MASTER

Integrated surgery scheduling

creating a tactical surgical schedule for the general surgery specialism considering the operating room usage and the bed occupations at different nursing wards

Jongen, T.P.H.

Award date:
2020

Integrated Surgery Scheduling

Creating a tactical surgical schedule for the general surgery specialism considering the operating room usage and the bed occupations at different nursing wards

In partial fulfillment of the requirements for the degree of

Master of Science in

Operations Management and Logistics

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

elkerliek ZIEKENHUIS

By T.P.H. (Thom) Jongen
Student number: 0898378
22-05-2020

Helmond, 22-5-2020

Creating a tactical surgical schedule for the general surgery specialism considering the operating room usage and the bed occupations at different nursing wards

By T.P.H. (Thom) Jongen
Student number: 0898378

In partial fulfillment of the requirements for the degree of

Master of Science in

Operations Management and Logistics

1st Supervisor TU/e:	Dr. M. (Murat) Firat
2nd Supervisor TU/e:	Dr. Ir. N. P. (Nico) Dellaert
3rd Supervisor TU/e:	Dr. Y. (Yingqian) Zhang
Supervisor Elkerliek:	M. (Mark) Varekamp
Supervisor Elkerliek:	W. (Willemijn) Hendrikx

Keywords

Integrated scheduling, Operating room, bed occupation, mixed integer quadratic programming, recourse leveling

Preface

The master thesis you are about to read is the final step in finishing my master Operational Management and Logistics at the Technical University of Eindhoven. Finishing this project does not only mean that I am done with my study, it also means that an important chapter of my life is over. I have always very much enjoyed my time studying and I am sure that I am going to miss it. On the other hand, I am excited to start a new chapter in my life. Before moving on to the content of the master thesis, I would like to thank everyone that has contributed in any way to making my student life the way it has been.

Furthermore, I would like to especially thank Murat Firat, my first supervisor and mentor, for his guidance during the project. He has spent a lot of time helping me finish this research in the best way possible. Even in the trying times due to the coronavirus, he continued to be of great support for me.

Several months ago I started my research at the Elkerliek hospital in Helmond. Looking back I can conclude that it has been an educational, though, strange, fun, and unique time. The Elkerliek has been very welcoming from the first day I started my research and it continued to be so. Every day, I have come to the Elkerliek with pleasure and I would like to thank everyone I have had the chance of meeting. In particular, I would like to thank Mark Varekamp and Willemijn Hendrikx for helping me with my research. They have been the supervisors at the hospital and were always available when I needed them. Furthermore, I would like to thank Stan Janssen for accepting my request to conduct a research at the Elkerliek and for being so involved and helpful. Finally, I would like to thank the other students at the Elkerliek, Jochem (who has already graduated), Laura, Anne, Cas, and Robert for making my time at the hospital so much fun. I really liked the atmosphere we have created in 'our' office and I appreciate the way we always tried to help each other. It has been a real shame that the last two months of my thesis had to be completed from home. Due to the Corona virus it was requested to work from home when possible. That is why I especially appreciate the (bi)weekly coffee meeting via zoom with Stan and the student team.

Finally, I would like to thank my family and friends. I would like to thank my parents for supporting me and helping me whenever they could, I really appreciate it. Next, I would like to thank my friends for helping me unwind during the weekends, this has been much needed. Finally, I would like to thank my girlfriend Lauri, for sticking by me and for always being there when I need it. .

Thom Jongen, May 2020

Executive summary

In the current healthcare environment, resource management plays an increasingly important role. Since providing healthcare is becoming increasingly expensive and government funding remains the same, the available resources should be used as effectively and efficiently as possible. The focus of this research is the optimization of the usage of the operating rooms (ORs) and the beds at different wards by creating a tactical OR schedule that takes both the OR and the downstream resource 'bed capacity' into account. These resources have been chosen because they are some of the most critical ones.

The research is conducted within the Elkerliek hospital in Helmond. The research is conducted because the Elkerliek experiences difficulties in aligning the available resources with the demand of the patients due to the highly fluctuating number of patients that are staying at the wards over time. To deal with these fluctuations, extra resources are needed to cope with peaks in demand. The output of the OR is partly responsible for the fluctuating demand for the wards. A fraction of this fluctuation is inevitable, due to the treatment of acute patients that cannot be scheduled far in advance. However, some of the fluctuations in demand come from the elective patient flow. This is the part on which this research has been conducted. At the moment, the Elkerliek focuses on the usage of the OR when creating an OR schedule. While creating this schedule, the downstream resources (beds and nurses among others) are seen more as a restriction than as resources that should also be used as efficiently as possible. The goal of this research is to create a tactical OR schedule that takes both the OR as well as the bed occupation at the different wards into account. The scope of the research is concentrated around the *general surgery department* and three wards: *the surgical ward (1BC), the day-care ward (2BC), and the short stay ward (2D)*. Based on the problems experienced by the Elkerliek, the following research question is formulated:

"How can statistical information on the provided datasets and the use of a MIQP model help the Elkerliek when developing a tactical surgical schedule for a selected department taking both the OR and the downstream resources at the wards into account?"

To answer this research question, the following consecutive steps were taken: Data acquisition (step 1), description of the current situation and determining the KPIs (step 2), creation of patient groups (step 3), creation of mixed integer quadratic programming (MIQP) models (step 4), testing the performance of the model (step 5), testing different scenarios (step 6) and finally, interpreting the results and drawing conclusions (step 7). These steps will be shortly explained in this summary.

Data acquisition
The first step of the research was the acquisition of the required data. For this research, data had to be gathered on the surgeries and the admission of patients of the general surgery specialism. The final dataset that was used in this research consisted of 5180 rows, each storing the information on one surgery and the corresponding admission. The dataset contained information of all surgeries performed by the general surgery specialism that occurred between 1-1-2018 and 24-9-2019 and the corresponding admissions to those surgeries.

Description of the current situation and determining the KPIs

For the next step, KPIs were determined and the current performance of the Elkerliek with regard to these KPIs was analysed. In total, 4 different KPIs were defined: The percentage of the available OR time that is used efficiently (KPI 1), the average overtime per OR-block in minutes (KPI 2), the percentage of days within the planning horizon that a deviation from the target bed occupation of 4 beds or more occurred (KPI 3) and finally, the average deviation from the target bed occupation in the number of beds (KPI 4). From the analysis of the current performance of the Elkerliek, it was concluded that the current OR-schedules are partly to blame for the high fluctuations at the wards. The current scores of the Elkerliek for the different KPIs are used as a baseline to compare the newly created schedules with.

Creation of patient groups

The third step was the creation of the patient groups. Patient groups were created because this research was conducted on the tactical level, which means that patients are not considered individually. The groups were created based on **what** resources each patient uses (the sub-specialism responsible for the patient and the ward the patient goes to) and **how much** each patient uses the recourses (the surgery duration of the patient and the length of stay (LOS) of the patient). To create patient groups, patients were first split based on what resources they have used. This means that for each sub-specialism/ward combination a set of patients was created. Then, the outliers in terms of surgery time and LOS were removed from each of these sets. Finally, with the help of a clustering tool, the patients of each sub-specialism/ward set were grouped based on the patients' surgery time and LOS. Eventually, 31 different patient groups were defined. The characteristics of these patient groups were used as an input for the MIQP models.

Creation of the MIQP models

The fourth step was the creation of the MIQP models. The scheduling problem was first defined as a MIQP problem that considered each patient individually. This individual model was used as a basis for the group MIQP model that makes use of the patient groups that were created in the previous step. The objective function of both MILQ models was to minimize the quadratic sum of under- and overutilization of the resources compared to the target utilization. This objective had to be achieved under certain restrictions like the maximum number of beds available on each ward. The decision variables for the model were: The number of patients operated on each day (for the individual model), the number of patients of each group operated on each day (for the group model) and the number of OR-blocks a sub-specialism can use on each day. The individual model was only used as a basis for the group model, to validate the group model and to compare the performance of the group model with. The group model was used to generate the results for this research.

Performance of the model

Testing the performance of the different models was the fifth step of the research. This performance was evaluated based on runtime and the gap towards optimality. Both the group model and the individual model were tested in the same way. This means that the performance of the two models could be compared to each other. This way it was possible to see what the effect was of considering groups instead of individual patients. Based on the results of the different performance tests, several conclusions were drawn. First, considering groups instead of individual patients reduces the complexity of the model. Second, having the model decide the OR-block allocation, as well as Patient scheduling, is way more complex to solve than having the model create a patient schedule for a predetermined OR-block schedule.

Third, increasing the time horizon the model has to create a schedule for exponentially increases the complexity of the model. Fourth, when considering a time horizon of 4 weeks (which is the time horizon for which the model is intended to be used) running the model for 1 hour will generate good enough results. Running the model for an additional hour leads to only slight improvements. Finally, when an OR-block schedule is known, the decision variables of the group model may be relaxed to instantly get relatively good results. When time is not of the essence, it is advised not to relax the decision variables.

Scenario analysis

The fifth step of the research was the testing of different scenarios. For this research, a total of 6 different scenarios were tested. By testing different scenarios alternative scheduling policies and OR-block allocation policies could be evaluated. Because there are several ways the OR schedule can be created, a scenario analysis was a good method to use. Furthermore, there might be multiple optimal solutions since there are several KPIs that are used to evaluate the schedules. There is a possibility that one policy leads to slightly better OR usage whilst another policy performs better when looking at the bed occupation. By using a scenario analysis, it was possible to give insight in such cases as well. Different scenarios were implemented in the model by changing input parameters, decision variables, and the objective function.

The first scenario considered a scheduling policy where only the usage of the OR was taken into consideration and the OR-block schedule was predetermined. This scenario was used to see what the best OR use would be when no other resources needed to be considered. The second scenario considered a policy under which both the OR as well as the ward were optimized, for this scenario the OR-block schedule was predetermined too. Next, the third scenario optimized the OR usage whilst stabilizing the arrivals at the ward, meaning that the LOS was not considered for scenario 3. The fourth scenario optimised both the OR as well as the bed occupation at the ward. This scenario differs from scenario 2 in that only 1 big surgical ward is considered instead of 3 separate ones. For scenario 5 and scenario 6 the model had to define the OR-block schedule as well as the patient schedule. For scenario 5 the model could only allocate OR-blocks to the different sub-specialisms that were allocated to the general surgery specialism. For scenario 6, the model was only restricted by the physical constraints of the hospital. For both scenarios 5 and 6, both the usage of the OR as well as the usage of the wards where optimised.

Conclusions

Based on the results of the scenarios, several conclusions were drawn. First of all, considering the bed occupation when creating an OR-schedule is highly recommended. Based on the results of scenarios 1 and 2 it was concluded that considering the wards as well as the OR only slightly decreases the OR usage. The second conclusion, based on the results of scenario 2 and scenario 4, is that it is important to take the LOS of patients into account for wards that are occupied by patients with a longer LOS. For the day-care ward and the short stay ward, stabilizing the arrivals at the wards without taking the LOS into consideration led to good results in only a short runtime. Doing the same thing for ward 1BC led to an improvement compared to the current situation, but more can be gained when the LOS is taken into consideration too. The next conclusion is that when all patients go to one big surgical ward, all of the findings stated above hold. This means that when only one ward is considered, the bed occupation should still be considered when creating an OR-schedule. Furthermore, when considering one big ward, it is hard to keep the bed occupation high during the weekend. It is recommended to set the target occupation higher in the workdays and lower in the weekends.

The conclusion regards the OR-block scheduling policy. It was found that the absolute best schedules in terms of expected KPI scores are created when the model gets complete freedom when allocating OR-blocks to sub-specialisms. However, these schedules perform only slightly better than the schedules created under the currently used, predetermined OR-block schedule. Because it would be disruptive to completely change the way the OR-block schedules are created and because this would only lead to a very small possible improvement, it is not recommended to change the allocation of OR-blocks amongst sub-specialisms.

The main research question: “How can statistical information on the provided datasets and the use of a MIQP model help the Elkerliek when developing a tactical surgical schedule for a selected department taking both the OR and the downstream resources at the wards into account?” can be answered as follows: By retrieving statistical information on the provided datasets, it can be determined how the mix of patients that are operated on each day should look like. By analysing the data and by using the MIQP model created for this research, a schedule will be generated in which it is defined what the best patient mix on each day would be. Finally, it can be found what the expected results are of certain schedules so that the choice can be made to actually implement the schedule or to adapt it before implementing it.

Abbreviations

OR: Operating room

ED: Emergency department

LOS: Length of stay

MSS: Master surgery schedule

LP: Linear programming

QP: Quadratic programming

KPI: Key performance indicator

MILP: Mixed integer linear programming

MIQP: Mixed integer quadratic programming

Table of content

1 Introduction

Currently, hospitals in the Netherlands are facing a challenge. Both demand for health care as well as the expenditures are steadily increasing while the government funding remains the same. This means that an increasing amount of patients must be treated with the current capacity of resources. In addition to that, there is a big shortage of health care professionals in the Netherlands. In the upcoming years, about 1100 operating assistants, 550 ED nurses, 800 intensive care nurses, and 500 child nurses need to be educated to cope with this shortage (Ligtvoet, 2019). In this healthcare environment, it is necessary to use the available resources as effectively and efficiently as possible. The focus of this research is the optimization of the usage of the operating rooms (OR) and the beds at the ward by creating a tactical OR schedule that takes both the OR and the downstream resource 'bed capacity' into account. This research will be conducted within the Elkerliek hospital located in Helmond.

1.1 The Elkerliek

The Elkerliek hospital has three different locations. First of all, there is the main hospital that is located in Helmond. At this location, long-term and intensive care patients are taken care of in addition to day-care patients. Furthermore, there is an operating department which consists of eight ORs. Finally, there is an emergency department (ED) and a first heart help department. The second location is situated in Deurne. At this location, patients are seen for eye surgery. These patients are all day-care patients, meaning that they do not stay the night. Finally, there is a small outpatient clinic located in Gemert. At this location, only small interventions are carried out. The focus of this research will be on the hospital located in Helmond. The name 'Elkerliek' will, therefore, refer to the hospital in Helmond from this point on unless it is stated otherwise.

To create a basic understanding of the patient flow at the Elkerliek, figure 1 was added. A clear distinction is made between the elective patient flow (green) and the acute patient flow (orange) because there is a significant difference in the way these two types of patients are treated as well as the way the treatment of these patients is planned. As can be seen in figure 1, there are multiple ways a patient can enter the ward and the operating theatre. The main difference between the acute patients and the elective patients is that elective patients can be accounted for a long time in advance whereas acute patients must be seen within 72 hours. On top of that, some acute patients have to be operated immediately upon arrival. This is one of many aspects that makes the development of OR schedules difficult.

The process an elective patient normally follows within the Elkerliek starts with a scheduled visit to the outpatient department. When the patient is cured after visiting the outpatient department a certain amount of times, he will be discharged from the hospital. In the case that an elective patient has to undergo surgery, the surgery of the patients is scheduled by the planners. On the day of surgery, the patient will be placed in the ward where some preparations will take place. After that, the patient is operated on in an OR. Finally, the patient will be placed in a bed on one of the wards, depending on the length of stay (LOS) of the patient. Finally, the patient is discharged from the hospital.

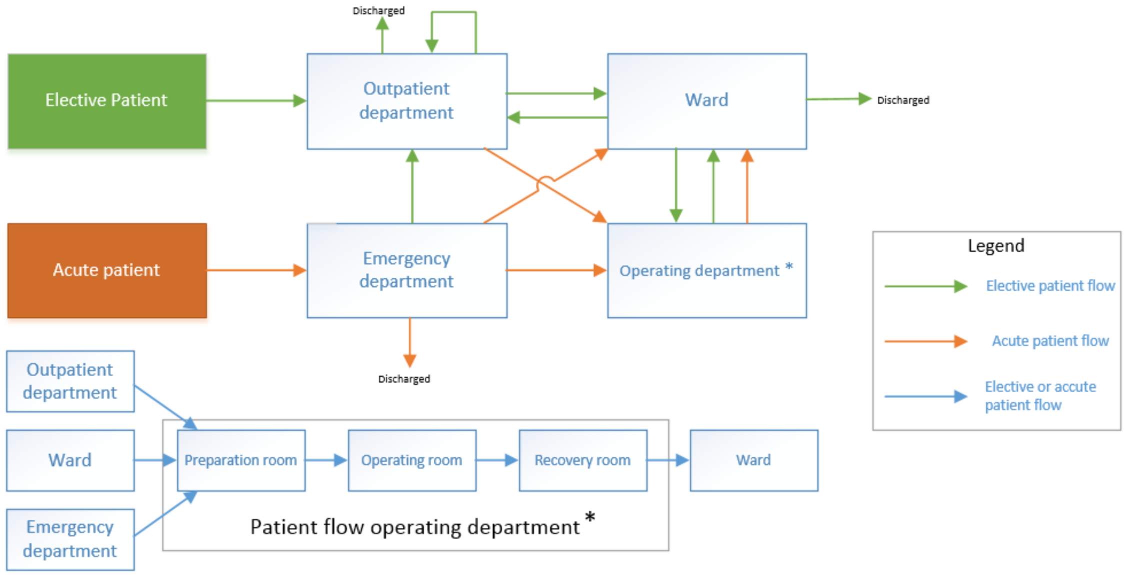

Figure 1: Patient flow at the Elkerliek

Acute patients can follow different routes through the hospital. For starters, acute patients can enter the hospital at the ED. Here, one of the following things can happen. Firstly, they can be discharged after being treated. Secondly, they can be sent home to be seen at the outpatient department at a later time where they will further be treated as an elective patient. Thirdly, the patient can be placed in a bed at a ward where they will be either be treated or where they wait to undergo surgery. Finally, in some extreme cases, the patient must be operated immediately. In these cases, the patient will undergo surgery as soon as possible and when necessary, the surgery of other patients will be cancelled. A final way for an acute patient to enter the system is through the outpatient clinic. It can be possible for an elective patient to visit the outpatient clinic where it becomes clear that the patient should be treated immediately.

1.2 The OR planning process

There are different steps that are taken at various moments on the planning horizon to come to a final operational planning for the ORs. The process of the OR scheduling at the Elkerliek is best explained according to the planning and control levels defined by Vissers and Beech (2005). In figure 2, the five different planning and control levels according to Vissers and Beech are shown. These five levels explain how decisions made at a higher level define the decision space of the lower levels. In the figure, it can be seen that the higher levels impose restrictions on the lower levels. Furthermore, there is a feed-forward and feed-backward connection between the different levels.

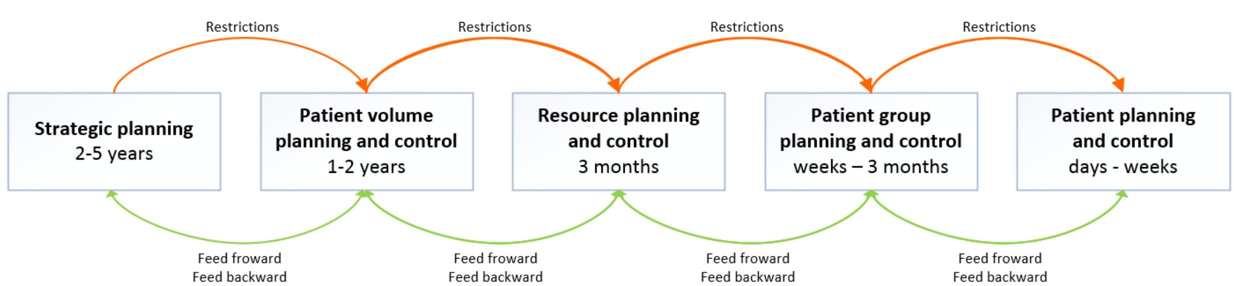

Figure 2: The five planning and control levels according to Vissers and Beech (2005)

The highest planning and control level is ‘strategic planning’. The time horizon on which the decisions on this level are made is several years. On this level, it is determined what range of services is offered by the Elkerliek, what the long-term resource requirements of the hospital are, and what region the hospital is responsible for, among other things. Decisions made by the Elkerliek on this level are, for example, the decision to move all eye surgeries to Deurne and the decision to focus a lot of resources on the practices surrounding abdomen surgeries.

The next planning and control level is ‘patient volume planning and control’. This level has a time horizon of approximately 1 year. Based on the strategic planning, the annual capacity for each specialism is determined, and with that, how many resources the specialism needs. This means that it has to be predicted what kind of care will be requested from the hospital. These predictions are made by the Elkerliek based on historical data of the demand of previous years, the seasonal patterns, and the appointments made with the insurance companies. Furthermore, there is a minimum amount of surgeries of certain types that need to be performed so that the specialists stay competent to perform said surgeries.

The third planning and control level is ‘resource planning and control’ which has a time horizon of several months. On this level, it is determined how the capacity of each specialism (which was determined on the previous level) is allocated over time. Here, the resources are divided among specialisms too. For the OR this means that a master surgery schedule (MSS) is created. In this MSS it is determined how the OR-blocks are divided among the specialisms over the week. OR-blocks are 4 hour time periods in which a surgeon can operate patients. In table 1 an example of a MSS for a hospital with 4 different specialisms and 3 ORs is shown. As table 1 presents, each day there is at least one OR-block reserved for emergency patients. For the Elkerliek, the MSS is revised every 3 months to adapt to the seasonal patterns and unexpected demand. When the MSS has been established, all specialisms allocate specialists to their assigned OR-blocks. The specialisms have a lot of freedom here to adhere to the personal preferences of the specialists. Since most of the specialists are not employed by the hospital, there might be a conflict in goals. One of the things that is investigated in this research is what the costs are of this possibly inefficient way of allocating specialists to OR-blocks.

Table 1: Example MSS of a hospital with 4 specialisms and 3 OR's

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Operating room 1	Spec 1	Spec 2	Spec 1	Spec 2	Spec 1
	Spec 1	Spec 2	Spec 1	Spec 2	Spec 1
Operating room 2	Spec 2	Spec 4	Spec 2	Spec 4	Spec 2
	Spec 3	Spec 4	Spec 3	Spec 4	Spec 3
Operating room 3	Spec 2	Spec 1	Spec 2	Spec 1	Spec 2
	Emergency	Emergency	Emergency	Emergency	Emergency

The fourth planning and control level is ‘patient group planning and control’. For this level, a time horizon of several weeks to months is used. On this level, service requirements and planning guidelines per patient group are determined. Within the Elkerliek, on this level, it is determined if any patient groups need to be prioritized over others due to an increasing waiting list. Furthermore, it is determined what targets the operational OR planners should strive for. Currently, the target is to fill the ORs as much as possible. On this level, it is also defined how the mix of patients that is operated on every day should look like. Currently the planners mix the day-care patients, short-stay patients, and the clinical patients in such a way that the bed limits of the wards are not surpassed. This research tries to come up with a patient mix for each day, based on the existing patient list, that does not only see the beds at the ward as a restriction. Instead, the wards will be treated as a downstream resource who’s use should be optimized too. The research will, therefore, be conducted on the third and fourth planning and control level.

The fifth and final level is ‘patient planning and control’. This level is also called the operational level and has a time horizon of several days to several weeks. On this level, the OR planners schedule the individual patients from the waiting list. They determine the room, the day, and the time the patients are operated on. Within the Elkerliek, the decision is made that patients have a say in the day they want to be operated too (determined on the strategic level). As time passes by, the planners fill the OR blocks more and more. This means that an OR-block that is still far away is likely to be empty whilst an OR-block that will take place tomorrow will most likely be already full of patients. It is the goal of the planners to fill the OR-blocks as much as possible whilst also taking the acute patients into account.

2 Problem statement

As stated in the introduction, it is increasingly important for the Elkerliek to use the available resources as efficiently as possible. Of those resources, the beds at the wards and the nurses working at the wards are some of the most critical and expensive. This means that the Elkerliek should align the availability of beds and nurses with the patient demand. In practice, this has proven to be very difficult due to the highly fluctuating number of patients that are staying at the wards over time. To deal with these fluctuations, extra resources are made available to cope with peaks in demand. Because of unforeseen peaks and dips in patient arrivals, situations occur where there are too few or too many beds and nurses available. Besides the costs of unused capacity or the cost of hiring extra capacity, the nurses experience high work pressure in cases of under capacity. To improve upon the current situation, the fluctuation of demand of the wards needs to be reduced and the peaks and dips in demand should be removed so that the resource use can more easily be aligned with the demand.

The output of the OR is directly related to the demand of the wards because all patients that undergo surgery are afterwards treated at one of the wards. Therefore, the output of the OR is partly responsible for the fluctuation in demand for the wards. A fraction of the fluctuation in demand is inevitable, due to the treatment of acute patients that cannot be scheduled far in advance. However, some of the fluctuations in demand come from the elective patient flow. This is the part of the fluctuation on which this research has been conducted. At the moment, the Elkerliek focuses on the usage of the OR when creating an OR schedule. While creating this schedule, the downstream resources (beds and nurses among others) are seen more as a restriction than as resources that should also be used as efficiently as possible. The goal of this research is to create a tactical OR schedule that takes both the OR as well as the bed occupation at the different wards into account.

The decision was made to focus this research on the general surgery specialism. This specialism was chosen because it makes the most use of the OR and the wards and, due to a wide variety in treatments, general surgery is the most difficult specialism to create a schedule for. With this knowledge, the following tactical scheduling problem has been defined:

> *How can a planning tool be developed for the general surgery department that minimizes current high fluctuations in resource requirements like beds and nurses at several wards while keeping the efficiency in OR occupancy at the desired level?*

As stated in the problem definition, the resources that have been taken into account are the OR time, the clinical beds, the short stay beds, and the day-care beds. The scope of the problem was restricted to the general surgery specialism, which can be divided into three sub-specialisms; trauma, vascular, and remaining.

The time horizon of the optimization problem is four weeks. This time horizon has been chosen because the amount and the mix of patients that arrive at the Elkerliek differs from month to month. In the summer months, the demand is lower than in the winter months for example. In addition, elective patients are often scheduled definitively several weeks in advance. Taking a time horizon of four weeks makes it possible to adhere to the seasonal influences and to come up with results that are useful for the planners.

The problem as described in this chapter has been solved via quadratic programming (QP) by formulating the aforementioned planning problem as a Mixed Integer Quadratic Programming (MIQP) model. A model has been created that puts out which patients should be treated on each day within the time horizon. The model compares possible solutions based on the difference between the predetermined target usage of each of the different resources and the expected resource usage. Furthermore, it is possible to use the currently applied allocation of OR blocks to the sub-specialism as an input for the model or to let the model choose what will be the best allocation of OR-blocks among sub-specialisms. In chapter 8, it is explicitly explained how the model works exactly and what the underlying mathematical formulation is.

3 Research questions

In chapter 2 the problem that this research focusses on is introduced. The goal of this research is to reduce the variability of the demand for the wards whilst keeping the performance of the OR high by revaluating the tactical surgical schedule. This leads to the following main research question:

How can statistical information on the provided datasets and the use of a MIQP model help the Elkerliek when developing a tactical surgical schedule for a selected department taking both the OR and the downstream resources at the wards into account?

Next to the main research question, several sub-questions are introduced below. These sub-questions will provide a path of sub-solutions towards the complete answer of the main research question. For each sub-question, the objectives are given in table 2. The sub-questions are defined as follows:

1. What is the current practice of scheduling the OR and what is the corresponding performance?
2. What are constraints and properties of the tactical surgical schedule that must be satisfied?
3. What are the proposed approaches in the literature for comparable scheduling problems?
4. How can the scheduling problem be represented in a mixed integer programming model?
5. How can the patients of the Elkerliek be grouped based on their resource usage?
6. How should the tactical surgical schedule look like when it is optimized for the OR only?
7. How should the tactical surgical schedule look like when it is optimized for both the OR and the downstream resources at the ward simultaneously?
8. What effect do different OR block allocation methods have on the best possible resource usage?

Table 2: Overview of the objectives for the sub-questions

Sub-question	Objectives
1	- Define key performance indicators with which the different scheduling methods can be evaluated. - Get initial performance indicators with which the new scheduling methods can be compared.
2	- Define the boundaries of the solution space. - Create a greater understanding of the scheduling problem.
3	- Create a greater understanding of how scheduling problems are solved - Create a foundation on which this research can continue
4	- Define the mixed integer quadratic programming model. - Create a greater understanding of the scheduling problem.
5	- Define how the patients of the Elkerliek can be grouped. - Define what the characteristics of each of the patient groups are. - Create groups that can be used as input for the programming model
6	- Find a (near) optimal tactical surgical schedule under the current scheduling policy. - Get an initial solution to compare the other methods to.
7	- Find a (near) optimal tactical surgical schedule where both the OR as well as the downstream resources at the ward are considered. - Identify the effect the tactical surgical schedule has on the fluctuating ward demands.
	- Find a (near) optimal tactical surgical schedule where both the OR as well as the downstream resources at the ward are considered when the model allocates the OR blocks to the different sub-specialisms. - Identify the effect the self-allocating of specialists to OR blocks has on the quality of the surgical schedule.

4 Significance

In the current literature, many articles have been written about the planning and scheduling of operating rooms. However, the majority of these articles did not take the downstream resources into account. Examples of downstream resources from the operating rooms are, among others: beds at the recovery rooms, beds at the wards, nurses, and medicine. Most articles have focused on optimising the use of the ORs and the resources that are required to perform the surgeries. However, in an article by Liu et al. (2019) it was stated that since hospital care is often delivered in successive stages, it will be beneficial to use integrated scheduling taking different hospital units into account. To identify how this study will contribute, existing literature that focusses on integrated scheduling has been studied. In paragraph 4.1 it is discussed how this research contributes to the knowledge on the topic of integral OR scheduling. In paragraph 4.2 the practical significance of this research for the Elkerliek will be discussed.

4.1 Scientific significance

This research aims to contribute to the existing literature on integrated OR scheduling. In this research, it is investigated how a tactical surgical schedule that takes both the OR and the bed occupation at several different wards into account, should look like. This research does not only considers several wards but also considers several different specialisms that each has an own set of surgeons and patients. Furthermore, different ways of allocating the available OR-blocks to the specialisms are studied. Articles exist in which the allocation of OR-blocks to specialists are discussed. One of those articles is the article by Essen et al. (2013). Here the number of required beds was reduced by creating an improved schedule with regard to the OR-block assignment. How patients are allocated to these OR-blocks was not considered. Articles also exist on the allocation of patients to days or OR-blocks. In the article by Adan et al. (2008) the patient mix was optimised with the same goal as this thesis research: to have both the OR usage as the bed occupation as close to the target values as possible. However, in this article, the available surgery time was predetermined for each day. Furthermore, in the article by Adan et al. a thorax centre is considered where only cardiothoracic surgeries are performed. This means that no different specialisms or sub-specialisms were considered. By taking different (sub)specialisms, wards and OR-block allocation policies into account this thesis research might lead to new insights into how an OR-schedule can influence the bed occupation.

4.2 Practical significance

As is stated in chapter 2, the Elkerliek experiences difficulty in efficiently using the beds and nurses because of the high fluctuations in patient demand. This research is of practical value for the Elkerliek because it will inform the Elkerliek whether the fluctuation in resource requirements at the wards can be prevented by using advanced planning techniques. For the Elkerliek it is of value to gain information on what should be considered when a surgical schedule is created. For this reason, several scenarios that are relevant for the Elkerliek have been tested. Based on the results of these scenarios the Elkerliek can make an informed decision on what scheduling policy should be used. Furthermore, the model that has been created can be used by the Elkerliek to create tactical surgical schedules. These tactical schedules can be used as input for how to schedule patients on the operational level. Finally, the information on the effect the self-allocating of OR-blocks by sub-specialism is valuable for the Elkerliek too. Currently, there are discussions on how the surgeons should be allocated to the available OR-blocks. This research can help by giving a theoretical background to these discussions.

5 Methodology

In this chapter the different steps that were taken to answer the research questions are described. In figure 3. The different steps are shown in the order in which they were completed. In the subparagraphs 5.1 to 5.7 the different steps are described.

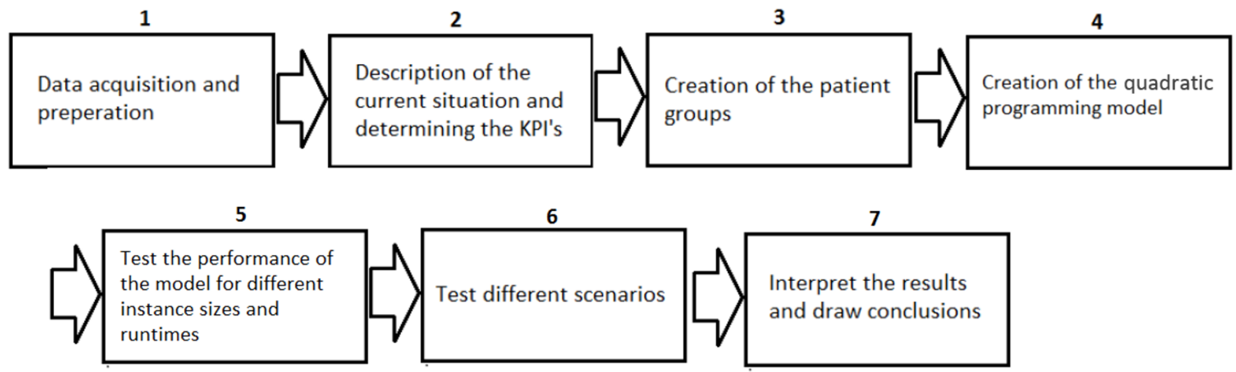

Figure 3: Different steps in order of completion

5.1 Data acquisition and preparation

The first step of the research was the acquisition of the required data. For this research, data had to be gathered on the surgeries and the admission of patients of the general surgery specialism. In chapter 6, a description of the data is given. The data was divided into two different data sets. One of the sets contained data on the surgeries and the other set contained data on the patient admissions. Before step 2 was performed these datasets were linked together based on the admission code that is unique for each admission. When the information of each patient admission was linked to the corresponding surgery it could be determined what resources each patient used and how long each patient used said resources. By using the 'patient enters OR' and 'patient leaves OR' rows, the exact time the patient has occupied the OR was determined. By using the 'admission time', 'discharge time', and 'ward location' rows, it was determined how long and on which ward each patient had occupied a bed. When this information was known, the next step was performed.

5.2 Description of the current situation

The second step that was performed was the description of the current situation. To evaluate the results that were generated in this research, it was necessary to know how the Elkerliek was performing before the research was conducted. The pre-research performance of the Elkerliek was set as a baseline with which the output of the models was compared. Also, information was found on the general surgery specialism so that informed decisions could be made with regard to the grouping of patients. Finally, key performance indicators (KPI) were defined so that solutions could be compared with each other. In chapter 7 the description of the current situation is discussed.

5.3 Creation of patient groups

The next step was the creation of patient groups. The different patient groups were defined by the patients' usage of the resources. This means that patients that make use of the same resources and use those resources a similar amount of time were grouped. The groups were created based on four criteria: the sub-specialism responsible for the patient, the ward the patient occupied a bed on, the surgery duration of the patient, and the LOS of the patient. To create patient groups, patients were first split based on what resources they have used. This means that for each sub-specialism/ward combination a set of patients was created. Then, the outliers in terms of surgery time and LOS were removed from each of these sets. Finally, with the help of a clustering tool, the patients of each sub-specialism/ward set were grouped based on the patients' surgery time and LOS. In chapter 8 the creation of patient groups is described in more detail.

After the creation of groups, a list of groups was created. For each of these groups the following information was known:

- What sub-specialism was responsible for the surgery of the patients in the group.
- What ward the patients in the group went to after surgery.
- The number of patients in the group.
- The expected surgery time for the patients in the group.
- The distribution of the LOS for the patients in the group.

This group list with the information as stated above was used as an input for the model.

5.4 Mixed integer quadratic programming model

To solve the tactical scheduling problem, a MIQP model was used. The scheduling problem was first defined as a MIQP problem that considered each patient individually. This individual model did not make use of patient groups and was used as a basis for the group MIQP model that makes use of patient groups instead of individual patients. The objective function of both MILQ models was to minimize the sum of under- and overutilization of the resources compared to the target utilization. This objective had to be achieved under certain restrictions like the maximum number of beds available on each ward. The decision variables for the model were:

- What patients are patients operated on each day (for the individual model)/ The number of patients of each group operated on each day (for the group model)
- The number of OR-blocks a sub-specialism can use on each day

In chapter 9 the workings of the MIQP models that were used are described in detail and the reasons for formulating the scheduling problem as a MIQP model are described.

As stated above, the individual model was created as a basis for the group model. Furthermore, it was used to validate whether the results of the model were realistic and gave a good representation of reality. Furthermore, the individual model was used to compare the runtime and the optimality gap of the group model with. Precision is lost when going from individual patients to patient groups because the exact surgery time and LOS of each patient are substituted by the expected surgery time and the LOS distribution of the patient group the patient belongs to. The group model is used because It was expected that considering each patient individually will greatly increase the runtime and the gap from the optimality. Below, other benefits of the group model are discussed.

The group model was used to generate the results for this research. This model made use of the patient groups created in the previous step and is on a higher tactical level than the individual model because no individual patients are considered. This model was used to generate results for several reasons. First of all, by having groups instead of individual patients the model consisted of fewer decision variables which requires less computational power. Second, using patient groups instead of individual patients makes it easier to match the operational level to the tactical level. An advantage of having groups instead of individual patients on the tactical level is that there is still some freedom on the operational level to choose what patients to schedule on what day. For example, patients have a say in what day their surgery takes place and cannot always be scheduled on a specific day. A patient belonging to patient group *x* can be scheduled on all days a patient of group *x* should be operated on according to the tactical model. When this patient is considered individually, the tactical model will only allow this patient to be scheduled on 1 specific day. When the patient does not agree to this day the operational planner has to stray from the tactical schedule.

5.5 Performance test of the model

After the model was defined and implemented, the performance of the model was tested. This performance was evaluated based on runtime and the gap towards optimality. Both the group model and the individual model were tested in the same way. This means that the performance of the two models could be compared to each other. This way it was possible to see what the effect was of considering groups instead of individual patients. In order to conduct the performance tests, different instances, and input parameters where used. To find out how the model behaves, the problem instance sizes were increased by expanding the time horizon considered. Furthermore, different runtimes were used to find solutions for the same instances in order to see how the quality of the solution improves as the model can run for a longer duration. Finally, instances, where the OR-block schedule was predetermined, were compared to instances where the OR-block schedule was defined by the model.

5.6 Test different scenarios

For the sixth step of this research, different scenarios were tested. By testing different scenarios alternative scheduling policies and OR-block allocation policies could be evaluated. Because there are several policies under which the OR schedule can be created, a scenario analysis was a good method to consider different methods. Furthermore, there might be multiple optimal solutions since there are several KPIs that are used to evaluate the schedules. There is a possibility that one policy leads to slightly better OR usage whilst another policy performs better when looking at the bed occupation. By using a scenario analysis, it was possible to give insight in such cases as well. Different scenarios were implemented in the model by changing input parameters, decision variables, and the objective function. These scenarios are discussed in chapter 11.

5.7 Analysing results

Finally, by evaluating the results of the scenario analysis, conclusions can be drawn. By studying the scenarios, insights were gained on what effect different scheduling policies have on the scores for the different KPIs. Not only the scores of the scenarios but also the schedules created were looked at. By looking at the schedules several conclusions about how the ideal patient mix should look like every day could be drawn. Finally, it could be seen how the OR-block allocation policy influences the quality of the surgical schedules for the Elkerliek. The research has been set up in such a way that it can be performed for other specialisms too. When the results of this research proof to be beneficial for the Elkerliek, other specialisms could be researched too, using the same method

6 Data description

The data that has been used for this research was originally divided into two different datasets. The first dataset contained all patient admissions at the wards from January 2016 to September 2019. The second dataset contained data about all surgeries performed by the general surgery specialism that have taken place from January 2018 until October 2019. All features of both datasets and the description of these features are given in appendix A. Before proceeding, these two datasets were merged using the admission number of the patient. By linking the datasets together, a dataset was created for which each row contains all available information of the surgery and admission of one patient. In this chapter, a basic description of this combined dataset is given.

The dataset used in this research consisted of 5180 rows with 56 features. Each row stored the information of 1 patient that has undergone surgery. The datasets stored only information about the surgeries that were performed by the general surgery specialism because this is the scope of the research. Both surgeries that have been performed in Helmond and Deurne are stored in the dataset because all surgeries that used to be performed in Deurne are now performed in Helmond. The dates the surgeries took place range from 1-1-2018 to 24-9-2019. Below, it is shown how the features were grouped and the important features are described.

The first group of features denotes the patient identification. This group consists of the patient-number, surgery-number, and the admission-number of a patient. The patient number is unique for each patient. The same number can occur multiple times in the dataset for patients that have undergone multiple surgeries. The surgery- and admission-number are unique for each surgery and admission. No duplicate surgery codes or admission codes can occur. The second group of features describes the patients. Features like the date of birth, gender, and how vulnerable the patient is, belong to this group. The third group of features stores information on who has treated the patient. Information on the specialism, the sub-specialism, the responsible surgeon, the amount of OR assistants involved in the surgery, and the surgeon responsible for the admission belong to this group. Fourth, there is the group of features that stores the location the surgery took place. The two features that belong to this group are the OR number and the location of the hospital. Closely related to this group is the fifth group. This group of features stores information on the patient's location in the ward. The ward, the room-number, and the bed-number are the features in this group.

Group six consists of the features on the waiting time of the patient. The features in this group are the date the surgery was requested, the date the surgery was scheduled to take place, and the date the surgery took place. The seventh group consists of a lot of features and stores the date and time that each step of the surgery took place. The different OR steps for which the start and end-time are stored are the following: the transport of the patient from ward to the OR department, the actual occupation of the OR, the actual surgery (the first incision until the moment the patient is stitched up), the recovery of the patient at the recovery room and finally, the transport of the patient from the OR department back to the ward. The eight group of features stores information on the expected duration of the surgery. This is split in the features expected cutting time, expected anaesthesia time, and expected sitting time. Group nine stores information on the LOS of the patient. This group consists of the date the patient is admitted to the ward, the date the patient is discharged, the warm bed minutes, and the actual nursing days. Finally, feature group 10 stores information on the reason the patient had to undergo surgery. The features belonging to this group are the diagnosis code, a description of the diagnosis, the treatment code, a description of the treatment, and additional treatment information.

7 The current situation

In this chapter, a description of the current situation in the Elkerliek is given. Firstly, a general overview of the general surgery specialism is presented in paragraph 7.1. Next, in paragraph 7.2 a quantitative analysis of the current situation is discussed. In this paragraph, KPI's are defined. Finally, the predictions for the OR time used by the Elkerliek are discussed in paragraph 7.4.

7.1 Overview of the general surgery specialism

As stated previously, this research focusses on the general surgery specialism. This subchapter is introduced to give some general information on this specialism. First of all, the patient demographic is discussed. Each year, about 3.000 patients are operated on by the general surgery specialism. About 80% of these patients are elective which means that 20% of the patients belong to the acute patient stream and are therefore not included in this research.

In figure 4, the average number of surgeries over the week is shown. Here it can be seen that most patients have been operated on Mondays. This because the Elkerliek hospital does not perform any elective surgeries on the weekend (which can also be seen in figure 4). To fill up the wards as fast as possible, most patients are operated on Monday. Over the other weekdays, the amount of arrivals seems quite stable.

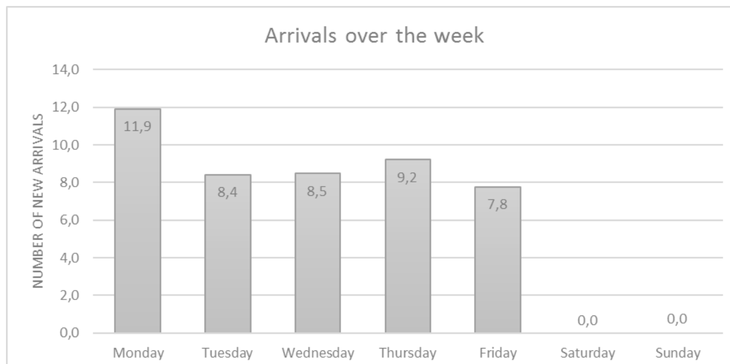

Figure 4: Surgeries over the week

In figure 5, the average number of surgeries over the year is shown. There are a few noteworthy months. First of all, the months July and August. These months are the months where the least patients are operated on. This can be explained by the summer holiday. The same holds for December and the Christmas holiday. Consequently, the months that occur directly after these vacation periods are slightly busier.

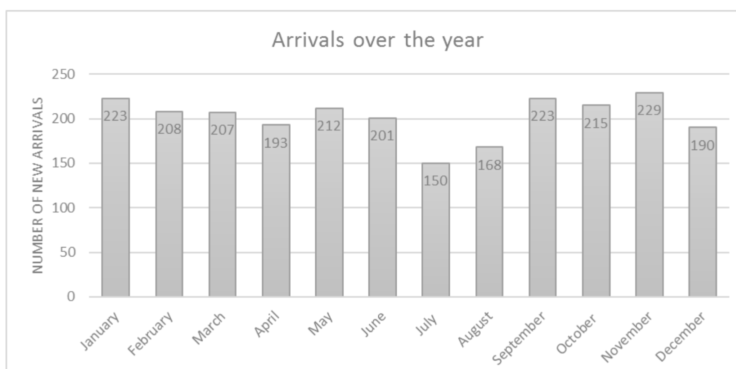

Figure 5: Surgeries over the year

The wards in the scope of the research were ward 1BC, 2BC, and 2D. These wards are very different from each other. Ward 1BC is the clinical ward. Here, patients stay for a longer amount of times. Often, patients who go here have undergone more impactful and bigger surgeries than patients who go to the other wards. Because patients can stay here for a long time, this ward is also used during the weekend. Ward 2BC is the day-care ward and the polar opposite of ward 1BC. At 2BC all patients leave on the day of arrival. Most patients that come here have undergone simpler surgeries that are not so impactful. Since no elective surgeries are performed at the weekend and because patients cannot spend the night, ward 2BC is closed on the weekend. Finally, there is ward 2D, the short stay ward. At this ward, which is physically connected with ward 2BC, patients can stay for a couple of days, but not longer than a week. This means that the average LOS is way shorter in this ward than on ward 1BC. This ward is open on Saturday but it is closed on Sunday. In table 3, the total ward capacity and the spread of elective and acute patients of the general surgery specialism on the 3 wards are shown. Note that the total bed capacity on these wards is shared by different specialisms.

Table 3: Ward capacity

Ward	Name ward	Bed capacity	Elective general surgery patients (%)	Acute general surgery patients (%)
1BC	Clinical	40	+- 58%	+- 42%
2BC	Day-care	40	+- 99%	+- 1 %
2D	Short-stay	23	+- 90%	+- 10%

The general surgery specialism is divided into 3 sub-specialisms, namely: vascular surgery, trauma surgery, and remaining surgery. There is a basic set of surgeries that all specialists can perform, regardless of the sub-specialism they belong to. Some surgeries have to be performed by a certain sub-specialism. Furthermore, most specialists prefer to operate on the patients that they themselves have requested the surgery for. When a patient is seen by a certain specialist in the outpatient clinic, it is very likely that this specialist will also perform the surgery. In the Elkerliek, there are 2 vascular surgeons. Next, 4 trauma surgeons belong to the general surgery specialism. There are also trauma surgeons that belong to the orthopaedics but they are not included in this research. Finally, 5 surgeons do not belong to either one of the previously mentioned sub-specialisms. This means that there are a total of 11 surgeons that will be considered in this research. In table 4 below it is shown how patients are divided among sub-specialisms and the different wards. The numbers in the table denote all elective patients that are represented in the data described in the previous chapter.

Table 4: Number of patients per sub-specialism and ward in the dataset described in chapter 6

Number of patients per sub-specialism and ward					
	1BC	**2BC**	**2D**	**Remaining**	**Total**
Vascular	378	228	58	137	**798**
Trauma	566	767	167	467	**1.964**
General	798	668	357	581	**2.401**
Total	**1.740**	**1.661**	**580**	**1.183**	**5.161**

The next table, shown below, shows the average surgery duration of the different sub-specialism/ward combinations. From table 5, it can be concluded that the sub-specialism that is responsible for the patient is clearly connected with the surgery time. Furthermore, when patients go to a ward where they have to stay for a longer period, the surgery duration is most likely longer too. For each combination of sub-specialism and ward, a histogram is made of the surgery times. These histograms are shown in appendix B.

Table 5: Average surgery duration per sub-specialism and ward

Average surgery duration in minutes					
	1BC	**2BC**	**2D**	**Remaining**	**Total average**
Vascular	140	60	50	75	**96**
Trauma	100	60	80	55	**65**
General	170	67	97	59	**97**
Total average	**148**	**63**	**88**	**59**	**86**

Finally, table 6 shows the average LOS in days of the different sub-specialisms/ward combinations. The results of this table are as expected. 1BC patients have the longest average LOS followed by 2D patients. The shortest LOS is reserved for 2BC patients that have an average LOS of exactly 1. The total average of trauma patients is the lowest, which can be explained because most trauma patients go to ward 2BC. The high average LOS of vascular patients can be explained in similar fashion: most vascular patients go to ward 1BC. In appendix C, histograms of the LOS for each combination of sub-specialism and ward 1BC and ward 2D are given. No histogram of ward 2BC is given because all patients stay exactly one day at this ward. No histogram of the remaining wards is given because these wards are not considered in this research.

Table 6: Average LOS per sub-specialism and ward

Average length of stay in days					
	1BC	**2BC**	**2D**	**Remaining**	**Total average**
Vascular	5,1	1,0	2,2	2,0	**3,0**
Trauma	4,2	1,0	2,1	1,2	**1,6**
General	6,2	1,0	2,0	1,5	**2,6**
Total average	**5,5**	**1,0**	**2,0**	**1,4**	**2,3**

7.2 Quantitative analysis

In order to evaluate the quality of an OR schedule, KPI's were determined. The two entities that are within the scope of the research are the OR and the ward. In this section, the KPI's that were used in this research to evaluate the current situation are introduced. These KPI's were also used to compare the current situation with the newly created schedules by the model.

7.2.1 KPI 1: Efficient use of OR time

The first KPI measures how much of the available OR time is actually used for surgery. Having available OR time left unused is expensive, especially in a time where there is a shortage of OR capacity. This first KPI is calculated by taking the percentage of the available OR time in an OR block (4 hours) that is used for operating on patients. The time between the start of the OR-block and the moment the first patient enters the room, the time the OR is empty between two surgeries and the time between the end of the OR block and the moment the last patient leaves the room are subtracted from the total OR-block time. To clarify, figure 6 is added. Technically, the maximum score for this KPI is 100%. However, it is not realistic to expect that this 100% will actually be achieved. Between surgeries, there is always some time lost because the room needs to be cleaned and preparations have to be made. On average, these switching times between surgeries take a total of 30 minutes per OR-block. Because these switching times cannot be removed, a score of 87% will be close to the actual maximum score for this KPI. To incorporate this into the MIQP models discussed in chapter 9, the target use of the OR-blocks is set to 210 minutes (which means that 30 minutes of the available 240 minutes will be reserved for switching times). Overtime will not be considered for this KPI because this is OR time that is used outside of the OR block.

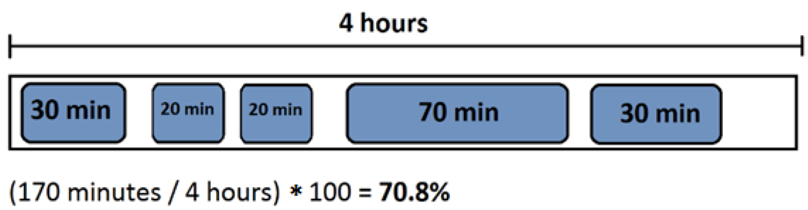

Figure 6: Example of KPI 1

In order to calculate the performance of the Elkerliek over the past two years, all regular OR blocks were considered. With regular OR blocks, all OR blocks scheduled between 8 AM and 5 PM on workdays in which elective patients were operated on are meant. It is still possible that within a normal OR-block, acute patients are operated. Emergency OR-blocks will not be considered for KPI 1. Emergency OR- blocks are blocks in which no patients are scheduled. These blocks are left empty to make sure that acute patients that need surgery right away can be treated without the cancelation of elective patient surgeries. Because these OR-blocks cannot be filled in advance and are used less efficiently than regular OR blocks, they distort the results for this KPI.

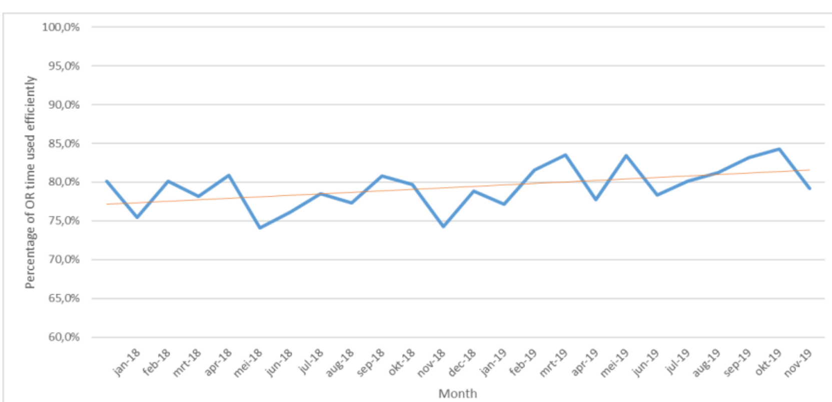

Figure 7: KPI 1 over time

In table 7, the scores for January 2018 to November 2019 and the total average score are shown. There is also a distinction made between the efficient use of the OR time with and without the inclusion of acute patients. The OR time that is used efficiently has been slowly increasing in the past two years, as can be seen in table 7 and figure 7.

Table 7: KPI 1, Percentage of OR time used for surgery

Year and month		Usage with acute surgery	Usage without acute surgery
2018	January	81,3%	76,1%
	February	79,1%	72,2%
	March	79,1%	75,9%
	April	77,7%	74,8%
	May	79,4%	76,1%
	June	78,1%	70,7%
	July	80,3%	72,9%
	August	82,0%	73,8%
	September	77,7%	73,5%
	October	82,5%	77,0%
	November	80,1%	74,6%
	December	75,6%	70,6%
2018 Average		**79,4%**	**74,0%**
2019	January	80,8%	76,2%
	February	80,0%	74,8%
	March	82,0%	78,3%
	April	82,9%	79,4%
	May	83,0%	75,4%
	June	83,2%	80,4%
	July	86,5%	73,3%
	August	83,3%	77,9%
	September	81,7%	76,6%
	October	84,2%	79,6%
	November	83,4%	78,4%
2019 Average		**82,8%**	**77,3%**

7.2.2 KPI 2: Average amount of overtime

The next KPI covers the overtime. For KPI 1 it is measured how much time is wasted. For KPI 2 the opposite is measured. There are cases where an OR is still being used after the OR-block has officially ended. When this happens in the morning, the following afternoon session in that room has to be postponed until the room is empty. When this happens for in the afternoon, overtime has to be paid to the OR employees which is very costly. This KPI will be determined by taking the average overtime per OR-block. The scores for this KPI are shown in table 8.

Table 8: Average overtime in minutes

Year and month		Average overtime in minutes		
		Total	**Morning**	**Afternoon**
2018	January	20,8	24,2	16,9
	February	12,6	12,6	12,8
	March	23,6	26,1	20,0
	April	13,5	11,6	15,6
	May	22,8	23,1	22,3
	June	13,1	12,1	14,4
	July	11,4	3,6	19,7
	August	22,2	25,2	18,2
	September	20,9	20,5	21,4
	October	18,9	19,2	18,4
	November	20,0	15,4	25,3
	December	20,8	24,0	16,9
2018 Average		**18,5**	**18,5**	**18,7**
2019	January	13,5	16,1	10,3
	February	10,6	10,6	10,7
	March	11,9	8,2	16,1
	April	11,3	5,8	18,2
	May	14,7	15,1	14,5
	June	15,9	16,0	15,8
	July	23,3	25,1	20,9
	August	11,2	12,1	10,2
	September	17,0	11,0	23,9
	October	12,2	9,7	14,8
	November	17,7	12,4	23,3
2019 Average		**14,1**	**12,8**	**15,8**

7.2.3 KPI 3: The number of extreme deviations in demand

The previous KPI’s focus on the OR. Since the ward is also within the scope of the research, KPI’s are defined for this entity too. The goal of the research is to reduce the peaks of demand in the ward and to make the required capacity more predictable. Two KPI’s were used to evaluate the quality of an OR schedule with regard to the ward. The first KPI for the ward and the third KPI overall is the number of extreme deviations in demand. An extreme deviation is defined as a deviation of 4 beds from the desired bed occupation. The number 4 has been chosen because there is a norm in the hospital that every 4 patients should be treated by 1 nurse. When 4 beds are occupied over the target, this means that a full extra nurse had to be scheduled which can be very costly when it is done at short time notice. Moreover, when the occupation is 4 beds lower than desired, this means that there is an extra nurse used that was not needed. In table 9 the overview of the extreme deviations is shown. For each ward, the percentage of days that an extreme deviation occurred is shown. Furthermore, the percentage of days that an extreme positive deviation (overuse) occurred and the percentage of days that an extreme negative deviation (underuse) occurred is shown.

Table 9: Extreme deviations in demand for the different wards

Extreme deviations in demand

Year and date		1BC Percentage of days with an extreme deviation from the target occupation			2BC Percentage of days with an extreme deviation from the target occupation			2D Percentage of days with an extreme deviation from the target occupation		
		Total	**Positive**	**Negative**	**Total**	**Positive**	**Negative**	**Total**	**Positive**	**Negative**
	January	26%	3%	13%	20%	10%	10%	6%	6%	0%
	February	28%	14%	14%	25%	14%	11%	8%	4%	4%
	March	32%	26%	6%	12%	6%	6%	9%	3%	6%
	April	20%	17%	3%	30%	10%	20%	6%	3%	3%
	May	6%	3%	3%	32%	16%	16%	9%	6%	3%
2018	June	6%	3%	3%	27%	20%	7%	0%	0%	0%
	July	26%	13%	3%	16%	3%	13%	6%	3%	3%
	August	10%	0%	10%	12%	6%	6%	3%	3%	0%
	September	43%	43%	0%	34%	17%	17%	0%	0%	0%
	October	45%	0%	45%	29%	13%	16%	0%	0%	0%
	November	27%	27%	0%	16%	13%	3%	3%	0%	3%
	December	0%	0%	0%	55%	29%	26%	19%	6%	13%
2018 Average		**20%**	**12%**	**8%**	**26%**	**13%**	**13%**	**6%**	**3%**	**3%**
	January	10%	0%	10%	20%	10%	10%	16%	6%	10%
	February	10%	0%	10%	7%	7%	0%	0%	0%	0%
	March	42%	19%	23%	16%	10%	6%	0%	0%	0%
	April	44%	17%	27%	43%	20%	23%	0%	0%	0%
2019	May	20%	10%	10%	33%	23%	10%	0%	0%	0%
	June	33%	23%	10%	36%	20%	13%	13%	13%	0%
	July	3%	0%	3%	16%	6%	10%	0%	0%	0%
	August	9%	6%	3%	16%	10%	6%	3%	3%	0%
	September	14%	7%	7%	7%	0%	7%	0%	0%	0%
2019 Average		**20%**	**9%**	**11%**	**21%**	**12%**	**9%**	**4%**	**3%**	**1%**

7.2.4 KPI 4: The average absolute deviation from the desired bed occupation

The second KPI for the ward and the fourth KPI overall is the average deviation from the desired bed occupation. The metric used for this KPI is the number of beds occupied over or under the desired bed occupation. When the desired number of patients that use a bed on a certain day is 10, an occupation of 8 (2 under the desired occupation) and an occupation of 12 (2 over the desired occupation) will both lead to a deviation of 2. This KPI is similar to the standard deviation. The main difference is that the standard deviation uses the mean as a central measure where the KPI uses the desired occupation.

Below in table 10, the average deviation per month is shown. The deviation is in the number of beds. For each of the three wards in the scope of this research, the absolute average deviation, the average over-usage, and the average underusage is given.

Table 10: Average deviation from the target occupation in number of beds

Average deviation from the target occupation in number of beds										
Year and date		1BC			2BC			2D		
		Total	Over	Under	Total	Over	Under	Total	Over	Under
2018	January	1,81	0,74	-1,06	1,71	0,97	-0,74	1,19	0,61	-0,58
	February	2,39	0,89	-1,50	1,68	0,82	-0,86	1,43	0,43	-1,00
	March	2,10	1,19	-0,90	1,65	0,77	-0,87	1,58	0,61	-0,97
	April	2,23	1,20	-1,03	2,30	0,93	-1,37	1,47	0,63	-0,83
	May	1,55	0,55	-1,00	2,13	1,13	-1,00	1,45	0,77	-0,68
	June	1,53	0,83	-0,70	2,10	1,17	-0,93	1,50	0,37	-1,13
	July	2,00	1,26	-0,74	1,71	0,84	-0,87	0,55	0,19	-0,35
	August	1,94	0,42	-1,52	1,42	0,61	-0,81	0,65	0,42	-0,23
	September	3,07	2,70	-0,37	2,23	0,93	-1,30	1,03	0,37	-0,67
	October	3,03	0,35	-2,68	2,06	0,94	-1,13	1,10	0,58	-0,52
	November	2,33	2,20	-0,13	1,53	1,00	-0,53	1,30	0,67	-0,63
	December	1,94	0,77	-1,16	3,10	1,65	-1,45	1,94	0,77	-1,16
2018 Average		**2,16**	**1,09**	**-1,07**	**1,97**	**0,98**	**-0,99**	**1,27**	**0,54**	**-0,73**
2019	January	1,84	0,55	-1,29	1,74	0,71	-1,03	1,74	0,65	-1,10
	February	1,84	0,55	-1,29	1,18	0,75	-0,43	1,07	0,61	-0,46
	March	3,48	1,97	-1,52	1,87	1,00	-0,87	1,06	0,48	-0,58
	April	2,97	1,27	-1,70	2,63	1,07	-1,57	1,10	0,50	-0,60
	May	2,10	0,90	-1,19	2,52	1,39	-1,13	1,29	0,52	-0,77
	June	2,67	1,67	-1,00	2,20	1,27	-0,93	1,67	0,80	-0,87
	July	1,74	0,39	-1,35	1,65	0,61	-1,03	0,19	0,13	-0,06
	August	1,74	1,16	-0,58	1,52	0,74	-0,77	0,68	0,48	-0,19
	September	2,00	0,73	-1,27	0,80	0,33	-0,47	1,20	0,47	-0,73
2019 Average		**2,26**	**1,02**	**-1,24**	**1,79**	**0,87**	**-0,91**	**1,11**	**0,51**	**-0,60**

In table 11 below, the maximum deviation per month is given. For each of the three wards in the scope of this research, the maximum overutilization and the maximum underutilization is given. The maximum deviation is an interesting metric because it tells the range between which the deviation from the target has fluctuated over the month. Take April 2018 for ward 2BC for example. On one day, the overutilization was 7 beds and, on another day, there was an underutilization of 10. This metric shows how much the demand for the wards can fluctuate.

Table 11: Maximum deviation in number of beds

The maximum deviation in number of beds							
Year and date		1BC		2BC		2D	
		Max over	Max under	Max over	Max under	Max over	Max under
2018	January	4	-6	7	-6	4	-3
	February	6	-6	6	-5	4	-4
	March	8	-4	6	-5	5	-4
	April	8	-4	7	-10	5	-4
	May	4	-4	8	-8	6	-4
	June	4	-4	8	-4	3	-3
	July	5	-5	4	-8	4	-4
	August	2	-6	8	-5	4	-3
	September	8	-3	5	-7	3	-3
	October	3	-6	7	-6	2	-3
	November	8	-1	5	-4	3	-4
	December	3	-3	9	-8	5	-4
2018 Average		**8**	**-6**	**8**	**-10**	**6**	**-4**
2019	January	3	-6	6	-6	5	-4
	February	3	-6	11	-3	3	-2
	March	7	-6	11	-6	2	-2
	April	6	-5	12	-12	3	-2
	May	4	-4	12	-7	3	-3
	June	6	-5	14	-12	5	-3
	July	3	-4	6	-4	1	-1
	August	5	-5	6	-8	4	-2
	September	4	-6	3	-5	3	-2
2019 Average		**7**	**-6**	**14**	**-12**	**5**	**-4**

7.3 The quality of the predicted OR time

To schedule patients in OR-blocks, a prediction of the time the patient will occupy the OR has to be made. To make a good schedule, it is important to predict this time as accurately as possible. When a surgery takes shorter in reality than predicted, the OR will not be used for a certain amount of time. When a surgery takes longer than in reality, this will lead to overtime and might even lead to cancelations of other surgeries. Currently, the prediction of the OR time is made automatically by HIX (the program used by the planners). When a surgery has to be scheduled, HIX looks at the last 10 surgeries of the same type and takes the average times of these surgeries. This average time is the predicted time. It is possible to override this time by the planners when a doctor states that he might need more or less time. This way of predicting the OR time is rather naïve and poor. It might be of importance to evaluate the predictions in order to possibly improve the quality of the schedules in the future. In previous research by T. Mul (2018) a method of improving the surgery time prediction was proposed. This research might be of value for the Elkerliek too.

In order to evaluate the quality of the predictions made for the general surgery specialism, two different measures are used. First of all, the coefficient of determination (R^2). This is a statistical measure of how well predictions approximate reality (Písař, 2019). R^2 is calculated according to the following formulas:

$$R^2 = 1 - \frac{SS_{residual}}{SS_{total}} \qquad SS_{total} = \sum(y_i - \bar{y})^2 \qquad SS_{residual} = \sum(y_i - \hat{y})^2$$

In the formulas, SS_{total} is calculated by summing the square of all realized OR times minus the average OR time. $SS_{residual}$ is calculated by summing the square of all realized OR times minus the predicted OR times. R^2 has a value between 0 and 1 where a score of 1 means that the prediction is completely correct. The predictions for the general surgery specialism used by the Elkerliek has an R^2 score of 0,59. Since the entirety of the OR-scheduling is based on the predictions for the OR time, this score is rather low.

The second measure that is used to evaluate the quality of the predictions is the root mean square error (RMSE). What this measure does is taking the absolute difference between all the predicted and realized OR times and taking the average of those absolute differences. A RMSE of 0 means that the prediction is completely congruent with reality, the higher the RMSE the more the predictions differ from the reality (Moody, 2019). The predictions for the general surgery specialism used by the Elkerliek has an RMSE of 19,82. This means that on average, a prediction differs 19,82 minutes from the realized time. This difference can be positive or negative. Given that the average surgery duration is 80,15 minutes, an average error of close to 20 minutes is significant.

8 The patient groups

In this chapter, the patient group generation is described. As stated in the problem description, the goal of the research was to find the ideal patient mix that should be operated on each day within the time horizon. Since this research was conducted on the tactical level, patients were not considered individually but they are categorised in different patient groups. The groups were created based on **what** resources each patient uses (the sub-specialism responsible for the patient and the ward the patient goes to) and **how much** each patient uses the recourses (the surgery duration of the patient and the LOS of the patient). In this chapter, the following topics are discussed: Firstly, the method of creating patient groups is discussed in paragraph 8.1. Next, the prepossessing of the data is explained in paragraph 8.2. Third, it is discussed how the number of groups has been chosen in paragraph 8.3. Finally, in paragraph 8.4, the patient groups that have been used for this research are described.

8.1 Method

The creation of groups occurred in two steps. Firstly, the patients were split based on what recourses they use. This leads to 12 intermediate patient groups as shown in table 12. In this table, the number of patients in the dataset of each sub-specialism/ward combination is shown. The division based on what resources patients use had been made because there is a clear distinction between what recourses patients use. It is not desirable that patients within one group use different recourses. Another reason why the patients were split based on what resources they use is the fact that patients of one sub-specialism/ward combination differ significantly from patients of other sub-specialism/ward combinations with regard to the expected surgery times and distributions of the LOS as is shown in chapter 7.

Table 12: Number of patients for each sub-specialism/ward combination in the data

Number of patients per sub-specialism and ward					
	1BC	**2BC**	**2D**	**Remaining wards**	**Total**
Vascular	378	228	58	137	**798**
Trauma	566	767	167	467	**1.964**
General	798	668	357	581	**2.401**
Total	**1.740**	**1.661**	**580**	**1.183**	**5.161**

The next step is to divide the subsets in the data, as described above, into the final patient groups. To create groups within the sub-specialism/ward subsets of the data, the surgery time and LOS are used. Based on these two features it was not possible to make a clear distinction between groups as in the first step. To tackle this problem a K-means clustering algorithm was used in this step as a structural grouping method. What clustering does, is grouping similar data points together. A simple example of how clustering can help in finding the patient groups within the data is shown in figure 8. Here, it can be seen that there are three distinct groups of patients when looking at the LOS and the surgery time. When using clustering, a higher quality will be achieved when three clusters are chosen as shown on the left then when 2 clusters are chosen. So, for this example, three groups should be chosen. For this simple example, the groups could have also been created by hand. However, for the real data sets, constructing the groups is more complex.

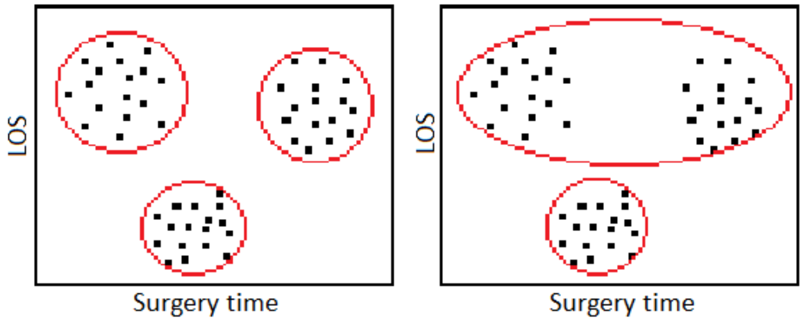

Figure 8: Clustering example

Before the grouping of patients took place, several data prepossessing steps should be taken to get the best results out of the clustering method. First of all, the outliers in the data were removed. This is discussed in paragraph 8.2 in detail. When the outliers were removed from the data, the clustering algorithm that was used to help identify the groups was chosen. There are various clustering algorithms, but for this research, K-means clustering has been used. The reasons are that K-means is very quick to implement and run. Furthermore, the K-means algorithm is fairly easy to interpret: it clusters points together that are closer to each other than to other clusters (Steinley, 2006). Finally, K-means was compared with other clustering methods (spectral and K-medoids) and it was found that K-means performed nearly as well or better than the other methods tested. The K-means clustering method operates in the following procedure (Steinley, 2006):

1. The number of clusters (K) should be defined.
2. K data points are randomly selected as cluster centres.
3. Assignment step.
4. Centroid update.
5. The centroids that are computed in step 4 will be the new cluster centres.
6. Finally, repeat steps 3 to 5 until convergence.

Finally, before the K-means clustering method could be used, the number of clusters should be defined. For this research, the Dunn-index is used to find the optimal number of clusters. The Dunn-index will be further explained in chapter 8.3.

8.2 Prepossessing of the data

As described in the previous chapter, some pre-processing of the data was needed. This means that the outliers needed to be removed from the data. Outliers were removed because they will have a negative influence on the group characteristics. The expected surgery times would have been higher when outliers were not removed. Furthermore, since the goal of the clustering was to create groups that can be used to create a tactical planning, the outliers are not important to take into consideration, because they occur only irregularly and mostly describe cases where unplanned complications occurred.

As described above the patients were separated based on sub-specialism, and on the ward they occupy after surgery. For each of these sets of patients, the first quartile (*Q1*), the third quartile (*Q3*), the interquartile range (*IQ*), the inner fences, and the outer fences were determined for the surgery times. The values for the fences of the surgery times are shown in table 13. The same metrics were determined for the LOS of all subsets containing 1BC and 2D patients. Patients from ward 2BC were not considered because all patients stay at this ward exactly one day and the remaining wards were not considered because they are out of the scope of this research.

The values of the fences of the LOS are shown in table 14. In appendix D, the boxplots of the datasets are shown. The lower inner fence and the upper inner fence are defined by $Q1 - 1.5 * IQ$ and $Q3 + 1.5 * IQ$ respectively. The lower- and the upper outer fence are defined by $Q1 - 3 * IQ$ and $Q3 + 3 * IQ$. A data point that is beyond either one of the inner fences but within the outer fences is defined as a mild outlier. A point that is beyond one of the outer fences is defined as an extreme outlier (U.S. Department of Commerce, 2013).

Table 13: Outlier metrics surgery times

1BC Patients		2BC Patients		2D Patients		Remaining	
Surgery times Vascular specialism in hours							
Lower inner fence	-2,45	Lower inner fence	-0,16	Lower inner fence	-0,45	Lower inner fence	0,00
Upper inner fence	6,08	Upper inner fence	2,11	Upper inner fence	2,08	Upper inner fence	2,05
Lower outer fence	-5,65	Lower outer fence	-1,01	Lower outer fence	-1,40	Lower outer fence	-0,77
Upper outer fence	9,28	Upper outer fence	2,96	Upper outer fence	3,03	Upper outer fence	2,82
Surgery times Trauma specialism in hours							
Lower inner fence	-0,48	Lower inner fence	-0,28	Lower inner fence	-0,07	Lower inner fence	-0,20
Upper inner fence	3,06	Upper inner fence	2,20	Upper inner fence	2,43	Upper inner fence	2,07
Lower outer fence	-1,80	Lower outer fence	-1,21	Lower outer fence	-1,00	Lower outer fence	-1,05
Upper outer fence	4,38	Upper outer fence	3,13	Upper outer fence	3,37	Upper outer fence	2,92
Surgery times General sub-specialism in hours							
Lower inner fence	-1,66	Lower inner fence	-0,21	Lower inner fence	-0,53	Lower inner fence	-0,26
Upper inner fence	6,19	Upper inner fence	2,35	Upper inner fence	3,46	Upper inner fence	2,08
Lower outer fence	-4,60	Lower outer fence	-1,18	Lower outer fence	-2,02	Lower outer fence	-1,13
Upper outer fence	9,13	Upper outer fence	3,32	Upper outer fence	4,95	Upper outer fence	2,95

For the surgery time, it was decided that both the extreme, as well as the mild outliers should be removed. The reason for this is that most of the patients that are identified as an outlier based on the surgery time had some sort of complication during surgery. For these patients, the realized surgery time was much higher than the expected surgery time. Since the planners cannot know when complications will occur, and since there is already time reserved in each OR-block for complications or the arrival of acute patients, it would be counterintuitive to create patient groups that are influenced by complications.

Table 14: Outlier metrics LOS

1BC Patients		2D Patients	
LOS Vascular specialism in days			
Lower inner fence	-1,50	Lower inner fence	-0,50
Upper inner fence	10,50	Upper inner fence	3,50
Lower outer fence	-6,00	Lower outer fence	-2,00
Upper outer fence	15,00	Upper outer fence	5,00
LOS Trauma specialism in days			
Lower inner fence	-7,00	Lower inner fence	-0,50
Upper inner fence	17,00	Upper inner fence	3,50
Lower outer fence	-16,00	Lower outer fence	-2,00
Upper outer fence	26,00	Upper outer fence	5,00
LOS General sub-specialism in days			
Lower inner fence	-5,50	Lower inner fence	-0,50
Upper inner fence	14,50	Upper inner fence	3,50
Lower outer fence	-13,00	Lower outer fence	-2,00
Upper outer fence	22,00	Upper outer fence	5,00

For the LOS of patients, it was decided to remove both the extreme as well as the mild outliers. For the patients staying at 1BC, this choice was made because the group of patients that are considered to be an outlier based on their LOS consists of very unique patients that only occur very irregularly or patients that have a long LOS because of complications. For patients staying at 2D, it made sense to remove patients that have a LOS that is greater than 3 days because this ward is mostly meant for patients that stay at the Elkerliek for 1 or 2 nights.

In total, 220 patients were removed from the data for being an outlier. This means that the groups will be created based on about 96% of the available data. Of the 220 outliers, 146 were removed because of their surgery times, and 74 patients were removed from the data because of their LOS.

8.3 Determining the number of groups

To determine what number of groups is appropriate to represent the patients in each of the sub-specialism/ward subsets the Dunn-index was used. In this paragraph, this method of qualifying clusters will be explained and the results will be shown.

The Dunn-index is a metric to evaluate the quality of groups. The higher the Dunn-index value the higher the quality of the group. This method assumes two things when assessing the quality of groups. First of all, it is assumed that good groups are well-separated from each other. Secondly, it is assumed that good groups are compact. The Dunn-index is calculated by taking the minimum inter-group distance and dividing this by the maximum group diameter. This means that higher group distances and a smaller group diameters lead to better scores (Pakhira, 2004). There are several ways to determine the inter-group distance. For this research, it was chosen to define the inter-cluster distance as the Euclidian distance between group centers. The group size is defined by taking the maximum Euclidian distance between two data points within one group.

In tables 15 to 17, the results of using the Dunn-index on the datasets are shown. As can be seen in the tables below, the Dunn-index gives unambiguous and clear information on what number of groups is appropriate for the different datasets. In the tables, the number of groups that were chosen are shown with bold numbers. For patients of the Vascular sub-specialism it was chosen to categorize all 2D patients together into one patient group. This choice was made because of the low number of vascular patients that go to the short-stay ward and because the Vascular/2D patients are all very similar with regard to their surgery times and LOS. The same choice was made for the remaining patients of the vascular sub-specialism.

Table 15: Dunn-index results for the Vascular sub-specialism

Dunn-index results Vascular sub-specialism			
1BC Patients		2BC Patients	
# Groups	Dunn-index	# Groups	Dunn-index
2	0,34	2	0,55
3	0,46	**3**	**0,61**
4	**0,48**	4	0,60
5	0,30	5	0,54
6	0,29	6	0,53
7	0,31	7	0,48
8	0,39	8	0,61

In table 15 above it can be seen that for 1BC patients of the vascular sub-specialism 4 groups should be used to categorize these patients. For 2BC patients of this specialism, the number of groups should be 3. As described above, the 2D patients and the remaining patients of the vascular specialism were both one group of their own. This means that the total set of patients of the vascular sub-specialism was grouped in 9 different groups.

Table 16: Dunn-index results for the Trauma sub-specialism

Dunn-index results Trauma sub-specialism							
1BC Patients		2BC Patients		2D Patients		Remaining Patients	
# Groups	Dunn-index	# Groups	Dunn-index	# Groups	Dunn-index	# Groups	Dunn-index
2	0,36	**2**	**0,66**	**2**	**0,70**	2	0,52
3	0,43	3	0,55	3	0,70	**3**	**0,61**
4	**0,52**	4	0,59	4	0,59	4	0,57
5	0,32	5	0,57	5	0,69	5	0,53
6	0,40	6	0,56	6	0,47	6	0,35
7	0,32	7	0,46	7	0,37	7	0,49
8	0,40	8	0,53	8	0,45	8	0,36

In table 16 above the Dunn-index results of the trauma, sub-specialism are shown. It can be seen that for 1BC patients of this sub-specialism 4 groups should be created. For 2BC patients of this specialism, the number of groups should be 2. For 2D trauma patients, the number of groups should either be 2 or 3 since the Dunn-index score was the same for both. The number of groups that was chosen is 2 because when three groups would have been chosen, the number of patients in each group would have been too small. Next, the remaining patients should be divided into 3 different groups. This means that the total set of patients of the trauma sub-specialism were grouped in 11 different groups.

Table 17: Dunn-index results for the General sub-specialism

Dunn-index results General sub-specialism							
1BC Patients		**2BC Patients**		**2D Patients**		**Remaining Patients**	
# Groups	Dunn-index	# Groups	Dunn-index	# Groups	Dunn-index	# Groups	Dunn-index
2	0,42	2	0,64	**2**	**0,63**	**2**	**0,62**
3	0,41	**3**	**0,67**	3	0,62	3	0,59
4	**0,42**	4	0,47	4	0,46	4	0,55
5	0,36	5	0,60	5	0,42	5	0,48
6	0,28	6	0,47	6	0,63	6	0,53
7	0,42	7	0,53	7	0,48	7	0,54
8	0,32	8	0,42	8	0,47	8	0,53

Finally, in table 17 the results for the general sub-specialism are shown. The patients of this sub-specialism going to ward 1BC should be grouped in either 2, 4, or 7 groups. It was chosen for 4 groups because 2 groups would be too general and 7 groups mean that some groups would have few patients. A total of 3 groups was used to categorize the 2BC patients. Both the 2D patients and the remaining patients can be divided into 2 groups. This led to a total of 11 groups for the general sub-specialism. The total number of groups for the entire general surgery specialism was 31. In the following paragraph, the groups are described and the values for the expected surgery time and LOS distribution of each group are given.

8.4 Final patient groups

After all the steps that are described in this chapter were completed, the actual grouping could take place. The created patient groups are shown in table 18 below. In this table, the group number, the average surgery time, the average LOS, and the number of patients of each group in the dataset are depicted. As can be seen, there were 31 different patient groups considered in this research. For the patients of groups 9, 19, 20, 30, and 31 no average LOS is shown. The reason for this is that these patients do not go to ward 1BC, 2BC, or 2D. The other wards are not within the scope of this research so their LOS had no significance. However, their surgery time was of importance because the patients of these groups still needed to undergo surgery and thus affected the usage of the OR.

Table 18: Final patient groups overview

	Ward	Group number.	Average Surgery time (minutes)	Average LOS (days)	# patients
Vascular sub-specialism	Ward 1BC	1	47	3	109
		2	55	7	33
		3	158	3	136
		4	224	6	52
	Ward 2BC	5	34	1	79
		6	64	1	100
		7	93	1	43
	Ward 2D	8	50	2	53
	Remaining wards	9	60	-	122
Trauma sub-specialism	Ward 1BC	10	53	3	121
		11	60	8	52
		12	120	4	221
		13	120	12	108
	Ward 2BC	14	41	1	463
		15	85	1	290
	Ward 2D	16	54	2	86
		17	100	2	55
	Remaining ward	18	29	-	90
		19	58	-	137
		20	89	-	214
General sub-specialism	Ward 1BC	21	70	2	115
		22	142	5	239
		23	146	10	274
		24	269	7	85
	Ward 2BC	25	42	1	307
		26	74	1	243
		27	110	1	108
	Ward 2D	28	60	2	183
		29	124	2	152
	Remaining wards	30	39	-	341
		31	79	-	204

Finally, for the patients that go to ward 1BC, the distribution of the LOS is shown in figure 9 and for the patients that go to ward 2D, the distribution of the LOS is shown in figure 10. In the figures, the probabilities of occupying a bed k days after surgery are shown for patients of each patient group. As can be seen, all patients have a probability of 100% of occupying a bed on the day of surgery (k=0). For patients of group 1, there is a probability of 92% of occupying a bed one day after surgery (k=1) which means that in 92% of the cases patients of group 1 have a LOS of 2 days. In appendix E two tables are added in which the distributions of the LOS are shown.

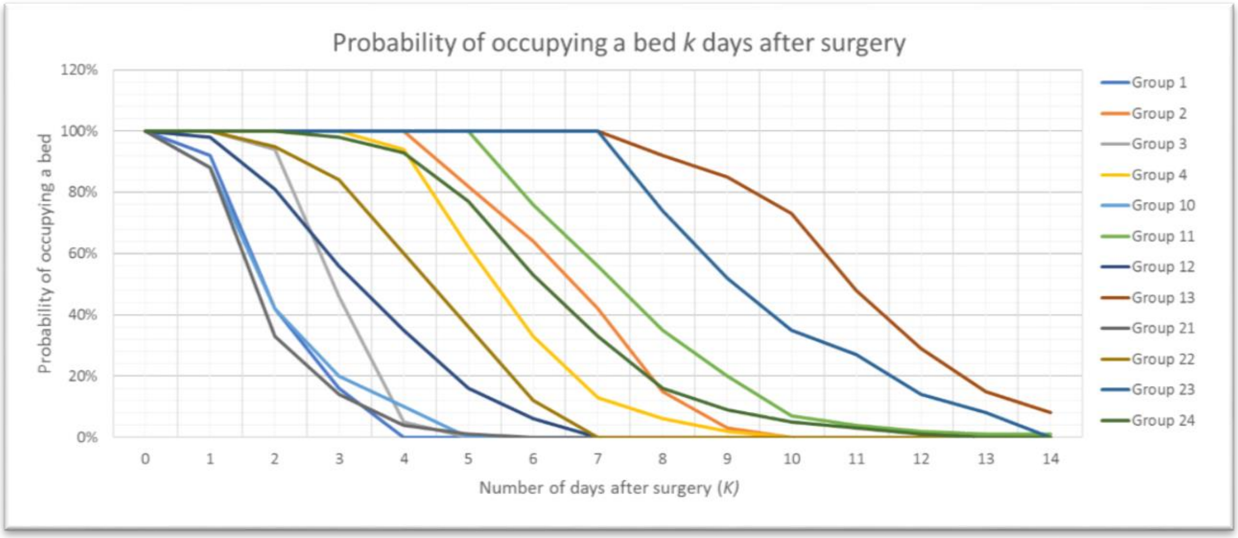

Figure 9: Probability of occupying a bed k days after surgery for 1BC patients

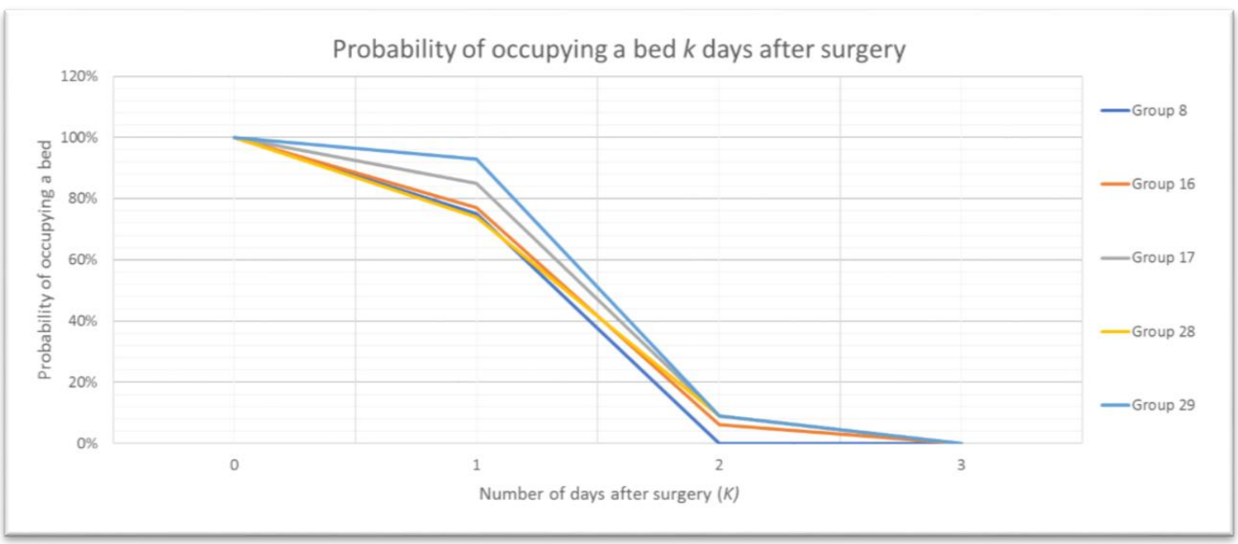

Figure 10: Probability of occupying a bed k days after surgery for 2D patients

9 Formulation and description of the individual MIQP model and the group MIQP model

In this chapter, the scheduling problem is formulated as a MIQP model. Two models are introduced in this chapter. One model considers each patient individually (the individual model) and one model uses patient groups (the group model). The individual model is created to schedule individual patients and to help make decisions on an operational level. This individual model is modified to help with decisions that are made on a longer-term. The individual model is modified such that patient profiles are considered instead of individual patients.

In this chapter, it will first be explained why the scheduling problem has been formulated as a MIQP model in paragraph 9.1. Next, the individual model is described first in paragraph 9.2. Firstly a short description of the individual model is given in this paragraph. Then, the entities that are present in the model are introduced in paragraph 9.2.1. Thirdly, the properties of the entities are translated to parameters in paragraph 9.2.2. Next, the decision variables, together with the other variables are discussed in paragraph 9.2.3. Finally, the objective function and the constraints of the individual model are introduced in paragraph 9.2.4. After that, the group model is described in paragraph 9.3. Firstly, a short description of the group model is given. Next, in paragraph 9.3.1, the differences between the individual- and group model are explicitly explained. After that, it is explained how the stochastics of the group model are dealt with in paragraph 9.3.2. Then the constraints of the group model are explained in paragraph 9.3.3. Finally, the verification and validation of the model is discussed in paragraph 9.4.

9.1 Reasons for formulating the scheduling problem as a MIQP model

Integer programming models are often used as planning/scheduling tools. This is because integer programming has many benefits. There is a big understanding of the power and scope of integer programming, reliable software is widely available, and complex problems can be solved among other things (Essays, 2018). Because of these reasons, and because mixed integer programming models are easily tuned so that solutions can be found for very specific cases (Beliën & Demeulemeester, 2004), the scheduling problem of this research has been translated into a mixed integer programming model. There are several types of mixed integer programming models. As stated before, this research uses a MIQP model. This means that the objective function is quadratic instead of linear, as is the case in the more commonly used mixed integer linear programming (MILP) models. In this paragraph, several reasons for choosing a quadratic objective function will be given.

9.1.1 Goal of the model according to stakeholders at the Elkerliek

First of all, when determining what the goals of the model should be, several interviews were conducted with the stakeholders at the Elkerliek. One of the things that came forward during these interviews, was that reducing the extreme deviations from the target occupations was more important than reducing the average deviation. It was stated by the stakeholders at the Elkerliek that the occurrence or multiple small deviations from the target occupation are less of an issue than the occurrence of a few big deviations. For example, having a bed occupation that only deviates 2 beds from the target is more easily dealt with than a deviation of 4 beds. A deviation of 2 beds over the target leads to a slightly higher workload for the nurses, but a deviation of 4 beds over the target occupation leads to a shortage of 1 full-time equivalent. Because the consequences of deviating from the target are not linear with the size of the deviation, it seems logical to use a quadratic objective function.

9.1.2 Quadratic programming in similar research

In the current literature, many articles have been written on OR scheduling. In the article "*A decision support system for cyclic master surgery scheduling*" by Beliën et al. (2008), a MIQP model is used to create an MSS that takes multiple objectives into account. In the article, one of the most important objectives of the model is to level the bed occupancy. This makes that the goal of the model created by Beliën et al. is very similar to the goal of the model used in this thesis research. In the article, several reasons for using an MIQP are offered. The most important reason for using a quadratic objective function is that the model explicitly tries to level the peaks as much as possible. Since reducing the peaks in bed occupation is one of the goals of this thesis research too, it seems a good choice to use a quadratic objective function instead of a linear one. Furthermore, in the article, the MIQP model outperformed several other mixed integer programming models in terms of both solution quality and computation time.

Another article by Bekker and Koeleman (2011) on the scheduling of patient admissions has as a goal to reduce the variability in bed demand. In this article, the choice has been made to use QP. The main reason for using a quadratic objective function is that the consequences of a deviation from the target bed occupation are not linear with the size of the deviation. In the article, it is stated that it is considerably more difficult for the personnel to deal with larger deviations than it is to deal with smaller deviations. This is the case for the Elkerliek too. Minimizing the average bed occupation is less important than reducing the peaks in demand.

9.1.3 Comparison of results found with a MILP model with results found with a MIQP model

Finally, to ensure that QP suits this research better than linear programming (LP), the MIQP models described in the next sections were adapted into MILP models. Since the constraints remain the same this was a fairly quick and easy step to make. The only thing that had to be done was substituting the quadratic objective function for a linear one. Both the MILP model and the MIQP model were used to generate some schedules. Comparing these schedules it was found that the average deviation from the target usage, for both the OR and the wards, was slightly lower under the schedules created by the MILP model. However, the number of bigger deviations was lower for the models created by the MIQP model.

In figure 11 the target utilization (black line) and the realized utilization (yellow line) of ward 2BC for the month of April are shown under a schedule created by the MILP model and under a schedule created by the MIQP model respectively. This figure is shown because in this figure it can be seen quite nicely how both models acquire different goals. This figure is representative of the achieved resource usages under both types of integer programming models. The schedule created by the MILP model for the month of April resulted in an average deviation of 0,5 beds per day while the schedule created by the MIQP model resulted in an average deviation of 0,55 beds per day. However, for the schedule created by the MILP model, 2 different deviations of 3 beds occurred while the schedule created by the MIQP resulted in a maximum deviation of 2 beds. As described above, the reduction of big deviations is more important than the reduction of the average deviation. Since the tests that were performed showed that the MIQP model does indeed lead to the biggest reduction in big deviations, it was chosen to use MIQP models for this research.

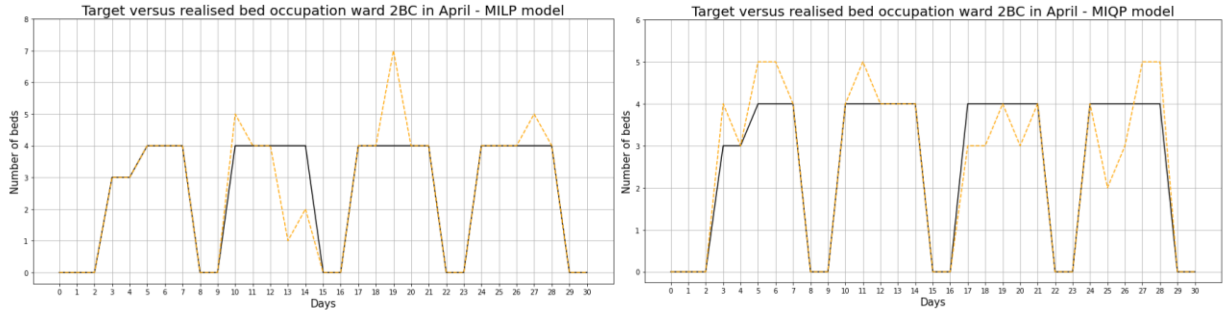

Figure 11: Target versus realized bed occupation for ward 2BC in April for a schedule created by the MILP model (left) and a schedule created by the MIQP model (right)

9.2 Description of the individual model

The individual model is an integrated scheduling model that simultaneously considers OR occupancy and ward bed occupation when scheduling patient surgeries. The problem instances are constructed from the information system of the Elkerliek hospital. In this research, the model that considers patients individually is called the 'individual model'. The individual model was built for several reasons. First of all, the model was used as a basis for the group model. Secondly, the individual model was used to compare the performance of the group model. By comparing the performance of these models it was possible to find the effect of considering groups instead of individual patients. Finally, the individual model was used to validate the results with the stakeholders at the Elkerliek.

The goal of the individual model is to find the best OR-schedule with regard to the usage of resources. What the model does is finding the patient allocations for which the resource utilization is as close to the predetermined target-utilization as possible whilst satisfying certain constraints. The output of the model is an allocation schedule in which it is specified for each patient on what day the patient's surgery takes place and how many OR blocks each sub-specialism gets on each day. An OR block is a 4 hour time period in which a surgeon is assigned to a room. Each 8 hour workday, one OR is split into two OR blocks of 4 hours each. The resources for which the utilization is optimized are the available OR time and the beds at the wards. The input for the model consists of all (elective) patients that have to be operated on within these 4 weeks amongst other things. Each individual patient has a few properties: the duration of the surgery, the sub-specialism that has to perform the surgery, the ward that the patient will stay at after surgery, and the length of stay of the patient.

The model used in this research is inspired by a model created by Adan et Al. in 2008 (Patient mix optimization and stochastic resource requirement: A case study in cardiothoracic surgery planning). The model introduced in this article is used as a basis for the one used in this thesis research. The model has been adapted so that different sub-specialisms can be considered.

Furthermore, the model used in this thesis research has a second decision variable that denotes how many OR-blocks each sub-specialism gets to use on each day. By introducing this decision variable and some extra constraints it is possible to have the model determine what the best OR-block allocation is.

9.2.1 The entities of the individual model

For the MIQP problem, three main entities are considered: The patient, the sub-specialism, and the ward. The set *P* which is indexed by *p*, represents the patients. The set C which is indexed by c, represents the sub-specialisms. Finally, the set *W* which is indexed by *w*, represents the wards. Each of these entities has several properties. These properties are discussed in the next paragraph. Furthermore, it is important to state that the time horizon of the model is *D* days where one day is denoted by *d.*

9.2.2 The parameters of the individual model

In this paragraph, the properties of the entities are described and the parameters of the model are introduced. Note that all parameters are denoted with capital letters.

Patient properties

Each patient has a set of properties. First, patients have a surgery-time which is denoted by T_p . This is the time the surgery of patient *p* takes in minutes. Second, patients have a LOS. This is translated in the model by $L_{p,w,k}$. This is a binary parameter that has value 1 when patient *p* occupies a bed on ward *w* after *k* days succeeding his operation where $k = 0,\dots,K$. Parameter $L_{p,w,k}$ has value 0 otherwise. The letter *k* denotes the number of days that have passed since surgery took place. So when $L_{1,1,5} = 1$, this means that patient 1 still occupies a bed on ward 1 after 5 days of succeeding the patient's surgery. Third, each patient can only be treated by surgeons of a certain sub-specialism. The parameter that denotes if patient *p* can be treated by a surgeon of sub-specialism *c* is $CP_{p,c}$. This binary parameter has value 0 when patient *p* cannot be treated by a surgeon of sub-specialism *c* and value 1 when patient *p* can be treated by a surgeon of sub-specialism *c*.

Sub-specialism properties

A sub-specialism has a set of properties too. First of all, surgeons of a certain sub-specialism can only treat some patients. This is denoted by the parameter $CP_{p,c}$ that is described above. Second, each sub-specialism has a certain capacity for each day. This is denoted by parameter $MC_{c,d}$, the maximum number of OR-blocks (4 hour time periods in which a surgeon performs surgeries) that can be occupied by sub-specialism *c* on day *d*. Since there is a limited number of OR blocks available each day for all specialisms together, the OR capacity of the hospital on a given day *d* is a parameter too. The parameter for this is MOR_d, the maximum number of available OR-blocks on day *d*. Finally, there is a maximum overutilization per OR-block which is denoted by *MT*. This maximum overutilization means that a sub-specialism can only use *MT* minutes of overtime for each OR-block said sub-specialism uses.

Ward properties

Next, the ward has some properties too. First of all, the ward has a maximum capacity. This is the maximum number of beds on ward *w* that can be occupied on day *d* which is denoted by $MW_{w,d}$. In addition to that, each ward *w* has a certain target occupation for each day *d* which is denoted by $TU_{w,d}$. This target occupation is the desired number of beds that should be occupied on day *d*. The desired occupation is known beforehand and can differ from day to day.

Model objectives

Finally, parameter $Weight_{or}$ and $Weight_{ward}$ denote the relative weight of the over- and underusage of the OR and the wards respectively. These weights are based on the costs of one OR hour compared to the costs of one occupied bed. The weights are needed to ensure that that the relative importance of the resources is taken into consideration.

9.2.3 The decision variables of the individual model

In this paragraph, the variables of the individual model are introduced. Note that all variables are denoted with lowercase letters. There are two decisions that the model has to make: The first decision is the day on which each individual patient is operated. The second decision is how many OR-blocks each sub-specialism uses on each day. The goal is to determine these decision variables whilst satisfying certain constraints and for which the resource use is as close to the target as possible.

As stated above, there are two decision variables. The first decision variable is $x_{p,d}$. This is a binary decision variable that has value 1 when patient *p* is operated on day *d* and has value 0 otherwise. The second decision variable is $y_{c,d}$. This is an integer decision variable that denotes the number of OR-blocks sub-specialism *c* uses on day *d*.

Depending on the values of the decision variables, the variables 'overutilization of the ward', 'underutilization of the wards', 'overutilization of the OR', and 'underutilization of the OR' will be affected. Overutilization is when a resource is used more than the predetermined target usage. Underutilisation is when a resource is used less than the predetermined target usage. Firstly, variables $ouw_{w,d}$ and $uuw_{w,d}$ are introduced to denote the over- and underutilization of ward *w* on day *d* in number of beds. Finally, the variables $ouc_{c,d}$ and $uuc_{c,d}$ are introduced to denote the over- and underutilization of OR-blocks used by sub-specialism *c* on day *d.* In table 19, all parameters and variables of the individual model are shown.

Table 19: Parameters and variables of the individual model

Parameter	Description
T_p	The planned time the surgery of patient *p* takes as found in HIX
$L_{p,w,k}$	$L_{p,w,k}$ = 1 when patient *p* is on ward *w* after *k* days after surgery $L_{p,w,k}$ = 0 otherwise
$CP_{p,c}$	$CP_{p,c}$ = 1 when patient *p* can be operated by a surgeon of sub-specialism *c* $CP_{p,c}$ = 0 otherwise
$MC_{c,d}$	The maximum number of OR blocks sub-specialism *c* can occupy on day *d*
MOR_d	Maximum number of OR blocks that can be used on day *d*
MT	Maximum overtime for every OR block
$MW_{w,d}$	The maximum number of beds on ward *w* that can be occupied on day *d*
$TU_{w,d}$	The target utilization (number of occupied beds) of ward *w* on day *d*
$Weight_{or}$	The relative weight of resource "OR"
$Weigth_{ward}$	The relative weight of resource "ward"

Variable	Description
$x_{p,d}$	$x_{p,c,d}$ = 1 when patient *p* is operated on day *d* $x_{p,c,d}$ = 0 otherwise
$y_{c,d}$	The number of OR blocks specialist *c* uses on day *d*
$ouc_{c,d}$	Overutilization of the OR by sub-specialism *c* on day *d*
$uuc_{c,d}$	Underutilization of the OR by sub-specialism *c* on day *d*
$ouw_{w,d}$	Overutilization of ward *w* on day *d*
$uuw_{w,d}$	Underutilization of ward *w* on day *d*

9.2.4 The objective function and constraints of the individual model

The objective function of the model is a minimization function. The function minimizes the sum of the over- and underusage of the OR for each sub-specialism and day, and the sum of the over- and underusage of the wards for each ward and each day. Both the sum of the over- and underusage of the OR time and the over- and underusage of the wards are squared. The reason for this is that it is more acceptable to have multiple days with a low deviation from the target values than to have one day with one large deviation from the target. For example, the model will choose an overuse of 1 bed for 3 days over an overuse of 3 beds for 1 day. The objective function looks as follows:

$$MIN: Weight_{or}\sum_{c=1}^{C}\sum_{d=1}^{D}(uuc_{c,d}+ouc_{c,d})^2 + Weight_{ward}\sum_{w=1}^{W}\sum_{d=1}^{D}(uuw_{w,d}+\ ouw_{w,d})^2$$

Now that the objective function and all parameters and variables have been introduced, the constraints of the model are introduced below. For each of the constraint, the mathematical formulation and a short description of why the constraint is used is given.

1) All patients that are put into the model should be assigned to exactly one day. To ensure this, the following equation was introduced:

$$\sum_{d=1}^{D} x_{p,d} = 1 \qquad\qquad p = 1, \dots, P$$

2) To denote the overutilization and the underutilization of the OR by sub-specialism *c*, variables $OUC_{c,d}$ and $UUC_{c,d}$ are introduced. The target utilization of the OR is calculated by taking the number of OR blocks sub-specialism *c* has on day *d* and multiplying this number with the desired number of minutes that should be used for elective surgeries. A full OR-block has 240 minutes and for each OR-block 30 minutes should be left unused for switching times and the possibility of the arrival of emergency patients. which leaves 210 minutes. Then we get for the utilization of the specialists:

$$210\ y_{c,d} + ouc_{c,d} \geq \sum_{p=1}^{P} T_p\ CP_{p,c}\ \ x_{p,d} \geq\ 210\ y_{c,d} - uuc_{c,d} \qquad c = 1, \dots, C, \quad d = 1, \dots, D$$

3) To determine the underutilization ($UUW_{w,d}$) and overutilization ($OUW_{w,d}$) of ward *w* on day *d,* the inequation below was introduced. To determine how many patients are on ward *w* on day *d*, *k* was introduced, where *k* = 0,…,*K* and *K* is the maximum LOS that has to be considered:

$$TU_{w,d} - uuw_{w,d} \leq \sum_{p=1}^{P}\sum_{k=0}^{K}\left(L_{p,w,k}\ x_{p,d-k}\right) \leq TU_{w,d} + ouw_{w,d} \qquad w = 1, \dots, W, \quad d = 1, \dots, D$$

4) In order to ensure that patient *p* can only be assigned to a day that a surgeon of sub-specialism *c* that can treat patient *p* is available, the following inequality was added:

$$x_{p,d}\ \leq CP_{p,c}\, y_{c,d} \qquad p = 1, \dots, P \quad c = 1, \dots, C, \quad d = 1, \dots, D$$

5) The number of OR blocks allocated to sub-specialism *c* on day *d* should always be smaller than or equal to the maximum number of OR block sub-specialism *c* can occupy on day *d*:

$$y_{c,d} \leq MC_{c,d} \qquad c = 1, \dots, C, \quad d = 1, \dots, D$$

6) To ensure that the maximum amount of overtime is not exceeded, the following inequality is introduced. The number of OR-blocks sub-specialism *c* gets on day *d* is multiplied with the maximum overtime per OR-block. This multiplication should always be greater than or equal to the overtime of sub-specialism *c* on day *d* :

$$MT\ y_{c,d} \geq ouc_{c,d} \qquad c = 1, \dots, C, \quad d = 1, \dots, D$$

7) To ensure that the total number of OR-blocks used on day d does not exceed the maximum number of OR-blocks available on day *d*, the following inequality was introduced:

$$\sum_{c=1}^{C} y_{c,d} \leq MOR_d \qquad d = 1, \dots, D$$

8) To ensure that the maximum number of beds isn't exceeded, the target utilization plus the over-utilization of ward *w* on day *d* should be smaller than the maximum number of available beds on ward *w* on day *d*:

$$TU_{w,d} + ouw_{w,d} \leq MW_{w,d} \qquad w = 1, \dots, W, \quad d = 1, \dots, D$$

9) The inequalities below were added to ensure that the underutilizations and the overutilizations are always greater than or equal to 0:

$$uuw_{w,d} \geq 0,\ ouw_{w,d} \geq 0,\ uuc_{c,d} \geq 0,\ ouc_{c,d} \geq 0,\ w = 1, \dots, W,\ c = 1, \dots, C,\ d = 1, \dots, D$$

9.3 Description of the group model

The main goal of this research is to propose a model that can help in making the planning decisions for a longer time horizon. These decisions fall between operational and tactical planning and control level. Therefore, the individual model is adapted to handle such cases by grouping patients to obtain patient profiles. This group model makes use of patient groups and is on a higher tactical level than the individual model because no individual patients are considered. This model was used to generate the results for this research for several reasons. First of all, by having groups instead of individual patients the model consisted of fewer decision variables which was expected to require less computational power. Second, using patient groups instead of individual patients makes it easier to match the operational level to the tactical level. An advantage of having groups instead of individual patients on the tactical level is that there is still some freedom on the operational level to choose what patients to schedule on what day.

The goal of the group model is to find the best OR schedule with regard to the usage of resources. The objective and constraints of the group model are similar to the individual model. The group model finds the mix of patient profiles for each day for which the resource utilization is as close to the predetermined target-utilization as possible whilst satisfying certain constraints. The output of the model is a schedule, in which it is specified for each day how many patients of each patient group should be operated on and how many OR blocks each sub-specialism gets. The resources for which the utilization is optimized are the available OR time of the sub-specialisms and the beds at the wards.

The input for the model consists of a list of all the patient groups that must be considered among other things. Each group has a few properties: the expected duration of the surgery, the sub-specialism that has to perform the surgery, the ward that the patients will stay at after surgery, and the LOS distribution of the patients. Furthermore, it is defined how many patients of each group should be operated within the time horizon the model makes a schedule for.

9.3.1 Differences between the group model and the individual model

As stated in the previous paragraph, the group model is similar to the individual model. However, there are some differences. The first significant difference has an impact on one of the entities: the patient. Since the patients are not considered individually anymore, the entity that will be considered for the group model will be the patient groups. For the group model, the set *G* which is indexed by *g* will represent the patient groups. This means that for all variables, parameters, constraints, and objectives where the entity 'patient' is used, this entity will be changed to patient groups.

The next difference between the model has to do with decision variable $x_{p,\mathrm{d}}$. For the individual model, this was a binary decision variable because each *p* denotes one individual patient that can only be scheduled once. For the group model this decision variable changes to $x_{g,d}$ which is an integer decision variable. This decision variable denotes the number of patients of group *g* that are operated on day *d*. Next, for the group model, an extra parameter NP_g is introduced. Parameter NP_p denotes the number of patients of group *g* that have to be scheduled within the time horizon. Furthermore, constraint 1 is different for the group model. This constraint is an aggregated version of the corresponding constraint in the individual model. Here, parameter NP_g is used. The constraint ensures for each group that the total number of patients that is scheduled is equal to the number of patients that needs to be scheduled. For the group model, constraint 1 will be defined as:

1 $$\sum_{d=1}^{D} x_{g,d} = NP_g \qquad\qquad g = 1, \dots, G$$

The final difference between the group- and the individual model is that the surgery times and the LOS of the individual patients is considered to be deterministic while the surgery times and the LOS of the groups is stochastic. The input of these parameters can be unique for each patient for the individual model. For the group model, these parameters are the expected value for all patients that belong to the group, but each individual patient in the group can deviate from this expected value. How the stochastics of the group model are dealt with is discussed in the next paragraph (paragraph 9.2.2). In table 21, all parameters and variables of the group model are shown.

9.3.2 The empirically obtained parameters in the group model

As stated in paragraph 9.1, integer programming has many benefits when it comes to solving scheduling problems. However, integer programming has one big shortcoming: numerical values have to be provided for each of the models' parameters. This is no problem for models that only include known parameters like the individual MIQP model described above, for which all parameters are assumed to be known. Unfortunately, models that must solve real-world problems often include parameters of which the true values can take on different values (J.M. de Reu, 2007). For the group MIQP model, the parameters 'surgery time' and 'LOS' are stochastic. Each group of patients has an expected surgery time and LOS, but the actual values of these parameters can vary for different patients that belong to the same group. Luckily, the empirical probability distributions of the unknown parameters can be estimated using the data of the Elkerliek. There are several ways to deal with stochastic parameters in integer programming. Some of these methods have been implemented in the model to find the best way to deal with stochastics for this research. In appendix F the findings of these implementations are shown. In the paragraphs below, it is explained how the stochastic surgery times and the stochastic LOS are dealt with in this research. It is also explained why certain choices are made.

Surgery times: Expected value

For the stochastic surgery times, it has been chosen to make use of the expected value. This means that for the surgery time, the population variance is assumed to be 0. The advantage is that the QP stays deterministic and that the complexity of the model does not increase as fast as the number of groups is increased as it does for other methods like 'chance constraints' or 'resource models' (these methods are explained in appendix F). The biggest disadvantage of this method is that the risk of possibly using more time than is maximally available is not addressed. This could be partly solved by setting the expected values higher than the actual expected values or by doing a sensitivity analysis (Rue, 2007).

There are several reasons for choosing to use the expected value for the surgery times of the groups. First of all, as stated above, when the number of groups increases, other methods that have been tested quickly become too big to solve. In order to apply these methods, the number of groups should remain quite low. This will mean that two patients within one group can possibly differ very much from each other which beats the purpose of having groups in the first place. For the Elkerliek and for the research itself, it will more insightful to have smaller, more detailed groups than to have big general groups so that the groups denote specific patient types. In addition to that, having bigger groups with a bigger range of surgery times increases the uncertainty that the model will have to deal with.

The second reason for using expected surgery times is the fact that the model is created on the tactical level. In a research by Adan et al. (2008) that has a similar goal and model as this research, using a deterministic surgery time is justified because the surgery durations are used to determine the number of surgeries of each group that is scheduled per day and with that the expected demand for the OR for each day. The further specify that the expected duration suffices because they are not interested in the overrunning of surgeries at the operational level. The same holds for this research. The main goal is to find the best mix of patients for each day and to show how the OR schedule influences the occupation of the different wards.

Feedback from the stakeholders on the expected value for the surgery time

Finally, the issue of stochastic surgery times has been discussed with stakeholders at the Elkerliek. From this discussion, it came forward that the tactical surgical schedule is used by the planners as a guideline to schedule the patients. The main goal for the Elkerliek is to find out where the planners should focus on when creating a surgical schedule and where the biggest improvements can be made. It was also stated that it is difficult to fully align the operational level with the tactical level. Last-minute cancellations leading to underuse of the OR occur regularly for example. On the other side, it can also occur that emergency patients arrive at the hospital which can lead to cancelations of elective patients or overuse of the OR. Furthermore, the OR department is very flexible. In cases where surgeries take longer than expected, it is possible to use extra personnel or to switch patients on the schedule to make sure the maximum overtime is not exceeded. The operational schedule is made on predictions of the surgery times which on average deviate from the realised surgery times with 20 minutes (more on the quality of the predictions for the surgery time was given in chapter 7.3). To deal with possible deviations from the expected surgery time on the tactical level will have only a few benefits because the schedule at the operational level is still experiencing much uncertainty.

LOS: Empirical distribution

For the uncertainty regarding the LOS, an empirical distribution for each patient group with an uncertain LOS will be created. For each ward and each day, the model sums the probabilities that patients who were operated *k* days ago occupy a bed.

This method has been chosen because in the research by Adan et al. (2008) it was shown that using this method, the quality of the tactical surgical schedule will greatly increase. Below, the mathematical representation of the bed occupation on day *d* for ward *w* is shown. Parameter $L_{g,w,k}$ denotes the probability that a patient of group *g* is still at ward *w* after *k* days after surgery and $x_{g,d\text{-}k}$ denotes the number of patients of group *g* that were operated on day *d-k*. The current day *d* is the same day as day *d-0* and day -1 is the same as day *d-1.*

$$Bed\ occupation\ on\ day\ d\ for\ ward\ w = \sum_{g=1}^{G}\sum_{k=0}^{K}\left(L_{g,w,k}\ x_{g,d-k}\right)$$

The empirical distribution of the LOS of all groups are shown in figure 9 and figure 10 in paragraph 8.4. Below in table 20, the empirical distributions of the LOS of two groups that make use of the same ward are shown in order to give a small example. Imagine that 1 patient of group 1 and one patient of group 2 is operated on day 0. When no other patients are operated, the bed occupation on day 0 would be 2 beds. The bed occupation on day 1 would be 1,52 beds and the bed occupation on day 2 would be 0,06 beds.

Table 20: Example empirical distribution LOS one patient group

Probability of occupying a bed *k* days after surgery				
Group nr.	*k* 0	1	2	3
1	100%	75%	0%	0%
2	100%	77%	6%	0%

Table 21: Parameters and variables of the group model

Parameter	Description
T_g	The expected surgery duration of patients of group *g*
$L_{g,w,k}$	The probability that a patient of group *g* is still at ward *w* after *k* days after surgery
$CP_{g,c}$	$CP_{g,c}$ = 1 when patients of group *g are* operated by sub-specialism *c* $CP_{g,c}$ = 0 otherwise
$MC_{c,d}$	The maximum number of OR blocks specialist *c* can occupy on day *d*
MOR_d	Maximum number of OR blocks that can be filled on day d
MT	Maximum overtime for every OR block in minutes
$MW_{w,d}$	The maximum number of beds on ward *w* that can be occupied on day *d*
$TU_{w,d}$	The target utilization (number of occupied beds) of ward *w* on day *d*
$Weight_{or}$	The relative weight of resource "OR"
$Weight_{ward}$	The relative weight of resource "ward"
NP_g	The number of patients of group *g* that have to be scheduled within the time horizon

Variable	Description
$x_{g,d}$	The number of patients of group *g* operated on day *d*
$y_{c,d}$	The number of OR blocks specialist *c* uses on day *d*
$ouc_{c,d}$	Overutilization of specialist c on day *d*
$uuc_{c,d}$	Underutilization of specialist c on day *d*
$ouw_{w,d}$	Overutilization of ward *w* on day *d*
$uuw_{w,d}$	Underutilization of ward *w* on day *d*

9.3.3 The objective function and constraints of the group model

The objective function of the model is a minimization function. The function minimizes the sum of the over- and underusage of the OR for each sub-specialism and day, and the sum of the over- and underusage of the wards for each ward and each day. Both the sum of the over- and underusage of the OR time and the over- and underusage of the wards are squared. The reason for this is that it is more acceptable to have multiple days with a low deviation from the target values than to have one day with one large deviation from the target. For example, the model will choose an overuse of 1 bed for 3 days over an overuse of 3 beds for 1 day. The objective function looks as follows:

$$MIN:\ Weight_{or}\sum_{c=1}^{C}\sum_{d=1}^{D}(uuc_{c,d}+ouc_{c,d})^2+Weight_{ward}\sum_{w=1}^{W}\sum_{d=1}^{D}(uuw_{w,d}+\ ouw_{w,d})^2$$

Now that the objective function and all parameters and variables are known, the constraints of the model are introduced below. For each of the constraints, the mathematical formulation and a short description of the constraint is given.

For each group, the total number of patients that is scheduled should be equal to the total number of patients that needs to be scheduled. To ensure this, the following equation was introduced:

1) $$\sum_{d=1}^{D} x_{g,d} = NP_g, \qquad\qquad g = 1, \dots, G$$

2) To denote the over- and the underutilization of the OR by sub-specialism *c*, variables $OUC_{c,d,}$ and $UUC_{c,d}$ are introduced. The target utilization of the OR is calculated by taking the number of OR blocks sub-specialism *c* has on day *d* and multiplying this number with the desired number of minutes that should be used for elective surgeries. A full OR-block has 240 minutes and for each OR-block 30 minutes should be left unused for switching times and the possibility of the arrival of emergency patients. which leaves 210 minutes. Then we get for the utilization of the specialists

$$210\, y_{c,d} + ouc_{c,d} \geq \sum_{g=1}^{G} T_g\, CP_{g,c}\, x_{g,d} \geq 210\, y_{c,d} - uuc_{c,d} \qquad c = 1, \ldots, C, \quad d = 1, \ldots, D$$

3) To determine the underutilization ($UUW_{w,d}$) and over-utilization ($OUW_{w,d}$) of ward *w* on day *d,* the inequation below is introduced. To determine how many patients a bed on ward *w* on day *d*, k is introduced, where *k* = 0,…,*K* and *K* is the maximum LOS that has to be considered:

$$TU_{w,d} - uuw_{w,d} \leq \sum_{g=1}^{G} \sum_{k=0}^{K} \left(L_{g,w,k}\, x_{g,d-k}\right) \leq TU_{w,d} + ouw_{w,d} \qquad w = 1, \ldots, W, \quad d = 1, \ldots, D$$

4) In order to ensure that a patient from group *g* can only be assigned to a day that a surgeon of sub-specialism *c* that can treat this patient is available, the following inequality was added:

$$x_{g,d} \leq CP_{g,c}\, y_{c,d} \qquad g = 1, \ldots, G \quad c = 1, \ldots, C, \quad d = 1, \ldots, D$$

5) The number of OR blocks a specialist *c* occupies on day *d* should always be smaller than or equal to the maximum number of OR block a specialist *c* can occupy on day *d*:

$$y_{c,d} \leq MC_{c,d} \qquad c = 1, \ldots, C, \quad d = 1, \ldots, D$$

6) To ensure that the maximum amount of overtime is not exceeded the following inequality was introduced. The number of OR blocks sub-specialism *c* gets on day *d* is multiplied with the maximum overtime per OR block. This multiplication should always be greater than or equal to the overtime of specialist *c* on day *d* :

$$MT\, y_{c,d} \geq ouc_{c,d} \qquad c = 1, \ldots, C, \quad d = 1, \ldots, D$$

7) To ensure that the total number of OR blocks used on day *d* does not exceed the maximum number of OR blocks available on day *d*, the following inequality was introduced:

$$\sum_{c=1}^{C} y_{c,d} \leq MOR_d \qquad d = 1, \ldots, D$$

8) To ensure that the maximum number of beds isn't exceeded, the target utilization plus the over-utilization of ward *w* on day *d* should be smaller than the maximum number of beds on ward *w* on day d:

$$TU_{w,d} + ouw_{w,d} \leq MW_{w,d} \qquad w = 1, \ldots, W, \quad d = 1, \ldots, D$$

9) The inequalities below were added to ensure that the underutilization and the overutilization are always bigger than or equal to 0:

$$uuw_{w,d} \geq 0,\ ouw_{w,d} \geq 0,\ uuc_{c,d} \geq 0,\ ouc_{c,d} \geq 0,\ w = 1, \ldots, W,\ c = 1, \ldots, C,\ d = 1, \ldots, D$$

9.4 Verification and validation of the individual and group model

Now that the models are defined, the validation and verification of the models are discussed in this paragraph.

9.4.1 Verification

The goal of verification of models is to evaluate if the model does what it is supposed to do (Thacker et al., 2004). The intention of the individual MIQP model is to find the mix of patients that should be operated on each day and the number of OR-blocks each sub-specialism uses on each day, for which the deviation from the target values is reduced as much as possible. For the group MIQP model, the intention of the model is the same, with the exception that it is not decided when each individual patient should be operated. The group model decides what the number of patients of each group that should be operated on each day should be. For both models, the intentions are met. The individual model shows what patients are operated on each day and what the number of OR-blocks for each sub-specialism on each day is that resulted in the lowest deviation from the target values. The group model shows the number of patients of each group that were operated on each day and what the number of OR-blocks for each sub-specialism on each day is that resulted in the lowest deviation from the target values. Since the verification step only looks at whether the intentions of the model are met, it can be concluded that the model is verified.

9.4.2 Validation

The goal of the validation is to evaluate if the model accurately represents reality from the perspective of the intended uses of the model (Thacker et al., 2004). This validation was done by having structural meetings with stakeholders at the Elkerliek, by comparing the output of the model with the actual performance of the Elkerliek and by changing parameters to see how the model behaves.

During meetings with stakeholders at the Elkerliek, both the output of the individual model and the group model was discussed. From these meetings, it can be concluded that the model represents the reality as it was intended by the Elkerliek. There was an agreement to validate the model by actually using a tactical schedule created by the model. Unfortunately, due to circumstances, this validation has not been performed. However, the tactical schedule was created and discussed with the stakeholders. From this meeting, it can be concluded that the output of the model is useful and represents reality as it is intended.

Next, the output of the models has been compared to the actual performance of the Elkerliek as suggested in the article by McCarl B.A. and Apland J. (1986). For this comparison, the total use of the OR and wards as predicted by the model has been compared with the actual use of the OR for the same period. On average, the model uses 1,64% more OR time than in reality. The ward usage was on average 0,36% higher for the model than it was in reality. Furthermore, it was checked whether the bed occupations did not exceed the maximum. There were never more beds occupied than the capacity of the wards at the Elkerliek. Finally, there were never more OR-blocks used for each day than possible.

Finally, to validate whether the model behaves like it is supposed to, different parameters were tried for which the outcome was known. For example, the number of patients that should be operated on within the time horizon was put to 0. This resulted in no usage of the OR and wards and no OR-blocks being allocated to sub-specialisms. When the number of patients that had to be operated on within the same time horizon was increased, the model reached a point that it did no longer find feasible solutions. While increasing the number of patients the model did never use more resources than it was allowed to. When increasing and decreasing the number of available beds and OR-blocks the model behaved as expected too.

10 Performance of the model

In this chapter, the performance of the model is discussed. The performance of the model was tested based on the runtime, the percentage of instances that were solved to optimality, and the optimality gap. Both the group model and the individual model were tested in the same way. This means that the performance of the two models can be compared to each other. In section 10.1 the setup of the different performance tests is discussed. In section 10.2 the results of the performance tests are explained.

10.1 Setup of the performance tests

In this section, the setups of the different performance tests are discussed. For both the individual model and the group model, instances were tested where the OR-block schedule was predetermined and where the OR-block schedule was defined by the model. This led to 4 different types of models that were used for the different performance tests:

- Patients are considered individually, the OR-block schedule is predetermined
- Patients are considered individually, the OR-block schedule is defined by the model
- Patients are considered in groups, the OR-block schedule is predetermined
- Patients are considered in groups, the OR-block schedule is defined by the model

Having a predetermined OR-block schedule means that the decision variable y_{cd} (number of OR-blocks allocated to sub-specialism *c* on day d) is changed to be an input parameter. The predetermined OR-block schedules used for the performance tests are created such that they closely match the actual OR-block schedules used by the Elkerliek. For this research, the performance of the model has been tested under different time horizons, under different runtimes, and under the relaxation of decision the variables. In paragraph 10.1.1 to 10.1.3, the setups of these tests are explained.

10.1.1 Setup performance tests different time horizons

In this paragraph, the setups of the tests regarding the different time horizons are explained. In table 22 the different setups for the tests are shown. These tests were performed for each of the 4 models. For each of the tests, the following input parameters were predetermined: the number of instances considered, the time horizon, and the maximum runtime for each instance. The number of instances means for how many different configurations of the parameters the model will be run. The time horizon is the number of weeks for which the model tried to find a schedule per instance. Each instance ran until optimality was reached or till the maximum runtime was exceeded.

Table 22: Setup of performance tests 1 to 4

Performance test	Number of instances	Time horizon	Max runtime per instance
1	52	1 week	15 minutes
2	52	2 weeks	15 minutes
3	17	3 weeks	30 minutes
4	12	4 weeks	60 minutes

For performance test 1, 52 different instances were used. Each instance consisted of a different week of the year. For performance test 2, 52 different instances were used. Each instance used the data of 2 weeks. When one instance finished running, the data that was used for the next instance shifted one week. This means that the first instance used data from week 1 and week 2, the second instance used data from week 2 and week 3, and so on. Next, for performance test 3 a total of 17 instances were used. Since it becomes more and more difficult to solve the model when the number of weeks increases, it was chosen to increase the runtime and to lower the number of instances considered.

For performance test 3, the first instance used data from weeks 1 to 3, the next instance used data from week 4 to 6, and so on. Finally, performance test 4 considered 12 instances of 4 weeks each. The maximum runtime per instance was once more increased. For each instance, data from a different month of the year was used. The total list of instances for performance tests 1 to 4 is shown in appendix G.

For each of these performance tests, the performance of the model was stored. After all tests had been completed, it was possible to see what happens with the performance of the model when the time horizon (and with that the number of patients) increases. Furthermore, the differences between the performance of the operational and tactical model could be compared with each other. Finally, the outcomes of the tests where the model determines the OR-block schedule could be compared with the outcomes of the tests where the actual OR-block schedule is used.

10.1.2: Setup performance tests different runtimes

Next, the effect of the runtime was tested. For this test, a time horizon of 4 weeks was chosen. The maximum runtime of performance test 5 will be set to 30 minutes. The maximum runtime of performance test 6 will be set to 120 minutes. Both tests considered 6 different instances. The instances for each of these tests consisted of the same 6 months of the year. Performance tests 5 and 6 used all four types of models that are described above. In table 23, the setup of these tests is shown. The complete list of instances of performance tests 5 and 6 is shown in appendix G.

Table 23: Setup of performance tests 5 and 6

Performance test	Number of instances	Time horizon	Max runtime per instance
5	6	4 weeks	30 minutes
6	6	4 weeks	120 minutes

10.1.3: Setup performance test relaxation of variables

Finally, the effect of the relaxation of the decision variables was tested. For this research, the decision variables of the group model are constrained to be integer. Relaxation of the decision variables means that the integrality constraints are removed, allowing the decision variables to be continuous. This changes the MIQP model into a QP model. In general, when the integer constraint is dropped, the model will be less complex to solve. (Agmon, 1954)

To test the effect of the relaxation of the decision variables, 2 performance tests were introduced. Both of these performance tests considered 12 different instances of 4 weeks each. For both tests, the maximum runtime was set to be 60 minutes. For performance test 7 only decision variable x_{gd} was relaxed. For performance test 8, the relaxed decision variables are x_{gd} and y_{cd}. This means that for this performance test, the decision variable that denotes the number of patients of group *g* that are operated on day *d* and the decision variable that denotes the number of OR-blocks that are assigned to sub-specialism *c* on day *d* are relaxed. In table 24, an overview of performance tests 7 and 8 is shown. In appendix G, the total list of instances considered for performance tests 7 and 8 is shown. The relaxation of decision variables has only been tested on the group model. Relaxation of the decision variables of the individual model was not possible because the decision variables are binary. Allowing a fractionizing of these binary variables lead to schedules that cannot be translated to realistic schedules. For performance test 8, only the group model without a predetermined OR-block was used because the group model with a predetermined OR-block schedule does not include decision variable y_{gd}. For performance test 7, both group models were used.

Table 24: Setup performance test 7 and 8

Performance test	Number of instances	Time horizon	Max runtime per instance	Relaxed decision variables
7	12	4 weeks	60 minutes	x_{gd},
8	12	4 weeks	60 minutes	x_{gd}, y_{cd}

For both tests, the performance measures of the model were stored. The decision variables that the model put out were rounded to the nearest integer number. These rounded decision variables were used to evaluate the quality of the schedules generated by the relaxed models.

10.2 Results of the performance test

After running the performance tests as described above, the output was analysed. In this section, the results and the conclusions regarding the performance of the models are discussed. In paragraph 10.2.1, the performance of the individual model is compared with the performance of the group model. In paragraph 10.2.2, the performance of the models under different time horizons is discussed. In paragraph 10.2.3, the performance of models with predetermined OR-block schedules is compared to the performance of models with an OR-block schedule that is determined by the model. The performance of the models under different runtimes is discussed in paragraph 10.2.4. Finally, in chapter 10.2.5, the effect the relaxation of decision variables has is discussed.

10.2.2 Performance of the individual model versus the group model

The first comparison that is made is the comparison of the individual model with the group model. The expectations before testing were that the group model would perform better than the individual model in terms of runtime. This means that it was expected that the gap towards optimality would be smaller for the group model compared to the individual model when both models are run the same amount of time. In order to keep the comparison of results fair, the results of the group model and the individual model for which the OR-block schedules were predetermined were only compared to each other. Consequently, the results of the group model and the individual model were the models determined the OR-block schedules were only compared to each other. This was done to make sure that the differences in results were the result of considering groups or individual patients.

Below in table 25 and table 26, the results of performance tests 1 to 4 are shown such that an easy comparison can be made between the group model and the individual model. In table 25 the results of the models with predetermined OR-schedules are represented. In table 26 the results of models with OR-block schedules that were defined by the models are shown.

Table 25: Results of performance test 1 to 4 for the group- and individual model with predetermined OR-block schedules

Predetermined OR-block schedule						
Performance test	**Runtime**		**Percentage solved to optimality**		**Optimality gap**	
	Group	*Individual*	*Group*	*Individual*	*Group*	*Individual*
1 (1 week)	0,6 seconds	7,65 min	98,08%	59,62%	0,0031%	0,017%
2 (2 weeks)	14,31 min	14,74 min	5,77%	1,92%	0,78%	0,67%
3 (3 weeks)	30,00 min	30,00 min	0,00%	0,00%	1,52%	1,90%
4 (4 weeks)	60,00 min	60,00 min	0,00%	0,00%	20,93%	26,97%

Table 26: Results of performance test 1 to 4 for the group- and individual model with OR-block schedules defined by the model

OR-block schedule determined by the model						
Performance test	**Runtime**		**Percentage solved to optimality**		**Optimality gap**	
	Group	*Individual*	*Group*	*Individual*	*Group*	*Individual*
1 (1 week)	8,34 min	15,00 min	65,38%	0,00%	5,24%	18,85%
2 (2 weeks)	15,00 min	15,00 min	0,00%	0,00%	23,34%	46,48%
3 (3 weeks)	30,00 min	30,00 min	0,00%	0,00%	24,94%	84,67%
4 (4 weeks)	60,00 min	60,00 min	0,00%	0,00%	77,36%	94,38%

From the results presented in the tables above, it can be concluded that the group model does indeed perform better than the individual model based on runtime and optimality gap. Overall, the group model found optimal solutions more often than the individual model. Furthermore, the average optimality gap of the group model is lower for all but one performance tests. The individual model only has a lower optimality gap for the performance test where the time horizon was 2 weeks and the OR-block schedule was predetermined. Only the runtime of the performance tests that considered a time horizon of 1 week can be compared since the other performance tests were stopped at the maximum runtime. For these performance tests, the runtime of the group model was significantly lower.

10.2.3 Performance under different time horizons

The second comparison is made based on the different time horizons considered. The expectations before running the performance test were that the complexity of the model would exponentially increase when the time horizon increased. To keep the comparison of the results fair, results are only compared when they are generated by the same model (group or individual) and under the same OR-block allocation policy (predetermined or defined by the model). This to make sure that differences are only due to changes in the time horizon. The results that are represented in tables 25 and 26 were also used to find the performance of the model under different time horizons.

First of all, it should be noted that the maximum runtimes of performance tests with a larger time horizon are increased. Regardless of the longer runtime, no case has occurred where the average optimality gap of a performance test is lower than a performance test that had a shorter time horizon. Looking at the results it can be concluded that the complexity of the model rapidly increases when the time horizon increases. When a short time horizon is considered, the average optimality gap is fairly low after a runtime of only 15 minutes. When the time horizon is increased to 4 weeks, the average optimality gap is significantly higher even though the runtime was increased to 1 hour.

10.2.4 Performance of models with predetermined OR-block schedules versus models with OR-block schedule determined by the model

Thirdly, the performance of models where the OR-block schedule was predetermined was compared with the performance of the models that defined the OR-block schedule themselves. To keep the comparison of the results fair, the results of the performance tests using the same type of model (group or individual) and the same time horizon are compared. This to ensure that the differences in results are due to the OR-block scheduling policies. The expectations before running were that the models that had to define the OR-blocks themselves would have a higher complexity leading to a longer runtime or a bigger optimality gap.

Below in table 27 and table 28, the results of performance tests 1 to 4 are shown such that an easy comparison can be made between the models with predetermined OR-block schedules and models that defined the OR-block schedule too. In table 27 the results of the group models are shown. In table 28 the results of the individual models are shown.

Table 27: Results of performance test 1 to 4 for the group model under different OR-scheduling policies

Group model						
Performance test	**Runtime**		**Percentage solved to optimality**		**Optimality gap**	
	Predetermined	*Defined by the model*	*Predetermined*	*Defined by the model*	*Predetermined*	*Defined by the model*
1 (1 week)	0,6 seconds	8,34 min	98,08%	65,38%	0,0031%	5,24%
2 (2 weeks)	14,31 min	15,00 min	5,77%	0,00%	0,78%	23,34%
3 (3 weeks)	30,00 min	30,00 min	0,00%	0,00%	1,52%	24,94%
4 (4 weeks)	60,00 min	60,00 min	0,00%	0,00%	20,93%	77,36%

Table 28: Results of performance test 1 to 4 for the individual model under different OR-scheduling policies

Individual model						
Performance test	**Runtime**		**Percentage solved to optimality**		**Optimality gap**	
	Predetermined	*Defined by the model*	*Predetermined*	*Defined by the model*	*Predetermined*	*Defined by the model*
1 (1 week)	7,65 min	15,00 min	59,62%	0,00%	0,017%	18,85%
2 (2 weeks)	14,74 min	15,00 min	1,92%	0,00%	0,67%	46,48%
3 (3 weeks)	30,00 min	30,00 min	0,00%	0,00%	1,90%	84,67%
4 (4 weeks)	60,00 min	60,00 min	0,00%	0,00%	26,97%	94,38%

From the results of the performance tests, it can be concluded that having the model determine the OR-block schedule greatly increases the complexity of the model. In almost all cases, the average optimality gap had more than quadrupled when the model had to determine the OR-block schedule. For the group model, the inclusion of creating the OR-block schedule seems to be manageable. For the individual model, it seems that creating the OR-block model becomes too complex when all patients are considered individually. Even when only 1 week is considered, the performance test resulted in an average optimality gap of almost 20% while the average optimality gap for the same test was 0,017% when the OR-block schedule was predetermined.

10.2.5 Performance under different runtimes

Next, the effect of the runtime is discussed. The results of performance tests 4 to 6 are shown in the table below. Additionally, figure 12 was generated. From the results and the figure, it can be concluded that the biggest optimality gap occurs when the runtime is the shortest. This is as expected. Furthermore, it can be seen that the difference between the average optimality gap is greater between the performance tests with a runtime of 30 minutes and 60 minutes than between the performance tests with a runtime of 60 minutes and 120 minutes. This means that it becomes increasingly difficult to minimize the optimality gap. Running the model for an extra hour only granted slight improvements.

Table 29: Results of performance test 4 to 6 (comparison of runtimes)

Group model with predetermined OR-block schedule					
Performance test	**Runtime**		**Percentage solved to optimality**	**Optimality gap**	
	Average	SD		Average	SD
4 (4 weeks)	60,00 min	0,00 min	0,00%	20,93%	5,53%
5 (4 weeks)	30,00 min	0,00 min	0,00%	23,78%	6,12%
6 (4 weeks)	120,00 min	0,00 min	0,00%	19,92%	2,46%
Group model with OR-block schedule defined by the model					
Performance test	**Runtime**		**Percentage solved to optimality**	**Optimality gap**	
	Average	SD		Average	SD
4 (4 weeks)	60,00 min	0,00 min	0,00%	77,36%	8,86%
5 (4 weeks)	30,00 min	0,00 min	0,00%	80,77%	5,37%
6 (4 weeks)	120,00 min	0,00 min	0,00%	75,12%	4,48%
Individual model with predetermined OR-block schedule					
Performance test	**Runtime**		**Percentage solved to optimality**	**Optimality gap**	
	Average	SD		Average	SD
4 (4 weeks)	60,00 min	0,00 min	0,00%	26,97%	2,81%
5 (4 weeks)	30,00 min	0,00 min	0,00%	30,72%	4,98%
6 (4 weeks)	120,00 min	0,00 min	0,00%	23,07%	3,91%
Individual model with OR-block schedule defined by the model					
Performance test	**Runtime**		**Percentage solved to optimality**	**Optimality gap**	
	Average	SD		Average	SD
4 (4 weeks)	60,00 min	0,00 min	0,00%	94,38%	2,30%
5 (4 weeks)	30,00 min	0,00 min	0,00%	96,68%	2,27%
6 (4 weeks)	120,00 min	0,00 min	0,00%	93,46%	2,04%

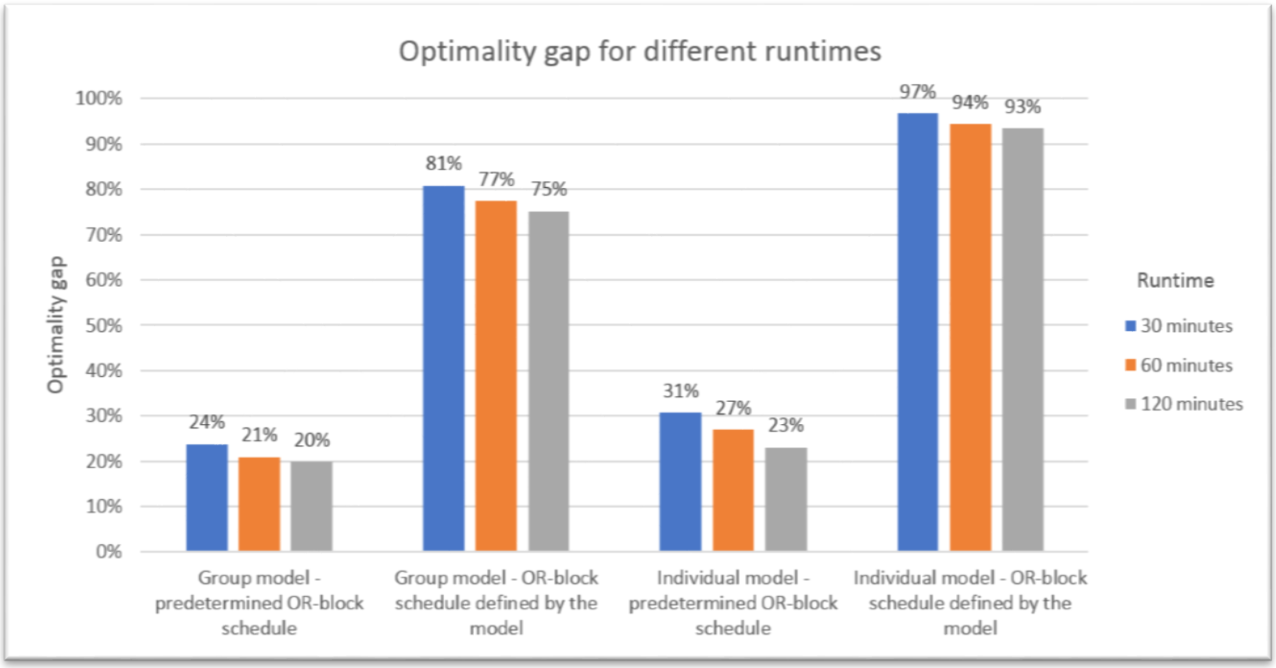

Figure 12: The optimality gap as a function of runtime

10.2.6 Performance under the relaxation of decision variables

Finally, the effect of the relaxation of the decision variables has on the performance of the model is discussed. In table 30 and table 31 below, the results of performance tests 7 and 8 are shown. These results will be discussed in more detail in the section below.

Table 30: Results of performance test 7 for the group model with predetermined OR-block schedules

Group model, predetermined OR-block schedule					
Performance test	**Runtime**		**Percentage solved to optimality**	**Optimality GAP**	
	Average	SD		Average	SD
7 (time horizon 4 weeks)	0,08 seconds	0,012 seconds	100%	0%	0,00%

Table 31: Results of performance test 7 for the group model with OR-block schedules defined by the model

Group model, OR-block schedule defined by the model					
Performance test	**Runtime**		**Percentage solved to optimality**	**Optimality GAP**	
	Average	SD		Average	SD
7 (time horizon 4 weeks)	60,00 min	0,00 min	0,00%	26,01%	14,62%
8 (time horizon 4 weeks)	0,14 seconds	0,001 seconds	100%	0,00%	0,00%

For the group model with a predetermined OR-block schedule, optimality is reached within a fraction of a second when decision variable x_{gd} is relaxed. For this model, the quality of the schedules created for performance test 7 was quite good too. The average KPI scores are shown in table 32 below. The percentage of the available OR time that is used efficiently is on average 76%, which is only a slight reduction compared to the actual efficient use of available OR time. Furthermore, the average overtime per session is expected to be about 6,5 minutes which is an improvement over the current situation. The biggest improvements are found when looking at the bed occupation. On ward 1BC, the percentage of days that an extreme deviation from the target bed occupation occurred was on average 1,78%. For the other wards, no extreme deviations occurred. The average deviation from the target bed occupation in number of beds was on average 1,35 for ward 1BC, 0,40 for ward 2BC and 0,37 for ward 2D.

Table 32: Average KPI scores for schedules created under performance test 7 by the group model with predetermined OR-block schedules

Group model, predetermined OR-block schedule								
Performance test	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)			KPI 4: Average deviation from the target (# beds)		
			1BC	2BC	2D	1BC	2BC	2D
7	76%	6,5	1,78%	0,00%	0,00%	1,35	0,40	0,37

Next, the results of performance test 7 for the group model that defines the OR-block schedule itself are discussed. The decision variable y_{cd} was restricted to be integer while the integrality constraint of decision variable x_{gd} was relaxed. Under these configurations, the model did not solve to optimality for any of the instances within an hour. The average gap towards optimality after running for 1 hour was 26%. The quality of the schedules created was quite good. The average KPI scores for the schedules created for performance test 7 are shown in table 33 below. The percentage of the available OR-time that is used efficiently is on average 82% while the average overtime is 9,7 minutes per session. These are both improvements over the current situation. The bed occupation is improved too. On ward 1BC, the percentage of days that an extreme deviation from the target bed occupation occurred was on average 4,22%. For the other wards, no extreme deviations occurred. The average deviation from the target bed occupation in number of beds was on average 1,79 for ward 1BC, 0,38 for ward 2BC and 0,62 for ward 2D.

Table 33: Average KPI scores for schedules created under performance test 7 and under performance test 8 by the group model with OR-block schedules defined by the model

Group model, OR-block schedule defined by the model								
Performance test	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)			KPI 4: Average deviation from the target (# beds)		
			1BC	2BC	2D	1BC	2BC	2D
7	82%	9,7	4,22%	0,00%	0,00%	1,79%	0,38%	0,62%
8	69%	13,85	1,98%	0,00%	0,00%	1,57	0,48%	0,38%

From the results of performance test 8, it can be concluded that relaxing decision variables x_{gd}, and y_{cd} leads to an enormous improvement with regard to the runtime of the model. With the relaxation of these variables, it only takes a fraction of a second to solve instances with a 4 week time horizon. However, based on the quality of the schedules created while running performance test 8, relaxing both decision variables is not recommendable. The average KPI scores for the schedules created for performance test 8 are shown in table 33. The quality of the schedules created for performance test 8 in terms of bed occupation was quite good. On ward 1BC, the percentage of days that an extreme deviation from the target bed occupation occurred was on average 1,98%. For the other wards, no extreme deviations occurred. The average deviation from the target bed occupation in number of beds was on average 1,57 for ward 1BC, 0,42 for ward 2BC and 0,38 for ward 2D. The costs of relaxing the decision variables are seen when the KPIs regarding the OR are evaluated. On average, only 69% of the available OR time was used which is a reduction of about 18% compared to the current situation. Furthermore, the average overtime per session is 13,85 minutes. While this does not seem very high, there are multiple days where the overuse of the OR exceeds the 30 minutes of overtime that is allowed.

From the results of performance tests 7 and 8, several conclusions are drawn. By relaxing the decision variables, the runtime of the model can be greatly reduced. When one of the decision variables remains integer, the gap towards optimality is greatly reduced. Based on the quality of the schedules generated under the different setups, it can be concluded that relaxing decision variables is only recommended when the OR-block schedule is predetermined and when there is a need to find solutions fast. When the OR-block schedule is predetermined and the decision variable x_{gd}, is not relaxed, better schedules can be found, but this requires a longer runtime (as shown in paragraph 11.3 where scenario 2 is discussed).

11 Scenario analysis

In this chapter, the different scenarios that were tested will be described and the corresponding results are presented. Different scenarios were created to test the effect of several scheduling policies on the quality of the tactical surgical schedule. For each scenario different input-parameters, decision variables or targets were used. Eventually, insightful information was gained by analysing the (near) optimal solutions for the different scenarios. The output of the model under the different scheduling policies was evaluated with the KPIs defined in chapter 7. By comparing the results of different scenarios, it was determined what ideally would be the best scheduling policy.

Below in table 34, an overview of all scenarios is given. In this table it is also specified why a certain scenario is tested and what information can be gained by it. In paragraph 11.1 the general information for all scenarios is given. In paragraphs 11.2 to 11.7 each scenario is discussed in more detail. In these paragraphs, the input for the model will be described and the results are discussed.

11.1: General information on the scenarios

For each of the scenarios, it has been chosen to test the scheduling policy for 4 different months. The months are March 2019, August 2019, November 2018, and an ‘average month’ (based on all available data). August has been chosen because this is the month with the lowest resource occupation. March has been chosen because in this month there was a medium resource occupation. November has been chosen because in this month the resource occupation was high. Finally, by taking the averages number of patients for all patient groups over the year an average month has been created. This average month is used to see how the tactical surgical schedule should look like when all months would look the same. By testing the scenarios for these months, it can be seen how the model performs under different occupation levels. Furthermore, by testing multiple months, the results are more reliable than when only one month would have been considered.

In table 35, the KPI scores for the months for which the scenarios will be tested are shown. These KPI scores will be referenced to as ‘the actual scores´ for the remainder of chapter 11. These KPI scores are based on the actual schedule used in these months by the Elkerliek. For KPI 1 and KPI 2, the expected surgery time has been used to generate the scores for these KPIs. These scores were based on the expected surgery times because the expected surgery times can differ from the actual surgery times and the arrival of emergency patients can increase overtime. Since schedules for the scenarios will be evaluated based on the expected surgery times too, it is only fair to evaluate the actual schedule in the same way. Because the LOS was not taken into consideration when creating the actual schedule, the actual LOS of the patients has been used to generate the scores for KPI 3 and KPI 4. The scores for KPI 1 and KPI 2 when the expected surgery time is used to calculate them are shown in appendix H.

Since the scheduling problem is complex to solve in terms of computational time, it has been chosen to run the model for at most 8 hours. Furthermore, the choice has been made to stop running the model when an optimality gap of 5% has been reached. These choices have been made because it would not be possible to run the model to completion for all 4 months for each of the 6 scenarios. Besides it taking too much time, finding a solution that is near the optimum or that significantly improves upon the current situation is good enough. For each of the scenarios, it will be stated for how long the model has run and what the achieved optimality gap was.

Table 34: Overview scenarios, descriptions, and goals

Scenario 1	
Description:	Only the OR usage will be optimized. The existing OR-block schedule will be used.
Goal:	To test how the optimal surgical schedule would look like for the OR without having to consider the bed occupation under the current OR-block schedule. This will be used as a baseline to compare other scenarios with.
Scenario 2	
Description:	Both the OR usage and the bed occupation at the wards will be optimized. The existing OR-block schedule will be used.
Goal:	To test how the optimal schedule would look like for both resources simultaneously under the current OR-block schedule.
Scenario 3	
Description:	Both the OR usage and the daily arrivals at each ward will be optimized. Instead of the bed occupation, the number of arrivals per day will be stabilized meaning that the LOS of patients will not be taken into consideration when creating the surgical schedule. The existing OR-block schedule will be used.
Goal:	To test the quality of schedules that are created without taking the LOS of patients into account. This scenario can be compared with scenario 2 to see how important it is to consider the LOS of patients.
Scenario 4	
Description:	Both the OR usage and the bed occupation of one big surgical ward will be optimized. Instead of considering each ward individually, one big ward is considered where all patients of the general surgery specialism go to. The existing OR-block schedule will be used.
Goal:	To test the quality of schedules that can be created when one single ward is considered. This will be interesting for the Elkerliek since the Elkerliek is planning on creating one big surgical ward.
Scenario 5	
Description:	Both the OR usage and the bed occupation at the wards will be optimized. The model will determine how the OR-blocks that are allocated to the general surgery specialism are divided among the sub-specialisms. How many OR-blocks can be used by the general surgery specialism on each day will be according to the existing OR-block schedule.
Goal:	To test the quality of schedules that are created when the available OR-blocks are allocated to the sub-specialisms by the model. It will be interesting to test if the surgical schedule can improve in quality when the OR-blocks are scheduled with the OR usage and bed occupation in mind.
Scenario 6	
Description:	Both the OR usage and the bed occupation at the wards will be optimized. The model will be completely free to create the OR-block schedule. The only constraint with regard to the OR-block schedule is the number of operating rooms at the Elkerliek.
Goal:	To test how the quality of the surgical schedules will improve when the model is free to determine the OR-block schedule. It is expected that this scenario will result in the greatest overall improvement.

Table 35: KPI scores for the schedule that was actually used by the Elkerliek when the expected surgery times are considered

	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)			KPI 4: Average deviation from the target (# beds)		
			1BC	2BC	2D	1BC	2BC	2D
March	81,2%	11,6	42%	16%	0%	3,48	1,87	1,06
August	73,8%	8,32	9%	16%	3%	1,74	1,52	0,68
November	81,1%	12,26	27%	16%	3%	2,33	1,53	1,30
Average over the year	77,8%	10,8	20%	23,5%	3%	2,21	1,88	1,10

Finally, the set-up of the model will be described. For each of the scenarios parameters, decision variables, or objective functions are changed to test different scheduling policies. Here the input of the base-model will be given. For each scenario, the input will be the same as described here, unless stated otherwise. For each of the scenarios, it will be explicitly stated what changed with regard to the base-model to test the scenario. The decision variable for the base model is $x_{g,d}$ which stands for the number of patients of group *g* operated on day *d*. The values for parameters T_g, $L_{g,w,k}$, *and* $CP_{g,c}$ will be the same as described in chapter 8 for each group *g*. The maximum overtime allowed by the model, *MT*, will be set to 30 minutes. This value is chosen because when more than 30 minutes of overtime is used in a morning session, it overlaps with the afternoon session which is not allowed.

When a scenario is used to create a schedule for a certain month. The actual number of patients of each group that has been operated on in that month is used as input for the model (parameter NP_g). The same holds for the OR-block schedule. When the trauma sub-specialism used 2 OR blocks on August 10th in reality, this will also be the case for the model. The number of OR-blocks each sub-specialism gets will be an input parameter. This means that parameter $MC_{c,d}$ will be set to the number of OR-blocks used by sub-specialism *c* on day *d.* Consequently, parameter MOR_d will be set to the number of OR-blocks used by the total general surgery specialism on day *d.* For the average month, the number of patients will be the average number of patients per month for each group. The OR-block schedule for the average month is created by taking the average number of OR-blocks each sub-specialism had for each day over the year.

The maximum number of beds on ward *w* that can be occupied on day *d* ($MW_{w,d}$) is set to be the total number of available beds at the Elkerliek. This means that for ward 1BC the maximum number of beds is 40 throughout the week. For ward 2BC this number is 40 from Monday to Friday and 0 on Saturday and Sunday. Finally, for ward 2D the maximum number of beds is 23 from Monday to Saturday and 0 on Sunday. Parameter $TU_{w,d}$ is chosen to be the sum of the LOS of all patients that go to ward *w* divided by the number of days that ward *d* is used. For example, when the sum of the LOS of all patients for ward 1BC is 14 and the ward is used 7 days, $TU_{w,d}$ is equal to 2. The target for the ward is chosen this way in order to stabilize the bed usage over the time horizon. The final parameters are $Weight_{or}$ and $Weight_{ward}$. The weight of the OR is set to 1 and the weight of the ward is set to 0,75. This choice is made based on how the costs of 1 OR hour compares to the costs of having a patient occupy 1 bed. (The costs of running the OR for 1 hour at the Elkerliek is about €170,-. The cost of 1 nurse per hour is equal to about €26,-. However, a nurse can take care of multiple patients. On average it costs about €127,- to have a patient occupying a bed for 1 full day.)

11.2: Scenario 1

For the first scenario, the model created tactical surgical schedules that only minimize the over- and underusage of the OR. The bed occupation at the different wards was not taken into account. This means that parameter $Weight_{or}$ was set to 1 and parameter $Weight_{ward}$ was set to 0. Furthermore, the OR-block schedule was predetermined. This means that the actual OR block schedule for the months March, August, and November was used. For the 'average month' the OR-block schedule described in paragraph 11.1 was used. It was expected for this scenario that the schedules created under this policy will have very good scores for KPI 1 and KPI 2. It was also expected that this policy leads to bad scores for KPI 3 and KPI 4.

In table 36 the scores of the KPIs are shown for each of the months. As can be seen, the surgical schedules for all months resulted in a good score for KPI 1 and KPI 2. The effective usage of the ward was on average 9,76% higher under the schedules created by the model than it was under the actually used schedules. The average overtime per OR-block had decreased with 5,2 minutes. Only looking at these KPI's, the scheduling policy used for scenario 1 would yield a large improvement. However, the scheduling policy of scenario 1 led to disastrous scores for KPI 3 and KPI 4. For ward 1BC the percentage of days within the time horizon that an extreme deviation occurred increased with more than 20% on average. For ward 2BC this increase was 2% and for ward 2D this increase was close to 5%. Compared to the actual KPI scores, the average deviation from the target bed occupation increased too. The average deviation increased with 1,64 beds for ward 1BC, 0,69 for ward 2BC and 0,62 for ward 2D.

Table 36: KPI scores for the schedules found using the scheduling policy of scenario 1

	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)			KPI 4: Average deviation from the target (# beds)		
			1BC	2BC	2D	1BC	2BC	2D
March	87,42%	7,26	53,6%	10,0%	0,0%	5,29	2,50	1,20
August	87,38%	4,57	28,6%	25,0%	10,0%	3,06	2,20	1,39
November	87,11%	7,09	42,9%	30,0%	10,0%	3,46	2,75	1,86
Average month	87,38%	3,29	53,6%	15,0%	10,0%	4,52	2,10	2,17

In table 37, the runtime of the model and the achieved optimality gap under scenario 1 are shown for each month. As can be seen, the scheduling policy for scenario 1 reached the target optimality gap faster than 8 hours. It was expected that the runtime for this scenario would be lower than the other scenarios because the model only takes the OR usage into account when creating a schedule.

Table 37: Runtime and achieved optimality gap for scenario 1

Month	**Runtime**	**Optimality gap**
March	1 hour 28 minutes	4,89%
August	56 minutes	4,87%
November	1 hour 3 minutes	4,92%
Average month	1 hour 50 minutes	4,77%

In figure 13 below, the expected occupations of the different wards for the schedules created under the policy used for scenario 1 are depicted for the average month (coloured dashes). In this figure, the target bed occupations are shown as well (black line). The figures depicting the expected bed occupations for the other months are shown in appendix I. The bed occupation for each of the tested months follows a similar pattern. When looking at the figures, it can be concluded that even though the policy used for scenario 1 improves upon the usage of the OR, this policy should not be implemented. The goal of this research was to find a way to reduce the variability of bed occupations. By creating a schedule that only optimizes the usage of the OR, the variability of the bed occupations only increases. However, what was learned from this scenario is what the optimal OR usage is. These results were used as a baseline to compare other scenarios with. When another scenario optimizes the usage of the ward as well as the usage of the OR, the results of scenario 1 can be used to see what the costs of considering the wards is in terms of OR efficiency.

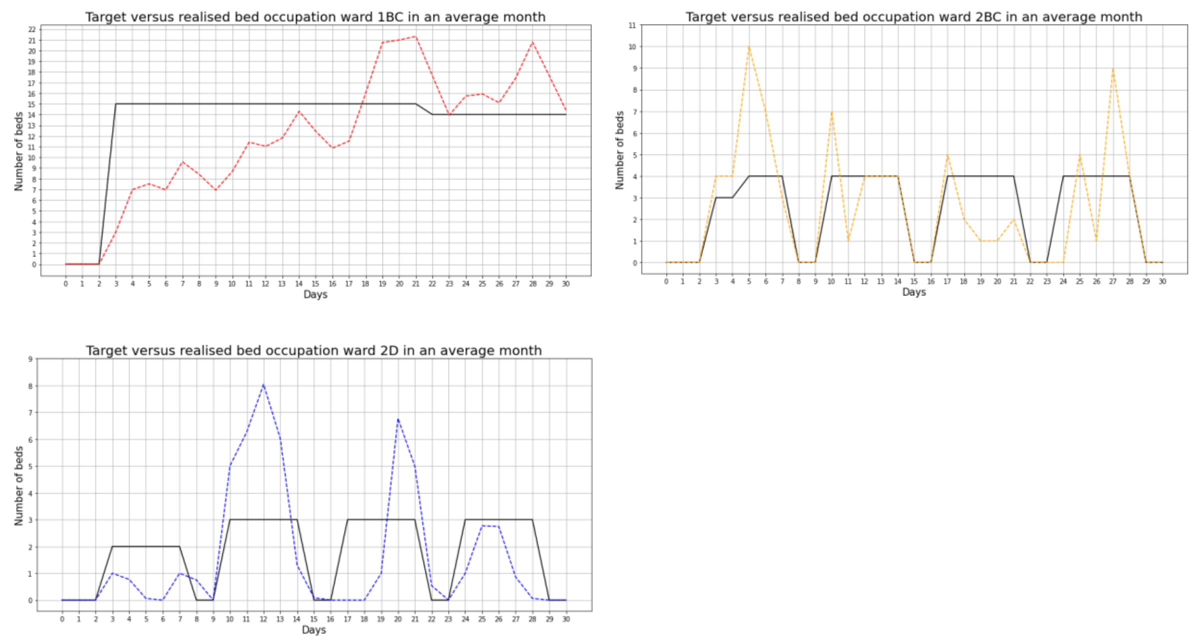

Figure 13: Target bed occupation (black line) versus realized bed occupancy (coloured dashes) for wards 1BC, 2BC and 2D for the schedule for the 'average month' created under the policy of scenario 1

11.3: Scenario 2

For the second scenario, the model created tactical surgical schedules that minimize the over- and underusage of both the OR as well as the wards. Furthermore, the OR-block schedule was predetermined. This means that the actual OR block schedule for the months March, August, and November were used. For the 'average month' the OR-block schedule described in paragraph 11.1 was used. For this scenario, the basic model as described in paragraph 11.1 was used without changes in any way. It was expected that the schedules created under the policy of scenario 2 would score lower for KPI 1 and KPI 2 than the schedules created for scenario 1. It is also expected that the scheduling policy used for scenario 2 leads to improved scores for KPI 3 and KPI 4 in comparison with both scenario 1 and the actual scores for these KPIs.

In table 38 the KPI scores for the schedules created under the scheduling policy of scenario 2 are shown. As expected, the scores for KPI 3 and KPI 4 were much better than the scores for these KPIs under scenario 1 and the actual scores for these KPIs. It was unexpected that the scores for KPI 1 and KPI 2 were only slightly worse than the scores for these KPIs for scenario 1. This means that taking the wards into account does not lead to huge costs in terms of OR efficiency. The efficient use of the OR decreased by less than 1% and the average overtime increased with less than 1 minute compared to scenario 1. Compared to the actual scores for these KPIs, the efficient use of the OR was increased with 9,52% on average. The average overtime per OR-block has decreased by almost 5 minutes on average. Based on the OR usage alone, scenario 2 is a big improvement upon the current situation. The biggest gain, however, was achieved for KPI 3 and KPI 4. There were almost no extreme deviations under the schedules created by the model when using the scheduling policy of scenario 2. Furthermore, the average deviation from the target bed occupation has decreased with 1,33 beds for ward 1BC, 1,30 beds for ward 2BC, and 0,40 beds for ward 2D compared to the actual score.

Table 38: KPI scores for the schedules found using the scheduling policy of scenario 2

	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)			KPI 4: Average deviation from the target (# beds)		
			1BC	2BC	2D	1BC	2BC	2D
March	86,80%	7,35	3,6%	0,0%	0,0%	1,35	0,30	0,51
August	87,26%	4,86	3,6%	5,0%	0,0%	1,18	0,92	0,74
November	87,07%	7,34	0,0%	0,0%	0,0%	0,88	0,20	0,71
Average month	87,21%	3,72	0,0%	0,0%	0,0%	1,03	0,20	0,57

In table 39 the runtime of the model and the achieved optimality gap for scenario 2 are shown for each month. As can be seen, the scheduling policy for scenario 2 did not reach the target optimality gap within 8 hours. It was expected that the runtime for this scenario would be higher than the runtime of scenario 1, since the model had to take multiple resources into account and had to calculate the bed occupation for each day. Still, the achieved optimality gaps were between 10% and 25%, which is fairly good. The policy used for scenario 2 is actually the same policy as was used for the performance tests of the group model with predetermined OR-schedules (chapter 10). When the model was run for 2 hours, an average optimality gap of 19,92% was reached. The average optimality gap of scenario 2 was 17,78%. Running the model for 6 extra hours resulted in only a slight improvement. Considering that finding the absolute best schedule is not important, it would suffice to run the model for 2 hours only in order to generate results when the scheduling policy of scenario 2 is used.

Table 39: Runtime and achieved optimality gap for scenario 2

Month	**Runtime**	**Optimality gap**
March	8 hour	15,60%
August	8 hour	25,66%
November	8 hour	10,71%
Average month	8 hour	18,15%

In figure 14 below, the expected occupations of the different wards for the schedules created under the policy used for scenario 2 are depicted for the average month (coloured dashes). In this figure, the target bed occupations are shown as well (black line). The figures depicting the expected bed occupations for the other months are shown in appendix I. When looking at the figures it can be concluded that using the policy of scenario 2 greatly improved upon the current situation with regard to the bed occupation. It can be seen that for ward 2BC, the target occupation was nearly equal to the achieved occupation. This was due to the fact that this ward is the easiest to stabilize, since all patients only stay one day. Even though ward 1BC and 2D are harder to stabilize, under the scheduling policy of scenario 2 the occupation on these wards was close to the targets as well. What is noticeable from the figure depicting the occupations at ward 1BC is that a certain pattern emerged. On each Friday, a slight peak occurred. This increase was due to the model trying to ensure that there are enough patients to keep the occupation high enough during the weekends. Overall, the scheduling policy of scenario 2 showed great promise. It improved the OR usage whilst also stabilizing the occupation at the wards.

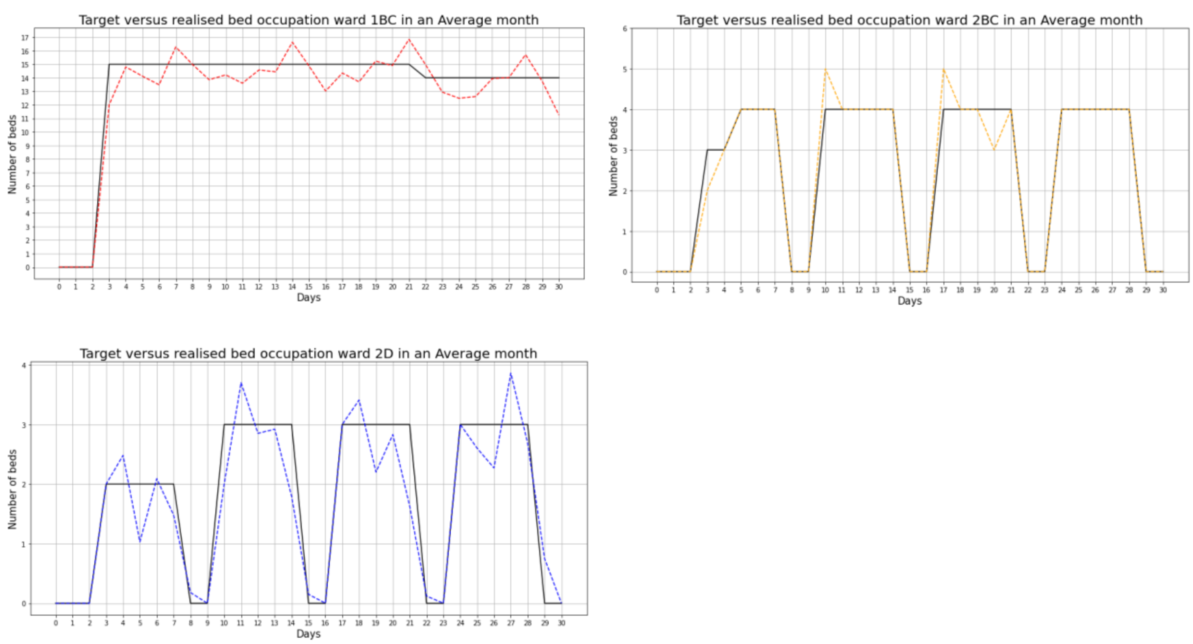

Figure 14: Target occupation (black line) versus realized bed occupancy (coloured dashes) for wards 1BC, 2BC and 2D for the schedule created for the 'average month' under the scheduling policy of scenario 2

11.4: Scenario 3

For the third scenario, the model created tactical surgical schedules that minimized the over- and underusage of the OR whilst stabilizing the daily arrivals to each of the wards. This means that the LOS of patients was not taken into consideration by the model when generating the surgical schedule. The goal of the model with regard to the wards was to have the same number of patients entering the wards each day from Monday to Friday (since no elective surgeries are performed in the weekend, no new patients can enter the wards on Saturday and Sunday). Furthermore, the OR-block schedule was predetermined. Meaning that the actual OR block schedules for the months March, August, and November were used. For the 'average month' the OR-block schedule described in paragraph 11.1 was used.

For the model, two adaptions were made: The parameter $L_{g,w,k}$ was set to 1 when k = 0 and this parameter value was set to 0 otherwise. The second parameter that changed, was the target utilisation of the ward ($TU_{w,d}$). This parameter was changed to denote the number of patients that should arrive each day to have the same number of arrivals at each day at each ward.

It was expected for this scenario that the schedules created under this policy would score lower for KPI 1 and KPI 2 then the schedules created for scenario 1. It was also expected that the scheduling policy of scenario 3 led to improved scores for KPI 3 and KPI 4 compared to the actual scores for these KPIs. However, a lower score for these KPIs was expected when compared with scenario 2. This scenario was interesting to run, because this scheduling policy is fairly easy to implement for the Elkerliek. When improvements are found this policy can easily be introduced. Furthermore, using this scenario, it was possible to test if it is actually necessary to take the LOS into account when creating schedules.

In table 40 the KPI scores for the schedules created under the scheduling policy of scenario 3 are shown. Again, the scores for KPI 1 and KPI 2 were improved compared to the actual scores for these KPIs. Compared to the scores achieved by using the scheduling policy of scenario 1, the scheduling policy of scenario 3 performed only slightly worse based on these OR focused KPIs. Looking at the scores for KPI 3 and KPI 4, an interesting observation could be made. For ward 1BC this scheduling policy led to bad scores for these KPIs. The percentage of days with extreme deviation from the target bed occupation increased with 7% compared to the current scores. Furthermore, the average deviation from the target bed occupation increased with about 1 bed for ward 1BC. Based on the scores for ward 1BC this scheduling policy seems to yield quite unfavourable results. However, using this scheduling policy, the KPI scores for wards 2BC and 2D greatly improved. This was due the fact that patients at these wards have a shorter LOS. Not taking the LOS into consideration has less of an effect for these wards. The percentage of days an extreme deviation occurred decreased to 0% for both ward 1BC and ward 2BC. The average deviation from the target bed occupation decreased with an average of 1,3 beds for ward 2BC and 0,37 for ward 2D.

Table 40: KPI scores for the schedules found using the scheduling policy of scenario 3

	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)			KPI 4: Average deviation from the target (# beds)		
			1BC	2BC	2D	1BC	2BC	2D
March	86,40%	7,28	35,7%	0,0%	0,0%	3,69	0,40	0,47
August	87,33%	4,69	39,3%	0,0%	0,0%	3,33	0,50	0,67
November	87,18%	8,05	17,9%	0,0%	0,0%	2,43	0,50	0,79
Average month	87,34%	3,42	35,7%	0,0%	0,0%	3,51	0,20	0,72

In table 41 the runtime of the model and the achieved optimality gap under the scheduling policy of scenario 3 are shown for each month. As can be seen, the scheduling policy for scenario 3 reached the target optimality gap within the maximum runtime of 8 hours. Another advantage of this scheduling policy was found by looking at the runtime. Compared to scenario 2, the runtimes were way shorter.

Table 41: Runtime and achieved optimality gap for scenario 3

Month	**Runtime**	**Optimality gap**
March	1 hour 10 minutes	4,76%
August	1 hour 2 minutes	4,89%
November	49 minutes	4,99%
Average month	53 minutes	4,84%

In figure 15 below, the expected occupations of the different wards for the schedules created under the policy used for scenario 3 are depicted for the average month (coloured dashes). In this figure, the target bed occupations are shown as well (black line). The figures depicting the expected bed occupations for the other months are shown in appendix I. These figures show what is also described in table 40. The bed occupation of ward 2BC and 2D were very close to the target occupations whilst the occupation at ward 1BC still varied heavily from day to day. However, a predictable pattern seemed to form for the occupation at ward 1BC. At the start of the week, the occupations were quite low and over the week the occupations increased till it reached its highpoint on Friday to once again decrease until Monday. Overall, the scheduling policy of scenario 3 seems to improve upon the current situation. The usage of the OR and the bed occupation at wards 2BC and 2D was greatly improved. For ward 1BC, this scheduling policy seems less useful. It seems that considering the LOS of patients is important for wards where patients stay for a longer amount of time. When the runtime of the model is not an issue, it would be better to create a surgical schedule under the scheduling policy of scenario 2, since better results were found under this policy.

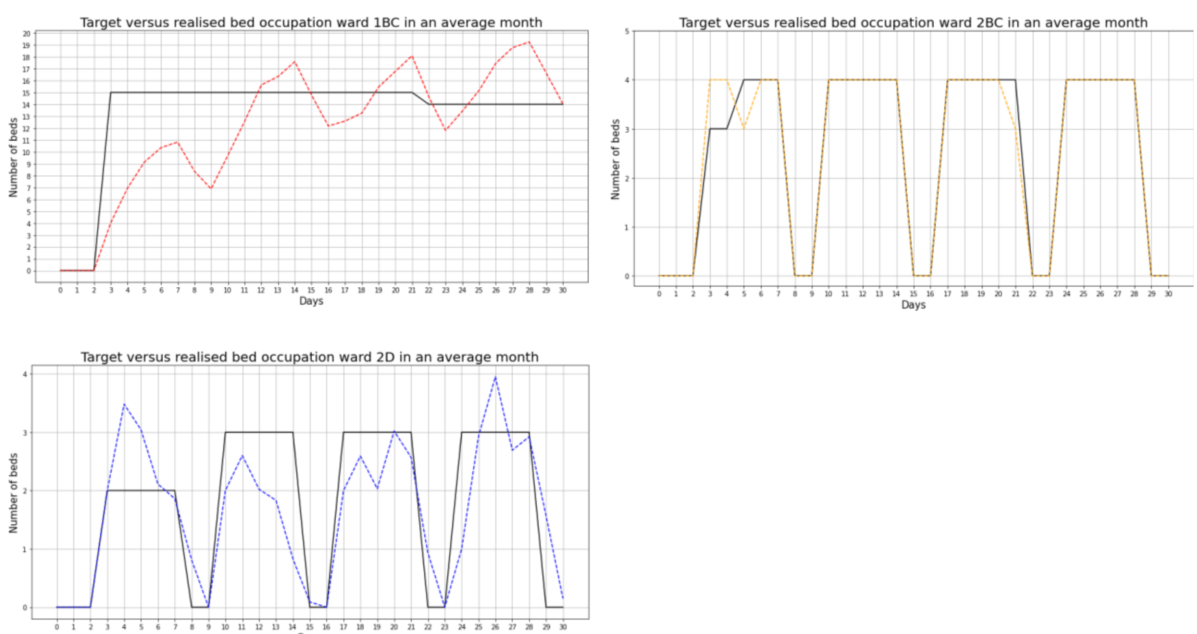

Figure 15: Target occupation (black line) versus realized bed occupancy (coloured dashes) for wards 1BC, 2BC and 2D for the schedule created for the 'average month' under the scheduling policy of scenario 3

11.5: Scenario 4

For the fourth scenario, the model created tactical surgical schedules that minimized the over- and underusage for both the OR and the wards. This is the same scheduling policy as the one used for scenario 2. The difference between these scenarios is that for scenario 4 all wards were combined into one big surgical ward. Furthermore, the OR-block schedule was predetermined. Meaning that the actual OR block schedules for the months March, August and, November were used. For the 'average month' the OR-block schedule described in paragraph 11.1 was used. For the model, all parameters that have to do with the ward changed. The parameter $MW_{w,d}$ (maximum bed occupation for ward w on day d) changed to MW_d because only one ward was considered. The value for this parameter was set to be the sum of all available beds of wards 1BC, 2BC and, 2D. The parameter $TU_{w,d}$ changed to TU_d because only one ward was considered. The value for this parameter was set to be the sum of the LOS of all patients, divided by the number of days in the time horizon.

This scenario was created to see what happens when all patients start going to one single ward instead of several wards. This scenario is currently being implemented in the Elkerliek. Since the same scheduling policy was used as in scenario 2, similar scores for the KPIs are expected.

In table 42, the KPI scores for the schedules created under the scheduling policy of scenario 4 are shown. The scores for KPI 1 and KPI 2 were very similar to the ones of scenario 2, as expected. Again, the scores for these KPIs were only slightly worse than the ones under the scheduling policy of scenario 1. Comparing the scores for KPI 3 and KPI 4 of this scenario with the scores of other scenarios is a bit difficult, since these scenarios consider only one big surgical ward. Overall, the scores for KPI 3 and KPI 4 as shown in table 42 seem quite good. Considering that all patients were considered in this single ward, an average deviation from the target bed occupation of 1,46 seems not to much. The percentage of days in which an extreme deviation occurred has decreases significantly compared to the current score for this KPI. Furthermore, as can be seen in figure 16 below, all these extreme deviations occurred during the weekend.

Table 42: KPI scores for the schedules found using the scheduling policy of scenario 4

	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)	KPI 4: Average deviation from the target (# beds)
			Surgical ward	Surgical ward
March	87,37%	7,33	7,1%	1,94
August	87,33%	4,69	7,1%	1,69
November	87,15%	7,11	0,0%	0,68
Average month	87,19%	3,74	10,7%	1,56

In table 43 the runtime of the model and the achieved optimality gap under the scheduling policy of scenario 4 are shown for each month. As can be seen, the scheduling policy for scenario 4 did not reach the predetermined optimality gap in most of the cases. Especially the average month seemed difficult to create a surgical schedule for. However, this scenario seems more easily solvable than scenario 2 because the achieved optimality gaps were lower for this scenario within 8 hours than the ones for scenario 2. This is probably due to the fact that fewer wards need to be considered.

Table 43: Runtime and achieved optimality gap for scenario 4

Month	**Runtime**	**Optimality gap**
March	8 hours	10,11%
August	8 hours	10,28%
November	7 hours 17 minutes	04,92%
Average month	8 hours	38,03%

In figure 16 below, the expected occupations of the different wards for the schedules created under the policy used for scenario 4 are depicted for each of the 4 months considered (red dashes). In this figure, the target bed occupations are shown as well (black line). As described above, the extreme deviations consisted purely of dips in the weekends. Apart from these points, the bed occupation seemed quite stable from Mondays to Fridays. By running this scenario, it was found that it would be better to set a lower target bed occupation during the weekends when only one ward is considered. Furthermore, it was shown that when considering one big ward, the tactical surgical schedule can still help in stabilizing the bed occupancy rate.

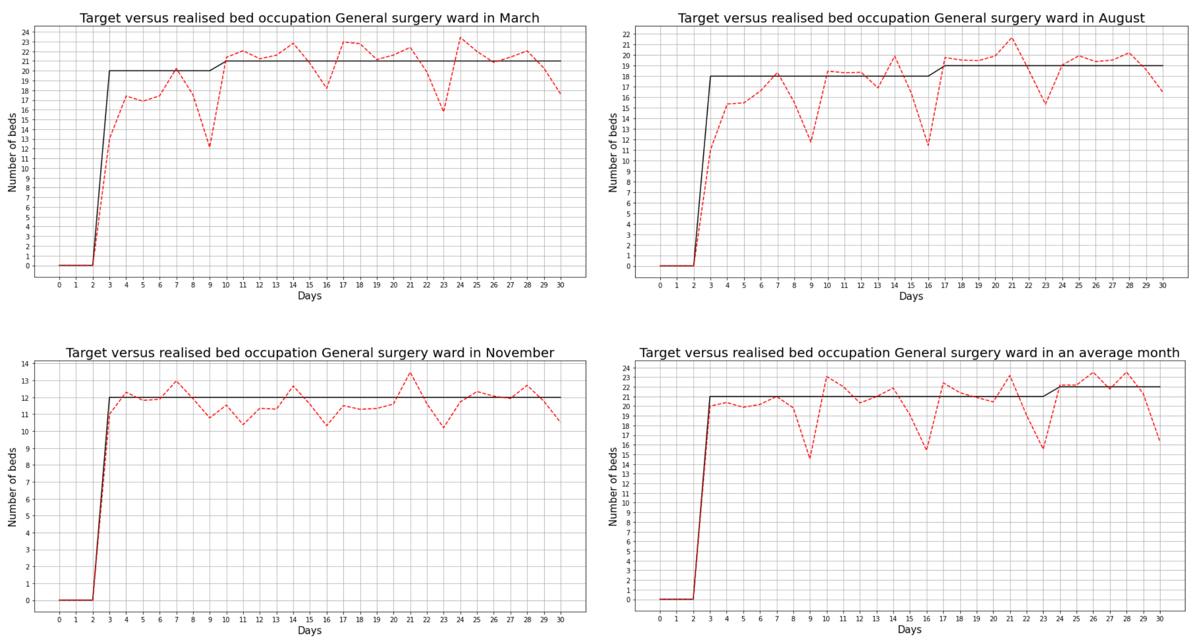

Figure 16: Target occupation (black line) versus realized bed occupancy (coloured dashes) for wards 1BC, 2BC and 2D for the schedule created for the months March, August, November and the 'average month' under the scheduling policy of scenario 4

11.6: Scenario 5

For the fifth scenario, the model created tactical surgical schedules that minimized the over- and underusage for both the OR and the wards. What was different for this scenario is that the allocation of OR-blocks amongst the different sub-specialisms was determined by the model. The number of OR-blocks available for the sub-specialisms combined was still the same as for the actual OR block schedules. What changed for the model is that decision variable $y_{c,d}$ was introduced. This variable denotes the number of OR-blocks sub-specialism *c* gets on day *d*. For this scenario, MOR_d does not change from the actual schedules used by the Elkerliek. This means that only the allocation of OR-blocks amongst the sub-specialisms was changed by the model. This scenario was created to see what happens when the OR-block allocation is no longer based on the preferences of the surgeons themselves, but based on what is best for the resource usage. Since the model got more freedom to schedule the patients, it was expected to see an improvement for all KPIs compared to scenario 2.

In table 44, the KPI scores for the schedules created under the scheduling policy of scenario 5 are shown. The scores for KPI 1 were quite similar to the scores for scenario 2. Compared to scenario 1, the scores for KPI 1 were slightly lower. Surprisingly, the scheduling policy of scenario 5 outperformed all policies of the first 4 scenarios. On average, the overtime for scenario 5 was 3 minutes lower compared to scenario 1. This is surprising because, for scenario 1, only the OR was taken into consideration and for scenario 5 the ward usage was optimized too. Scenario 5 scored the same on KPI 3 as scenario 2 did. This means that there was a significant improvement over the current situation with regard to extreme deviations from the target bed occupation. Scenario 5 and scenario 2 both scored about the same for KPI 4. The main gain of having the model decide what sub-specialism can use what OR-block, was unexpectedly seen in the usage of the OR.

Table 44: KPI scores for the schedules found using the scheduling policy of scenario 5

	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)			KPI 4: Average deviation from the target (# beds)		
			1BC	2BC	2D	1BC	2BC	2D
March	86,92%	1,33	3,6%	0,0%	0,0%	1,02	0,50	0,51
August	87,28%	3,40	3,6%	5,0%	0,0%	1,35	1,00	0,88
November	86,80%	1,27	0,0%	0,0%	0,0%	0,63	0,25	0,71
Average month	87,04%	4,11	0,0%	0,0%	0,0%	1,07	0,10	0,47

In table 45 the runtime of the model and the achieved optimality gap under the scheduling policy of scenario 5 are shown for each month. As can be seen, the scheduling policy for scenario 5 did not reach the predetermined optimality gap within 8 hours. The optimality gap was still quite large after running the model for 8 hours. This was as expected since the model had a second decision variable.

Table 45: Runtime and achieved optimality gap for scenario 5

Month	**Runtime**	**Optimality gap**
March	8 hours	54,36%
Augusts	8 hours	66,50%
November	8 hours	87,60%
Average	8 hours	72,83%

In figure 17 below, the expected occupations of the different wards for the schedules created under the policy used for scenario 3 are depicted for the average month (coloured dashes). In this figure, the target bed occupations are shown as well (black line). The figures depicting the expected bed occupations for the other months are shown in appendix I. What can be seen in these figures is that the bed occupation under scenario 5's policy was very similar to the bed occupation for the schedules created for scenario 2. Letting the model decide what sub-specialism can use what available OR-block had a lower impact on the ward usage as expected. Changing the entire way the OR-blocks are divided among sub-specialism for this slight improvement does not seem worth it.

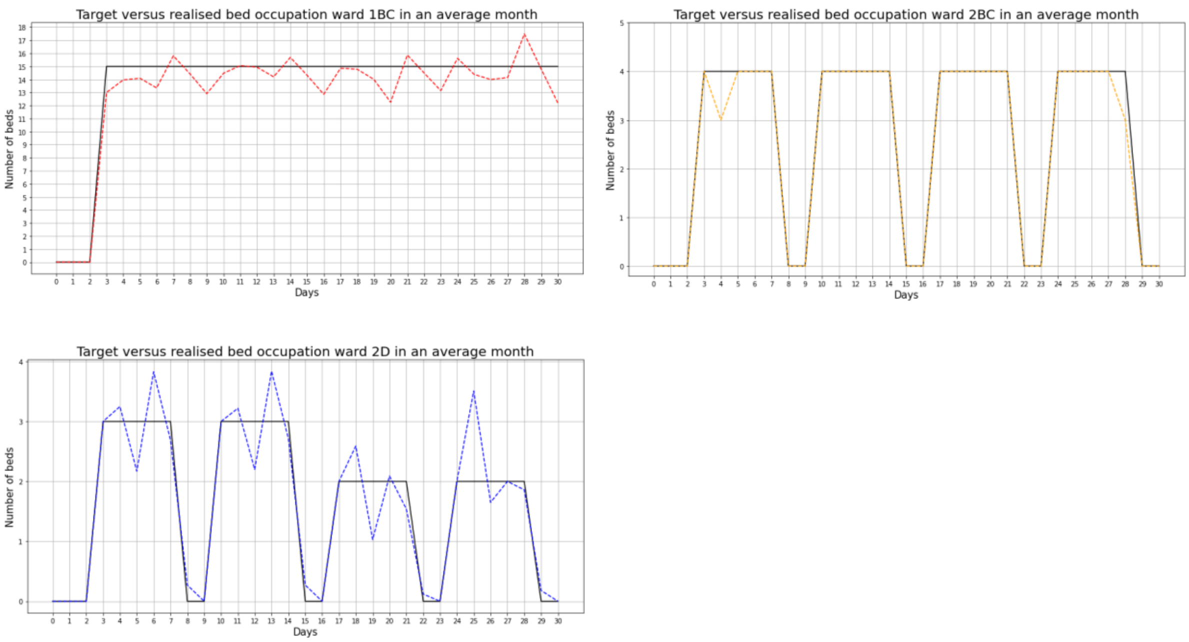

Figure 17: Target occupation (black line) versus realized bed occupancy (coloured dashes) for wards 1BC, 2BC and 2D for the schedule created for the 'average month' under the scheduling policy of scenario 5

11.7: Scenario 6

For the sixth and final scenario, the model created tactical surgical schedules that minimized the over- and underusage for both the OR and the ward. However, for this scenario, the allocation of OR-blocks amongst the different sub-specialisms was determined by the model. For scenario 5, this OR-block allocation was restricted by the number of OR-blocks that was available for the general surgery specialism on each day. For this scenario, the model was completely free to allocate OR-blocks to the sub-specialisms. The only restriction with regard to the OR-block allocation was the maximum number of ORs and surgeons that are available at the Elkerliek (meaning that Saturday and Sunday still no surgery can take place for example). What changed compared to the basic model is that variable $y_{c,d}$ was introduced. This variable denotes the number of OR-blocks sub-specialism c gets on day d. Furthermore, parameter $MC_{c,d}$ was set to be equal to the total number of OR-blocks that all surgeons of sub-specialism c can fill for each day d. For this scenario, MOR_d was set to the maximum number of OR-blocks that is available on day d. For the Elkerliek MOR_d was set to 16, since 8 rooms are divided into 2 OR-blocks for each day.

This scenario was created to see what would be the best possible tactical surgical schedule that can be achieved based on the resource usage. Because the model gets complete freedom to allocate OR-blocks and patients to each day it was expected that the best scores for the KPIs will be found under the scheduling policy of this scenario.

In table 46, the KPI scores for the schedules created under the scheduling policy of scenario 6 are shown. The results for KPI 1 were decreased by 1,26% compared to the scores achieved for scenario 1. Still, scenario 6 shows a huge improvement compared to the actual scores. KPI 2 is the first KPI for which a big improvement was found. The average overtime per OR-block under the scheduling policy of scenario 6 was the lowest of all scenarios. The average overtime per OR-block was lower than 20 seconds for each of the months that were tested. This was not only a huge improvement over the actual average overtime, but also a big improvement over the average overtime achieved under the scheduling policy of scenario 1. Under the scheduling policy of scenario 6, there were no extreme deviations from the target bed occupation. This means that with regard to this KPI, the absolute optimum was reached. For KPI 4, the scheduling policy used for scenario 6 outperformed all other scheduling policies too. The average deviation from the target bed occupation was less than 1 bed in all cases. Based on the scores for the KPIs it would be best to create a surgical schedule under the policy used for scenario 6.

Table 46: KPI scores for the schedules found using the scheduling policy of scenario 6

	KPI 1: Effective use OR on average (%)	KPI 2: Average overtime per OR-block (minutes)	KPI 3: Percentage of days an extreme deviation occurred (%)			KPI 4: Average deviation from the target (# beds)		
			1BC	2BC	2D	1BC	2BC	2D
March	86,43%	0,24	0,0%	0,0%	0,0%	0,84	0,10	0,45
August	86,38%	0,15	0,0%	0,0%	0,0%	0,84	0,10	0,54
November	85,52%	0,11	0,0%	0,0%	0,0%	0,72	0,15	0,67
Average month	85,80%	0,22	0,0%	0,0%	0,0%	0,93	0,20	0,46

In table 47 the runtime of the model and the achieved optimality gap under the scheduling policy of scenario 6 are shown for each month. As can be seen, the scheduling policy for scenario 6 did not reach the predetermined optimality gap within 8 hours. The optimality gap was still quite large after running the model for 8 hours. However, this was expected since the model had very many decision variables and a big feasible region to consider. Even though the optimality gaps were quite large, the results after running the model were quite good.

The configurations of the model used for scenario 6 were the same as for the performance tests of the group model with OR-block schedules defined by the model (chapter 10). The average optimality gap of this model after a runtime of 2 hours was 75,12%. Running the model for 8 hours for the months considered for scenario 6 resulted in an average optimality gap of 76.38%. This average score is actually higher than the average score achieved in 2 hours. This is possible because for the performance test, more months were tested. The months that were considered for scenario 6 appear to be the more complex iterations. These results suggest that running the model for 2 hours will suffice.

Table 47: : Runtime and achieved optimality gap for scenario 6

Month	**Runtime**	**Optimality gap**
March	8 hours	76,35%
Augusts	8 hours	72,01%
November	8 hours	82,83%
Average	8 hours	74,34%

In figure 18 below, the expected occupations of the different wards for the schedules created under the policy used for scenario 3 are depicted for the average month (coloured dashes). In this figure, the target bed occupations are shown as well (black line). The figures depicting the expected bed occupations for the other months are shown in appendix I. In figure 18, it can clearly be seen that the achieved bed occupations were very close to the target occupations for each of the wards. Just like the scores for KPI 3 denoted, there were no big peaks or dips in demand to be seen. For both the usage of the OR and the usage of the wards, the scheduling policy under scenario 6 yielded the best results. However, to achieve these results, the current way of creating the master surgical schedule and the current way of allocating OR-blocks to specialists has to be adapted. What scenario 6 shows is that there is still room for improvement and that the surgeries can be scheduled on the tactical level in such a way that the usage of both OR and wards can be aligned with each other.

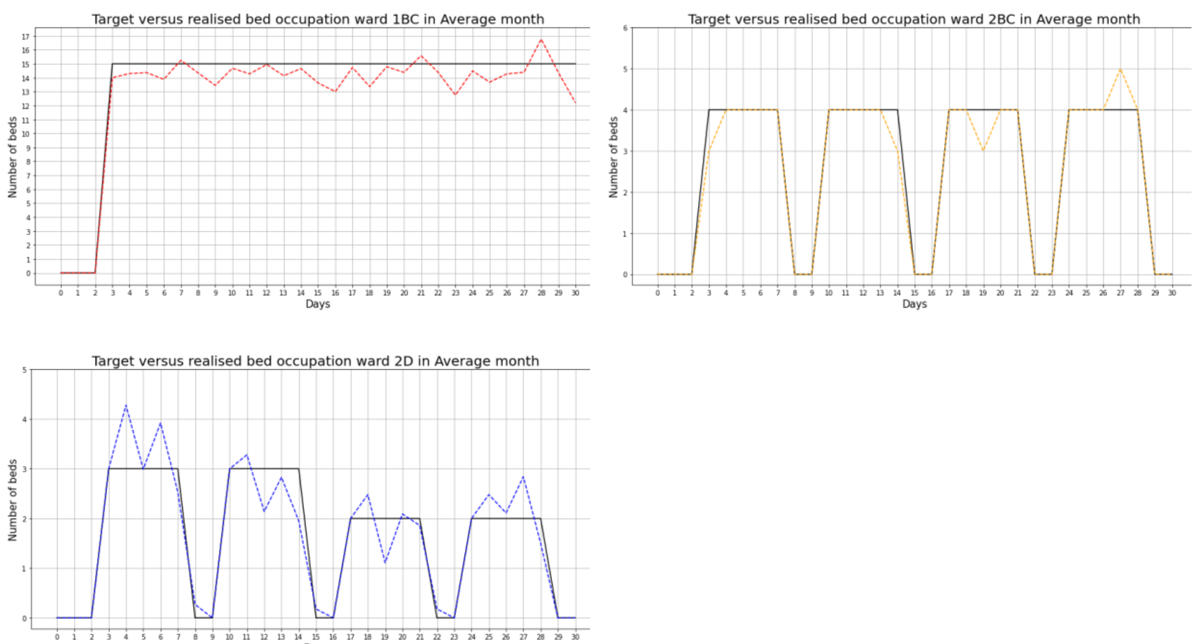

Figure 18: Target occupation (black line) versus realized bed occupancy (coloured dashes) for wards 1BC, 2BC and 2D for the schedule created for the 'average month' under the scheduling policy of scenario 6

12 Conclusions

This research has been conducted to find an answer to the question: "How can statistical information on the provided datasets and the use of a MIQP model help the Elkerliek when developing a tactical surgical schedule for a selected department taking both the OR and the downstream resources at the wards into account?". In this chapter, the conclusions are presented. At the end of this chapter, the research question will be answered.

The first conclusion that can be drawn from this research is that it is beneficial to take the usage of the wards into account when creating an OR schedule. This follows from the results of scenario 1 and scenario 2. Currently, the OR schedule is focused mainly on the OR usage, with downstream resources acting as constraints. In this research, it has been found that the OR usage under such a policy can only perform slightly better than a policy that takes the bed occupation at the different wards into account too. The difference in efficient OR usage between a policy where only the OR usage is taken into consideration and a policy where both resources are taken into account was less than 1%. The difference in average overtime per OR-block between these two policies was only 1 minute on average. Opposite to this, the difference in ward usage was significant. Under the policy that took the ward usage into account too, there were almost no days that there was a deviation from the target occupation of 4 or more beds. Under this policy, the average deviation from the target bed occupation has been greatly reduced too. For ward 1BC and ward 2D, the average deviation has more than halved. For ward 2BC, the average deviation has been reduced even more in most cases. In short, having a reduction of less than 1% of efficient use of OR time and an increase of 1 minute in overtime is a low price to pay for the big improvements that can be made for the ward usage.

The second conclusion is based on the performance tests described in chapter 10. First of all, when the model is used to create schedules for a time horizon of 4 weeks, it would suffice to run the model for about 1 hour. Based on the performance tests it can be stated that running the model more than 1 hour leads to only minor improvements. Furthermore, it was found that the complexity of the model rapidly increases when individual patients are considered. It is suggested to only use the individual model for a time horizon of up to 2 weeks in order to find solutions that are close to the optimum. Next, having the model come up with an OR-block schedule increases the complexity of the model significantly. Finally, the quality of the schedules created when decision variables are relaxed is lower than when the decision variables are restricted to be integer However, the runtime of the model when the decision variables are relaxed is only a fraction of a second. When there is enough time, the decision variables should not be relaxed, but when time is of the essence, relaxing the decision variables can be helpful. Furthermore, when the OR-block schedule is defined by the model, relaxing the decision variables leads to very poor schedules. In conclusion, relaxation of the decision variables is only helpful when the OR-block schedule is known and a solution is needed fast.

The third conclusion regards the consideration of the bed occupation by the model. There are several ways to consider the bed occupation. Three main distinctions can be made here. Firstly the bed occupation can be ignored altogether, like in scenario 1. Secondly, the bed occupation can be considered where the LOS of each patient is taken into account as well, as in scenario 2. Finally, the bed occupation can be taken into account by stabilizing the number of new patients that enter the different wards each day, like in scenario 3.

Based on the analysis of the scenarios it can be concluded that for wards for short stay and day-care patients, it is good enough to only stabilize the arrivals at the wards. Even though using this method for the ward where patients stay for a longer duration works better than ignoring the bed occupation altogether, it is not optimal. For wards where patients can stay for a longer duration it is best to consider the LOS when creating an OR-schedule. This means that a scheduling policy as described in scenario 2 will result in the best resource usage.

The fourth conclusion will be based on the results of scenario 4. For this scenario, only one big surgical ward was considered. Having all patients go to one big ward can have several benefits. First of all, it will be easier to adjust the personnel to the demand of the patients. Secondly, for one big ward, less spare capacity is needed as a safety net compared to several small wards. In terms of the results found for scenario 4, it can be concluded that it is difficult to keep the bed occupation stable for the entire week. In the weekends, dips in bed occupation occur because there are no surgeries. This is more of a problem for the big surgical ward than it is for ward 1BC because for the big surgical ward a large part of the patients are short stay and day-care patients. The advice here is to not try to keep the bed occupation high during the weekend, but to set a higher target occupation during workdays and a lower target occupation in the weekends. Another solution would be to start performing surgeries in the weekend.

The fifth conclusion will be on the OR-block scheduling policy. Three different scenarios have been compared to gain insight into this topic. Firstly a scenario where the current OR-block schedule is used (scenario 2). Secondly, a scenario where the total number of OR-blocks available for the general surgery specialism on each day stays the same, but the distribution of these OR-blocks towards the different sub-specialisms is defined by the model (scenario 5). And finally, a scenario where the model is only bound by the physical constraints of the hospital when creating an OR-block schedule (scenario 6). The absolute best results with regard to the OR usage as well as the bed occupation are found under the scheduling policy of scenario 6. For this scenario, no extreme deviations from the target bed occupation occurred. Furthermore, the average overtime per OR-block was reduced to almost 0 minutes. Scenario 5 performed only slightly better than scenario 2 based on OR usage and about the same as scenario 2 based on the bed occupation. In a situation where the general surgery specialism is the only specialism that makes use of the OR, scenario 6 would have been the best scenario to implement. However, other specialisms restrict the freedom to allocate OR-blocks. It is not realistic to expect the Elkerliek to implement this method. Looking at the differences between scenario 2 and scenario 5, having the model optimize the OR-block schedule bounded by the available OR-blocks for the general surgery specialism does not lead to the big improvement that was originally expected. It turns out that under the current OR-block allocation it is still possible to create a good tactical surgical schedule that takes both the OR and the usage of the wards into account.

The main research question: “How can statistical information on the provided datasets and the use of a MIQP model help the Elkerliek when developing a tactical surgical schedule for a selected department taking both the OR and the downstream resources at the wards into account?” can be answered as follows: By retrieving statistical information on the provided datasets, it can be determined how the mix of patients that are operated on each day should look like. Furthermore, the best scheduling policy can be found by analysing the data and by using the model. Finally, it can be found what the expected results are of certain schedules so that the choice can be made to actually implement the schedule or to adapt it before implementing it.

13 Discussion

In this chapter, three points will be discussed. First of all, the limitations of the research will be discussed in subsection 13.1. Next, the recommendations for the Elkerliek are discussed in subsection 13.2. Finally, in subsection 13.3 the recommendations for further research are discussed.

13.1 Limitations of the research

In this paragraph, the limitations of the research are described. First of all, this research does not take acute patients into account. When acute patients arrive at the Elkerliek it is possible that they need to be operated on a time that an elective patient was scheduled. The realization of the schedules found in this thesis is subjected to disturbances due to acute patients. Hence, the results are not directly applicable to operational decisions, but are supposed to provide helpful insights and to be a good starting point for the hospital. Furthermore, for the surgery times of the patient groups, the expected surgery times are used by the model. Creating a schedule for the OR always happens based on a predicted surgery time. In reality, these predicted times can vary quite a bit from the realised surgery times. For the LOS of the patient groups, the distribution of the LOS is used by the model. Still, it is possible that in reality patients stay longer or shorter than expected. Because of the uncertainty, it might be possible that the actually achieved results differ from the ones in this research. However, it is expected that this variation is not high.

Another assumption made is that all surgeons of the same sub-specialism are equal. The only distinction that is made is based on the sub-specialism of the surgeons. When all patients have to be treated by a specific surgeon instead of a specific sub-specialism the model could have more trouble creating a good tactical schedule under the current OR-block schedule. What the model stabilizes is the number of patients that occupy a bed on a certain day. The model assumes the bed occupation to occur as discrete: a bed is either occupied or empty on a certain day. In reality, the bed occupations are not discrete. It is very much possible that a patient leaves the hospital early in the morning. When this happens, the model assumes a bed to be occupied for the entire day, while in reality, the bed will be empty for a big portion of the day.

A final limitation of the research is that patients are considered to only go to one ward. In reality, patients can sometimes be transferred to different wards. For example, patients that stay at the short stay ward can be transferred to the clinical ward when the patient cannot be discharged before Sunday.

13.2 Recommendations for the Elkerliek

Based on this research several things can be recommended for the Elkerliek. First of all, the Elkerliek hospital should take the bed occupation into account when scheduling patients. The current practice of only optimising the OR occupancy leads to a bed occupation that is very variable. It has been proven that taking the bed occupation into account only results in a very small efficiency loss for the OR usage. In order to take the bed occupation into account, the models created for this research can be used as a helpful tool. It will be difficult to change the way the OR is scheduled overnight because it has to be learned how the bed occupation should be taken into consideration. The first step for the Elkerliek would be to consider integrated planning with some pilot cases. This way, the planners as well as the specialists and nurses can get used to the idea and the hospital can move slowly towards a better resource utilisation.

The second recommendation has to do with the patient mix each day. Currently, it often occurs that in an OR-block only one type of patient is scheduled which has bad consequences for the bed occupation. Looking at the solutions the model created, it can be concluded that it is better to have a diverse mix of patients that are operated each day. This mix should consist of patients with a different LOS that stay at different wards. To be able to create a good patient mix, OR planners should get more freedom to determine what patients should be operated on what day. This means that there should be fewer restrictions imposed on them about what types of patients should be scheduled on certain days. The more freedom there is to schedule patients, the easier it becomes to create good schedules. To generate more freedom it is recommended to allow patients to be scheduled in any OR-block allocated to a surgeon that can perform the patients' surgery. This means that surgeons do not only treat 'their own' patients but that they can operate on all patients that they are competent for. Before implementing such actions, this should be communicated with the doctors as well.

The final recommendation is to not change the way the OR-blocks are divided amongst specialisms and sub-specialisms. It has been found that there is some improvement possible when the model allocates the OR-blocks. However, the trouble that the Elkerliek has to go through to change the way the OR-blocks are divided outweighs the small improvements that can be made. Next to the inconvenience for the surgeons that will not operate on fixed days which means that the outpatient-sessions of the surgeons should also be scheduled differently, other specialisms also make use of the ORs. Under the current OR-block schedule there are enough improvements possible.

13.3 Recommendations for future research

In this paragraph, some recommendations for future research will be given. First of all, since this research focusses on the tactical surgical schedule, a good next step would be to study how to go from this tactical schedule to an operational one. This means that also the acute patients should be considered. It can also be found what the effect is of having acute patients and what would be the best way to deal with the uncertainty that comes with them. Secondly, it would be interesting to study how the model used in this research can be used hospital-wide. For this research, only one specialism was considered. However, the beds at the wards are occupied by patients of different specialisms. It would be interesting to find out how all specialisms can be taken into account. It should be noted that when the number of specialisms considered increases, it is likely that the complexity of the model in terms of computational time increases too.

Next, future research can be focused on better ways to deal with the stochastic parameters of the model. For the model, the expected surgery time and an empirical distribution of the LOS of patients are used. Both are still quite rough depictions of the reality. Finding patient-specific surgery time predictions and incorporating good-quality predictions into planning or scheduling algorithms is a promising follow-up topic of the current research. Similar to this, future research could be targeted towards the improvement of the expected surgery time and the expected LOS. Since schedules are created based on the predicted surgery time and LOS it will be beneficial for the quality of the schedule to improve the predictions.

Finally, the model is created to improve the usage of the OR and the bed occupation at different wards. What has not been taken into consideration is the happiness of the surgeons. It has a reason that they impose certain restrictions on the planners. Future research could see how the decisions made by the model influence the happiness of the surgeons and how good schedules can be made whilst keeping the surgeons happy. Surgeons prefer to operate on their 'own' patients for example. Checking how a good schedule can be made whilst keeping requests like this into account could be an interesting topic.

14 References

Adan, I., Bekkers, J., Dellaert, N., Vissers, J., & Yu, X. (2008). Patient mix optimisation and stochastic resource requirements: A case study in cardiothoracic surgery planning. *Health Care Management Science*, 12(2), 129–141. https://doi.org/10.1007/s10729-008-9080-9

Agmon, S. (1954). The Relaxation Method for Linear Inequalities. *Canadian Journal of Mathematics*, *6*, 382–392. https://doi.org/10.4153/cjm-1954-037-2

Bekker, R., & Koeleman, P. M. (2011). Scheduling admissions and reducing variability in bed demand. *Health Care Management Science*, *14*(3), 237–249. https://doi.org/10.1007/s10729-011-9163-x

Beliën, J., & Demeulemeester, E. (2004). *Integer programming for building robust surgery schedules*. Retrieved from https://www.researchgate.net/publication/46429935_Integer_programming_for_building_robust_surgery_schedules

Beliën, J., Demeulemeester, E., & Cardoen, B. (2008). A decision support system for cyclic master surgery scheduling with multiple objectives. Journal of Scheduling, 12(2), 147–161. https://doi.org/10.1007/s10951-008-0086-4

Cikit-learn developers. (n.d.). *Clustering*. Retrieved 31 March 2020, from https://scikit-learn.org/stable/modules/clustering.html

Essays, UK. (November 2018). Linear Programming: Advantages, Disadvantages and Strategies. Retrieved 15 April 2020, from https://www.ukessays.com/essays/management/i-linear-programming.php?vref=1

Essen, J. T. V., Bosch, J. M., Hans, E. W., Houdenhoven, M. V., & Hurink, J. L. (2013). Reducing the number of required beds by rearranging the OR-schedule. *OR Spectrum*. https://doi.org/10.1007/s00291-013-0323-x

Gupta, S., Kumar, R., Lu, K., Moseley, B., & Vassilvitskii, S. (2017). *Local search methods for k-means with outliers*. *Proceedings of the VLDB Endowment, 10(7*), 757–768. https://doi.org/10.14778/3067421.3067425

Ligtvoet, F. (2019, March 28). Ziekenhuizen speelden grote rol in eigen personeelstekort. Retrieved 5 November 2019, from https://nos.nl/nieuwsuur/artikel/2277973-ziekenhuizen-speelden-grote-rol-in-eigen-personeelstekort.html

Liu, N., Truong, V.-A., Wang, X., & Anderson, B. R. (2019). Integrated Scheduling and Capacity Planning with Considerations for Patients' Length-of-Stays. *Production and Operations Management*, 28(7), 1735–1756. https://doi.org/10.1111/poms.13012

Masud, M. A., Huang, J. Z., Wei, C., Wang, J., Khan, I., & Zhong, M. (2018). I-nice: A new approach for identifying the number of clusters and initial cluster centres. *Information Sciences*, *466*, 129–151. https://doi.org/10.1016/j.ins.2018.07.034

McCarl, B. A., & Apland, J. (1986). Validation Of Linear Programming Models. *Southern Journal of Agricultural Economics*, *18*(2), 1–10. https://doi.org/10.1017/S0081305200006208

Meskens, N., Duvivier, D., & Hanset, A. (2013). Multi-objective operating room scheduling considering desiderata of the surgical team. *Decision Support Systems*, *55*(2), 650–659. doi: https://doi.org/10.1016/j.dss.2012.10.019

Moody, J. (2019, September 5). What does RMSE really mean? Retrieved 18 January 2020, from https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48e

Mul, T. (2018). *How the prediction quality of estimated surgery duration change along the process of information collection in a hospital*. Retrieved from https://research.tue.nl/en/studentTheses/how-the-prediction-quality-of-estimated-surgery-duration-change-a

NOS. (2019, 28 maart). Ziekenhuizen speelden grote rol in eigen personeelstekort. Retrieved 13 November 2019, from https://nos.nl/nieuwsuur/artikel/2277973-ziekenhuizen-speelden-grote-rol-in-eigen-personeelstekort.html

Pakhira, M. K., Bandyopadhyay, S., & Maulik, U. (2004). Validity index for crisp and fuzzy clusters. *Pattern Recognition*, *37*(3), 487–501. https://doi.org/10.1016/j.patcog.2003.06.005

Patlolla, C. R. (2018). *Understanding the concept of Hierarchical clustering Technique*. Retrieved 31 march 2020, from https://towardsdatascience.com/understanding-the-concept-of-hierarchical-clustering-technique-c6e8243758ec

Pisař, P. (2019). European SMEs' value management based on controlling, financial analysis and ratios – empirical study. *Investment Management and Financial Innovations*, *16*(4), 277–289. https://doi.org/10.21511/imfi.16(4).2019.24

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

de Rue, J. M. (2007). *Stochastic Programming in Health Care Planning (*student thesis*)*. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjLhubPyO_oAhVODOwKHftkBlgQFjAAegQIAxAB&url=https%3A%2F%2Fwww.math.vu.nl%2F~sbhulai%2Fpapers%2Fpaper-derue.pdf&usg=AOvVaw1TuDnXKbQuXRFY9dwNbpNU

Steinley, Douglas. (2006). K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology*, *59*(1), 1–34. https://doi.org/10.1348/000711005x48266

Thacker, Ben & S.W.Doebling, & Hemez, Francois & Anderson, Mark & Pepin, J.E. & Rodriguez, Edward. (2004). Concepts of Model Verification and Validation. 10.2172/835920.

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, *18*(4), 267–276. https://doi.org/10.1007/bf02289263

U.S. Department of Commerce. (2013, 10 maart). *What are outliers in the data?* Retrieved 4 March 2020, from https://www.itl.nist.gov/div898/handbook/prc/section1/prc16.htm

Vali-Siar, M. M., Gholami, S., & Ramezanian, R. (2018). Multi-period and multi-resource operating room scheduling under uncertainty: A case study. *Computers & Industrial Engineering*, 126, 549–568. doi: https://doi.org/10.1016/j.cie.2018.10.014

Vissers, J. M. H., & Beech, R. (2005). *Health operations management : patient flow logistics in health care*. (Routledge health management series). London: Routledge Taylor & Francis Group.

Appendix A: Features of the datasets

Table 48: Features patient admission dataset

Features patient admission dataset		
Category	**Data feature**	**Type**
Patient identification	Patient number	Categorical
	Admission number	Categorical
Length of stay (LOS)	Date and time of admission	Numerical – continuous
	Date and time of discharge	Numerical – continuous
	Warm bed minutes	Numerical – discrete
	Actual nursing days	Numerical – discrete
Patient location	Department	Categorical
	Room number	Categorical
	Bed number	Categorical
Practitioner of the patient	Practitioner code	Categorical
	Specialism	Categorical
	Sub-specialism	Categorical
Reason for admission	Elective or acute	Categorical
	Diagnosis code	Categorical
	Diagnosis description	Categorical
	Treatment description	Categorical
DBC code	DBC code	Categorical
	DBC start-date	Numerical – continuous
	DBC end-date	Numerical – continuous
Patient personal information	Date of birth	Numerical – continuous
	Gender	Categorical
	Postal code	Categorical

Table 49: Features surgery dataset

Features surgery dataset		
Category	**Data feature**	**Type**
Patient identification	Patient number	Categorical
	Admission number	Categorical
	Surgery number	Categorical
Surgery location and OR-block	OR code	Categorical
	Location code (Helmond or Deurne)	Categorical
	OR session code	Categorical
	OR daypart code	Categorical
Practitioner of the patient	Specialism	Categorical
	Specialist	Categorical
	Number of assistants	Numerical – discrete
	Amount of substituted	Numerical – discrete
Treatment of the patient	OR surgery code	Categorical
	Treatment in words	Categorical
	Used anesthesia	Categorical
	Wellbeing of the patient	Categorical
	Urgency	Categorical
	DBC code	Categorical
Planned occupation of resources	Planned surgery time	Numerical – continuous
	Planned occupation of the OR	Numerical – continuous
Surgery date	Date the surgery was requested	Numerical – continuous
	Planned surgery date	Numerical – continuous
	Actual surgery date	Numerical – continuous
Date and time (DT) OR process steps take place	DT transport ward to OR starts	Numerical – continuous
	DT transport ward to OR done	Numerical – continuous
	DT patient enters the OR	Numerical – continuous
	DT sedating of patient starts	Numerical – continuous
	DT sedating of patient done	Numerical – continuous
	DT first incision	Numerical – continuous
	DT surgeon finished	Numerical – continuous
	DT patient leaves OR	Numerical – continuous
	DT recovery at the recovery room starts	Numerical – continuous
	DT recovery at the recovery room done	Numerical – continuous
	DT transport from OR to ward starts	Numerical – continuous

Appendix B: Histograms of the surgery times per specialism per ward

Vascular sub-specialism

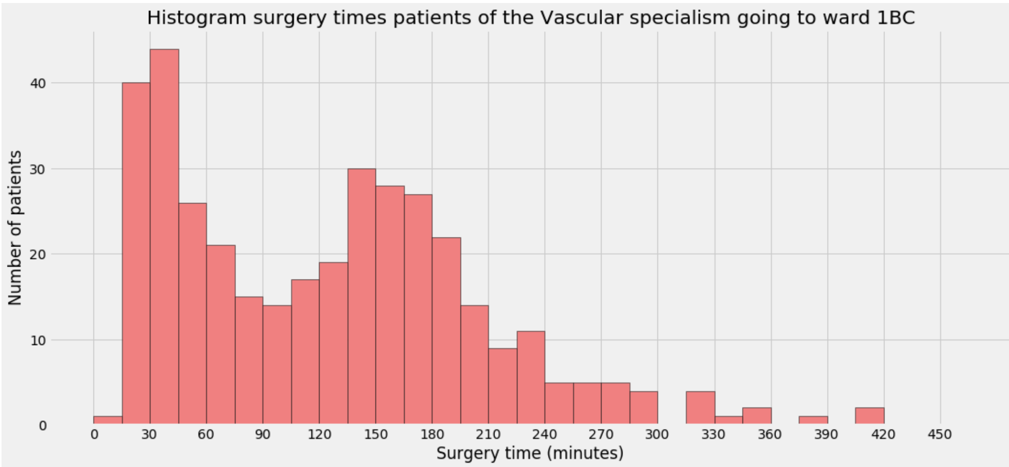

Figure 19: Histogram surgery times Vascular sub-specialism/1BC

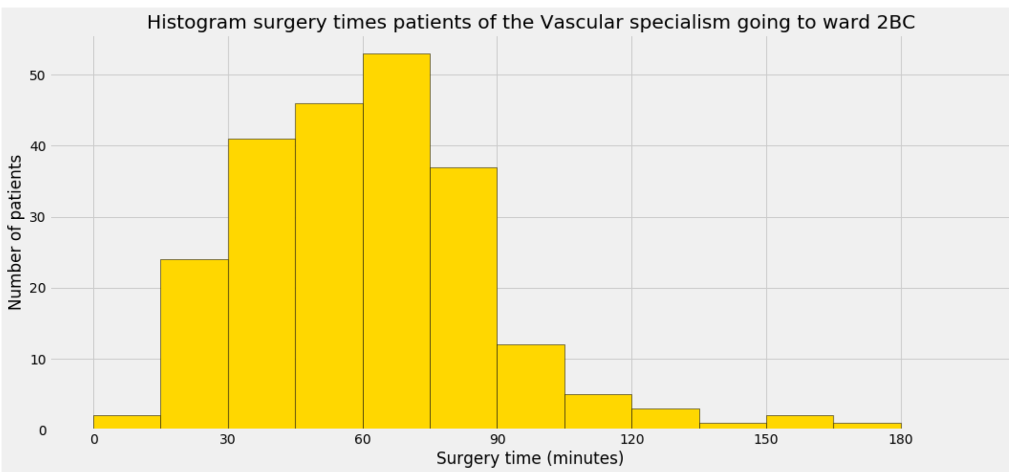

Figure 20: Histogram surgery times Vascular sub-specialism/2BC

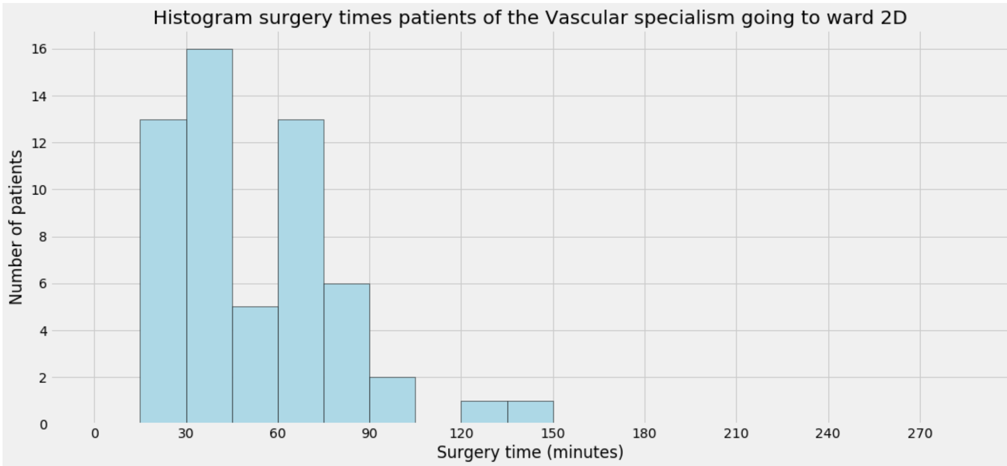

Figure 21: Histogram surgery times Vascular sub-specialism/2D

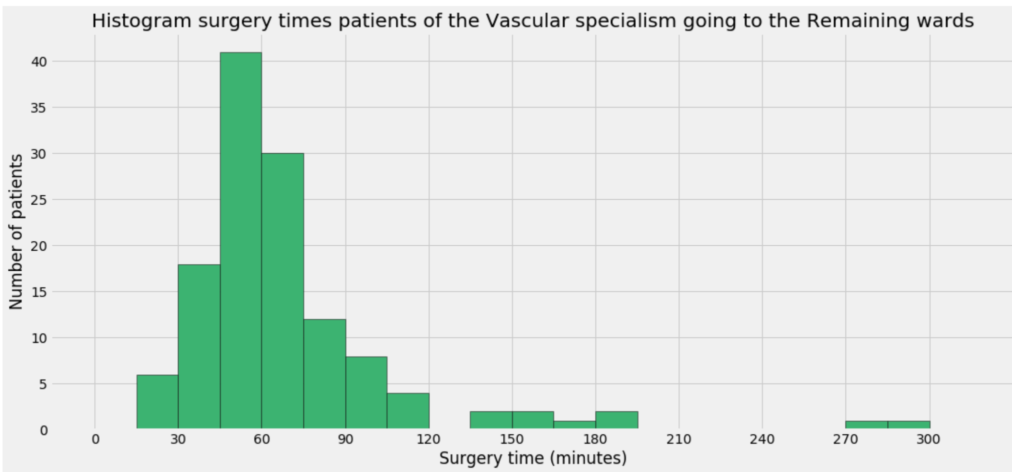

Figure 22: Histogram surgery times Vascular sub-specialism/remaining wards

Trauma sub-specialism

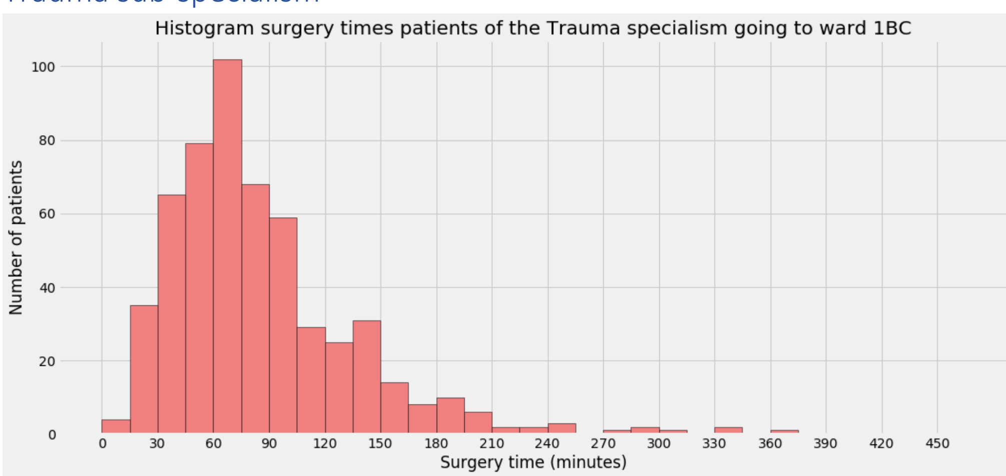

Figure 23: Histogram surgery times Trauma sub-specialism/1BC

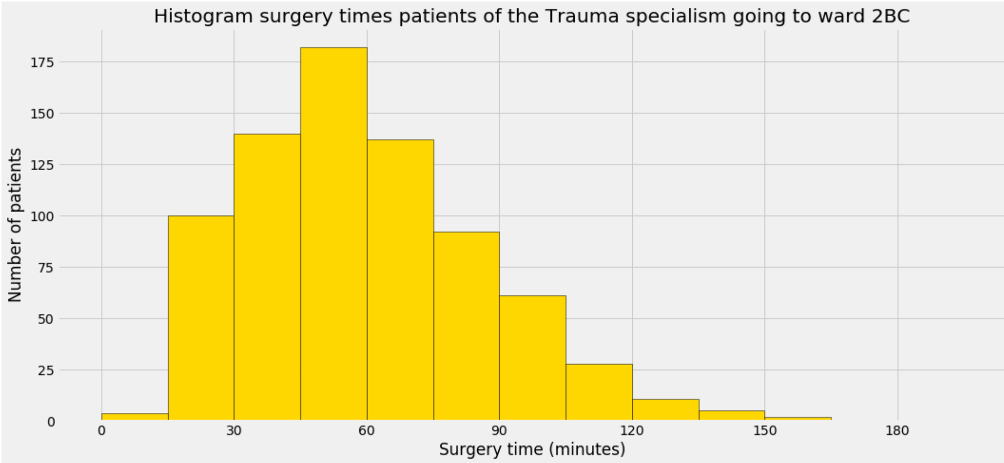

Figure 24: Histogram surgery times Trauma sub-specialism/2BC

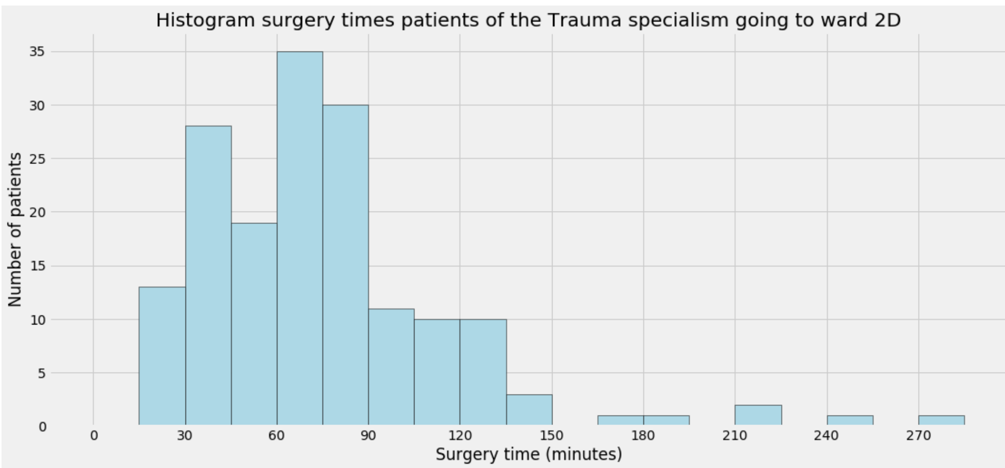

Figure 25: Histogram surgery times Trauma sub-specialism/2D

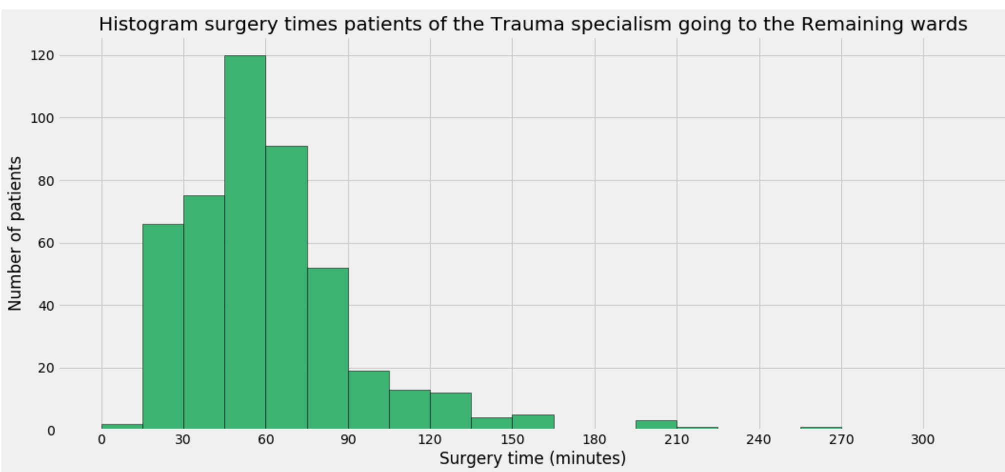

Figure 26: : Histogram surgery times Trauma sub-specialism/remaining wards

General sub-specialism

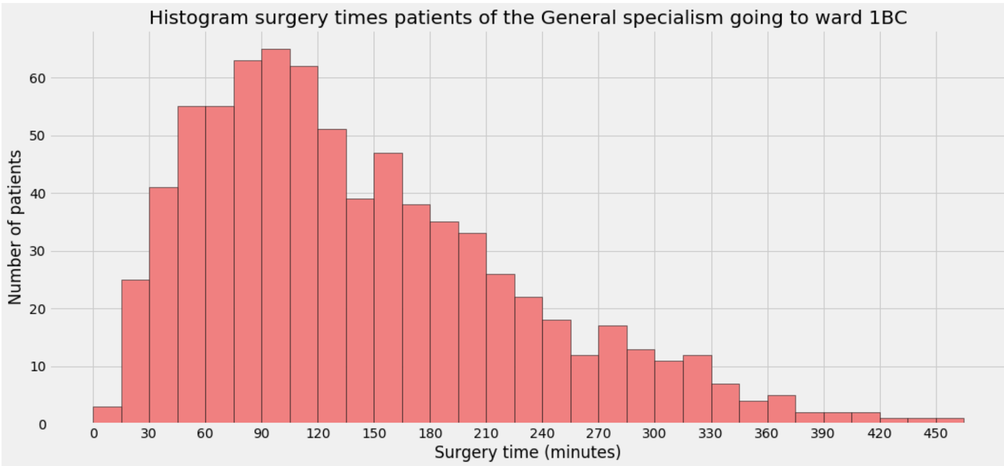

Figure 27: Histogram surgery times General sub-specialism/1BC

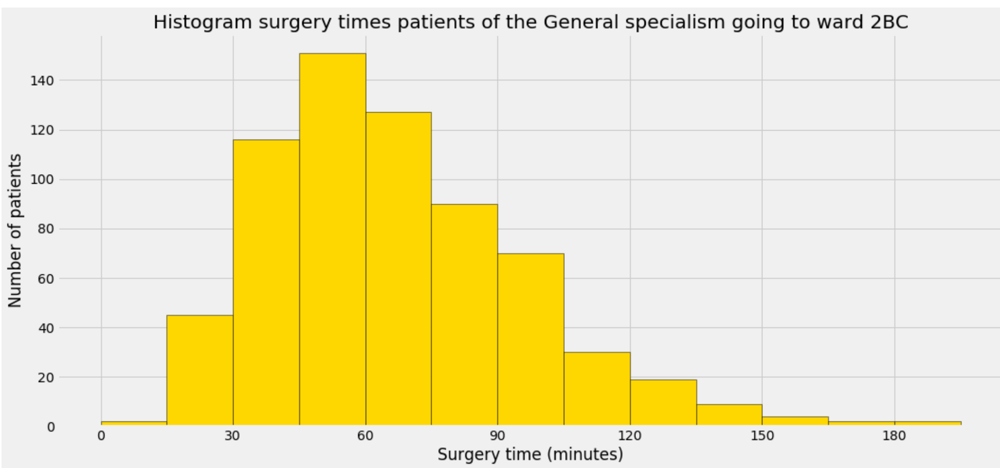

Figure 28: Histogram surgery times General sub-specialism/2BC

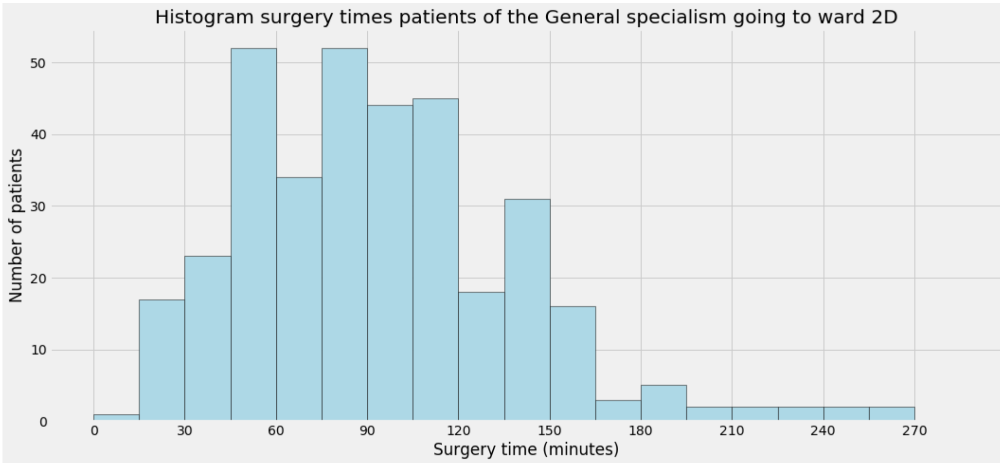

Figure 29: Histogram surgery times General sub-specialism/2D

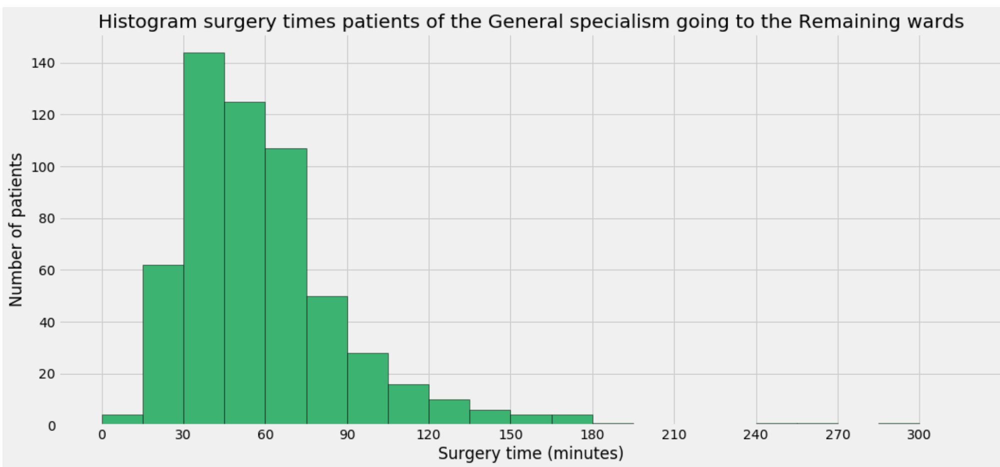

Figure 30: Histogram surgery times General sub-specialism/remaining wards

Appendix C: Histograms of the LOS

Vascular sub-specialism

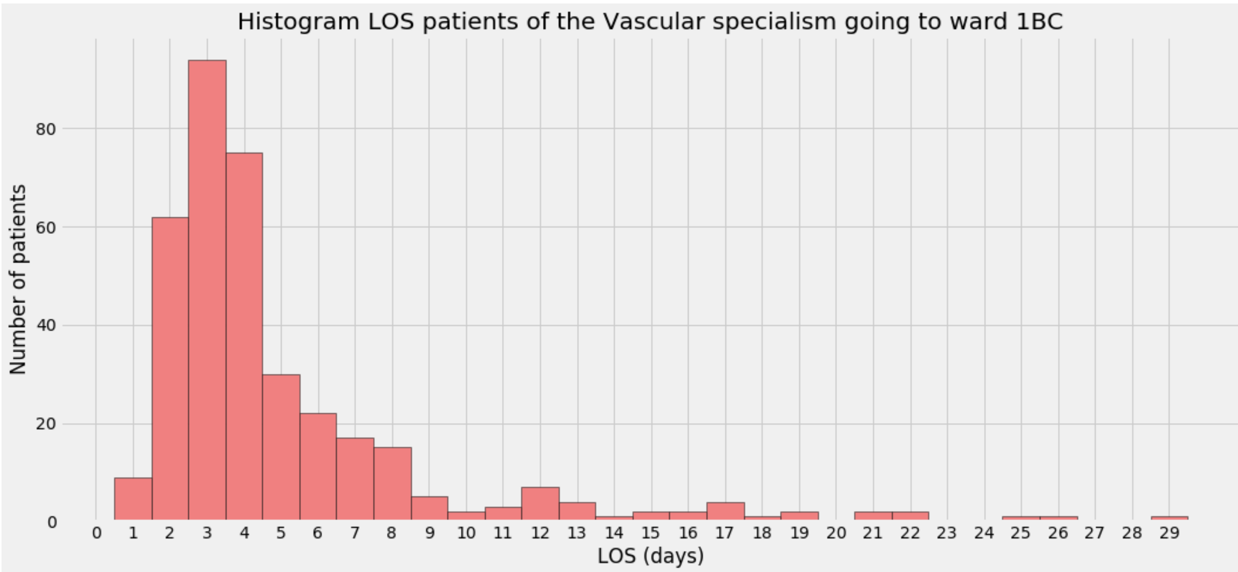

Figure 31: Histogram of the LOS of the Vascular sub-specialism/1BC

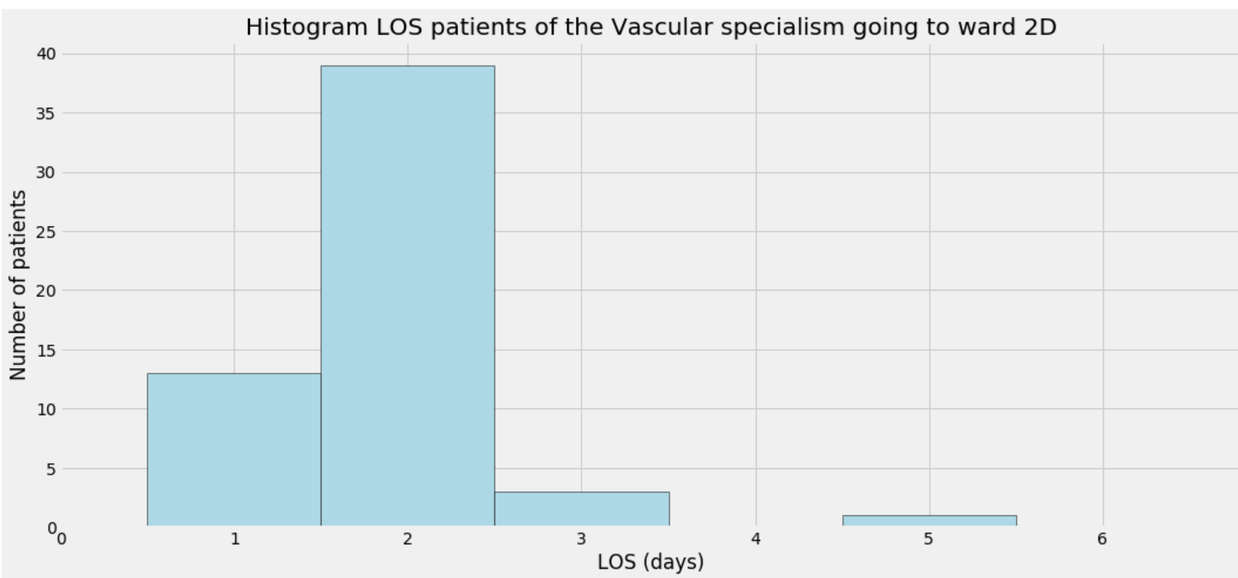

Figure 32: Histogram of the LOS of the Vascular sub-specialism/2D

Trauma sub-specialism

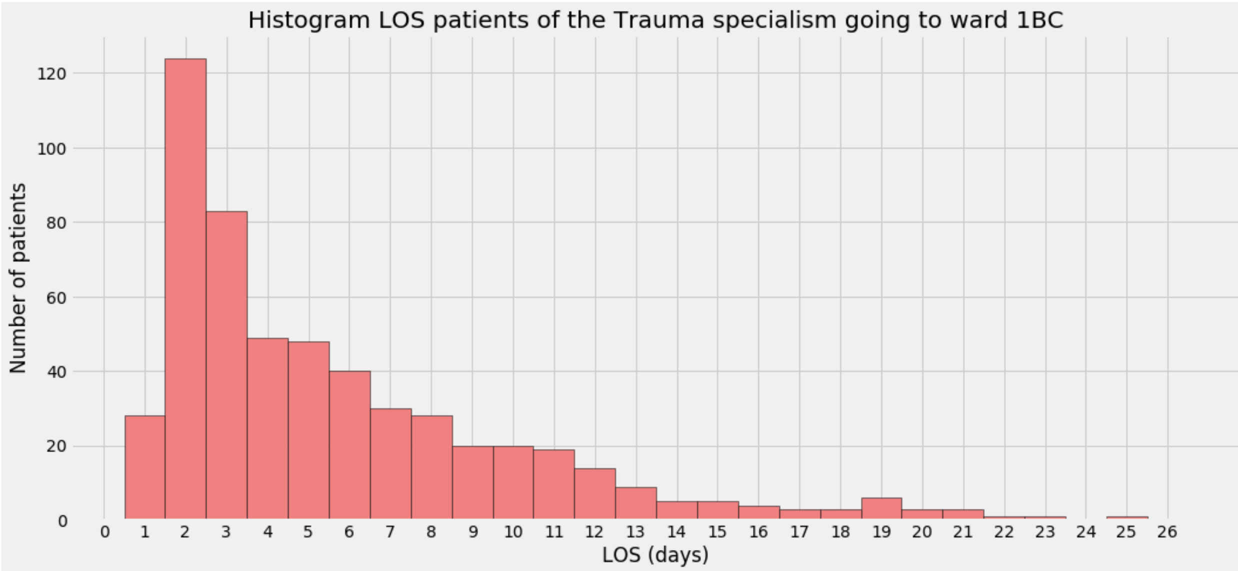

Figure 33: Histogram of the LOS of the Trauma sub-specialism/1BC

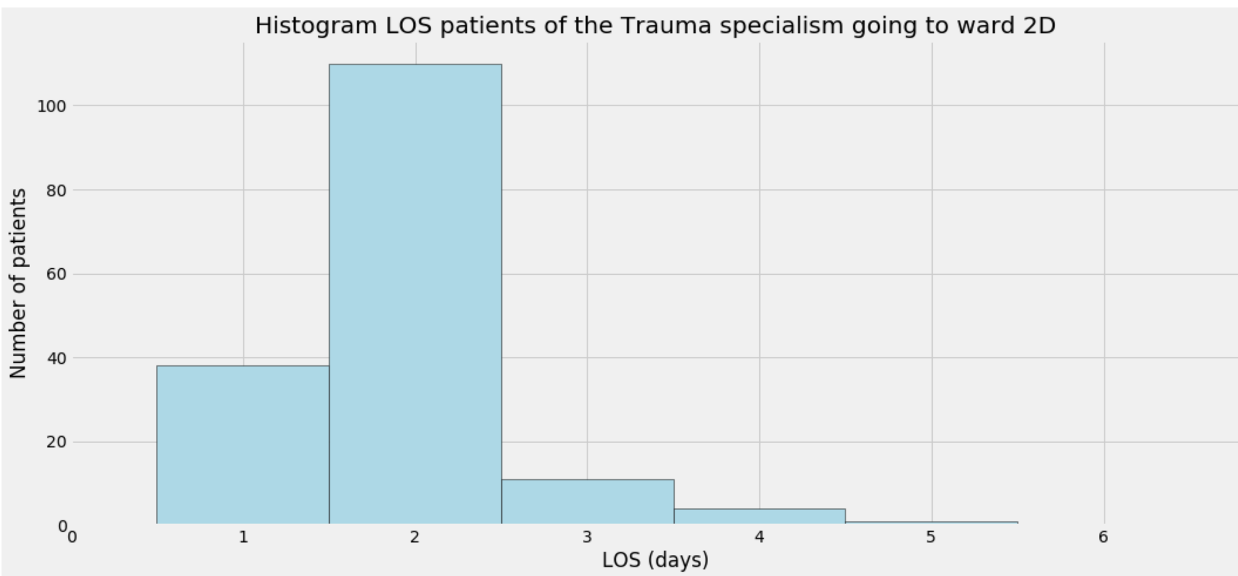

Figure 34: Histogram of the LOS of the Trauma sub-specialism/2D

General sub-specialism

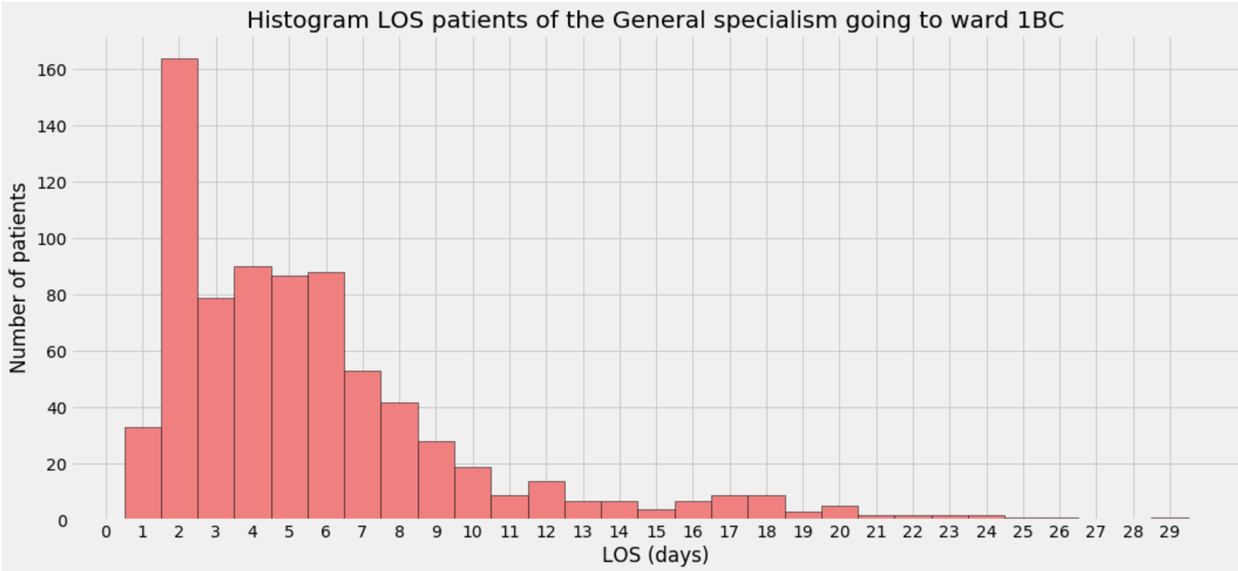

Figure 35: Histogram of the LOS of the General sub-specialism/1BC

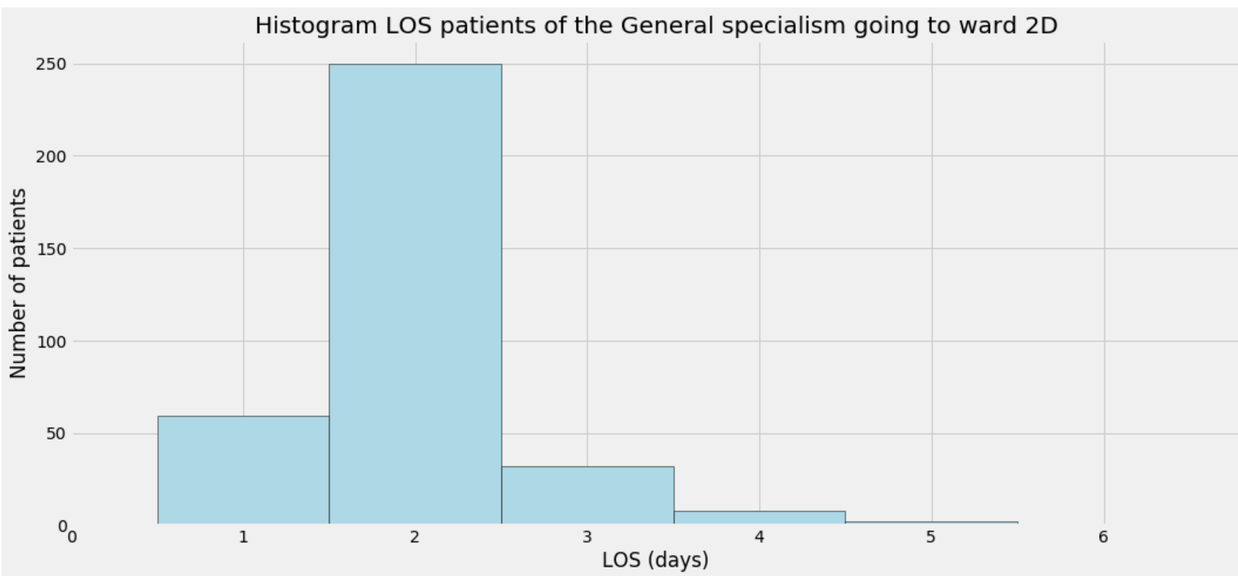

Figure 36: Histogram of the LOS of the General sub-specialism/2D

Appendix D: Boxplots surgery times and LOS

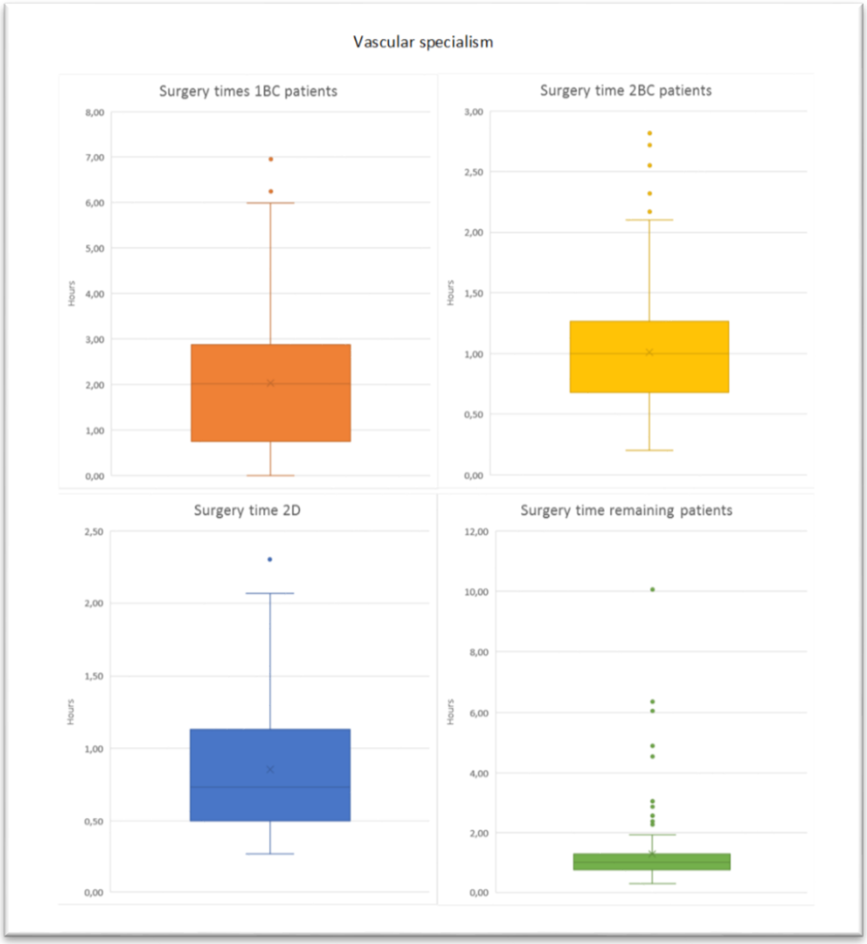

Figure 37: Boxplots surgery time of the vascular sub-specialism

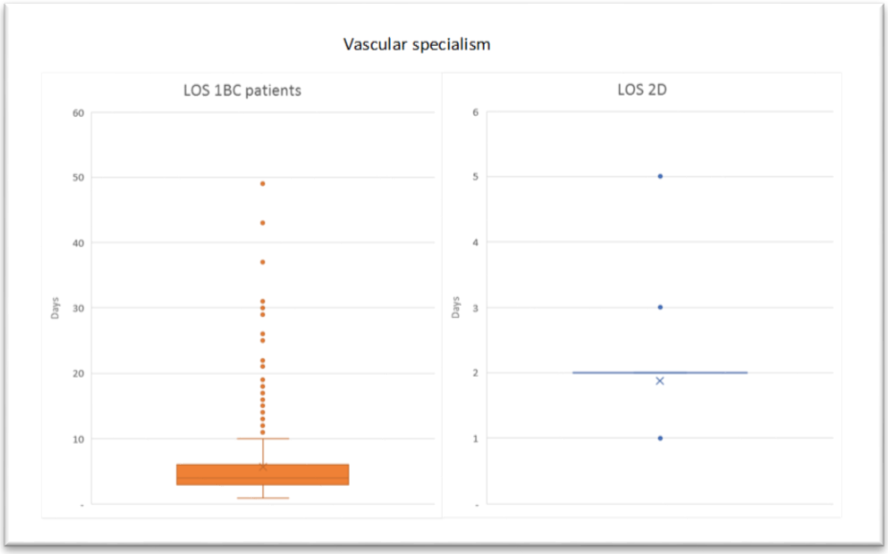

Figure 38: Boxplots LOS of the vascular sub-specialism

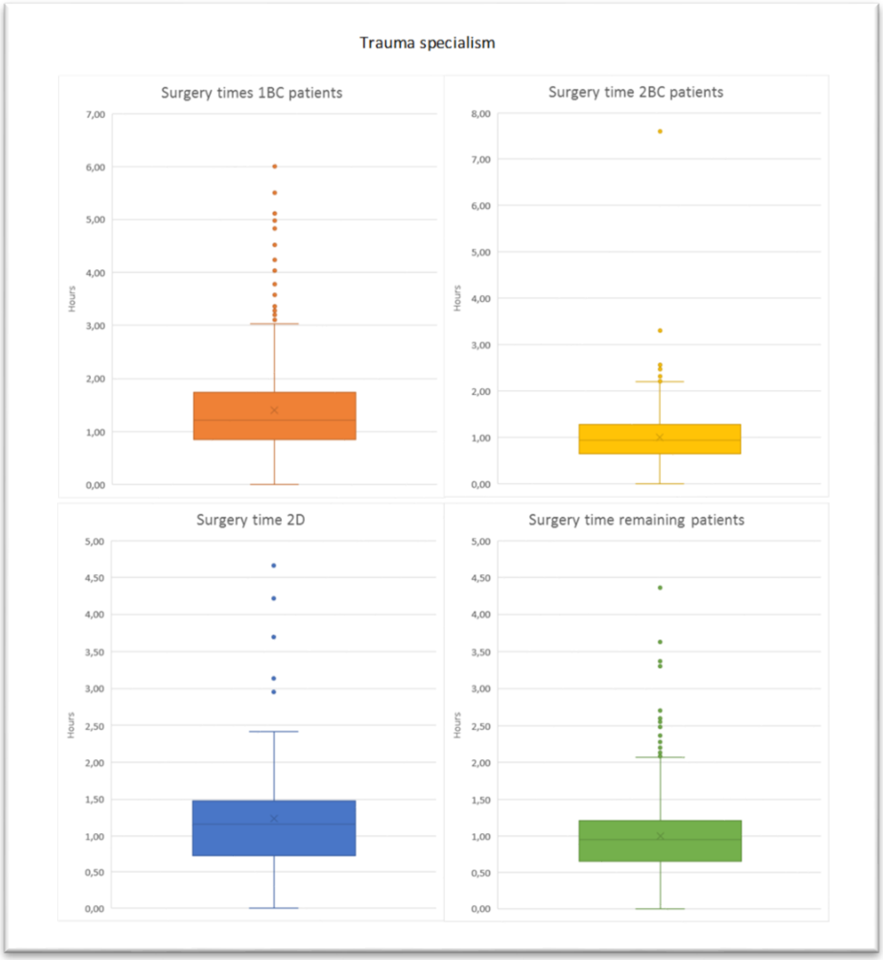

Figure 39: Boxplots surgery times of the trauma sub-specialism

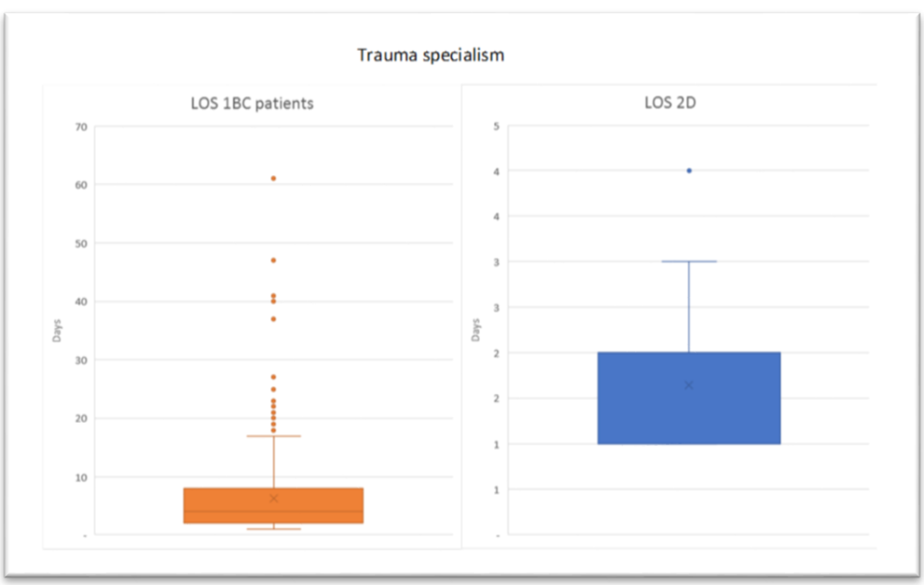

Figure 40: Boxplots LOS of the trauma sub-specialism

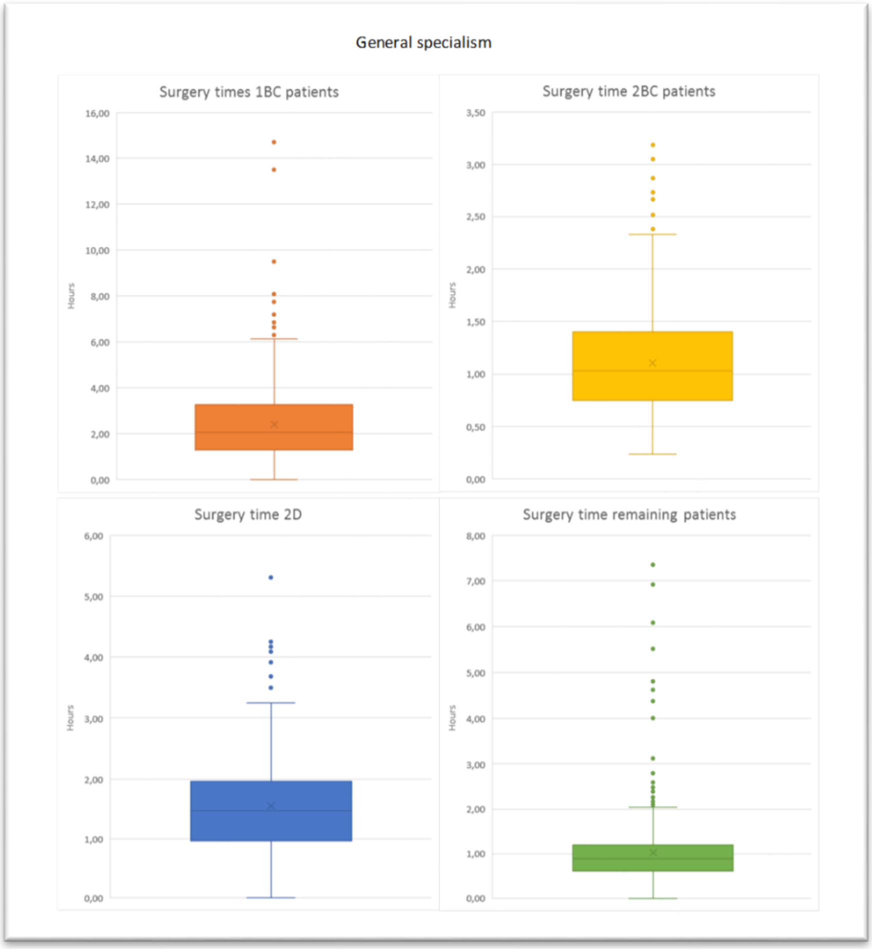

Figure 41: Boxplots surgery times of the general sub-specialism

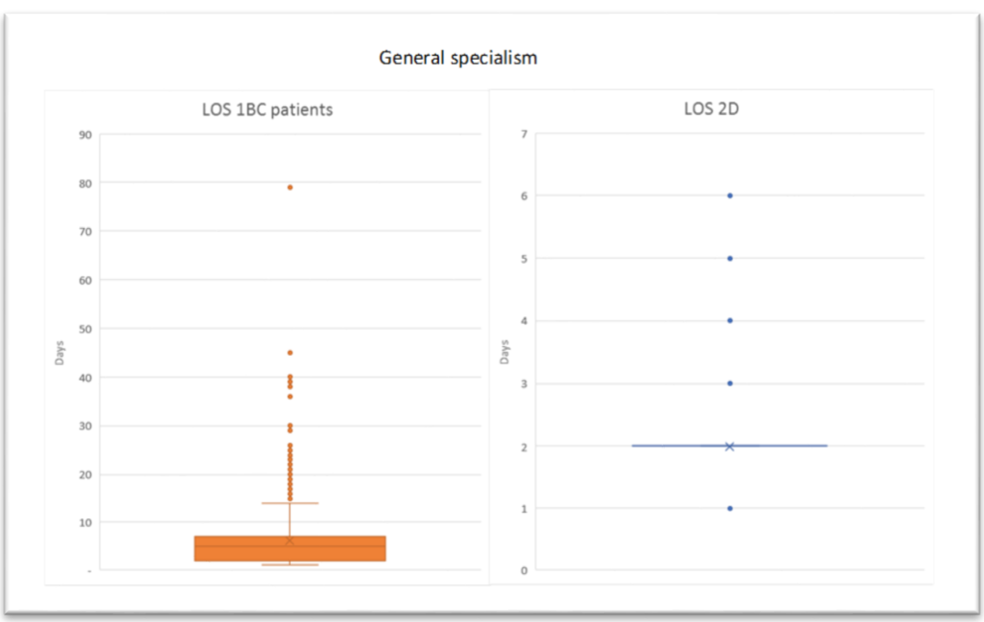

Figure 42: Boxplots LOS of the general sub-specialism

Appendix E: Distribution of the LOS for 1BC and 2D patients

Table 50: Distribution of the LOS for 1BC patients

Probability of occupying a bed *k* days after surgery															
Group nr.	*k*														
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	100%	92%	42%	16%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
2	100%	100%	100%	100%	100%	82%	64%	42%	15%	3%	0%	0%	0%	0%	0%
3	100%	100%	94%	46%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
4	100%	100%	100%	100%	94%	62%	33%	13%	6%	2%	0%	0%	0%	0%	0%
10	100%	88%	42%	20%	10%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
11	100%	100%	100%	100%	100%	100%	76%	56%	35%	20%	7%	4%	2%	1%	1%
12	100%	98%	81%	56%	35%	16%	6%	0%	0%	0%	0%	0%	0%	0%	0%
13	100%	100%	100%	100%	100%	100%	100%	100%	92%	85%	73%	48%	29%	15%	8%
21	100%	88%	33%	14%	4%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
22	100%	100%	95%	84%	60%	36%	12%	0%	0%	0%	0%	0%	0%	0%	0%
23	100%	100%	100%	100%	100%	100%	100%	100%	74%	52%	35%	27%	14%	8%	0%
24	100%	100%	100%	98%	93%	77%	53%	33%	16%	9%	5%	3%	1%	0%	0%

Table 51: Distribution of the LOS for 2D patients

Probability of occupying a bed *k* days after surgery				
Group nr.	*k*			
	0	1	2	3
8	100%	75%	0%	0%
16	100%	77%	6%	0%
17	100%	85%	9%	0%
28	100%	74%	9%	0%
29	100%	93%	9%	0%

Appendix F: Stochastic parameter implementations

F.1 Chance constraints

The first way to deal with stochastic parameters is with the use of chance constraints. First, a reliability level must be determined. This reliability level can be determined by the problem owner. For the constraint that ensures that the overtime of specialist c on day d is lower than the maximum overtime:

$y_{c,d} * MT \geq ouc_{c,d}, \qquad\qquad c = 1, \ldots, C, \quad d = 1, \ldots, D,$

The corresponding chance constraint with a reliability level of 95% becomes:

$P * \left(y_{c,d} * MT \geq \widetilde{ouc}_{c,d}\right) \geq 95\%, \qquad\qquad c = 1, \ldots, C, \quad d = 1, \ldots, D,$

This chance constraint means that in 95% of the cases, random variable $\widetilde{ouc}_{c,d}$ (overtime of specialist c on day d) should be less than the maximum allowed overtime. The advantage of chance constraints is that it can be specified by the problem owner how much risk is allowed. The disadvantage of this method is that the problem quickly becomes too big to solve when the number of random variables increases. Kall (1991) states that considering *N* random independent variables and *i* realizations for each variable, the number of different possible realizations is i^N. For the constraint described in this paragraph, the parameter $ou_{c,d}$ includes the sum of the surgery time of all patient groups that can be treated by a specific sub-specialist. When these surgery times are represented by 5 descriptive points this means that having only 10 different patient groups leads to more than 9,5 million different realisations of $ou_{c,d}$ that need to be considered. Describing the surgery times of 10 groups with only 3 descriptive points leads to about 60.000 different realisations of parameter $ou_{c,d}$. In both cases, the models will become too big to solve for the number of groups that is needed to be useful for the Elkerliek.

F.2 Resource models

Resource models are another way to deal with stochastics. For this method, a slack variable is introduced. This variable denotes how much the maximum value is exceeded. Resource models allow an exceeding of the maximum capacity, but make it very costly to do so (Sen & Higle, 1999, p. 41). Take the same constraint as in the previous paragraph that ensures that the overtime of specialist c on day d is lower than the maximum overtime:

$y_{c,d} * MT \geq ouc_{c,d}, \qquad\qquad c = 1, \ldots, C, \quad d = 1, \ldots, D,$

The corresponding resource model constraint becomes:

$y_{c,d} * MT \geq \widetilde{ouc}_{c,d,i} - Z_i, \qquad\qquad c = 1, \ldots, C, \quad d = 1, \ldots, D, \quad i = 1, \ldots, I,$

In this new constraint the slack variable Z_i is introduced. This variable denotes the overuse of the OR for realisation *i* of variable $\widetilde{ouc}_{c,d,i}$. What this means is that when there are 10 different possible realisations for $\widetilde{ouc}_{c,d}$, slack variables Z_1 to Z_{10} will denote the overuse of the OR for each of the realisations. These slack variables are then added to the objective function where each Z_i is multiplied with chance P_i (the chance that $\widetilde{ouc}_{c,d,i}$ happens) and a cost for exceeding the maximum capacity. Again, all possible realizations of the variable $ou_{c,d}$ must be considered. This means that this method too will become too big when the number of groups increases.

Both chance constraints and the resource model methods have been implemented into the group model described in chapter 9.2. The maximum number of groups that could be considered before the model became too big to solve was 6 groups for 5 descriptive points per group and 8 groups for 3 descriptive points per group.

F.3 Convolutional sum

Another way to deal with stochastics is to combine the distributions of all random variables into one distribution. For example, the three different distribution of three surgeries in one OR-block will be combined into one distribution for the total surgery time. This single distribution can be used in the methods described above. Meaning that only one random variable needs to be considered. The problem here is that calculating the convolution itself is very time consuming. The number of calculations needed to come to the convolutional sum grows with the same speed as the methods above. It is possible to do offline calculations to get the convolutional sum for different number of patients of the same patient group. For example, it can be calculated before running the model what the convolutional sum is when 3 patients of the same group are operated on one day. To get the total distribution of all patient groups scheduled on a certain day for a specific specialism, all possible combinations of surgery times for those groups should be considered again. Having i points that describe the distribution of the surgery times of each group, the number of calculations needed to get to the convolutional sum is i^N where N is the number of groups.

Appendix G: Instances to test the performance of the model

Table 52: Instances for performance test 1

Instances for performance test 1			
Data from weeks	Number of patients	Time horizon	Max runtime
Week 1	34	1 week	15 minutes
Week 2	116	1 week	15 minutes
Week 3	118	1 week	15 minutes
Week 4	110	1 week	15 minutes
Week 5	90	1 week	15 minutes
Week 6	114	1 week	15 minutes
Week 7	84	1 week	15 minutes
Week 8	120	1 week	15 minutes
Week 9	96	1 week	15 minutes
Week 10	80	1 week	15 minutes
Week 11	114	1 week	15 minutes
Week 12	100	1 week	15 minutes
Week 13	94	1 week	15 minutes
Week 14	93	1 week	15 minutes
Week 15	118	1 week	15 minutes
Week 16	104	1 week	15 minutes
Week 17	51	1 week	15 minutes
Week 18	57	1 week	15 minutes
Week 19	95	1 week	15 minutes
Week 20	107	1 week	15 minutes
Week 21	85	1 week	15 minutes
Week 22	103	1 week	15 minutes
Week 23	115	1 week	15 minutes
Week 24	69	1 week	15 minutes
Week 25	97	1 week	15 minutes
Week 26	113	1 week	15 minutes
Week 27	91	1 week	15 minutes
Week 28	60	1 week	15 minutes
Week 29	68	1 week	15 minutes
Week 30	55	1 week	15 minutes
Week 31	50	1 week	15 minutes
Week 32	44	1 week	15 minutes
Week 33	52	1 week	15 minutes
Week 34	116	1 week	15 minutes
Week 35	100	1 week	15 minutes
Week 36	106	1 week	15 minutes
Week 37	93	1 week	15 minutes
Week 38	82	1 week	15 minutes
Week 39	52	1 week	15 minutes
Week 40	52	1 week	15 minutes
Week 41	53	1 week	15 minutes
Week 42	28	1 week	15 minutes
Week 43	51	1 week	15 minutes
Week 44	51	1 week	15 minutes
Week 45	50	1 week	15 minutes
Week 46	50	1 week	15 minutes
Week 47	54	1 week	15 minutes
Week 48	54	1 week	15 minutes
Week 49	43	1 week	15 minutes
Week 50	60	1 week	15 minutes
Week 51	63	1 week	15 minutes
Week 52	17	1 week	15 minutes

Table 53: Instances for performance test 2

Instances performance test 2			
Data from weeks	**number of patients**	**Time horizon**	**Max runtime**
Week 1 and 2	150	2 weeks	15 minutes
Week 2 and 3	234	2 weeks	15 minutes
Week 3 and 4	228	2 weeks	15 minutes
Week 4 and 5	200	2 weeks	15 minutes
Week 5 and 6	204	2 weeks	15 minutes
Week 6 and 7	198	2 weeks	15 minutes
Week 7 and 8	204	2 weeks	15 minutes
Week 8 and 9	216	2 weeks	15 minutes
Week 9 and 10	176	2 weeks	15 minutes
Week 10 and 11	194	2 weeks	15 minutes
Week 11 and 12	214	2 weeks	15 minutes
Week 12 and 13	194	2 weeks	15 minutes
Week 13 and 14	187	2 weeks	15 minutes
Week 14 and 15	211	2 weeks	15 minutes
Week 15 and 16	222	2 weeks	15 minutes
Week 16 and 17	155	2 weeks	15 minutes
Week 17 and 18	108	2 weeks	15 minutes
Week 18 and 19	152	2 weeks	15 minutes
Week 19 and 20	202	2 weeks	15 minutes
Week 20 and 21	192	2 weeks	15 minutes
Week 21 and 22	188	2 weeks	15 minutes
Week 22 and 23	218	2 weeks	15 minutes
Week 23 and 24	184	2 weeks	15 minutes
Week 24 and 25	166	2 weeks	15 minutes
Week 25 and 26	210	2 weeks	15 minutes
Week 26 and 27	204	2 weeks	15 minutes
Week 27 and 28	151	2 weeks	15 minutes
Week 28 and 29	128	2 weeks	15 minutes
Week 29 and 30	123	2 weeks	15 minutes
Week 30 and 31	105	2 weeks	15 minutes
Week 31 and 32	94	2 weeks	15 minutes
Week 32 and 33	96	2 weeks	15 minutes
Week 33 and 34	168	2 weeks	15 minutes
Week 34 and 35	216	2 weeks	15 minutes
Week 35 and 36	206	2 weeks	15 minutes
Week 36 and 37	199	2 weeks	15 minutes
Week 37 and 38	175	2 weeks	15 minutes
Week 38 and 39	134	2 weeks	15 minutes
Week 39 and 40	104	2 weeks	15 minutes
Week 40 and 41	105	2 weeks	15 minutes
Week 41 and 42	81	2 weeks	15 minutes
Week 42 and 43	79	2 weeks	15 minutes
Week 43 and 44	102	2 weeks	15 minutes
Week 44 and 45	101	2 weeks	15 minutes
Week 45 and 46	100	2 weeks	15 minutes
Week 46 and 47	104	2 weeks	15 minutes
Week 47 and 48	108	2 weeks	15 minutes
Week 48 and 49	97	2 weeks	15 minutes
Week 49 and 50	103	2 weeks	15 minutes
Week 50 and 51	123	2 weeks	15 minutes
Week 51 and 52	80	2 weeks	15 minutes
Week 52 and 1	51	2 weeks	15 minutes

Table 54: Instances for performance test 3

Instances performance test 3			
Data from weeks	number of patients	Time horizon	Max runtime
Week 1, 2 and 3	268	3 weeks	30 minutes
Week 4, 5 and 6	314	3 weeks	30 minutes
Week 7, 8 and 9	300	3 weeks	30 minutes
Week 10, 11 and 12	294	3 weeks	30 minutes
Week 13, 14 and 15	305	3 weeks	30 minutes
Week 16, 17 and 18	212	3 weeks	30 minutes
Week 19, 20 and 21	287	3 weeks	30 minutes
Week 22, 23 and 24	287	3 weeks	30 minutes
Week 25, 26 and 27	301	3 weeks	30 minutes
Week 28, 29 and 30	183	3 weeks	30 minutes
Week 31, 32 and 33	146	3 weeks	30 minutes
Week 34, 35 and 36	322	3 weeks	30 minutes
Week 37, 38 and 39	227	3 weeks	30 minutes
Week 40, 41 and 42	133	3 weeks	30 minutes
Week 43, 44 and 45	152	3 weeks	30 minutes
Week 46, 47 and 48	158	3 weeks	30 minutes
week 49, 50 and 51	166	3 weeks	30 minutes

Table 55: Instances for performance test 4

Instances performance test 4			
Data from month	number of patients	Time horizon	Max runtime
January	215	4 weeks	1 hour
February	204	4 weeks	1 hour
March	218	4 weeks	1 hour
April	185	4 weeks	1 hour
May	199	4 weeks	1 hour
June	207	4 weeks	1 hour
July	163	4 weeks	1 hour
August	159	4 weeks	1 hour
September	185	4 weeks	1 hour
October	213	4 weeks	1 hour
November	230	4 weeks	1 hour
December	190	4 weeks	1 hour

Table 56: Instances for performance test 5

Instances for performance test 5			
Data from weeks	Number of patients	Time horizon	Max runtime
January	215	4 weeks	30 minutes
March	218	4 weeks	30 minutes
June	207	4 weeks	30 minutes
August	159	4 weeks	30 minutes
October	213	4 weeks	30 minutes
November	230	4 weeks	30 minutes

Table 57: Instances for performance test 6

Instances for performance test 6			
Data from weeks	Number of patients	Time horizon	Max runtime
January	215	4 weeks	1 hour
March	218	4 weeks	1 hour
June	207	4 weeks	1 hour
August	159	4 weeks	1 hour
October	213	4 weeks	1 hour
November	230	4 weeks	1 hour

Table 58: Instances for performance test 7

Instances performance test 7				
Data from month	number of patients	Time horizon	Max runtime	Relaxed decision variables
January	215	4 weeks	1 hour	x_{gd}
February	204	4 weeks	1 hour	x_{gd}
March	218	4 weeks	1 hour	x_{gd}
April	185	4 weeks	1 hour	x_{gd}
May	199	4 weeks	1 hour	x_{gd}
June	207	4 weeks	1 hour	x_{gd}
July	163	4 weeks	1 hour	x_{gd}
August	159	4 weeks	1 hour	x_{gd}
September	185	4 weeks	1 hour	x_{gd}
October	213	4 weeks	1 hour	x_{gd}
November	230	4 weeks	1 hour	x_{gd}
December	190	4 weeks	1 hour	x_{gd}

Table 59: Instances for performance test 8

Instances performance test 8				
Data from month	number of patients	Time horizon	Max runtime	Relaxed decision variables
January	215	4 weeks	1 hour	x_{gd}, y_{cd}
February	204	4 weeks	1 hour	x_{gd}, y_{cd}
March	218	4 weeks	1 hour	x_{gd}, y_{cd}
April	185	4 weeks	1 hour	x_{gd}, y_{cd}
May	199	4 weeks	1 hour	x_{gd}, y_{cd}
June	207	4 weeks	1 hour	x_{gd}, y_{cd}
July	163	4 weeks	1 hour	x_{gd}, y_{cd}
August	159	4 weeks	1 hour	x_{gd}, y_{cd}
September	185	4 weeks	1 hour	x_{gd}, y_{cd}
October	213	4 weeks	1 hour	x_{gd}, y_{cd}
November	230	4 weeks	1 hour	x_{gd}, y_{cd}
December	190	4 weeks	1 hour	x_{gd}, y_{cd}

Appendix H: Scores for KPI 1 and KPI 2 for the actual schedule used by the Elkerliek using the expected surgery times

Table 60: Scores KPI 1 and KPI 2 for the actual schedule and expected surgery times

Year and month		Score KPI 1 for the actual schedule based on the predicted surgery times	Score KPI 2 for the actual schedule based on the predicted surgery times
2018	January	76,3%	9,36
	February	78,8%	12,43
	March	77,1%	12,24
	April	73,6%	6,01
	May	77,2%	15,96
	June	77,6%	14,59
	July	74,5%	11,18
	August	72,0%	5,93
	September	76,0%	9,33
	October	79,8%	10,01
	November	81,1%	12,26
	December	81,3%	15,07
2018 Average		**77,1%**	**11,24**
2019	January	82,6%	10,3
	February	79,8%	13,5
	March	81,2%	11,6
	April	82,7%	9,80
	May	78,1%	9,41
	June	76,7%	9,76
	July	73,8%	6,81
	August	73,8%	8,32
	September	79,7%	12,46
2019 Average		**78,9%**	**10,25**

Appendix I: Results for scenario 1 to 6

H.1 Scenario 1

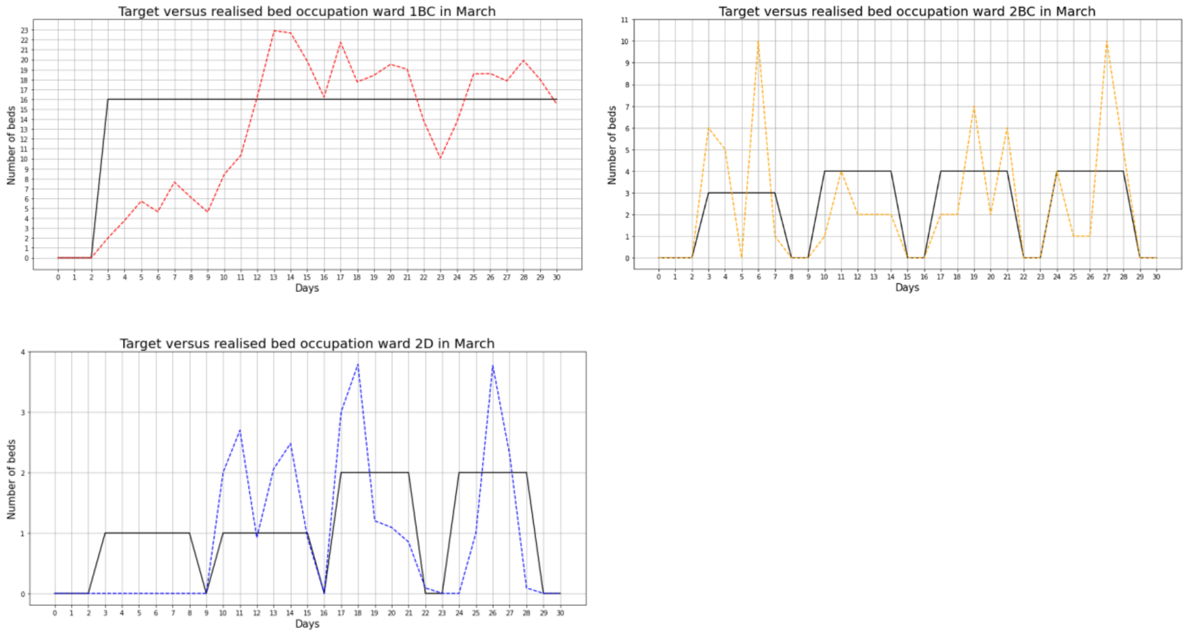

Figure 43: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month March under the scheduling policy of scenario 1

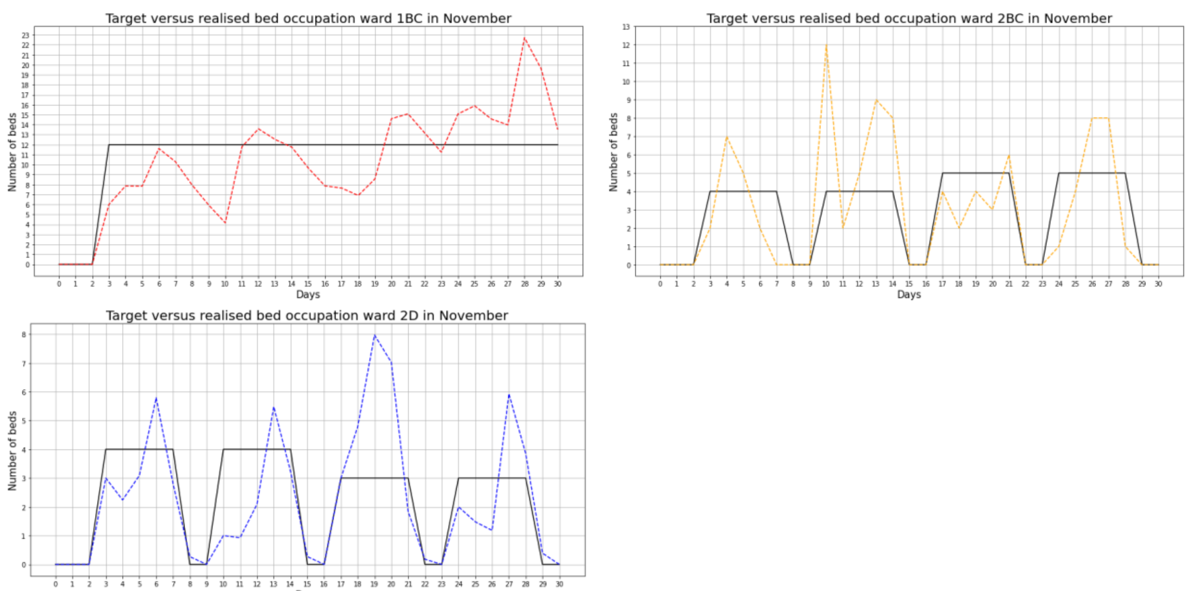

Figure 44: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month November under the scheduling policy of scenario 1

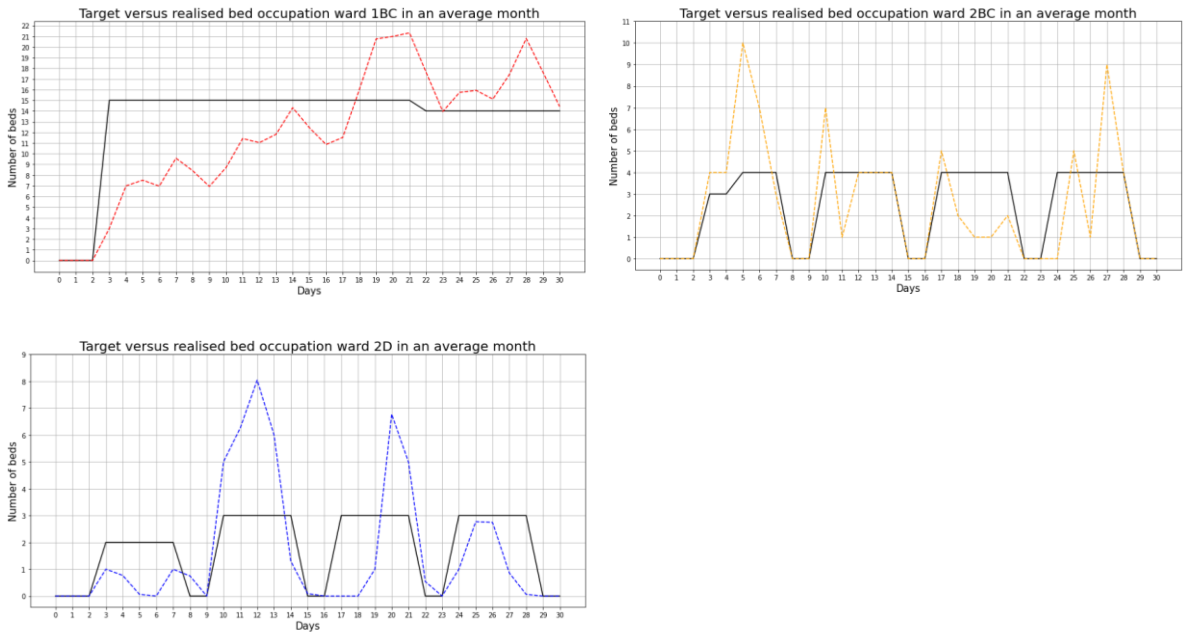

Figure 45: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for an average month under the scheduling policy of scenario 1

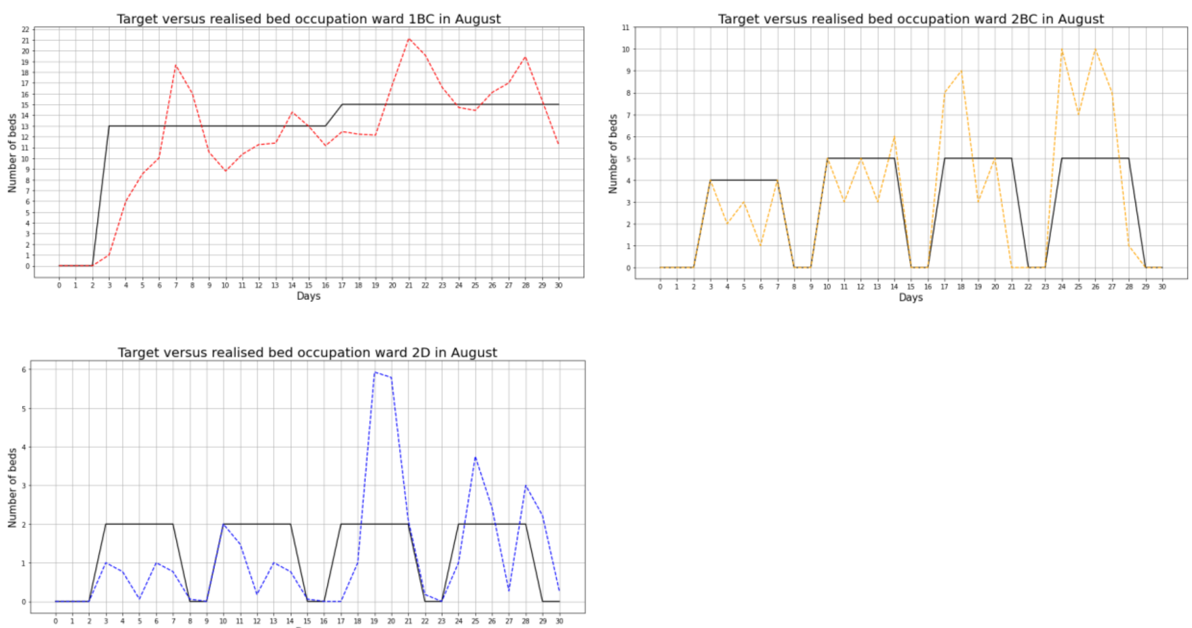

Figure 46: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month August under the scheduling policy of scenario 1

H.2 Scenario 2

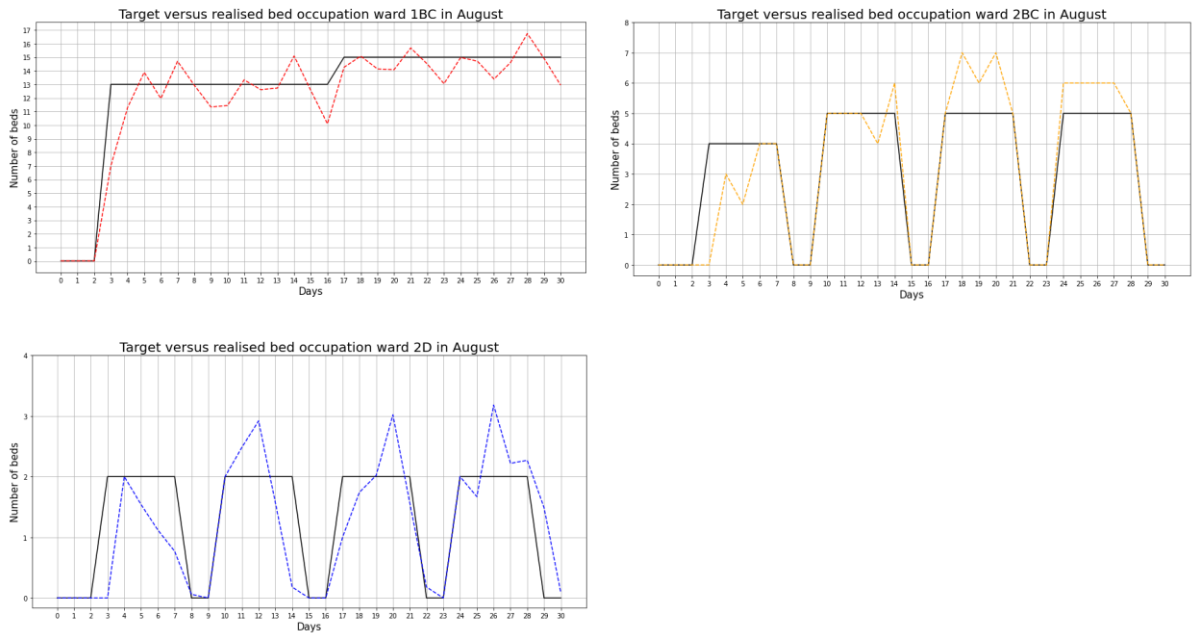

Figure 47: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month August under the scheduling policy of scenario 2

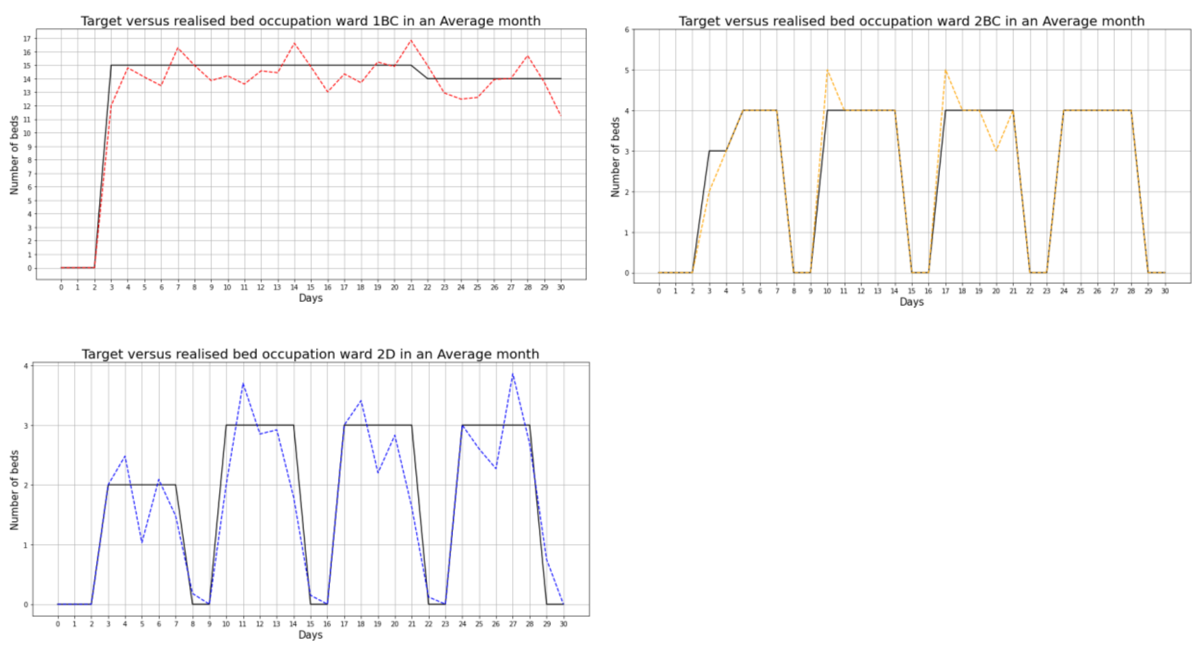

Figure 48: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for an average month under the scheduling policy of scenario 2

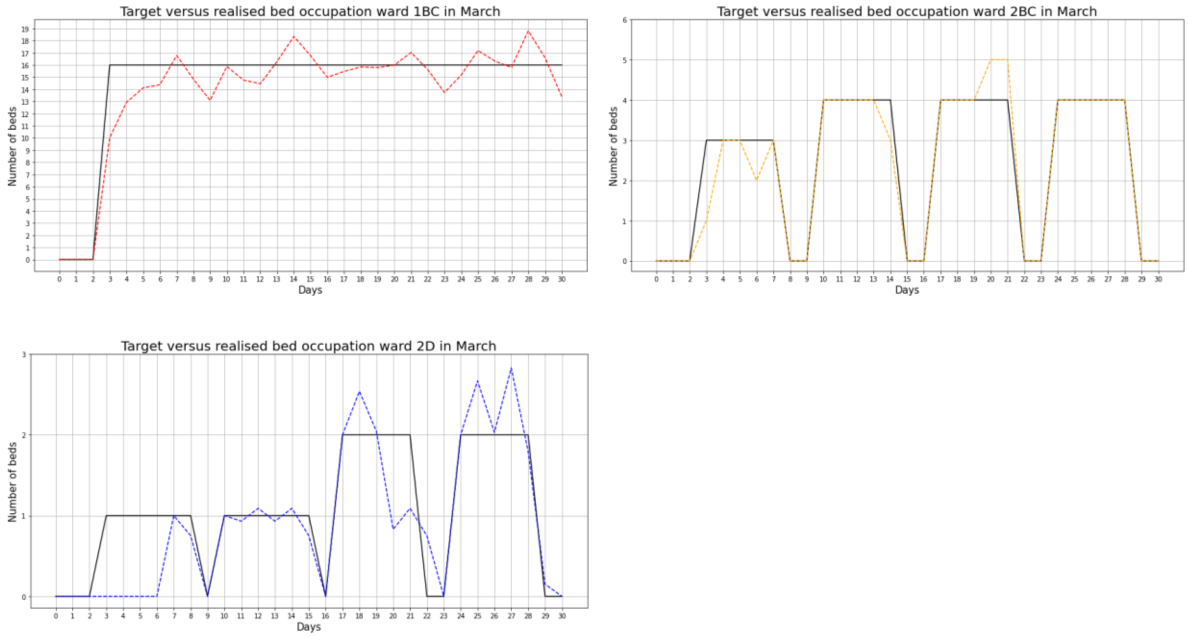

Figure 49: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month March under the scheduling policy of scenario 2

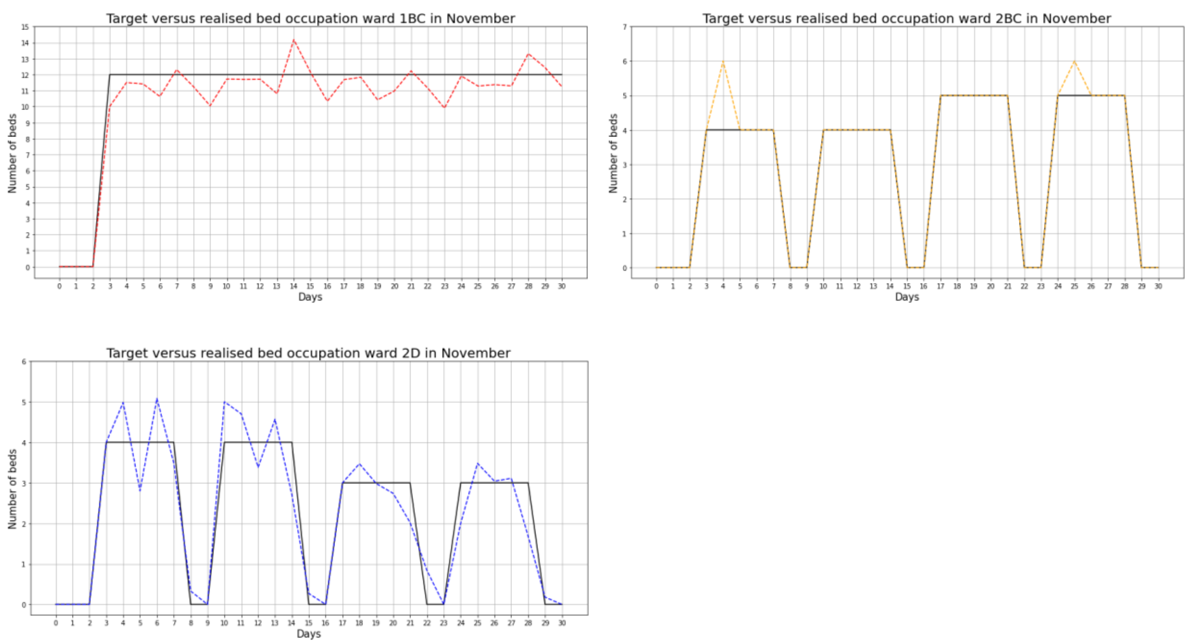

Figure 50: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month November under the scheduling policy of scenario 2

H.3 Scenario 3

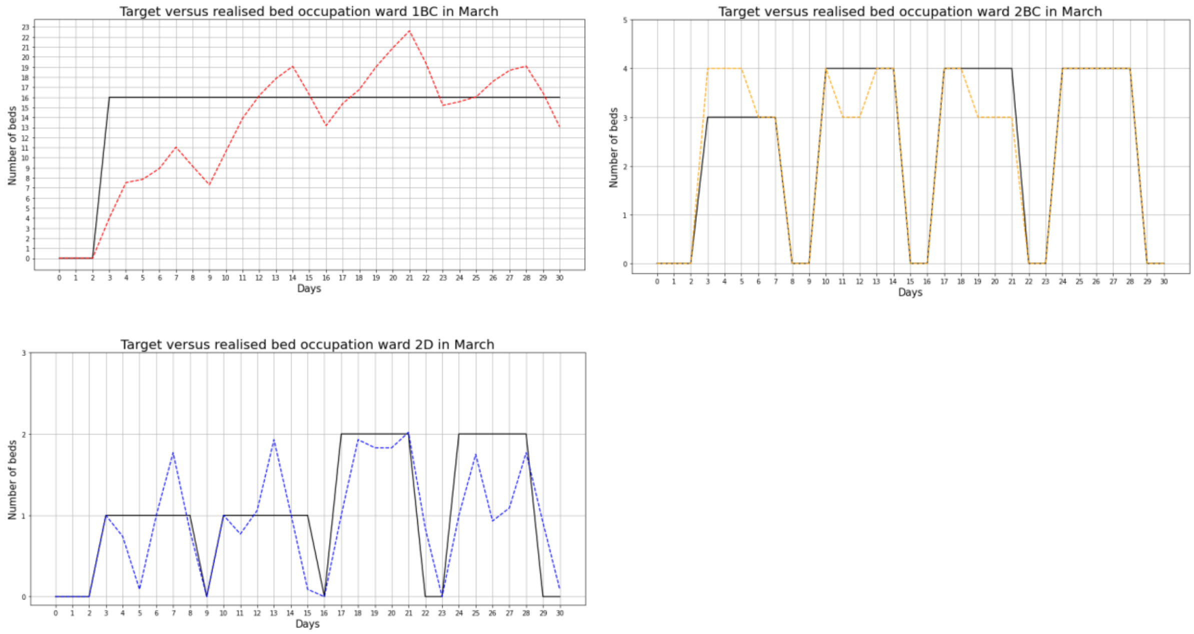

Figure 51: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month March under the scheduling policy of scenario 3

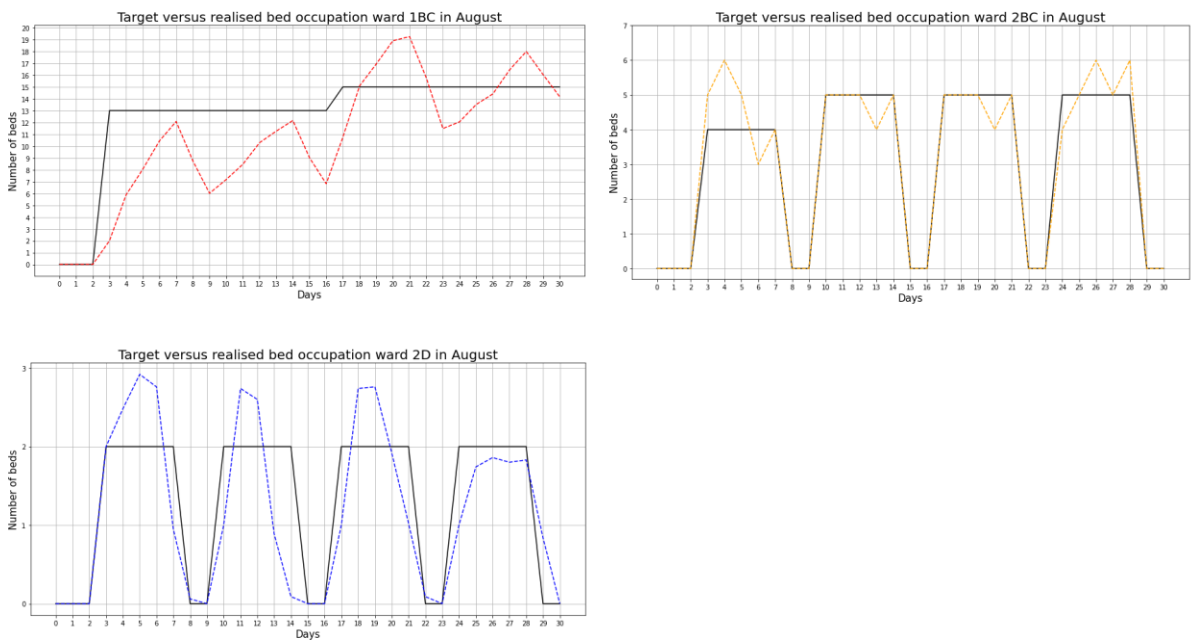

Figure 52: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month August under the scheduling policy of scenario 3

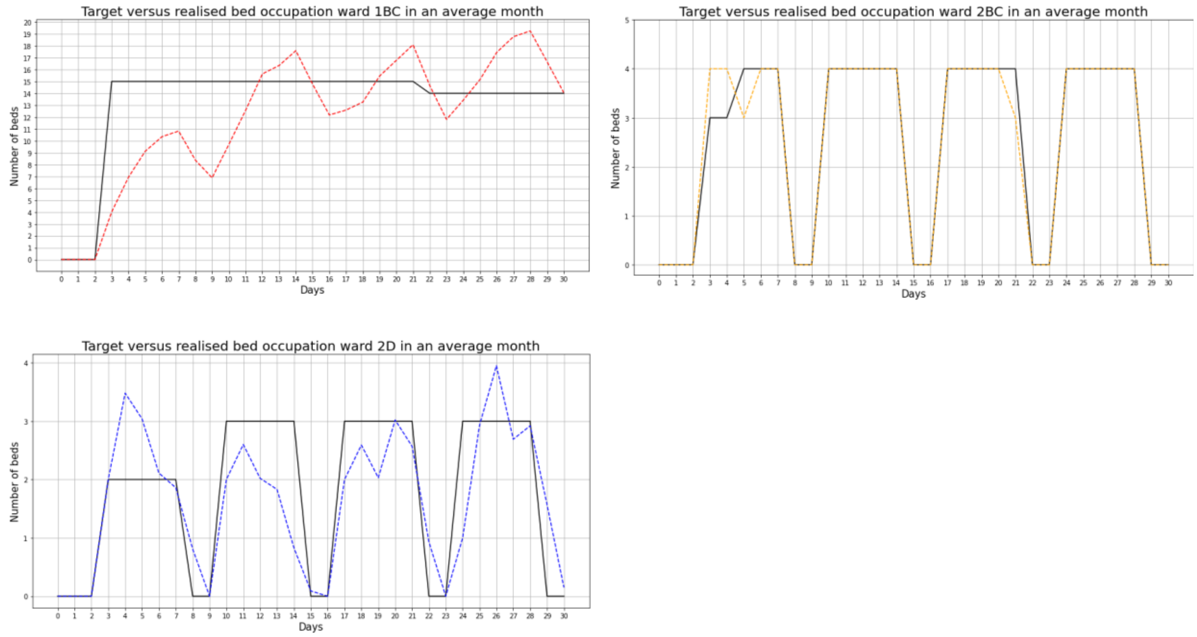

Figure 53: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for an average month under the scheduling policy of scenario 3

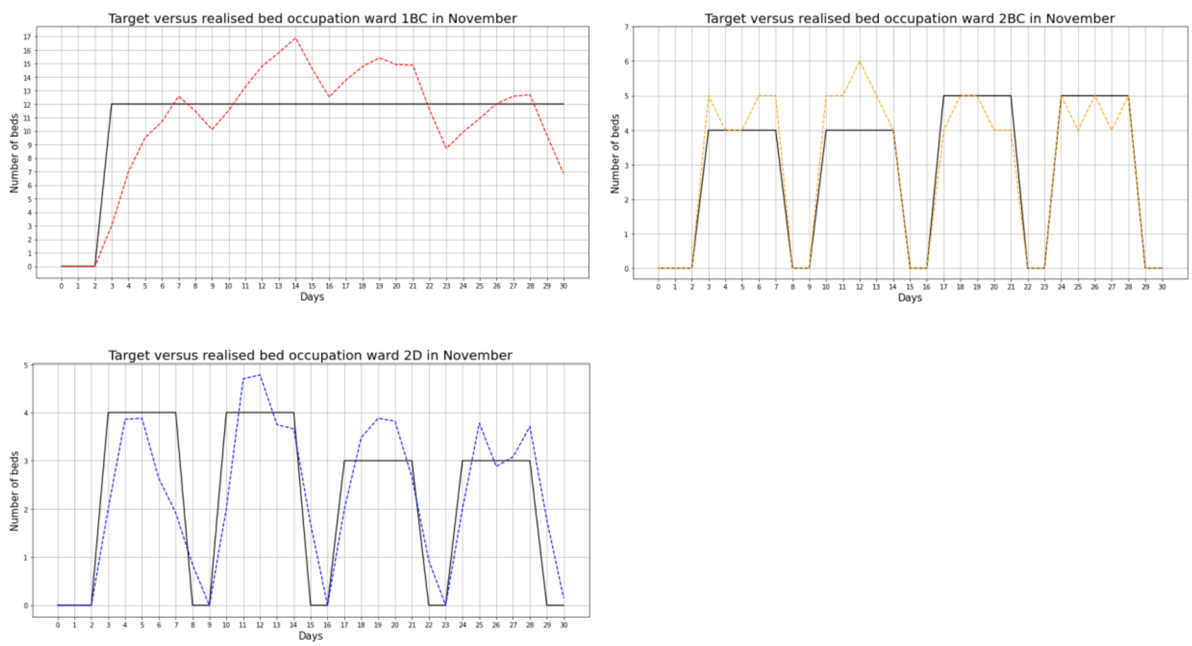

Figure 54: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month November under the scheduling policy of scenario 3

H.4 Scenario 4

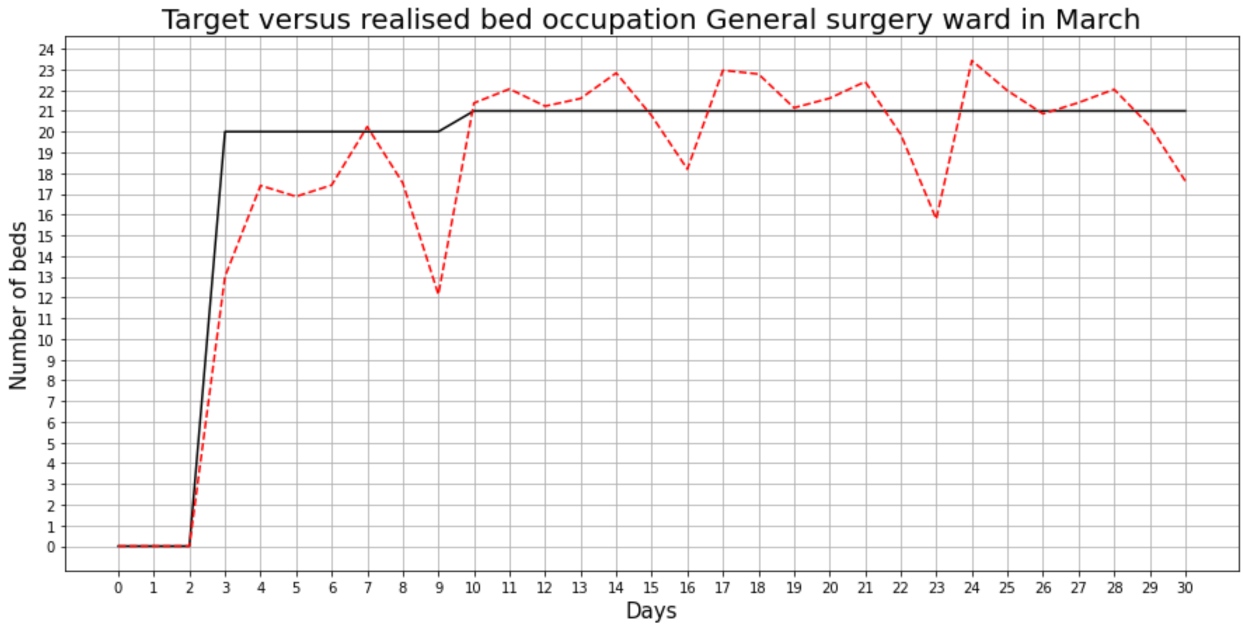

Figure 55: Target versus realized bed occupancy for the surgical ward for the schedule created for the month March under the scheduling policy of scenario 4

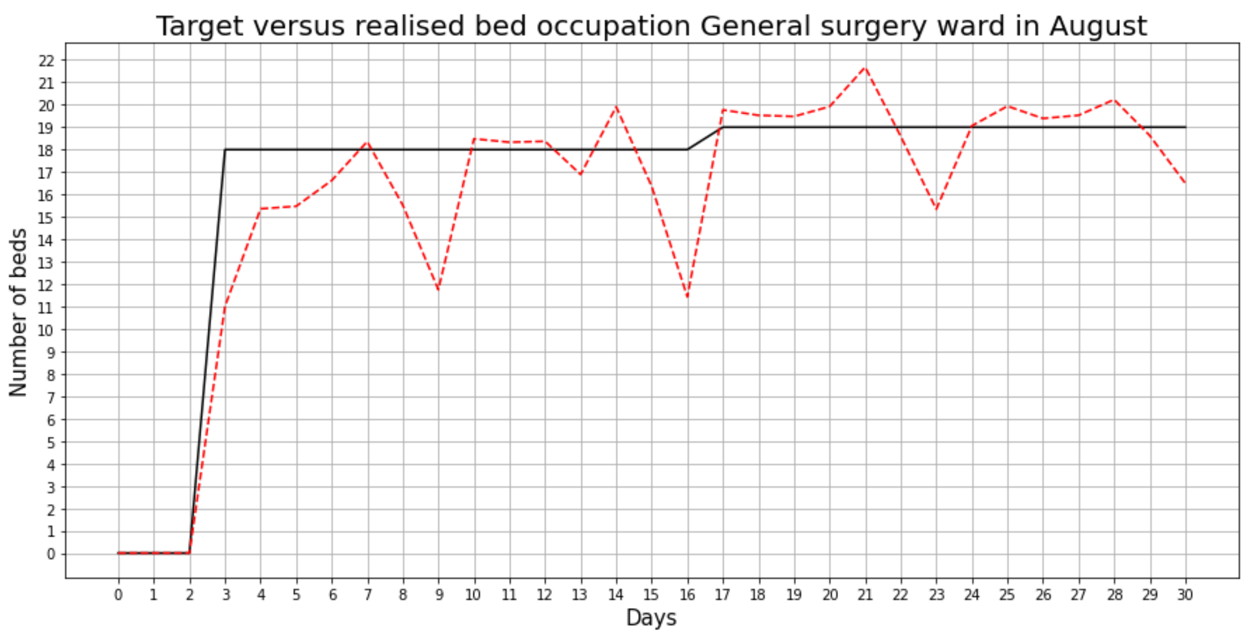

Figure 56: Target versus realized bed occupancy for the surgical ward for the schedule created for the month August under the scheduling policy of scenario 4

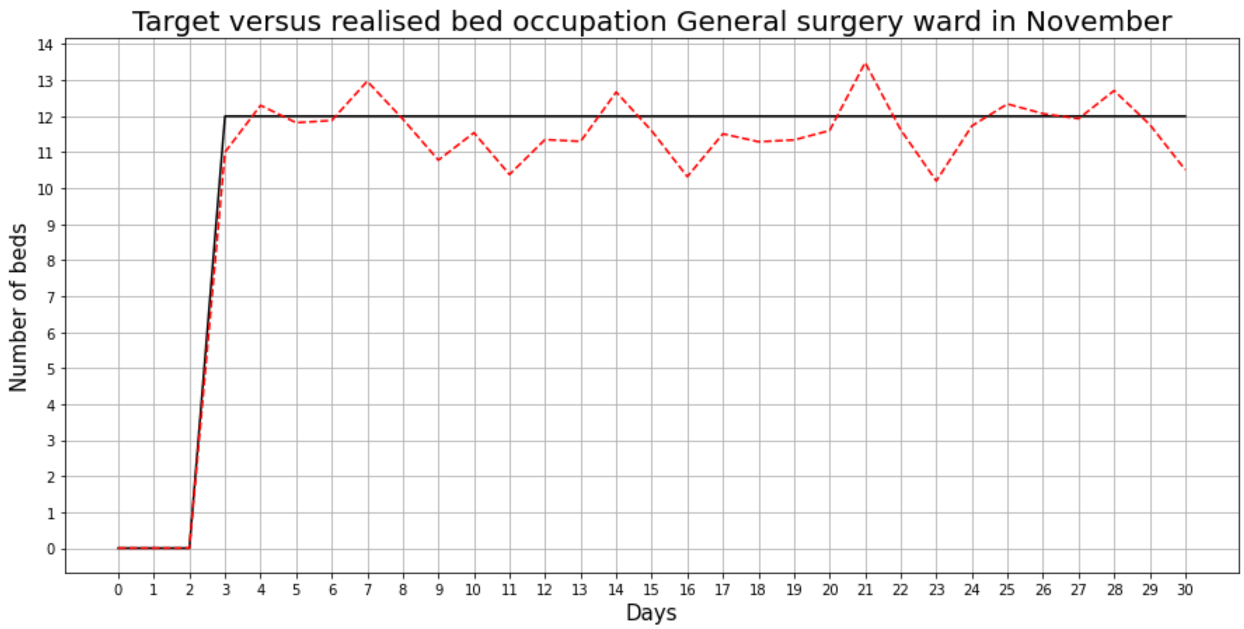

Figure 57: Target versus realized bed occupancy for the surgical ward for the schedule created for the month November under the scheduling policy of scenario 4

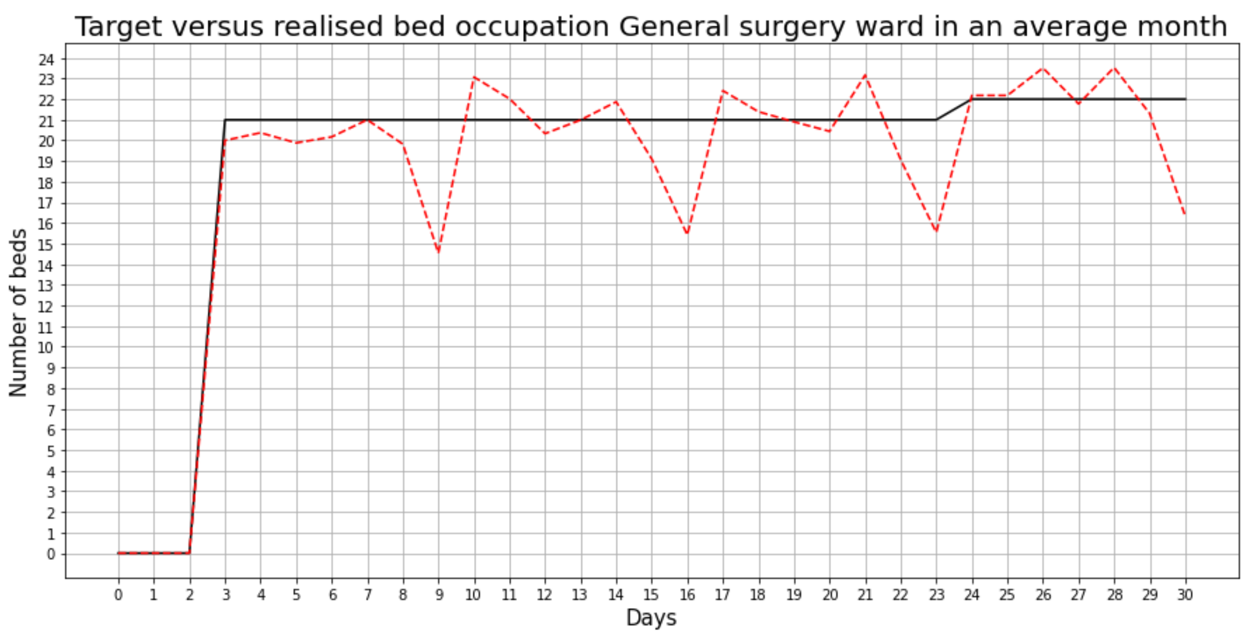

Figure 58: Target versus realized bed occupancy for the surgical ward for the schedule created for an average month under the scheduling policy of scenario 4

H.5 Scenario 5

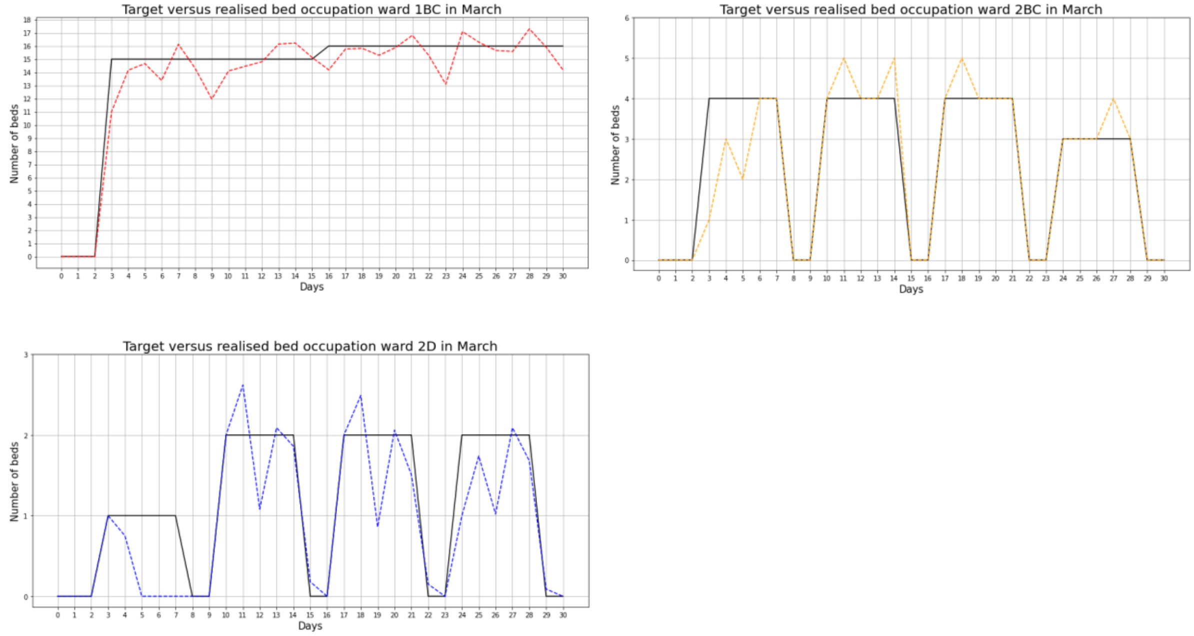

Figure 59: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month March under the scheduling policy of scenario 5

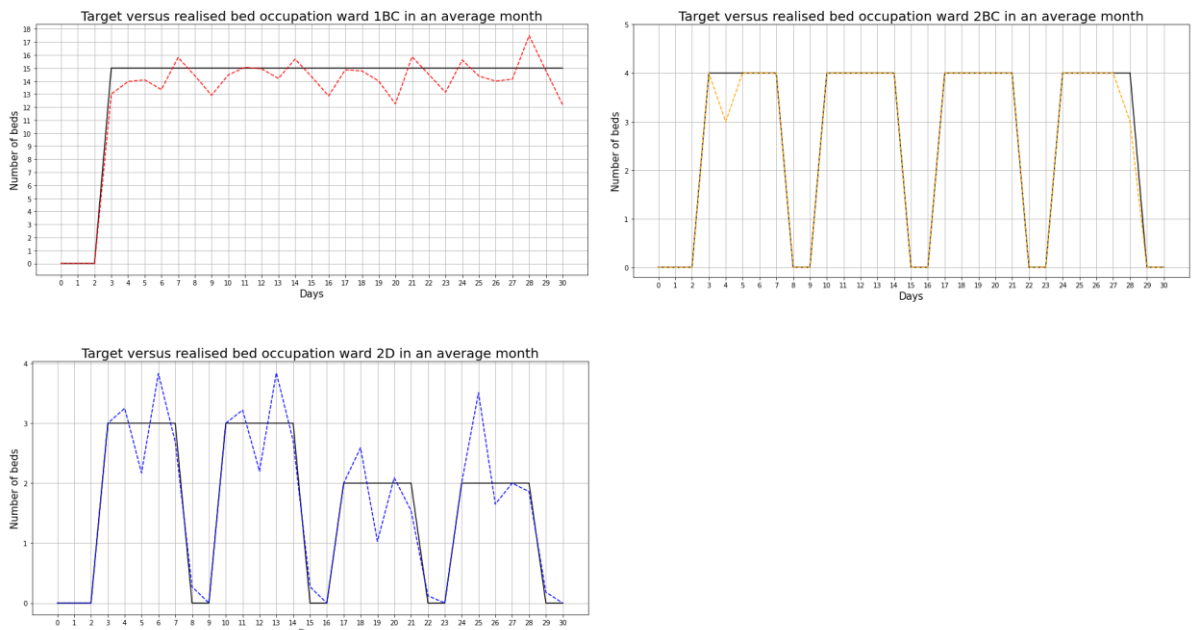

Figure 60: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for an average month under the scheduling policy of scenario 5

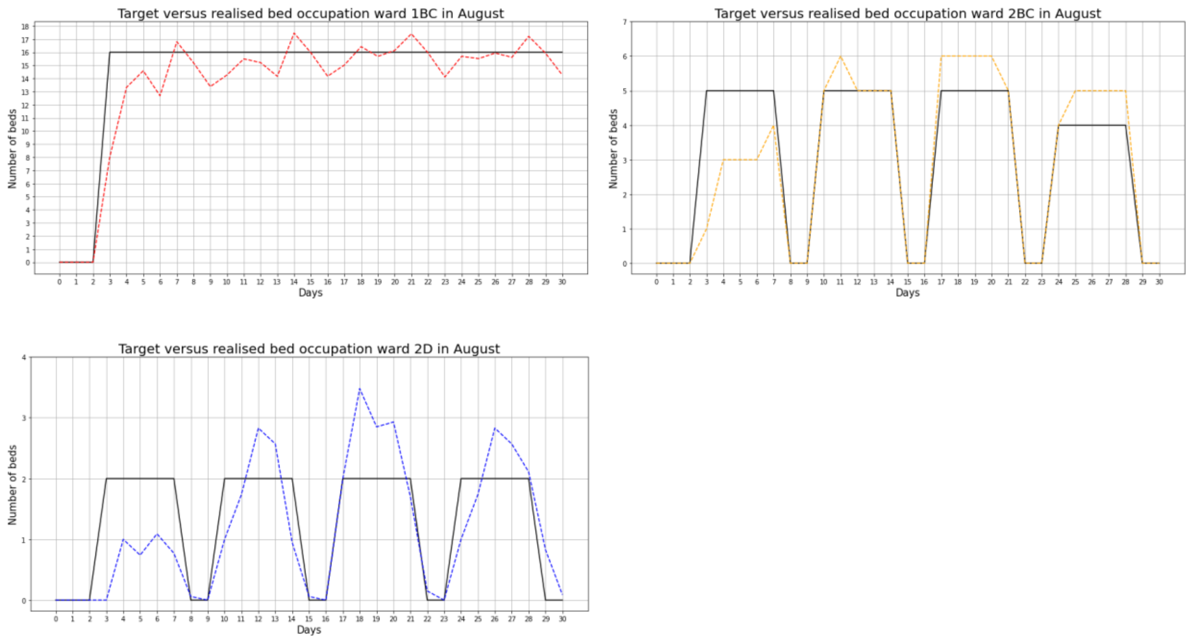

Figure 61: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month August under the scheduling policy of scenario 5

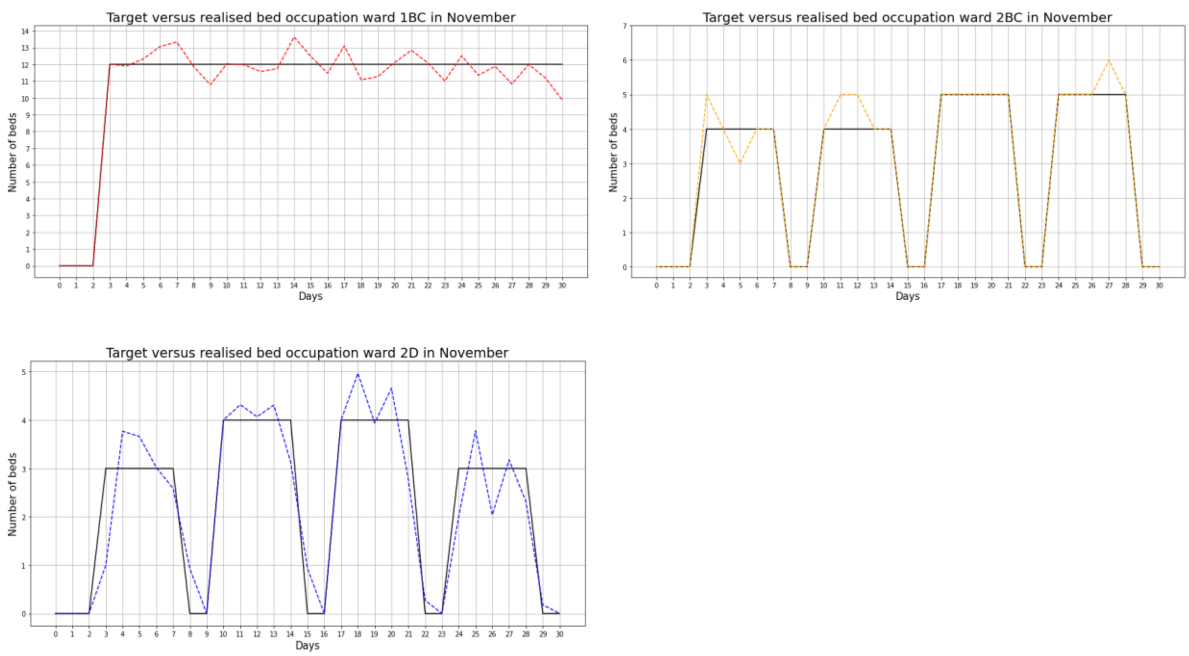

Figure 62: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month November under the scheduling policy of scenario 5

H.6 Scenario 6

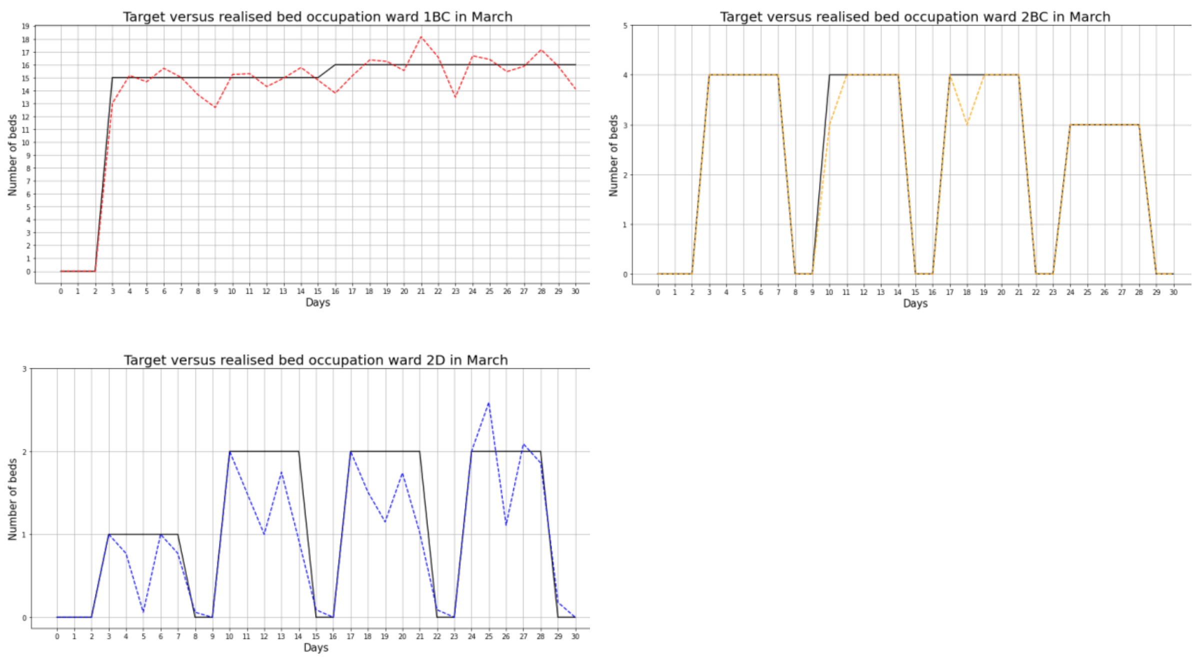

Figure 63: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month March under the scheduling policy of scenario 6

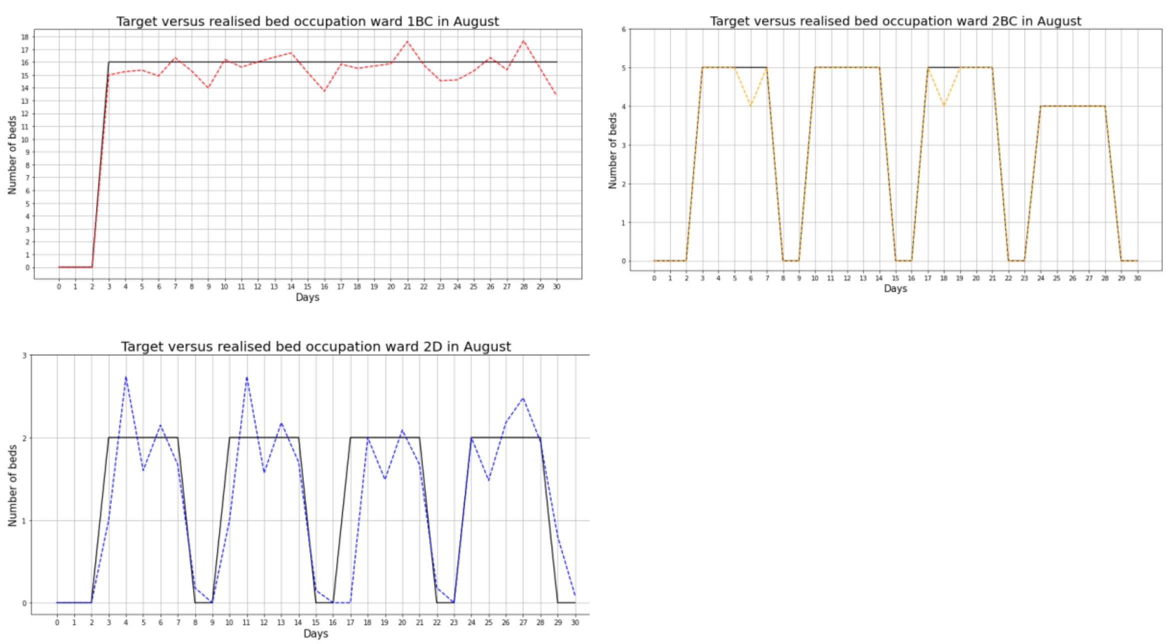

Figure 64: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month August under the scheduling policy of scenario 6

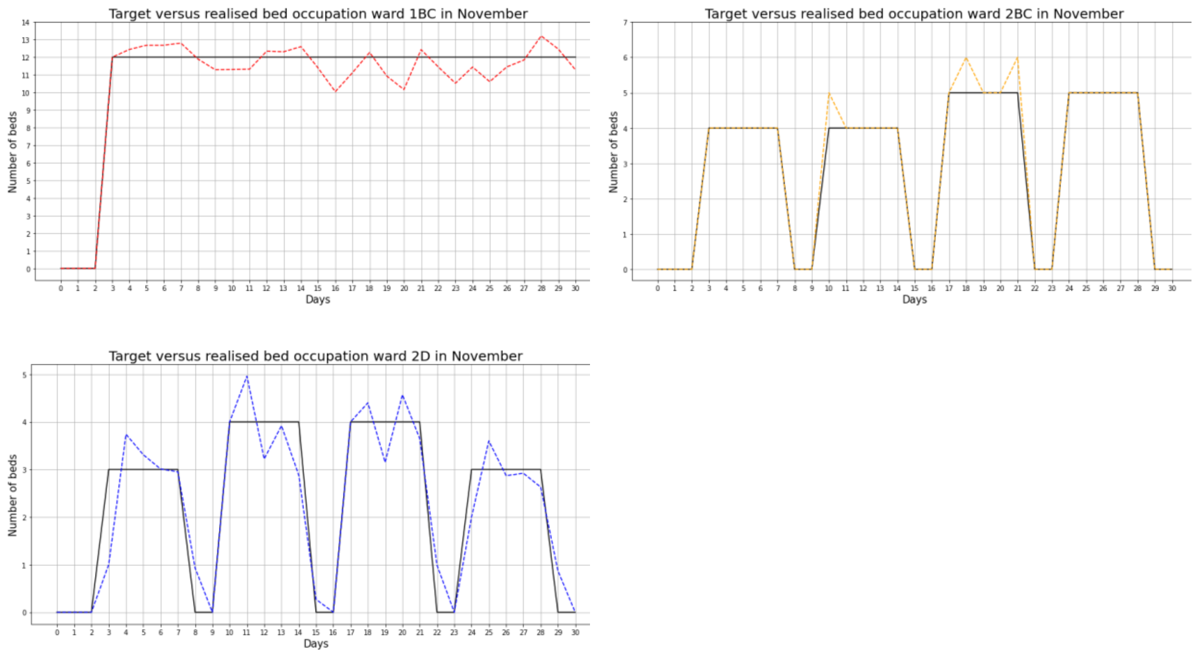

Figure 65: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for the month November under the scheduling policy of scenario 6

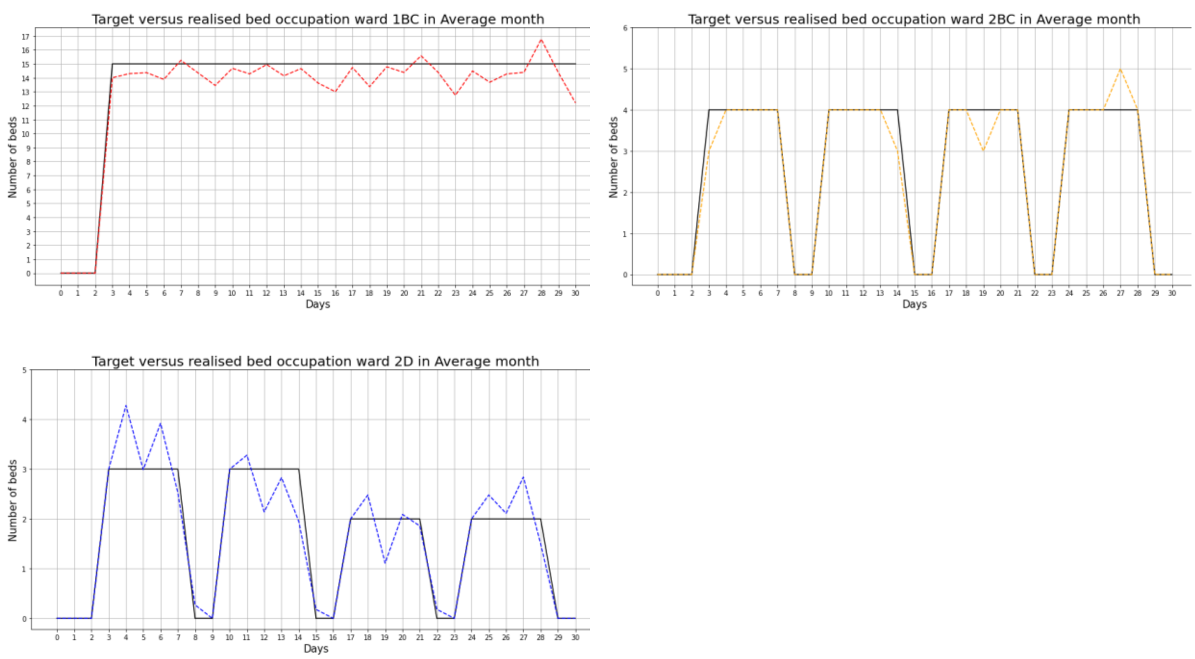

Figure 66: Target versus realized bed occupancy for wards 1BC, 2BC and 2D for the schedule created for an average month under the scheduling policy of scenario 6