Scheduling surgical specialties

Leveling the bed occupancy through stochastic master surgery scheduling

Thao Nguyen

Technische Universiteit Delft

TU Delft

Scheduling surgical specialties

Leveling the bed occupancy through stochastic master surgery scheduling

by

Thao Nguyen

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on 30-08-2023.

Student number:	4704924	
Thesis committee:	Dr. ir. J. T. van Essen,	TU Delft, supervisor
	Dr. L. M. Staals,	Erasmus MC Sophia Children's Hospital, supervisor
	Prof. dr. ir. K. I. Aardal,	TU Delft
	Dr. K. P. Hart,	TU Delft

TU Delft

Abstract

This research addresses the operational challenges faced by the Sophia Children's Hospital through a comprehensive analysis of its current state, literature review, and mathematical modeling. A model is created that produces a master surgery schedule, allowing for the allocation of patients to specific specialties, operating rooms, and days. Our aim is to maximize the utilization of the OR while also striving for a leveled bed occupancy and a balanced relative OR assignment for the specialties.

To address the uncertainty of future patient characteristics, we consider the surgery durations and the downstream to the nursing wards in a probabilistic manner. For the latter, we follow the approach of Schneider et al. (2020). For the first aspect, we devised a column generation based approach in which, assuming that individual surgery durations follow a log-normal distribution, we employ the Fenton-Wilkinson method to estimate the distribution of the total sum of individual surgery durations. When this distribution is known, it becomes feasible to identify pairs of specialties with corresponding surgery counts that can be scheduled within our overtime restriction. The resulting model that includes this incorporation is referred to as the Log-normal Column model.

For our research, we use historical data provided by the Sophia Children's Hospital. The data included properties about the patients' surgeries and bed assignments. Due to the presence of errors in the data, we conducted preprocessing before utilizing it as input in our modeling. Additionally, we conducted goodness of fit tests to assess whether adopting the log-normal distribution for surgery duration was genuinely superior to the normal distribution. Our analysis revealed that, for the majority of instances, the log-normal distribution outperformed the normal distribution. This was the case for individual surgeries, as well as the Fenton-Wilkinson approximation for the duration of multiple surgeries.

We compared the performance of our Log-normal Column model to two other models which assume normality for the surgery durations. One is, similar to the Log-normal Column model, created with the column generation based approach, while the other is the model described by Schneider et al. (2020). The two column generation based approach models performed significantly better than the model proposed by Schneider et al. (2020). Furthermore, we compared our Log-normal Column model to the real-life situation with the help of a simulation.

Acknowledgements

This thesis marks the end of my time as an Applied Mathematics student at the Delft University of Technology. Six years ago, my journey in Delft began, and throughout this time, I have experienced significant growth, both in my mathematical abilities and in shaping the person I am today. I could not have accomplished this journey on my own, and within this section, I wish to extend my gratitude to those who have contributed to shaping this thesis as well as my personal growth.

Firstly, I would like to thank the staff at the Sophia Children's Hospital. Especially, Lonneke, Bert, and René, for sharing their expertise and always taking the time to answer my questions.

Next, my heartfelt appreciation goes to Theresia for her invaluable guidance throughout this project. Her patience, understanding, and unwavering support extended beyond the confines of academia. She even went above and beyond by taking care of my boyfriend's baby niece for an hour or so.

I am also thankful for Karen and KP. Since the beginning of my studies, I have considered them as my mentors. I could always pass by to ask for advice or just a chat, and it brings me great joy to conclude my time in Delft with them in my committee.

Furthermore, I would like to thank my parents. Their hard work has paved the way for the opportunities I now enjoy, opportunities that I occasionally overlook. My gratitude towards them knows no bounds, and it is thanks to them that I have achieved my educational milestones.

Lastly, I want to thank my friends and TU Delft staff. Engaging in regular social interactions during the course of my master's project significantly shortened my days and enhanced their overall enjoyment. I would like to express my appreciation to Joop, for steadfastly standing by my side and even going the extra mile by staying up late to keep my motivation alive. To Björn, for consistently lending an ear to my mathematical ramblings. To my study buddies, which started with Joop, Björn, Ruby, Stan, occasionally joined by Leonie, Célio and Kristy, and ended with Jenny, Lander and his 'fourth-floor gang', for all the study sessions in EWI. To 'my hermanito Roro', for the occasional hot chocolate breaks on the eighteenth floor, that were anything but consistent. Certainly, I extend my gratitude to my other friends as well. Although their assistance was not directly related to my writing, they provided invaluable support in various other aspects of my life.

Thao Nguyen
Delft, August 2023

Contents

1 Introduction **1**
1.1 Context . . . 1
1.2 Problem description . . . 1
1.3 Thesis outline . . . 2

2 Situational analysis **3**
2.1 Surgical specialties . . . 3
2.2 Hospital departments . . . 4
2.3 Pathway of a surgical patient . . . 5
2.4 Healthcare planning . . . 6
2.5 Master surgery schedule . . . 7

3 Literature review **9**
3.1 OR scheduling regarding downstream resources . . . 9
3.2 Stochastic master surgery scheduling . . . 10

4 Mathematical model **13**
4.1 Model restrictions . . . 13
4.2 Post-operative length of stay distribution . . . 15
4.2.1 Step 1: Specialty specific distribution for one planned surgery . . . 17
4.2.2 Step 2: Specialty specific distribution for one surgical patient within the MSS . . . 19
4.2.3 Step 3: Distribution bed occupancy . . . 20
4.3 Objective function . . . 21

5 Solution methods **23**
5.1 Linearization overtime constraint under a log-normal distribution . . . 23
5.1.1 Fenton-Wilkinson method . . . 24
5.1.2 Schwartz-Yeh method . . . 24
5.1.3 Mehta approximation . . . 26
5.1.4 Method selection . . . 28
5.1.5 Column generation based approach. . . . 30
5.2 Linearization overtime constraint under a normal distribution . . . 30
5.3 Linearization objective function . . . 33
5.4 Resulting linear problem under log-normality . . . 35
5.5 Resulting linear problem under normality . . . 36

6 Data analysis **37**
6.1 Data description . . . 37
6.2 Preprocessing surgery durations . . . 38
6.3 Preprocessing bed assignment . . . 40
6.4 Fitting surgery durations . . . 40
6.4.1 One surgery. . . . 41
6.4.2 Multiple surgeries . . . 42

7 Results **45**
7.1 Performance under different objective weights . . . 45
7.2 Conversion to time schedule . . . 49
7.3 Performance of Log-normal Column model . . . 50
7.4 Performace of adjusted model . . . 53

8 Conclusion and recommendations **57**
8.1 Conclusion . . . 57
8.2 Recommendations . . . 59
8.2.1 Recording of data . . . 59
8.2.2 Future research . . . 59
A Model parameters and variables **61**
B Explicit formulas considered paths **65**

1

Introduction

Operation management has become more and more important in hospitals, due to an aging population, increasingly scarce resources and a shortage in personnel (Berden et al., 2016). One of the most significant expenses in the hospital is the operation room resources. As the surgical waiting lists at hospitals have only increased during the COVID-19 pandemic, the question of how to make optimal use of the hospital resources has only become more prominent. At the Sophia Children's Hospital, it is expected that optimizing the assignment of operating rooms to the surgical specialties based on historical patient data could be an option to reduce the number of patients on the waiting list while keeping the bed occupancy as leveled as possible. In this thesis, we investigate whether this is indeed the case and study the effect of the proposed schedule.

1.1. Context

The Sophia Children's Hospital (or in short: Sophia) is the first and largest children's hospital in the Netherlands, located in Rotterdam. It is part of the Erasmus Medical Center, the biggest Dutch academic teaching hospital. Sophia is one of the eight hospitals in the Netherlands that specializes in child health care. Due to their specialization, complicated pediatric cases from the area are referred to Sophia. Hence, the patients are typically children with complex or rare disorders.

1.2. Problem description

In order to schedule children's surgeries, each quarter of the year a 'Master Surgery Schedule' (MSS) is determined. The MSS states when and which operating room (OR) is assigned to a certain specialty. For convenience, the MSS is an OR schedule with a repetition for every four weeks within the quarter. Based on the MSS, patients are scheduled in the OR time belonging to the specialty of their surgeon.

In theory, the MSS can be changed each quarter to suit the situation. However, this is quite difficult as many factors have to be taken into account. First of all, many surgeons have requirements outside the OR and are hence not available on each day. They have appointments at the outpatient clinic or they could partly work at other locations besides Sophia. Secondly, there are only a limited number of beds for patients to stay in after their surgery. Certain specialties also share a ward and assigning those specialties' ORs around the same time may create an impossible situation to handle downstream. Thirdly, some ORs are specially equipped for certain surgeries and are thus reserved for specific (sub)specialties.

In this research, we disregard the first two constraints presented above and investigate what the best MSS would be. Instead of having a limitation on the number of beds for surgical patients, we want to predict the number of beds that are needed at the medium care and intensive care for surgical patients. As the medium care and intensive care are not used exclusively for surgical patients, taking into account the number of beds in the objective function of our model instead of in a constraint prevents the overestimation of the number of beds for non-surgical patients.

From historical patient data, predictions can be made about the number of elective and non-elective surgeries each specialty can expect to carry out within one quarter and about the duration of a surgery. Also, the type of nursing ward and the length of stay at a nursing ward after a surgery of a certain specialty can be predicted. These features are represented as probability distributions and are taken into account as input variables for our model.

There are many objectives we can optimize for. At Sophia, they recognize the availability of beds at the nursing wards as a major bottleneck for processing the waiting list. As the MSS has a direct effect on the downstream departments (Vanberkel et al., 2011), taking the downstream resources into account when creating an MSS can prevent undesired consequences such as large variations in the number of used beds. In our research, the goal is to make an MSS that leads to a leveled bed occupancy at each nursing ward, without compromising the OR utilization and the OR assignment of specialties.

In this thesis, we developed a model which takes all three objectives into account. We assume both the surgery duration and the length of stay are given by log-normal distributions. In contrast to the normal distribution, the log-normal distribution does not have a fixed form for its sum. To address this challenge, we explored other research fields for suitable approximation methods and developed a model based on column generation. As the log-normal distribution can be a better fit for surgery durations due to their skewness and positive valued nature, our model could give a better result regarding overtime.

1.3. Thesis outline

This thesis is structured in the following way. We start by providing an overview of the current situation in the Sophia Children's Hospital in Chapter 2. Here we give some specific information about the Hospital by detailing various components and processes within the hospital. Among other things, we describe the various surgical specialties and departments in the hospital and explain the MSS in more detail. Furthermore, we give a description of the patients' pathways and identify the most prominent pathways.

Next, Chapter 3 gives a summary of pertinent literature in surgery planning. We look at research with a focus on scheduling specialties or patient groups and optimizing bed occupancy levels. We further concentrate on research considering patient characteristics in a stochastic manner.

In Chapter 4, a mathematical model is presented, which is inspired by the preceding chapter's literature. This model has two non-linear components: a non-linear overtime constraint and a non-linear sub-objective concerning the bed variation. Chapter 5 presents solution approaches for the model of Chapter 4. We present two possible linearizations for the overtime constraint: one for the assumption that surgery durations are log-normally distributed and one for the assumption that they are normally distributed. We developed the linearization for log-normally distributed surgery durations ourselves. Moreover, we also give a linearization of the objective function.

For our model, we use data provided by the Sophia Children's Hospital. In Chapter 6, this data is described and preprocessed. Additionally, we conducted goodness of fit tests to assess whether the surgery durations are more likely to follow the log-normal distribution compared to the normal distribution. The results of our model are thereafter described in Chapter 7 where we compare its performance to other models. This thesis is concluded with the conclusion and recommendations in Chapter 8.

2

Situational analysis

As this thesis features a case study at the Sophia Children's Hospital, some specific information from the hospital is required. In this chapter, we describe various components and processes. Firstly, the surgical specialties and hospital departments are listed in Sections 2.1 and 2.2, respectively. Thereafter, a description of the pathway of a surgical patient is given in Section 2.3. In Section 2.4, we give a brief explanation of healthcare planning and conclude the chapter with a description of the master surgery schedule in Section 2.5.

2.1. Surgical specialties

In an MSS of the Sophia Children's Hospital, the specialties mentioned in Table 2.1 are scheduled.

Table 2.1: Abbreviation of surgical specialties.

Abbreviation	Specialty
FLEX	Flex program (non-elective)
GYN	Gynaecology
NSP	Pediatric neurology
PDC	Pediatric cardiology
PDD	Pediatric dermatology
PDE	Pediatric dentistry
PDO	Pediatric otolaryngology
PG	Pediatric gastroenterology
PMFS	Pediatric maxillofacial surgery
PNS	Pediatric neurological surgery
PDP	Pediatric pulmonology
PDR	Pediatric radiology
PDS	Pediatric (general) surgery
PO	Pediatric ophthalmology
PORS	Pediatric orthopaedic surgery
PPS	Pediatric plastic surgery
PU	Pediatric urology

The odd one out of this table is PDR, as this does not involve a surgical procedure. However, as the patients at Sophia are generally children, they have their scans while anesthetized. Furthermore, the procedures for the specialties PPS and PORS can be split naturally into two categories and these two specialties are hence divided into two subspecialties in the MSS we generate. For PPS, the subspecialties are craniofacial surgery (FPS) and hand surgery (HPS). For PORS, we make a distinction between open spine surgery (PORS2) and other orthopedic surgical procedures (PORS).

2.2. Hospital departments

Hospitals have different departments with each their own function. In this subsection, we describe the departments that a surgical patient typically visits during their treatment(s).

Outpatient clinic: In the outpatient clinic, patients are treated who do not require a bed during their hospital visit. A patient typically visits the outpatient clinic to talk to a specialist for a consultation or to have a small examination or treatment.

Operating rooms: At Sophia, there are eleven ORs available. One of these ORs is located apart from the others in the Obstetrics department. As this OR is not provided with sufficient airflow quality for most surgeries, it is exclusively used for cesarean sections and is not taken into account in our research.

Due to the requirement for children to remain still during an MRI, patients undergoing this procedure are also administered anesthesia. Consequently, the MRI room is considered and treated as an operating room within the MSS.

For elective surgeries, the ORs are open from 8.00-15.30 on weekdays. Some ORs are provided with special medical equipment and are thus reserved for specific procedures or (sub)specialties. In Table 2.2, the ORs reservations can be found and they are included as constraints in our model.

Table 2.2: ORs reserved for special usage.

OR	Specialty	Additional information
OR1	GYN	Reserved for emergency cesarean section
OR2	PDP, PDO	Available bronchoscopy
OR 8	e.g. PNS, PORS	Better ventilation for sterility
OR 9	PDC	Reserved for cardiac catheterization
OR 10	e.g. PNS, PORS	Better ventilation for sterility

Recovery unit: After a surgery, the patient stays at the recovery unit. Here, the patient is closely monitored while the patient recovers from the anesthesia. Afterwards, the patient is transferred to one of the nursing wards. If the patient was assigned beforehand to the intensive care (IC) nursing ward, the patient is taken to the IC immediately after the surgery without visiting the recovery unit. The duration a patient spends at the recovery is approximately 30 minutes. At Sophia, there are eight places available for patients. As it rarely happens that the bed occupancy at the recovery becomes a threshold for the downstream of patients, we do not take the recovery unit into account in our research.

Nursing wards: After a patient's stay in the recovery room, the patient is moved to a bed at a nursing ward. At Sophia, there are 3 different types of wards:

- day care (DC),
- medium care (MC),
- intensive care (IC).

At the day care, there are ten to fifteen beds available for surgical patients to stay. The patients are assigned to the DC when they are expected to be discharged on the same day as their surgery. The DC is open on Monday till Friday from 7.30 to 18.00.

The medium care is divided into five sections, which we denote by MC1, MC2, MC3, MC4 and MC5. Multiple specialties share the first four sections. Not all patients at the MC have to undergo surgery to stay at the MC. Most children that stay in the first three sections of the MC are surgical patients, while the patients of MC4 are not. Table 2.3 provides an overview of the medium care sections and their associated specialties. The layout presented is not strictly binding and can be subject to variations.

Children in intensive care are treated with (acute) life-threatening conditions requiring intensive monitoring and care. Similar to the patients in the medium care, these children did not have to undergo surgery in order to stay at the IC. Children who had surgery, skip the recovery unit and go straight to the IC. Each day, three beds are available for elective surgical patients and it is desired to always set aside at least 1 bed for emergencies at the IC. There are eight intensive care units: four neonatal intensive care units (NIC1, NIC2, NIC3, NIC4) and four 'normal' intensive care units (IC1, IC2, IC3, IC4). Each intensive care unit comprises a hall housing six beds and a nursing station. IC1 is typically allocated to patients who need a bit more care than an MC can provide and is hence nicknamed MC+. IC2 is used primarily for patients with cardiac and pulmonary conditions, while IC3 is used primarily for neurology patients. IC4 is considered a high care unit: while patients in this unit require less intensive care compared to those in IC2 and IC3, they still require close monitoring.

The non-elective patients who had their surgeries during the FLEX are moved to the nursing unit corresponding to the specialty of their surgeon and hence, FLEX is not included in Table 2.3. For our research, we look at the FLEX as a separate specialty and allocate their patients to the wards similarly as found in the historical data.

Table 2.3: Medium care sections and their specialties.

MC1	MC2	MC3	MC4	MC 5
NSP	PDS	PDC	PG	GYN
PDO	PU	PDP		
PMFS	PORS	PDD		
PNS				
PPS				
PDE				
PO				

2.3. Pathway of a surgical patient

There are three ways for a (surgical) patient to get into the Sophia Children's Hospital for the first time:

- during the patient's birth,
- for a visit to the outpatient clinic due to health problems (possibly by referral of an external pediatrician),
- as an emergency patient.

A surgical elective patient is first seen by a physician at the outpatient clinic for consultation. Generally, this physician is also the surgeon who performs the patient's surgery or a surgeon of the same specialty as the performing surgeon. During the consultation(s), the physician makes a diagnosis and a plan of action for the treatment of the patient. When the physician decides that the patient needs surgery, the patient is put on the waiting list with an indication of the latest date for performing the surgery. Before the surgery, the patient is seen by an anesthetist one last time at the outpatient clinic for a preoperative screening.

On the day of the surgery, the patient is prepared for the surgery (changed in suitable clothes and small checks are executed) before the patient is taken to the OR. One parent of the patient is allowed to accompany the patient to the OR and can stay until the patient has fallen asleep from the anesthesia.

After the surgery, the patient is moved to either the recovery unit or the IC depending on their condition. Once the anesthesia wears off, patients at the recovery unit are moved to their corresponding nursing ward which is the DC or the MC. Due to unforeseen circumstances, a patient at the DC may need to stay longer than expected and is hence moved to the MC. Similarly, patients may be moved from the MC to the IC if a patient's condition worsens, or from the IC to the MC if a patient's condition improves. However, for patients going to the DC after their surgery, almost all are discharged without being transferred to another nursing ward. Hence, we assume that there is no displacement between the DC and other nursing wards. Figure 2.1 gives a visualization of the described pathway.

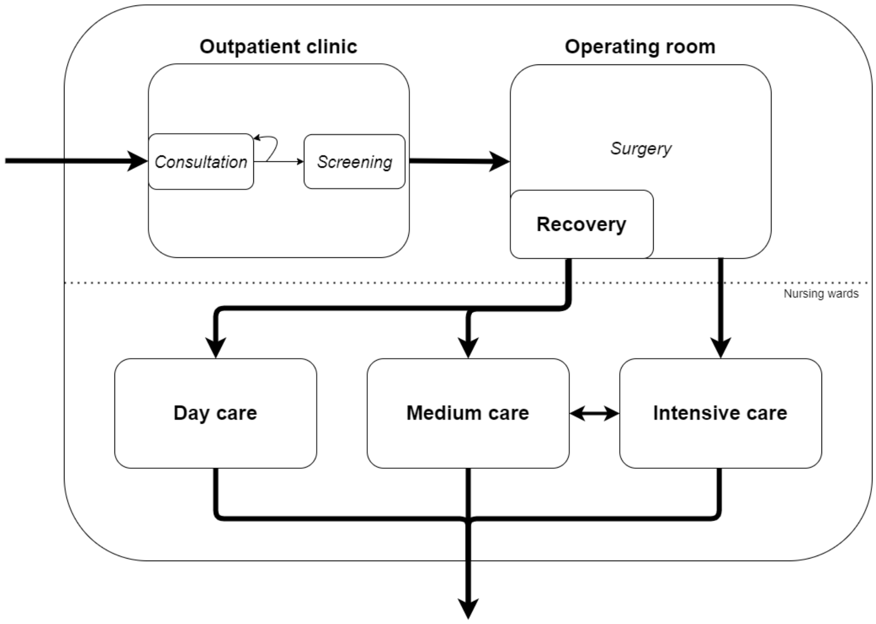


Figure 2.1: General path of a surgical patient.

On rare occasions, for example when complications occur, the patient's route differs from the possible pathways shown in Figure 2.1. To simplify our problem, we only look at the pathways that are most common at Sophia and only take these in account when making our model.

From the data, we see that 89% of the patients go to only one nursing ward during their hospitalization. The percentages for the DC, MC and IC are 57%, 31% and 1%, respectively. When patients move between the MC and IC, this most often happens in the pattern MC - IC - MC, followed by the iterations MC - IC and IC - MC. These patterns make up respectively 5%, 2% and 1% of the total patterns. The six mentioned pathways are the only ones we consider. They can be generalized by just three paths based on the starting nursing wards, i.e., DC, IC-MC, MC-IC-MC. When a patient is discharged halfway through one of the two latter paths, we let the length of stay at the remaining wards in the path be equal to zero days. However, in addition to the fact that staying solely at the MC before getting discharged is much more common than MC-IC and MC-IC-MC, patients are more likely to stay longer at the MC if they do not get transferred. Hence, we decided to make a distinction between going solely to the MC and the other two paths starting at the MC. The resulting paths are depicted in Figure 2.2.

2.4. Healthcare planning

In management, there are various levels of planning: strategic, tactical and operational. Of the three levels, the strategic one is the highest planning level. On a strategic level, ultimate goals and long-term directions are set. Meanwhile, on an operational level, short-term plans for the implementation of specific processes are made. The tactical level is located between the strategic and operational levels. To support and achieve the goals described at the strategic level, intermediate-term plans of the operational level are made on a tactical level (Hans et al., 2012).

As our research concerns the allocation of the ORs to the surgical specialties in the form of a master surgery scheduling, our problem belongs to the tactical planning level. Based on the MSS, surgeries are scheduled during the OR session of the corresponding specialty. The scheduling of the surgeries happens on the operational level.

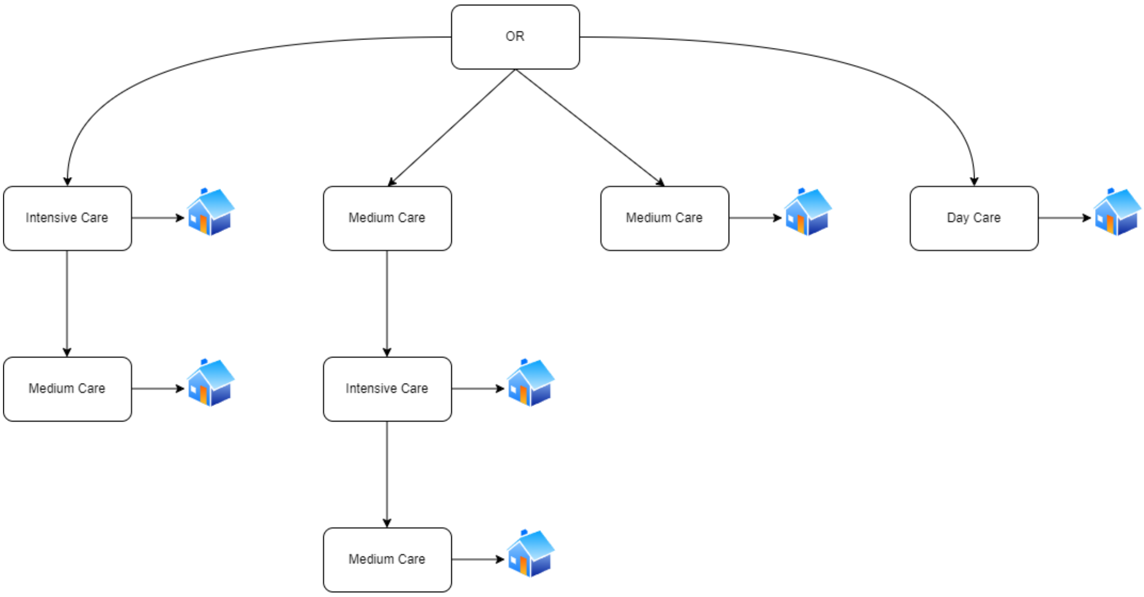


Figure 2.2: Considered patient pathways from OR to discharge.

2.5. Master surgery schedule

The master surgery schedule is an OR schedule for a quarter of a year, which states when and which OR is assigned to a certain specialty. For convenience, the MSS at the Sophia has a repetition every four weeks within the quarter. Based on the MSS, patients are scheduled in the OR time appointed to the specialty of their surgeon. ORs are scheduled between 8.00 and 15.30 from Monday till Friday. In Table 2.4, the assignments for one week of the current MSS at the Sophia Children's Hospital can be found. The MSS also provides the time intervals the OR is available for the assigned specialty. This usually spans a duration ranging from 1.5 hours to 7.5 hours.

Table 2.4: OR assignment for the first week of the current MSS.

Day	OR1	OR2	OR3	OR4	OR5	OR6	OR7	OR8	OR9	OR10	MRI
Monday	GYN	PDP	PDS FLEX		PU PU	PPS PPS		PORS PORS	PDC PDC	PNS PNS	PDR PDR
Tuesday	GYN	PDO PDO	PG FLEX	 NSP	PDO PDO		PDS PDS	PORS PORS		PPS PPS	PDR PDR
Wednesday	GYN			PDO PDO	PU PU	PMFS PMFS	PDS PDS	PORS PORS	PDC PDC	PPS PPS	
Thursday	GYN	PDP	PDS FLEX	 PORS	PU PU	PPS PPS	PDS PDS	PORS PORS		PNS PNS	
Friday	GYN		PDO PDO	PO PO	PU PU	PDS FLEX	PDS PDS	PORS PORS		PORS PORS	PDR PDR

3
Literature review

Extensive research has been conducted regarding OR planning. Section 3.1 provides an overview of previous studies focusing on scheduling specialties or patient groups and optimizing bed occupancy levels. Section 3.2 is dedicated to studies with a specific focus on considering patient characteristics in a stochastic manner.

3.1. OR scheduling regarding downstream resources

As mentioned previously, a lot of research has been dedicated to operating room scheduling. Many studies have focused on creating MSS's by arranging specialties or patient groups with similar properties. Although we do not consider distinct patient groups within a specialty, it is still worthwhile to consider these research findings, as the solution methods proposed can potentially be adapted for scheduling various specialties.

In their literature review, Wang et al. (2021) identify publications regarding operation room scheduling and its effect on the downstream resources. They make a distinction between patients who leave the hospital on the same day and patients that stay overnight. The former are referred to as outpatients, while the latter are referred to as inpatients. In our research, we aim to incorporate both considerations, specifically focusing on inpatient cases. In this chapter, we have chosen some relevant articles from Wang et al. (2021) to elaborate on.

Beliën and Demeulemeester (2007) were one of the first to propose models for developing cyclic master surgery schedule that take the bed occupancy into account. In a master surgery schedule, they planned fixed blocks of surgeries of the same type. They had three different objectives regarding the bed leveling: they looked at minimizing the highest expected number of beds, minimizing the highest variance in bed occupancy and a combination of the two. Furthermore, they assumed the length of stay to be following a multinomial distribution and the number of surgical patients per block to be deterministic. Beliën and Demeulemeester (2007) considered only two types of constraints: surgery demand constraints and OR capacity constraints. They recognized that for real-life applications extra restrictions may be needed.

Beliën et al. (2009) made an extension of their work in Beliën and Demeulemeester (2007). They assumed that both the number of surgical patients per OR block and the length of stay followed multinomial distributions. Furthermore, they also took into account multiple wards and assigned the OR blocks to individual surgeons instead of surgeon groups. Instead of fixed OR blocks, they allowed the blocks to have varying sizes. The model not only aims to level the bed occupancy, but also aims to share ORs as little as possible between specialties and make the master surgery schedule as simple and repetitive as possible. All these objectives are relatable for the Sophia Children's Hospital as well. For our research, we have not explicitly taken into account the sharing of ORs between specialties. Instead, we take it into account implicitly by requiring extra cleaning time on the day when different specialties are sharing an OR.

As the operating rooms have a direct effect on the workload in the wards, Vanberkel et al. (2011) investigated how to calculate the expected workload so that nursing ward managers could use this as a decision support tool. Vanberkel et al. (2011) describe an exact approach to calculate the distributions for the bed occupancy, patient admission and discharge.

Ma and Demeulemeester (2013) propose a three-stage integrative approach. In the first stage, they looked from a financial standpoint of the hospital and propose a model where the patient set was found that brings the maximum total financial contribution within the specified resource capacity. In the second stage, they build a cyclic master surgery schedule in which the total expected bed shortage is minimized and the length of stay is described by a discrete probability distribution. In the last stage, the performance of the model is evaluated through a simulation analysis.

Schneider et al. (2020) developed a model for scheduling surgery groups with the aim of optimizing operating room (OR) capacity while minimizing variations in bed occupancy across different wards. The researchers made the assumption that patients can follow only two distinct paths through various ward types. To calculate the expected number of beds, Schneider et al. (2020) relied on the methodology used in Vanberkel et al. (2011) and Fügener (2015) in their integrated approach. The researchers employed two different strategies: a global approach utilizing mixed integer linear programming and a simulated annealing approach. Simulated annealing is a probabilistic optimization algorithm inspired by the annealing process in metallurgy. Upon comparing the performance of these two approaches in a real-life setting, the study found that the global approach outperformed the simulated annealing approach. This suggests that the global approach was more effective in achieving the desired objectives of optimizing OR capacity and minimizing variation in bed occupancy.

3.2. Stochastic master surgery scheduling

In their literature review, Wang et al. (2021) also give special attention to articles that incorporate the uncertainty in future patient characteristics and how they do so. They found that the log-normal distribution was utilized in most studies to represent the surgery duration. 46 out of 67 studies that consider the surgery duration in a stochastic manner do this by using the log-normal distribution. This happens because it is widely recognized that durations display a strong adherence to the log-normal distribution (see, e.g. Stepaniak et al. (2009), Kayış et al. (2015), Sagnol et al. (2018)).

Jebali and Diabat (2015) and M'Hallah and Visintin (2019) both propose stochastic models to schedule surgery groups within a given MSS. They take into account surgery durations and length of stay stochastically and assume that both types are log-normally distributed. However, the setting Jebali and Diabat (2015) consider is simpler and they focus on minimizing the costs while M'Hallah and Visintin (2019) focus on maximizing the number of scheduled patients. The model of M'Hallah and Visintin (2019) consists of two stages. The first stage optimizes the expected number of elective surgical patients for each specialty. When a solution (i.e., a scenario with a certain number of patients for each specialty) is found in the first stage, they determine how many surgeries would actually be carried out for different samples of the surgery durations and length of stays. Afterwards, they take the average of each sample to get an approximation for the expected number of elective surgical patients for each specialty. We could also make such a two-stage stochastic program for the assignment of specialties to ORs. The downside of using stochastic programming would, however, be that we would typically get a set of scenarios or the average of the set, instead of one concrete MSS.

While Schneider et al. (2020) model the length of stay for patients by their discrete empirical distributions, they do consider the surgery durations by probability distributions. They mention that Stepaniak et al. (2009) had shown that when fitting the normal distribution and log-normal distributions, the log-normal distribution performed better. However, they decided to utilize the normal distribution for the duration of surgeries, as it allowed for an exact expression when considering the distribution of the overall sum of surgeries. This was used in their constraint to restrict the allowed overtime. In order to linearize this constraint, the constraint was rewritten and the square root that appeared was approximated by a piecewise linear function. In Section 5.2, we elaborate on this further.

4

Mathematical model

In this chapter, we describe the model constructed for our research. In the first section, the constraints of our model are formulated. In the next section, we describe an analytical model for determining the bed occupancy. Lastly, in Section 4.3, the objective function of our model is given. In Appendix A, a summary of all parameters and variables considered in this chapter is given.

4.1. Model restrictions

Let S be the set of specialties and O the set of ORs. The elements in set S are ordered by their index within the set. The index of specialty $s \in S$ is indicated by $\hat{s}$. Define $S_o \subset S$ as the set of specialties that can operate in OR $o \in O$. Let Q be the number of days the MSS covers including the days the elective surgeries are not scheduled (in our case: the weekends) and let K denote the set of days that the ORs are open.

We use three types of variables:

- Z_{oks}: an integer decision variable representing the number of surgeries planned in OR $o \in O$ for specialty $s \in S_o$ on day $k \in K$,
- Y_{oks}: a binary auxiliary variable denoting for OR $o \in O$ which specialty $s \in S_o$ is assigned on day $k \in K$,
- X_{ok}: a binary variable denoting whether OR $o \in O$ is open on day $k \in K$.

These variables are connected with each other by the following constraints:

$$Z_{oks} \leq M_s \cdot Y_{oks}, \quad \forall o \in O, \forall k \in K, \forall s \in S_o, \tag{4.1}$$

where M_s denotes the maximum number of patients specialty $s \in S$ can schedule on one day in one OR.

Moreover, OR $o \in O$ can only be assigned to at most two specialties from S_o on day $k \in K$ when it is open on day $k \in K$. We only permit this when two specialties can be 'combined' in one OR. Let sets I_s, for each $s \in S$, be the set of specialties s_2 such that $\hat{s}_2 < \hat{s}$ and s_2 can be combined with $s \in S$. This is established with the following constraints:

$$\sum_{s \in S_o} Y_{oks} \leq 2 \cdot X_{ok}, \qquad \forall o \in O, \forall k \in K, \tag{4.2}$$

$$Y_{oks_2} + Y_{oks} \leq 1, \qquad \forall o \in O, \forall k \in K, \forall s \in S_o, \forall s_2 \notin I_s. \tag{4.3}$$

Besides this, we want that on each day $k \in K$ at most χ_k ORs are used in order to keep some rooms and staff in reserve in case of emergency. Let $\bar{O}_k \subset O$ be the subset of ORs that cannot be closed on day k. Then, the following constraint needs to be satisfied:

$$\sum_{o \in O \setminus \bar{O}_k} X_{ok} \leq \chi_k - |\bar{O}_k|, \quad \forall k \in K. \tag{4.4}$$

Each specialty has a different number of available surgeons. We have to restrict the number of ORs a specialty gets assigned on a day by the number of surgeons they have. Denoting the number of surgeons available for specialty $s \in S$ as C_s, we construct the following constraint:

$$\sum_{o \in O : s \in S_o} Y_{oks} \leq C_s, \quad \forall s \in S, \forall k \in K. \tag{4.5}$$

Furthermore, we wish to ensure that our MSS has a cycle of T days. The following constraints are added to ensure this:

$$\begin{aligned} Y_{oks} &= Y_{o(k+T)s}, && \forall o \in O, \forall s \in S_o, k \in \{1, ..., T\}, \\ Y_{oks} &= Y_{o(k+2T)s}, && \forall o \in O, \forall s \in S_o, k \in \{1, ..., T\}, \\ &\vdots \\ Y_{oks} &= Y_{o\left(k+\left\lfloor \frac{Q}{T} \right\rfloor \cdot T\right)s}, && \forall o \in O, \forall s \in S_o, k \in \{1, ..., (Q \bmod T)\}. \end{aligned} \tag{4.6}$$

The cyclic property only has to hold for the assignment of specialties to ORs, i.e., for the variables Y_{oks}, and not for Z_{oks}. Hence, the number of surgeries that a specialty schedules on their assigned OR day can fluctuate.

When planning the surgeries in an OR, there is a possibility that the total surgery duration takes longer than planned. While overtime can happen, it is not desirable for it to happen too often. Let g_{ok} denote the probability distribution of the total surgery duration that is scheduled in OR $o \in O$ on day $k \in K$ and let β_{ok} denote the duration that OR $o \in O$ was planned to be open on day $k \in K$. However, after each surgery we would like to reserve some time of length κ for cleaning and an additional time of length κ if there is a change in specialty in the OR. For this, we introduce the following auxiliary variables:

- R_{oks}: a binary variable indicating whether in OR $o \in O$ specialty $s \in S_o$ has at least one surgery on day $k \in K$,
- W_{ok}: a binary variable indicating whether in OR $o \in O$ exactly two specialties in S_o had surgeries planned on day $k \in K$.

The following constraints ensure that R_{oks} can only be equal to one if Z_{oks} is greater or equal to one:

$$R_{oks} \geq \frac{Z_{oks}}{M_s + 1}, \quad \forall o \in O, \forall k \in K, \forall s \in S_o, \tag{4.7}$$

Furthermore, the variables R_{oks} and W_{ok} are connected with each other by the following constraints:

$$W_{ok} \geq \sum_{s \in S_o} R_{oks} - 1, \quad \forall o \in O, \forall k \in K. \tag{4.8}$$

Now, we can use the newly introduced variables W_{ok} to indicate if extra time should be reserved for the change in specialty in OR $o \in O$ on day $k \in K$. Let $\mathbf{g}_{ok}$ be the corresponding stochastic variable of g_{ok}. Then, the following constraint guarantees that the probability of overtime is less than α.

$$\mathbb{P}\left(\mathbf{g}_{ok} \geq \beta_{ok} - \kappa \cdot \left(\sum_{s \in S_o} Z_{oks} + W_{ok}\right)\right) \leq \alpha, \quad \forall o \in O, k \in K. \tag{4.9}$$

Note that Constraint (4.9) is not a linear constraint. In Section 5.1, it is shown how the constraint can be linearized.

Now, let $S_o^C \subset S_o$ be the subset of specialties that can only have surgeries planned in OR $o \in O$ in combination with another specialty. Then, the following constraints ensure that surgeries of specialties $s \in S_o^C$ are not planned alone in an OR:

$$\sum_{s_2 \in S_o,\ s_2 \neq s} R_{oks_2} \geq R_{oks}, \quad \forall o \in O, \forall k \in K, \forall s \in S_o^C. \tag{4.10}$$

Lastly, it is unfavorable for surgeons to have to go to an OR for just a short amount of time. Hence, a constraint is added that for specialty $s \in S$ the number of scheduled surgeries is at least m_s. Thus,

$$Z_{oks} \geq m_s \cdot R_{oks}, \quad \forall o \in O, \forall k \in K, \forall s \in S_o. \tag{4.11}$$

4.2. Post-operative length of stay distribution

In this section, a model is described to compute the distribution of the number of post-surgical patients in the nursing wards for a given MSS. To calculate the post-surgical patient distribution over the nursing wards, we follow the approach described by Schneider et al. (2020) with two main differences. The approach of Schneider et al. (2020) consists of 3 steps. In the first step, the patient distribution for the nursing wards after one OR session is calculated for each specialty. This step is expanded in our case to include more patient pathways. Thereafter, the patient distribution of one OR session within the MSS is calculated for each specialty. In this step, the fact that a patient can be discharged in a later cycle than the cycle of their surgery is taken into account. Contrary to Schneider et al. (2020), we do not appoint surgeries to only one cycle within the MSS but to all cycles of the MSS. Hence, this step would not be necessary. However, we are going to assume that our schedule of Q days is also repeated in order to investigate the bed occupancy during a steady state. Lastly, the patient distributions of the specialties are combined in order to compute the total bed occupancy of the nursing wards.

Let I be the set of ICs, W the set of MC wards and let D denote the daycare department. Let P represent the set of considered post-surgical patient subpaths through the nursing ward departments. As introduced in Section 2.3, we consider six different paths for patients to follow after their surgery up until their discharge. In this section, we use ‘path’ if the nursing department itinerary can only be followed by discharge and ‘subpath’ if the patient can still be transferred. We consider every possible subpath emanating from the OR and denote these subpaths by the abbreviations presented in Table 4.1. Generally, the individual letters represent the departments the patient has visited after their surgery (S) and the last letter represents the current nursing ward. As we consider the patients who have stayed at a ward of W and are discharged afterwards separately from the ones who are transferred to an IC of I after their stay in W, we define their subpath abbreviations as $SW1$ and $SW2$, respectively. $SW2$ is therefore the subpath which always preceeds subpath SWI.

Table 4.1: Subpaths.

Subpaths $p \in P$	Description
SI	Subpath in which a patient is transferred to an IC in I after surgery
SWI	Subpath in which a patient is transferred to an IC in I after spending time at some ward in W after surgery
$SW1$	Path in which the patient is transferred to a ward in W after surgery and will be discharged after the stay
$SW2$	Subpath in which the patient is transferred to a ward in W after surgery and will be transferred to the IC after the stay
SIW	Path in which the patient is transferred to a ward in W after spending time at the IC after surgery
$SWIW$	Path in which the patient is transferred to a ward in W after spending first time at a ward in W, followed by the IC after surgery
SD	Path in which a patient goes to D after surgery

Define the partition of P by sets P^I, P^W and P^D, where the superscript denotes the last nursing ward department of the subpath. In other words, $P^I = \{SI, SWI\}$ and $P^W = \{SW1, SW2, SIW, SWIW\}$ and $P^D = \{SD\}$. We now define the probabilities for the transferal to another nursing ward type. We do this for each subpath of Table 4.1. This results in Figure 4.1, which is an extension of Figure 2.2. Followed by the figure, an explanation of the parameters is given.

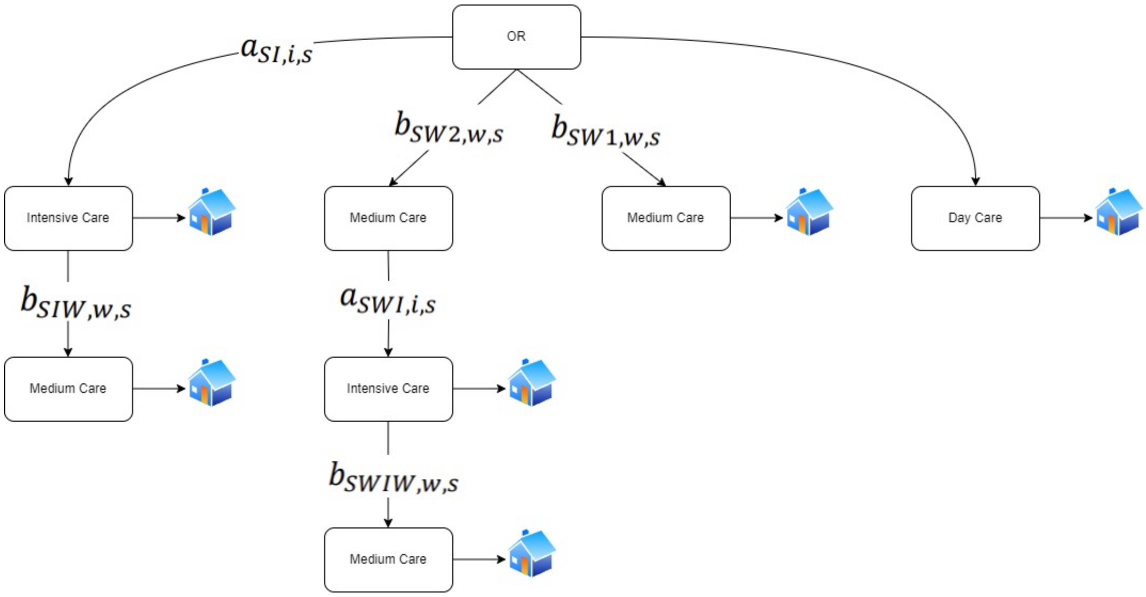


Figure 4.1: Considered patient pathways from OR to discharge with their probabilities.

Parameters a_{pis} represent the probability of a patient of specialty $s \in S$ being transferred to IC $i \in I$ as the last nursing ward of subpath $p \in P^I$.

- $a_{(SI)is}$: probability that a patient of specialty $s \in S$ is transferred to IC $i \in I$ after surgery.
- $a_{(SWI)is}$: probability that a patient of specialty $s \in S$ is transferred to IC $i \in I$ after spending time at some nursing ward of W after surgery, i.e., after having followed subpath $SW2$.

Similarly, parameters b_{pws} represent the probability of a patient of specialty $s \in S$ being transferred to ward $w \in W$ as the last nursing ward of subpath $p \in P^W$.

- $b_{(SW1)ws}$: probability that a patient of specialty $s \in S$ is transferred to ward $w \in W$ after surgery and will be discharged after the stay at ward $w \in W$.
- $b_{(SW2)ws}$: probability that a patient of specialty $s \in S$ is transferred to ward $w \in W$ after surgery and will be transferred to the IC after the stay at ward $w \in W$.
- $b_{(SIW)ws}$: probability that a patient of specialty $s \in S$ is transferred to ward $w \in W$ after spending time at the IC after surgery, i.e., after having followed subpath SI.
- $b_{(SWIW)ws}$: probability that a patient of specialty $s \in S$ is transferred to ward $w \in W$ after spending the first time at a ward in W followed by the IC after surgery, i.e., after having followed subpath SWI.

We can find the probability of a patient of specialty $s \in S$ being transferred to D after surgery, i.e. following the only path $p \in P^D$, by calculating

$$1 - \sum_{i \in I} a_{(SI)is} - \sum_{w \in W} b_{(SW1)ws} - \sum_{w \in W} b_{(SW2)ws}.$$

Define $S_w \subset S$ as the subset of specialties that can be transferred to ward $w \in W$. Similarly, define $S_i \subset S$ as the subset of specialties that are transferred to IC $i \in I$. Parameters a_{pis} and b_{pws} are equal to zero for $s \notin S_i$ and $s \notin S_w$, respectively.

For the duration of a patient occupying a bed in the nursing ward, we use the data of the hospital and perform log-normal distribution fittings to find the probabilities c_{psn} that a patient of specialty $s \in S$ stays n days in the last nursing ward of subpath $p \in P$. This choice of the log-normal distribution accounts for the positive valuedness and skewness usually observed in the length of stay data. In models seen in our literature review, the empirical distribution was used for the length of stay (Fügener et al., 2014;

Schneider et al., 2020). This has the advantage of potentially being more accurate. We, however, wanted to use a standard distribution for the length of stay for its simplicity. In their paper, Faddy et al. (2009) fitted the length of stay with the log-normal distribution, the gamma distribution and a phase-type distribution. They found that although the log-normal was inferior to the phase-type, it performed significantly better than the gamma distribution and the estimates obtained with the log-normal were relatively close to that of the phase-type.

As the log-normal is a continuous probability distribution while we wish c_{psn} to be discrete, we approximate c_{psn} by integrating the probability density function f over the domain $[n, n+1)$:

$$c_{psn} = \int_{n}^{n+1} f(x)\, dx. \tag{4.12}$$

Notice that no distinction between wards $w \in W$ are made as well as between ICs $i \in I$ for the parameters c_{psn}. We assume that the length of stay at a nursing ward only depends on the specialty $s \in S$ and the general ward department order.

Starting from Subsection 4.2.1, we only take into account the patients that go to wards in W and I, i.e. subpaths in P^I and P^W. As the patients in D are discharged on the same day as their surgery, we only consider those patients on the day of their surgery. We estimate the bed occupancy in D for each day $k \in K$ by summing all respective lengths of stay expectations of the patients in D on day $k \in K$ and dividing the total sum by the total opening duration of D on day $k \in K$.

4.2.1. Step 1: Specialty specific distribution for one planned surgery

Assume that the surgery takes place on day 1. From the parameters c_{psn}, we use conditional probability to calculate the probabilities $d_{ps(n+1)}$ that a patient, after they have stayed n days in the last ward, is discharged from the last ward of their subpath $p \in P$ on day $n+1$. Using the definition of conditional probability we obtain that $d_{ps(n+1)} = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. In our case, A equals the event that a patient stays exactly n days in their ward, i.e., $\mathbb{P}(A) = c_{psn}$, while B corresponds to the event that the patient has already stayed n days in the ward and hence, is surely discharged after exactly n days or more, i.e., $\mathbb{P}(B) = 1 - \sum_{k=1}^{n-1} c_{psk}$.

Let N_{ps} denote the maximum length of stay for a patient of specialty $s \in S$ in the latest ward in the subpath $p \in P$. These are taken as a certain high percentile of the fitted log-normal distribution as used for the probabilities c_{psn}. Then $\mathbb{P}(B) = 1 - \sum_{k=1}^{n-1} c_{psk} \approx \sum_{k=n}^{N_{ps}} c_{psk}$. Using the above, we find:

$$d_{ps(n+1)} = \frac{c_{psn}}{1 - \sum_{k=1}^{n-1} c_{psk}} \approx \frac{c_{psn}}{\sum_{k=n}^{N_{ps}} c_{psk}} \quad s \in S, n \in \{0, \dots, N_{ps}\}. \tag{4.13}$$

Now using $d_{ps(n+1)}$, we compute e_{psn} representing the probabilities that a patient of specialty $s \in S$ is in the last ward of the subpath $p \in P$ on day n. When the patient has been in one different ward before, we add parameter m as the number of days stayed in the former ward resulting in probabilities e_{psnm}. In the case of two former wards, we define probabilities $e_{psnm_1m_2}$ in a similar manner in which parameter m_1 denotes the number of days stayed in the first ward in path $p \in P$ whilst m_2 represents this for the second ward in path $p \in P$. Contrary to Schneider et al. (2020), we do not assume that a patient has to stay at least one day when they are transferred.

We compute the probabilities e_{psn} in a recursive manner. For the subpaths $p \in P$ which only contain one ward, the probability for $n = 1$ is equal to the probability to be transferred respectively to an IC or a ward after surgery. For $n \in \{2, \dots, N_{ps}+1\}$, e_{psn} is equal to the probability that the patient was there the day before times the probability that the patient is not transferred on day n, i.e., $n-1$ days after the surgery.

In case $p = SI$, this results in:

$$e_{(SI)sn} = \begin{cases} \sum_{i\in I} a_{(SI)is} & \text{for } n = 1 \\ \left(1 - d_{(SI)s(n-1)}\right) e_{(SI)s(n-1)} & \text{for } n \in \{2, \dots, N_{(SI)s} + 1\} \\ 0 & \text{otherwise.} \end{cases} \tag{4.14}$$

In case $p = SW1, SW2$:

$$e_{psn} = \begin{cases} \sum_{w\in W} b_{pws} & \text{for } n = 1 \\ \left(1 - d_{ps(n-1)}\right) e_{ps(n-1)} & \text{for } n \in \{2, \dots, N_{ps} + 1\} \\ 0 & \text{otherwise.} \end{cases} \tag{4.15}$$

When the subpath contains multiple wards, we again start by considering the probability that the patient is in the ward on the same day as their arrival at the ward, i.e., $n = m$ or $n = m_2$. Let $\tilde{p}$ denote the preceding subpath of p by excluding the last nursing department, and define $\hat{p}$ as the preceding subpath of $\tilde{p}$ in case $p = SWIW$. For instance, for $p = SWIW$, $\tilde{p} = SWI$ and $\hat{p} = SW2$. This probability is equal to the probability that the patient is not discharged or transferred on the first day in their current ward $(1 - d_{ps1})$, multiplied by the probabilities of staying m or m_2 days and being discharged after m or m_2 days in the preceding subpath $\tilde{p} \in P$. For $n > m, m_2$, the probability is computed in a similar fashion as before, by multiplying the probability that the patient was there the day before with the probability that the patient is not transferred on day n, i.e., $n - 1$ days after the surgery.

In the case that the patient visits two nursing wards:

$$e_{psnm} = \begin{cases} \left(1 - d_{ps1}\right) e_{\tilde{p}sm} d_{\tilde{p}sm} & \text{for } m \in \{1, \dots, N_{\tilde{p}s} + 1\}, n = m \\ \left(1 - d_{ps(n-m+1)}\right) e_{ps(n-1)m} & \text{for } m \in \{1, \dots, N_{\tilde{p}s} + 1\}, n \in \{m + 1, \dots, m + N_{ps}\} \\ 0 & \text{otherwise.} \end{cases} \tag{4.16}$$

In the case that the patient visits three nursing wards:

$$e_{psnm_1m_2} = \begin{cases} \left(1 - d_{ps1}\right) e_{\tilde{p}sm_2m_1} d_{\tilde{p}s(m_2-m_1+1)} & \text{for } m_1 \in \{1, \dots, N_{\hat{p}s} + 1\}, m_2 \in \{m_1, \dots, m_1 + N_{\tilde{p}s}\}, n = m_2 \\ \left(1 - d_{ps(n-m_2+1)}\right) e_{ps(n-1)m_1m_2} & \text{for } m_1 \in \{1, \dots, N_{\hat{p}s} + 1\}, m_2 \in \{m_1, \dots, m_1 + N_{\tilde{p}s}\}, \\ & n \in \{m_2 + 1, \dots, m_2 + N_{ps}\} \\ 0 & \text{otherwise.} \end{cases} \tag{4.17}$$

Let f_{isn} be the probability distributions that a patient of specialty $s \in S$ is in IC $i \in I$ on day n. Suppose that $\mathbf{f}_{isn}$ are the corresponding discrete stochastic variables and define N_s^I as the total maximal length of stay for the whole subpath of subpaths in P^I, i.e., $N_s^I = \max\{N_{(SI)s}, N_{(SW2)s} + N_{(SWI)s}\}$. The probability that a patient of specialty $s \in S$ is in IC $i \in I$ on day n, $\mathbb{P}(\mathbf{f}_{isn} = 1)$, is given by the sum of the probabilities that a patient of specialty $s \in S$ is in IC $i \in I$ on day n given their subpath $p \in P^I$ multiplied by the probability of subpath $p \in P^I$. The probability of the opposite event, $\mathbb{P}(\mathbf{f}_{isn} = 0)$, is calculated as $1 - \mathbb{P}(\mathbf{f}_{isn} = 1)$.

$$\begin{aligned} \mathbb{P}\left(\mathbf{f}_{isn} = 1\right) &= \sum_{p\in P^I} \mathbb{P}(p \wedge \text{last ward is } i) \cdot \mathbb{P}(\mathbf{f}_{isn} = 1 | p \wedge \text{last ward is } i) && \text{for } n \in \{1, \dots, N_s^I + 1\} \\ &= \frac{a_{(SI)is}}{\sum_{i\in I} a_{(SI)is}} e_{(SI)sn} + \frac{a_{(SWI)is}}{\sum_{i\in I} a_{(SWI)is}} \sum_{m=1}^{n-1} e_{(SWI)snm} && \text{for } n \in \{1, \dots, N_s^I + 1\} \\ &= \frac{a_{(SI)is}}{\sum_{i\in I} a_{(SI)is}} e_{(SI)sn} + a_{(SWI)is} \sum_{m=1}^{n-1} e_{(SWI)snm} && \text{for } n \in \{1, \dots, N_s^I + 1\}. \end{aligned} \tag{4.18}$$

Note that $\sum_{i\in I} a_{(SWI)is} = 1$ as subpath SWI is always followed by the patients that followed its preceding subpath $SW2$. Furthermore, remark that $\mathbb{P}(\mathbf{f}_{isn} = 1 | p \wedge \text{last ward is } i) = \mathbb{P}(\mathbf{f}_{isn} = 1 | p)$ as the length of stay in a ward is only dependent on the ward type.

Similarly, the probability that a patient of specialty $s \in S$ is in ward $w \in W$ on day n is given by the sum of the probabilities that a patient of specialty $s \in S$ is in ward $w \in W$ on day n given their subpath $p \in P^W$ multiplied by the probability of subpath $p \in P^W$. Let f_{wsn} be the probability distributions that a patient of specialty $s \in S$ is ward $w \in W$ on day n. Suppose that $\mathbf{f}_{wsn}$ are the corresponding discrete stochastic variables. The probability of the opposite event, $\mathbb{P}(\mathbf{f}_{wsn} = 0)$, is calculated as $1 - \mathbb{P}(\mathbf{f}_{wsn} = 1)$. Let N_s^W denote the total maximal length of stay for the whole subpath of subpaths in P^W, i.e., $N_s^W = \max\{N_{(SW1)s}, N_{(SW2)s}, N_{(SI)s} + N_{(SIW)s}, N_{(SW2)s} + N_{(SWI)s} + N_{(SWIW)s}\}$. Then, we can calculate $\mathbb{P}(\mathbf{f}_{wsn} = 1)$ as follows:

$$\begin{aligned}\mathbb{P}(\mathbf{f}_{wsn} = 1) = &\sum_{p \in P^W} \mathbb{P}(p \wedge \text{last ward is } w) \cdot \mathbb{P}(\mathbf{f}_{wsn} = 1 | p \wedge \text{last ward is } w) \quad \text{for } n \in \{1, \dots, N_s^W + 1\} \\ = &\frac{b_{(SW1)ws}}{\sum_{w \in W} b_{(SW1)ws}} e_{(SW1)sn} + \frac{b_{(SW2)ws}}{\sum_{w \in W} b_{(SW2)ws}} e_{(SW2)sn} + \frac{b_{(SIW)ws}}{\sum_{w \in W} b_{(SIW)ws}} \sum_{m=1}^{n} e_{(SIW)snm} \\ &+ \frac{b_{(SWIW)ws}}{\sum_{w \in W} b_{(SWIW)ws}} \sum_{m_2=1}^{n} \sum_{m_1=1}^{m_2} e_{(SWIW)snm_1m_2} \quad \text{for } n \in \{1, \dots, N_s^W + 1\}.\end{aligned} \tag{4.19}$$

4.2.2. Step 2: Specialty specific distribution for one surgical patient within the MSS

In this subsection, we take into account the fact that the length of stay of a patient can exceed the time period that one MSS covers. We assume that the schedule that is produced, is the only MSS used in order to look at the bed occupancy during a steady state.

Depending on the maximum length of stay and the day of arrival, it is possible for patients to occupy a bed in a later MSS time period of Q days than the MSS time period of their surgery. Hence, when determining the bed occupancy of a certain day, the patients staying at the ward from previous MSS time periods have to be taken into account as well. The number of Q-length periods that have to be taken into account for a specialty $s \in S$ is dependent on the maximal length of stay at the IC given by N_s^I or the maximal length of stay at the MC given by N_s^W. For a day q in the Q-length period ($q \in \{1, \dots, Q\}$) and assuming the patient's surgery was scheduled on the first day of the time period, the number of MSS time periods to take into account is:

$$\left\lfloor \frac{N_s^I - q}{Q} \right\rfloor + 1 \quad \text{for the IC,} \tag{4.20}$$

$$\left\lfloor \frac{N_s^W - q}{Q} \right\rfloor + 1 \quad \text{for the MC.} \tag{4.21}$$

Let F_{isq}^I denote the distribution for the number of patients of specialty $s \in S$ occupying a bed in IC $i \in I$ on the qth day of a Q-length period, assuming that the surgery of specialty $s \in S$ was scheduled on the first day of the time period. Then, F_{isq}^I is equal to the convolution of the probabilities that a patient is in the IC on the qth day, resulting from all $\left\lfloor \frac{N_s^I - q}{Q} \right\rfloor + 1$ past Q-length periods. The distribution F_{wsq}^W for the number of patients of specialty $s \in S$ occupying a bed in ward $w \in W$ on the qth day of a Q-length period, assuming that the surgery of specialty $s \in S$ was scheduled on the first day of the time period, can be constructed similarly. This results in the following equations:

$$F_{isq}^I = f_{isq}^I * f_{is(q+Q)}^I * \dots * f_{is\left(q + \left(\left\lfloor \frac{N_s^I - q}{Q} \right\rfloor\right) \cdot Q\right)}^I \qquad i \in I, s \in S, q \in \{1, \dots, Q\}, \tag{4.22}$$

$$F_{wsq}^W = f_{wsq}^W * f_{ws(q+Q)}^W * \dots * f_{ws\left(q + \left(\left\lfloor \frac{N_s^W - q}{Q} \right\rfloor\right) \cdot Q\right)}^W \qquad w \in W, s \in S, q \in \{1, \dots, Q\}. \tag{4.23}$$

4.2.3. Step 3: Distribution bed occupancy

Now, we calculate the bed occupancy in the nursing wards for the resulting MSS. For each day $k \in K$, the created schedule gives us integer values for the variables Z_{oks} which represent the number of surgeries of specialty $s \in S$ scheduled in OR $o \in O$ on day $k \in K$. The distribution of the bed occupancy in IC $i \in I$ or ward $w \in W$ on day q of the MSS when a surgery of specialty $s \in S$ is scheduled on day $k \in K$ in OR $o \in O$ is represented by G^I_{ioksq} and G^W_{woksq}, respectively. For this, the distributions F^I_{isq} and F^W_{wsq} are shifted to correspond with the correct length of stay. For $q \geq k$, the smallest number of days a patient who had surgery on day k could have stayed in the hospital is $q - k + 1$ as we assume the surgery takes place on day 1. For the days preceding k, i.e. $q < k$, the patient distribution can only be dependent on the patients from preceding MSS schedules. Hence, we shift by an additional Q days. This results in the following equations:

$$G^I_{ioksq} = \begin{cases} F^I_{is(q-k+1)} \mathbb{1}_{Z_{oks}>0}, & q \geq k \\ F^I_{is(q-k+1+Q)} \mathbb{1}_{Z_{oks}>0}, & \text{otherwise} \end{cases} \qquad i \in I, o \in O, k \in K, q \in \{1, \dots, Q\}, s \in S_i, \tag{4.24}$$

$$G^W_{woksq} = \begin{cases} F^W_{ws(q-k+1)} \mathbb{1}_{Z_{oks}>0}, & q \geq k \\ F^W_{ws(q-k+1+Q)} \mathbb{1}_{Z_{oks}>0}, & \text{otherwise} \end{cases} \qquad w \in W, o \in O, k \in K, q \in \{1, \dots, Q\}, s \in S_w. \tag{4.25}$$

Note that we have not taken into account the actual number of surgical patients yet. In order to take this into account, we have to apply convolution on the corresponding distribution G^I_{ioksq} or G^W_{woksq} Z_{oks} times. We introduce the following notations that we use in the remainder of the section. For some functions h, $h_1, \dots, h_m$ and positive integer m:

$$\begin{aligned} \underbrace{h * h * \cdots * h}_{m} &= h^{*m}, \quad h^{*0} = \delta_0, \\ h_1 * \cdots * h_m &= \mathop{\circledast}_{i=1}^{m} h_i, \end{aligned} \tag{4.26}$$

where δ_0 denotes the Dirac delta distribution.

So we can calculate the total patient distribution $\hat{G}^I_{ioksq}$ and $\hat{G}^W_{woksq}$ for IC $i \in I$ and ward $w \in W$, respectively, on day q of the MSS resulting from specialty $s \in S$ being assigned to OR $o \in O$ on day $k \in K$ as follows:

$$\hat{G}^I_{ioksq} = \left(G^I_{ioksq}\right)^{*Z_{oks}} \qquad i \in I, o \in O, k \in K, s \in S_i, q \in \{1, \dots, Q\}, \tag{4.27}$$

$$\hat{G}^W_{woksq} = \left(G^W_{woksq}\right)^{*Z_{oks}} \qquad w \in W, o \in O, k \in K, s \in S_w, q \in \{1, \dots, Q\}. \tag{4.28}$$

To estimate the distribution for the bed occupancy at IC $i \in I$ or ward $w \in W$ on day q resulting from all surgeries of OR $o \in O$ on day $k \in K$, we have to convolve over $\hat{G}^I_{ioksq}$ or $\hat{G}^W_{woksq}$, respectively, for all specialties $s \in S$. Define these distributions as H^I_{iokq} and H^W_{wokq}, respectively. This results in the following equations:

$$H^I_{iokq} = \mathop{\circledast}_{s \in S_i} \hat{G}^I_{ioksq} \qquad i \in I, o \in O, k \in K, q \in \{1, \dots, Q\}, \tag{4.29}$$

$$H^W_{wokq} = \mathop{\circledast}_{s \in S_w} \hat{G}^W_{woksq} \qquad w \in W, o \in O, k \in K, q \in \{1, \dots, Q\}. \tag{4.30}$$

In order to calculate the bed occupancy for the whole MSS, we still need to convolve over every OR-day $k \in K$ and over every OR $o \in O$. Suppose $\hat{H}^I_{iq}$ and $\hat{H}^W_{wq}$ represent the distribution of occupied beds on day q of the MSS for IC $i \in I$ and ward $w \in W$, respectively. $\hat{H}^I_{iq}$ and $\hat{H}^W_{wq}$ can thus be calculated in the following way:

$$\hat{H}^I_{iq} = \circledast_{o\in O, k\in K} H^I_{iokq} \qquad i \in I, q \in \{1, \dots, Q\}, \tag{4.31}$$

$$\hat{H}^W_{wq} = \circledast_{o\in O, k\in K} H^W_{wokq} \qquad w \in W, q \in \{1, \dots, Q\}. \tag{4.32}$$

Similarly as Schneider et al. (2020), we define $\hat{H}^I_{iq}[n]$ and $\hat{H}^W_{wq}[n]$ to be the probabilities of having n patients in IC $i \in I$ and ward $w \in W$, respectively. Let ψ represent the given MSS schedule. Now, define $\gamma_{iq}(\psi)$ as the number such that the number of occupied beds on day $q \in K$ in IC $i \in I$ for schedule ψ is at most $\gamma_{iq}(\psi)$ with probability at least $1-\epsilon$. This leads to the following formula for $\gamma_{iq}(\psi)$:

$$\gamma_{iq}(\psi) = \min\left\{ n \Big| \sum_{m=0}^{n} \hat{H}^I_{iq}[m] \geq 1-\epsilon \right\}. \tag{4.33}$$

Similarly, we define $\gamma_{wq}(\psi)$ as the number such that the number of beds on day $q \in K$ in ward $w \in W$ for schedule ψ is at most $\gamma_{wq}(\psi)$ with probability at least $1-\epsilon$. So,

$$\gamma_{wq}(\psi) = \min\left\{ n \Big| \sum_{m=0}^{n} \hat{H}^W_{iq}[m] \geq 1-\epsilon \right\}. \tag{4.34}$$

In our research, we are interested in calculating the variation in bed occupancy. For an IC $i \in I$ and ward $w \in W$, we calculate the variation γ_i and γ_w as the difference between the maximum and minimum number of beds needed during the MSS. On days $q \in \{1, \dots, Q\} \setminus K$, the number of beds for elective patients can only decrease as no elective surgeries are scheduled on these days. Hence, we neglect these days when calculating the variation:

$$\gamma_i(\psi) = \max_{q\in K} \gamma_{iq}(\psi) - \min_{q\in K} \gamma_{iq}(\psi), \tag{4.35}$$

$$\gamma_w(\psi) = \max_{q\in K} \gamma_{wq}(\psi) - \min_{q\in K} \gamma_{wq}(\psi). \tag{4.36}$$

As the minimum and maximum operators in the above equations are non-linear operators, we resolve this by a linearisation in the next chapter.

4.3. Objective function

For our model, we wish to obtain two main things. First of all, we would like to schedule each specialty with equal relative frequency. This is done by comparing the number of surgeries a specialty has scheduled to the average number of surgeries carried out by the specialty during a Q-length period. Besides this, we want to minimize the variation in bed occupancy during the Q-length period. However, this should not be at the expense of the OR utilization. Hence, our last goal is to maximize the OR utilization.

So, firstly, it is desirable that every specialty has relatively the same time in the OR with respect to the OR time they were expected to utilize. For this, we look at the number of surgeries the specialties on average carry out within a Q-length period. Let L_s denote the average number of surgeries for specialty $s \in S$. Then, we introduce auxiliary variables V_s to represent the percentage of less scheduled surgeries with respect to the expected number of surgeries for specialty $s \in S$:

$$V_s = \frac{L_s - \sum_{o\in O, k\in K} Z_{oks}}{L_s} = 1 - \frac{1}{L_s} \sum_{o\in O, k\in K} Z_{oks}, \quad \forall s \in S. \tag{4.37}$$

In order to give every specialty relatively the same time, we would like to have the largest and smallest V_s close to each other. For this, we add auxiliary variables $V_{\max}$ and $V_{\min}$, as well as the constraints:

$$V_{\max} \geq V_s, \quad \forall s \in S, \tag{4.38}$$
$$V_{\min} \leq V_s, \quad \forall s \in S. \tag{4.39}$$

Hence, the first sub-objective becomes:

$$\min V_{\max} - V_{\min}, \tag{4.40}$$

which is equivalent to

$$\max -(V_{\max} - V_{\min}). \tag{4.41}$$

Secondly, we wish to utilize the OR capacity as much as possible. In order to do this, we estimate each surgery duration by its expected value. Let the expected duration of a surgery from specialty $s \in S$ be denoted by $\mathbb{E}[DS_s]$. Then we get the following sub-objective:

$$\max \sum_{o \in O, k \in K, s \in S_o} \mathbb{E}[DS_s] \cdot Z_{oks}. \tag{4.42}$$

Lastly, we wish to minimize the variation in bed occupancy for each IC $i \in I$ and ward $w \in W$. Thus, we want to minimize each γ_i and γ_w for the given MSS schedule ψ. Equivalently to this is maximizing each $-\gamma_i$ and $-\gamma_w$ for the given MSS schedule ψ.

As each objective is of different importance and different order of magnitude, we include non-negative weights θ when combining the sub-objectives in the objective function. The resulting objective function is then given by:

$$\max -\theta_V \cdot (V_{\max} - V_{\min}) + \theta_Z \cdot \sum_{o \in O, q \in Q, s \in S_o} \mathbb{E}[DS_s] \cdot Z_{oqs} - \sum_{i \in I} \theta_i \cdot \gamma_i(\psi) - \sum_{w \in W} \theta_w \cdot \gamma_w(\psi). \tag{4.43}$$

5

Solution methods

This chapter introduces linearization techniques for both the overtime constraint (Constraint (4.9)) and the objective function (4.43) from the previous chapter. The initial two sections are dedicated to the linearization of the overtime constraint. The first section outlines a method for linearizing the overtime constraint under the assumption that surgery durations follow log-normal distributions, while the subsequent section presents a linearization procedure based on the assumption that surgery durations follow normal distributions. Subsequently, we provide the linearization approach for the objective function. To conclude the chapter, we describe the two resulting linear problems (under log-normality, as well as normality).

5.1. Linearization overtime constraint under a log-normal distribution

In this subsection, we describe an approach to linearize Constraint (4.9), where we assume that the surgery durations follow log-normal distributions. This constraint was introduced to take into account overtime in our model. The constraint is given as follows:

$$\mathbb{P}\left(\mathbf{g}_{ok} \geq \beta_{ok} - \kappa \cdot \left(\sum_{s \in S_o} Z_{oks} + W_{ok}\right)\right) \leq \alpha, \quad \forall o \in O, \forall k \in K.$$

in which

- $\mathbf{g}_{ok}$ is the stochastic variable corresponding to the probability distribution g_{ok} of the cumulative surgery duration in OR $o \in O$ on day $k \in K$,
- β_{ok} denotes the duration that OR $o \in O$ was planned to be open on day $k \in K$,
- κ is the time length reserved for cleaning between the surgeries and for the change in specialty in an OR,
- $\sum_{s \in S_o} Z_{oks}$ is the total number of surgeries scheduled in OR $o \in O$ on day $k \in K$,
- W_{ok} is the variable that indicates whether there are two specialties having surgeries in OR $o \in O$ on day $k \in K$ or not,
- α is the desired upper bound for the probability of overtime.

The probability distribution g_{ok} is dependent on the individual durations of the planned surgeries on day $k \in K$ in OR $o \in O$. In the review of Wang et al. (2021), it is found that the most recurring distribution to fit the surgery duration is the log-normal as previous research has shown that surgery durations appear to be best described by a log-normal distribution (Stepaniak et al., 2009; Kayış et al., 2015). Unfortunately, up to now, there has not been a closed form solution for the sum of independent log-normal random variables (RVs) (Mehta et al., 2007). In fields such as telecommunication and stock pricing, various approximations are used to approximate the sum of log-normal RVs by one log-normal RV, e.g., the Fenton-Wilkinson (FW) (Fenton, 1960) and Schwartz-Yeh (SY) approximations (Schwartz and Yeh, 1982). As these approximations are not linear methods, we use inspiration from the column generation method by generating all possible combinations for the number of surgeries for each pair of specialties that certainly satisfy Constraint (4.9). This is further explained in Subsection 5.1.5.

As these approximation methods are generally used in a telecommunication setting, it is not clear how accurate the methods are for our use. Therefore, we have chosen three approximation methods to investigate: the FW method, the SY method and the method described in Mehta et al. (2007) which we refer to as the 'Mehta approximation'. All three approximations attempt to estimate the sum of independent log-normal random variables by one log-normal random variable. If a method seems like an applicable candidate for our use, we compare the performance to simulations and choose the most fitting method for our approximation. First, the three mentioned methods are briefly explained.

5.1.1. Fenton-Wilkinson method

The Fenton-Wilkinson method approximates the sum of independent log-normal random variables by one log-normal random variable by matching the first and second moments about zero. Let $X_1, \dots, X_R$ denote R independent two-parameter log-normal RVs with parameters $(\mu_1, \sigma_1), \dots, (\mu_R, \sigma_R)$, respectively. In other words, $X_i = \exp(W_i)$ where W_i is normally distributed with mean μ_i and standard deviation σ_i for every $i \in \{1, \dots, R\}$. Assume that p_{X_i} denotes the probability distribution function for RV X_i for $i \in \{1, \dots, R\}$. Now, let X denote the log-normal RV that the FW method uses to approximate $\sum_{i=1}^R X_i$. The approximation method determines the parameters μ_X and σ_X of RV X as follows.

The Fenton-Wilkinson method finds μ_X and σ_X^2 by matching the first and second moments of X about the origin to those of $\sum_{i=1}^R X_i$. So,

$$\mathbb{E}[X] = \int_0^\infty x p_X(x) dx = \sum_{i=1}^R \int_0^\infty x p_{X_i}(x) dx, \tag{5.1}$$

$$\mathbb{E}[X^2] = \int_0^\infty x^2 p_X(x) dx = \sum_{i=1}^R \int_0^\infty x^2 p_{X_i}(x) dx. \tag{5.2}$$

Note that Equation (5.1) is equivalent to $\mathbb{E}[X] = \sum_{i=1}^R \mathbb{E}[X_i]$. Furthermore, as the variance is related to the second moment by $\mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, instead of Equation (5.2) one can also match the variance of X to the variance of $\sum_{i=1}^R X_i$. This results in

$$\mathrm{Var}(X) = \int_0^\infty (x - \mu_X)^2 p_X(x) dx = \sum_{i=1}^R \int_0^\infty (x - \mu_{X_i})^2 p_{X_i}(x) dx. \tag{5.3}$$

When the expectation and variance of X are known, its parameters μ_x and σ_X^2 can be calculated in the following way. We start with the expression of the expectation and variance for a log-normal random variable.

$$\mathbb{E}[X] = \exp\left(\mu_X + \frac{\sigma_X^2}{2}\right) \tag{5.4}$$

$$\mathrm{Var}(X) = (\exp(\sigma_X^2) - 1)\exp(2\mu_X + \sigma_X^2) = (\exp(\sigma_X^2) - 1)(\mathbb{E}[X])^2 \tag{5.5}$$

Now, it is easy to see that σ_X^2 is equal to $\log\left(\frac{\mathrm{Var}(X)}{\mathbb{E}[X]^2} + 1\right)$ and $\mu_X = \log(\mathbb{E}[X]) - \frac{\sigma_X^2}{2}$.

5.1.2. Schwartz-Yeh method

Similarly to the FW method, the Schwarts-Yeh method tries to find one suitable log-normal variable for the sum of log-normals by matching the first and second moments. Let again $X_1, \dots, X_R$ denote R independent two-parameter log-normal RVs with parameters $(\mu_1, \sigma_1), \dots, (\mu_R, \sigma_R)$, respectively. Let p_{X_i} denote the probability distribution function for RV X_i for $i \in \{1, \dots, R\}$ and let X denote the log-normal

RV that we use to approximate $\sum_{i=1}^{R} X_i$. However, in contrast to the FW method, the SY approximation does not use the moments of X and X_i but rather those of $\ln X$ and $\ln\left(\sum_{i=1}^{R} X_i\right)$.

This results in the following two equations:

$$\int_0^\infty \ln(x) p_X(x)dx = \int_0^\infty \ln(x) p_{\sum_i X_i}(x)dx, \tag{5.6}$$

$$\int_0^\infty (\ln(x) - \mu_X)^2 p_X(x)dx = \int_0^\infty \left(\ln(x) - \mu_{\sum_i X_i}\right)^2 p_{\sum_i X_i}(x)dx. \tag{5.7}$$

In the above equations, $p_{\sum_i X_i}$ denotes $\circledast_{i=1}^{R} p_{X_i}$. Instead of working with these convolutions, we find an approximation for $p_{\sum_i X_i}$ (although it is exact for $R = 2$). For $R > 2$, $p_{\sum_i X_i}$ is determined in a nested fashion. In the first iteration, the p_Y is determined where $Y = X_1 + X_2$. Next, p_{Y+X_3} is calculated in a similar way, continuing this until the convolution of the desired sum is approximated. We continue by giving the idea behind one iteration. For the exact details, we refer to Schwartz and Yeh (1982).

Assume that Y, X_1 and X_2 are log-normal random variables for which the parameters of Y are unknown and the parameters of X_i are μ_i and σ_i for $i \in \{1, 2\}$. Suppose that the associated normal RVs are W, W_1 and W_2, respectively. Then,

$$W = \ln(X_1 + X_2) \tag{5.8}$$
$$= \ln(\exp(W_1) + \exp(W_2)) \tag{5.9}$$
$$\mu_X = \mathbb{E}[W] \tag{5.10}$$
$$= \mathbb{E}[\ln(\exp(W_1) + \exp(W_2))] \tag{5.11}$$
$$= \mathbb{E}[\ln(\exp(W_1)(1 + \exp(W_2 - W_1)))] \tag{5.12}$$
$$= \mathbb{E}[W_1] + \mathbb{E}[\ln(1 + \exp(W_2 - W_1))] \tag{5.13}$$

As W_1 and W_2 are normally distributed, their difference $\tilde{W} = W_2 - W_1$ is again normally distributed with mean $\mu_{\tilde{W}} = \mu_2 - \mu_1$ and variance $\sigma_{\tilde{W}}^2 = \sigma_1^2 + \sigma_2^2$. Hence, the last term of Equation (5.13) can be rewritten as:

$$\mathbb{E}[\ln(1 + \exp(\tilde{W}))] = \int_{-\infty}^{\infty} [\ln(1 + \exp(x))] p_{\tilde{W}}(x)dx. \tag{5.14}$$

We can now expand the logarithm term with the power series expansion

$$\ln(1 + x) = \sum_{j=1}^{\infty} C_j x^j, \quad \text{where } C_j = \frac{(-1)^{j+1}}{j}. \tag{5.15}$$

This expansion's radius of convergence is $|x| < 1$ and is hence valid when $x < 0$ for $\exp(x)$. Hence, if we split the integral of Equation (5.14) as follows, we can use the power series expansion on the left integral.

$$\int_{-\infty}^{\infty} [\ln(1 + \exp(x))] p_{\tilde{W}}(x)dx = \int_{-\infty}^{0} [\ln(1 + \exp(x))] p_{\tilde{W}}(x)dx + \int_{0}^{\infty} [\ln(1 + \exp(x))] p_{\tilde{W}}(x)dx. \tag{5.16}$$

To use the expansion on the whole integral of Equation (5.14), we rewrite the integral on the positive domain as follows:

$$\int_0^\infty [\ln(1 + \exp(x))] p_{\tilde{W}}(x)dx = \int_0^\infty [\ln(\exp(x) \cdot \exp(-x) + \exp(x))] p_{\tilde{W}}(x)dx, \tag{5.17}$$
$$= \int_0^\infty [\ln(\exp(x)) + \ln(\exp(-x) + 1)] p_{\tilde{W}}(x)dx, \tag{5.18}$$
$$= \int_0^\infty [x + \ln(\exp(-x) + 1)] p_{\tilde{W}}(x)dx. \tag{5.19}$$

Using the expansion, $\mathbb{E}[\ln(1+\exp(\tilde{W}))]$ can be denoted as a function $G_a(\mu_{\tilde{W}}, \sigma_{\tilde{W}})$. So, substituting this into Equation (5.13) gives

$$\mu_X = \mu_1 + G_a(\mu_{\tilde{W}}, \sigma_{\tilde{W}}). \tag{5.20}$$

Something similar can be done for the second moment, where functions $G_b(\mu_{\tilde{W}}, \sigma_{\tilde{W}})$ and $G_c(\mu_{\tilde{W}}, \sigma_{\tilde{W}})$ are introduced. Then, σ_X^2 can be represented by

$$\sigma_X^2 = \sigma_1^2 - G_a^2(\mu_{\tilde{W}}, \sigma_{\tilde{W}}) - 2\left(-\frac{\sigma_1}{\sigma_{\tilde{W}}}\right)^2 G_c(\mu_{\tilde{W}}, \sigma_{\tilde{W}}) + G_b(\mu_{\tilde{W}}, \sigma_{\tilde{W}}). \tag{5.21}$$

In favor of the computational performance, the functions G_i with $i \in \{a, b, c\}$ are approximated by low-order polynomials using a least-square fit, which are used instead in Equations (5.20) and (5.21).

In the end, we refrain from using the SY method for our simulation tests. According to Mehta et al. (2007), the SY method approximates the tail of the cumulative distribution less accurately than the other methods. As we choose α in Constraint (4.9) to be small, we favor a good approximation method for the tail part of the cumulative distribution. Not only is the SY method less accurate, it is also quite complicated which is also unfavorable.

5.1.3. Mehta approximation

Mehta et al. (2007) have another opinion on how to find a suitable log-normal RV to approximate the sum of log-normal RVs. Mehta et al. (2007) explain in their article that both FW and SY methods can be seen as weighted integrals of probability density functions (PDF). Due to the difference in the weights of the integrals (in Equations (5.1), (5.2), (5.6), (5.7)), the FW method would be better in approximating the tail of the PDF while the SY method gives a better estimation of the head of the PDF. Mehta et al. (2007) have thought of a method that can both approximate the head and tail of the PDF by choosing different variables. While the FW and SY methods compare the first and second moments of X and $\sum_i X_i$, Mehta et al. (2007) compare the moment generating functions (MGF) for two different values. These values are dependent on what area of the PDF we are interested in.

Another thing they do differently is that they do not assume an exponential relation with base e between the log-normal RV and its normal RV. Instead they assume that $X_i = 10^{\frac{1}{10}W_i}$, where W_i is a normal random variable with mean μ_i and standard deviation σ_i.

The MGF of an RV X is given as

$$\Psi_X(t) = \int_0^\infty \exp(-tx) p_X(x) dx, \quad (\mathrm{Re}(t) \geq 0). \tag{5.22}$$

The MGF of a sum of independent RVs has the useful property that it is equivalent to the product of the MGFs of the original RVs:

$$\Psi_{\left(\sum_{i=1}^R X_i\right)}(t) = \prod_{i=1}^R \Psi_{X_i}(t), \quad (\mathrm{Re}(t) \geq 0). \tag{5.23}$$

Again, these can be seen as weighted functions of the PDF. By using different values for t, more weight can be put on either the head or tail portion depending on the area of interest. For our research, we are interested in large percentiles of the cumulative distribution function of the sum and hence, the tail portion is more important for us. The appropriate two values for t that Mehta et al. (2007) propose are $(t_1, t_2) = (0.001, 0.005)$. Note that we solve two equalities to be able to estimate two unknown values.

We now give a more elaborate description of the Mehta approximation. Note that when $X = \exp(\tilde{W})$ with $\tilde{W} \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$, $p_X(x) = \frac{1}{x\tilde{\sigma}\sqrt{2\pi}} \exp\left(-\frac{(\ln(x)-\tilde{\mu})^2}{2\tilde{\sigma}^2}\right)$.

If $X = 10^{\frac{1}{10}W} = \exp\left(\frac{\ln(10)}{10}W\right)$ with $W \sim \mathcal{N}(\mu, \sigma^2)$. Let $\xi = \frac{10}{\ln(10)}$. Then

$$p_X(x) = \frac{1}{x\frac{\sigma}{\xi}\sqrt{2\pi}} \exp\left(-\frac{\left(\ln(x) - \frac{\mu}{\xi}\right)^2}{2\left(\frac{\sigma}{\xi}\right)^2}\right) \tag{5.24}$$

$$= \frac{\xi}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{\xi^2\left(\ln(x) - \frac{\mu}{\xi}\right)^2}{2\sigma^2}\right) \tag{5.25}$$

$$= \frac{\xi}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi\ln(x) - \mu)^2}{2\sigma^2}\right). \tag{5.26}$$

This we can substitute in the MGF:

$$\Psi_X(t) = \int_0^\infty \exp(-tx) p_X(x) dx \tag{5.27}$$

$$= \int_0^\infty \exp(-tx) \frac{\xi}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi\ln(x) - \mu)^2}{2\sigma^2}\right) dx. \tag{5.28}$$

Now substitute $x = \exp\left(\frac{\sqrt{2}\sigma z + \mu}{\xi}\right)$, i.e., interchange x with $z = \frac{\xi\ln(x) - \mu}{\sqrt{2}\sigma}$ in the integral above, resulting in:

$$\Psi_X(t) = \int_0^\infty \exp(-tx) \frac{\xi}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi\ln(x) - \mu)^2}{2\sigma^2}\right) dx \tag{5.29}$$

$$= \int_{-\infty}^\infty \frac{1}{\sqrt{\pi}} \exp\left[-t\exp\left(\frac{\sqrt{2}\sigma z + \mu}{\xi}\right)\right] \exp(-z^2) dz \tag{5.30}$$

We now use the Gauss-Hermite series expansion. It approximates the value of integrals of certain forms in the following way:

$$\int_{-\infty}^\infty \exp(-x^2) f(x) dx \approx \sum_{n=1}^N w_n f(a_n), \tag{5.31}$$

where N is the integration order, a_n are the roots of the Hermite polynomial $H_N(x)$ and w_n are the weights given by $\frac{2^{N-1}N!\sqrt{\pi}}{n^2[H_{N-1}(a_i)]^2}$. Mehta et al. (2007) found that with an integration order of $N = 12$, the MGF was approximated accurately. The values for a_n and w_n for $N = 12$ can be found in Abramowitz and Stegun (1972).

Thus,

$$\Psi_X(t) \approx \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \exp\left[-t\exp\left(\frac{\sqrt{2}\sigma a_n + \mu}{\xi}\right)\right]. \tag{5.32}$$

So, if X is the log-normal RV corresponding to the approximated sum and the X_i's are the original log-normal RVs, we want to find μ_X and σ_X such that:

$$\begin{aligned}\sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \exp\left[-t_m\exp\left(\frac{\sqrt{2}\sigma_X a_n + \mu_X}{\xi}\right)\right] = \\ \prod_{i=1}^R \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \exp\left[-t_m\exp\left(\frac{\sqrt{2}\sigma_{X_i} a_n + \mu_{X_i}}{\xi}\right)\right], \quad \text{for } m = 1, 2.\end{aligned} \tag{5.33}$$

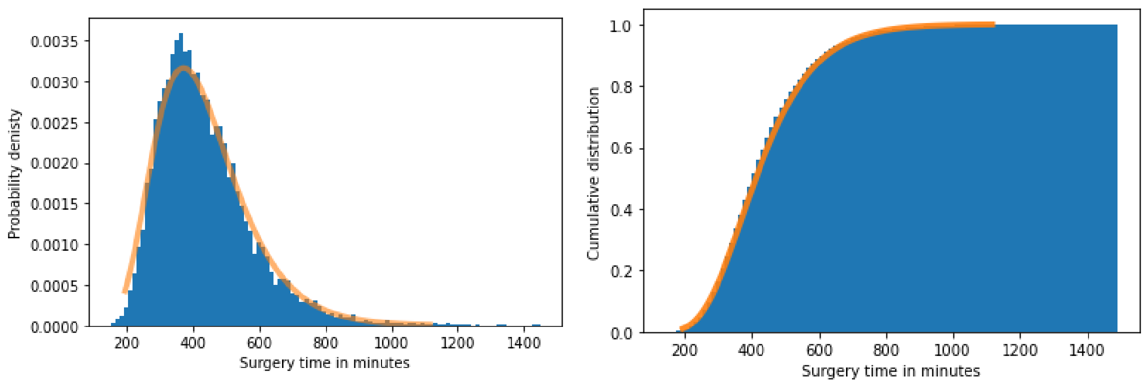
These equations are solved by using a nonlinear equations solver like numpy.optimize.fsolve to find the root of the difference between the left-handside and right-handside of Equation (5.33).

5.1.4. Method selection

As previously stated, both the FW method and the Mehta approximation are techniques commonly used in the field of telecommunication. Hence, it is uncertain whether they are applicable for our specific data. In this subsection, we inspect if the two methods give reasonable results when applied to our data. Should a method demonstrate satisfactory results, it will be selected for subsequent use in our log-normal-based model. In Subsection 6.4.2, we employ goodness of fit tests on the chosen method to further assess its suitability.

As explained in Section 4.1, only two specialties can be 'combined' in one OR. Furthermore, for each specialty $s \in S$, we know its maximum number of patients M_s. Now, suppose that X_s denotes the log-normal RV with parameters μ_s and σ_s for the surgery duration of specialty $s \in S$. For two distinct specialties s_1 and s_2, we may use the approximation methods to find suitable log-normal distributions for $n_1 X_{s_1} + n_2 X_{s_2}$ for $n_1 \in \{1, \ldots, M_{s_1}\}$ and $n_2 \in \{1, \ldots, M_{s_2}\}$. Starting from this point forward, the various potential values of n_1 and n_2 for all the compositions of s_1 with s_2 are referred to as 'combinations.'

Upon examination, it seems that the FW method provides a reasonable approximation for all combinations. To exemplify this, let us consider the combination with two PDP surgeries and three FLEX surgeries. We generated a sample consisting of 10000 total surgery durations based on the fitted log-normal distributions of individual surgery durations. In Figure 5.1, the sample's duration is visualized through a histogram, along with the approximated distribution derived from the FW method. As the curves from the approximated distribution follow the histograms quite well and as this is the case for all combinations, we conclude that the FW method works well.



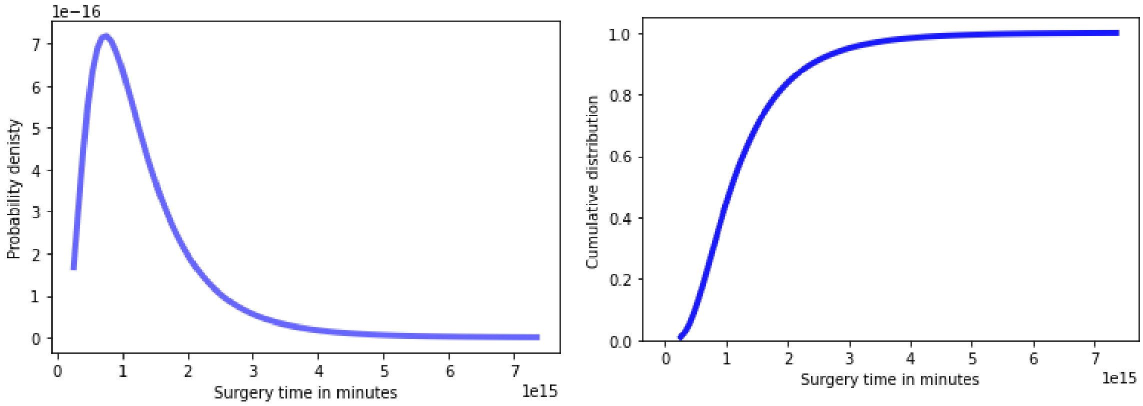
(a) Probability density function.

(b) Cumulative distribution function.

Figure 5.1: Histogram depicting the total surgery duration across 10000 samples along with the corresponding FW method approximation, in a scenario where two surgeries of PDP and three surgeries of FLEX are scheduled.

We can perform a similar analysis for the Mehta approximation. However, the outcomes for the Mehta approximation were unsatisfactory. The ranges of both the domain and codomain of the Mehta approximations did not align with the scale we anticipated. In Figure 5.2, we have exclusively illustrated the Mehta approximation for the scenario involving two PDP surgeries and three FLEX surgeries. Combining this with a generated sample leads to plots with unreadable representations due to significant scale differences.

As mentioned before, it is hard to predict beforehand how the approximation methods would perform on our data. It could be that the Mehta approximation performs better with other t-values. In order to check if the proposed t_1 and t_2 are ill-suited for our data, we solve Equation (5.33) with the approximation of σ_X and μ_X resulting from the FW method in the hope to find suitable values for t_1 and t_2. Using all possible combinations for the surgery appointment on one OR day, we find that 90% of the values for t_1 lie between $(0.114, 2.01)$ and for t_2 between $(0.323, 5.755)$.

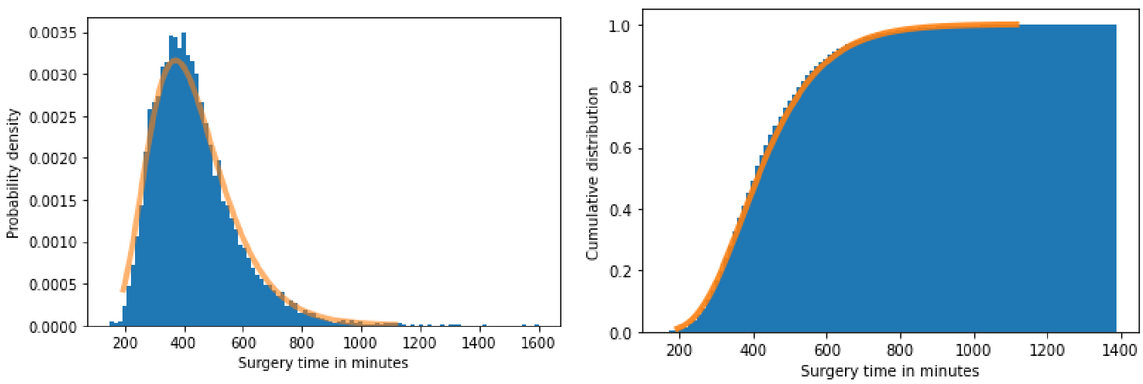


(a) Probability density function.

(b) Cumulative distribution function.

Figure 5.2: Plots resulting from the Mehta approximation with $(t_1, t_2) = (0.001, 0.005)$, in a scenario where two surgeries of PDP and three surgeries of FLEX are scheduled.

If we set $(t_1, t_2) = (1.014, 1.825)$, this gives a much better approximation. In Figure 5.3, a visualization is again given for the combination where two surgeries of PDP and three surgeries of FLEX are scheduled.



(a) Probability density function.

(b) Cumulative distribution function.

Figure 5.3: Histogram depicting the total surgery duration across 10000 samples along with the corresponding Mehta method approximation with $(t_1, t_2) = (1.014, 1.825)$, in a scenario where two surgeries of PDP and three surgeries of FLEX are scheduled.

Unfortunately, this method does not prove effective for all combinations. In instances where the combinations involve the PDO specialty, the Mehta approximation could not be computed due to an exponential overflow. Experimenting with various t-values yielded similar unfavorable outcomes for at least one specialty. Since we could not identify any (t_1, t_2)-pair that functioned for all our combinations and given the more consistent performance of the more simple FW method, we resolved to exclusively employ the FW method moving forward.

In summary, our selection for the sum of log-normals approximation method is the FW method. With this approximation technique in place, we can now determine the combinations of surgeries that satisfy the overtime constraint. While generating the viable combinations, it became apparent that no suitable arrangement could be identified for the specialty PORS2 for our selected α value in the context of the overtime constraint. Consequently, we had to make an exemption for this specific specialty, allowing it to have a single surgery in an operating room on one day. It is noteworthy that this exemption was extended to all models explored in Chapter 7 for the same reason.

5.1.5. Column generation based approach

Using the Fenton-Wilkinson approximation, we can find a good log-normal fit for the sum of surgery durations. As the Fenton-Wilkinson method is not linear, we generate all the possible combinations of surgeries from at most two specialties in one OR on one day such that Constraint (4.9) is satisfied. To model this, we introduce sets J_{ok} as the set of feasible combinations on day $k \in K$ for OR $o \in O$. We introduce binary variables U_{jok} which are one when combination $j \in J_{ok}$ is scheduled in OR $o \in O$ on day $k \in K$ and zero otherwise. We also introduce parameters v_{js} that give the number of surgeries of specialty $s \in S$ included in combination $j \in J_{ok}$. Our new approach is to choose one combination U_{jok} with $j \in J_{ok}$ for each day $k \in K$ and for each OR $o \in O$. Now, decision variables Z_{oks} can be replaced by $\sum_{j \in J_{ok}} U_{jok} \cdot v_{js}$.

The sets J_{ok} can be partitioned into two types of subsets: the subsets J_{oks} for each $s \in S$, which contain the combinations that only schedule surgeries of specialty $s \in S$ in OR $o \in O$ on day $k \in K$, and the subsets $J_{oks_1s_2}$ for $s_1, s_2 \in S$ and $\hat{s}_1 < \hat{s}_2$, which contain the combinations of surgeries of specialties s_1 and $s_2 \in S$ in OR $o \in O$ for day $k \in K$.

When two specialties s_1 and $s_2 \in S$ are scheduled in OR $o \in O$ on day $k \in K$, indicated by binary variables Y_{oks}, only surgeries of these specialties can be scheduled in that OR on that day. So, only one combination of $J_{oks_1s_2}$ can be picked, resulting in the following constraints:

$$\sum_{j \in J_{oks_1s_2}} U_{jok} \leq \frac{1}{2}\left(Y_{oks_1} + Y_{oks_2}\right), \quad \forall o \in O, \forall k \in K, \forall s_1, s_2 \in S \text{ and } \hat{s}_1 < \hat{s}_2. \tag{5.34}$$

It is also possible that only one specialty $s \in S$ is scheduled in OR $o \in O$ on day $k \in K$. Then, a combination from subset J_{oks} has to be picked. This results in the following constraints:

$$\sum_{j \in J_{oks}} U_{jok} \leq Y_{oks}, \quad \forall o \in O, \forall k \in K, \forall s \in S. \tag{5.35}$$

Furthermore, only one combination of J can be picked for each OR day. Hence, Constraints (5.36) are added.

$$\sum_{j \in J_{ok}} U_{jok} \leq 1, \quad \forall o \in O, \forall k \in K. \tag{5.36}$$

Using this column generation based approach in our model, we can neglect some constraints from our original model, as these constraints are inherently implied by the combinations. Hence, we can disregard Constraints (4.7) through (4.11).

5.2. Linearization overtime constraint under a normal distribution

In this subsection, we describe an approach to linearize Constraint (4.9) where we assume that the surgery durations follow normal distributions. This approach is similar to the one of Schneider et al. (2020). Again we look at Equation (4.9), which is given by:

$$\mathbb{P}\left(\mathbf{g}_{ok} \geq \beta_{ok} - \kappa \cdot \left(\sum_{s \in S_o} Z_{oks} + W_{ok}\right)\right) \leq \alpha, \quad \forall o \in O, k \in K.$$

in which

- $\mathbf{g}_{ok}$ is the stochastic variable corresponding to the probability distribution g_{ok} of the cumulative surgery duration in OR $o \in O$ on day $k \in K$,
- β_{ok} denotes the duration that OR $o \in O$ was planned to be open on day $k \in K$,
- κ is the time length reserved for cleaning between the surgeries and for the change in specialty in an OR,
- $\sum_{s \in S_o} Z_{oks}$ is the total number of surgeries scheduled in OR $o \in O$ on day $k \in K$,
- W_{ok} is the variable that indicates whether there are two specialties having surgeries in OR $o \in O$ on day $k \in K$ or not,

- α is the probability of overtime.

Suppose that the individual surgery duration of specialty $s \in S$ follows the normal distribution with mean μ_s and variance σ_s^2. Then, $\mathbf{g}_{ok}$ is normally distributed with mean $\mu_{ok} = \sum_{s\in S} \mu_s \cdot Z_{oks}$ and variance $\sigma_{ok}^2 = \sum_{s\in S} \sigma_s^2 \cdot Z_{oks}$.

We can rewrite the overtime constraint to the standard normal distribution form:

$$\mathbb{P}\left(\frac{\mathbf{g}_{ok} - \mu_{ok}}{\sigma_{ok}} \leq \frac{\beta_{ok} - \kappa \cdot \left(\sum_{s\in S_o} Z_{oks} + W_{ok}\right) - \mu_{ok}}{\sigma_{ok}}\right) = \Phi\left(\frac{\beta_{ok} - \kappa \cdot \left(\sum_{s\in S_o} Z_{oks} + W_{ok}\right) - \mu_{ok}}{\sigma_{ok}}\right), \quad \forall o \in O, \forall k \in K, \tag{5.37}$$

where Φ denotes the cumulative distribution function of the standard normal distribution.

So, we have rewritten the constraint to $\Phi\left(\frac{\beta_{ok} - \kappa \cdot \left(\sum_{s\in S_o} Z_{oks} + W_{ok}\right) - \mu_{ok}}{\sigma_{ok}}\right) \geq 1 - \alpha, \forall o \in O, \forall k \in K.$

Now, using Φ^{-1} and the sum versions of μ_s and σ_s,

$$\mu_{ok} + \Phi^{-1}(1-\alpha) \cdot \sigma_{ok} \leq \beta_{ok} - \kappa \cdot \left(\sum_{s\in S_o} Z_{oks} + W_{ok}\right), \quad \forall o \in O, \forall k \in K, \tag{5.38}$$

$$\left(\sum_{s\in S} \mu_s \cdot Z_{oks}\right) + \Phi^{-1}(1-\alpha) \cdot \sqrt{\sum_{s\in S} \sigma_s^2 \cdot Z_{oks}} \leq \beta_{ok} - \kappa \cdot \left(\sum_{s\in S_o} Z_{oks} + W_{ok}\right), \quad \forall o \in O, \forall k \in K. \tag{5.39}$$

The only part in Equation (5.39) that is not yet linear, is the square root. To linearly approximate the square root function $f(x) = \sqrt{x}$, we approximate the square root by piece-wise linear functions. We now briefly describe the piece-wise linear functions. For the exact details, we refer to Bosch (2011).

First, determine the interval $[x_{\min}, x_{\max}]$ for which we want to estimate $f(x)$. In our case, we try to find suitable lower and upper bounds for $\sum_{s\in S} \sigma_s^2 \cdot Z_{oks}$. Hence, we choose $x_{\min} = 0$ and $x_{\max} = \max_{s\in S}\{\sigma_s^2\} \cdot \max_{s\in S}\{M_s\}$. The interval $[x_{\min}, x_{\max}]$ is split by $N+1$ breakpoints. The breakpoints are denoted by $x_0, x_1, \dots, x_N$ where $x_0 = x_{\min}$ and $x_N = x_{\max}$. The sub-intervals between each pair of successive breakpoints are the intervals for the piece-wise linear function. Each linear approximation function between interval $[x_{n-1}, x_n]$ can be written as $h_n(x) = a_n + b_n \cdot x$ for $n \in \{1, \dots, N\}$.

In each interval $[x_{n-1}, x_n]$, there exists a point t_n such that $h_n(x)$ is exactly the tangent line in $x = t_n$. Hence, $\sqrt{t_n} = a_n + b_n \cdot t_n$ and $b_n = (\sqrt{t_n})' = \frac{1}{2}\sqrt{\frac{1}{t_n}}$. Consequently, $a_n = \frac{1}{2}\sqrt{t_n}$ and we get the following expression for h_n:

$$h_n(x) = \frac{1}{2}\sqrt{t_n} + \frac{1}{2}\sqrt{\frac{1}{t_n}} \cdot x. \tag{5.40}$$

For each breakpoint x_n with $n \in \{1, \dots, N\}$, let y_n be its corresponding function value (i.e., $y_n = h_n(x_n)$) and let $y_0 = h_1(x_0)$. After identifying the breakpoints, we can employ the λ-formulation, a widely used technique for modeling piece-wise linear functions, as described in Bisschop (2006). This tells us that the function value of any point between two breakpoints is equal to the weighted sum of the function values at these two breakpoints. Let λ_n denote the nonnegative weights corresponding to breakpoint x_n. The expression for the piecewise linear approximation of the square root function is as follows:

$$\sum_{n=0}^{N} \lambda_n y_n = h_n(x), \tag{5.41}$$

$$\sum_{n=0}^{N} \lambda_n x_n = x, \tag{5.42}$$

$$\sum_{n=0}^{N} \lambda_n = 1. \tag{5.43}$$

Additionally to the characteristic described in Equation (5.43), at most two consecutive λ_n can obtain positive values. For this, we introduce binary variables δ_n and the following constraints:

$$\sum_{n=0}^{N-1} \delta_n = 1, \tag{5.44}$$

$$\lambda_0 \leq \delta_0, \tag{5.45}$$

$$\lambda_N \leq \delta_{N-1}, \tag{5.46}$$

$$\lambda_n \leq \delta_{n-1} + \delta_n, \forall n \in \{1, \dots, N-1\}. \tag{5.47}$$

Going back now to our overtime constraint, Equation (5.39) is given for each OR $o \in O$ and for each day $k \in K$. In order to linearize each constraint, we introduce decision variables λ_{okn} and δ_{okn} instead of λ_n and δ_n in the described square root linearization. Then, our overtime constraint and its additional constraint are given by:

$$\left(\sum_{s\in S} \mu_s \cdot Z_{oks}\right) + \Phi^{-1}(1-\alpha) \cdot \sum_{n=0}^{N} \lambda_{okn} y_n \leq \beta_{ok} - \kappa \cdot \left(\sum_{s\in S_o} Z_{oks} + W_{ok}\right), \quad \forall o \in O, \forall k \in K, \tag{5.48}$$

$$\sum_{n=0}^{N} \lambda_{okn} x_n = \sum_{s\in S} \sigma_s^2 \cdot Z_{oks}, \quad \forall o \in O, \forall k \in K, \tag{5.49}$$

$$\sum_{n=0}^{N} \lambda_{okn} = 1, \quad \forall o \in O, \forall k \in K. \tag{5.50}$$

$$\sum_{n=0}^{N-1} \delta_{okn} = 1, \quad \forall o \in O, \forall k \in K, \tag{5.51}$$

$$\lambda_{ok0} \leq \delta_{ok0}, \quad \forall o \in O, \forall k \in K, \tag{5.52}$$

$$\lambda_{okN} \leq \delta_{ok(N-1)}, \quad \forall o \in O, \forall k \in K, \tag{5.53}$$

$$\lambda_{okn} \leq \delta_{ok(n-1)} + \delta_{okn}, \quad \forall o \in O, \forall k \in K, \forall n \in \{1, \dots, N-1\}. \tag{5.54}$$

However, as the piece-wise linear approximation is always an overestimation (due to the concavity of the square root), we have that

$$y_0 = h_1(x_0) = \frac{1}{2}\sqrt{t_n} > 0. \tag{5.55}$$

This causes a problem when OR $o \in O$ is closed. Then, both β_{ok} and Z_{oks} are equal to zero, while $\lambda_{ok0} = 1$. This results in the following form for Equation 5.48:

$$\Phi^{-1}(1-\alpha) \cdot y_0 \leq 0. \tag{5.56}$$

For $\alpha < \frac{1}{2}$, $\Phi^{-1}(1-\alpha) > 0$ and the constraint is not fulfilled. To address this issue, we introduce an extra breakpoint between the existing breakpoints x_0 and x_1. Subsequently, we assign the value of $h_1(x_0)$ as the y-value of the new breakpoint and set $y_0 = 0$.

5.3. Linearization objective function

To linearize the objective function, we use the same method as Schneider et al. (2020). As mentioned in Section 4.2, the minimum and maximum operators are non-linear operators. Rather than using γ_i and γ_w, we use the expected value of the number of beds at ward $w \in W$ and IC $i \in I$ on day $q \in K$ which we denote by $\bar{\gamma}_{wq}$ and $\bar{\gamma}_{iq}$, respectively. In this section, we recall distributions from Subsections 4.2.2 and 4.2.3 and define corresponding stochastic variables for these distributions. The stochastic variables have the same variable name, but they are presented in bold.

$\bar{\gamma}_{iq}$ is computed by taking the expected value of the random variables $\hat{\mathbf{H}}^I_{iq}$ associated with the distributions for the bed occupancy $\hat{H}^I_{iq}$ for IC $i \in I$ on day $q \in K$. Let $\mathbf{H}^I_{iokq}$ represent the stochastic variables associated with the distributions H^I_{iokq} for bed occupancy in IC $i \in I$ on day $q \in K$. These variables result from the surgeries performed in OR $o \in O$ on day $k \in K$.

Similarly, we define $\hat{\mathbf{G}}^I_{ioksq}$ as the stochastic variables associated with the distribution $\hat{G}^I_{ioksq}$ for the bed occupancy in IC $i \in I$ on day $q \in K$, arising from the surgeries of specialty $s \in S$ scheduled on day $k \in K$ in OR $o \in O$.

Additionally, let the stochastic variables $\mathbf{G}^I_{ioksq}$ be associated with the distributions G^I_{ioksq}. These variables represent the bed occupancy in IC $i \in I$ on day $q \in K$, resulting from one surgery of specialty $s \in S$ scheduled on day $k \in K$ in OR $o \in O$.

Lastly, we define $\mathbf{F}^I_{isq}$ as the stochastic variables associated with the distributions F^I_{isq} for the number of recovering patients of specialty $s \in S$ occupying a bed in IC $i \in I$ on the qth day. This is when the surgery occurred on the first day of the Q-length period for the MSS.

Using the fact that the expected value of a convolution of density functions is equal to the sum of the expected value of the individual density functions, we can rewrite $\bar{\gamma}_{iq}$ as follows:

$$\begin{aligned}
\bar{\gamma}_{iq} &= \mathbb{E}\left[\hat{\mathbf{H}}^I_{iq}\right] && (5.57)\\
&= \mathbb{E}\left[\sum_{o\in O}\sum_{k\in K}\mathbf{H}^I_{iokq}\right] && (5.58)\\
&= \sum_{o\in O}\sum_{k\in K}\mathbb{E}\left[\mathbf{H}^I_{iokq}\right] && (5.59)\\
&= \sum_{o\in O}\sum_{k\in K}\sum_{s\in S_I}\mathbb{E}\left[\hat{\mathbf{G}}^I_{ioksq}\right] && (5.60)\\
&= \sum_{o\in O}\sum_{k\in K}\sum_{s\in S_I}\mathbb{E}\left[\mathbf{G}^I_{ioksq}\right]\cdot Z_{oks} && (5.61)\\
&= \sum_{o\in O}\sum_{k\in K, q\geq k}\sum_{s\in S_I}\mathbb{E}\left[\mathbf{F}^I_{is(q-k+1)}\right]\cdot Z_{oks} + \sum_{o\in O}\sum_{k\in K, q<k}\sum_{s\in S_I}\mathbb{E}\left[\mathbf{F}^I_{is(q-k+1+Q)}\right]\cdot Z_{oks} && (5.62)\\
& && (5.63)
\end{aligned}$$

Similarly, let $\mathbf{F}^W_{wsq}$ be the stochastic variables associated with the distributions F^W_{wsq} for the number of recovering patients of specialty $s \in S$ occupying a bed in ward $w \in W$ on the qth day when the surgery occurred on the first day of the Q-length period for the MSS. Then, the following equation can be found for $\bar{\gamma}_{wq}$:

$$\bar{\gamma}_{wq} = \sum_{o\in O}\sum_{k\in K, q\geq k}\sum_{s\in S_W}\mathbb{E}\left[\mathbf{F}^W_{ws(q-k+1)}\right]\cdot Z_{oks} + \sum_{o\in O}\sum_{k\in K, q<k}\sum_{s\in S_W}\mathbb{E}\left[\mathbf{F}^W_{ws(q-k+1+Q)}\right]\cdot Z_{oks}. \quad (5.64)$$

As we are interested in the variation in bed occupancy, we want to calculate the difference between the maximum and minimum number of beds for each ward and IC. Hence, we introduce variables $\hat{\gamma}_i$

for the variation in IC $i \in I$ and $\hat{\gamma}_w$ for the variation in ward $w \in W$ as:

$$\hat{\gamma}_i = \bar{\gamma}_i^{\max} - \bar{\gamma}_i^{\min}, \qquad \forall i \in I, \tag{5.65}$$

$$\hat{\gamma}_w = \bar{\gamma}_w^{\max} - \bar{\gamma}_w^{\min}, \qquad \forall w \in W, \tag{5.66}$$

and add the following constraints for $\bar{\gamma}_{iq}^{\max}$, $\bar{\gamma}_{wq}^{\max}$, $\bar{\gamma}_{iq}^{\min}$ and $\bar{\gamma}_{wq}^{\min}$:

$$\bar{\gamma}_i^{\max} \geq \bar{\gamma}_{iq}, \qquad \forall i \in I, \forall q \in K, \tag{5.67}$$

$$\bar{\gamma}_w^{\max} \geq \bar{\gamma}_{wq}, \qquad \forall w \in W, \forall q \in K, \tag{5.68}$$

$$\bar{\gamma}_i^{\min} \leq \bar{\gamma}_{iq}, \qquad \forall i \in I, \forall q \in K, \tag{5.69}$$

$$\bar{\gamma}_w^{\min} \leq \bar{\gamma}_{wq}, \qquad \forall w \in W, \forall q \in K. \tag{5.70}$$

The resulting objective function is given by

$$\max -\theta_V \cdot (V_{\max} - V_{\min}) + \theta_Z \cdot \sum_{o \in O, q \in Q, s \in S_o} \mathbb{E}[DS_s] \cdot Z_{oqs} - \sum_{i \in I} \theta_i \cdot \hat{\gamma}_i - \sum_{w \in W} \theta_w \cdot \hat{\gamma}_w. \tag{5.71}$$

In the following two sections, we present the resulting integer linear problem models for overview.

5.4. Resulting linear problem under log-normality

In this section, we present the resulting linear problem we obtain using the linearization of Sections 5.1 and 5.3.

$$
\begin{aligned}
\text{maximize} \quad & -\theta_V \cdot (V_{\max} - V_{\min}) + \theta_Z \cdot \sum_{\substack{o\in O,\, q\in Q,\\ s\in S_o}} \exp\left(\mu_s + \frac{\sigma_s^2}{2}\right) \cdot Z_{oqs} - \sum_{i\in I} \theta_i \cdot \hat{\gamma}_i - \sum_{w\in W} \theta_w \cdot \hat{\gamma}_w \\
\text{subject to} \quad & Z_{oks} \le M_s \cdot Y_{oks}, && \forall o\in O, \forall k\in K, \forall s\in S_o \\
& \sum_{s\in S_o} Y_{oks} \le 2\cdot X_{ok}, && \forall o\in O, \forall k\in K \\
& Y_{oks_2} + Y_{oks} \le 1, && \forall o\in O, \forall k\in K, \forall s\in S_o, \forall s_2 \notin I_s \\
& \sum_{o\in O\setminus \bar{O}_k} X_{ok} \le \chi_k - \left|\bar{O}_k\right|, && \forall k\in K \\
& \sum_{o\in O: s\in S_o} Y_{oks} \le C_s, && \forall s\in S, \forall k\in K \\
& Y_{oks} = Y_{o(k+T)s,} && \forall o\in O, \forall s\in S_o, k\in\{1,\dots,T\} \\
& Y_{oks} = Y_{o(k+2T)s,} && \forall o\in O, \forall s\in S_o, k\in\{1,\dots,T\} \\
& \vdots \\
& Y_{oks} = Y_{o\left(k+\left\lfloor \frac{Q}{T}\right\rfloor\cdot T\right)s,} && \forall o\in O, \forall s\in S_o, k\in\{1,\dots,\ (Q \bmod\ T)\} \\
& Z_{oks} = \sum_{j\in J} U_{jok}\cdot v_{joks}, && \forall o\in O, \forall k\in K, \forall s\in S \\
& \sum_{j\in J_{s_1s_2}} U_{jok} \le \frac{1}{2}\left(Y_{oks_1} + Y_{oks_2}\right), && \forall o\in O, \forall k\in K, \forall s_1, s_2\in S \text{ and } \hat{s}_1 < \hat{s}_2 \\
& \sum_{j\in J_s} U_{jok} \le Y_{oks}, && \forall o\in O, \forall k\in K, \forall s\in S \\
& \sum_{j\in J} U_{jok} \le 1, && \forall o\in O, \forall k\in K \\
& \bar{\gamma}_{iq} = \sum_{o\in O}\sum_{k\in K, q\ge k}\sum_{s\in S_I} \mathbb{E}\left[\mathbf{F}^I_{is(q-k+1)}\right]\cdot Z_{oks} \\
& \qquad + \sum_{o\in O}\sum_{k\in K, q<k}\sum_{s\in S_I} \mathbb{E}\left[\mathbf{F}^I_{is(q-k+1+Q)}\right]\cdot Z_{oks}, && \forall i\in I, \forall q\in K \\
& \bar{\gamma}_{wq} = \sum_{o\in O}\sum_{k\in K, q\ge k}\sum_{s\in S_W} \mathbb{E}\left[\mathbf{F}^W_{ws(q-k+1)}\right]\cdot Z_{oks} \\
& \qquad + \sum_{o\in O}\sum_{k\in K, q<k}\sum_{s\in S_W} \mathbb{E}\left[\mathbf{F}^W_{ws(q-k+1+Q)}\right]\cdot Z_{oks}, && \forall w\in W, \forall q\in K \\
& \hat{\gamma}_i = \bar{\gamma}_i^{\max} - \bar{\gamma}_i^{\min}, && \forall i\in I \\
& \hat{\gamma}_w = \bar{\gamma}_w^{\max} - \bar{\gamma}_w^{\min}, && \forall w\in W \\
& \bar{\gamma}_i^{\max} \ge \bar{\gamma}_{iq}, && \forall i\in I, \forall q\in K \\
& \bar{\gamma}_w^{\max} \ge \bar{\gamma}_{wq}, && \forall w\in W, \forall q\in K \\
& \bar{\gamma}_i^{\min} \le \bar{\gamma}_{iq}, && \forall i\in I, \forall q\in K \\
& \bar{\gamma}_w^{\min} \le \bar{\gamma}_{wq}, && \forall w\in W, \forall q\in K
\end{aligned}
$$

5.5. Resulting linear problem under normality

In this section, we present the resulting linear problem we obtain using the linearization of Sections 5.2 and 5.3.

$$\text{maximize} \quad -\theta_V \cdot (V_{\text{max}} - V_{\text{min}}) + \theta_Z \cdot \sum_{\substack{o\in O,\, q\in Q,\\ s\in S_o}} \mu_s \cdot Z_{oqs} - \sum_{i\in I} \theta_i \cdot \hat{\gamma}_i - \sum_{w\in W} \theta_w \cdot \hat{\gamma}_w$$

$$
\begin{aligned}
\text{subject to} \quad & Z_{oks} \le M_s \cdot Y_{oks}, && \forall o \in O, \forall k \in K, \forall s \in S_o \\
& \sum_{s\in S_o} Y_{oks} \le 2 \cdot X_{ok}, && \forall o \in O, \forall k \in K \\
& Y_{oks_2} + Y_{oks} \le 1, && \forall o \in O, \forall k \in K, \forall s \in S_o, \forall s_2 \notin I_s \\
& \sum_{o\in O\setminus \bar{O}_k} X_{ok} \le \chi_k - \left|\bar{O}_k\right|, && \forall k \in K \\
& \sum_{o\in O: s\in S_o} Y_{oks} \le C_s, && \forall s \in S, \forall k \in K \\
& Y_{oks} = Y_{o(k+T)s,} && \forall o \in O, \forall s \in S_o, k \in \{1, \dots, T\} \\
& Y_{oks} = Y_{o(k+2T)s,} && \forall o \in O, \forall s \in S_o, k \in \{1, \dots, T\} \\
& \vdots && \\
& Y_{oks} = Y_{o\left(k+\left\lfloor \frac{Q}{T}\right\rfloor \cdot T\right)s,} && \forall o \in O, \forall s \in S_o, k \in \{1, \dots,\ (Q \bmod T)\} \\
& R_{oks} \ge \frac{Z_{oks}}{M_s + 1}, && \forall o \in O, \forall k \in K, \forall s \in S_o \\
& W_{ok} \ge \sum_{s\in S_o} R_{oks} - 1, && \forall o \in O, \forall k \in K \\
& \sum_{s_2\in S_o,\, s_2\ne s} R_{oks_2} \ge R_{oks}, && \forall o \in O, \forall k \in K, \forall s \in S_o^C \\
& Z_{oks} \ge m_s \cdot R_{oks}, && \forall o \in O, \forall k \in K, \forall s \in S_o \\
& \left(\sum_{s\in S} \mu_s \cdot Z_{oks}\right) + \Phi^{-1}(1-\alpha) \cdot \sum_{n=0}^{N} \lambda_{okn} y_n \le \beta_{ok} - \kappa \cdot \left(\sum_{s\in S_o} Z_{oks} + W_{ok}\right), && \forall o \in O, \forall k \in K \\
& \sum_{n=0}^{N} \lambda_{okn} x_n = \sum_{s\in S} \sigma_s^2 \cdot Z_{oks}, && \forall o \in O, \forall k \in K \\
& \sum_{n=0}^{N} \lambda_{okn} = 1, && \forall o \in O, \forall k \in K \\
& \sum_{n=0}^{N-1} \delta_{okn} = 1, && \forall o \in O, \forall k \in K \\
& \lambda_{ok0} \le \delta_{ok0}, && \forall o \in O, \forall k \in K \\
& \lambda_{okN} \le \delta_{ok(N-1)}, && \forall o \in O, \forall k \in K \\
& \lambda_{okn} \le \delta_{ok(n-1)} + \delta_{okn}, && \forall o \in O, \forall k \in K, \forall n \in \{1, \dots, N-1\} \\
& \bar{\gamma}_{iq} = \sum_{o\in O} \sum_{k\in K, q\ge k} \sum_{s\in S_I} \mathbb{E}\left[\mathbf{F}^I_{is(q-k+1)}\right] \cdot Z_{oks} && \\
& \qquad + \sum_{o\in O} \sum_{k\in K, q<k} \sum_{s\in S_I} \mathbb{E}\left[\mathbf{F}^I_{is(q-k+1+Q)}\right] \cdot Z_{oks}, && \forall i \in I, \forall q \in K \\
& \bar{\gamma}_{wq} = \sum_{o\in O} \sum_{k\in K, q\ge k} \sum_{s\in S_W} \mathbb{E}\left[\mathbf{F}^W_{ws(q-k+1)}\right] \cdot Z_{oks} && \\
& \qquad + \sum_{o\in O} \sum_{k\in K, q<k} \sum_{s\in S_W} \mathbb{E}\left[\mathbf{F}^W_{ws(q-k+1+Q)}\right] \cdot Z_{oks}, && \forall w \in W, \forall q \in K \\
& \hat{\gamma}_i = \bar{\gamma}_i^{\text{max}} - \bar{\gamma}_i^{\text{min}}, && \forall i \in I \\
& \hat{\gamma}_w = \bar{\gamma}_w^{\text{max}} - \bar{\gamma}_w^{\text{min}}, && \forall w \in W \\
& \bar{\gamma}_i^{\text{max}} \ge \bar{\gamma}_{iq}, && \forall i \in I, \forall q \in K \\
& \bar{\gamma}_w^{\text{max}} \ge \bar{\gamma}_{wq}, && \forall w \in W, \forall q \in K \\
& \bar{\gamma}_i^{\text{min}} \le \bar{\gamma}_{iq}, && \forall i \in I, \forall q \in K \\
& \bar{\gamma}_w^{\text{min}} \le \bar{\gamma}_{wq}, && \forall w \in W, \forall q \in K
\end{aligned}
$$

6 Data analysis

This chapter provides a description of the two datasets provided by the Sophia Children's Hospital, as well as the preprocessing of these. The first section is dedicated to the description, while the second and third segments delve into the preprocessing of the surgery properties dataset and the bed assignment dataset, respectively. As mentioned in Subsection 5.1.4, we performed goodness of fit tests to assess the adequacy of our log-normal distribution approximation in contrast to the normal distribution fit in the last section.

6.1. Data description

For our research, we use data provided by the Sophia Children's Hospital. Interconnected by the patient's admission number, the patient information encompasses two datasets. The first dataset is a comprehensive list of surgeries conducted during a specific time period, providing detailed information on each procedure. This includes properties such as the treating specialty, start and end times of the surgeries. The second dataset, on the other hand, comprises a list of bed assignments that occurred within the same time frame.

Below, examples can be found of the two datasets.

Table 6.1: Example surgery properties dataset.

Admission number	Specialty	Start time of surgery	End time of surgery
1	PDS	2018-01-04 12:12	2018-01-04 13:20
2	PDO	2018-01-05 9:18	2018-01-05 11:03

Table 6.2: Example bed assignment dataset.

Admission number	Ward	Start date	End date
1	MC2	2018-01-04	2018-01-04
1	IC1	2018-01-04	2018-01-06
1	MC2	2018-01-06	2018-01-10
2	MC1	2018-01-05	2018-01-06

In total, the dataset provided by Sophia Children's Hospital encompasses information about 20364 surgical procedures, of which 13775 were elective surgeries. These surgeries were conducted between the years 2019 and 2021. Furthermore, the data concerning patients' bed assignments includes a record of 36902 bed-switches for surgical patients.

From the surgery properties dataset, we collected the following characteristics:

- admission number,
- specialty code,
- whether the surgery was elective or not,
- surgeon,
- type of surgery (for the division in subspecialties as explained in Section 2.1),
- start time of the surgery,
- start time of the induction,
- start time of the surgeon,
- end time of the surgeon,
- end time of the surgery,
- arrival time at the recovery.

From the bed assignment dataset, we collected the characteristics

- admission number,
- type of hospitalization (corresponding to the type of ward),
- ward of the bed assignment,
- start date of the bed assignment,
- end date of the bed assignment.

6.2. Preprocessing surgery durations

During a surgery, there are some significant moments for which the time can be measured. Given in chronological order, these moments are:

- start time of the surgery,
- start time of the induction,
- start time of the surgeon,
- end time of surgeon,
- end time of surgery,
- arrival time at the recovery.

For the calculation of a surgery duration, we look at the difference between the start and end time of a surgery. However, it can happen that one of the two times is missing or that the end time is before the start time. To get the largest reliable data set of our model, we make some adjustments when there are inconsistencies. The inconsistencies that we adjust for, are described in this section. We want to note that in our data, not all time moments are registered. The adjustments are only made if the moments that are needed for the adjustments are available. The adjustments detailed in this section were made in collaboration with experts from the Sophia Children's Hospital.

As mentioned before, the start time of a surgery can be missing. However, if the start time of the induction is known, the start time of the surgery is set to five minutes prior the start time of the induction. If the start time of the induction is not known, we cannot give a good estimation for the surgery start and do not take into account this data input.

If the end time of a surgery is missing, we proceed in the following way. If the arrival time at the recovery is known, we set the end time of the surgery five minutes prior to this arrival time. If this is not the case, we again do not take into account this data input.

It can occur that the end time of a surgery appears to have taken place before the start time of the surgery. We first look at the case where also the dates of the start and end time of the surgery differ. If the start time of the induction is available and it occurs before the end time of the surgery, then we change the date of the start time of the surgery to the date of the induction start. If the start time of the induction is missing, we try to do the same thing but instead use the start time of the surgeon.

While ending the surgery before starting the surgery indicates a clear error in at least one of the two times, it can also happen that the date of one of the two times is incorrect but not indicated by the order. These mistakes can for example happen when the surgery started before midnight and ended after midnight. This can lead to unrealistic long surgery durations. We fix this inconsistency with the following three steps if the dates for the start and end of the surgery differ but their order is correct:

- In case the start time of the induction differs more than two hours from the start time of the surgery, we proceed in one of the following two ways. Firstly, if the date of the start time of the induction differs from the date of the start time of the surgery and changing the date of the start time of the surgery into the date of the induction start results in a positive surgery duration, we match the date of the start of the surgery to the date of the start of the induction. However, if this would have resulted in a negative surgery duration, we look at whether changing the date of the surgery start to one day before the date of the induction results in a positive surgery duration.

- In case the start time of the surgeon differs more than four hours from the start time of the surgery, we proceed in one of the following two ways. Firstly, if the date of the start time of the surgeon differs from the date of the start time of the surgery and changing the date of the start time of the surgery into the date of the surgeon's start results in a positive surgery duration, we match the date of the start of the surgery to the date of the start of the surgeon. However, if this would have resulted in a negative surgery duration, we look at whether changing the date of the surgery start to one day before the date of the start of the surgeon results in a positive surgery duration.

- In case the end time of the surgeon and the end time of the surgery differ more than three hours, we proceed in one of the following two ways: Firstly, if the date of the end time of the surgeon differs from the date of the end time of the surgery and changing the date of the end time of the surgery into the date of the surgeon's end results in a positive surgery duration, we match the date of the end of the surgery to the date of the end of the surgeon. However, if this would have resulted in a negative surgery duration, we look at whether changing the date of the surgery ending to one day after the date of the end of the surgeon results in a positive surgery duration.

Prior, we have looked at the cases when the dates of the start and end time of a surgery differ. Now, we tackle the inconsistency when the dates are the same but the order is incorrect. For this, we have two possible fixes.

- When the start time of the surgery is before the arrival time at the recovery, we change the end time of the surgery to five minutes prior the arrival time at the recovery if this results in a positive surgery duration.

- When the end time of the surgery is after the start time of the induction, we change the start time of the surgery to five minutes prior the start time of the induction if this results in a positive surgery duration.

Next, we examine small order differences. In practice, the end time of the surgery should fall after the end time of the surgeon and before the patient's arrival at the recovery. If this condition is not met, it might suggest that the recorded surgery duration is inaccurate due to an error while noting the surgery's end time. To address this, we first want to eliminate the possibility that the error in surgery duration could be linked to the start time of the surgery. As a first step, we assess if the difference between the start time of the surgery and the start time of the induction is reasonable. At Sophia, they consider a difference of down to negative five minutes between these times as acceptable (i.e., start time of induction - start time of surgery > -5 minutes), accounting for potential mistakes of up to five minutes. Now, there are two ways the placement of the end time of the surgery can be wrong compared to the end time of the surgeon and the patient's arrival at the recovery, where the order of the latter two moments are correct:

- The end time of the surgery could be after both moments. Again, we take into account potential mistakes of up to five minutes. If (arrival time at the recovery - end time of surgery < -5 minutes) and (end time of surgery - end time of surgeon > -5 minutes), we change the end time of the surgery to five minutes before the arrival time at the recovery if this results in a positive surgery duration.
- The end time of the surgery could be before both moments. If (arrival time at the recovery - end time of surgery > -5 minutes) and (end time of surgery - end time of surgeon < -5 minutes), we change the end time of the surgery to five minutes before the arrival time at the recovery if this results in a positive surgery duration.

Following the completion of all adjustments, each entry in the dataset should represent a realistic surgery. If an entry does not meet this criterion, it indicates that we are unable to retrieve representative data from the available information. Specifically, surgeries that have a duration of over twenty hours are considered unrealistic and are subsequently removed from the dataset. Additionally, entries lacking an admission number or specialty code are also deleted. This is necessary to ensure that surgeries can be appropriately linked to patient information on the ward or specialty. To ensure a focused analysis, we only consider patients whose complete treatment paths can be found in the dataset.

6.3. Preprocessing bed assignment

Following the surgery, patients undergo recovery in a designated ward. In our dataset, each patient is assigned a bed in this ward. The bed assignment should occur either on the day of the surgery or prior to it. Additionally, the last day of the bed assignments should be on or after the day when the surgery concluded. Unfortunately, these conditions are not consistently met within our dataset.

For some patients, the date of the earliest surgery is after their last bed assignment. When this happens, we shift the bed data assignments so that the first bed assignment has the same date as the first surgery.

Furthermore, within our dataset, certain patients were observed to have undergone their final surgery after the last recorded end date of their bed assignments. By once again adjusting the bed assignment data to align the initial bed assignment date with the date of the first surgery, we resolved most of these inconsistencies.

6.4. Fitting surgery durations

In this section, we first test if the individual surgery durations from the data are fitted best by a log-normal or normal distribution. In the second subsection, we test if the simulated total duration of the combinations in J are fitted well by the proposed log-normal distribution of the FW method or the sum of the estimated normal distribution.

We use the following goodness of fit tests for this:

- the Kolmogorov-Smirnov test (KS test),
- the Cramer-Von Mises test (CvM test),
- the Anderson-Darling test (AD test).

The above tests work as follows. Let the hypothesized cumulative distribution be denoted by F and the empirical distribution function by F_n.

Thus,

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(\infty,x]}(X_i), \tag{6.1}$$

where X_i, $i \in \{1, \ldots, n\}$, are n independent and identically distributed ordered observations.

The KS test looks for the largest difference between F and F_n. Its test statistic is given by:

$$T = \sup_x |F_n(x) - F(x)|. \tag{6.2}$$

The CvM test looks at the square difference between the two cumulative distributions:

$$T = n\int_{-\infty}^{\infty} \left(F_n(x) - F(x)\right)^2 dF(x). \tag{6.3}$$

The AD test is similar to the CvM test but has an additional weighting function $(F(x)(1 - F(x)))^{-1}$. By adding this factor, the CvM test takes into account the fact that the tails are estimated more precisely as a sample estimate near 0 or 1 has a lower variance than one near 0.5. One can intuitively think about this as follows. The value $F(0)$ is the smallest value and hence, $F_n(0) - F(0)$ can only be non-negative. Similarly, $F(1)$ attains the maximum value and hence, $F_n(1) - F(1)$ can only be non-positive, while there is no such restriction for the values of $F_n(x) - F(x)$ when x approaches 0.5.

The AD test is given by:

$$T = n\int_{-\infty}^{\infty} \frac{\left(F_n(x) - F(x)\right)^2}{F(x)\,(1 - F(x))} dF(x). \tag{6.4}$$

These tests evaluate the following hypotheses:
H_0 : the cumulative distribution of X_i, $i \in \{1, \dots, n\}$ is distribution F
H_1 : the cumulative distribution of X_i, $i \in \{1, \dots, n\}$ is not distribution F.

For any of the test statistics, we reject the null hypothesis when the value of the test is large. Usually, the null hypothesis is rejected by comparing the p-values to some chosen significance level. Suppose we obtained the value t for the test statistic T. Then, the p-value is equal to $2\min\{\mathbb{P}(T \geq t|H_0), \mathbb{P}(T \leq t|H_0)\}$.

6.4.1. One surgery

In this subsection, we check if the surgery durations from our data are more likely to be fitted better by the log-normal distribution than by the normal distribution, as described in the literature. This is done with the test statistics provided above. Table 6.3 shows the obtained test statistics of the individual surgery duration of a specific specialty for log-normality and normality, while Table 6.4 provides the corresponding p-values.

Table 6.3: Goodness of fit statistics for the distribution of one surgery.

Specialty	# observations	KS test		CvM test		AD test	
		log-normal	normal	log-normal	normal	log-normal	normal
FLEX	901	0.02035	0.13748	0.03955	6.60245	0.29112	38.41107
FPS	1181	0.03167	0.10955	0.27706	4.66944	1.90910	28.37696
GYN	843	0.06566	0.11335	0.84787	3.36769	5.34450	18.99519
HPS	151	0.06727	0.15253	0.11616	0.88965	0.62134	4.94314
NSP	134	0.08335	0.10992	0.15598	0.38620	0.86520	2.37873
PDC	494	0.03428	0.09060	0.06463	1.38486	0.50663	8.42488
PDD	71	0.12286	0.21259	0.20597	0.94122	1.28068	5.18531
PDE	76	0.05123	0.104251	0.03581	0.19212	0.36810	1.56269
PDO	2008	0.04780	0.19982	1.54986	31.12626	11.48949	173.99654
PG	510	0.07702	0.15393	0.47930	3.77578	2.54998	Inf
PMFS	419	0.04209	0.09962	0.07283	1.34141	0.43425	8.46222
PNS	134	0.08335	0.10992	0.15598	0.38620	0.86520	2.37873
PDP	227	0.06999	0.13589	0.22844	1.37098	1.34670	8.08425
PDR	1806	0.10465	0.13155	4.72898	10.59022	28.65940	Inf
PDS	2025	0.08187	0.19434	3.39050	26.68019	18.18100	143.86595
PO	241	0.09390	0.050647	0.55703	0.09993	3.48845	0.82681
PORS	1205	0.03065	0.11436	0.14824	5.95037	1.44353	Inf
PORS2	228	0.14083	0.08472	1.49652	0.50482	8.76845	3.16163
PU	1569	0.03049	0.11260	0.26498	6.70828	1.83199	41.80685

From Table 6.4, we see that for 15 out of 19 specialties, the p-values are greater for the log-normal distributions than for the normal distributions for all three goodness of fit tests. This would imply that it is more likely for those specialties to be log-normal distributed. For specialties PO and PORS2, the normal distribution seems more likely. For the remaining specialties PDR and PDS, both fits were very unlikely and there was not one more favorable.

Table 6.4: Goodness of fit p-values for distribution of one surgery.

Specialty	KS test		CvM test		AD test	
	log-normal	normal	log-normal	normal	log-normal	normal
FLEX	0.84963	0.00000	0.93557	0.00000	0.94466	0.00000
FPS	0.18693	0.00000	0.15709	0.00000	0.10312	0.00000
GYN	0.00139	0.00000	0.00561	0.00000	0.00197	0.00000
HPS	0.50147	0.00178	0.51180	0.00440	0.62778	0.00306
NSP	0.30960	0.07846	0.37219	0.07846	0.43609	0.05748
PDC	0.60715	6e-04	0.78509	0.00031	0.73997	7e-05
PDD	0.23417	0.00327	0.25649	0.00325	0.23864	0.00236
PDE	0.98845	0.38067	0.95443	0.28361	0.87933	0.16226
PDO	0.00021	0.00000	0.00013	0.00000	0.00000	0.00000
PG	0.00472	0.00000	0.04493	0.00000	0.04666	0.00000
PMFS	0.44788	0.00049	0.73457	0.00039	0.81413	7e-05
PNS	0.30960	0.07846	0.37219	0.07846	0.43609	0.05748
PDP	0.21609	0.00046	0.21857	0.00033	0.21762	1e-04
PDR	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PDS	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PO	0.02854	0.56681	0.02853	0.58561	0.01560	0.46197
PORS	0.20765	0.00000	0.39502	0.00000	0.19055	0.00000
PORS2	0.00024	0.07577	0.00017	0.03863	5e-05	0.02271
PU	0.10819	0.00000	0.17026	0.00000	0.11385	0.00000

6.4.2. Multiple surgeries

In this subsection, we do the same as in the previous subsection but for multiple surgeries. In the prior chapter, we have seen two methods:

- approximating the sum of log-normals by one log-normal with the FW method,
- fitting the individual surgery durations by normal distributions and use the normal distribution corresponding to the sum.

For each combination of $j \in \bigcup_{o\in O,k\in K} J_{ok}$, we create a sample of size 100000 for the total surgery durations of the combination. Subsequently, we calculate the p-values for the approximation methods of both techniques. To enhance clarity, rather than presenting an extensive table of values, we present the values using boxplots in Figure 6.1. Outliers have been omitted from the boxplots for improved readability. The p-values of the log-normal approximation appear to be generally greater than the p-values for the normal approximation, indicating that the first method is more likely to give a better approximation.

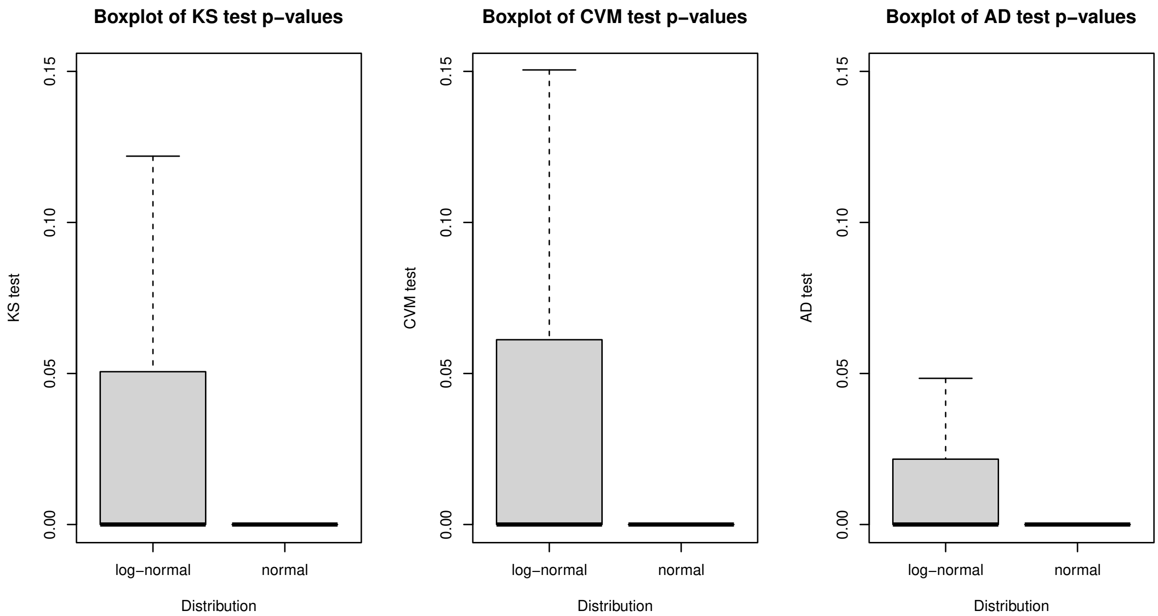


Figure 6.1: Boxplots of the p-values for KS test, CvM test and AD test for all possible combinations.

7 Results

In this chapter, we present our results when using the data from the previous chapter in the models described in Sections 5.4 and 5.5. We additionally define a model similar to the one from Section 5.4, but assume that the individual surgery durations are normally distributed. In the first section, we investigate how the models perform under different objective weights. After this investigation, we decided to investigate the log-normal model from Section 5.4 further and compare its performance with the help of a simulation and historical data. We conclude this chapter with results from an adjusted model that better reflects the wishes of the Sophia Children's Hospital.

All the models were implemented in Python with Gurobi Optimizer as solver. We additionally make use of the Delft High Performance Computing Centre to run our models.

7.1. Performance under different objective weights

In this section, we look at the results we get when choosing different objective weights for the objective function (4.43). Table 7.1 presents the different investigated sets of objective weights. Note that when the entire objective function is scaled, the same relation between the sub-objectives is given and hence, the optimization direction is the same. Therefore, we decided to fix objective weight θ_Z to 1. Thus, our objective function now looks as follows:

$$\max \sum_{o\in O, q\in Q, s\in S_o} \mathbb{E}[DS_s]\cdot Z_{oqs} - \theta_V \cdot (V_{\text{max}} - V_{\text{min}}) - \sum_{i\in I}\theta_i \cdot \hat{\gamma}_i - \sum_{w\in W}\theta_w \cdot \hat{\gamma}_w. \tag{7.1}$$

In this chapter, we refer to the parts of the objective function in the following way:

- $\sum_{o\in O, q\in Q, s\in S_o} \mathbb{E}[DS_s]\cdot Z_{oqs}$ is referred to as sub-objective 1,
- $-\theta_V \cdot (V_{\text{max}} - V_{\text{min}})$ is referred to as sub-objective 2,
- $-\sum_{i\in I}\theta_i \cdot \hat{\gamma}_i - \sum_{w\in W}\theta_w \cdot \hat{\gamma}_w$ is referred to as sub-objective 3.

Table 7.1: Table with the different objective weight sets.

Set number	θ_V	θ_i	θ_w
0	0	0	0
1	50000	0	0
2	500000	0	0
3	50000	100	100
4	500000	100	100

We test on three different models. The first model uses the linearization described in Section 5.4 and is referred to as the Log-normal Column Model. The second model is similar to the Log-normal

Column Model, but instead of taking into account the surgery duration distribution as a log-normal, the normal distribution is used. We refer to this model as the Normal Column Model. The last model we consider is the model where the overtime constraint is solved in the same way as in Schneider et al. (2020). This model is referred to as the Schneider Normal Model and is described in Section 5.5. All models use the linearization of the objective function as described in Section 5.3.

To evaluate the suitability of the sets presented in Table 7.1, we examine the performance of the three models using the specified objective weights after running for a duration of twelve hours (note that the optimal value for all three models was attained within the runtime for weight sets 0 and 2). In this analysis, we assess the resulting expected OR utility, $V_{\max}$ and $V_{\min}$ values, and bed occupancy variation. The expected OR utility is expressed as a percentage, representing that achieving 100% utility would necessitate the complete occupation of all available ORs throughout their operating hours for each day $k \in K$. Furthermore, we delve into the values related to the three sub-objectives.

The first weight set, with set number 0, only considers the OR utilization. We expect that for this weight set, each model tries to schedule specialties with on average short surgeries, consequently resulting in $V_{\max} = 100\%$ and an extremely negative $V_{\min}$. This is indeed the case, see Table 7.2. As the second and third sub-objectives were assigned a weight value equal to 0, the first sub-objective is equal to the objective function value. For weight set 0, all three models performed very similarly.

Table 7.2: Results weight set 0.

Model	Obj. function	OR utility [%]	$V_{\max}$[%]	$V_{\min}$ [%]	Gap [%]	Runtime [sec]
Log-normal Column	177199.79	68.24	100.00	-4300.00	0.00	160.5
Normal Column	176118.63	67.82	100.00	-4300.00	0.00	144.5
Schneider Normal	176118.63	67.82	100.00	-4275.19	0.00	2828.4

In weight sets 1 and 2, we incorporate a positive weight for θ_V, while still not factoring in the beds in the objective function ($\theta_i = 0$ and $\theta_w = 0$). The performance of the models with these weight sets can be found in Tables 7.3 and 7.4.

Table 7.3: Results weight set 1.

Model	Obj. function	Sub-obj. 1	Sub-obj 2	Sub-obj. 3	Runtime [hr]
Log-normal Column	126904.66	134482.37	-7577.71	0.00	12.0
Normal Column	126928.62	134502.57	-7573.95	0.00	12.0
Schneider Normal	125108.47	132637.04	-7528.58	0.00	12.0

Model	OR utility [%]	$V_{\max}$[%]	$V_{\min}$ [%]	Gap [%]
Log-normal Column	51.79	0.56	-14.59	0.49
Normal Column	51.80	0.56	-14.58	0.49
Schneider Normal	44.32	0.56	-14.49	13.04

Table 7.4: Results weight set 2.

Model	Obj. function	Sub-obj. 1	Sub-obj 2	Sub-obj. 3	Runtime [sec]
Log-normal Column	101604.80	120981.64	-19375.84	0.00	121.7
Normal Column	101794.40	121170.24	-19375.84	0.00	44.6
Schneider Normal	101794.40	121170.24	-19375.84	0.00	2752.2

Model	OR utility [%]	$V_{\max}$[%]	$V_{\min}$ [%]	Gap [%]
Log-normal Column	46.59	0.56	-3.31	0.00
Normal Column	46.66	0.56	-3.31	0.00
Schneider Normal	46.66	0.56	-3.31	0.00

For both weight sets 1 and 2, the objective function and sub-objective function values are quite similar for the models. We observe that for weight set 2, which has a greater θ_V, the difference between $V_{\max}$ and $V_{\min}$ gets smaller as expected. However, when doing so, the OR utilization decreases, which is unfavorable. Furthermore, we observe that while the Schneider Normal model yields comparable outcomes when it attains the optimal value (for weight sets 0 and 2), its execution time is significantly longer. In contrast, the runtimes of the other two models are relatively comparable. The extended runtime of the Log-normal Column model can be attributed to the slightly larger size of the combination sets J_{ok}, contributing to the increased complexity of the model.

In weight sets 3 and 4, we additionally give the weights θ_i and θ_w a positive value. This increases the complexity of our models. Hence, we expect the optimality gaps of our models to increase. The optimality gap represents the difference between the best known solution and the bound for the best possible solution. The performance of the models with these weight sets can be found in Tables 7.5 and 7.6.

Table 7.5: Results weight set 3.

Model	Obj. function	Sub-obj. 1	Sub-obj 2	Sub-obj. 3	Runtime [hr]
Log-normal Column	123613.81	133179.00	-7577.77	-1987.42	12.0
Normal Column	124684.99	134604.91	-8124.88	-1795.04	12.0
Schneider Normal	19795.84	71697.88	-50045.79	-1856.26	12.0

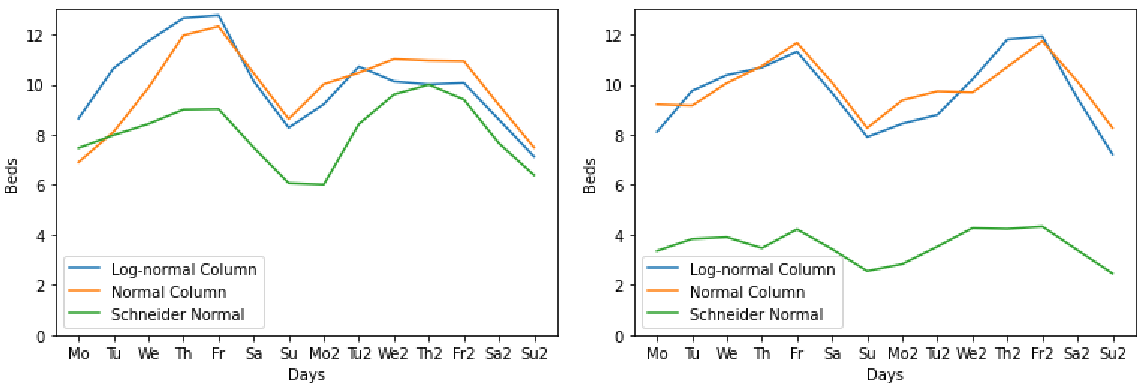
Model	OR utility [%]	$V_{\max}$[%]	$V_{\min}$ [%]	Gap [%]
Log-normal Column	51.29	0.56	-14.59	9.13
Normal Column	51.83	0.56	-15.69	6.66
Schneider Normal	27.61	100.00	-0.09	630.93

Table 7.6: Results weight set 4.

Model	Obj. function	Sub-obj. 1	Sub-obj 2	Sub-obj. 3	Runtime [hr]
Log-normal Column	55978.18	132934.58	-74982.33	-1974.07	12.0
Normal Column	100226.81	121170.24	-19375.84	-1567.59	12.0
Schneider Normal	-370869.21	80645.61	-449556.54	-1958.28	12.0

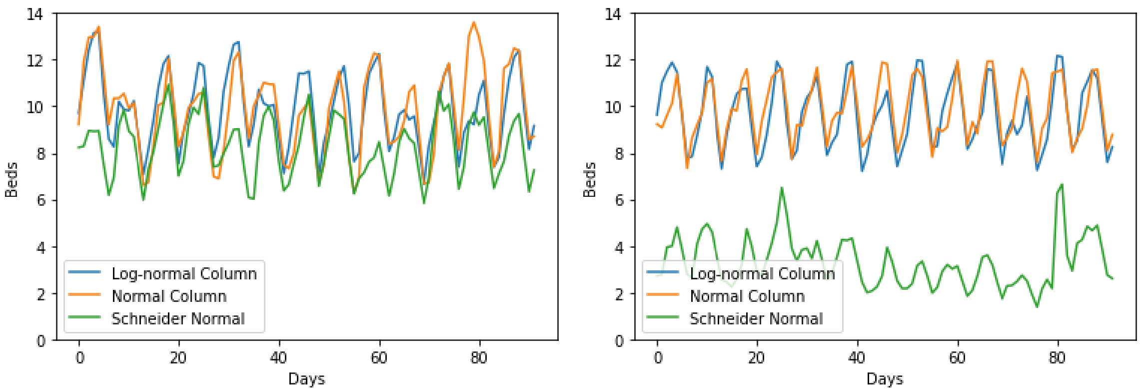
Model	OR utility [%]	$V_{\max}$[%]	$V_{\min}$ [%]	Gap [%]
Log-normal Column	51.19	0.56	-14.43	82.21
Normal Column	46.66	0.56	-3.31	1.80
Schneider Normal	33.06	89.91	-0.00	128.01

Introducing weights for sub-objective 3 should result in a more leveled bed occupancy. This can also be seen in Figure 7.1 and Figure 7.2. In these figures, the expected bed assignments for ward MC2 are visualized. For Figure 7.2, the days in the x-axis are numbered. Mondays are presented by integers that are congruent to $1 \pmod 7$, Tuesdays by integers that are congruent to $2 \pmod 7$, etc. We observe that the graphs obtained for each model with weight set 3 exhibit a shallower slope compared to those at weight set 1, which is as expected. Moreover, the outcome of the Schneider Normal model for weight set 3 can also be rationalized by its subpar performance after a runtime of twelve hours, indicated by the substantial optimality gap observed.



(a) Expected bed assignment for weight set 1. (b) Expected bed assignment for weight set 3.

Figure 7.1: Expected bed assignment for the three models in MC2 with or without weights associated with the bed leveling (i.e., sub-objective 3) when looking at two weeks within the period.



(a) Expected bed assignment for weight set 1. (b) Expected bed assignment for weight set 3.

Figure 7.2: Expected bed assignment for the three models in MC2 with or without weights associated with the bed leveling (i.e., sub-objective 3) when looking at the total period.

It is indeed noticeable that the optimality gap expands for each model in comparison to their performance with the preceding two weight sets. The Schneider Normal model consistently exhibits inferior performance in contrast to the column models. Concerning both column models, the increased value for θ_V does contribute to narrowing the gap between $V_{\max}$ and $V_{\min}$. However, the value of $V_{\max}$ remains unaffected, and since this metric carries greater significance by indicating the proportion of surgical patients that certain specialties might be unable to schedule, a higher θ_V than that of weight sets 1 and 3 does not appear necessary. Furthermore, a more negative $V_{\min}$, and thus, a greater difference between $V_{\max}$ and $V_{\min}$, may even be more favorable as the OR utilization is likely to benefit from this. For the Log-normal Column, it also seems to increase the complexity resulting in an optimality gap more comparable to the Schneider Normal model.

It might still be interesting to look at what happens when we personalize θ_w and θ_i for the wards and ICs. For this, we introduce parameters B_i and B_w for the maximum number of beds in our received dataset for $i \in I$ and $w \in W$, respectively. To investigate this, we have tested the Log-normal Column model with the weight sets 5 and 6 that can be found in the following table.

Table 7.7: Table with the special θ_w and θ_i objective weight sets.

Set number	θ_V	θ_i	θ_w
5	50000	$100/B_i$	$100/B_w$
6	50000	$500/B_i$	$500/B_w$

Table 7.8: Results Log-normal Column model with weight set 3, 5 and 6.

Model	Obj. function	Sub-obj. 1	Sub-obj 2	Sub-obj. 3	Runtime [hr]
Log-normal Column 3	123613.81	133179.00	-7577.77	-1987.42	12.0
Log-normal Column 5	126103.71	133793.46	-7528.58	-2124.80	12.0
Log-normal Column 6	125711.13	134043.49	-7577.77	-10212.40	12.0

Model	OR utility [%]	$V_{\max}$[%]	$V_{\min}$ [%]	Gap [%]
Log-normal Column 3	51.29	0.56	-14.59	9.13
Log-normal Column 5	51.52	0.56	-14.50	7.42
Log-normal Column 6	51.62	0.56	-14.59	7.37

Upon examining bed occupancy alongside the $\hat{\gamma_w}$ and $\hat{\gamma_i}$ values of sub-objective 3, no evident distinction emerged. Consequently, we proceeded to investigate a unified θ_w and θ_i value.

At this point, we are confident that if we choose one model to investigate further, it will not be the Schneider Normal due to its consistently worse performance compared to the other two models. Although the Normal Column performed slightly better for weight set 3 and significantly better for weight set 4 than the Log-normal Column, we believe that the Log-normal Column is still more favorable for the following reasons:

- In Section 6.4, it has been demonstrated that the approach employed for generating combinations in J_{ok} for the Log-normal Column model holds a higher likelihood of accuracy than that of the Normal Column model.
- As previously mentioned, an increased θ_V does not appear to align more favorably with our scheduling objectives. Consequently, a detailed consideration of the performances for weight sets 2 and 4 may not be necessary.

Furthermore, as mentioned before, the underperformance of the Log-normal Column model can be attributed to the larger size of the combination sets J_{ok}, contributing to the increased complexity of the model.

Finally, it is important to discuss the relatively modest OR utility percentage. This can be attributed to the approach of grouping surgeries of the same specialty together, instead of making clusters. While maintaining compliance with the overtime constraint for longer surgeries is a necessity, even if they are less frequent, our models naturally lean towards a more cautious approach in terms of scheduling surgeries. This tendency is influenced by the need to ensure that the overtime constraint remains satisfied.

7.2. Conversion to time schedule

Up till now, our resulting schedule has been the assignment of specialties to ORs on days in a period, accompanied by the number of surgeries the specialty could perform on that specific day. However, as explained in Chapter 2, an MSS gives the time a surgical specialty is allowed to perform surgeries on that day rather than the number of surgeries it can perform. In this section, we explain how we can assign a time duration to a specialty instead.

Again, we make use of the FW method. Suppose an OR is used on a day by specialties $s_1, s_2 \in S$ and our schedule lets s_1 perform n_1 surgeries and s_2 gets to perform n_2 surgeries. Then, by using the FW method we can approximate both the total surgery duration distribution for s_1 with n_1 surgeries, as for s_2 when it has n_2 surgeries. By using the inverse cumulative distribution, we can determine a time duration which the total surgery time does not exceed φ percent of the cases. Furthermore, when we add the cleaning times, we can assign the resulting duration as the assigned duration for that specific specialty in that specific operating room on that particular day.

Generally, it is undesirable for a specialty to exceed its designated OR time, particularly if it leads to delays for other specialties. As a consequence, φ is usually selected to be relatively high. Moreover, for the sake of convenience, we round the assigned durations up to intervals of fifteen minutes.

7.3. Performance of Log-normal Column model

Within this section, we delve deeper into the Log-normal Column model's performance. This investigation involves a comparison of its outcomes with both the provided historical data and a simulation. For this section, we use the Log-normal Column model with weight set 3 from Section 7.1.

With a simulation, we schedule surgeries, as they emerge in historical data, into the MSS resulting from our model. We also take the bed assignments and lengths of stay from the data itself. In order to make a good evaluation, we asked the Sophia Children's Hospital for patient data that we had not used when developing the model. By using data that the model has not seen before, we can assess how well the model generalizes to new, unseen examples. If the same data was used for the simulation, the model's performance could be overly optimistic. Sophia provided us with additional patient data of the year 2022. This data is preprocessed in the same way as described in Chapter 6. For the simulation, we use the data from the first quarter of 2022.

For simplicity, we assume that each patient only has one surgery. This is true for 97.5% of the elective surgical patients from 2019 to 2022. In instances where a patient has multiple surgeries, we create a new virtual patient for each subsequent surgery. The bed assignments for the first surgery (and hence, associated with the original patient number) are the first bed assignments till the date of the next surgery. The corresponding bed assignments for the 'new patients' are the bed assignments starting on the date of the surgery till the date of what would have been the next surgery. In case of the last surgery, the date that the patient went home. In our simulation, we schedule the patients in the same order as in the data.

Firstly, we compare the number of patients that are scheduled by our model and that were scheduled in the data for the first quarter of 2022. In Table 7.9, these numbers are given. Consistent with the anticipated value of $V_{\max} = 0.56$ presented in Table 7.5, the patient count our model schedules aligns with or surpasses the average number of patients in the dataset. When we contrast the model's numbers with the initial quarter's scheduled appointments and assess the percentage of shortfall, we notice that the specialty PDE experiences the highest relative loss. However, it is worth noting that given PDE's overall lower patient count, even minor fluctuations can lead to a high relative loss.

Using the time schedule conversion from the previous section, we can now assign a certain time duration for each OR assignment of a specialty. After scheduling patients in our simulation, we can compare the total duration with the assigned duration. For our conversion, we have chosen $\varphi = 90\%$. On average, compared to the given time, the utilization percentage is 57.12%. This is lower than what they have at the Sophia Children's Hospital, which is 67.29%. The utilization percentages for each specialty's operating room usage can be found in Table 7.10. In our study, we opted to examine the entire range of surgery durations within a single specialty at a given time. Hence, when given the time schedule, a duration is given that is satisfied 90% of the time. This should hold for a specialty's longer surgeries as well as its smaller surgeries. Hence, the model tends to schedule more slack, resulting to a lower OR utilization.

Next, we investigate the extent of overtime that occurred in our simulation. When generating our schedule, we used an $\alpha = 0.1$ for the overtime constraint. Hence, we allow overtime for less than 10% of the time. On average, in contrast to the duration we allocated to a specialty following the time schedule conversion explained in the preceding section, we encountered overtime in 6.5% of instances. In comparison, from the data, we concluded that at the Sophia Children's Hospital, an overtime rate of 14.3% can be found.

Table 7.9: Table with the number of patients that are scheduled in by our model and in the data.

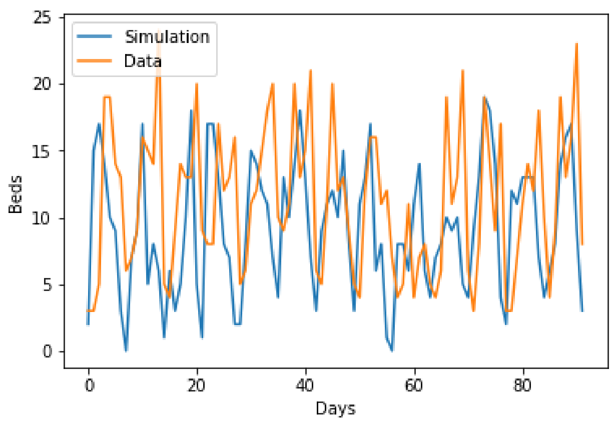
Specialty	Average in data	In first quarter of 2022	In model and simulation
FLEX	69.38	68	75
FPS	90.76	98	104
GYN	65.08	67	74
HPS	11.62	12	13
NSP	10.79	4	12
PDC	40.29	38	46
PDD	17.00	7	19
PDE	6.03	10	6
PDO	154.46	191	177
PG	39.56	35	40
PMFS	32.44	41	36
PNS	37.38	49	42
PDP	17.48	23	20
PDR	138.87	139	155
PDS	155.63	186	176
PO	18.80	15	20
PORS	92.61	119	105
PORS2	17.67	21	20
PU	120.58	148	138

Table 7.10: Table with the OR utilization in the simulation.

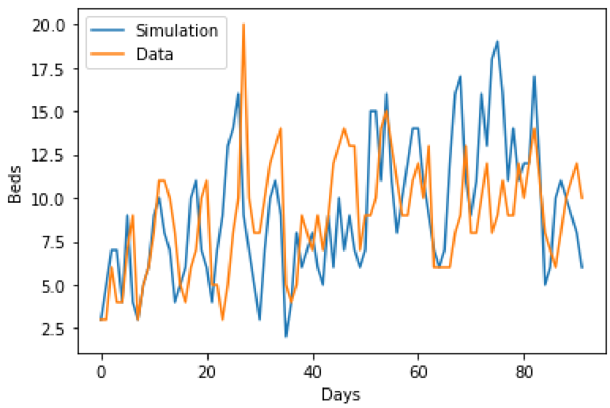
Specialty	OR utilization
FLEX	52.30
GYN	57.13
NSP	34.62
PDC	61.69
PDD	68.34
PDE	75.38
PDO	49.10
PG	56.29
PMFS	69.32
PNS	13.61
PDP	62.06
PDR	69.02
PDS	60.20
PO	56.72
PORS	57.82
PORS2	62.74
FPS	53.61
HPS	48.07
PU	69.05

Considering the relatively low overtime percentage alongside the comparatively modest utilization of operating rooms within the assigned time, it is plausible that the chosen value of φ for the time schedule conversion might be too large. A smaller φ could potentially be more appropriate in this context. Another possibility is to let φ be specialty specific. In our simulation, we observe that among our nineteen specialties, eight consistently avoid overtime, whereas five specialties experience overtime more than 10% of the time, surpassing α. This might indicate that we assign specialties from the first category excessive slack, while the latter category needs more flexibility. On the other hand, the low overtime percentage also means that our schedule gives room for more surgeries. If used in combination with a better planning strategy for the simulation, we could make better use of the OR time and improve the number of scheduled patients without changing the allocated time from our model.

Lastly, we take a look at the bed occupancy for our simulation. Since we schedule the patients in the same order, but not in the same numbers per day, we expect to do a worse job than the planners at Sophia, who take into account the surgery type and its expected duration. In Figure 7.3, the bed occupancy can be found for wards MC1 and MC2 resulting from our simulation and the historical data. For ward MC1, the simulation seems to have fewer beds in use, but the variation in beds has not necessarily improved. The decrease in bed use could be explained by the specialties that are typically assigned to MC1. From Table 2.3, we have seen that specialties PDO, PMFS, PNS and PE are assigned to MC1. From Table 7.9, we observe that these specialties were allocated a smaller number of patients in our simulation compared to real-life occurrences. For ward MC2, the simulation seems to do a better job of leveling the bed occupancy for the first half of the quarter, but the difference between the maximum and minimum bed occupancy is comparable for the simulation and the historical data.



(a) Bed occupancy for MC1.



(b) Bed occupancy for MC2.

Figure 7.3: Bed occupancy for wards MC1 and MC2 resulting from our model expectation, simulation and historical data.

7.4. Performace of adjusted model

In this section, we look at a model more to the preference of the Sophia Children's Hospital. For the convenience of the surgical staff, we set the cycle length at one week ($T = 7$) and additionally, whenever a specialty is assigned, it should have at least one surgery ($Z_{oks} > Y_{oks}$). They also believe that a more selected range of specialties should have the capability to share an OR on a single day. However, it could be argued that a certain degree of flexibility is necessary for improvement. For the adjusted model, only the subsequent specialties are eligible to share an OR with each other: PDP, FLEX, PG, PDD, PO, NSP and PDE.

In Section 7.1, we have seen that increasing the weight for the bed occupancy (i.e., sub-objective 3) is likely to add to the complexity of our model. Adding the constraint mentioned in the previous paragraph does so as well. Hence, we decide to use the following weights when running the adjusted model:

Table 7.11: Table with the weight set for the adjusted model.

θ_V	θ_i	θ_w
50000	20	20

Furthermore, the runtime was set to two full days (i.e., 48 hours), and the resulting values of this weight set can be found in Table 7.12.

Table 7.12: Results for adjusted model.

Obj. function	Sub-obj. 1	Sub-obj 2	Sub-obj. 3
104421.83	132639.57	-27864.97	-352.78

OR utility [%]	$V_{\max}$[%]	$V_{\min}$ [%]	Gap [%]
51.08	2.79	-52.94	12.97

The most comparable model to our adjusted model is the Log-normal Column model with weight set 3. In Table 7.5, we observed a smaller and more favorable $V_{\max}$ for the Log-normal Column model, while this model also considers bed leveling more extensively, attributed to its higher weight for sub-objective 3. The reason for the increased $V_{\max}$ could be the restriction in specialties that can share an OR. In addition to the shorter cycle length, the constraint gives less room for flexibility in the assignment of specialties to an OR.

We, furthermore, investigate if the runtime of 48 hours was too excessive in comparison to the 12 hours we used beforehand. In Figure 7.4, we plotted the objective function value and the best bound value against the runtime in hours. We observe that the found objective function is approached relatively fast after just a few hours. It is mostly the best bound that improves over time. Hence, we might find an acceptable answer after 12 hours, but this would not be reflected in the optimality gap. The optimality gap after 12 hours was equal to 27.0%.

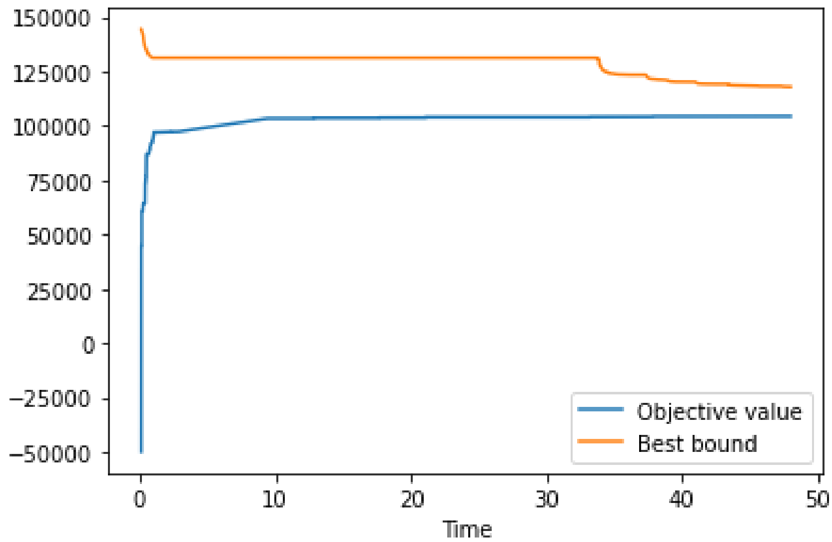


Figure 7.4: Objective function value and the best bound against the runtime in hours for the adjusted model.

The number of patients that are scheduled in the adjusted model and the OR utilization for each specialty can be found in Table 7.13 and Table 7.14, respectively. For the number of patients, we see the same trend as in the previous section: the numbers in the adjusted model are comparable to the values of the average in the historical data, which served as the base for our model. However, the patients can deviate a lot from the average, and for some specialties the actual value for V_s is much higher than expected.

Table 7.13: Table with the number of patients that are scheduled in by the adjusted model and in the data.

Specialty	Average in data	In first quarter of 2022	In the adjusted model
FLEX	69.38	68	78
FPS	90.76	98	91
GYN	65.08	67	99
HPS	11.62	12	14
NSP	10.79	4	13
PDC	40.29	38	52
PDD	17.00	7	26
PDE	6.03	10	8
PDO	154.46	191	152
PG	39.56	35	56
PMFS	32.44	41	42
PNS	37.38	49	39
PDP	17.48	23	25
PDR	138.87	139	135
PDS	155.63	186	158
PO	18.80	15	28
PORS	92.61	119	92
PORS2	17.67	21	27
PU	120.58	148	132

Regarding the utilization of the operating rooms, we observe that in 6.4% of cases, the total duration exceeded the initially assigned duration. On average, our simulation for the adjusted model gives an OR utilization percentage of 60.31%, which is better than the percentage of the Log-normal Column model with weight set 3. As the number of scheduled patients is different for the adjusted model, it is plausible that specialties scheduled more frequently in the adjusted model receive more appropriate time durations. This, in turn, contributes to improved operating room utilization during those assignments.

Table 7.14: Table with the OR utilization in the simulation for the adjusted model.

Specialty	OR utilization
FLEX	46.43
GYN	60.32
NSP	34.47
PDC	62.38
PDD	70.16
PDE	77.43
PDO	45.86
PG	57.09
PMFS	69.75
PNS	14.06
PDP	52.76
PDR	70.94
PDS	58.34
PO	60.74
PORS	53.77
PORS2	64.34
FPS	53.91
HPS	68.00
PU	48.82

8 Conclusion and recommendations

The objective of our research was to develop a model for the Sophia Children's Hospital that provides an MSS and has the objective to not only maximize operating room utilization but also to level the bed occupancy and balance the relative OR assignment fir the specialties. To tackle the uncertainty of the characteristics of future patients, we have taken the length of stay and surgery duration into account in a stochastic manner. This chapter serves as the culmination of our research, presenting both our conclusions and offering recommendations for future studies.

8.1. Conclusion

After analyzing the current situation at the Sophia Children's Hospital in Chapter 2 and reviewing existing literature in Chapter 3, we developed a mathematical model in Chapter 4 which models the length of stay similarly as Schneider et al. (2020). The model produces a schedule from which we can determine the number of patients whose surgeries can be arranged for a given specialty on a specific day and in a specific operating room.

The model had two non-linear components: a non-linear overtime constraint, as well as a non-linear sub-objective concerning the bed variation. In order to solve the model, we provided linearization methods for these components in Chapter 5. For the overtime constraint, we provided two methods: a linearization under the assumption that surgery durations are normally distributed as proposed by Schneider et al. (2020), and a linearization where the log-normal distribution is assumed for the individual surgery durations. The latter is a creation of our own. It uses an approximation method to find one suitable log-normal distribution for the sum of the surgery durations. In order to find a suitable approximation method, we looked into three methods commonly used in the field of telecommunication: the Fenton-Wilkinson method, the Swartz-Yeh method and the Mehta approximation. After further examination, we concluded that the Fenton-Wilkinson method was most suitable for our research.

Using the Fenton-Wilkinson approximation, we can calculate which pairs of specialties with corresponding number of surgeries, could be assigned an OR together without violation of the overtime constraint. After generating all possible pairs, in our research referred to as combinations, we incorporated them into our model with a column generation based approach. Furthermore, we can utilize the Fenton-Wilkinson approximation to transform the schedule presented by our model into a schedule based on time duration rather than the number of surgeries a specialty can perform.

In Chapter 6, we analyzed the data that the hospital provided us. The data included properties about the patients' surgeries and bed assignments. Due to the presence of errors in the data, we conducted preprocessing before utilizing it as input in our modeling. Additionally, we conducted goodness of fit tests to assess whether adopting the log-normal distribution for surgery durations was genuinely superior to the normal distribution. Our analysis revealed that, for the majority of instances, the log-normal distribution outperformed the normal distribution. This was the case for individual surgeries, as well as the Fenton-Wilkinson approximation for the duration of multiple surgeries.

As previously stated, we had two approaches to linearize the overtime constraint: one was based on the work of Schneider et al. (2020), and the other involved a column generation approach utilizing the Fenton-Wilkinson approximation. We refer to them as the Schneider Normal model and the Log-normal Column model, respectively. Moreover, as the column generation based approach could be easily adapted to accommodate normally distributed surgery durations, we also developed a Normal Column model.

In Chapter 7, we first compared the performance of the three models under different objective weights. The objective weight θ_V influenced the proportion of surgery assignments for different specialties, while θ_w and θ_i affected the variation in bed occupancy. When these objective weights were set to zero, the models prioritized achieving a high expected OR utilization exclusively and only part of the specialties gets scheduled. Appointing a positive value to weight θ_V ensures that all specialties are assigned to an OR and that this happened relatively fair. However, appointing a very large value for θ_V does not necessarily increase OR assignment for the specialty that is relatively scheduled the least.

When the objective weights for the bed occupancy were added, we observed that the expected bed occupancy variation became smaller. However, this is at the expense of the complexity of the problem. We investigated weights that were both dependent and independent for their respective wards but could not find a significant difference.

Upon evaluating the performance of the three models, it becomes evident that in instances where optimality was achieved, the two column-based models exhibited runtimes of up to 3 minutes. In contrast, the Schneider Normal model required approximately 47 minutes to converge and find the solution. When optimality was not achieved, the Schneider Normal model underperformed in comparison to the other models as well. For most weight sets, the two column based models performed similarly. Only for the weight set with the largest objective weight values did the Log-normal Column model underperform compared to the Normal Column model. This could be attributed to the larger complexity the Log-normal Column model due to its larger combinations sets. As the combinations of the Log-normal Column model are likely to be more accurate than the Normal Column model, and we deemed a large value for θ_V not necessary, we decided to continue with the Log-normal Column model for the model comparison with a simulation in Section 7.3.

In a given quarter, the simulation schedules the same number of patients for a specialty into the schedule as determined by our model. Since our model's patient scheduling for a specialty relies on the historical quarterly average for that specific specialty and the actual number could deviate greatly from this, we noticed instances where certain specialties had fewer surgery assignments than what they actually needed.

As mentioned before, we can use the Fenton-Wilkinson method to convert the schedule produced by our model into a time schedule. This can be done by using the inverse cumulative distribution to determine a suitable duration for which a specialty finishes its assigned surgeries in time for φ percent of the instances. After scheduling patients in our simulation, we can determine the OR utilization by comparing the sum of the patients' surgery duration to the assigned duration. We had chosen φ to be equal to 90%, which resulted in an average OR utilization percentage of 57.12% and an overtime occurrence of 6.5% of the instances. Lastly, when looking at the bed occupancy for our simulation and the data, no significant differences could be found.

8.2. Recommendations

This section offers suggestions for the Sophia Children's Hospital, along with recommendations for potential future research.

8.2.1. Recording of data

Due to the presence of errors in the data that was provided by the Sophia Children's Hospital, we conducted preprocessing before utilizing it as input in our modeling. At the Sophia Children's Hospital, the program HiX is used for the registration of surgery data. The data is registered manually without any automated checks for mistakes. To enhance future data analyses, we suggest the implementation of automated checks. For instance, if errors occur in vital information fields leading to incorrect sequencing, a notification system could be established. It could be as simple as a change in field color or a pop-up alert.

Another suggestion concerning data registration at the Sophia Children's Hospital is to standardize the various systems. For instance, there exist different abbreviations for the same specialty in the MSS and the exported data from HiX. Additionally, data entered into HiX is labeled differently in the export sheet. To avoid confusion, we suggest maintaining consistency across all hospital data.

8.2.2. Future research

As mentioned before, when scheduling real patients' data in our schedule, we obtain a relatively low OR utilization percentage while the overtime constraint is not near its bound. One possible explanation for this could stem from our consideration of all surgeries within a given specialty. Since we do not differentiate between types of surgeries, the duration assigned to, say, a collection of commonly short surgeries in one specialty is treated the same as a group of typically longer surgeries within the same specialty, as long as their group size remains comparable. If a distinction was made between surgery types, the model could provide more accurate time duration estimations.

Another explanation could be that the chosen φ used for the conversion to a time schedule is chosen too large. Opting for a smaller value, perhaps even making it dependent on the specialty, could possibly result in a better master surgery schedule with less redundant slack. This choice would lead to assigning shorter total durations, thereby increasing OR utilizations and permitting more controlled overtime within the specified limit.

Additionally, a low OR utilization rate indicates the potential for accommodating more surgeries. In our simulation, we allocate precisely the number of patients as indicated by the schedule. This allocation is carried out in the order of data appearances, without employing any sophisticated scheduling strategies. If the time schedule would be used by a planning expert in the Sophia Children's Hospital, the performance of our schedule will most likely improve significantly. While arranging surgeries, scheduling experts also consider the probable ward assignment for each patient. By employing a scheduling approach that imitates the strategy employed at the Sophia Children's Hospital, we can observe the real impact our model has on bed occupancy.

Next to suggestions to improve the evaluation of our model, we also have some suggestions on the model itself. The model utilizes the average patient count as a base for the number it will schedule. However, the actual number of patients can significantly deviate from this. When the actual number is greater than the number we schedule, the variation negatively impacts the waiting list. Therefore, it is valuable to examine the trend in incoming patients to consider using it as the base for our model instead. Another improvement might involve considering both the average and the standard deviation in patient planning to mitigate potential patient scheduling shortfalls. Although, this can negatively impact the OR utilization due to increased slack.

Furthermore, we have only tried the model for a selection of weight sets. It could be that a different choice of weight set suits the needs of the hospital better.

A

Model parameters and variables

Table A.1: Parameters of the model.

Parameter	Description
Q	Number of days that the MSS covers
M_s	Maximum number of patients specialty $s \in S$ can schedule on one day
χ_k	Maximal number of open ORs on day $k \in K$
C_s	Number of surgeons from specialty $s \in S$
T	Length of a cycle in the MSS
κ	Cleaning time between surgeries
m_s	Minimum number of surgeries for specialty $s \in S$
N_{ps}	Maximum length of stay for a patient of specialty $s \in S$ in the latest ward in the subpath $p \in P$
N_s^I	Maximal length of stay for the all subpaths of P^I
N_s^W	Maximal length of stay for the all subpaths of P^W
β_{ok}	Duration OR $o \in O$ was planned to be open on day $k \in K$
α	Allowed probability of overtime
$\gamma_{iq}(\psi)$	number such that the number of occupied beds on day $q \in K$ in IC $i \in I$ for schedule ψ is at most $\gamma_{iq}(\psi)$ with probability at least $1-\epsilon$
$\gamma_{wq}(\psi)$	number such that the number of occupied beds on day $q \in K$ in ward $w \in W$ for schedule ψ is at most $\gamma_{wq}(\psi)$ with probability at least $1-\epsilon$
$\gamma_i(\psi)$	Variation in number of required beds on IC $i \in I$ for a given schedule ψ
$\gamma_w(\psi)$	Variation in number of required beds on ward $w \in W$ for a given schedule ψ
L_s	Average number of surgeries specialty $s \in S$ had during the past Q-length periods

Table A.2: Variables of the model.

Variable	Description
Z_{oks}	Integer decision variables representing the number of surgeries scheduled on OR $o \in O$ on day $k \in K$ by specialty $s \in S_o$
Y_{oks}	Binary auxiliary variables indicating whether any surgeries are scheduled in OR $o \in O$ on day $k \in K$ for specialty $s \in S_o$
X_{ok}	Binary variable denoting whether OR $o \in O$ is open on day $k \in K$
R_{oks}	Binary variable indicating whether in OR $o \in O$ specialty $s \in S_o$ has at least one surgery on day $k \in K$
W_{ok}	Binary variable indicating whether in OR $o \in O$ exactly two specialties in S_o had surgeries planned on day $k \in K$
V_s	Auxiliary variables representing the percentage not scheduled surgeries with respect to the expected number of surgeries of specialty $s \in S$
V_{max}	Variable associated to the largest V_s
V_{min}	Variable associated to the smallest V_s

Table A.3: Sets for the post-operative patient model.

Set	Description
S	Specialties
O	ORs
S_o	Specialties compliant to OR $o \in O$
K	Days when the OR is open
I_s	Specialties s_2 such that $\hat{s}_2 < \hat{s}$ and s_2 can be combined with specialty $s \in S$
$\bar{O}_k$	ORs that cannot be closed on day k
S_o^C	Specialties that can only have surgeries planned in OR $o \in O$ in combination with another specialty.
I	Nursing wards of IC
W	Nursing wards of MC
D	Day care department
P	Post-surgical patient subpaths through the nursing wards
P^I	Post-surgical patient subpaths ending with an IC $i \in I$
P^W	Post-surgical patient subpaths ending with a ward $w \in W$
P^D	Post-surgical patient subpathss ending with the DC
S_w	Specialties compliant to ward $w \in W$
S_i	Specialties complaint to IC $i \in I$

Table A.4: Probabilities for the post-operative patient model.

Notation	Probability
g_{ok}	Probability distribution of total OR time of OR $o \in O$ on day $k \in K$
a_{pis}	Probability that a patient of specialty $s \in S$ is being transferred to IC $i \in I$ as the last nursing ward of subpath $p \in P^I$
b_{pws}	Probability that a patient of specialty $s \in S$ is being transferred to ward $w \in W$ as the last nursing ward of subpath $p \in P^W$
c_{psn}	Probability that a patient of specialty $s \in S$ stays n days in the last nursing ward of subpath $p \in P$
$d_{ps(n+1)}$	Probability that a patient of specialty $s \in S$ is discharged after staying exactly n days in the last ward of subpath $p \in P$
e_{psn}	Probability that a patient of specialty $s \in S$ is in the last ward of subpath $p \in P$ on day n in case the patient has visited one nursing department after surgery
e_{psnm}	Probability that a patient of specialty $s \in S$ is in the last ward of subpath $p \in P$ on day n and stayed m days in the preceding nursing department in case of visiting 2 nursing departments
$e_{psnm_1m_2}$	Probability that a patient of specialty $s \in S$ is in the last ward of path $p \in P$ on day n and stayed m_1 and m_2 days in the preceding nursing departments in case of visiting 3 nursing departments
f_{isn}	Probability distributions that a patient of specialty $s \in S$ is in IC $i \in I$ on day n
f_{wsn}	Probability distributions that a patient of specialty $s \in S$ is in ward $w \in W$ on day n
$\hat{H}^I_{iq}[n]$	Probability of having n patients in IC $i \in I$
$\hat{H}^W_{wq}[n]$	Probability of having n patients in ward $w \in W$

Table A.5: Distributions.

Notation	Distribution
F^I_{isq}	Distribution of the number of recovering patient of specialty $s \in S$ occupying a bed in IC $i \in I$ on the qth day when the surgery occurred on the first day
F^W_{wsq}	Distribution of the number of recovering patient of specialty $s \in S$ occupying a bed in ward $w \in W$ on the qth day when the surgery occurred on the first day
G^I_{ioksq}	Distribution of the bed occupancy in IC $i \in I$ on day q resulting from one surgery of specialty $s \in S$ being scheduled on day $k \in K$ in OR $o \in O$
G^W_{woksq}	Distribution of the bed occupancy in ward $w \in W$ on day q resulting from a surgery of specialty $s \in S$ being scheduled on day $k \in K$ in OR $o \in O$
$\hat{G}^I_{ioksq}$	Distribution of the bed occupancy in IC $i \in I$ on day q resulting from the surgeries of specialty $s \in S$ scheduled on day $k \in K$ in OR $o \in O$
$\hat{G}^W_{woksq}$	Distribution of the bed occupancy in ward $w \in W$ on day q resulting from the surgeries of specialty $s \in S$ scheduled on day $k \in K$ in OR $o \in O$
H^I_{iokq}	Distribution of the bed occupancy in IC $i \in I$ on day q resulting from all surgeries of OR $o \in O$ on day $k \in K$
H^W_{wokq}	Distribution of the bed occupancy in ward $w \in W$ on day q resulting from all surgeries of OR $o \in O$ on day $k \in K$
$\hat{H}^I_{iq}$	Distribution of bed occupancy in IC $i \in I$ on day q of the schedule
$\hat{H}^W_{wq}$	Distribution of bed occupancy in ward $w \in W$ on day q of the schedule

B

Explicit formulas considered paths

- $c_{(SI)sn}$: probability that a patient of specialty $s \in S$ stays n days in the IC after surgery.
- $c_{(SWI)sn}$: probability that a patient of specialty $s \in S$ stays n days in the IC after spending time at some nursing ward after surgery.
- $c_{(SW1)sn}$: probability that a patient of specialty $s \in S$ stays n days at one of the wards in W after surgery and is discharged after the stay at a ward of W.
- $c_{(SW2)sn}$: probability that a patient of specialty $s \in S$ stays n days at one of the wards in W after surgery and is transferred to the IC after the stay at a ward W.
- $c_{(SIW)sn}$: probability that a patient of specialty $s \in S$ stays n days at one of the wards in W after spending time at the IC after surgery.
- $c_{(SWIW)sn}$: probability that a patient of specialty $s \in S$ stays n days at one of the wards in W after spending first time at a ward in W, followed by the IC after surgery.
- $c_{(SD)sh}$: probability that a patient of specialty $s \in S$ stays h hours at the DC after surgery.

$$d_{(SI)s(n+1)} = \frac{c_{(SI)sn}}{1 - \sum_{k=1}^{n-1} c_{(SI)sk}} \approx \frac{c_{(SI)sn}}{\sum_{k=n}^{N_{(SI)s}} c_{(SI)sk}} \qquad s \in S, n \in \{0, \dots, N_{(SI)s}\} \tag{B.1}$$

$$d_{(SWI)s(n+1)} = \frac{c_{(SWI)sn}}{1 - \sum_{k=1}^{n-1} c_{(SWI)sk}} \approx \frac{c_{(SWI)sn}}{\sum_{k=n}^{N_{(SWI)s}} c_{(SWI)sk}} \qquad s \in S, n \in \{0, \dots, N_{(SWI)s}\} \tag{B.2}$$

$$d_{(SW1)s(n+1)} = \frac{c_{(SW1)sn}}{1 - \sum_{k=1}^{n-1} c_{(SW1)sk}} \approx \frac{c_{(SW1)sn}}{\sum_{k=n}^{N_{(SW1)s}} c_{(SW1)sk}} \qquad s \in S, n \in \{0, \dots, N_{(SW1)s}\} \tag{B.3}$$

$$d_{(SW2)s(n+1)} = \frac{c_{(SW2)sn}}{1 - \sum_{k=1}^{n-1} c_{(SW2)sk}} \approx \frac{c_{(SW2)sn}}{\sum_{k=n}^{N_{(SW2)s}} c_{(SW2)sk}} \qquad s \in S, n \in \{0, \dots, N_{(SW2)s}\} \tag{B.4}$$

$$d_{(SIW)s(n+1)} = \frac{c_{(SIW)sn}}{1 - \sum_{k=1}^{n-1} c_{(SIw)sk}} \approx \frac{c_{(SIW)sn}}{\sum_{k=n}^{N_{(SIW)s}} c_{(SIW)sk}} \qquad s \in S, n \in \{0, \dots, N_{(SIW)s}\} \tag{B.5}$$

$$d_{(SWIW)s(n+1)} = \frac{c_{(SWIW)sn}}{1 - \sum_{k=1}^{n-1} c_{(SWIw)sk}} \approx \frac{c_{(SWIW)sn}}{\sum_{k=n}^{N_{(SWIW)s}} c_{(SWIW)sk}} \qquad s \in S, n \in \{0, \dots, N_{(SWIW)s}\} \tag{B.6}$$

$$e_{(SI)sn} = \begin{cases} \sum_{i \in I} a_{(SI)i,s} & \text{for } n = 1 \\ (1 - d_{(SI)s(n-1)}) e_{(SI)s(n-1)} & \text{for } n \in \{2, \dots, N_{(SI)s} + 1\} \\ 0 & \text{otherwise} \end{cases} \tag{B.7}$$

$$e_{(SW1)sn} = \begin{cases} \sum_{w \in W} b_{(S1)ws} & \text{for } n = 1 \\ (1 - d_{(SW1)s(n-1)}) e_{(SW1)s(n-1)} & \text{for } n \in \{2, \dots, N_{(SW1)s} + 1\} \\ 0 & \text{otherwise} \end{cases} \tag{B.8}$$

$$e_{(SW2)sn} = \begin{cases} \sum_{w \in W} b_{(S2)ws} & \text{for } n = 1 \\ (1 - d_{(SW2)s(n-1)}) e_{(SW2)s(n-1)} & \text{for } n \in \{2, \dots, N_{(SW2)s} + 1\} \\ 0 & \text{otherwise} \end{cases} \tag{B.9}$$

$$e_{(SWI)snm} = \begin{cases} (1 - d_{(SWI)s1}) e_{(SW)sm} d_{(SW)sm} & \text{for } m \in \{1, \dots, N_{(SW)s} + 1\}, n = m \\ (1 - d_{(SWI)s(n-m)}) e_{(SWI)s(n-1)m} & \text{for } m \in \{1, \dots, N_{(SW)s} + 1\}, n \in \{m + 1, \dots, m + N_{(SWI)s} + 1\} \\ 0 & \text{otherwise} \end{cases} \tag{B.10}$$

$$e_{SIW,s,n,m} = \begin{cases} (1 - d_{SIW,s,1}) e_{SI,s,m} d_{SI,s,m} & \text{for } m \in \{1, \dots, N_{SI,s} + 1\}, n = m \\ (1 - d_{SIW,s,n-m}) e_{SIW,s,n-1,m} & \text{for } m \in \{1, \dots, N_{SI,s} + 1\}, n \in \{m + 1, \dots, m + N_{SIW,s} + 1\} \\ 0 & \text{otherwise} \end{cases} \tag{B.11}$$

$$e_{(SWIW)snm_1m_2} = \begin{cases} (1 - d_{(SWIW)s1}) e_{(SWI)sm_2m_1} d_{(SWI)s(m_2-m_1+1)} & \begin{array}{l} \text{for } m_1 \in \{1, \dots, N_{(SW)s} + 1\}, \\ m_2 \in \{m_1, \dots, m_1 + N_{(SWI)s}\}, n = m_2 \end{array} \\ (1 - d_{(SWIW)s(n-m2+1)}) e_{(SWIW)s(n-1)m_1m_2} & \begin{array}{l} \text{for } m_1 \in \{1, \dots, N_{(SW)s} + 1\}, \\ m_2 \in \{m_1, \dots, m_1 + N_{(SWI)s}\}, \\ n \in \{m_2 + 1, \dots, m_2 + N_{(SWIW)s}\} \end{array} \\ 0 & \text{otherwise} \end{cases} \tag{B.12}$$

Bibliography

Milton Abramowitz and Irene A Stegun. Handbook of mathematical functions with formulas, graphs, and mathematical tables. In *Appl*, 1972.

Jeroen Beliën, Erik Demeulemeester, and Brecht Cardoen. A decision support system for cyclic master surgery scheduling with multiple objectives. *Journal of scheduling*, 12(2):147, 2009.

Jeroen Beliën and Erik Demeulemeester. Building cyclic master surgery schedules with leveled resulting bed occupancy. *European Journal of Operational Research*, 176:1185–1204, 01 2007. doi: 10.1016/j.ejor.2005.06.063.

Bart Berden, Leo Berrevoets, and Windi Winasti. *Capaciteitsplanning in de zorg*. Springer, 2016.

Johannes Bisschop. *AIMMS optimization modeling*. Lulu. com, 2006.

J.M. Bosch. Better utilisation of the or with less beds : a tactical surgery scheduling approach to improve or utilisation and the required number of beds in the wards, August 2011. URL http://essay.utwente.nl/61486/.

Malcolm Faddy, Nicholas Graves, and Anthony Pettitt. Modeling length of stay in hospital and other right skewed data: Comparison of phase-type, gamma and log-normal distributions. *Value in Health*, 12 (2):309–314, 2009. ISSN 1098-3015. doi: https://doi.org/10.1111/j.1524-4733.2008.00421.x. URL https://www.sciencedirect.com/science/article/pii/S1098301510607097.

L. Fenton. The sum of log-normal probability distributions in scatter transmission systems. *IRE Transactions on Communications Systems*, 8(1):57–67, 1960. doi: 10.1109/TCOM.1960.1097606.

Andreas Fügener. An integrated strategic and tactical master surgery scheduling approach with stochastic resource demand. *Journal of Business Logistics*, 36(4):374–387, 2015.

Andreas Fügener, Erwin W Hans, Rainer Kolisch, Nikky Kortbeek, and Peter T Vanberkel. Master surgery scheduling with consideration of multiple downstream units. *European journal of operational research*, 239(1):227–236, 2014.

Erwin W Hans, Mark van Houdenhoven, and Peter JH Hulshof. A framework for healthcare planning and control. In *Handbook of healthcare system scheduling*, pages 303–320. Springer, 2012.

Aida Jebali and Ali Diabat. A stochastic model for operating room planning under capacity constraints. *International Journal of Production Research*, 53(24):7252–7270, 2015.

Enis Kayış, Taghi T Khaniyev, Jaap Suermondt, and Karl Sylvester. A robust estimation model for surgery durations with temporal, operational, and surgery team effects. *Health care management science*, 18:222–233, 2015.

Guoxuan Ma and Erik Demeulemeester. A multilevel integrative approach to hospital case mix and capacity planning. *Computers & Operations Research*, 40(9):2198–2207, 2013.

Neelesh B. Mehta, Jingxian Wu, Andreas F. Molisch, and Jin Zhang. Approximating a sum of random variables with a lognormal. *IEEE Transactions on Wireless Communications*, 6(7):2690 – 2699, 2007. doi: 10.1109/TWC.2007.051000. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-34547485292&doi=10.1109%2fTWC.2007.051000&partnerID=40&md5=495ef7cb164717a66bc9c952f3cf5832. Cited by: 278.

Rym M'Hallah and Filippo Visintin. A stochastic model for scheduling elective surgeries in a cyclic master surgical schedule. *Computers & Industrial Engineering*, 129:156–168, 2019.

Guillaume Sagnol, Christoph Barner, Ralf Borndörfer, Mickaël Grima, Matthes Seeling, Claudia Spies, and Klaus Wernecke. Robust allocation of operating rooms: A cutting plane approach to handle lognormal case durations. *European Journal of Operational Research*, 271(2):420–435, 2018. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2018.05.022. URL https://www.sciencedirect.com/science/article/pii/S037722171830420X.

AJ Thomas Schneider, J Theresia van Essen, Mijke Carlier, and Erwin W Hans. Scheduling surgery groups considering multiple downstream resources. *European journal of operational research*, 282 (2):741–752, 2020.

S. C. Schwartz and Y. S. Yeh. On the distribution function and moments of power sums with lognormal components. *The Bell System Technical Journal*, 61(7):1441–1462, 1982. doi: 10.1002/j.1538-7305.1982.tb04353.x.

Pieter S Stepaniak, Christiaan Heij, Guido HH Mannaerts, Marcel de Quelerij, and Guus de Vries. 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*, 109(4):1232–1245, 2009.

Peter T Vanberkel, Richard J Boucherie, Erwin W Hans, Johann L Hurink, Wineke AM Van Lent, and Wim H Van Harten. An exact approach for relating recovering surgical patient workload to the master surgical schedule. *Journal of the Operational Research Society*, 62(10):1851–1860, 2011.

Lien Wang, Erik Demeulemeester, Nancy Vansteenkiste, and Frank E. Rademakers. Operating room planning and scheduling for outpatients and inpatients: A review and future research. *Operations Research for Health Care*, 31:100323, 2021. ISSN 2211-6923. doi: https://doi.org/10.1016/j.orhc.2021.100323. URL https://www.sciencedirect.com/science/article/pii/S2211692321000394.