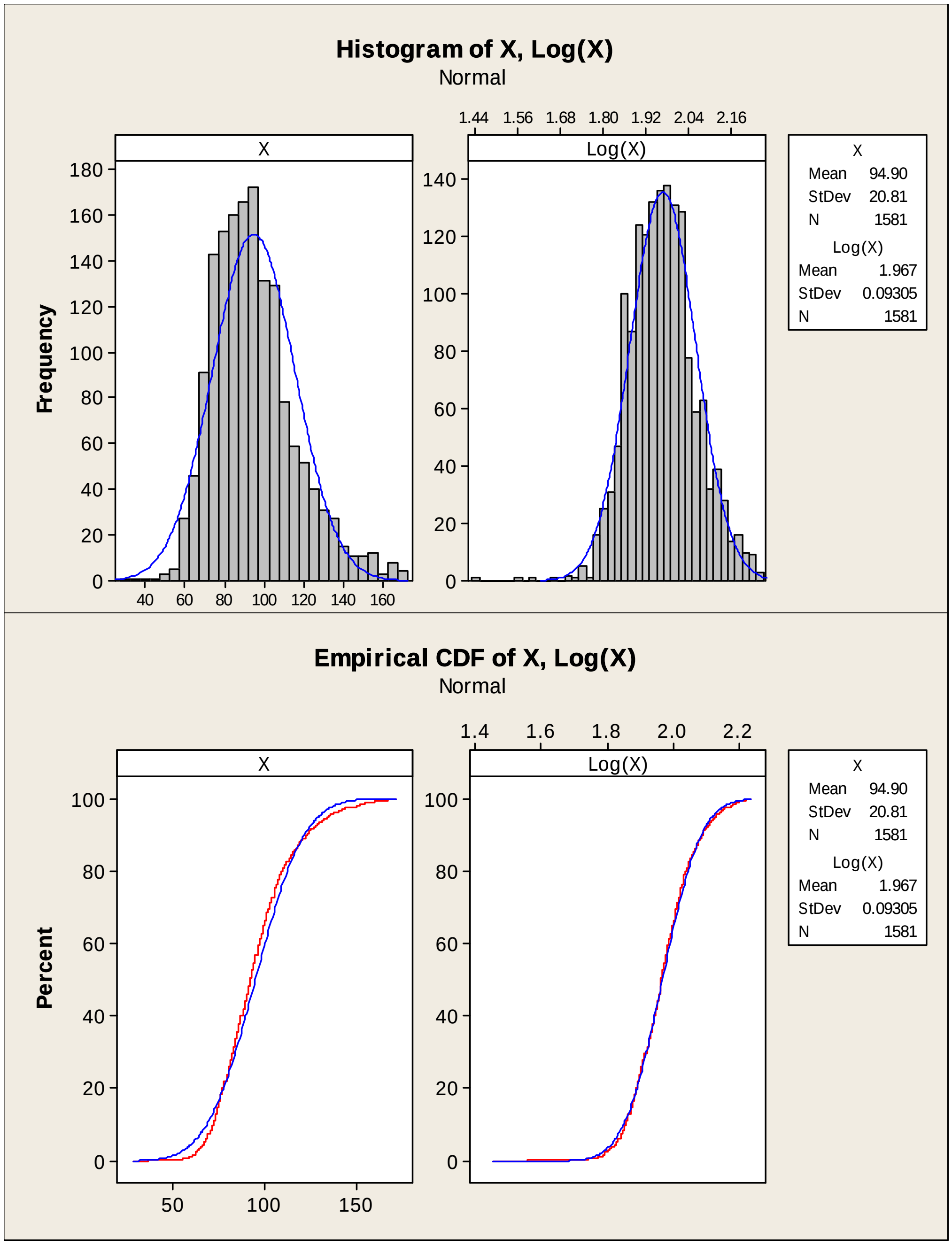


Figure 10 Histograms and cumulative density function for best-fit normal and lognormal of the procedure duration (X) before and after applying the log transformation. The unit for the horizontal axis is in minutes

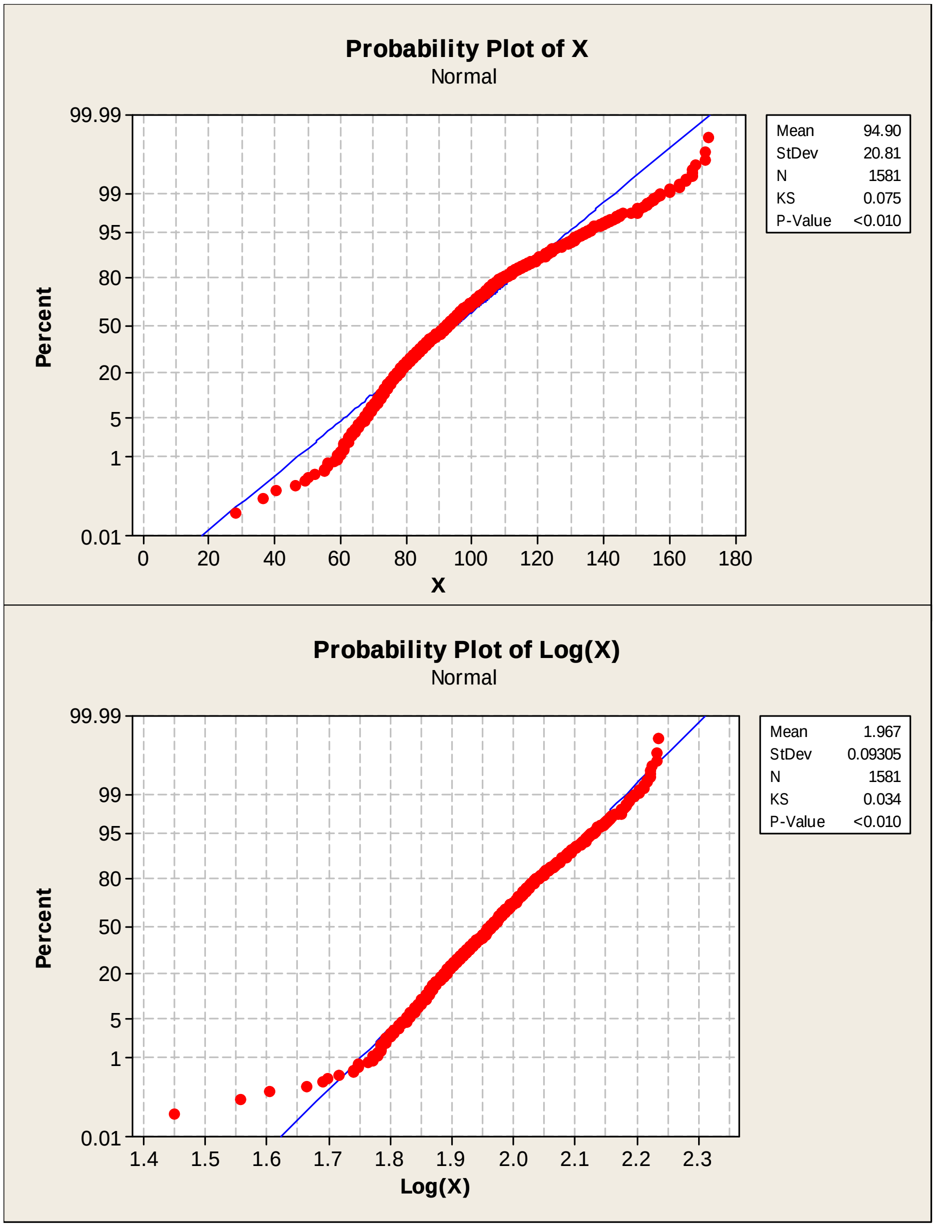


Figure 11 Probability plot of procedure duration (X) and log(X) with their corresponding KS statistic and p-values

6.1.3.2 Triangular Distribution

Typically, the uncertainty associated with non-uniform continuous probability functions, can be modeled by three-point estimates (optimistic, most likely, and pessimistic values) when data is scarce [52]. Also, the three-point estimates are used to modeling expert's opinion in several fields and to build an approximate probability distribution for the outcome of a possible future event when the available information is limited [52, 53].

In this work, this approach is used to model the procedure duration of "infrequent" operations and for modeling random variables based on very limited information, which is the case of operations belonging to dataset 2 (705 operations).

The mean and standard deviation are estimated from its three parameters as shown in Equations (4) and (5) [54].

$$\bar{x} = \frac{a+b+c}{3} \tag{4}$$

$$\sigma = \sqrt{\frac{a^2+b^2+c^2-ab-ac-bc}{18}} \tag{5}$$

where

a=Minimum value (lower limit)

b=Maximum value (upper limit)

c=Most likely value (mode)

6.1.4 Use of Generalized Linear Model to Predict the Surgical Procedure Duration

Generalized linear methods (GLM) are defined in [55]. The basic model uses the least squares method to fit a linear equation between one or more independent variables (predictors) and a dependent variable (response). According to [56], the assumptions of the linear model include: (1) homogeneity of variance; (2) simplicity of structure for the expected value of the response; and (3) at least approximate normality of the additive errors. It is also assumed that the errors are independent. Additionally, if it is not possible to satisfy these requirements in the original scale of measurement of the response (variable of interest) it may be useful to apply a transformation to the response to stabilize the variance and produce at least approximate normality [57].

A traditional linear model is defined in Equation (6).

$$y_i = X_i'\beta + \epsilon_i \tag{6}$$

where

y_i=Response variable of the ith observation

X_i'=Column vector of independent variables (or predictors) for observation ith

β=Vector of unknown regression coefficients

ϵ=Vector of the errors for the ith observation, which is assumed to be an independent normal random variables with zero mean and constant variance

The unknown parameters in the model described by Equation (6) are the regression coefficients β and the error. Thus, for estimating β, the least-squares criterion is used, that is,coefficients are estimated by a least-squares fit to the data y_i. Then, the expected value of y_i, denoted by μ_i, is defined in Equation (7).

$$\mu_i = X_i' \beta \tag{7}$$

where

X_i'=Column vector of independent variables (or predictors) for observation ith

β=Vector of unknown regression coefficients

As highlighted in the literature review chapter, linear methods have been extensively used to model the surgical procedure duration [14, 22, 23, 42, 58, 59]. Therefore, the general process followed in this work identifies a set of independent variables or predictors affecting the procedure duration to fit a general linear model. The most significant factors identified based in the literature are the following:

- Patient characteristics such as age, and gender
- Case information such as operation type, and anesthesia type
- Surgeon factors such as years of experience, and work rate

6.1.4.1 GLM Dataset

For the purpose of fitting a generalized linear model to predict the procedure duration of new surgical cases we selected a subset of surgical cases from dataset 1 (operation types with more than 30 cases). Since the performance of a generalized linear model will be compared with Support Vector Regression, we selected a reduced subset of surgical records to avoid the issue of "course of dimensionality". According to [37], "high dimensionality of the input (that is, the number of variables in a model) increases the size of the search space in an exponential manner", and thus increases the chance that the model being tested could find spurious solutions that are generally invalid. Thus, we have selected operations in dataset 1 that were classified under the groups of high frequency, low variability (F_HV_L) and low frequency, low variability (F_LV_L) for fitting the models to be compared. Specifically, the reduced dataset is composed of 5,342 surgical cases belonging to 38 operations (see Figure 7). Each surgical record in the regression dataset is composed by the following seven variables or predictors

- Patient age (P_age),
- Patient sex (P_sex),
- American Society of Anesthesiologist (ASA) physical status classification system (ASA_code),
- Emergency ASA code (E_code),
- Surgical service (Serv_code),
- Surgeon speed score (SSS), and
- Surgical case duration score (SCS).

Before proceeding with the construction of the model we split the regression dataset into three random subsets. The first subset is used for training the model (60%, 3,206 surgical cases) and the remaining surgical cases were assigned to the test subset and validation subset (20% each, 1068 surgical cases).

6.1.4.2 GLM Attribute Selection

The impact of the seven variables previously described was investigated by using the GLMSELECT procedure available in SAS v.2. The GLMSELECT procedure performs effect selection in the framework of general linear models [60]. Specifically, we used the stepwise option to select the best subset of predictors based on a predetermined significant alpha level of 0.05. The stepwise option implemented in procedure GLMSELECT, is a modification of the forward selection technique. In this technique, effects that initially are in the model, can be removed [61]. The stepwise process ends with the selection stage, when none of the effects (predictors) outside of the model, has an F statistic equal to 0.05 and every effect in the model is significant at the SLSTAY equal to 0.05. After applying the stepwise process, three predictors were removed (E_code, ASA_code, and P_age) from the set of potential informative predictors.

6.1.4.3 General Linear Model (GLM)

For the variable of interest which is the total procedure duration, a generalized linear model using the regression train dataset (3,206 surgical cases) was built. The resulting linear model was tested using the test regression dataset (1,068 surgical cases). The model building process was carried out in SAS v.2 (SAS Institute Inc., Cary, NC, USA). Three significant predictors with p-values less than or equal to 0.05 were considered as independent variables. Table 12shows the relevant information about the full-fitted linear model using the training set.

Table 12 Summary of the full fitted linear model to predict procedure duration using the training subset (n=3,206 surgical records)

The SAS System

Model: MODEL1
Dependent Variable: logY

Number of Observations Read	3206
Number of Observations Used	3206

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	7	996.18858	142.31265	552.35	<.0001
Error	3198	823.96939	0.25765		
Corrected Total	3205	1820.15796			

Root MSE	0.50759	**R-Square**	0.5473
Dependent Mean	4.39789	**Adj R-Sq**	0.5463
Coeff Var	11.54175		

Parameter Estimates							
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	Intercept	**1**	-2.60687	0.14539	-17.93	<.0001	0
SEX_CODED	SEX_CODED	**1**	-0.02639	0.01799	-1.47	0.1425	1.00456
SERV_CODED	SERV_CODED	**1**	0.0345	0.01411	2.45	**0.0145**	1.2289
P_AGE	P_AGE	**1**	-0.0011	0.00056309	-1.95	0.0511	1.42933
SPS	SPS	**1**	0.72097	0.13377	5.39	**<.0001**	1.55961
PDS	PDS	**1**	6.36369	0.13256	48.01	**<.0001**	1.53682
ASA_CODED	ASA_CODED	**1**	0.01273	0.01454	0.88	0.3815	1.36121
EMERG_CODE	EMERG_CODE	**1**	-0.00236	0.05524	-0.04	0.966	1.04731

Table 13 shows the fitted linear model considering only the significant predictors and the training subset. From Table 13, we concluded that the fitted model explains a statistically significant proportion of the variance (F-test: p-value <.0001). The proportion of the total variance explained by the model is 54.6% (RMSE equal to 0.50777). The table also displays the variance inflation factor (VIF) for each fitted coefficient of the regression model. In general, if the VIF values are less than 5 it can be assumed that the model does exhibit high multicollinearity [62].

Once the linear model was fitted, we proceeded to assess its performance on unseen data. For that purpose, we used the test subset (1,068 surgical records).

Table 13 Summary of the fitted linear model to predict procedure duration using only the significant predictors and the training subset (n=3,206 surgical records)

The SAS System

Model: MODEL 1 (significant predictors)
Dependent Variable: logY

Number of Observations Read	3206
Number of Observations Used	3206

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	3	994.59349	331.53116	1285.86	<.0001
Error	3202	825.56448	0.25783		
Corrected Total	3205	1820.15796			

Root MSE	0.50777	**R-Square**	0.5464
Dependent Mean	4.39789	**Adj R-Sq**	0.546
Coeff Var	11.5457		

Parameter Estimates							
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	Intercept	**1**	-2.58754	0.14156	-18.28	<.0001	0
SERV_CODED	SERV_CODED	**1**	0.03432	0.01378	2.49	0.0128	1.1717
SPS	SPS	**1**	0.71851	0.13287	5.41	<.0001	1.53776
PDS	PDS	**1**	6.3067	0.1284	49.12	<.0001	1.44084

Table 14 displays the relevant information regarding the performance of the fitted model on unseen data (test subset). The predictive performance (R^2) for the test subset is 52.51% (RMSE equal to 0.51869). It would appear that the fitted model generalizes well to unseen data and does not exhibit overfitting.

Table 14 Summary of the fitted linear model to predict procedure duration using the test subset (n=1,068 surgical records)

The SAS System

The REG Procedure
Model: MODEL1 (test dataset)
Dependent Variable: logY

Number of Observations Read	1068
Number of Observations Used	1068

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	3	318.16393	106.05464	394.2	<.0001
Error	1064	286.25389	**0.26904**		
Corrected Total	1067	604.41782			

Root MSE	0.51869	**R-Square**	0.5264
Dependent Mean	4.4008	**Adj R-Sq**	**0.5251**
Coeff Var	11.78618		

6.1.4.4 Checking Assumptions of Linearity

While adjusting a general linear model, various graphical and analytical procedures were used to assess the validation of the linear assumptions. Figure 12 and Figure 13 respectively, illustrate an approximate normal distribution of the errors and their corresponding normal probability plot. Figure 14 shows that residuals are randomly dispersed around zero, which indicates that the assumption of constant variance of the residual holds.

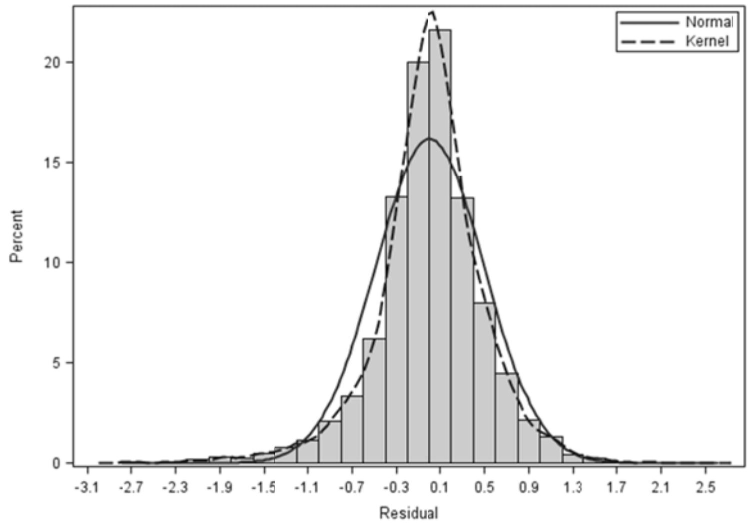


Figure 12 Histogram of the residuals.

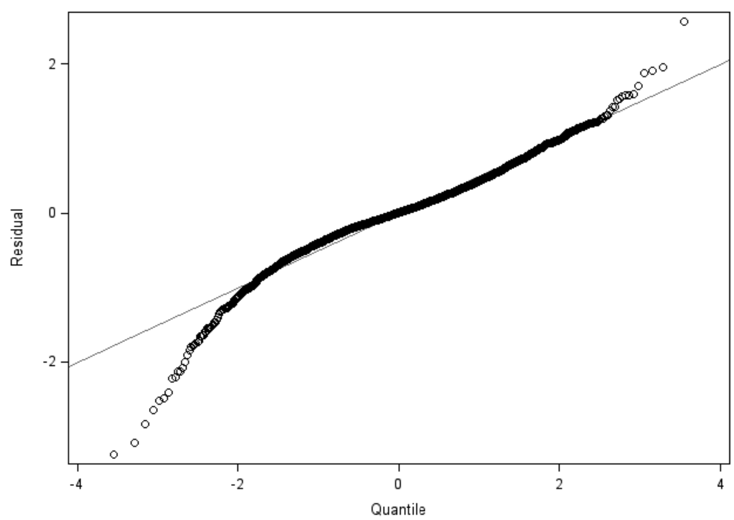


Figure 13 Residual normal quantiles plot

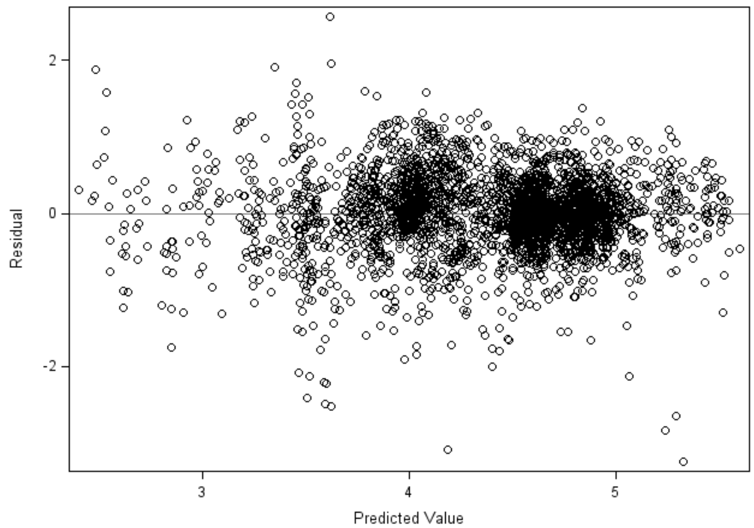


Figure 14 Plot of standardized residuals versus predicted values

6.1.5 Use of Support Vector Regression to Predict the Surgical Procedure Duration

In this section we sketch the Support Vector Regression (SVR) algorithm and summarize its motivation. A more detailed description of SVR can be found in [63].

SVMs were initially developed for classification purposes by Vapnik [64]. Essentially, SVMs are supervised methods that use learning algorithms to analyze data and recognize patterns [65]. The main goal of SVMs is to produce a model (based on training data) which predicts the target values of the test data given only the test data attributes [66]. SVM uses an implicit mapping Φ of the input data to a higher dimensional space and employs a kernel function, which returns the inner product between two points in a suitable feature space. Thus, the learning process takes place in the feature space and the data points only appear inside dot products with other points [67].

This mapping process is also known as the "*kernel trick*" according to [37], because it is a method for computing similarity in the transformed space, using the original attribute set. The similarity function k, which is computed in the original attribute space, is also referred to as the kernel function. Table 15 contains the kernels most commonly used with SVMs.

Table 15 Functions commonly used with SVMs

Type of kernel	**Kernel function**	**Kernel hyperparameter**
Linear	$k(x,x') = \langle x,x' \rangle$	---
Gaussian Radial Basis Function (RBF)	$k(x,x') = \exp(-\gamma\|x - x'\|^2)$	γ
Polynomial	$k(x,x') = (scale \cdot \langle x,x' \rangle + offset)^p$	p
Laplace Radial Basis Function	$k(x,x') = \exp(-\gamma\|x - x'\|)$	γ

The principles of SVM can also be used for non-discrete outcomes, resulting in a regression model instead of a classification model [63] known as Support Vector Regression (SVR). SVR is described in [68] as follows: given a regression training dataset$D$ that satisfies the conditions in Equation (8), the goal of SVR is to find a function f that has minimum ω less than ε deviation from the target y_i, as described in Equation (9). Here, the parameter ε controls the $\varepsilon - insensitive\ zone$, used to fit the training data.

$$D=\{(x_i,y_i),(x_2,y_2),\ldots,(x_N,y_l)\}\subseteq X\ x\ R \tag{8}$$

where

N =Number of examples (data points)

x=Input vector

y =Output vector

X =Input space

$$f(x) = \langle \omega, x \rangle + b \text{ with } \omega \in \text{X}, \text{b} \in \text{R} \tag{9}$$

where

$\langle .,. \rangle$ denotes the dot product in X

In order to compute the optimal SVR model, a convex optimization problem is defined, to find the minimum ω as shown in Equation (10).

$$Min \frac{1}{2}\|\omega\|^2 \tag{10}$$

$$s.t. \begin{cases} y_i - \langle \omega, x_i \rangle - b \leq \varepsilon \\ \langle \omega, x_i \rangle + b - y_i \leq \varepsilon \end{cases}$$

where

y =Output vector

ε =Threshold or radius of the hypertube used to fit the training data

Equation (10) assumes that it is possible to fit all the observations (data points) into a hypertube of width 2ε. Therefore, to be able to consider all the training data points (including those outside the strict hypertube), it is necessary to introduce a correction term or slack variable (ξ) as suggested in [63]. Slack variables account for the correction or deviation of each observation that lies outside the hypertube of width 2ε. Equation (11) is equation (10) modified to include the slack variables.

$$Min\,\frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{11}$$

$$s.t.\begin{cases} y_i - \langle \omega, x_i \rangle - b \le \varepsilon + \xi_i \\ \langle \omega, x_i \rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases}$$

where

C =Penalty constant or regularization parameter

ξ_i =Absolute value necessary to move the observation to the hypertube setting $\xi_i^*$to zero

ξ_i^*=Absolute value necessary to move the observation to the hypertube setting $\xi_i$to zero

ε =Threshold or radius of the hypertube used to fit the training data

By applying the Lagrangian function we can solve the Equation (11) in its dual optimization formulation. Then, via the kernel function, the data is mapped into a higher dimensional space. Thus, the Equation (11) is expanded into Equations (12), (13), (14) and (15) [63, 64, 69, 70].

$$Min\left\{-\frac{1}{2}\sum_{i,j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)k(x_i, x_j) - \varepsilon\sum_{i=1}^{N}(\alpha_i - \alpha_i^*) + \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\right. \tag{12}$$

$$s.t.\begin{cases} \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases}$$

$$\omega = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*)\, \Phi(\mathrm{x_i}) \tag{13}$$

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*)\, k(x_i, x) + b \tag{14}$$

$$k\langle x, x' \rangle \coloneqq \langle \Phi(x), \Phi(x') \rangle \tag{15}$$

where

Φ =Mapping function

k =Kernel function

In summary, SVR ignores errors that are smaller than a certain threshold $\varepsilon > 0$, creating a n-dimensional hypertube around the true output (target) and penalizing points falling outside the hypertube through a regularization parameter C training points that fall outside the hypertube [63].

6.1.5.1 SVR Dataset

For the purpose of modeling the procedure duration of operations using the SVR model, we used the same dataset of section 6.1.4.1, in order to compare the performance of both methods in equal terms and conditions. As described before in section (6.1.4.1), each surgical record is composed by seven predictors.

6.1.5.2 Building the SVR Model

Although a generic algorithm for modeling SVR does not exist, we used a variation of the algorithm for Support Vector Classification problem as proposed in [66]. Table 16 contains a summary of the general steps followed to train and test the SVR model. The model building process was carried out in R version 2.13.0 (Free Software Foundation's GNU project). Specifically, we used the e1071 package which is an implementation of SVM in R [71]. The svm() function in e1071, fit the regression model using the training subset and predicts the test subset values. The details concerning the hyperparameter tuning process are presented in the next section.

Table 16 Algorithm followed to build the SVR model

Step 1: Transform the data to the format of an SVM/SVR package
Step 2: Conduct a simple scaling of the data
Step 3: For $\varepsilon = 10^{x}\ with\ x = -3, -2, \ldots..3$ conduct a kernel selection process as described as follows
3.1 Consider the kernel types: linear and RBF
3.2 Conduct a coarse grid search using 10-fold cross validation to find the values of C (penalty constant) and any parameters the kernel function may depend on
3.3 Repeat 1 and 2 for all the values of ε being considered
Step 4 : Select the SVR model with the minimum MSE
Step 5: For the SVR model selected (ε, C, and kernel type) conduct a refined grid search on the neighborhood of the corresponding hyperparameters
Step 6: Use the selected ε and the refined hyperparameters values to train the final SVR model using the whole training set
Step 7: Test the model using the test dataset

6.1.5.3 SVR Hyperparamenter Tuning Process

As is the case with other predictive methods, the performance of SVR may be very sensitive to the proper choice of hyperparameters [37, 71-73]. Therefore, we used the function tune.svm() included in the R package e1071 [71] for the purpose of tuning the hyperparameters of the model being constructed. Since we could not find prior applications of SVR for predicting surgical times, a kernel selection process using a grid search algorithm was conducted empirically as specified in Table 16. In addition, a 10-fold cross validation was used inside the tune.svm() R function. The hyperparameter settings selected from the grid had the lowest minimum RMSE. Since doing an exhaustive grid-search may be computationally expensive, a coarse grid search was conducted to identify a "better" region on the grid. After that, a search into a reduced grid area was done. The coarse grid search was performed for $\varepsilon = 0.1$ and the RBF kernel with the parameter $\mathbf{C}$ ranging from 10^{-3} up to10^{3} and γ varying from 2^{-15} up to 2^{6}. Then, after a "better" region was identified, a refined grid search for $\varepsilon = 0.1$ and the RBF was conducted in the neighborhood of $C = 1\ and\ \gamma = 0.5$ was conducted. This process was repeated until the RMSE stabilized its value at a minimum for the selected kernel and the ε value.

Lastly, after applying the grid search algorithm presented in Table 16, we selected the Radial Basis Function (RBF) as the kernel function and combined it with the hyperparameter setting presented in Table 17.

Table 17 Grid search hyperparameter setting with the lowest RMSE

C	γ	ε	RMSE 10-fold CV Train subset
0.8	0.07	0.1	0.4829±0.0101

6.1.5.4 SVR Model

Once the kernel function and the SVR hyperparameter settings were selected, we proceeded to train the model using the training subset. As a result, the final SVR model was generated and used to predict the target value over an unseen test subset. Table 18 contains the performance details of the SVR model on training and test subsets. The performance of the SVR model on the test subset decreases by 1.5% considering the mean value of RMSE as the training accuracy.

Table 18 Root-mean-square error (RMSE) results of SVR model on training and test subsets

RMSE 10-fold CV (train subset)	RMSE (test subset)
0.48293±0.01011	0.49012

6.1.6 Comparison between GLM and SVR Models

We had fit two models to predict the surgical procedure duration as shown in sections 6.1.4.3 and 6.1.5.4. For the purpose of comparing the performance of both models on the test subset (1,068 cases), we used the root-mean-square-error (RMSE) as an estimator of the predictive accuracy. Table 19 displays the RMSE values for each predictive model.

Table 19 Root-mean-square error (RMSE) results on test subset according to the predictive models

Predictive Model	RMSE
GLM	0.51869
SVR	0.49012

Comparing the RMSE values shown in Table 19, we noticed that the SVR model performs better on the testing dataset than GLM. Specifically, SVR has improved the prediction by 5.5%.

Based on the RMSE value, it is recommended to use SVR with these types of data. In addition, since SVR does not depend on normality, it is considered as a suitable technique to minimize the generalization error in the presence of outliers which is an issue that affects the performance of GLM [68].

After selecting SVR as the predictive method that provides the higher accuracy on unseen data, we proceeded to evaluate its performance on the validation subset (20%, 1068 surgical cases).

7. SCHEDULING OPERATING ROOMS UNDER UNCERTAINTY

As described in chapter 2, scheduling elective surgical cases is a process that could be initiated as far as six months prior to the planned day of surgery. A common approach used to schedule operating rooms is to estimate the duration of surgical cases using the expected duration of an operation using historical data. However, since case duration is a random variable, the variation around the expected value may cause delays in the start time of subsequent cases or even worse could cause the cancellation of delayed cases [14]. Furthermore, the arrival of unplanned surgical cases (emergency cases) may worsen the aforementioned scenario.

One of the key aspects of planning and scheduling operations in ORs is the coordination of several activities and actors under an uncertain environment [6]. As a result, most of the published work focuses on deterministic approaches to avoid dealing with the computational complexity required to incorporate, in the modeling process, the stochastic nature of healthcare activities [6, 11].

As pointed out in [11] several modeling approaches have been applied to analyze operating room planning and scheduling problems. Most of the published work uses combinatorial optimization techniques, and a lesser proportion uses scenario analysis and benchmark studies to compare multiple settings or options through performance criteria [11]. Thus, simulation is one of the suitable techniques to investigate the stochastic nature of healthcare activities in the process of OR scheduling under uncertainty. As a criterion for application, simulation will be justified by the degree of detail and the amount of stochasticity being considered [6, 11].

In this work we developed a discrete event simulation model (DES) using the simulation software Flexsim (Flexsim Software Inc., Orem, UT, USA), to compare various performance measures of the behavior of an imaginary virtual hospital OR system using four levels of three different control factors: (1) Scheduling policy, (2) Scheduling percentile, and (3) Emergency arrival rate. In the next sections, we describe the simulation model and the virtual hospital OR system in details.

7.1 Case Study: A Virtual Hospital OR System

As we mentioned before, we created an imaginary virtual hospital OR system upon which to design and run a discrete event simulation model. One of our main assumptions is that the ORs operates inside a Level I trauma center. As mentioned in section 2.4, Level I trauma centers offers comprehensive medical services to trauma patients 24 hours a day (365 days per year). Therefore, the virtual hospital OR system needs to have at least one dedicated trauma OR. For that purpose, we are considering that the virtual hospital OR system has nine operative ORs, one of which is a dedicated trauma/emergency OR. Thus, the remaining eight ORs are used to schedule elective operations on a daily basis.

We assumed that the elective ORs are identically equipped, and can be scheduled 8 hrs daily (or 480 minutes) five days a week. These parameters can be modified during the simulation, but expansions would greatly increase the computational complexity.

Therefore, we are focusing on one set of pre-specified, deterministic time constraints, drawing inferences regarding longer hours and more operating rooms.

We further assumed that the virtual hospital OR system only performs elective surgical cases that belong to three different surgical disciplines: Orthopedics, General surgery and Ear-Nose & Throat. These surgical disciplines represent the core business of a medium size U.S. hospital and also characterize the three most typical prototypes of surgical procedures. Orthopedics (ORTHO) represents high volume and high income, General Surgery (GS) is a "must have" surgical discipline within a hospital and, Ear-Nose & Throat (ENT) is a narrow but important surgical specialty. These three surgical specialties capture 20,414 of all surgical cases listed on Table 2. These cases include 390 different types of elective operations.

After analyzing the available dataset, it was found that 112 operations behave according to a lognormal probability distribution. The remaining 278 operations were modeled using the triangular distribution, mainly because there was not enough available data to model otherwise ($n < 30$ surgical cases). Detailed information about the modeling aspects of emergency operations is provided in the next section.

7.2 Experimental Design and Solution Approach

Initially, we constructed preliminary OR schedules considering a roster of 390 elective operations. Each preliminary schedule corresponds to a "one day" OR schedule, although operations can be "prescheduled" many months in advance. For the purpose of simulating the behavior of the virtual hospital OR system under uncertain conditions, and to assess the impact of these conditions in a predefined set of outcomes (performance indicators), we only tested the various levels within Spo and Spe control factors shown in Table 20. Afterwards, during the simulation process, the third control factor was incorporated to account for the randomness of the arrival of emergency cases.

Table 20 Control factors and levels, virtual hospital OR system experiment

	Level			
Control Factor	**1**	**2**	**3**	**4**
Spo: scheduling policy	SP1	SP2	SP3	SP4
Spe: scheduling percentile	60^{th}	70^{th}	80^{th}	90^{th}
L_e: emergency arrival rate (emergencies/hr)	1/2	1/4	1/6	1/8

Simulation were ran for one thousand independent days. A total of 16 "base scenarios" (or base treatments) resulting from the interaction of the levels for each control factor were used to generate a set of feasible preliminary OR schedules.

Each "base scenario" was obtained randomly, selecting cases from the set of operations available in the dataset. Each level of the control factor scheduling policy considers the frequency of each operation by comparing a random number with a cumulative distribution function. Details, regarding the levels of the control factor scheduling policy, are presented in Table 21.

The four levels from the scheduling policy control factor were tested against the four levels of the scheduling percentile factor (percentile of the expected duration of each operation from patient-in to patient-out). Consequently, 16 sets (4x4) of OR schedules that correspond to each of those combinations were generated. Numbers and shown in Table 22.

Table 21 Description of the levels of the control factor scheduling policy (Spo)

Spo Control Factor Level	Scheduling Approach	Description
Sp1	Random	Schedules a random operation until filling each OR each day. After not finding a feasible operation to be scheduled in five consecutive attempts, it declares the OR full and continues scheduling the next OR.
Sp2	Best Fit	Uses the Best Fit[74] by allocating operations to the OR with the least amount of additional time available. It stops when no operation can be scheduled in any of the ORs after five consecutive attempts.
Sp3	Best Fit and Shortest Processing Time (SPT)	Applies the Shortest Processing Time (SPT) rule to schedules, generated under SP2, meaning that operations are rescheduled from the shortest one to the longest one within the same OR.
Sp4	Modified-block OR time	Uses a similar strategy to the one utilized in SP1 but considers that some ORs are pre-assigned to the surgical specialties. Thus, OR1 and OR2 are exclusively used to allocate ORTHO operations, OR3 and OR4 are used for GS operations, OR 5 is used for ENT operations, and OR6, OR7, and OR8, are scheduled following a first-come first-served policy.

As expected, since Sp3 is a variation of Sp2, the number of operations scheduled using these factor levels, is the same. Table 22 shows that using a higher scheduling percentile (Spe), results in a lower number of operations scheduled.

Table 22 Number of operations scheduled for each "base scenario"

Scheduling policy (Spo)	Scheduling percentile (Spe)			
	60th	70th	80th	90th
Sp1	20,665	18,707	16,777	12,786
Sp2	21,106	18,906	16,915	14,524
Sp3	21,106	18,906	16,915	14,524
Sp4	20,818	18,896	16,962	14,267

Once the 16 "base scenarios" and their respective ORs schedules were created, we expanded the quantity of scenarios to be explored during the simulation process, by including the levels of the third control factor, which is the emergency arrival rate (L_e). Particularly, we modeled the arrival of surgical emergencies using a Poisson process as described in [75], which means that emergencies arrive at the virtual hospital OR system at a L_e rate (see Table 20). Inter-arrival times are independent and exponentially distributed. The duration of emergency operations was modeled using a procedure-specific triangular distribution with 76 different types of emergency operations and the operational times of 371 emergency cases.

Finally, the simulation model for the 64 scenarios generated was ran and compared through the performance metrics below as suggested in [76]:

- Average OR system utilization (ORsyst_ut): average of the ratio between the total used OR system time for elective cases and the available regular OR system time.
- Average OR system Overtime (ORsyst_ov): average of the times, used to perform operations after the regular available OR system time has ended
- Elective case average waiting times (Elec_wt): the difference between the planned and actual start time of an elective surgical case.
- Emergency case average waiting time (Emer_wt): the total delay that an emergency case undergoes before entering an OR.

7.3 Simulation Model

Based on the experimental design presented in the previous section, we proceeded to evaluate the robustness of each "base scenario" running a simulation model for the virtual hospital OR system.

As previously mentioned, the virtual hospital OR system has nine ORs. Eight of them are called "elective ORs" and are being scheduled using four scheduling policies and four scheduling percentiles during a rush hour period from 8am to 4pm. The ninth OR is a trauma/emergency dedicated OR, and does not participate in the elective schedule. Therefore, if an emergency case arrives it is assigned to the dedicated OR. In the situation where multiple emergencies arrive together (or close to each other) and the dedicated OR is occupied, the pending emergency case(s) will take precedence over regular elective scheduled operations and are assigned to the next available OR(s). That is, surgical emergency have priority over the preliminary surgical OR schedule. An assumption made for the simulation is that emergency cases are treated in a first-come-first-served basis. In addition, no delays due to a shortage of surgeons or OR staff were considered.

During the surgical case loading process, elective patients are assigned to ORs each day. In addition, operations are expected to start at a tentative pre-assigned hour. As pointed in [77], during the surgical case loading process the main objectives are to maximize OR utilization and to minimize overtime considering capacity constraints.

The OR schedules generated in section 7.2 were used as the baseline schedule. Each operation performed at the virtual OR system was simulated considering the stochastic duration of three stages: pre-surgery, surgery, and post-surgery. Details about the probability distribution of the duration of an elective operation were provided in section 7.2. Note that the turnover time was added after the post-surgery stage.

Initially, during the construction of the "base scenarios", turnover times were estimated to last 30 minutes, mimicking the considerations followed by many hospitals. However, during the simulation runs, we considered an empirical distribution of these times following the turnover times values reported along with the variables included in the virtual hospital dataset. The OR is considered busy until the turnover time is over, and the next operation starts during the pre-surgery state.

8. EXPERIMENTAL RESULTS

The simulation model was implemented in Flexsim (Flexsim Software Inc., Orem, UT, USA) using the reference documentation as described in [78]. After model verification, several validation test runs were performed. A total of 1,000 replicates were executed for each of the 64 scenarios under consideration (4 scheduling policies, 4 scheduling percentiles, and 4 emergency arrival rates), to comply with the half-width of the 97.5% confidence intervals. In addition, as mentioned in section 7.3, we use common random numbers by assigning a separate stream of random numbers (CRN), to each source of randomness in the model. The use of this randomness (duration of operations and the arrival of emergency cases), will increase the precision of our comparisons.

For each scenario, the respective performance measures (Emerg_wt, Elec_wt, ORsyst_ut, and ORsyst_ov) that allow comparisons across scenarios were collected.

8.1 Simulation Output Measures

For the finite-horizon simulation, we computed the confidence intervals using the approach suggested in [79]. Table 23 to Table 26 display the values of the outcome measures (performance indicators) described in section 7.3 and their respective 95% half width broken down by the control factors, scheduling policy, and scheduling percentile. It is seen that 95% half widths are small compared to the values at the center, which means that 1,000 replicates were enough to obtain a precise estimation of the mean values.

In Table 23 and Table 24 we noted that the aggregate average of elective waiting time (95 min), almost doubled the time for the emergency cases (56 min). With regard to the virtual hospital OR system utilization and OR system overtime, we noticed in Table 25 and Table 26, that it values fluctuate between 67% and 88%, and 97 and 111 minutes, respectively.

Table 23 Final output for average elective waiting times from 1,000 replications for the virtual hospital OR system experiment (elec_wt in minutes, 95% half width)

Spo	**Spe**			
	60th	**70th**	**80th**	**90th**
Sp1	89.84(2.02)	94.42(2.28)	103.35(2.62)	117.26(3.60)
Sp2	83.98(1.85)	83.08(1.98)	87.80(2.27)	103.50(2.89)
Sp3	79.87(1.83)	82.99(2.04)	87.43(2.32)	103.86(2.96)
Sp4	88.5(1.02)	93.16(1.25)	102.12(2.61)	120.18(3.27)

Table 24 Final output for average emergency waiting times from 1,000 replications for the virtual hospital OR system experiment (emer_wt in minutes, 95% half width)

Spo	**Spe**			
	60th	**70th**	**80th**	**90th**
Sp1	56.49(2.67)	50.39(2.64)	51.79(3.03)	47.88(2.96)
Sp2	83.08(3.18)	52.40(2.72)	66.64(2.90)	50.17(3.13)
Sp3	82.91(3.74)	63.12(3.13)	51.20(2.79)	45.69(3.78)
Sp4	60.18(2.72)	50.15(2.52)	45.07(2.61)	46.88(2.88)

Table 25 Final output for average utilization from 1,000 replications for the virtual hospital OR system experiment (ORsyst_ut in minutes, 95% half width)

Spo	Spe			
	60th	**70th**	**80th**	**90th**
Sp1	85.78(1.92)	83.67(2.04)	77.24(2.18)	66.78(2.90)
Sp2	87.98(1.78)	85.16(2.20)	80.49(2.34)	72.36(2.78)
Sp3	85.67(1.65)	83.98(1.99)	79.45(1.87)	71.86(2.03)
Sp4	85.36(1.98)	82.28(1.56)	77.56(2.32)	67.56(2.12)

Table 26 Final output for average overtime from 1,000 replications for the virtual hospital OR system experiment (ORsyst_ov in minutes, 95% half width)

Spo	Spe			
	60th	**70th**	**80th**	**90th**
Sp1	99.77(2.49)	111.88(3.49)	99.18(3.33)	96.01(4.44)
Sp2	99.65(2.29)	97.36(2.66)	97.93(2.19)	99.97(3.89)
Sp3	98.99(2.35)	98.99(2.75)	100.40(2.27)	97.58(3.87)
Sp4	98.47(2.87)	99.41(3.29)	99.42(3.29)	97.25(3.12)

On average, waiting times for elective cases are almost twice as long as the waiting times for emergency cases. Figure 15 shows the values for elective waiting times (Elec_wt), and emergency waiting times (Elec_wt), broken down by the control factors. In the upper panel of Figure 15, we observed that there is not a clear effect of how a more conservative metric in scheduling within each scheduling policy impacts the elective waiting times.

In contrast, in the lower panel, it is shown that there is a clear effect of how the control factors impacts the emergency waiting times. Thus, as the percentile increase so does the waiting times for emergency cases under scheduling policies Sp2 and Sp3, being less of an impact for scheduling policies Sp1 and Sp4.

With regard to the performance measure OR system utilization, and OR system overtime, Figure 16 and Figure 17 show that they both depend on the levels of the control factors scheduling policy and scheduling percentile. For instance, using higher scheduling percentiles, strongly decreases the OR utilization and OR overtime. Additionally, it was noticed that as the emergency arrival rate is increased, both performance measures are strongly affected.

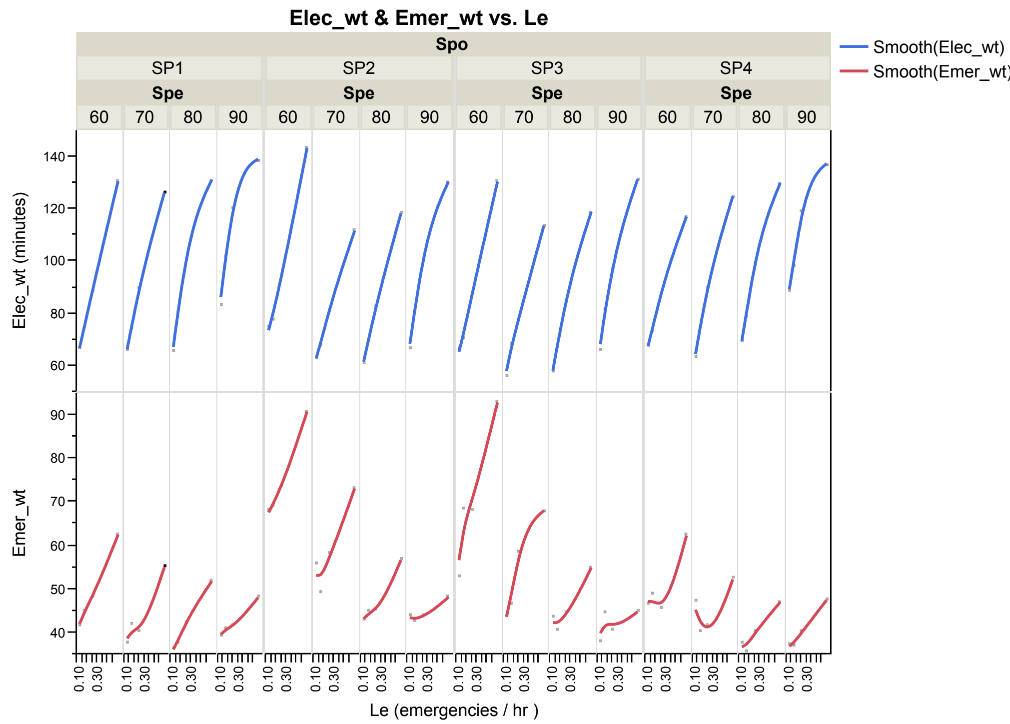


Figure 15 Elective and emergency wait times by control factors

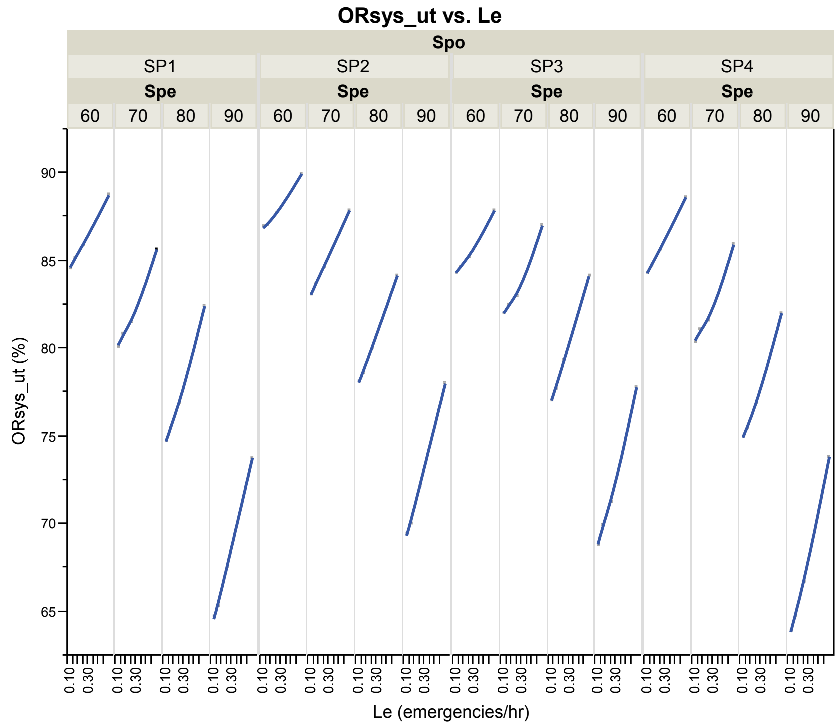


Figure 16 OR system utilization by control factors

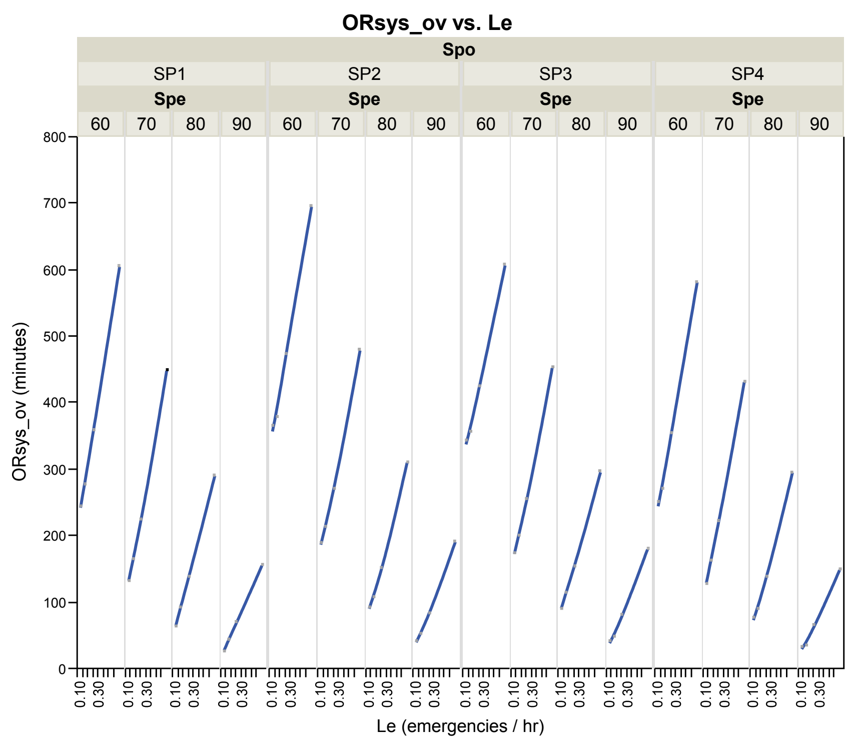


Figure 17 OR system overtime by control factors

8.2 Simulation Output Statistical Analysis

As pointed out in [79] “A primary goal of most simulations studies is the approximation of prescribed system parameters with the objective of identifying parameter values that optimize some system performance measures”.

Since each replication of a simulation with random input processes produces a random output, the performance measures (outcomes) obtained are estimates of the parameters of interest [80]. Generally, these estimates are used to understand the behavior of the system and to predict the outcomes if the input parameter changes.

In the next sections we use multivariate statistical techniques to analyze the set of outcomes obtained from the virtual hospital simulation experiment. Initially. we conducted a multivariate analysis of variance (MANOVA) to explore to what extent the set of outcomes can be predicted or “explained” by the set of control factors. It was also used to identify which control factors were statistically significant. Following the multivariate analysis, we performed a multivariate regression analysis to fit a multivariate-multiple-regression model using the simulation outcomes as the dependent variables.

8.2.1 Multivariate Analysis of Variance (MANOVA)

Multivariate analysis of variance (MANOVA) is used to examine group differences (control factors) on linear combinations of several dependent variables [81]. MANOVA allows the identification of significant difference among groups, and aids in the process of determining which factors are statistically significant.

The MANOVA was carried out in SAS v.2 (SAS Institute Inc., Cary, NC, USA). The macro-level results are shown in Table 27. The F-test indicates an evident group difference among the linear combination of the dependent variables. Then, follow-up analyses were conducted at the micro level. Four ANOVAs were reported by SAS, one for each dependent variable. These results are presented in Table 28. The F-test for all the dependent variables, is significant with R^2 values above 90%. Thus, it appears that the outcomes are significantly different across the control factor levels (main and first order interactions).

The Tukey test (HSD) was used to find means that were significantly different from each other. Table 29 provides the micro-level Tukey test for each of the dependent variables. For instance, we observe that the Tukey test between pair of groups indicates two significant differences for the outcome emergency waiting time (Log_emer_wt). Thus, the mean of level factors Sp1 and Sp2 are statistically different from the mean of the level factor Sp3 and Sp4 (SAS use letters to indicate significant differences). On a similar fashion, the remaining outcomes were tested for difference among pairs.

Table 27 Macro-levels results for MANOVA, virtual hospital OR system experiment

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Spo_c Effect
H = Type III SSCP Matrix for Spo
E = Error SSCP Matrix

S=3 M=0 N=21.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.48301799	3.15	12	119.35	0.0006
Pillai's Trace	0.60834983	2.99	12	141	0.0009
Hotelling-Lawley Trace	0.88322198	3.25	12	74.547	0.0008
Roy's Greatest Root	0.55861724	6.56	4	47	0.0003

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Spe Effect
H = Type III SSCP Matrix for Spe
E = Error SSCP Matrix

S=1 M=1 N=21.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.023767	462.1	4	45	<.0001
Pillai's Trace	0.976233	462.1	4	45	<.0001
Hotelling-Lawley Trace	41.07514205	462.1	4	45	<.0001
Roy's Greatest Root	41.07514205	462.1	4	45	<.0001

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Le Effect
H = Type III SSCP Matrix for Le
E = Error SSCP Matrix

S=1 M=1 N=21.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.3297257	22.87	4	45	<.0001
Pillai's Trace	0.6702743	22.87	4	45	<.0001
Hotelling-Lawley Trace	2.03282397	22.87	4	45	<.0001
Roy's Greatest Root	2.03282397	22.87	4	45	<.0001

Table 28 Micro-level ANOVA results for each outcome variable, virtual hospital OR system experiment

The GLM Procedure

Dependent Variable: ***Log_Emer_wt***

Source	DF	Sum of	Mean Square	F Value	Pr > F
Model	15	0.54189636	0.03612642	37.5	<.0001
Error	48	0.04623616	0.00096325		
Corrected	63	0.58813253			

R-Square	Coeff Var	Root MSE	Log_Emer_wt
0.921385	1.842583	0.031036	1.684392

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Spo_c	3	0.01238416	0.00412805	4.29	0.0093
Spe	1	0.01935468	0.01935468	20.09	<.0001
Spe*Spo_c	3	0.00765755	0.00255252	2.65	0.0594
Le	1	0.02286651	0.02286651	23.74	<.0001
Le*Spo_c	3	0.00484611	0.00161537	1.68	0.1844
Spe*Le	1	0.0090778	0.0090778	9.42	0.0035
Spe*Le*Spo_c	3	0.00422007	0.00140669	1.46	0.2371

Dependent Variable: ***Log_Elec_wt***

Source	DF	Sum of	Mean Square	F Value	Pr > F
Model	15	0.74464718	0.04964315	29.16	<.0001
Error	48	0.0817181	0.00170246		
Corrected	63	0.82636528			

R-Square	Coeff Var	Root MSE	Log_Elec_wt
0.901111	2.120249	0.041261	1.94604

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Spo_c	3	0.00675817	0.00225272	1.32	0.2778
Spe	1	0.01543211	0.01543211	9.06	0.0041
Spe*Spo_c	3	0.00976106	0.00325369	1.91	0.1404
Le	1	0.02830011	0.02830011	16.62	0.0002
Le*Spo_c	3	0.00084963	0.00028321	0.17	0.9185
Spe*Le	1	0.00239899	0.00239899	1.41	0.241
Spe*Le*Spo_c	3	0.00118983	0.00039661	0.23	0.873

Dependent Variable: ***Log_ORsyst_ut***

Source	DF	Sum of	Mean Square	F Value	Pr > F
Model	15	0.09209968	0.00613998	52.96	<.0001
Error	48	0.00556471	0.00011593		
Corrected	63	0.0976644			

R-Square	Coeff Var	Root MSE	Log_ORsyst_u
0.943022	0.567083	0.010767	1.89869

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Spo_c	3	0.00022953	0.00007651	0.66	0.5807
Spe	1	0.0287785	0.0287785	248.24	<.0001
Spe*Spo_c	3	0.00039526	0.00013175	1.14	0.3438
Le	1	0.00069796	0.00069796	6.02	0.0178
Le*Spo_c	3	0.00000245	0.00000082	0.01	0.9992
Spe*Le	1	0.00185215	0.00185215	15.98	0.0002
Spe*Le*Spo_c	3	0.00000752	0.00000251	0.02	0.9956

Dependent Variable: ***Log_ORsyst_ov***

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	7.85036922	0.52335795	244.79	<.0001
Error	48	0.10262408	0.002138		
Corrected Total	63	7.9529933			

R-Square	Coeff Var	Root MSE	Log_ORsyst_ov Mean
0.987096	2.071713	0.046239	2.231899

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Spo_c	3	0.00661333	0.00220444	1.03	0.3873
Spe	1	2.08866448	2.08866448	976.92	<.0001
Spe*Spo_c	3	0.00162136	0.00054045	0.25	0.859
Le	1	0.02361945	0.02361945	11.05	0.0017
Le*Spo_c	3	0.00330499	0.00110166	0.52	0.6737
Spe*Le	1	0.14894641	0.14894641	69.67	<.0001
Spe*Le*Spo_c	3	0.00151563	0.00050521	0.24	0.8706

Table 29 Micro-level Tukey test for ANOVA on the outcomes, virtual hospital OR system experiment

The SAS

The GLM Procedure

I. Tukey's Studentized Range (HSD) Test for Log_Emer_wt

Alpha	0.05
Error	48
Error Mean	0.000963
Critical	3.76375
Minimum	0.0292

Means with the same letter are not significantly different.			
Tukey Group	**Mean**	**N**	**Spo_c**
A	1.74137	16	SP2
A			
A	1.71234	16	SP3
B	1.64254	16	SP1
B			
B	1.64131	16	SP4

II. Tukey's Studentized Range (HSD) Test for Log_Elec_wt

Alpha	0.05
Error	48
Error Mean	0.001702
Critical	3.76375
Minimum	0.0388

Means with the same letter are not significantly different.				
Tukey Grouping		**Mean**	**N**	**Spo_c**
	A	1.96836	16	SP1
	A			
	A	1.96416	16	SP4
	A			
B	A	1.93441	16	SP2
B				
B		1.91722	16	SP3

III. Tukey's Studentized Range (HSD) Test for Log_ORsyst_ut

Alpha	0.05
Error	48
Error Mean	0.000116
Critical	3.76375
Minimum	0.0101

Means with the same letter are not significantly different.			
Tukey Group	**Mean**	**N**	**Spo_c**
A	1.908972	16	SP2
A			
A	1.902801	16	SP3
B	1.891974	16	SP1
B			
B	1.891014	16	SP4

IV. Tukey's Studentized Range (HSD) Test for Log_ORsyst_ov

Alpha	0.05
Error	48
Error Mean	0.002138
Critical	3.76375
Minimum	0.0435

Means with the same letter are not significantly different.			
Tukey Group	**Mean**	**N**	**Spo_c**
A	2.28774	16	SP2
A			
A	2.26581	16	SP3
B	2.18779	16	SP1
B			
B	2.18626	16	SP4

8.2.2 Multivariate Regression Model

Commonly, a simulation model produces more than one outcome. As a result, trying to predict the behavior of the real system under "slightly" different experimental settings using the simulation outcomes requires the use of multivariate techniques [82-84]. Since simulation models are simplified representations of real systems, their outputs are usually correlated. In the literature, this problem has been referred as the "multiple response problem" [83]. Some authors consider that to accurately deal with multiple outcomes, "metamodels" should be built and validated [79, 83].

Multivariate regression is a technique that estimates a single regression model when considering more than one outcome variable. Thus, a multivariate-multiple-regression model involves several predictors and multiple outcomes.

The basic assumptions of a multivariate-regression-model are very similar to the Univariate Regression Model [81]: (1) multivariate normality of the residuals, (2) conditional homoscedasticy on predictors, (3) common variance structure among observations, and (4) independent observations.

The structure of a multivariate-multiple-regression model is defined in Equation (16).

$$Y_1 = \beta_{01} + \beta_{11} z_1 + \cdots + \beta_{r1} z_r + \epsilon_1 \tag{16}$$

$$Y_2 = \beta_{02} + \beta_{12} z_1 + \cdots + \beta_{r2} z_r + \epsilon_2$$

$$\vdots$$

$$Y_p = \beta_{0p} + \beta_{1p} z_1 + \cdots + \beta_{rp} z_r + \epsilon_p$$

$$\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_p)'$$

where

Y_p=Response variables ($with\ p > 1$)

z_r=Matrix of coefficients of predictors (control factors)

β=Matrix of unknown regression coefficients

ϵ=Matrix of the errors, which is assumed to have zero mean and variance matrix $\sum_{pxp}$.

The multivariate-multiple regression model for this research, was carried out in SAS v.2 (SAS Institute Inc., Cary, NC, USA) using the procedure PROC REG. The PROC REG statement not only produces four univariate models, but also allows us to test the hypothesis in a multivariate regression. Table 30 shows the SAS output for testing the null hypothesis, which states that all estimated parameters, except the intercept, are zero. The F-test is significant for all four statistics, indicating that the overall model is statistically significant ($p < .0001$). That is, the multivariate tests indicate that for the dependent variables the set of predictors accounts for a statistically significant portion of the variance.

PROC REG and PROC GLM can be used to fit the Univariate models (one for each outcome). Initially, we fit the four models in their original scale but after conducting a residual plot inspection important departures from normality were detected. Therefore, a log-transformation of the response was conducted to stabilize the variance and approximate it to normality.

Table 30 Multivariate test, virtual hospital OR system experiment

The REG Procedure
Model: MODEL1
Multivariate Test: parameters except the intercept
are the same for all the outcomes

Multivariate Statistics and Exact F Statistics					
S=1 M=0.5 N=29					
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
Wilks'	0.056367	334.81	3	60	<.0001
Pillai's	0.943633	334.81	3	60	<.0001
Hotelling-	16.74074	334.81	3	60	<.0001
Roy's	16.74074	334.81	3	60	<.0001

Since our research deals with the outcome variables simultaneously, we need to account for interrelationships. Therefore, the SAS statement MTEST in PROC REG was used to simultaneously fit multivariate regression models and statistically test the significance of various terms corresponding to control factors (independent variables) using multivariate methods as suggested in [85].

The complete second order multivariate model would involve a total of ten terms: three main control factors, three first-order interactions, three quadratic terms, and the intercept. The multivariate hypotheses to test with MTEST in PROC REG are defined below.

- $H_o^{(1)}$= The multivariate model contains only linear terms plus an intercept
- $H_o^{(2)}$= The multivariate model is quadratic without interaction terms
- $H_o^{(3)}$= The multivariate has only linear, first order terms, and an intercept but not quadratic terms

The SAS output from the MTEST, reported significant p-values (<0.0001) for the three hypotheses tested. Therefore, we proceeded to fit a multivariate-multiple-regression model with main effect, first order interactions, and quadratic terms. Table 31 displays the equations of the four estimated response surfaces and their respective R^2, obtained from the output corresponding to the Univariate analysis.

Table 31 Regression model for each outcome variable, virtual hospital OR system

Outcome (Response)	Univariate Model Fit	R^2_{adj} (%)
Emergency waiting times	$Y_{emer} = 2.49129 - 0.02743 * Spe + 0.76718 * Le - 0.04247 * Sposq + 0.0001682 * Spesq - 0.00732 * Spe\_Le$	83.11
Elective waiting times	$Y_{elec} = 2.95291 - 0.12127 * Spo - 0.03535 * Spe + 2.03579 * Le + 0.02022 * Sposq + 0.0002502 * Spesq - 1.59884 * Lesq$	93.99
OR system utilization	$Y_{ut} = 1.65011 + 0.03424 * Spo + 0.00939 * Spe - 0.18541 * Le - 0.00720 * Sposq - 0.00008859 * Spesq + 0.00331 * Spe\_Le$	97.39
OR system overtime	$Y_{ov} = 2.74128 + 0.21742 * Spo - 0.04488 * Sposq - 0.00023176 * Spesq - 1.85288 * Lesq + 0.02964 * Spe\_Le$	99.16

$Sposq = Spo \times Spo, Spesq = Spe \times Spe, Spe\_Le = Spe \times Le, Lesq = Le \times Le$

Lastly, once the “metamodel” had been fitted, the approach pointed out in [85], which suggest to produce plots to check if the data are close to being multivariate normal was followed. Figures 18 to 21 show the diagnosis for the four outcome models conducted in SAS v9.2. The plot of the RSTUDENT residuals shows externally studentized values that take into account heterogeneity in the variability of the residuals. RSTUDENT residuals, that exceed the threshold values of ±2, often indicate outlying observations. Only a small amount of this cases is noted. The residual-by-leverage plots, shows the observations that have high leverage. For instance, the residual-by-leverage plot show in Figure 18, allows us to identify four observations with high leverage values. The normal probability Q-Q plot in the second panel of Figure 18 to Figure 21 shows that the normality assumption for the residuals is reasonable for all the univariate models except for the utilization model (Figure 20). Finally, the third panel shows a histogram of the distributions of the residuals and the corresponding box-plots.

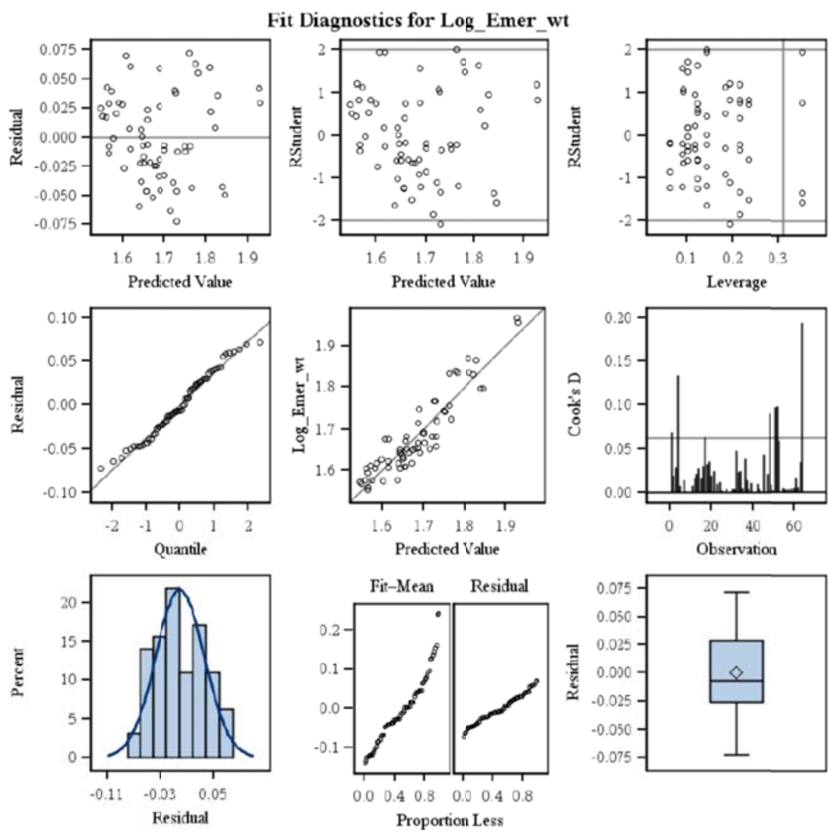


Figure 18 Plot diagnostic for linear fit of emergency wait times

.

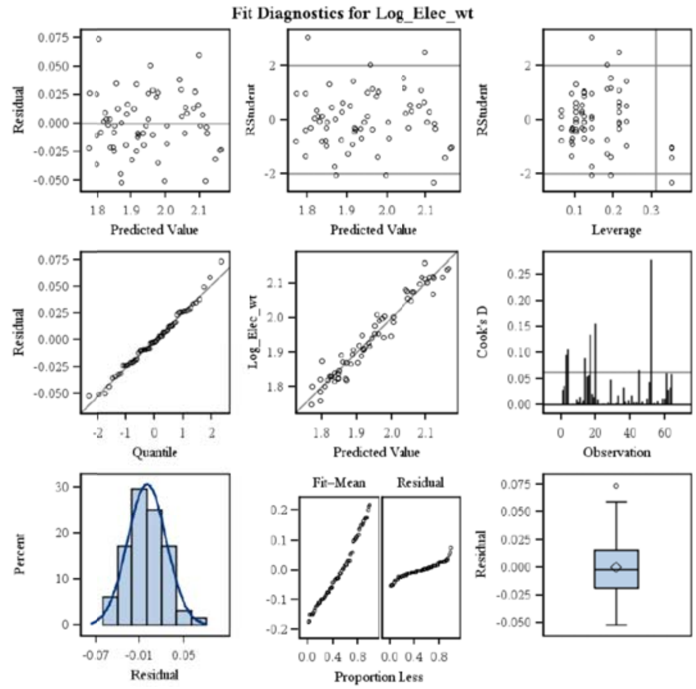


Figure 19 Plot diagnostic for linear fit of elective wait times

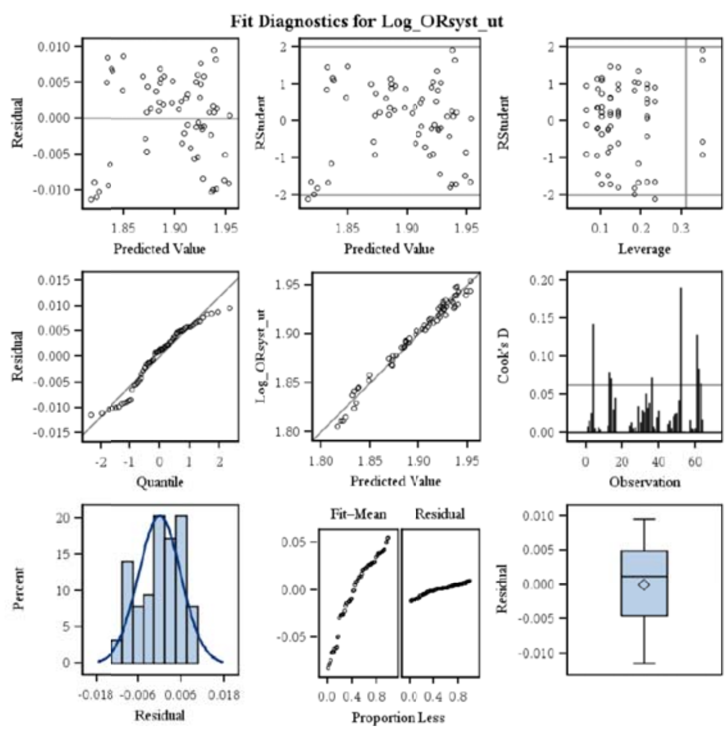


Figure 20 Plot diagnostic for linear fit of OR utilization

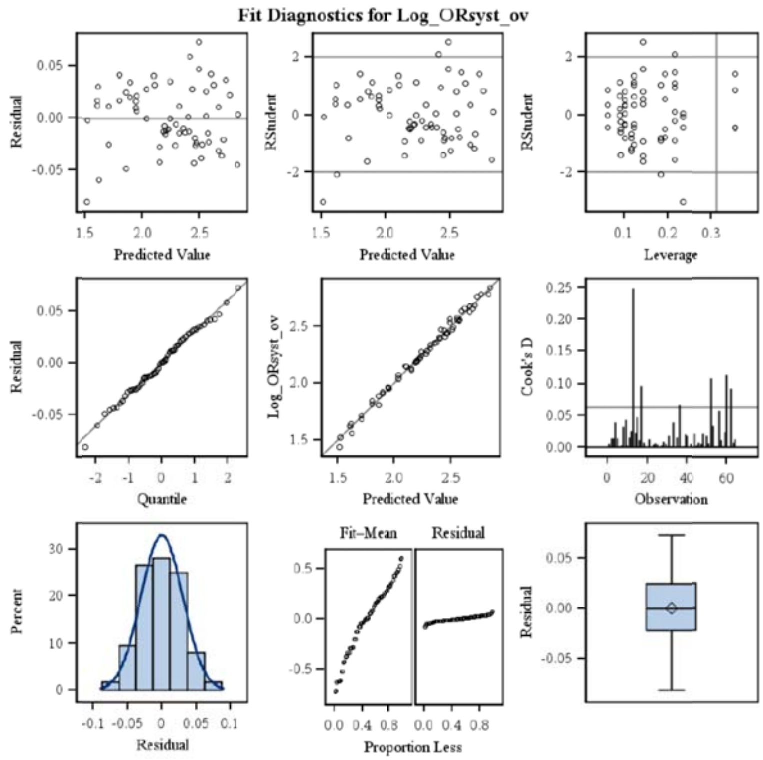


Figure 21 Plot diagnostic for linear fit of OR overtime

9. CONTRIBUTION OF RESEARCH AND FUTURE DIRECTIONS

The operating room department is one of the most costly functional areas within hospitals as well as their major profit center. The management of ORs is a complex task, involving several risks, requiring the simultaneous integration of many actors (e.g., patients, surgeons, nurses, technicians) which may have conflicting interests and different priorities. Furthermore, OR departments have to cope with the scarcity of expensive technological resources even when the demand of surgical services shows an incremental trend. The unpredictability associated with the limitation to predict the arrival time and the number of emergencies, adds on to the complexity, resulting on frequent mismatches between OR times and human resources management. Consequently, costly inefficiencies increase.

This dissertation focused on the development of a simulation based methodology, for scheduling operating rooms under uncertainty that could address one of the most important sources of ORs complexity: the incertitude and variability pertaining to the planning of the surgical operations. To accomplish this task, the process of scheduling ORs under uncertainty was separated into two components. In the first component, a real surgical dataset to design a research roadmap for modeling surgical case was used considering the volume and variability of the operations performed at a large teaching hospital in Florida. In the second component, a simulation based methodology for scheduling ORs was developed through a case study. Findings and conclusions are presented in the next section.

9.1 Summary and Conclusions

The surgical case duration was modeled using real surgical data through parametric (Generalized Linear Models, GLM) and free-distribution (Support Vector Regression, SVR) predictive methods. It was found that Support Vector Regression did better than Generalized Linear Models, increasing the prediction accuracy by at least 5.5 %. In addition, since SVR does not depend on normality, it is considered a suitable technique to minimize generalization error in the presence of outliers, an issue that affects the performance on GLM.

Next, a simulation based methodology to assist the multi-objective decision making and analysis in OR scheduling was developed. The simulation model was used to compare various performance measures of the virtual hospital OR system, using four levels of three different control factors: (1) Scheduling policy, (2) Scheduling percentile, and (3) Emergency arrival rate. Given that the virtual hospital OR system operates inside a Level I trauma center, it was required for the simulation to include a variable dedicated to the trauma/emergency OR.

Initially, schedules were built combining three levels from the scheduling policy, and the scheduling percentile control factors. Then, during the simulation process, the stochastic duration of operations and the arrival rate of emergencies control factor were considered.

Subsequently, the different scenarios were compared through four performance metrics: (1) emergency waiting times, (2) elective waiting times, (3) OR utilization, and

(4) OR overtime. It was noticed that the aggregate average of elective waiting time (95 min), almost doubled the time for the emergency cases (56 min). Regarding the virtual hospital OR system utilization and OR system overtime, values fluctuated between 67% and 88%, and 97 and 111 minutes, respectively.

An impact on the elective waiting times by changing from conservative to non-conservative scheduling policies was not apparent. In contrast, there was a clear effect on emergency waiting times when varying the control factors associated with scheduling policy and scheduling percentile. Specifically, as the percentile increased, the waiting times for emergency cases notably increased for scheduling policies Sp2 and Sp3; but had a lower impact under scheduling policies Sp1 and Sp4.

With regard to the OR system utilization and OR system overtime, it was noticed that both depend on the levels of the control factors scheduling policy, and scheduling percentile. For instance, using higher scheduling percentiles strongly decreases the OR utilization and OR overtime. However, as the emergency arrival rate is increased, both performance measures also increase.

To analyze the set of outcomes obtained from the simulation experiment, multivariate statistical techniques were used. Initially, the multivariate analysis of variance (MANOVA) reported significant difference among the means of the group of control factors on the linear combinations of the outcomes. Thus, a multivariate-multiple-regression model, often referenced in the context of the analysis of simulation outputs as a “metamodel”, was conducted. Metamodels facilitate the identification of the parameter values that optimize some system performance metrics. Thus, they were incorporated to enhance the understanding of the intrinsic dynamics of the real system under study.

9.2 Future Research Opportunities

This work provides several directions for future research. First, the prediction of individual surgical procedure duration should be based on cases of the same operation, performed by the same surgeon. Unfortunately, the majority of the surgical procedures in the data used were performed by a small number of experienced surgeons, which forced us, in order to generate reliable estimates, to merge valuable surgeon-related case duration data. Access to a larger database should prevent this limitation.

Second, other methods should be explored for estimating surgical case duration to reduce the variability within individuals. To accomplish this, data that include a broader set of factors related to patient's specific information is recommended.

Finally, the methodology presented used optimization techniques and simulation models in succession. An immediate extension would be to develop a simulation-based optimization model that integrates optimization techniques into simulation analysis.

REFERENCES

1. Litvak, E. and M. Bisognano, *More patients, less payment: increasing hospital efficiency in the aftermath of health reform.* Health Aff (Millwood), 2011. **30**(1): p. 76-80.

2. Brandeau, M.L., W.P. Pierskalla, and F. Sainfort, *Operations research and health care: a handbook of methods and applications / edited by Margaret L. Brandeau, François Sainfort, William P. Pierskalla*. International series in operations research & management science ; 70 FU FMFIU. 2004, Boston, Mass: Kluwer Academic.

3. Squires, D.A., *Explaining high health care spending in the United States: an international comparison of supply, utilization, prices, and quality.* Issues in International Health Policy, 2012. **10**: p. 1-14.

4. CMS. *National Health Expenditures 2010 Highlights*. 2012 09/25/2012]; Available from: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/downloads/highlights.pdf.

5. Vargas, L.G., et al., *Operating Room Scheduling and Capacity Planning*, in *Anesthesia Informatics*, K.R. J. Stonemetz, Editor. 2008, Springer New York. p. 361-392.

6. Guerriero, F. and R. Guido, *Operational research in the management of the operating theatre: a survey.* Health Care Manag Sci, 2011. **14**(1): p. 89-114.

7. CDC. *Inpatient Surgery National Hospital Discharge Survey, U.S. 2009.* 2012; Available from: http://www.cdc.gov/nchs/fastats/insurg.htm.

8. Harper, P.R., *A framework for operational modelling of hospital resources.* Health Care Manag Sci, 2002. **5**(3): p. 165-73.

9. HFMA, *Achieving operating room efficiency through process integration.* Healthcare Financial Management, 2003. **57**(3).

10. Cardoen, B., *Operating room planning and scheduling: solving a surgical case sequencing problem.* 4or-a Quarterly Journal of Operations Research, 2010. **8**(1): p. 101-104.

11. Cardoen, B., E. Demeulemeester, and J. Belien, *Operating room planning and scheduling: A literature review.* European Journal of Operational Research, 2010. **201**(3): p. 921-932.

12. Ozkarahan, I., *Allocation of surgical procedures to operating rooms.* J Med Syst, 1995. **19**(4): p. 333-52.

13. Tancrez, J.-S., et al., *How Stochasticity and Emergencies Disrupt the Surgical Schedule*, in *Intelligent Patient Management*. 2009. p. 221-239.

14. Eijkemans, M.J.C., et al., *Predicting the Unpredictable: A New Prediction Model for Operating Room Times Using Individual Characteristics and the Surgeon's Estimate.* Anesthesiology, 2010. **112**(1): p. 41-49.

15. Wullink, G., et al., *Closing Emergency Operating Rooms Improves Efficiency.* Journal of Medical Systems, 2007. **31**(6): p. 543-546.

16. Van Houdenhoven, M., et al., *Improving Operating Room Efficiency by Applying Bin-Packing and Portfolio Techniques to Surgical Case Scheduling.* Anesthesia & Analgesia, 2007. **105**(3): p. 707-714.

17. CDC. *Access to Trauma Care*. [cited 2012 12/01]; Available from: http://www.cdc.gov/traumacare/access_trauma.html.

18. Hans, E.W. and P.T. Vanberkel, *Operating Theatre Planning and Scheduling*, in *Handbook of Healthcare System Scheduling*, R. Hall, Editor. 2012, Springer US. p. 105-130.

19. Hancock, W.M., et al., *Operating room scheduling data base analysis for scheduling.* Journal of Medical Systems, 1988. **12**(6): p. 397-409.

20. Denton, B., J. Viapiano, and A. Vogl, *Optimization of surgery sequencing and scheduling decisions under uncertainty.* Health Care Management Science, 2007. **10**(1): p. 13-24.

21. van Oostrum, J., et al., *A master surgical scheduling approach for cyclic scheduling in operating room departments.* OR Spectrum, 2008. **30**(2): p. 355-374.

22. Stepaniak, P.S., C. Heij, and G. De Vries, *Modeling and prediction of surgical procedure times.* Statistica Neerlandica, 2010. **64**(1): p. 1-18.

23. Li, Y., et al., *Predicting surgical case durations using ill-conditioned CPT code matrix.* IIE Transactions, 2009. **42**(2): p. 121-135.

24. Strum, D.P., J.H. May, and L.G. Vargas, *Modeling the uncertainty of surgical procedure times: comparison of log-normal and normal models.* Anesthesiology, 2000. **92**(4): p. 1160-7.

25. Stepaniak, P.S.M., et al., *Modeling Procedure and Surgical Times for Current Procedural Terminology-Anesthesia-Surgeon Combinations and Evaluation in Terms of Case-Duration Prediction and Operating Room Efficiency: A Multicenter Study.* Anesthesia & Analgesia, 2009. **109**(4): p. 1232-1245.

26. Combes, C., et al., *Using a KDD process to forecast the duration of surgery.* International Journal of Production Economics, 2008. **112**(1): p. 279-293.

27. Devi, S., K. Rao, and S. Sangeetha, *Prediction of Surgery Times and Scheduling of Operation Theaters in Optholmology Department.* Journal of Medical Systems, 2012. **36**(2): p. 415-430.

28. Krupta, D.C., Shashikant Sathaye, *Reducing non-operative time: methods and impact on operating rooms economics.* International Journal of Healthcare Technology and Management, 2008. **9**(4).

29. Seim, A., B. Andersen, and W.S. Sandberg, *Statistical Process Control as a Tool for Monitoring Nonoperative Time.* Anesthesiology, 2006. **105**(2): p. 370-380.

30. Leslie, M., et al., *Using lean methods to improve OR turnover times.* AORN journal, 2006. **84**(5): p. 849-855.

31. Stepaniak, P.S., et al., *Working with a fixed operating room team on consecutive similar cases and the effect on case duration and turnover time.* Arch Surg, 2010. **145**(12): p. 1165-70.

32. Lamiri, M., F. Grimaud, and X. Xie, *Optimization methods for a stochastic surgery planning problem.* International Journal of Production Economics, 2009. **120**(2): p. 400-410.

33. Zhang, B., et al., *A Mixed Integer Programming Approach for Allocating Operating Room Capacity.* The Journal of the Operational Research Society, 2009. **60**(5): p. 663-673.

34. Van Houdenhoven, M., et al., *A Norm Utilisation for Scarce Hospital Resources: Evidence from Operating Rooms in a Dutch University Hospital.* Journal of Medical Systems, 2007. **31**(4): p. 231-236.

35. Bowers, J. and G. Mould, *Managing uncertainty in orthopaedic trauma theatres.* European Journal of Operational Research, 2004. **154**(3): p. 599-608.

36. Lamiri, M., et al., *A stochastic model for operating room planning with elective and emergency demand for surgery.* European Journal of Operational Research, 2008. **185**(3): p. 1026-1037.

37. Pang-Ning Tan, M.S.a., Vipin Kumar, *Introduction to Data Mining*. 2006, Boston: Pearson Addison Wesley.

38. Fernandez, G., *Data Mining Using SAS Applications*. 2003: Chapman&Hall/CRC.

39. Spangler, W.E., et al., *Estimating procedure times for surgeries by determining location parameters for the lognormal model.* Health Care Management Science, 2004. **7**(2): p. 97-104.

40. Strum, D.P., et al., *Estimating times of surgeries with two component procedures: comparison of the lognormal and normal models.* Anesthesiology, 2003. **98**(1): p. 232-40.

41. Franklin, D. 09/22/2012 10/03/2012]; Available from: http://www.franklindexter.net/bibliography_PredictingDuration.htm.

42. Strum, D.P., et al., *Surgeon and type of anesthesia predict variability in surgical procedure times.* Anesthesiology, 2000. **92**(5): p. 1454-66.

43. Tiwari, V.a.B., D., *Elective surgery scheduling*. 2010: INFORMS 2010 Annual Meeting.

44. Macario, A. and F. Dexter, *Estimating the duration of a case when the surgeon has not recently scheduled the procedure at the surgical suite.* Anesth Analg, 1999. **89**(5): p. 1241-5.

45. Dexter, F., A. Macario, and S.M. Cerone, *Hospital profitability for a surgeon's common procedures predicts the surgeon's overall profitability for the hospital.* J Clin Anesth, 1998. **10**(6): p. 457-63.

46. May, J.H., et al., *The Surgical Scheduling Problem: Current Research and Future Opportunities.* Production and Operations Management, 2011. **20**(3): p. 392-405.

47. Spangler, W.E., et al., *A data mining approach to characterizing medical code usage patterns.* J Med Syst, 2002. **26**(3): p. 255-75.

48. Wilcox, R.R., *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*. 2010: Springer.

49. Johnson, N.L., S. Kotz, and N. Balakrishnan, *Continuous univariate distributions*. 1995: Wiley & Sons.

50. Casella, G. and R.L. Berger, *Statistical inference*. 2002: Thomson Learning.

51. Quinn, G.P. and M.J. Keough, *Experimental Design and Data Analysis for Biologists*. 2002: Cambridge University Press. 557p.

52. Bartlett, J., *Project Risk Analysis and Management Guide 2nd edition*. 2004: Association for Project Management.

53. Young Hoon, K. and L. Ingall, *Exploring Monte Carlo Simulation Applications for Project Management.* Risk Management, 2007. **9**(1): p. 44-57.

54. Forbes, C., et al., *Statistical Distributions*. 2010: John Wiley & Sons.

55. Nelder, J.A. and R.W.M. Wedderburn, *Generalized Linear Models.* Journal of the Royal Statistical Society. Series A (General), 1972. **135**(3): p. 370-384.

56. Atkinson, A.C., *Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis*. 1985: Clarendon Press.

57. Wu, C.F.J. and M. Hamada, *EXPERIMENTS: PLANNING, ANALYSIS, AND PARAMETER DESIGN OPTIMIZATION*. 2009: Wiley India Pvt. Limited.

58. Devi, S.P., K.S. Rao, and S.S. Sangeetha, *Prediction of surgery times and scheduling of operation theaters in ophthalmology department.* J Med Syst, 2012. **36**(2): p. 415-30.

59. Cassera, M.A., et al., *Surgical time independently affected by surgical team size.* Am J Surg, 2009. **198**(2): p. 216-22.

60. O'Rourke, N., L. Hatcher, and E.J. Stepanski, *A Step-by-step Approach to Using SAS for Univariate & Multivariate Statistics*. 2005: SAS Institute, Incorporated.

61. Publishing, P.S. and S. Institute, *SAS/Stat 9.2 User's Guide: The Glm Procedure (Book Excerpt)*. 2008: SAS Publishing.

62. McLain, A.C., *An Investigation Into Multicollinearity Using Variance Inflation Factor*. 2005: University of South Carolina.

63. Hamel, L.H., *Knowledge Discovery with Support Vector Machines*. 2011: Wiley.

64. Vapnik, V., *The Nature of Statistical Learning Theory*. 1999: Springer.

65. Meshram, K.N.a.B.B. (2012) *Data Classification using Support Vector Machines*. International Journal of Engineering Research and Aplications.

66. Chin-Wei Hsu, C.-C.C., and Chih-Jen Lin. *A Practical Guide to Support Vector Classification*. 2003; Available from: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.

67. Alexandros Karatzoglou, D.M., and Kurt Hornik, *Support Vector Machines in R.* Journal of Statistical Software, 2006. **15**(9).

68. Seo, S.T., et al., *Support Vector Regression-based Model to Analyze Prognosis of Infants with Congenital Muscular Torticollis.* Healthc Inform Res, 2010. **16**(4): p. 224-30.

69. Schölkopf, B., C.J.C. Burges, and A.J. Smola, *Advances in Kernel Methods: Support Vector Learning*. 1999: Mit Press.

70. Scholkopf, A.J.S.a.B. *A Tutorial on Support Vector Regression*. 2003; Available from: http://alex.smola.org/papers/2003/SmoSch03b.pdf.

71. David Meyer, E.D., Kurt Hornik, Andreas Weingessel, and Friedrich Leisch. *Package 'e1071'*. 10/30/2012]; Available from: http://cran.r-project.org/web/packages/e1071/.

72. Gaber, M.M., *Scientific data mining and knowledge discovery [electronic resource] : principles and foundations / Mohamed Medhat Gaber, editor*. 2010: Heidelberg ; New York : Springer, c2010.

73. Meyer, D., *Support Vector Machines: The interface to libsvm in package e1071*. 2012.

74. Dexter, F., et al., *An operating room scheduling strategy to maximize the use of operating room block time: computer simulation of patient scheduling and survey of patients' preferences for surgical waiting time.* Anesth Analg, 1999. **89**(1): p. 7-20.

75. Persson, M. and J. Persson, *Analysing management policies for operating room planning using simulation.* Health Care Management Science, 2010. **13**(2): p. 182-191.

76. Hall, R.W., *Handbook of healthcare system scheduling [electronic resource] / Randolph Hall, editor*. International series in operations research & management science: v. 168. 2012: New York : Springer Science+Business Media, LLC, c2012.

77. Jebali, A., A.B. Hadj Alouane, and P. Ladet, *Operating rooms scheduling.* International Journal of Production Economics, 2006. **99**(1–2): p. 52-62.

78. Malcolm Beaverstock, A.G., Eammon Lavery, and William Nordgren, *Applied Simulation, Modeling and Analysis using FlexSim*. 2nd Edition ed. 2011: FlexSim Software Products, Inc.

79. Alexopoulos, C. *A Comprehensive Review of Methods for Simulation Output Analysis*. in *Proceedings of the 2006 Winter Simulation Conference*. 2006. Atlanta, GA, U.S.A.

80. Altiok, T. and B. Melamed, *Simulation Modeling and Analysis with ARENA [electronic resource]*. 2007: Burlington : Elsevier, 2007.

81. Harlow, L.L., *Essence Of Multivariate Thinking: Basic Themes And Methods*. 2005: Taylor & Francis Group.

82. Friedman, L.W., *Exploring Relationships in Multiple_response Simulation Experiments.* Omega, 1986. **14**(6): p. 4.

83. Friedman, L.W., *The multivariate metamodel in queuing system simulation.* Computers & Industrial Engineering, 1989. **16**(2): p. 329-337.

84. Reis Dos Santos, M.I. and A.M.O. Porta Nova. *The main issues in nonlinear simulation metamodel estimation*. in *Simulation Conference Proceedings, 1999 Winter*. 1999.

85. Khattree, R. and D.N. Naik, *Applied multivariate statistics with SAS software / Ravindra Khattree, Dayanand N. Naik*. 2nd ed ed. 1999: Cary, NC : SAS Institute; [New York] : J. Wiley & Sons, c1999.