POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Pediatric Surgery Scheduling in a Hub & Spoke Hospitals Network

Tesi di Laurea Magistrale in
Mathematical Engineering - Ingegneria Matematica

Author: **Luigi Umana**

Student ID: 10614228
Advisor: Professor Giuliana Carello
Co-advisors:
Academic Year: 2022-23

Abstract

There is an increasing need in clinical sites for optimized planning of surgical procedures. To meet this need, the Hub&Spoke network model, a system of organization that allows a clinical center (hub) the possibility of using operating rooms of hospitals (spoke), is becoming more widespread. This collaborative model is proposed as a solution to two common problems in hospital management: on the one hand, the harm, also economical, of not making the best use of the availability of valuable resources, such as operating rooms, and on the other hand, the risk of having overcrowded clinical centers, causing inconvenience to patients, due to delays and extended waiting times. Over the past decades, much research has been conducted on optimizing the scheduling of surgical procedures, differentiating them according to patient type and hospital needs. This Thesis considers a new model of collaboration between hospitals and addresses the problem of planning pediatric surgeries, considering the related restrictions and specific requirements in a Hub&Spoke setting. Criteria considered for efficient scheduling include the urgency of patients, the penalization in the case of failure to meet the due date for operations, and the preference for surgeries in external operating rooms (spokes) over the referring clinical center (hub), or consider another formulation that aims exclusively to complete all clinical cases on the waiting list in the shortest number of days. The Thesis assumes a deterministic planning problem, which does not consider emergencies, formulated as an integer linear programming problem. However, given the computational onerousness of the ILP model, several model relaxations, alternative formulations, and heuristic solution methods will be proposed. Once the model with the best formulation in terms of solution quality and convergence speed has been identified, its advantages and disadvantages will be analyzed. The proposed approaches were tested on realistic data to evaluate the planning obtained and formulate reflections on possible difficulties in scheduling, points to be improved, and possible strategic decisions that the clinic could take to improve the service offered to patients.

Keywords: Health care management, Surgery scheduling, Hub-and-spoke hospital networks, Optimization, Integer Linear Programming

Abstract in lingua italiana

Nei centri clinici è sempre più necessaria una pianificazione ottimale degli interventi chirurgici. Per venire incontro a questa esigenza sta diventando maggiormente diffuso il modello di rete Hub&Spoke, un sistema di organizzazione che permette ad un centro clinico (hub) la possibilità di usufruire sale operatorie di ospedali (spoke). Questo modello di collaborazione si propone come una soluzione a due problemi comuni nella gestione ospedaliera: da un lato il danno, anche economico, di non sfruttare al meglio la disponibilità di risorse pregiate, quali le sale operatorie, e dall'altro il rischio di avere centri clinici sovraffollati, causando disagi ai pazienti, dovuti ai ritardi e ai tempi di attesa prolungati. Negli ultimi decenni sono state condotte numerose ricerche sull'ottimizzazione della pianificazione degli interventi chirurgici, differenziandoli in base al tipo di paziente e alle esigenze ospedaliere. La Tesi considera un nuovo modello di collaborazione tra ospedali e affrontare il problema di pianificare gli interventi chirurgici pediatrici, considerando le relative restrizioni e le specifiche richieste in un contesto Hub&Spoke. Tra i criteri presi in considerazione per una calendarizzazione efficiente vi sono l'urgenza dei pazienti, la penalizzazione nel caso di mancato rispetto della "due date" per le operazioni, e la preferenza per gli interventi chirurgici nelle sale operatorie esterne (spokes) rispetto al centro clinico di riferimento (hub), o considerare un'altra formulazione che mira esclusivamente a completare tutti i casi clinici in lista d'attesa nel minor numero di giorni possibile. La tesi considera un problema di pianificazione deterministico, che non considera le emergenze, formulato come un problema di programmazione lineare intera. Tuttavia, data l'onerosità computazionale del modello di ILP, verranno proposti diversi rilassamenti del modello, formulazioni alternative e metodi risolutivi euristici. Una volta identificato il modello con la formulazione migliore in termini di qualità della soluzione e velocità di convergenza, verranno analizzati i vantaggi e gli svantaggi dello stesso. Gli approcci proposti sono stati testati su dati realistici per valutare la pianificazione ottenuta e formulare riflessioni sulle eventuali difficoltà nella calendarizzazione, i punti da migliorare ed eventuali decisioni strategiche che la clinica potrebbe adottare per migliorare il servizio offerto ai pazienti.

Parole chiave: Gestione dei Servizi e Sistemi Sanitari, Pianificazione degli Interventi, Reti di Ospedali Hub&Spoke, Ottimizzazione, Programmazione Lineare Intera

Contents

1 | Introduction

This Thesis aims to find an optimization method in health care management. We want to propose a model which optimizes surgical scheduling, in terms of benefits to the patients, but also on a wise utilization of resources. In fact, effective resource management in healthcare is crucial, especially when these ones are limited. It involves prioritizing and strategically allocating personnel, supplies, equipment, and especially facilities. Decision support techniques and data analysis help to optimize resource utilization and ensure that critical cases receive immediate attention. Ultimately, effective resource management enables healthcare providers to deliver quality care, maintain patient safety, and achieve equitable outcomes.

This thesis's goal can be divided into two parts: the first one is oriented toward finding a good model in order to locate optimized scheduling, respecting preferences, requests, and constraints given by the clinic, and the second part, indeed is related to some data analysis on the selected model, to highlight some key aspects or hidden information, and give some basis for future strategical decisions on medical management.

1.1. Surgical Scheduling: Importance, Challenge and Benefits

In the fast-paced world of healthcare, efficient surgery scheduling plays a critical role in maximizing patient care, optimizing resource allocation, and enhancing overall operational efficiency. The process of scheduling surgical procedures involves intricate coordination among various stakeholders, including patients, surgeons, anesthesiologists, nurses, and support staff. It is a complex task that requires careful consideration of numerous factors such as urgency, operating room availability, surgical team availability, patient preferences, preoperative preparations, and many others.

Efficient surgical scheduling is a critical aspect of hospital management, ensuring that surgical procedures are performed in a timely manner while considering various constraints and urgencies. Traditionally, scheduling surgery operations is a complex and time-consuming task, prone to errors and sub-optimal utilization of resources. However, over 15-20 years, advancements in technology have paved the way for automated scheduling systems that can find the optimal schedule, taking into account hospital constraints and patients' urgencies. In the next paragraphs, it will be discussed some main intricacies involved in setting up a schedule for surgery operations, the challenges faced by hospitals, and the benefits of adopting automated scheduling solutions.

Constraints and Features

- One of the constraints in surgical management is the limited number of operating rooms. These need to be scheduled efficiently to maximize utilization while accounting for various factors such as room setup time, cleaning and sterilization requirements, and equipment availability. Underutilization of Operating Rooms (OR) takes with it generally two main disadvantages: economic loss and surgery delays, with the related patients' inconvenience.
  But the operating theatre is not the only resource to consider. In fact, it is also necessary to take into counting the beds and related staff for the patients' Post Anesthesia Care Unit (PACU) i.e. the time required for the patient case to recover after a specific intervention.
- Surgeon and Anesthesiologist Availability. Coordinating the availability of surgeons and eventual anesthesiologists is crucial for a successful surgical scheduling. Each surgeon may have preferred time slots or specific days for performing certain procedures. Coordinating these preferences, considering vacation and conference schedules, and ensuring adequate coverage for emergencies poses a significant challenge.
- Patient Urgencies and Prioritization. Patients with urgent or emergent surgical needs require immediate attention. Balancing the urgency of such cases with the scheduling of elective surgeries is essential to prioritize patient care effectively. Urgent cases may require rescheduling or adjusting the schedule to accommodate their needs while maintaining efficient utilization of operating room time. Further, the different priorities of elective patients must be considered as well
  There are many other factors to consider in setting up the schedule, like the number of waiting days for each case, and the kind of surgery.
- Equipment and Resource Allocation. Surgical procedures often require specialized equipment, such as imaging devices, surgical instruments, and support staff. Scheduling surgeries in a way that optimizes the allocation of these resources is crucial to avoid conflicts and ensure smooth operations. Moreover, considering the

maintenance and repair schedules of equipment is necessary to minimize disruptions during surgical procedures.

- Specialized Surgeries. Every surgical intervention needs specific medical equipment, to better accommodate the kind of surgery. So a factor that cannot be overlooked is the compatibility of a patient to be operated on in a certain Operating Theatre (OT). Generally, each hospital facility does not have a unique specialization and related equipment, but it is designated to be used for different kinds of surgical operations.

Benefits of optimized scheduling

By employing various techniques to assist decision-making and leveraging advanced data analysis methods, we can successfully tackle the complex task of managing operating theatres. This approach offers a multitude of benefits, making the entire process more efficient and effective. It enables healthcare professionals to optimize resource allocation, enhance patient scheduling, streamline surgical procedures, and ultimately improve overall patient care and outcomes. With the aid of decision support systems and robust data analysis, hospitals and healthcare facilities can make informed decisions. The integration of decision support techniques and data analysis empowers healthcare organizations to optimize their operating theatre management processes by looking at many factors

- Algorithms and optimization techniques allow to allocate efficiently resources, time slots, and personnel. By applying optimization techniques to the scheduling process, hospitals can minimize human errors and reduce the time required to generate an optimal schedule. This efficiency leads to increased productivity, improved patient flow, and reduced waiting times, ultimately enhancing overall hospital operations.
- Optimization approaches consider the availability of operating rooms, surgeons, anesthesiologists, and other resources simultaneously. By analyzing multiple variables and constraints, these approaches can generate an optimized schedule that maximizes resource utilization, minimizes or even eliminates conflicts, and reduces idle time. The result is improved efficiency, cost-effectiveness, and a higher volume of surgeries performed with the available resources.
- Optimization methods for scheduling consider patient urgencies and prioritize cases accordingly. Urgent and emergent cases can be integrated into the schedule, reducing delays and ensuring timely access to necessary surgical interventions. This

prioritization improves patient satisfaction, reduces patient waiting times, and enhances overall patient safety by minimizing the risk of delayed treatments.

- Different methods for optimization can be integrated with data analysis tools and performance monitoring. They provide valuable data that can be analyzed to identify trends, optimize workflows, and improve future scheduling processes. Key performance indicators, such as on-time starts, turnover times, and resource utilization, can be monitored to identify areas for improvement, leading to enhanced efficiency and better patient outcomes. But these aspects will be discussed in the next sections.

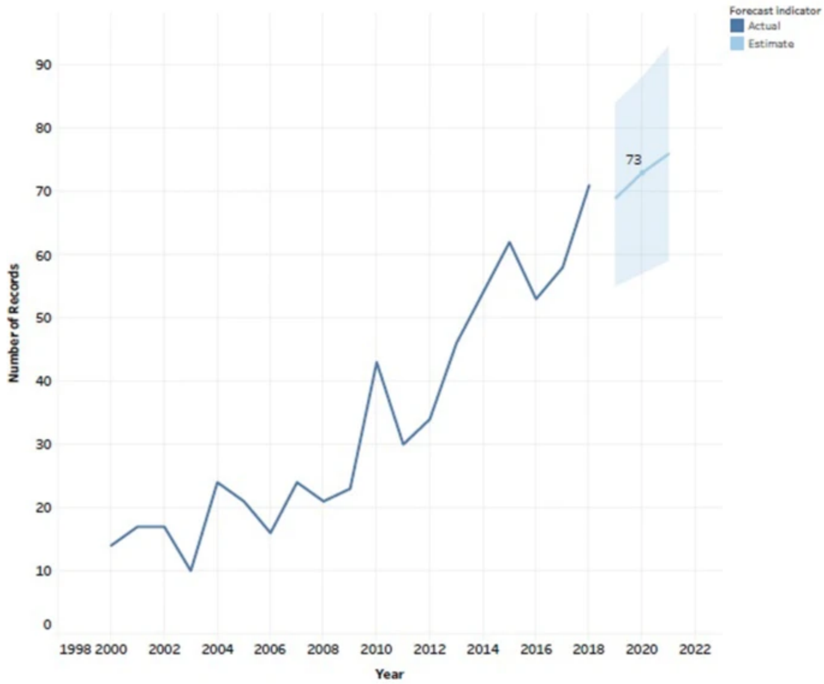
To sum up: applying optimization techniques in surgery scheduling can noticeably help, or even revolutionize surgical scheduling in hospitals, addressing the complex constraints and urgencies involved. By optimizing resource allocation, prioritizing patient care, and increasing operational efficiency, these systems offer significant benefits. The adoption of optimized scheduling not only streamlines the scheduling process but also improves patient satisfaction, enhances resource utilization, and ultimately leads to better overall hospital performance.

1.2. State of the Art and Literature

The operating room scheduling problem is a complex research topic widely studied. Literature reviews presented before 2020 have provided a classification system for organizing technical articles on operating room schedules until 2014. However, with the increasing number of articles published each year, a new literature review is necessary to capture emerging trends more promptly. As Harris, Sean and David Claudio show in their article [25] 246 technical operating room scheduling articles have been published from 2015 to 2020. The review highlights current trends and identifies areas for future research. The fact that this problem is becoming more and more necessary and an important case of study can be easily from Figure 1.1 as reported in the article. It shows the number of essays recorded about this topic from 2000 to 2018 and gives a prediction of the number of them in the next three years

Two major themes stand out: ongoing development and innovation across various categories, and the significant challenge of implementing these models in real-life settings.

Figure 1.1: Number of articles in OR Scheduling [25] [25]



1.2.1. Decision Levels

According to Rahimi, Iman, and Amir H. Gandomi [24] OR scheduling and planning involve different decision levels: strategic, tactical, and operational.
At the strategic level, long-term decisions such as capacity planning and allocation are made. The *Case Mix Problem* (CMP) focuses on optimizing profit/cost over a long time period by determining the allocation of OR time to surgical specialties.
Tactical-level problems, like the *Master Surgery Scheduling* (MSS) problem, deal with cyclic OR schedules and aim to assign surgical specialties to OR time slots to optimize resource utilization. The output of the tactical level, in the form of a cyclic timetable, informs decision-making at the operational level.
This latter one, the operational level, involves shorter-term decisions such as resource allocation, surgical case scheduling, and advanced scheduling. The *Surgical Process Scheduling* (SPS) problem is divided into two sub-problems: advance scheduling (planning future dates for surgical cases) and allocation scheduling (determining start time and resource allocation for cases over a short time horizon). Various literature reviews have been published on OR planning and scheduling, covering different decision levels [24].

In this thesis project, SPS problems will be addressed,

1.2.2. Methods and techniques

Abdalkareem and Zahraa [2] provide a comprehensive survey of recent literature on healthcare scheduling problems, covering various areas such as patients' admission scheduling, nurse rostering, operation room assignment, surgery scheduling, and other healthcare scheduling problems. The research in this medical topic, as we largely said ahead, is crucial for optimizing costs, improving patient flow, and effectively utilizing hospital resources.
In recent decades, there has been a proliferation of healthcare scheduling methods that employ metaheuristic techniques to optimize resource management in hospitals. However, the reported results in the literature are often fragmented due to the independent solving of specific problems and the availability of different problem definitions and datasets.

To address this, the paper integrates existing results by conducting a comprehensive review and analyzing 190 articles based on problem definitions, formulations, datasets, and methods. The review focuses on patients' admission scheduling, nurse enrolling, and operation room reserving problems. It gives a perfect overview of recent developments and also helps to identify new trends for future research directions in healthcare scheduling.

Demeulemeester and his team presented in [29] the importance of efficient scheduling in hospital operating rooms, a clear summary of the recent techniques used in the last two decades in the academic literature. To assist researchers and practitioners in identifying relevant articles, the models are classified based on various factors such as patient type, performance measures, decisions made, facilities, research methodology, and testing phase.

The survey identifies trends and promising topics based on these classifications.
Additionally, it identifies three common pitfalls that hinder the adoption of research results by stakeholders, including the lack of a clear target audience (researchers or practitioners), the use of inappropriate performance measures, and the failure to clearly report the hospital setting and method-related assumptions. The paper provides guidelines to select appropriate performance measures and the importance of including relevant assumptions in articles to help readers determine the relevance of the research presented.

The survey gives a review of papers related to this topic, and each section analyzes articles from a different perspective, which can be either problem or technically oriented, which means proposing methods and formulations whose goals are to optimize the problem (for example the scheduling) or to optimize or to propose new techniques for solving them.
In particular, they distinguish between seven sections:

- Problem classification: distinguish the problem to solve on the basis of the time horizon chosen;

- Patient characteristics: reviewing the literature according to the elective (inpatient, outpatient) or non-elective (urgency, emergency) status of the patient;
- Performance measures: discussing the performance measures (PM) such as utilization, idle time, waiting time, preferences, throughput, financial value, makespan, and patient deferral;
- Decision delineation: indicating what type of the decision has to be made (date, time, room, and capacity) and whether this decision applies to a medical discipline, a type of surgeon or patient up- and downstream facilities: discussing whether an approach includes other units (e.g., PACU);
- Uncertainty: indicating to what extent researchers incorporate uncertainty (stochastic versus deterministic approaches);
- Operations research methodology: describing the sort of analysis carried out and the approach used for the solution or evaluation;
- Testing phase and application: covering the information on the testing data of the research and its implementation in practice.

Every section explains clearly the importance of each feature, the huge amount of variables to take into consideration, and the needs, where some of them are considered as goals and preferences, and others are seen as inviolable constraints.
Now we spotlight carefully all of them.

Patient characteristic

Every patient has specific characteristics and requirements, some of which depend on individual characteristics while some of which depend on the specific kind of operations.

In the first group, we consider all the personal information like age and eventual disabilities. Age in a pediatric environment is a very sensitive topic. In fact, in a general medical ambient, differences in age of a few years or even a few months do not change the accommodation criteria. But this does not hold in a pediatric setting, where patients who are 6 months, 1 year, or 3 years old, have largely different requirements and attention needs.

The second group is composed of all the pieces of information regarding the type of surgery. Firstly we have to see which kind of time horizon planning the patient should be inserted, and so understand if it will be an elective or non-elective patient, but this

argument will be deeply discussed in the next paragraph.
Then, every intervention has an urgency factor, that indicates a sort of priority in the schedule: the bigger the earlier.
Every surgical operation has also indications about its average duration, and eventually the standard deviation of it, since no one operation last the exact expected surgical times.
After the intervention, there should be a time for the correct recovery, the PACU, which can be low or high, depending for example if an intervention has had local or total anesthesia. Some of the recoveries can last also for one entire night or even more, and in such a case they cannot be treated as day-hospital surgeries and must be adequately ministered.

Problem classification

The first consideration that must be done is the type of schedule we want to organize, which depends on the time horizon considered. All surgical cases can be scheduled in a time range of months, weeks, days, or even a few hours. This actually depends on the type of surgical intervention, and relatively at his urgency and emergency.

So the literature to differentiate these kinds of problems used to group patients, on the basis of their characteristics, in elective and non-elective patients.

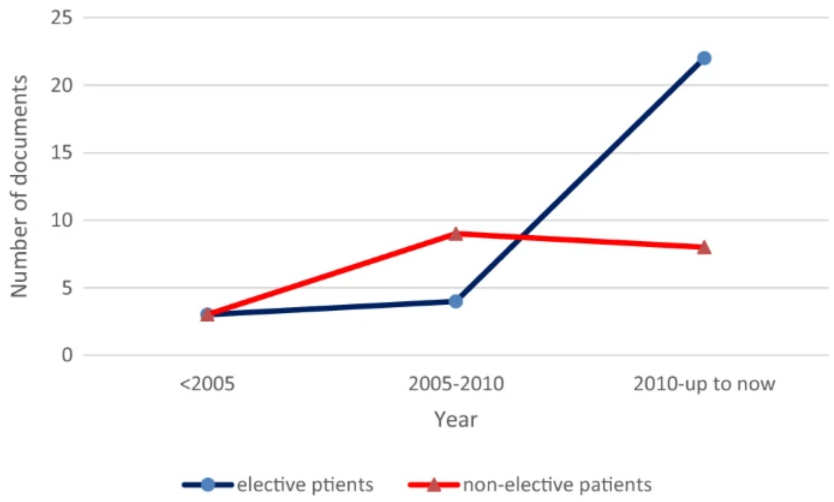
There are many comparisons between articles on elective and articles on non-elective patients and about inpatients and outpatients. Different classes of patients require appropriate scheduling for surgery.

As Figure 1.2 shows, in recent years there has been a notable shift in attention towards elective-patient surgery operation scheduling compared to non-elective ones. Before showing the several reasons that explain this fact, let us see what an elective patient is, in medical terms.

An elective patient refers to an individual who undergoes a planned medical procedure or surgery that is scheduled in advance and is not considered to be an emergency. Elective procedures are typically non-life-threatening and can be scheduled at a time that is convenient for both the patient and the healthcare provider.
Elective surgeries can encompass a wide range of medical interventions, and treatments that aim to improve a patient's quality of life or address non-emergency health conditions. These procedures are often recommended by healthcare professionals based on the patient's medical condition, symptoms, and the expected benefits of the intervention.

Unlike emergency or urgent surgeries, elective procedures allow patients and healthcare providers to plan and prepare adequately. Surgeries are conducted in a controlled environment, such as an operating theatre, and require the involvement of various health-

Figure 1.2: Comparison between the number of documents on elective patients and non-elective patients scheduling [25]



care professionals, including surgeons, anesthesiologists, nurses, and support staff. By scheduling elective procedures, and so make planning in advance, hospitals can optimize the utilization of their resources and streamline the delivery of care.

The reasons why literature pushed significantly the elective patient OR scheduling are several. First of all elective surgeries are planned in advance and can be more easily scheduled, allowing hospitals to better allocate their resources, including operating rooms, staff, and equipment. By optimizing the scheduling of elective surgeries, hospitals can achieve higher resource utilization and efficiency.

Elective surgeries are often less time-sensitive and can be rescheduled if necessary, allowing hospitals to balance the workload and utilize resources more effectively. By focusing on optimizing elective-patient surgery operation scheduling, hospitals can reduce costs associated with overtime, under-utilization of resources, and cancellations.

Secondly, they involve patients who could have the option to choose the timing, or period, of their procedure. By organizing elective-patient surgery operation scheduling, hospitals can provide better patient satisfaction by accommodating their preferences and minimizing waiting times. This can lead to improved patient experience and outcomes.

Elective surgeries are typically predictable, allowing hospitals to plan and optimize their schedules in advance. On the other hand, non-elective surgeries, such as emergency or very urgent procedures, often require immediate attention and cannot be easily, or at all, scheduled in advance. The unpredictable nature of non-elective surgeries makes their scheduling more challenging and more difficult to optimize.

While non-elective surgeries remain critical and time-sensitive, the focus on elective-patient surgery operation scheduling in recent years can be attributed to the potential for greater efficiency, cost-effectiveness, improved patient experience, and the ability to plan and optimize resources. However, it is important to recognize the ongoing importance of non-elective surgeries and the need for efficient scheduling practices to ensure timely and appropriate care for all patients.

Elective patients can also be divided into two different groups: inpatients and outpatients. *Inpatients are hospitalized patients who have to stay overnight, whereas outpatients typically enter and leave the hospital on the same day* [24].

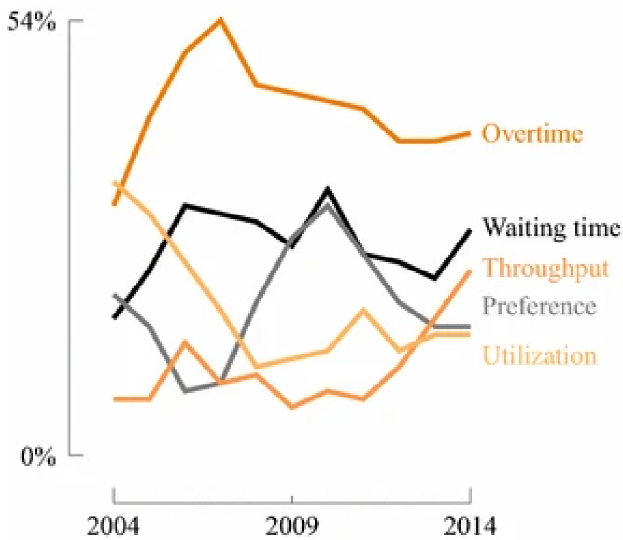
Actually, outpatient care (also called ambulatory care) has increased in literature, since, according to the Milliman Medical Index [1], outpatient care costs have grown by 9.9% on average over time. This rise is mostly related to rising costs for both current services and pricey new services, although linked to a proportionate rise in admissions for outpatient treatment [22]. As an illustration, to delineate differences between the two groups, routine operations, and less invasive techniques are frequently used in outpatient surgery. Furthermore, the precise timing of the arrival of outpatients is unpredictable because they are not already residing in a hospital ward prior to the operation. The selection of the scheduling approach may be significantly influenced by these and other factors.

The planning of outpatient surgeries so is a procedure that does not require hospitalization and has lower costs and risks of infection. Non-elective surgeries, since are urgent but not life-threatening cases that may disrupt the elective schedule, have different ways of handling. Break-in moments, dedicated ORs, and disaster scenarios are some of the methods and policies that have been proposed or implemented to deal with non-electives. The article also provides a comprehensive overview of the current challenges and opportunities in outpatient surgery management.

Performance measures

The paper focuses on performance measures (PMs) used in the literature related to operating room (OR) planning and scheduling. Different stakeholders prioritize different PMs based on their interests and goals. Hospital administrators may prioritize high utilization [7][13][11], and low costs [12], while some may prioritize short waiting times of patients [7][14] .
To find a compromise, many authors include multiple PMs and often use a weighted sum approach. The major PMs discussed in the literature include waiting time [7][14], utilization [7][11], leveling (balancing the utilization of units connected to the OR, for a more balanced workload for the medical staff) [19][20][18], idle time [14], throughput

Figure 1.3: Performance measures used over time, from 2004 to 2014 [29]



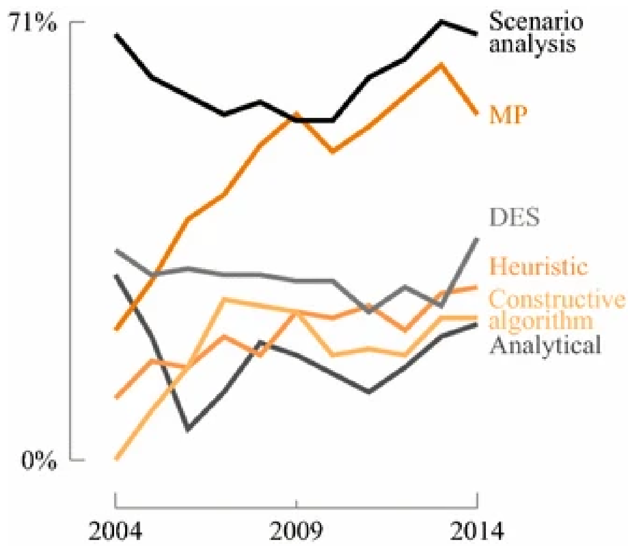
[7][11][13], preferences [5][6][10], financial measures [15], [18] makespan [8] and patient deferral [9][16][18]. Direct and indirect waiting times are frequently used as the main goal to minimize. Surgeon waiting time is less commonly included. Minimizing overtime is a popular objective due to its impact on staff satisfaction, costs, surgery cancellations, and downstream disruptions. Regular OR utilization and patient throughput are also considered, but less frequently. A quick overview to see which PM was preferred, and taken into analysis is given in Figure 1.3. It shows how many surveys in the literature, have taken into consideration specific criteria for the measurement of performance, and how they change in the decade.
The article also provides examples of studies and discusses the importance of each PM, for each different problem. Overall, selecting appropriate PMs is crucial for balancing the interests of different stakeholders in OR planning and scheduling.

Operation research metodology

The literature on OR planning and scheduling encompasses various methodologies within the domain of operations research. Table 1.4 provides an overview of the techniques used to solve OR setup problems. Scenario analysis, where different scenarios or options are compared, is one of the most popular approaches, particularly in discrete-event simulation (DES) modeling. Simulations were used to affect patient flow and resource allocation, such as patient scheduling and routing, resource scheduling, and the sizing and planning of beds, rooms, and staff. Most of the simulation work is done at the department or clinic level.

Some notable studies which include an integrated DES model were conducted by Steins

Figure 1.4: Methods used for the OR scheduling problems, from 2004 to 2014 [29]



et al. [31], considering preoperative care and a PACU, or through an analytical approach using a Markov model by Tancrez et al. [32], determining OR capacity for non-elective patients. Olivares et al. [28] indeed analyzed the decision-making process of reserving OR capacity using the newsvendor model.

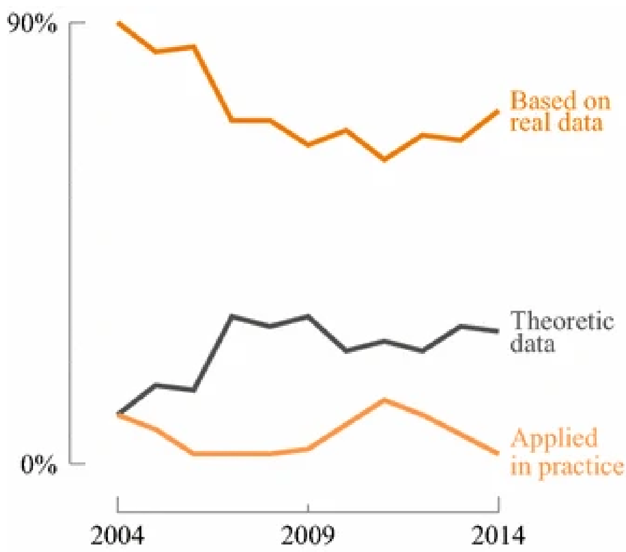
The use of mathematical programming (MP), constructive algorithms, and improvement heuristics is prevalent in the literature. MPs, such as mixed-integer programming (MIP), deal with combinatorial optimization problems. Multiple objectives, including under-/overtime or under/over-utilization, are often considered in the optimization models. Elective patients are the focus of many MP models, but stochastic versions incorporating random variables are also used.

Heuristics are proposed when MPs become computationally challenging. Column generation-based heuristics and genetic algorithms are employed to solve patient scheduling problems. Additionally, methods like Six Sigma and screening for the economic impact of improving first-case starts are introduced in certain articles [21].

For future studies, simulation-optimization is regarded as a promising method for solving complex optimization problems while incorporating the complexities of the OR scheduling process. Traditional methods can also provide valuable insights, but efforts should be made to make them applicable to a broader range of realistic problem settings by accommodating multiple sources of variability and expanding the supported distributions.

Testing phase and application

Figure 1.5: Different methods proposal, from 2004 to 2014 [29]



Researchers often conduct thorough testing phases to demonstrate the applicability of their research. They aim to show computational efficiency or the extent to which objectives can be achieved, requiring a substantial amount of data. Most of the data used in these tests are based on real healthcare practices. However, how it is reported in Figure 1.5, it has been found that despite the availability of real data for testing, less than 7% of the methods proposed in the literature are actually applied in practice. This contradiction highlights the gap between research and implementation in the practical domain of OR planning and scheduling. While some progress has been made in implementing research in health services, the majority of articles still lack reports on actual implementation.

There are a few exceptions where research has been successfully implemented. For instance, Wachtel and Dexter [34] introduce a website used by different US hospitals to determine patient arrival times for surgery appointments, considering factors such as early starts and patient availability. Another example is the decision support system developed by van Essen et al. [33] for daily rescheduling problems where is evaluated with a simulation modeling tool in British healthcare organizations. Understanding key issues in practice helps researchers build models that better reflect reality and solve problems closer to real-world scenarios.

Authors often provide limited detail about the implementation process. It is essential to provide additional information on the behavioral factors associated with implementation, as identifying the causes of success or failure can benefit the research community.

Many articles define problems specific to a single hospital, and the extent to which methods are applicable to other settings is unclear. Generalizable methods should be introduced to facilitate the spread and implementation of effective operations research practices across multiple hospitals. Van Essen et al. [33] address this by surveying several hospitals to justify the generality of their modeling assumptions.

Future research can focus on assessing the current use of planning and scheduling expertise in hospitals. Surveys, such as the one conducted by Sieber and Leibundgut [30] in Switzerland, provide insights into the state of OR management. Developing guidelines on how scheduling data should be made publicly available, along with a standardized format for describing hospital settings, would be beneficial for future research.

1.3. Multihospital

One recent innovation that holds immense importance in the medical and hospital sector is the multihospital (MH) strategy. This approach involves the integration and collaboration of multiple hospitals to deliver comprehensive, high-quality care. We briefly show the latest innovation surrounding the multihospital model and highlight its significance in the current moment for medical and hospital reasoning.

The multihospital approach involves the formation of networks or systems comprising multiple hospitals and healthcare facilities. These institutions come together to share resources, expertise, and best practices, with the aim of providing coordinated, patient-centered care across a broader geographic area. The multihospital model fosters collaboration, standardization of care, and the optimization of resources, ultimately leading to improved healthcare outcomes and patient experiences.

An important issue that MH faces up is enhanced access to specialized care. In many regions, access to specialized healthcare services can be limited, requiring patients to travel long distances or face lengthy waiting periods. The multi-hospital approach addresses this challenge by leveraging the expertise of multiple hospitals within a network. By coordinating services and sharing resources, specialized care can be made more accessible to patients, even in underserved areas. This approach reduces the burden on individual hospitals, ensures equitable access to quality care, and minimizes the need for patients to seek care far from their communities.

With this model care coordination and continuity are improved. The multihospital approach places in fact a strong emphasis on care coordination and continuity, particularly for patients with complex medical needs. Through interconnected electronic health record systems and standardized protocols, healthcare providers within the network can share patient information, test results, and treatment plans. This enhanced communication and collaboration enable a comprehensive understanding of patients' medical histories, ensuring continuity of care across different hospitals and specialties. As a result, patients experience smoother transitions between providers, reduced duplication of tests, and improved patient safety.

The multihospital strategy demonstrates its critical importance in times of crisis, such as pandemics or natural disasters. The interconnectedness and shared resources of the network allow for effective coordination and response. In the face of a surge in patient volume, hospitals within the network can collaborate to allocate resources, staff, and expertise where they are most needed. This approach enhances the healthcare system's resilience, improves emergency preparedness, and ensures that patients receive the care they require, even in the midst of challenging circumstances.

The MH approach represents a transformative innovation in the medical and hospital sector, and maybe every clinic will be cooperative and build an enormous web of social welfare, and this will become, hopefully, a standard method in the future.

1.3.1. Hub & Spoke

The Hub & Spoke approach is one of the huge multihospital system.

In recent years, the healthcare industry has witnessed significant transformations, driven by advancements in technology and an increasing focus on efficiency and patient-centric care. One innovative approach that has gained traction and holds immense promise is the Hub & Spoke model. This model is revolutionizing medical and hospital systems, offering a plethora of benefits that are particularly crucial in today's ever-evolving healthcare landscape.
We explore the latest innovation surrounding the Hub & Spoke model and shed light on its importance in the medical and hospital sector.

The Hub & Spoke model is a strategic framework that involves a central hub institution collaborating with satellite facilities, known as spokes, to provide specialized healthcare services. The central hub typically houses the expertise, resources, and infrastructure required for complex procedures and diagnoses, while the spokes offer primary care and routine medical services. This approach enables effective coordination, streamlined patient flow, and optimized utilization of resources, leading to improved healthcare outcomes.

The importance of the Hub & Spoke model in the current moment

In many regions, access to specialized medical services is limited, resulting in patients traveling long distances or facing prolonged waiting times. The Hub & Spoke model addresses this challenge by decentralizing specialized care. Spokes, located in local communities, provide primary care and basic diagnostic services, reducing the burden on the central hub and improving accessibility for patients. This approach ensures that quality healthcare is available closer to patients' homes, making it more affordable and convenient. This is a Hub & Spoke approach especially used in vastly extended regions, where the geographical facility location has a relevant importance.

Efficient Resource Utilization

This method is used, and it has a big potential utility, also in metropolitan areas. Healthcare facilities, especially hospitals, often face resource constraints such as limited beds, specialized equipment, and skilled medical personnel. By adopting the Hub & Spoke model, hospitals can optimize resource allocation. Routine cases and low-complexity treatments can be handled at the spokes, while the central hub focuses on complex cases and specialized services. This efficient distribution of resources reduces overcrowding at the central hub, enhances cost-effectiveness, and enables optimal utilization of expensive medical equipment, creating a web of collaborative hospitals.

Collaboration and Knowledge Exchange

The Hub & Spoke model fosters collaboration between the central hub and spokes, creating a network of healthcare providers who can share expertise and exchange knowledge. The central hub serves as a hub of excellence, providing training and mentorship to the spokes, which, in turn, gain exposure to advanced medical practices. This collaboration boosts the overall competence of the healthcare system and ensures that patients receive the best possible care.

Disaster Preparedness

In times of emergencies, such as pandemics or natural disasters, healthcare systems face unprecedented challenges. The Hub & Spoke model, with its decentralized structure, provides inherent resilience in the face of such crises. The spokes can act as decentralized healthcare centers, handling routine cases and triaging patients, while the central hub concentrates on critical cases and resource mobilization. This approach facilitates a coordinated response, ensures efficient allocation of resources, and enhances the overall healthcare system's ability to handle emergencies.

To conclude, the Hub & Spoke model represents a significant innovation in healthcare, offering numerous advantages for medical and hospital systems. From improving accessibility and affordability to optimizing resource utilization [27] [26].

1.3.2. Post pandemic

The impact of the COVID-19 pandemic on surgical scheduling and the subsequent surge of articles and studies on this topic cannot be overlooked. The pandemic brought about unprecedented challenges to healthcare systems worldwide, causing disruptions in surgical services and necessitating the development of innovative solutions to address the backlog of postponed surgeries. As a result, numerous articles have been published in recent years, highlighting the impact of the pandemic on surgical scheduling and proposing strategies to mitigate its effects. These articles shed light on the importance of efficient scheduling practices in the face of emergencies and the need for adaptive scheduling systems.

The COVID-19 pandemic has resulted in the postponement of elective surgeries to prioritize resources and minimize the risk of virus transmission by sending patients to specific spokes. This delay has led to a significant backlog of surgeries, exacerbating the challenges associated with surgical scheduling. Researchers and experts have recognized the urgent need for efficient scheduling methodologies to manage the backlog effectively and ensure timely access to necessary surgical interventions [17].

The impact of the pandemic has also highlighted the need for better prioritization strategies in surgical scheduling. With limited resources and urgent cases arising from COVID-19 complications, hospitals have had to make difficult decisions regarding which surgeries to prioritize. Articles [23][26] have explored the development of algorithms and decision-making frameworks that consider factors such as patient urgency, surgical complexity, and available resources to optimize scheduling decisions and maximize patient outcomes.

Furthermore, the pandemic has accelerated the adoption of digital technologies and artificial intelligence in surgical scheduling. Articles [17][23] have discussed the integration of automated scheduling systems with electronic health records, data analytics, and predictive modeling to enhance efficiency and optimize resource allocation. These technologies enable hospitals to analyze large datasets, identify patterns, and generate optimized schedules based on real-time information, ultimately improving patient care and mitigating the impact of the pandemic on surgical services.

Kasivisvanathan and Ramanathan [26] studied surgical services during the COVID-19 pandemic based on the model and experience of the RMCancerSurgHub. This proposed model is based on the hub and spoke idea, focusing on establishing local and regional hub centers to provide urgent surgical treatment in a setting protected from the burden of COVID-19. It utilizes an extended multidisciplinary team (MDT) approach, delivering core NHS services in clean sites that can adapt to surges in viral cases. The model incorporates a clinical prioritization process to ensure equitable access within and between specialties, prioritizing patients based on the severity of their condition while minimizing exposure to the virus for those whose treatment can be safely delayed. Moreover, this model has the capacity to increase surgical activity and guide units and networks through the recovery phase.

In summary, the COVID-19 pandemic has spurred a significant increase in the number of articles addressing surgical scheduling, highlighting the challenges posed by the crisis and the importance of efficient scheduling practices. These articles emphasize the need for adaptability, prioritization strategies, and the integration of advanced technologies to optimize surgical schedules, manage backlogs, and ensure timely access to essential surgical care. The pandemic has acted as a catalyst for innovation in surgical scheduling, leading to a wealth of research and practical solutions aimed at improving healthcare system resilience and patient outcomes in times of crisis.

1.4. Aims and Goals

This Thesis aims to find a new optimization method for OR scheduling, in particular in pediatric health care. The proposed model wants to optimize surgical scheduling for day-hospital, in terms of benefits to the patients, but also on a wise utilization of resources.

The key factors which have been taken into consideration are urgency, the control of the waiting time, and the penalization of delays. To these principal aspects, the model is built based on a Hub & Spoke approach, by preferring surgical operations allocated in spoke sites.

This big problem is composed of two factors: proposing a model that respects all the requirements and the constraints given by the clinic that commissioned this research and then finding a formulation, or a solving method, that allows getting the optimal solution in a manageable amount of time. The model was structured to solve an integer programming problem and since the scheduling problem in our case is NP-hard, it is very time demanding to solve it in an analytical way, so alternative paths are required, proposing new meta-heuristic solutions.

The final step is, based on data analysis on the chosen model to draw attention to certain important details or hidden information and provide some foundation for the next strategic medical care decisions.

2 | Description and Formulations

The problem has been suggested by Ospedale dei Bambini Vittore Buzzi (V. Buzzi Children Hospital) of Milano.

We consider the problem of scheduling pediatric elective surgeries in a Hub & Spoke environment. The considered surgeries are day hospital ones, namely the intervention and the recovery time must start and finish on the same day.

We consider a set of patients P, and a time horizon, that goes from day 1 to day $|D|$. Every day has been divided into many time slots, presented in a set T. The starting time for scheduling, each of the patients and surgeons, is given by the time slot. The set of time slots T has been made for the simplicity of the modeling. It represents the working hours. Those have been divided into time slots to define a specific start time for the operation. The thicker is the partition the more accurate is the schedule, and maybe it is possible to wedge in more surgeries, but on the contrary, we significantly increase the number of variables to find, and secondly, it proposes a schedule very labile to eventual delays.

Since it is a scheduling problem, we have to decide if a patient would be operated on in the time range, where, and when, and all the necessary information to give to the patient for the surgical intervention.
Whenever a patient is not scheduled in the given time horizon, the child must be considered in the following planning horizon. For this reason, the non-allocation of the patient would be considered as an assignment to a dummy facility $fictitious$, without any specific kind of constraints to be respected. A patient is assigned to it if is not possible to schedule for any reason. For example, there could be too many surgical cases for a small time horizon, and it is impractical for any possible schedule. Or simply the best scheduling is to avoid operating on the patient and postpone the operation to the next time horizon.

Some patients are already on a waiting list at the beginning of the planning horizon. With wl_p we denote the number of days from the patient's p registration. All the operations are intended as day-hospital and emergencies are not evaluated. Any stochastical factor is not taken into consideration, for example, the variability of the operation time. Every model proposed will be built in order to find a totally deterministic schedule.

Every surgical case p also has personal information, given as parameters, regarding the age_p and the due date dd_p, which is strongly preferred to be respected, otherwise if a patient p is allocated after the dd_p, a positive integer variable δ_p will assume the value given by the difference of the scheduled day and dd_p.
There are further data related to the kind of surgery, in terms of urgency u_p, the number of necessary surgeons $ndoc_p$, the average length of the operation time τ_p, and the $PACU_p$. We recall that this last one stays for Post Anesthesia Care Unit, which means the time a patient must be in a recovery bed immediately after the surgery, in order to be monitored typically by anesthesiologists, and other medical staff.

In addition to the main sites, hubs, which could have different operating theatres, patients can be operated on in some Operating Theatres (OTs) in other hospitals, spokes, on certain days, letting some doctors move from the hub to the spokes. In the previous chapter, we have seen how important is a collaboration between hospitals from different points of view.
Let us consider a set of available hospitals Hr, each of which has its own set of OT OH_h. With set H we will denote the real facilities with the dummy one. All the OTs are in unique sets: Or if we consider all the real operating rooms, while with O if we count also one of the dummy hospitals.
It has been separated into the sets of H and Hr, and analogously O and Or because we want to clearly distinguish the real sites from the dummy ones. This decision will be very clear later on since some strict constraints must be respected in true hospitals but can be, sometimes must be, relaxed in dummy medical facilities.

Then every patient that needs surgery must be allocated to a specific operating room o, in a specific hospital center h on a specific day d, and the operation is scheduled to start at a specific time slot t. For this reason it has been proposed binary variables zy_{phd} and y_{podt} that assume value 1 if the patient p on day d is allocated in hospital h, and respectively starts his surgical operation at time t in operating room o.

The principal hub site, which forms the set HUB, would prefer to move different surgical operations to the spoke sites, since on the one hand the hub is relieved, to make it less busy, and on the other hand, having more available facilities, the patients waiting list can be completed in a reduced amount of days. It also strives to facilitate the increment of usage of O.T. that otherwise would be rarely used. But not every patient can be allocated to the spoke sites. Indeed very small children, who are younger than a threshold, that in our case is 5 years old, must be operated on in the hub, since there is more appropriate medical equipment to be used on more delicate, and physically smaller patients.

Not every day the clinic can rely on the spoke site. They are available only on certain days, for example, in our case, the Buzzi pediatric center can rely on external facilities only once a week. For this reason with a binary matrix a_{hd} we can set the availability of each facility on each day in the planning horizon.

Due to the characteristic of the different interventions, each patient p requires specific equipment, but not every medical site has all the possible types of tools for every kind of surgery. Due to these ad hoc requirements, we describe this compatibility with a binary matrix mPH_{ph}, that assumes value 1 if the machinery needed for patient p is present in hospital h and 0 if it is not.
We assume that we have a number of specific instruments in a hospital greater or equal to the number of its O.T.s. Otherwise, the schedule found could allocate two or more different cases in different O.T.s of the same hospital, who can be operated at the same instant, and there is not enough available required machinery.

Finally, every ambulatory care center provides a number of available beds b_h. Obviously, each bed for the recovery, as also the OT, can be used only for one person at a time. With a binary variable w_{phdt} we assure a positive value if patient p, in hospital h, on day d starts to use a bed for the recovery at time t

Not only the patient must be associated with a definite place and time, but also the surgeons. Surgeons, who are collected in a set S must be considered in the model since the doctors are not indistinguishable. Each of them has a particular specialization and habilitations for certain surgeries. For this reason, the mastery of the doctors must be taken into account, and possibly their availability. For the first issue, we defined a binary matrix mPS_{ps} that assumes value 1 if patient p needs a kind of surgery that can be accomplished by surgeon s, 0 otherwise. Then it must be considered the fact if a surgeon operates in a specific medical facility, the surgeon will spend the entire day in that center, and can not move to others. To allocate surgeons, it has been used the same idea for patients' schedules: it has been introduced binary variables zx_{shd} and x_{sodt} that assume value 1 if surgeon s on day d is allocated in hospital h, and respectively starts his surgical intervention at time t in the operating room o.
To have also a correct assignment of the surgeon to the medical case, it has been introduced also a binary variable l_{psdt} which is set to 1 whenever the patient p starts to be operated by doctor s on day d at time t.

In addition to the points discussed before, the hospital's scheduling program is built to respect some priorities, like urgency u_p. While in the daily scheduling, there must be respected specific ordering. The precedence on each operating day is given by the patients' ages *age*: the youngest child has priority. This is due to the fact that the younger child since he is more delicate and weak, it is preferred to let the patient stay at the hospital the less as possible, and so let the child not stay out of the home more than necessary, avoiding the incident of a delay of the ahead operations.

Sets, parameters, and variables are recapped, with a brief description, in Table 2.1, Table 2.2, and Table 2.3 respectively.

2.1. Approach

It has been decided to consider three different variations of the problem. The models we have been instructed to build are required to adhere to certain specifications.
The models are built in order to propose an OT, a date, and a starting time for every patient, recalling that an association with the dummy hospital means that the patient was not scheduled in the specific planning horizon.
The proposed schedule must respect all the constraints described previously, so:

- patients must be allocated to facilities that are available on the required day, and which possess the necessary equipment
- each patient must be operated on by a number of surgeons required by the kind of operation, and every doctor of the specific interventions must have the competencies to operate on that kind of medical case
- each bed and each OR can be used at most by only one patient at a time
- since are considered day hospital cases, the PACU, and so the relative assignment to a bed, must be done in the same instant the surgical intervention ends in the same hospital, and the discharge time must be on the same day of entry
- for each working day, patients must be scheduled in increasing order by their age
- different capacities which cannot be disrespected: the closure hour per day, and the available number of beds for each facility, in which can not be allocated a number of patients who would require a bed at the same time more than the available number of beds
- each surgeon cannot operate on in different medical facilities on the same day
- each surgeon can operate on only one patient at a time.

Respecting all these constraints we want to build a model whose aim is to optimize the scheduling under 3 main KPIs:

- operate on the patients the first as possible, especially if they have a high urgency factor
- penalize delays. These are calculated as the exceeding days of the proposed schedule from the operation's due date. Even these ones are weighted with the respective urgency factor of each patient
- strongly privileges the usage of spokes OTs instead of the hub's ones, to maximize the usage of spoke resources that are not widely used, and to allocate fewer people to the hub to relieve the crowding of a big and multi-specialized medical center.

Following also these requests we decided to approach the problem as an Integer Linear Programming problem, through three different formulations, of Surgery Planning with the Hub and Spoke (SPHS)

First Problem

This formulation is bonded by all of the constraints aforementioned.

And all three preferences requested by the clinic (minimize the waiting days, penalize delays, and prefer spoke utilization) are inserted in the objective function, which we want to minimize, in the form of different summations on weighted variables.

The models are built in order to propose a schedule, and so for each patient an OT, a date, and a starting time, are found by a binary variable, which is set to 1 if the patient is allocated there.

It is the most complete formulation, including all the constraints related to normal scheduling, the ones associated with the hospital's requirements, and some indeed related to pediatric necessities.

The presence of the constraints that assure the correct ordering on a working day is only in this formulation. It will be the main and only difference with the second model.

Second Problem

We propose a second variant in order to simplify the model, removing one constraint, which is detachable, and not strongly necessary.

We want to explore how this simpler model, without this constraint, behaves.
Comparison between these two problems will be made, analyzing the differences, to see if this second formulation brings benefits to the schedule setup, and, if so, quantify them through different Performance Measures.

Does this restriction penalize the overall scheduling, or is it roughly the same?

Third Problem

We decided to study also a different problem with respect to the original requests made by the clinic center.
It has a similar formulation but it is a separate problem since it has a different aim. Here the goal is to minimize the makespan, which is the earliest date in order to schedule all the cases.
However, despite it is built to obtain a different target, this result could help on our analysis to have a general overview of the minimum number of facilities the clinic requires and of the makespan to satisfy all the demands.

The solution found by this last formulation represents also a lower bound of the completion of all the patients on the waiting list, and so it is the minimum date to be communicated to all patients, to inform them that they will certainly be operated on by that date. This is possible only if the medical center does not regard at patients' urgency and their desired operation's due date. Here in fact the KPI would be entirely focused on the clinical interests, not considering any patient's potential benefits. The attention is totally focused on the completion time.

This problem also does not include any required ordering, like the one required in the first model.

It is not considered a fictitious hospital, where to allocate patients that cannot be scheduled along the proposed time horizon since this model aims to show how many days all the patients can be listed, and so we cannot admit the possibility of non-operability.

2.2. Common elements of the three formulations

Therefore here are listed the sets, parameters, variables defined, and constraints.

2.2.1. Common Sets

Set	Description
$\boldsymbol{Hr}$	Set of available hospitals
$\boldsymbol{H}$	Set of all hospitals, including the fictitious one, i.e. $H = Hr \cup \{fictitious\}$
$\boldsymbol{HUB}$	Subset of Hr, the hospitals that are hubs
$\boldsymbol{OH}_h$	Set of available operating theatres for each hospital h
$\boldsymbol{O}$	Set of available operating theatres, i.e. $O = \bigcup_{h \in H} OH_h$
$\boldsymbol{Or}$	Set of available operating theatres
$\boldsymbol{P}$	Set of patients
$\boldsymbol{S}$	Set of Surgeons
$\boldsymbol{D}$	Days of the time horizon
$\boldsymbol{T}$	Time slots, to partition the working day

Table 2.1: Sets

2.2.2. Common Parameters

Parameter	Description
$\boldsymbol{mPS}_{ps}$	Coverage matrix Patient-Doctor. Binary: 1 if doctor $s \in S$ can operate patient $p \in P$, 0 otherwise
$\boldsymbol{mPH}_{ph}$	Coverage matrix Patient-Hospital. Binary: 1 if the surgical case $p \in P$ can be operated on facility $h \in H$
$\boldsymbol{\tau}_p$	Operating time of surgical case $p \in P$
$\boldsymbol{numdoc}_p$	Number of doctors required to operate patient $p \in P$
$\boldsymbol{dd}_p$	Deadline of surgical case $p \in P$ in days; i.e. the day before which this surgical case must be performed
$\boldsymbol{u}_p$	Urgency. The higher the more urgent
$\boldsymbol{wl}_p$	How many days patient $p \in P$ is in the waiting list
$\boldsymbol{age}_p$	Age, in months, of the patient $p \in P$
$\boldsymbol{PACU}_p$	Post Anesthesia Care Unit of surgical case $p \in P$, i.e. time the patient requires a bed after his operation
$\boldsymbol{a}_{oh}$	Availability of a hospital $h \in H$ on day $d \in D$. Binary: 1 if it is available, 0 otherwise
$\boldsymbol{b}_h$	Number of beds for PACU for hospital $h \in H$
$\boldsymbol{mha}$	Maximum patients' age that must be operated on in a hub center

Table 2.2: Parameters

2.2.3. Common Variables

Variable	Description
$\boldsymbol{x}_{sodt}$	binary: 1 if doctor s goes to O.T. o, on day d, and starts an operation at time t, 0 otherwise. t is in $0...\lvert T\rvert - \tau_p$
$\boldsymbol{y}_{podt}$	binary: 1 if patient p starts to be operated on at O.T. o, on day d, at working time t, 0 otherwise. t is in $0...\lvert T\rvert - \tau_p - PACU_p$
$\boldsymbol{zx}_{shd}$	binary: 1 if doctor s goes to hospital h, on day d, 0 otherwise
$\boldsymbol{zy}_{phd}$	binary: 1 if patient p goes to hospital h, on day d, 0 otherwise
$\boldsymbol{w}_{phdt}$	binary: 1 if a bed is used by patient p, at hospital h, on day d, from working time t, until he finishes his recovery, 0 otherwise
$\boldsymbol{l}_{psdt}$	binary: 1 if surgeon s operates on patient p on day d starting at time t, 0 otherwhise.
$\boldsymbol{\delta}_p$	integer ≥ 0: delay with respect to the due date of the operation of patient p
closure_day	integer ≥ 0: the last day of the entire schedule

Table 2.3: Variables

2.2.4. Common Constraints

$$\sum_{h\in H, d\in D} zy_{phd} = 1 \qquad \forall p \in P \tag{2.1}$$

$$zy_{phd} = 0 \qquad \forall p \in P, \forall h \in H, \forall d \in D : mPH_{ph} = 0 \tag{2.2}$$

$$\sum_{h\in H} zx_{shd} \leq 1 \qquad \forall s \in S, \forall d \in D \tag{2.3}$$

$$x_{sodt} \leq zx_{shd} \qquad \forall s \in S, \forall h \in H, \forall o \in OH_h, \forall d \in D, \forall t \in T \tag{2.4}$$

$$\sum_{o\in OH_h, t\in T} y_{podt} = zy_{phd} \cdot a_{hd} \qquad \forall p \in P, \forall h \in H, \forall d \in D \tag{2.5}$$

$$\sum_{s\in S,} x_{sodt} \cdot mPS_{ps} \geq numdoc_p \cdot y_{podt}$$
$$\forall p \in P, \forall h \in Hr, \forall o \in OH_h, \forall d \in D, \forall t \in T \tag{2.6}$$

$$\sum_{p\in P, i\in\{max(t-\tau_p+1,0),\dots,t\}} y_{podi} \leq 1 \qquad \forall o \in Or, \forall d \in D, \forall t \in T \tag{2.7}$$

$$w_{phdt+\tau_p} = \sum_{o\in OH_h} y_{podt} \qquad \forall p \in P, \forall h \in Hr, \forall d \in D, \forall t \in T \tag{2.8}$$

$$\sum_{p\in P, i\in\{max(t-PACU_p+1,0),\dots,t\}} w_{phdi} \leq b_h \cdot a_{hd} \qquad \forall h \in Hr, \forall d \in D, \forall t \in T \tag{2.9}$$

$$y_{podt} + x_{sodt} - 1 \leq l_{psdt} \qquad \forall p \in P, \forall s \in S, \forall o \in O, \forall d \in D, \forall t \in T \tag{2.10}$$

$$\sum_{p\in P, i\in\{max(t-\tau_p+1,0),\dots,t\}} l_{psdi} \leq 1 \qquad \forall s \in S, \forall d \in D, \forall t \in T \tag{2.11}$$

$$\delta_p \geq \sum_{o\in O, d\in D, t\in T} (d - dd_p) \cdot y_{podt} \qquad \forall p \in P \tag{2.12}$$

$$(t + \tau_p + PACU_p) \cdot y_{podt} \leq |T| \cdot a_{hd} \qquad \forall p \in P, \forall h \in Hr, \forall o \in OH_h, \forall d \in D, \forall t \in T \tag{2.13}$$

$$zy_{phd} = 0 \qquad \forall p \in P, \forall h \in H \setminus HUB, \forall d \in D : age_p > mha \tag{2.14}$$

The first two constraints (2.1) and (2.2) impose not only must the patients be operated on only once, but they must be operated on in a suitable OT due to the presence of the necessary equipment for the specific kind of intervention.

Constraints (2.3) guarantee each surgeon to not work in more than one hospital center on the same day.

Equations (2.4) and (2.5) are consistency constraints, for variables x_{sodt} and zx_{shd}, and

y_{podt} with zy_{yhd} respectively. They also guarantee to not use hospitals when they are not available. The same is also (2.10), which connects the variable l_{psdt} with the variables of the starting operation of the patient, y_{podt}, and the surgeon x_{sodt}. Restriction (2.5) should be completed with multiplication with the availability parameter a_{hd} since we must control that a patient is allocated to an open hospital.

Equations (2.6) state that each patient must be operated on by as many doctors as the number required from the specific treatment.

Equations (2.7) impose that in each operating room, there can not be more than one operation at each time instant. The same is also forced for surgeons by (2.11), and analogously to the maximum available beds for Post Anesthesia Care Unit for each hospital by (2.9).

The number of used beds for post-surgery recovery, to avoid going over the total beds, is set by equation (2.8).

Inequalities (2.14) are extra constraints that follow this policy: patients younger than 5 years old ($mha = 60$ months), must be operated on in a hub site, for the reason already described.

Equations (2.13) state that the termination of each surgery, which consists of operation time and PACU, must terminate before the hospital closure hour, if the clinic is open, i.e. $a_{hd} = 1$.

2.3. Problems and Models

Hereafter are proposed three different versions of the problem described above. The first is aimed at finding the best scheduling according to the clinic's requests.

The second one is very similar to the previous one, but we want to evaluate the impact of the constraints patient's age. Could this relaxation have a substantial improvement on the complete schedule? Or on the efficiency and speed of the solution?

The third variant has a completely different approach, we want to compute the minimum number of days to operate on all the patients.

2.3.1. Formulation Problem 1: $SPHS$

For this formulation, we need to add two more parameters presented in Table 2.4, which are two multiplicative factors, needed to give weights to different parts of the objective function: α_h is a vector that for each hospital gives a "cost" of utilization. Since the clinic strongly prefers to use spoke resources, α_h will assume a low value if the hospital h is a spoke, and a higher value if it is a hub. Then we want to penalize postponing the surgical interventions in the next time horizon, so if h is the dummy hospital, the cost will be much higher.

Parameter	**Description**
$\boldsymbol{\alpha}_h$	Multiplicative factor for the objective function to identify spokes and hubs
$\boldsymbol{\gamma}$	Multiplicative factor for the objective function which controls the delays

Table 2.4: Other parameters for $SPHS$

But the biggest difference that distinguishes $SPHS$ from $SPHS\_noYF$ is the presence of constraints (2.15), which impose an ordering on a specific day, such that patients are operated on in increasing order by their age.

$$\begin{aligned} &\sum_{t\in T}(t+\tau_i)\cdot y_{iodt} \leq \sum_{t\in T} y_{jodt} + 2|T|\cdot(2-\sum_{t\in T} y_{iodt} - \sum_{t\in T} y_{jodt}) \\ &\forall o\in O, \forall d\in D, \forall i,j\in P : age_i \leq age_j \end{aligned} \tag{2.15}$$

This formulation has the aim to find the best scheduling to minimize the following objective function:

$$\min\ objfun2 = \sum_{p\in P}(u_p \cdot \sum_{o\in Or, d\in D, t\in T} d\cdot y_{podt}) + \gamma\cdot\sum_{p\in P} u_p\cdot\delta_p + \sum_{p\in P, h\in H, d\in D} zy_{phd}\cdot\alpha_h \tag{2.16}$$

It is essentially composed of 3 summations: the first two are referred to the urgency of each patient. Specifically, the first one is about the preference to avoid letting a patient not wait too long to be operated on, especially if the child has a higher urgency. While the second one refers to the delay time: the more a specific surgery is urgent, the more its operation day tends to be not after the due date, or at least to have a small delay.
The last one specifies the preference for using spoke sites over the hubs, which means paying a higher cost in terms of objective function if the hospital h where patient p is allocated, is a hub, and paying an even more expensive price if the patient is assigned to

the fictitious medical facility.

These parts could be multiplied with parameters, to give more weight to specific components, i.e. more importance to a specific section, regarding the clinic's priority.

2.3.2. Formulation Problem 2: $SPHS\_noYF$

The second variant is very similar to the first one. It has got the same formulation, the same objective function (2.16), and all the constraints from (2.1) to (2.14), but without the added constraint (2.15), that from now on will be noted as $YoungerFirst$ constraints. We have proposed this second problem, as a little variation of the previous model since we want to examine if there is a potentially significant improvement in the overall schedule due to relaxing the constraint related to the daily priority given by the age of patients.

Since the objective function does not account for the starting time of the scheduled operation, we can consider reordering the daily schedule at a second moment. But the ordering of the surgeries given by the solution may not be feasible for (2.13).

Looking at this simple example. Consider two patients, A and B, with these parameters: $\tau_A = 3$, $PACU_A = 2$, $\tau_B = 2$, $PACU_B = 1$. Suppose also $|T| = 6$, and a fixed day and an O.T.
We essentially have two different ways to schedule them: A before B, or vice versa, as presented in Figure 2.1 and Figure 2.2.

Figure 2.1: Example scheduling 1

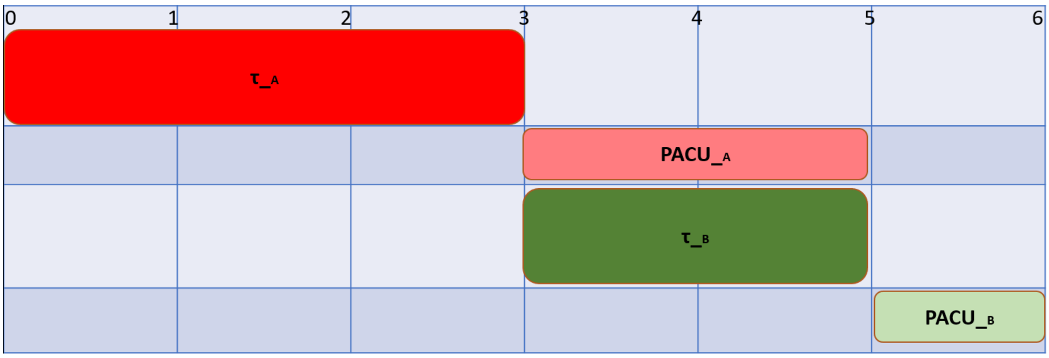
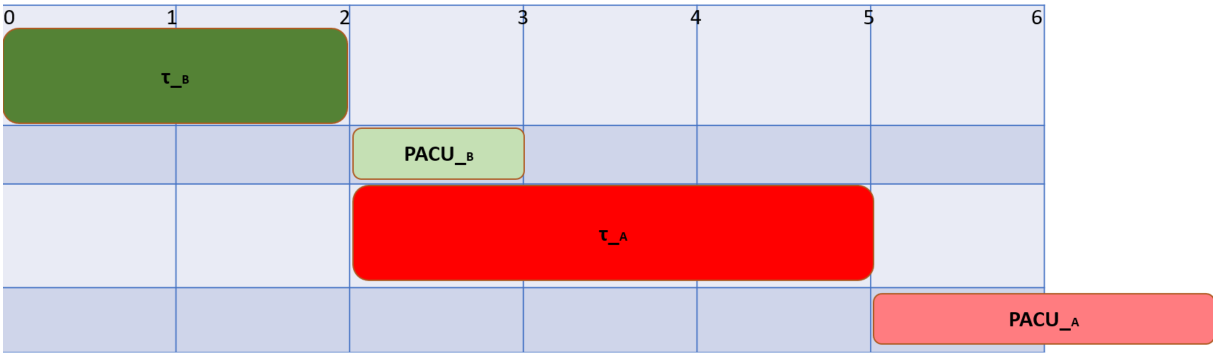


Figure 2.2: Example scheduling 2



As we can easily see, while the first schedule is feasible, the second one shows how, the same surgeries, but in a different order, violate constraints (2.13), and make the setup impossible.

2.3.3. Formulation Problem 3: $SPHS\_MS$

This third formulation is not so different from the previous ones, despite the goal is different. All the structure is the same proposed in Section 2.2, but with two new elements.

The first one is a new variable $closure\_day$, a positive integer value that indicates the last day of the entire schedule.

Variable	Description
closure_day	integer ≥ 0: the last day of the entire schedule

Table 2.5: Added Variables for $SPHS\_MS$

This variable essentially stands for the makespan, which means that every surgical operation must be scheduled before the $closure\_day$. And this is given by the equation (2.17) which guarantees the definition of the makespan of the planning.

$$closure\_day \geq d \cdot zy_phd \qquad \forall p \in P, \forall h \in H, \forall d \in D \tag{2.17}$$

This could have been seen as a lower bound. It means that if the clinic center does not consider any economic and space advantages, and treats all the care recipients in the same manner, how many days are required to operate on all the patients? We are essentially calculating the makespan of setting up all the cases, with the above-mentioned hypothesis.

Model name	Objective function	s.t constraints
$SPHS$	(2.16)	(2.1)-(2.14),(2.15)
$SPHS\_noYF$	(2.16)	(2.1)-(2.14)
$SPHS\_MS$	(2.20)	(2.1)-(2.14),(2.17),(2.18)

Table 2.6: Formulation of all models

In this way, the clinic could give eventually to the patient a maximum number of waiting days, from the specific time horizon.

Since we want to know in how many days the net of hospitals could operate on all the patients we do not consider any fictitious hospital where to allocate patients that cannot be scheduled along the proposed time horizon.
For this reason, constraints (2.18) that force patients to not be allocated to a dummy hospital have been added to previous ones:

$$y_{podt} = 0 \qquad \forall p \in P, \forall h \in H \setminus Hr, \forall o \in OH_h, \forall d \in D, \forall t \in T \tag{2.18}$$

The second main difference is the objective function. Here, in fact, the clinic center disregards the spoke preference, urgency, and delays, to treat all care recipients equally. The attention is totally focused on the completion time.
The objective function so is constructed as follows:

$$\min objfun\_of\_cl = closure\_day \tag{2.19}$$

However with this formulation, after several tests, we recognized that it was very time demanding. For this reason, a valid alternative to the objective function. has been proposed. The urgent patient precedence criteria have been included but with a lower weight with respect to the value of closure day. The new objective function became:

$$\min objfun3 = \sum_{p \in P} (u_p \cdot \sum_{o \in Or, d \in D, t \in T} d \cdot y_{podt}) + \beta \cdot closure\_day \tag{2.20}$$

with β a multiplicative factor that gives more weight to the total objective function.

This last formulation of the objective function brings two advantages simultaneously: reducing the solve time and giving already a schedule, really similar to the one we look for. From now on, this last formulation will be used when referring to $SPHS\_MS$.

2.4. Other Problems

When it comes to optimizing the operation schedule, various criteria are employed, as previously discussed while considering the diverse key performance indicators (KPIs) utilized by hospitals. This underscores the fact that there is no one-size-fits-all approach to achieving an ideal operation schedule. Instead, hospitals employ different criteria and metrics based on their unique priorities and objectives.

These optimization criteria can encompass a wide range of factors, such as minimizing patient waiting times, maximizing the utilization of operating rooms and resources, reducing surgeon idle time, optimizing the allocation of doctors and surgical teams, and ensuring efficient use of equipment and facilities. Each criterion represents a distinct aspect of operational efficiency and effectiveness within the context of hospital scheduling.

Furthermore, the selection of specific criteria is influenced by various factors, especially by the overall goals and priorities of the healthcare institution. As a result, different hospitals may prioritize different KPIs and adopt distinct optimization approaches based on their specific needs and circumstances.

By acknowledging the existence of diverse optimization criteria and KPIs, healthcare professionals and administrators can tailor their operation schedules to align with their organizational objectives and deliver optimal patient care.

For each criterion we want to adopt, we will propose alternative objective functions, and give a general overview of all of the proposals in the next Chapter, since for some of them, new variables are required.

3 | Relaxation and Alternative Formulations

The formulations described in Chapter 2 turned out to be very time-consuming even with a hundred patients.

So before trying to solve completely this with some approaches, like heuristics or matheuristic ones, we tried to compute relaxations of the problems in Section 3.1.
We also tried to modify the formulations in Section 3.2.
Indeed in Section 3.3 will be shown different objective function proposals, to essentially improve the solve time of specific models.

3.1. Relaxation: Neglecting Physicians

So far, the model has been built in order to find the best scheduling considering both the patients and the doctors. Surgeons must be taken into because every surgery has different characteristics, and they can be treated only by specialized doctors. The first relaxation proposed is to remove all the sets, variables, and related constraints regarding the physicians. It means variables x_{sodt}, zx_{shd}, l_{psdt} and constraints (2.3), (2.4), (2.6), (2.10) and (2.11).

However, surgeons and related associations with the patients and OTs cannot be simply discarded, for the reason above mentioned. For this reason, a relaxation and a method of two phases have been proposed: relaxation and then an assignment of surgeons a posteriori.
Since it is a heuristic approach, it will be largely discussed in the next chapter.

Before showing the results, and the eventual improvements, let us see how much impact this relaxation has in terms of the number of variables. This last formulation has an order of $\mathcal{O}(P \cdot (ODT + HD + HDT + 1))$, instead of $\mathcal{O}(P \cdot (ODT + HD + HDT + 1) + S \cdot (ODT + HD + PDT))$.

By neglecting doctors we decrease the number of nodes but effectively we eliminate all those nodes which present symmetries between them. In fact not only some doctors can be exchangeable, but can be considered more doctors than necessary.

From now on, whenever this relaxation is used, it will be made explicit with the suffix *nodoc* attached to the name's model.

Model name	Objective function	s.t constraints
$SPHSnoc$	(2.16)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14), (2.15)
$SPHS\_noYFnodoc$	(2.16)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14)
$SPHS\_MSnodoc$	(2.20)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.17),(2.14),(2.17),(2.18)

Table 3.1: General Formulation of $SPHSnodoc$

3.2. Bin-Packing-based enriched formulation

We now propose an alternative formulation of the problems, enriched with variables and constraints.

Every OT, each day, can be seen as a bin, with some capacity limit, for instance, the working day hours. So we can look at which day d, which operating theatre o should be open, by setting a variable yBP_{od}. Consequently, we must see the allocation of the patient p to the bin opened, through a variable xBP_{pod}.

For this reason, it has been decided to add variables and constraints correctly associated with the Bin Packing formulation to the models already written and described above, instead of substituting them. And the same holds also for the list of constraints used for every model proposed.

To the formulations in section 2.2.3 ($SPHS$,$SPHS\_noYF$ and $SPHS\_MS$) these two variables:

Variable	Description
$\boldsymbol{yBP}_{od}$	binary: 1 if the bin, i.e. the OT o on day d is open, 0 otherwise
$\boldsymbol{xBP}_{pod}$	binary: 1 if the patient p is allocated in the bin identified by OT o and day d, 0 otherwise

While the restrictions to add to the list of the already present constraints for a specific model are the following:

$$\sum_{o \in OH_h, d \in D} xBP_{pod} = 1 \qquad \forall p \in P \tag{3.1}$$

$$\sum_{p \in P, t \in T} y_{podt} \leq yBP_{od} \cdot |T| \qquad \forall o \in O, \forall d \in D \tag{3.2}$$

$$\sum_{p \in P} \tau_p \cdot xBP_{pod} \leq yBP_{od} \cdot |T| \qquad \forall o \in O, \forall d \in D \tag{3.3}$$

$$\sum_{p \in P} PACU_p \cdot xBP_{pod} \leq yBP_{od} \cdot |T| \qquad \forall o \in O, \forall d \in D \tag{3.4}$$

Equations 3.1 state that every patient must be allocated to exactly one bin.

Constraints 3.2, 3.3 and 3.4 guarantee three different capacities: the number of patients allocated in a specific bin cannot exceed the total number of working time slots, the sum o operation times τ_p, and $PACU$ of all patients assigned to that bin can not overcome the overall amount of labor hours. These last two can be improved by substituting $|T|$ with $|T|/(\min \tau)$, and respectively $|T|/(\min PACU)$.

We have added two variables xBP_{pod} and xBP_{pod}, and appropriate constraints.

While for the objective function, obviously depends on which model we are using, and so which the aim is. If we use the obj. fun. considering the classic Bin Packing problem, we would write:

$$\min objfunBP = \sum_{o \in O, d \in D} yBP_{od} \tag{3.5}$$

which is the goal of the $SPHS\_MS$, since we want to minimize the number of days for completing the entire set of clinical cases.

But this formulation presents an error if the problem to solve is $SPHS\_MS$. With this objective function, it is minimized the number of bins, but there's any imposition about the minimization of the makespan. For the solver each bin, on each day has the same weight, while we want to reduce as much as possible the number of days.

Also computationally speaking, it could be very onerous, since the solver explores different nodes which bring the same result. We essentially have to break the symmetries given by the (almost) equality between the various bins. And so maybe give a weight directly proportional to the day.

Breaking symmetries is a convenient approach to enhance the efficiency and effectiveness of the learning process. Symmetries refer to situations where multiple hypotheses or solutions exist that are essentially equivalent or indistinguishable from each other.

In this way we are essentially reducing the search space, allowing the solver to explore a more focused and manageable set of hypotheses.

Secondly, redundant solutions are avoided. Since most of the bins are indistinguishable, breaking symmetries can be eliminated redundant solutions and focus on finding unique and informative hypotheses.

So the first intuitive new objective function proposed is the following:

$$\min objfunBP22 = \sum_{o \in O, d \in D} yBP_{od} \cdot d \tag{3.6}$$

But it has been proposed also different formulations which will be shown more in detail in the next section, describing different proposals of the objective function.

Model name	Objective function	s.t constraints
$SPHS\_BP$	(3.6)	(2.1) - (2.14),(2.15),(3.1)-(3.4)
$SPHS\_noYF\_BP$	(3.6)	(2.1)-(2.14),(3.1)-(3.4)
$SPHS\_MS\_BP$	(3.6)	(2.1) - (2.9), (2.10) - (2.14),(2.17),(2.18),(3.1)-(3.4)

Table 3.2: General Formulation of $SPHS\_BP$

The pure Bin Packing problem can be seen as the lower bound of the solution if the question is to find the makespan. In fact, if the clinic center makes the deliberate choice to prioritize equal treatment for all care recipients, disregarding individual preferences, urgency, and potential delays and the primary focus is solely on minimizing the overall completion time, every patient will be treated in a maximum $closure\_day$ days.

It is important to note that while this approach promotes equality and efficiency, it may result in varying waiting times for individual patients. By focusing on the completion time and considering the maximum number of days a patient may have to wait, the clinic strives to strike a balance between fairness and operational effectiveness within the given time horizon.

From now on, whenever it is used the Bin Packing approach, it will be made explicit with the suffix $\_BP$ attached to the model's name.

3.3. Alternative Objective Functions

Before guessing and trying the best objective function, for each model, it is necessary to understand what is important to prioritize for the hospital, if the avoidance of delays, the number of days used to meet up all the patients' requests, or anything else. Then since some objective function formulations take into account many factors, like the ones aforementioned, it is important to calibrate the weights associated with every part of the o.f.
Moreover, after understanding carefully which are the preferences of the management, it can be discussed whether some little editing can or not help the solver to find faster a good solution, without changing significantly the overall results.

In previous chapters, these decisions have been presented as Performance Measures (PM), also known as a key performance indicators (KPI). The performance measures refer to the metrics used to evaluate the planning and scheduling processes in operations research. These measures assess various aspects such as computational efficiency and computational time

Performance Measures play a crucial role in OR planning as it has been discussed before, as they provide a basis for comparing different methodologies, algorithms, and models. These measures are also used by researchers to evaluate the performance of their proposed solutions and to demonstrate the effectiveness of their approaches in addressing real-world problems.

Here we focus essentially on some of the aspects described in the literature, each with its own specific weights and importance.
Objective achievement, which evaluates the extent to which the desired objectives, such as minimizing costs, maximizing resource utilization, or optimizing schedules.
Utilization measures the employment of resources and assesses whether they are effectively and efficiently utilized. Here, since we want to build a model based on Hub & Spoke approach we want to widely take advantage of the OT borrowed from other's medical facilities.
Waiting times. They quantify times experienced by patients in the scheduling process, providing insights into the efficiency of the system and patient flow.

Here are reported some proposals for different objective functions: especially for the third formulation, and some of them referred to the Bin Packing formulation:

$$\min\ objfuncIday = closure\_day \tag{3.7}$$

This o.f. 3.7 essentially aims to minimize the makespan, without considering the patient's or hospital's benefits.

$$\min\ objfunBinPacking = \beta \cdot \sum_{o\in O, d\in D} yBP_{od} + \sum_{p\in P} u_p \cdot (\sum_{o\in O, d\in D, t\in T} d \cdot y_{podt}) \tag{3.8}$$

$$\begin{aligned}\min\ objfunBinPacking\_m12 = &\sum_{p\in P} u_p \cdot (\sum_{o\in Or, d\in D, t\in T} d \cdot y_{podt}) + \gamma \cdot \sum_{p\in P} u_p \cdot \delta_p + \\ &\sum_{p\in P, h\in H, d\in D} zy_{phd} \cdot \alpha_h + \\ &\beta \cdot \sum_{o\in O, d\in D} yBP_{od}\end{aligned} \tag{3.9}$$

$$\min\ objfunBinPacking2 : \sum_{o\in O, d\in D} f(d) \cdot yBP_{od} \tag{3.10}$$

These last o.f.s refer to the Bin Packing formulations.
The equation (4.3) takes into consideration the urgency and prioritization of the patients, and the number of bins used, recalling that a bin is uniquely identified by the OT and the day. The first part, which is multiplied by factor β, does not determine the makespan. In fact, it does not give any preference about the day, to minimize the completion time, but only the number of opening OTs. The variable yBP_{od} can be multiplied by the index d of the first summation, in order to minimize the sum of the days when bins are open, and so, given for free, the maximum day of opening. But it has been noticed by comparing these two variations that they are the same. Effectively the second part already contains the day factor and implicitly has got the aim to minimize the makespan.

The $objfunBinPacking\_m12$ in (3.9), is a little bit more complete, taking into account four different factors. Both the first two parts take as Performance Measure the urgency of the patients in two different forms: one, as the same as the previous formulation, prefers to not allocate patients too far in the time horizon, and the other tends to keep control of delays, avoiding the deferral of the interventions respect to the due date dd_p.

The third component is referred to the Hub & Spoke modeling, preferring to allocate surgical operations in the spokes sites, with respect to the hubs' ones. While the last one is completely equal to the part described in (4.3), describing the minimization of the working OT used.
This formulation has the same structure of the objective function (2.16) presented in $SPHS$ and $SPHS\_noYF$ in Section 2.3, with only the addition of the part regarding the reduction as much as possible of the number of bins used.

The o.f. (3.10) is a variation of the one proposed in *model_clday*. Here our PM is only the makespan, using the Bin Packing approach. We want to discover if it would be a valid alternative to the objective function proposed in (2.19) which has been revealed very time-demanding, even with a small dataset. To avoid the problem well described above, and so to consider the bin all equals, without imposing the preference of minimizing the number of used days, we have to multiply the variable yBP_{od} to a factor, that depends on the day. It was already proposed a linear weighting, related to the day. Now we propose different functions $f(d)$, not linear anymore. Polynomial functions have been tried, at different orders, and exponential ones, to see if and how at the variation of f the solving time differs.
The functions tested, and with reported results, were essentially three: two polynomials of different orders, to see what happens to the compile time at the increasing order of the function, and an exponential function. They are here proposed in expanding order of growth, to spot if there are evident variations in terms of solving time. It has been chosen $f(d) = d^2$, $f(d) = d^5$ and finally $f(d) = 2^d$.

As it has been said: when it comes to optimizing the operation schedule, various criteria are employed. To recap the criteria that we wanted to adopt, were the minimization of patient waiting times, weighted with his/her urgency, in the time horizon (1), the penalization of eventual delays, also weighted with the urgency (2), the Spokes predilection to the Hubs (3), the total make-span, i.e. the closure day (4), and the total number of OTs used along the time horizon (5). A recap of all the different objective functions proposed for each is presented in Table 3.3.

Each distinct objective function introduces specific enhancements based on the individual requirements it aims to address.

OBJECTIVE FUNCTION	FORMULATION	CRITERIA
objfun2	$\sum_{p\in P} u_p \cdot (\sum_{o\in Or, d\in D, t\in T} d \cdot y_{podt})$ $+\sum_{p\in P} u_p \cdot \delta_p$ $+\sum_{p\in P, h\in H, d\in D} zy_{phd} \cdot \alpha_h$	(1),(2),(3)
objfun3	$\sum_{p\in P} u_p \cdot (\sum_{o\in O, d\in D, t\in T} d \cdot y_{podt})$ $+\gamma$*closure_day	(1),(4)
objfunBinPacking	$\sum_{o\in O, d\in D} \beta \cdot yBP_{od}$ $+\sum_{p\in P} u_p \cdot (\sum_{o\in O, d\in D, t\in T} d \cdot y_{podt})$	(1)
objfunBinPacking2	$\sum_{o\in O, d\in D} d \cdot yBP_{od}$	(5)
objfunBinPacking2*	$\sum_{o\in O, d\in D} f(d) \cdot yBP_{od}$	(5)
objfunBinPacking2**	$\sum_{o\in O, d\in D} 2^d \cdot yBP_{od}$	(5)
objfunclday	closure_day	(4)
objfunBinPacking_m12	$\sum_{p\in P} u_p \cdot (\sum_{o\in Or, d\in D, t\in T} d \cdot y_{podt})$ $+\sum_{p\in P} u_p \cdot \delta_p$ $+\sum_{p\in P, h\in H, d\in D} zy_{phd} \cdot \alpha_h$ $+\sum_{o\in O, d\in D} \beta \cdot yBP_{od}$	(1),(2),(3),(5)

Table 3.3: Different Objective Functions

4 | Heuristic Approaches

It has been decided to find enriched formulations, with the hope to reach the optimal solution, sometimes just feasible ones, in a fewer amount of time. In this part, various possible improvements are essentially studied, especially of the third model, which aim is to reduce the makespan combined with optimal scheduling to avoid too many delays.

So in this chapter, some heuristic methods will be presented, such as the greedy in Section 4.1, and some matheuristics ones in Sections 4.2 and 4.3. It has been developed a base relax heuristic in Section 4.4.

4.1. Greedy

Particularly for the third formulation, an initial approach to achieve a favorable makespan could be through the implementation of a greedy algorithm. This method involves sequentially considering patients and, for each day, hospital, operating theater (O.T.), and time slot in a specific order. The objective is to assess whether adding the surgical case at those particular indexes would result in available space, ensuring that all capacities and constraints are respected. It has been used a First Fit heuristic approach.

The process begins with the first patient and proceeds iteratively. For each day, hospital, O.T., and time slot in the specified order, the algorithm checks if there is sufficient space to accommodate the surgical case by verifying if all capacities, such as the maximum number of surgeries allowed per day or the availability of O.T. time slots, are adhered to. If the conditions are met, the patient is scheduled for the specific O.T. at the designated date. The algorithm then moves on to the next patient, repeating the process.

However, if the proposed indexes do not satisfy the capacity constraints, the algorithm rejects those specific indexes and continues to the next available option in the reverse order of the indexes. This ensures that alternative scheduling options are explored to find feasible solutions.

To have a priority rank, hoping for a better schedule, we ordered patients in a decreasing way of their urgency coefficient.

By employing this greedy algorithmic approach, a rudimentary yet efficient starting point, in terms of makespan, can be achieved. It allows for systematic scheduling based on available space while respecting the capacity limitations of each day, hospital, O.T., and time slot. This initial scheduling arrangement forms the foundation upon which further optimization and adjustments can be made to refine the overall scheduling process and minimize the makespan.

Listing 4.1: Greedy

```
let a flag=0
for each p in P
    if age[p] <= mha
        for each d in D
            for each h in HUB
                for each o in OH[h]
                        for each t in T
                        if setting y[p,o,d,t]=1 are respected:
                            capacity of each OT,
                            time limit of closure hour, and
                            beds availability for patient PACU time
                        then add the patient in the schedule, namely:
                        let y[p,o,d,t]=1
                        let zy[p,h,d]=1
                        let w[p,h,d,t+tau[p]] = w[p,h,d,t+tau[p]]+1
                        let flag:=1
                        break and start with the next patient;
    if age[p] > mha
        for each d in D
            for each h in H
                for each o in OH[h]
                        for each t in T
                        if setting y[p,o,d,t]=1 are respected:
                            capacity of each OT,
                            time limit of closure hour, and
                            beds availability for patient PACU time
                        then add the patient in the schedule, namely:
                        let y[p,o,d,t]=1
                        let zy[p,h,d]=1
```

```
let w[p,h,d,t+tau[p]] = w[p,h,d,t+tau[p]]+1
let flag:=1
break and start with the next patient;
```

4.2. Decomposing

From the very first result analyzed and discussed in the next Chapter, and as we could expect, the more the number of patients grows, the more the compiler is time demanding. And it does not raise linearly, but exponentially, since it is solved with a branch and bound method. So the first improvement proposal is to divide the dataset into two different partitions, to solve sequentially.

A first approach, only to see if it could have a little improvement, the partitions are based simply only on the order of the ID of the patients, without taking into consideration any other features. The initial part is the first half of the total instance, while the second takes the other half.

It is essentially the same problem we have discussed so far, but divided into two steps, in order to have fewer parameters for each of them. We solve the first one, find the schedule, and attach to it the setup found with the remaining surgical cases.

The next analysis has been done with different partitions, considering for example the due date and/or the level of urgency, clearly the more urgent the patient's condition, the first is allocated, and analogously with the smaller due date.

We could think about how many partitions the dataset should be divided in order to obtain the best trade-off between optimum and compiling time. After careful consideration and analysis, the decision has been reached to avoid dividing the dataset into numerous partitions and keep the number of splits equal to two. While it is acknowledged that this approach may result in a reduction of compilation time, as there would be fewer variables to schedule, it has been recognized that the overall optimization of the entire scheduling process would be compromised due to the decreased number of patients involved.

Dividing the dataset into multiple partitions initially seemed appealing as a means to streamline the scheduling system. With fewer variables to consider, the compilation time would naturally decrease, potentially leading to more efficient scheduling operations. This reduction in compilation time could be perceived as a notable advantage, as it would expedite the process and allow for quicker decision-making.

However, upon deeper evaluation, it became apparent that such partitioning would have unintended consequences. By reducing the number of patients within each partition, the total optimization of the scheduling system would suffer. A smaller pool of patients means that the system would have limited flexibility and fewer opportunities to find optimal scheduling solutions.

Optimizing the entire scheduling process relies on analyzing and considering a comprehensive dataset, encompassing a diverse range of patient characteristics, requirements, and preferences. By encompassing a larger pool of patients, the scheduling system gains access to a wider array of variables, allowing for a more comprehensive and effective optimization process. This holistic approach ensures that the scheduling system can consider various factors to generate the most optimized schedules.

In summary, while dividing the dataset into multiple partitions may offer a reduction in compilation time, it would compromise the overall optimization of the scheduling process. By maintaining a larger dataset and avoiding such partitioning, the scheduling system can capitalize on the full range of variables, leading to more effective and comprehensive scheduling solutions.

4.2.1. Sprint reducing the number of days

Since we are decomposing the dataset, the number of patients is reduced, so it is not necessary to explore all the days in the time horizon. For this reason, to improve the speed of reaching the best solution, the number of days of the time horizon has been reduced, decreasing significantly the number of parameters.

Listing 4.2: Decompose and Sprint

```
let c multiplicative factor, that depends on the split
let days=|D|
Decompose the dataset into two partitions: P1 and P2
Solve the problem P1 in a small time
Save the closure_day found so far, and save bestday=closure_day
For the model with partition P1 set a time limit of 1000,
 and update the set D={1..bestday}
Solve the model with P1
if {bestday*c < days}
    set D={closure_day+1..bestday*c}$
else
    set D={closure_day+1..days}$
For the model with partition P2 set a time limit of 1000,
```

```
  and update the set D decided in the previous step
Solve the model with P2
```

4.3. Associations Surgeons

As we previously said, the doctors' scheduling cannot be overlooked. To reach the entire feasibility of the problem either patients and surgeons must be correctly allocated.

Before showing the methods utilized to assign medical staff, in our case only doctors, to OTs and to clinical cases, let us see why this step can not be ignored, or given to it less importance than actually it requires.

The methods utilized to assign medical staff, in our case, only doctors, to OTs and to clinical cases will be discussed in chapter 4, but firstly let us see why this step can not be ignored, or given to it less importance than actually, it requires.

In fact, if we consider also the availability of medical personnel, the entire scheduling, divided into these two sequential steps, an unfeasibility can incur.
Let us see how this inconvenience can happen through this example. Consider a solution to the first part of the problem, i.e. the surgical cases scheduling. Let us assume that the result assigns on a specific day, in a precise operating room, patients with the same kind of operation, maybe rare, let us call it kind R. Suppose that every doctor specialized in interventions of kind R is unforthcoming on that day, or, in another scenario, the solution decided to allocate surgeries of this kind of operation simultaneously in many different OTs, but there is no sufficient availability of specialized staff. Since the solution proposed in the first part did not take into account this factor, it is feasible with this restriction, but the entire problem turns out to be unfeasible.

If this error should occur, there are two possible ways to solve this contradiction. One depends on the clinic's operational decision management and the other one from the model formulation. In the first case, the medical center has to look carefully at the costs-benefits in order to understand if it is convenient to make internal proposals to fill in that time slot.
The other option is to add to the model constraints that violate the allocation of patients that cannot be satisfied by doctors and retrain the first part followed by the second one until a feasible solution is found.

The approach that has been used for solving this challenge is the following: to partition the formulation into two different autonomous problems.

The first one has variables y_{podt}, zy_{phd}, w_{phdt}, δ_p and $closure\_day$
The constraints are all the ones specific for each model, that do not have the variables regarding the doctors. The objective function is the same used for the entire problem, whatever formulation we choose, since in every o.f. proposed so far there is any element that keeps control of the medical staff.

The second step is a sort of complementary problem: our variables and constraints are all the ones previously discarded and the same holds for the list of constraints. This has been approached as a facility location problem, in fact, we have to associate surgeons with scheduled medical cases, where each of them has specific necessities and requirements, like the number of medical staff, or the kind of specialist, on the basis of the surgical intervention.

For what concerns the objective function of this latter problem, it has been proposed the minimization of the total number of doctors, for each hospital and day, used to satisfy all the demands, i.e. :

$$\min objfun\_mindoc = \sum_{s\in S, h\in H, d\in D} zx_{shd} \tag{4.1}$$

It has been used variable zx_{shd} instead of x_{sodt} since the idea is to minimize the number of surgeons in a specific working day, and not only the number of them for meeting every patient's need.

From now on we consider this model as $model_S A$, for indicating the surgeons association.

4.4. Forcing the solution of $SPHS\_noYF$ with $youngerfirst$ constraints

We have already talked about the difference between $SPHS$ and $SPHS\_noYF$, and the advantages that this relaxation brings. As the results in the next chapters will show, the second formulation, which excludes the $youngerfirst$ 2.15 constraints from the original formulation, has a better optimal solution, and it is also found in a lower time than the first one.

The proposed approach involves finding a solution to the formulation without incorporating the $youngerfirst$ constraints initially. Subsequently, in the second step, the solution is modified to ensure compliance with the $youngerfirst$ restriction.

The intention is to explore a solution space without the initial limitations imposed by the

$younger first$. By doing so, we can identify potential solutions that may not have been apparent or accessible under strict adherence to those constraints.

Once an initial solution is obtained, the subsequent step involves enforcing the $younger first$ constraints upon it. This means adjusting or modifying the solution to align with the specific requirements and limitations outlined by the pediatric needs.

By adopting this two-step approach, we enable a more comprehensive exploration of the solution space, potentially discovering innovative and effective strategies that would have otherwise been overlooked. It allows for flexibility, striking a balance between exploring unconstrained possibilities and eventually incorporating the necessary constraints to meet the desired objectives.

Inevitably, when initially exploring the solution space without the $younger first$ constraints, there is a possibility that the obtained solution may not satisfy all the restrictions, just think about the capacity of the working hours, recalling the example proposed in Figure 2.1 and Figure 2.2. In such cases, an assessment is required to determine the number of additional hours needed to make the solution feasible while ensuring compliance with the constraints.

After applying the $younger first$ constraints to the solution obtained in the first step, a careful analysis is conducted to identify any violations or discrepancies. By comparing the modified solution with the constraints, it becomes possible to quantify the extra time required to render the solution feasible.

This evaluation involves pinpointing the time capacities that initially succeeded and calculating the additional hours needed to rectify the situation. By quantifying the extra time required, the clinic or organization can make informed decisions regarding resource allocation, scheduling adjustments, adding extra working time, or any other necessary measures to ensure that the solution becomes feasible while still adhering to the $younger first$ constraints.

In summary, by assessing the solution obtained without the $younger first$ constraints and determining the violations, the calculation of extra time needed provides valuable insights into the adjustments necessary to make the solution feasible. This analysis enables the clinic to address any shortcomings, optimize scheduling, and ultimately reach a solution that satisfies all the constraints while minimizing additional time requirements.

Technically we set the formulation in the following way: we run the $SPHS\_noYF$ first, then save the results of variables zy_{phd}, fixing for each patient, the date, and the medical center found by the solution. Then the second step is to run the $SPHS$, with the values of zy_{phd} unchangeable, and consider them as a parameter instead of variables. Then the problem is only to give the perfect order of the new schedule, imposing the $youngerfirst$ constraints, and eventually rescheduling the medical staff, to avoid unfeasible overlaps. Before doing this it is necessary to consider a new set of time slots, say $T1$, such that $|T1| > |T|$ since we have to allow patients, to be reallocated, following the ordering given by their age, to avoid the unfeasibility of the problem.

We could run the $SPHS$ as a second step, to assure the wanted order. Anyway, it has been found a better approach for this last step. A more efficient method is to run a new model, with the same variables and constraints of $SPHS$, and the variables zy_{phd} fixed, but with a new variable, a new constraint, and a different objective function. We actually indicate that we want to minimize the eventual extra hours.

We add a new integer variable c_{od} that says the closure hour of each OT on each day, by setting this constraint

$$c_{od} \geq (t + \tau_p + PACU_p) \cdot y_{podt} \qquad \forall p \in P, \forall o \in O, \forall d \in D, \forall t \in T \tag{4.2}$$

And with the objective function:

$$\min \ objfun\_clhour = \sum_{o \in O, d \in D} c_{od} \tag{4.3}$$

The last step is to count and quantify the exceeded hours, of this forced solution. Then the clinic, following its appropriate KPI, will decide what is the best solution to adopt, according to its Performance Measure, if to assume the possibility to give its medical staff some extra working time, to do not consider the new order, and keep the solution found by $SPHS\_noYF$, redo the problem, considering all the constraints all in one, or to find a trade-off.

A possible compromise is to make a relaxation of the second step. After finding the first proposed solution, and fixing zy_{phd}, instead of solving the $SPHS\_noYF$ problem, we can consider a Lagrangian relaxation: the ordering imposed by patients' age can be seen as a strong penalization to the objective function instead of inviolable constraints. In this way, the medical center does not consider the possibility to make overtime operations, but at least, it reschedules the order of interventions, to operate the younger the first, compatible with the working hours limit.

5 | Results

In this Chapter the computational results are described. All the approaches have been implemented in AMPL, Version 20230124 (MSVC 19.29.30147.0, 64-bit. The models have been solved with Gurobi 10.0.0 which proved to be the most effective solver in preliminary tests.
Computational tests have been run on a machine with AMD Ryzen 3 3250U Processor, 8,00 GB of RAM installed (5,88 GB usable); 64-bit Operating System. Windows 11 Home edition, 22H2 version.

We tested the approaches on a set of instances described in Section 5.1.

Then in Section 5.2 the very first results of the $SPHS$, $SPHS\_noYF$, and $SPHS\_MS$ are shown.
The results obtained by the alternative formulations and relaxations are described in Section 5.3, and the ones obtained by the heuristic approaches, in Section 5.4.
In Section 5.4.5 deeper analysis on the schedule found by the best approach discovered so far are conducted.
Finally in Section 5.5 are studied the behavior on the delays, by considering a shorter time horizon.

In Table 5.1 there is a map that shows, for each formulation, the objective function used, and the constraints present.

MODEL	OBJECTIVE FUNCTION	CONSTRAINTS
SPHS	(2.16)	(2.1) - (2.14),(2.15)
SPHSnodoc	(2.16)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14), (2.15)
SPHSnodoc_BP	(3.6)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14), (2.15),(3.1)-(3.4)
SPHS_BP	(3.6)	(2.1) - (2.14),(2.15),(3.1)-(3.4)
SPHSnodoc_BP_of_BP12	(3.9)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14), (2.15),(3.1)-(3.4)
SPHS_noYF	(2.16)	(2.1)-(2.14)
SPHS_noYFnodoc	(2.16)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14)
SPHS_noYFnodoc_BP	(3.6)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14) ,(3.1)-(3.4)
SPHS_noYF_BP	(3.6)	(2.1)-(2.14),(3.1)-(3.4)
SPHS_noYFnodoc_BP_of_BP12	(3.9)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14) ,(3.1)-(3.4)
SPHS_MS_BPnodoc	(2.20)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MS_nodoc	(2.20)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.17),(2.14),(2.17),(2.18)
SPHS_MS_BP	(3.6)	(2.1) - (2.9), (2.10) - (2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MS	(2.20)	(2.1) - (2.9), (2.10) - (2.14),(2.17),(2.18)
SPHS_MSnodoc_of_clday	(2.19)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.17),(2.14),(2.17),(2.18)
SPHS_MS_BP_of_dxnbin	(3.10)	(2.1) - (2.9), (2.10) - (2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MSnodoc_BP_of_d2xnbin	(3.10) (d^2)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.17),(2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MS_BP_of_d2xnbin	(3.10) (d^2)	(2.1) - (2.9),(2.10) - (2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MSnodoc_BP_of_d2xnbin	(3.10) (d^2)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.17),(2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MS_BP_of_d5xnbin	(3.10)(d^5)	(2.1) - (2.9), (2.10) - (2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MSnodoc_BP_of_2dxnbin	(3.10) (2^d)	(2.1),(2.2),(2.6),(2.7),(2.8),(2.9),(2.12),(2.13),(2.17),(2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MS_BP_of_2dxnbin	(3.10) (2^d)	(2.1) - (2.9), (2.10) - (2.14),(2.17),(2.18),(3.1)-(3.4)
SPHS_MS_of_clday	(2.19)	(2.1) - (2.9), (2.10) - (2.14),(2.17),(2.18)
SPHS_MS_BP_of_dxnbin	(3.10)	(2.1) - (2.9), (2.10) - (2.14),(2.17),(2.18),(3.1)-(3.4)

Table 5.1: Recap all models

5.1. Instances

In this Section we describe the instances which are partially derived from Buzzi Children's Hospital's staff suggestions and partially from realistic data.

5.1.1. Hospitals

Buzzi acts as the hub and it is open every day. Two sites, acting as spokes, are available only once a week.
The hub has got two operating theatres, while the spoke has one OT which can be used by the hub each. To all these, it has been added a fourth facility, which is the dummy hospital.
From now on we consider the hospital n. 1 as the Hub, hospitals 2 and 3 as Spokes, and 4 as the dummy one.

To be able to run the models explained hereafter, we added different parameters, so far just to have a general idea of the behaviors of the different models and their relative variations. Then all the parameters will be given by the medical center or chosen adequately.
We decided to consider a time horizon of 32 days, i.e. $|D| = 32$, and each working day is divided into 10 time slots, one for each hour. The daily working hour will be partitioned

further considering a slot of 30, 20, or 15 minutes. So let a new parameter $slots\_per\_h$ which indicates in how many partitions an hour is divided. So, for now, we have 10 working hours per day, and letting $slots\_per\_h = 1$, we have $|T|$ equal to 11 since we start from slot 0.

We consider a set of 10 surgeons. The matrix mPS_{ps} is randomly generated so as to guarantee that each surgeon s can operate on 70% of the cases. The same is also done to make the coverage matrix, considering the probability for a patient to be operated on in that hospital h, namely, it has the necessary equipment for patient p, equals 0.9.

It has been considered a bed capacity of beds equal to 5, if the facility is a hub, and equal to 3 of instead it is a spoke. While if it is the dummy hospital, we should assure to have enough beds to potentially allocate all those patients who could not be allocated in the time horizon, so we set it to 140.

5.1.2. Patients

We tested the approaches on 24 sets of patients, [4][3]. The dataset is composed of 8 instances with 80 patients, 8 instances with 120 patients, and 8 instances with 140 patients, for a total of 24 instances.

All the features of the patient are given:

- ID
- urgency coefficient (u_p)
  - Five urgency classes are defined with maximum waiting time set at 8, 30, 60, 180, and 360 days, respectively, and corresponding urgency coefficients u equal to 45, 12, 6, 2, 1. Namely, 45 is the most urgent, while 1 is the least urgent.
- numbers of days on the waiting list (wl_p)
- average surgery time [min] (τ_p)

Since the duration of the surgical intervention is given in minutes, we have to transform τ_p, considering as unit of measurement the $slots\_per\_h$, and so simply re-scaling it, dividing the original τ_p by 60(minutes in 1h) and multiply it by $slots\_per\_h$, obtaining the new correct value of τ_p.

The age of each patient [month] has been generated by a Gaussian distribution with a mean $\mu = 67$, and standard deviation $sd = 32$. But since the hospital operates on children older than 3 months, it is chosen the maximum between 3 and the age randomly

generated.
The number of doctors required for the surgery is set to 1 for each operation.
PACU (Post Anesthesia Care Unit) was considered as the ceiling of the operation time τ_p divided by 3.
The due date dd_p was obtained by the value of the urgency factor u_p, by taking the integer part of $\frac{60}{u_p}$.
The mPH_{ph} has been generated in this way: if the patient p is younger than the threshold mha, it must be operated on in a hub site (in our case $h = 1$) and so $mPH_{p1} = 1$. Otherwise, if $age_p > mha$, the value of mPH_{ph} is randomly generated, following a Bernoulli distribution with a probability equal to 0.9.

To all those, it has been chosen to give a high value of β, setting it to 500, to have a big impact on the objective functions where it is present.

To have a general recap of all the parameters we have set, we report in Table 5.2 the list of the assigned values to Sets and Parameters.

5.2. Results

A general overview of the very first results is presented in Table 5.3. Each row is composed, in this order, by: the name of the run instance, the computational time, the objective function found, the lower bound of the solution, and the relmipgap.
This latter one indicates the gap between the value of the objective function found so far and the lower bound, it is calculated as $relmipgap = \frac{objfun - lowerbound}{objfun}$.

Then in Table 5.7 are reported the results regarding the name of the run instance, the mean and the maximum waiting times of all the patients in the schedule found as a solution, the number of patients scheduled to be operated on before the due date, and the number who did not, and the mean and the maximum days of delay (considered only on the number of people scheduled on delay).

It has compiled all the models written in their original formulation, with all the variables and constraints proposed in Chapter 2; the number present in the name of the id in Tables 5.3 and 5.7 represents how many patients compose the instance.

There's an important element to pay attention to. In different instances the solver did not find any feasible solution even after 6000 seconds, then it could not give all the pieces of information, (it happened for one instance with 120 patients and for four instances

Set or Parameter	Value
Hr	$[1, 2, 3]$
H	$[Hr, 4]$
$OH[1]$	$[1, 2]$
$OH[2]$	$[3]$
$OH[3]$	$[4]$
$OH[4]$	$[5]$
$\|S\|$	10
$\|D\|$	32
$\|T\|$	11
mPS_{ps}	$Bernoulli(p = 0.7)$
mPH_{ph}	1 if $h = 4$ or ($age_p \leq mha$ and $h = 1$)
	$Bernoulli(p = 0.9)$ elsewhere
τ_p	$\frac{old\_\tau_p}{60} \cdot slots\_per\_h$
$numdoc_p$	1
age_p	$\max(Normal(\mu = 67, sd = 32), 3)$
$PACU_p$	$\lceil \frac{\tau_p}{3} \rceil$
dd_p	$\lfloor \frac{60}{u_p} \rfloor$
a_{hd}	1 if h=1 , everyday
	1 if $h = 2$, only the 1^{st} day of every week
	1 if $h = 3$, only the 3^{rd} day of every week
	1 if $h = 4$, only day 32
	0 otherwise
α_h	$[100, 1, 1, 1000]$
β	500
γ	100
b_h	$[5, 3, 3, 140]$
mha	60

Table 5.2: Assigned Values to Sets and Parameters

with 140 in $SPHS\_MS$). So, whilst analyzing the results, particular attention needs to be paid to take this factor into consideration, especially when we describe the general overview giving the average values. The difficulty, sometimes impossibility, to find any feasible solution in some instances, especially if there are a big number of patients to set up, already provides a glimpse of how difficult it is for the solver to find a solution with an exact resolution method, and therefore how necessary it is to try heuristic methods to solve the problem.

To give a directed reading key, it has been collected in Tables 5.4-5.6 a synthesized recap of Table 5.3, grouping all the results by the type of model utilized, and the number of patients of the instance.

Then all these Tables are followed by images that recap visually some of the previous results and other analyses, for all models which found a solution.

5.2.1. Computational Results

id1	computational_time	objfun	lowerbound	relmipgap
SPHS				
80_t_01:	6676.102	7400	7400	0
80_t_02:	1992.183	7052	7052	0
80_t_03:	5380.21	8082	8080	0
80_t_04:	5848.041	7773	7773	0
80_t_05:	1533.057	6161	6161	0
80_t_06:	4212.849	8193	8176	0.2
80_t_07:	2811.724	8493	8493	0
80_t_08:	5521.619	8094	8094	0
120_t_01:	3952.732	14410	13946	3.2
120_t_02:	3854.296	20552	12669	38.3
120_t_03:	3935.627	30989	14440	53.4
120_t_04:	8492.44	13603	13268	2.4
120_t_05:	3854.372	19877	10990	44.7
120_t_06:	3981.867	25314	14609	42.2
120_t_07:	4227.534	32917	15788	52
120_t_08:	>6000			
140_t_01:	>6000			
140_t_02:	4309.344	46809	15505	66.8
140_t_03:	4088.848	44521	17999	59.5
140_t_04:	4268.575	17862	16383	8.2
140_t_05:	4028.631	34767	13438	61.3
140_t_06:	>6000			
140_t_07:	>6000			
140_t_08:	>6000			
SPHS_noYF				
80_t_01:	3674.263	7400	7400	0
80_t_02:	616.5198	7052	7052	0
80_t_03:	1456.896	8082	8082	0
80_t_04:	534.9277	7773	7773	0
80_t_05:	258.8708	6161	6161	0
80_t_06:	757.6649	8178	8178	0
80_t_07:	826.4519	8493	8493	0
80_t_08:	1026.172	8094	8094	0
120_t_01:	6522.347	14021	14020	0
120_t_02:	3772.925	12875	12828	0
120_t_03:	5644.817	14612	14611	0
120_t_04:	3942.537	13579	13579	0
120_t_05:	1688.104	11112	11111	0
120_t_06:	3684.494	18249	14713	19.3
120_t_07:	3682.615	18939	15906	16
120_t_08:	>6000			
140_t_01:	>6000			
140_t_02:	3689.717	17453	15853	9.1
140_t_03:	3689.627	30102	18137	39.7
140_t_04:	1833.705	16719	16719	0
140_t_05:	1521.686	13593	13593	0
140_t_06:	>6000			
140_t_07:	>6000			
140_t_08:	>6000			
SPHS_MS				
80_t_01:	8756.676	9347	347	3.7
80_t_02:	5794.424	8978	588	6.5
80_t_03:	3671.905	11199	1682	15
80_t_04:	3673.414	10126	1102	11
80_t_05:	7468.536	7703	406	5.2
80_t_06:	6743.682	10168	493	4.8
80_t_07:	5961.324	10494	277	2.6
80_t_08:	8989.888	10069	359	3.5
120_t_01:	5249.529	16785	546	3.2
120_t_02:	3683.105	19436	4550	23.4
120_t_03:	3685.52	23916	7139	29.8
120_t_04:	3689.848	22167	6779	30.5
120_t_05:	3689.086	18384	5690	30.9
120_t_06:	3686.364	21689	4677	21.5
120_t_07:	3684.27	25645	7318	28.5
120_t_08:	>6000			
140_t_01:	>6000			
140_t_02:	3691.504	26095	7966	30.5
140_t_03:	3688.18	34246	13694	39.9
140_t_04:	3695.138	24526	5617	22.9
140_t_05:	4729.969	25722	10314	40
140_t_06:	>6000			
140_t_07:	>6000			
140_t_08:	>6000			

Table 5.3: Computational Results of the three formulations on the 24 instances*

*whenever some of them are not present is because any solution was found in at least 6000 seconds

# of patients	average solve time [sec]	average percentage gap	max percentage gap	# of instances without a feasible solution within 6000 sec	# of instances solved to optimality	average CPU time to reach the optimum
80	4247	0	0.2	0	6	5003
120	4542	35.2	53.4	1	0	
140	4249	52.2	66.8	4	0	

Table 5.4: $SPHS$: Computational results

# of patients	average solve time [sec]	average percentage gap	max percentage gap	# of instances without a feasible solution within 6000 sec	# of instances solved to optimality	average CPU time to reach the optimum
80	1144	0	0	0	8	1143.38
120	4134	11.9	19	1	1	3942
140	3604	12.2	39	4	2	3689

Table 5.5: $SPHS\_noYF$: Computational results

# of patients	average solve time [sec]	average percentage gap	max percentage gap	# of instances without a feasible solution within 6000 sec	# of instances solved to optimality
80	6370	5.1	15	0	0
120	5542	25.1	31	1	0
140	4549	36.2	40	4	0

Table 5.6: $SPHS\_MS$: Computational results

The correlation between the number of patients and the effectiveness of the solution becomes increasingly evident. As the patient count rises, the quality of the solution obtained diminishes. Even with 80 patients, the solution achieved is close to optimal; however, it demands a significant amount of time to reach that point. Moreover, whenever the time horizon for scheduling is extended or the time slots are expanded, the number of parameters escalates significantly, thereby exerting a substantial influence on the overall computational time. The computational challenge of the problem intensifies as the number of patients and parameters grow, making it increasingly challenging to find an efficient solution within a reasonable time frame.

5.2.2. Analytical Results

Now let us analyze some of the most relevant results about the scheduling, including the waiting times of patients, the number of patients scheduled after their due date, and the time of delays.

Figure 5.1 represents the average of the mean and maximum total waiting time (days in the schedule plus the number of days on the waiting list). The mean value of the total waiting time is more or less the same for the 3 models, but we could not say the same for the maximum. The third model has got a lower value. It means as one could expect, looking at their objective functions, that since we do not prefer spokes, which are not always available, we could operate earlier, reducing some max delays. But the explanation for the fact that for the mean waiting time, there is no evident difference between the models, while in the max waiting time yes, is that only a few patients were "sacrificed" by delaying their operation, in order to have a better overall schedule
Figure 5.2 shows the average number of patients scheduled before and respectively after the due date dd_p. Since in $SPHS\_MS$ (m3), there is no indication regarding the delays, and we want to minimize the total makespan, patients are more embedded, looking essentially at the feasibility more than the possible delay.
Then Figure 5.3 represents the average waiting time, in terms of the number of days patients on delay are scheduled after the due date, reflecting the same considerations done above.

id1	mean_wt	max_wt	on_time	on_delay	mean_delay	max_delay
SPHS						
80_t_01:	69.775	364	73	7	3.285667	7
80_t_02:	69.0375	373	77	3	4.999833	5
80_t_03:	69.6125	366	76	4	4.999875	5
80_t_04:	69.0625	373	76	4	4.999875	5
80_t_05:	68.0625	364	78	2	4.99975	5
80_t_06:	70.0375	378	74	6	4.666589	7
80_t_07:	70.6375	378	69	11	3.363606	7
80_t_08:	70.4	366	71	9	3.444406	7
120_t_01:	79.40833	380	85	35	4.057131	17
120_t_02:	81.29167	378	81	39	10.71792	26
120_t_03:	83.59167	378	60	60	11.48331	26
120_t_04:	78.29167	380	111	9	5.666604	7
120_t_05:	81.34167	371	79	41	11.85363	24
120_t_06:	82.55	380	66	54	10.25924	27
120_t_07:	83.28333	369	58	62	10.64514	27
120_t_08:						
140_t_01:						
140_t_02:	90.1	376	50	90	12.63332	30
140_t_03:	90.25714	376	55	85	13.05881	31
140_t_04:	83.80714	380	116	24	5.958309	17
140_t_05:	87.27857	376	83	57	12.91226	31
140_t_06:						
140_t_07:						
140_t_08:						
SPHS_noYF						
80_t_01:	69.8	364	73	7	3.285667	7
80_t_02:	69.0375	371	77	3	4.999833	5
80_t_03:	69.6125	366	76	4	4.999875	5
80_t_04:	69.0625	371	76	4	4.999875	5
80_t_05:	68.275	364	77	3	4.999833	5
80_t_06:	69.7625	371	76	4	4.999875	5
80_t_07:	70.6	378	69	11	3.363606	7
80_t_08:	70.375	366	71	9	3.444406	7
120_t_01:	78.9	371	88	32	3.656239	8
120_t_02:	78.09167	378	103	17	3.235275	7
120_t_03:	78.84167	378	94	26	2.730759	7
120_t_04:	78.275	378	111	9	5.444384	7
120_t_05:	76.49167	373	116	4	4.999875	5
120_t_06:	80.89167	380	79	41	6.682911	26
120_t_07:	81.375	378	73	47	6.212753	22
120_t_08:						
140_t_01:						
140_t_02:	83.63571	380	102	38	4.052621	24
140_t_03:	87.15	380	73	67	8.895509	27
140_t_04:	83.18571	380	116	24	2.333324	5
140_t_05:	81.80714	375	136	4	4.999875	5
140_t_06:						
140_t_07:						
140_t_08:						
SPHS_MS						
80_t_01:	68.85	364	74	6	1.999967	3
80_t_02:	68.1	366	79	1	4.9995	5
80_t_03:	69.3375	366	73	7	5.285639	7
80_t_04:	68.6	366	77	3	3.999867	6
80_t_05:	66.6125	357	80	0	0	0
80_t_06:	68.925	364	77	3	1.333289	2
80_t_07:	69.375	364	72	8	2.124973	4
80_t_08:	69.325	357	73	7	2.428537	5
120_t_01:	78.4	373	87	33	3.757564	10
120_t_02:	82.59167	372	88	32	5.499983	22
120_t_03:	80.68333	364	75	45	7.088873	20
120_t_04:	79.96667	371	84	36	7.527757	24
120_t_05:	78.075	371	96	24	7.708301	22
120_t_06:	80.53333	373	73	47	6.319135	17
120_t_07:	81.68333	371	66	54	8.42591	27
120_t_08:						
140_t_01:						
140_t_02:	85.08571	377	93	47	7.978706	27
140_t_03:	87.26429	370	73	67	9.283568	31
140_t_04:	84.90714	371	99	41	8.390223	21
140_t_05:	84.07857	368	102	38	9.289449	27
140_t_06:						
140_t_07:						
140_t_08:						

Table 5.7: Analytical Results of the three formulations on the 24 instances*

*whenever some of them are not present is because any solution was found in at least 6000 seconds

# of patients	average waiting time (day)	average max w.t. (days)	average # of patients on time	average # of patients on delay	average delay (days)	average max delay
80	69.58	370	74.25	5.75	4.34	6
120	81.57	375	75.12	44.87	9.39	22.5
140	88.93	377	67.37	72.62	11.91	28.5

Table 5.8: $SPHS$: Results about analysis

# of patients	average waiting time (day)	average max w.t. (days)	average # of patients on time	average # of patients on delay	average delay (days)	average max delay
80	61.12	368	74.37	5.63	4.35	5.75
120	78.85	377	94.86	25.14	5.01	11.71
140	83.7	378	106.75	33.25	11.91	15.25

Table 5.9: $SPHS\_noYF$: Results about analysis

# of patients	average waiting time (day)	average max w.t. (days)	average # of patients on time	average # of patients on delay	average delay (days)	average max delay
80	68.64	363	75.63	4.37	2.77	4
120	80.28	371	81.29	38.71	6.62	20.3
140	85.33	371.5	91.75	48.25	8.74	26.5

Table 5.10: $SPHS\_MS$: Results about analysis

Figure 5.1: Bar plot waiting time

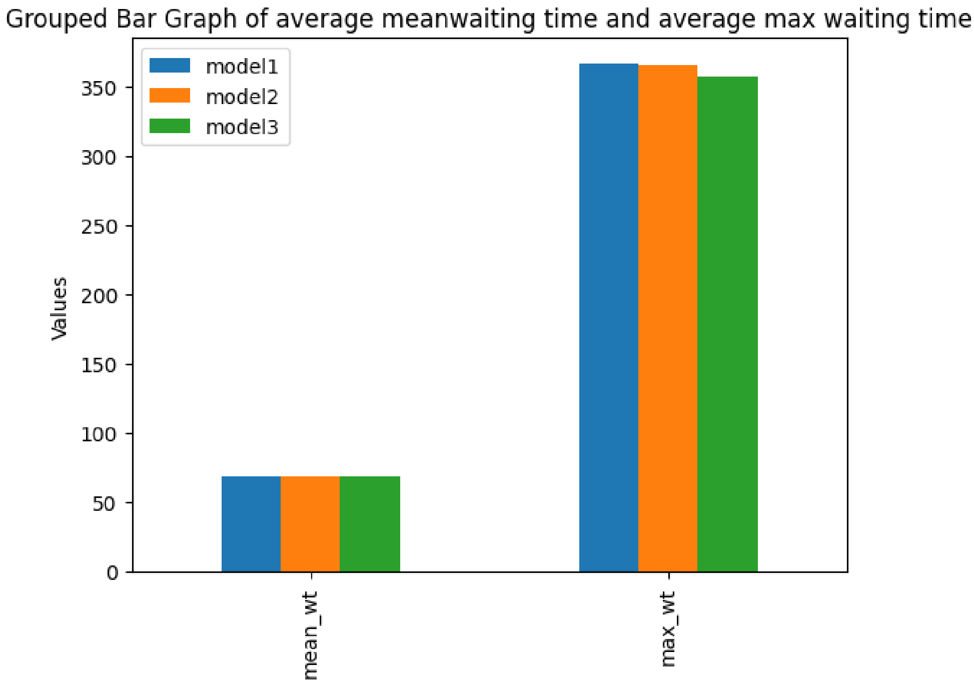


Figure 5.2: Bar plot operation before due date

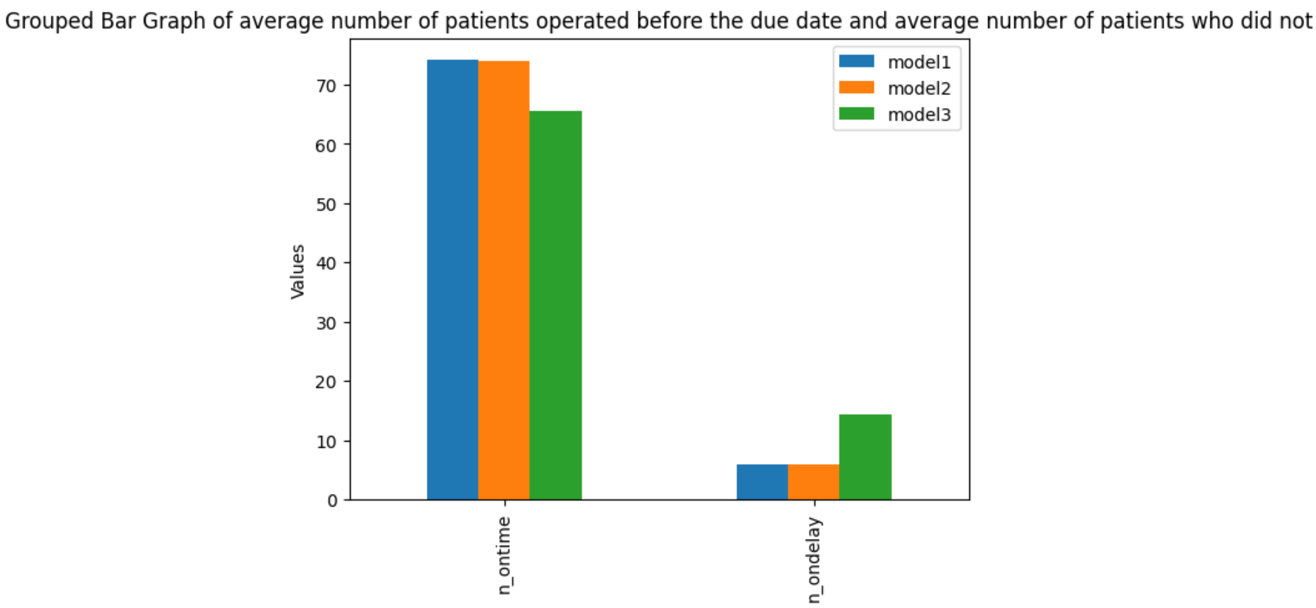


Figure 5.3: Bar plot delay

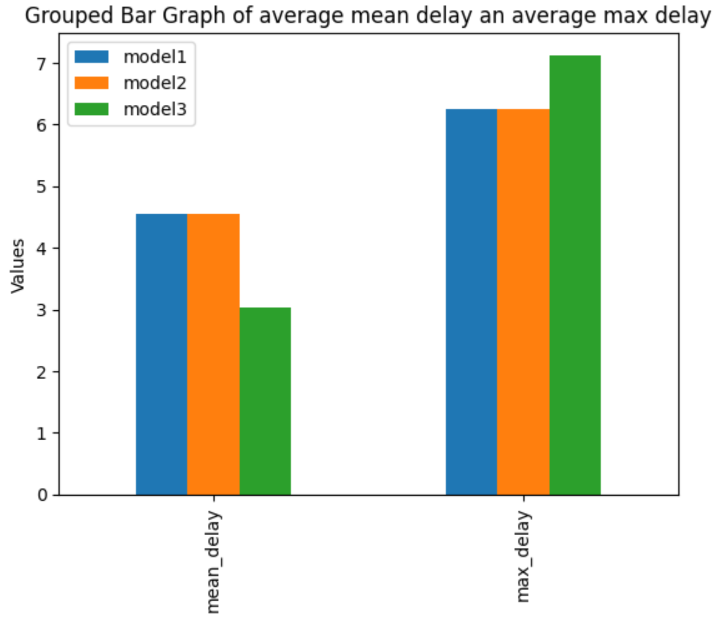
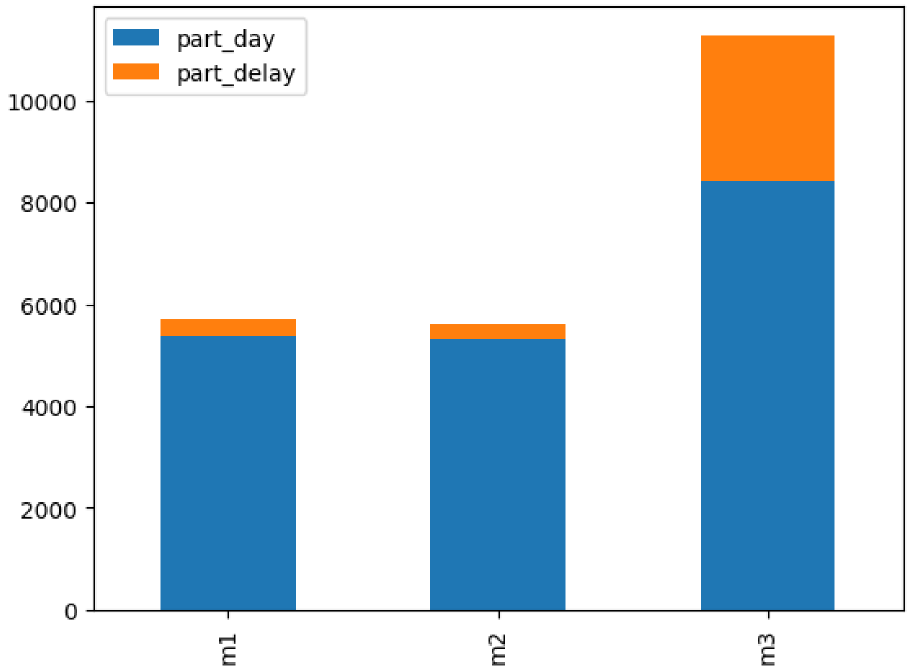


Figure 5.4: Objective function's (2.16) weights



Finally, Figure 5.4 represents the operation days and the delays weighted with urgency. It is actually the first and the second part of the weighted sum of the objective function (2.16) of $SPHS$ and $SPHS\_noYF$ models. From this last image, we can notice two key

points.
The first one is that, by comparing $SPHS$ (m1) and $SPHS\_noYF$ (m2) we hardly see an improvement of the second model, which does not have the constraints of respecting the interventions order on a working day. So this problem despite does not have these extra constraints, and has a bigger feasible region to find a better schedule, does not seem to reach a significant improvement, except for the $closure_day$, but which is not actually a goal for the first two problems.
The second noteworthy observation is between the first two models and the third $SPHS\_MS$ (m3). In fact, we can notice that operating all patients to finish them as soon as possible does not mean operating them well. Despite in the $SPHS\_MS$ model we actually have fewer days in the total scheduling, what gives a higher weight in the objective function is the day of the operation multiplied by the urgency u_p, making the objective function

5.3. Results for Variants

From now on, to show the behavior of the different formulations, we run and report only the result of the instance 140_t_01, which counts 140 patients, as a representative instance.
For every run, a time limit of 1000 seconds has been set, and a memory limit of 512 MB.

5.3.1. Bin Packing Enrichment

In Tables 5.11-5.16 are presented the computational and analytical results of the run model.

$SPHS$:

MODEL	OBJ FUN	NOTES	SOLVE TIME	RELMIPGAP[%]	CLOSURE DAY
$SPHS$	objfun2	run out of mem	660	NO FOUND	NO FOUND
$SPHS$_BP	objfunBinPacking		1007	92	32

Table 5.11: BP Variations of $SPHS$ - Computational results

MODEL	OBJ FUN	mean_wt	max_wt	delays/$\|P\|$	mean_dlt	max_dlt
$SPHS$	objfun2	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS$_BP	objfunBinPacking	91.97	379	0.3	4.38	30

Table 5.12: BP Variations of $SPHS$ - Analysis results

$SPHS\_noYF$:

MODEL	OBJ FUN	SOLVE TIME	RELMIPGAP[%]	CLOSURE DAY
$SPHS\_noYF$	objfun2	891	0.41	31
$SPHS_noYF$_BP	objfunBinPacking	654	0	31

Table 5.13: BP Variations of $SPHS\_noYF$ - Computational results

MODEL	OBJ FUN	mean_wt	max_wt	delays/$\|P\|$	mean_dlt	max_dlt
$SPHS\_noYF$	objfun2	83.71	377	0.29	4.51	10
$SPHS_noYF$_BP	objfunBinPacking	84.22	377	0.31	4.40	11

Table 5.14: BP Variations of $SPHS\_noYF$ - Analysis results

$SPHS\_MS$:

MODEL	OBJ FUN	SOLVE TIME	RELMIPGAP[%]	CLOSURE DAY
$SPHS\_MS$	objfun3		NO FOUND	NO FOUND
$SPHS_MS$_BP	objfunBinPacking	526	NO FOUND	NO FOUND
$SPHS_MS$_BP_of_dxnbin	objfunBinPacking2	676	78	32
$SPHS_MS$_BP_of_d2xnbin	objfunBinPacking2(d^2)	743	77	29
$SPHS_MS$_BP_of_d5xnbin	objfunBinPacking2(d^5)	694	NO FOUND	NO FOUND
$SPHS_MS$_BP_of_2dxnbin	objfunBinPacking2 (2^d)	1003	99.9	31
$SPHS_MS$_of_clday	objfunclday	644	NO FOUND	NO FOUND
$SPHS_MS$_BP_of_dxnbin	objfunBinPacking2		NO FOUND	NO FOUND

Table 5.15: BP Variations of $SPHS\_MS$ - Computational results

MODEL	OBJ FUN	mean$_{w}t$	max_wt	delays/$\|P\|$	mean_dlt	max_dlt
$SPHS\_MS$	objfun3	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS_MS$_BP	objfunBinPacking	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS_MS$_BP_of_dxnbin	objfunBinPacking2	86.17	361	0.45	11.50	30
$SPHS_MS$_BP_of_d2xnbin	objfunBinPacking2(d^2)	85.15	363	0.53	7.67	24
$SPHS_MS$_BP_of_d5xnbin	objfunBinPacking2(d^5)	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS_MS$_BP_of_2dxnbin	objfunBinPacking2 (2^d)	85.15	367	0.50	7.90	21
$SPHS_MS$_of_clday	objfunclday	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS_MS$_BP_of_dxnbin	objfunBinPacking2	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND

Table 5.16: BP Variations of $SPHS\_MS$ - Analysis results

Differences with a Bin-Packing formulation

Upon proposing this formulation, we initially lacked knowledge of the magnitude of its impact. However, through meticulous analysis, it becomes evident that in the case of both $SPHS$ and $SPHS\_noYF$, in terms of velocity, this formulation brings improvements. First of all in some cases, like $SPHS$, $SPHS\_MS$ and many of its variants.

If we want to compare analytical differences, keeping in mind they are different problems, improvements in terms of solution quality are negligible and sometimes nonexistent.

Similar observations can be made regarding the third model, $SPHS\_MS$, and its respective variations. However, it is important to delve further into this matter. The sensitivity of the solution found is heavily influenced by the presence or absence of variables associated with doctors, depending on the chosen objective function. Their inclusion or exclusion can significantly impact the final solution's quality.

It has been made different variations of this last formulation since a Bin Packing approach was better suited for a problem whose aim is to minimize the make-span. As we discussed in previous chapters, minimizing the total number of days used to complete all the surgical cases, and using as few operating theatres as possible to reach the same goal, are two sides of the same coin, so we decided to explore extensively this approach for this model.

5.3.2. Relaxed Formulations

Keeping in mind that model $SPHS\_noYF$ is for all intents and purposes a relaxation of $SPHS$, in this Subsection we analyze differences of the relaxation of the models given by removing the surgeons from the formulations.

In Tables 5.17-5.22 are presented the computational and analytical results of the run model.

$SPHS$:

MODEL	OBJ FUN	NOTES	SOLVE TIME	RELMIPGAP[%]	CLOSURE DAY
$SPHS$	objfun2	run out of mem	660	NO FOUND	NO FOUND
$SPHS$nodoc	objfun2		814	3.8	31
$SPHS$nodoc_BP	objfunBinPacking		506	0.2	24
$SPHS$_BP	objfunBinPacking		1007	92	32
$SPHS$nodoc_BP_of_BP12	objfunBinPacking_m12		501	1.3	31

Table 5.17: _$nodoc$ Variations of $SPHS$ - Computational results

MODEL	OBJ FUN	mean_wt	max_wt	delays/$\|P\|$	mean_dlt	max_dlt
$SPHS$	objfun2	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS$nodoc	objfun2	83.93	377	0.3	4.57	10
$SPHS$nodoc_BP	objfunBinPacking	82.92	370	0.28	4.57	11
$SPHS$_BP	objfunBinPacking	91.97	379	0.3	4.38	30
$SPHS$nodoc_BP_of_BP12	objfunBinPacking_m12	83.77	377	0.30	4.34	10

Table 5.18: _$nodoc$ Variations of $SPHS$ - Analysis results

$SPHS\_noYF$:

MODEL	OBJ FUN	SOLVE TIME	RELMIPGAP[%]	CLOSURE DAY
$SPHS\_noYF$	objfun2	2057.91	0.4	31
$SPHS_noYF$nodoc	objfun2	65	optimal	31
$SPHS_noYF$nodoc_BP	objfunBinPacking	52	optimal	24
$SPHS_noYF$_BP	objfunBinPacking	1827.02	4.2	24
$SPHS_noYF$nodoc_BP_of_BP12	objfunBinPacking_m12	56	optimal	31

Table 5.19: $\_nodoc$ Variations of $SPHS\_noYF$ - Computational results

MODEL	OBJ FUN	mean_wt	max_wt	delays/$\|P\|$	mean_dlt	max_dlt
$SPHS\_noYF$	objfun2	83.71	377	0.29	4.46	10
$SPHS_noYF$nodoc	objfun2	83.62	377	0.29	4.46	10
$SPHS_noYF$nodoc_BP	objfunBinPacking	82.95	370	0.28	4.34	11
$SPHS_noYF$_BP	objfunBinPacking	82.95	370	0.28	4.34	11
$SPHS_noYF$nodoc_BP_of_BP12	objfunBinPacking_m12	83.62	377	0.29	4.46	10

Table 5.20: $\_nodoc$ Variations of $SPHS\_noYF$ - Analysis results

$SPHS\_MS$:

MODEL	OBJ FUN	SOLVE TIME	RELMIPGAP[%]	CLOSURE DAY
$SPHS_MS$_BP_nodoc	objfun3	1077	3.3	22
$SPHS_MS$_nodoc	objfun3	1068	2.3	22
$SPHS_MS$_BP	objfunBinPacking	526	NO FOUND	NO FOUND
$SPHS\_MS$	objfun3		NO FOUND	NO FOUND
$SPHS_MS$nodoc_of_of3	objfun3	1133	2.3	22
$SPHS_MS$nodoc_of_clday	objfunclday	26	optimal	21
$SPHS_MS$_BP_of_dxnbin	objfunBinPacking2	676	78	32
$SPHS_MS$nodoc_BP_of_d2xnbin	objfunBinPacking2(d^2)	39	optimal	21
$SPHS_MS$_BP_of_d2xnbin	objfunBinPacking2(d^2)	743	77	29
$SPHS_MS$nodoc_BP_of_d2xnbin	objfunBinPacking2(d^5)	29	optimal	21
$SPHS_MS$_BP_of_d5xnbin	objfunBinPacking2(d^5)	694	NO FOUND	NO FOUND
$SPHS_MS$nodoc_BP_of_2dxnbin	objfunBinPacking2 (2^d)	10	optimal	21
$SPHS_MS$_BP_of_2dxnbin	objfunBinPacking2 (2^d)	1003	99.9	31
$SPHS_MS$_of_clday	objfunclday	644	NO FOUND	NO FOUND
$SPHS_MS$_BP_of_dxnbin	objfunBinPacking2		NO FOUND	NO FOUND

Table 5.21: $\_nodoc$ Variations of $SPHS\_MS$ - Computational results

MODEL	OBJ FUN	$mean_w t$	max_wt	delays/$\|P\|$	mean_dlt	max_dlt
$SPHS_MS$_BP_nodoc	objfun3	84.07	369	0.29	4.36	10
$SPHS_MS$_nodoc	objfun3	83.09	368	0.29	4.36	9
$SPHS_MS$_BP	objfunBinPacking	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS\_MS$	objfun3	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS_MS$nodoc_of_of3	objfun3	83.09	368	0.29	4.36	9
$SPHS_MS$nodoc_of_clday	objfunclday	84.35	368	0.47	6.81	20
$SPHS_MS$_BP_of_dxnbin	objfunBinPacking2	86.17	361	0.45	11.50	30
$SPHS_MS$nodoc_BP_of_d2xnbin	objfunBinPacking2(d^2)	84.27	364	0.492857	6.72	16
$SPHS_MS$_BP_of_d2xnbin	objfunBinPacking2(d^2)	85.15	363	0.53	7.67999	24
$SPHS_MS$nodoc_BP_of_d2xnbin	objfunBinPacking2(d^5)	84.14	368	0.49	5.99	16
$SPHS_MS$_BP_of_d5xnbin	objfunBinPacking2(d^5)	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS_MS$nodoc_BP_of_2dxnbin	objfunBinPacking2 (2^d)	84.25	365	0.49	6.56	17
$SPHS_MS$_BP_of_2dxnbin	objfunBinPacking2 (2^d)	85.15	367	0.50	7.90	21
$SPHS_MS$_of_clday	objfunclday	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND
$SPHS_MS$_BP_of_dxnbin	objfunBinPacking2	NO FOUND	NO FOUND	NO FOUND	NO FOUND	NO FOUND

Table 5.22: $\_nodoc$ Variations of $SPHS\_MS$ - Analysis results

Differences with $\_nodoc$ relaxation

The main observation to be noted concerns the substantial enhancement achieved in this relaxation. This improvement becomes apparent through the removal of variables and constraints associated with surgeons, leading to advancements across various aspects.

Regarding the optimality of the solution, we observe a noteworthy improvement in the relmipgap, particularly in certain formulations such as the $SPHS\_MS$ and its variations, where the enhancement surpasses 99 percentile points, illustrating a significant upgrade.

In terms of quality, several analytical elements put forth demonstrate a decrease in the mean and maximum waiting time, mean and maximum delay, and other relevant metrics, thereby indicating an overall improvement in quality.

Furthermore, when considering the computational time aspect, there have been enhancements in this area as well. In some formulations, like $SPHS\_MSnodoc\_of\_clday$ or $SPHS\_MSnodoc\_BP\_of\_2dxnbin$, the optimal solution is reached in just a few seconds.

The explanation for this behavior can be elucidated by considering two key factors. The first one, as previously mentioned, revolves around the reduction in the number of parameters. This reduction undoubtedly contributes to the observed improvements. However, it is important to note that this factor alone does not entirely account for the substantial disparities witnessed in terms of computational time and solution quality.

The second crucial facet that deserves consideration is the absence of any explicit indication or incorporation of the surgeons' presence in the objective functions proposed thus far. By omitting any guidance or specific directions for enhancing the allocation of doctors, it can be surmised that the solver, which employs a Branch and Bound technique in its computational method, is confronted with a more extensive and intricate search space.

In essence, the absence of explicit guidance regarding the optimization of surgeons' allocation, combined with the solver's reliance on the Branch and Bound technique, results in a significantly larger number of nodes that need to be explored. Consequently, this can substantially contribute to the observed disparities in both computational time and the quality of the obtained solutions.

Analysis on different Objective Functions

From the previous results and Tables, having an overview of the different formulations, some considerations regarding the differences between various objective functions can be now done.
Before giving a deep analysis of the results, let us recap all the differences between the various obj. functions, presented in Table 3.3.

We set, as parameters γ equal to 100 and β equal to 500.

Each distinct objective function introduces specific enhancements based on the individual requirements it aims to address. Consequently, our attention is primarily directed towards two key aspects: the relmipgap and the computational time. These metrics provide us with a comprehensive overview of the model's behavior, allowing us to gauge its performance more broadly, rather than solely relying on its overall outcomes.

The relmipgap serves as an essential measure of the optimality of the solution obtained. By examining this metric, we can assess the model's ability to approximate the optimal solution, thereby providing valuable insights into its effectiveness and efficiency.

Simultaneously, the computational time offers crucial insights into the computational efficiency of the model. This metric reflects the duration required by the solver to explore the search space, evaluate potential solutions, and ultimately converge on an optimized schedule. By monitoring the computational time, we can evaluate the model's computational performance, allowing us to make informed decisions regarding its feasibility and practicality.

By focusing on these specific metrics, we gain a comprehensive understanding of the model's behavior and suggestions on the convergence speed, enabling us to assess its efficacy in addressing the stated objectives and fulfilling the desired requirements.

$objfunBinPacking$ seems to give an important sprint for the optimal solution. Given a high value of β parameter, the solver first decided which is the minimum number of OTs required to accommodate all the surgical cases and then tries to optimize the order of the scheduling, by prioritizing the more urgent patients.

For what concerns the models aimed to find the minimum make-span, it has been proposed two different ways to reach this goal. The first was simply to minimize the $closure\_day$, ($objfuncIday$). The second approach was to minimize the number of OTs used. Since itself alone does not give any indications on the completion day, the opening of the bin, i.e. the variable yBP_{od}, is multiplied by its day. So objective function as $objfunBinPacking2$, $objfunBinPacking2*$, and $objfunBinPacking2$ were added. We also wanted to see if, instead of multiplying yBP_{od} by its day, we could have possible improvements by using a function of the index d.
We tested essentially three different functions: two polynomials of order 2 and order 5, and an exponential function. We spot controversial results. If we consider formulations with the presence of doctors, the more we increase the growth of the function, the worse is the solution found within a certain time limit, whenever it is got. While the discussion and the results are quite the opposite if instead, we do not consider also the surgeons' schedule. We, in fact, obtain, at an increasing growth of the function adopted, a decreasing compiling time, to reach the optimum.

5.4. Heuristic and Matheuristic Methods

Now we check the goodness of the heuristic approaches by comparing them with the optimal solutions of the problems, by letting them run, someone even for several hours.

5.4.1. Greedy

For 3 instances with respectively 80, 120, and 140 patients, it has been run a greedy algorithm. The numerical values of the solution with this approach were found in just a few seconds and they are reported in Table 5.23.

# of patients	waiting time (day)	max w.t. (days)	# of patients on delay	delay (days)	max delay	objective function 2.20	closure day
80	69.94	364	4	3.25	8	9971	13
120	79.15	368	32	6.62	12	17867	19
140	89.86	369.5	47	8.08	16	19777	23

Table 5.23: Results about greedy

# of patients	waiting time (day)	max w.t. (days)	# of patients on delay	delay (days)	max delay	objective function 2.20	closure day
80	68.05	360	8	1.99	3	8973	12
120	78.15	364	46	3.62	10	16113	18
140	87.86	368.5	67	4.08	15	18749	21

Table 5.24: Results about the optimal solutions of $SPHS\_MS$

We could compare it with the optimal solution of $SPHS\_MS$ model, presented in Table 5.24.
But, we would look at some considerations about its effects rather than numerical values. The outcomes, as we could have imagined, are not so good, in terms of a wide allocation of patients. However, in a few seconds, we have a feasible solution. So one could think of starting the analytical resolution method from this greedy solution. From this initial point, the solver can only improve the objective function, and in some formulations, a good starting point could be very decisive for the convergence to the optimum.

Regardless, the schedule found by the greedy algorithm, with that specific ordering, is not so appropriate as a starting point for the resolution. Actually, the relevant pieces of information are given by the number of days used to complete the waiting list, as in the $SPHS\_MS$ goal, or the number, and the kind, of OTs used to allocate all the surgical cases, for the first two formulations $SPHS$ and $SPHS\_noYF$, where is given big importance to spokes' OTs utilization.
This kind of information which has been found in a greedy manner can be used as a starting point, or even as constraints to respect. In this latter way, we can cut a significant space in the feasible region, reducing the number of nodes to explore. On the other hand, giving so much strength to only one aspect of the objective function 2.16 for $SPHS$

and $SPHS\_noYF$, we could eliminate some solutions that for example allocate patients on more OTs but reduces the overall delay for most urgent cases, making the value of the $objfun$ lower than all the possible schedules that must satisfy the constraints of the number of ORs used.

5.4.2. Force the solution of $SPHS\_noYF$ with the YF constraints

Comparing Table 5.4 with Table 5.5, and Table 5.8 with Table 5.9, it is evident that the restriction of the day ordering compromises not only the quality of the solution, as we could have imagined since we have a wider feasible area, but also on the computational results.
This latter result is the opposite of our expectations. In fact, by removing constraints, the feasible region would be expanded, which means the search space may become larger, making it more challenging and time-consuming to find an optimal solution. But for roughly the same amount of time, the relmipgap is always better for $SPHS\_noYF$ than $SPHS$, despite the potentially bigger number of nodes to explore.

A possible explanation for this phenomenon could be the fact that in $SPHS\_noYF$ surgical operations are more interchangeable, and the solver could have a faster convergence speed.

However, we would like to take advantage of these improvements with this relaxation, understanding the impact on the quality of the solution. As we discussed in Section 4.4 this is not a fundamental constraint, but by guaranteeing it, the clinic assures benefits and welfare to the patients, especially to the younger babies, reducing the risks of occurring in delays, and leaving them in an uncomfortable environment.
So the possible questions are: what if the hospital does not respect this constraint? Do the solution's improvements bring significant medical benefits, paying them with patients' inconveniences?

Now suppose that the clinic does not want to lose prestige, or strongly wants to respect this order set by the patient's age. How does it cost, to maintain the solution found by $SPHS\_noYF$, which is better than the one found by $SPHS$.

	SAME DAY	EARLIER THAN MODEL1	AFTER THAN MODEL 1
COUNT OPERATING DAY	61	43	36
MEAN OPERATING DAY	0	-3.09	3.02
WEIGTHED MEAN OP. DAY WITH u_p	0	-18.46	19.5
COUNT DELAY	101	21	18
MEAN DELAY	0	-18.46	19.5
WEIGTHED MEAN DELAY WITH u_p	0	-20	20.66

Table 5.25: Differences of the results of $SPHS\_noYF$ with $youngerfirst$ constraints added after and the ones of $SPHS$

To compare the eventual advantages we have run $SPHS\_noYF$ and $SPHS\_noYF$ for 1000 seconds each. For $SPHS\_noYF$ we saved the solution, add the constraints that impose the operation day and the hospital for each patient, namely save variables zy_{phd}, and then run $SPHS$ with these new variables restriction, with a larger set of $|T|$.
Then to compare the results of the two models, we simply compute the difference of values of the $SPHS\_noYF$ with the addition of $youngerfirst$ constraints, with the original $SPHS$, and some notable values have been reported in Table 5.25. Whenever we find positive values of the averages, it means that the values from the result of $SPHS\_noYF$ with the $youngerfirst$ constraints are bigger than the ones of $SPHS$ results. As we can notice there are no relevant improvements in terms quality of the solution, especially if we have a high cost to pay. In fact, we have a total of 21 OTs where OTs go on overtime, for a total of 54 hours accumulated as extra time.

5.4.3. Decomposing

Looking at the results in Table 5.7, and in all the Tables 5.4-5.10, the number of surgical cases of the instance strongly influences the quality of the solution. So in the same way we have observed significant improvements by reducing the number of variables related to doctors, we could imagine, and prove by the obtained outcomes, that we have improvements by reducing the number of patients, and their related variables.

To look at some examples, the improvement given by a smaller dataset is much more evident in the first model.
In Table 5.4 we can see that with more or less the same computational time, the relmipgap decreases from a value of 0.5226 when 140 patients are considered, to 0.3518 if they are 120, reaching a value of 0.0003 with 80 patients, which means that many of the solutions found with those instances were optimal, or very close to the optimum.

We used this approach essentially for $SPHS\_MS$, since it was a more suitable approach.

But we found very bad results. We put as a time limit half of the previous time limit, in order to have generally the same computational time as the previous simulations and to be able to compare adequately the results.

We obtained very bad results, most of them unfeasible. Actually, the compiler, in that amount of time (500 seconds), allocates the first slot of patients in, let us say, the first d_1 days. Then the other partition cannot be scheduled in a time horizon of $|D| - d_1$ days because maybe they are too few.

So, for this reason, we decided to give a sprint to the search for the solution, by reducing the number of available days, since it is not needed to explore all the nodes, causing a long search, a slow convergence, and an insufficient solution.

Decomposition reducing the time horizon for each partition

We decided to give manually some improvements to show how the reduction of the number of days in the time horizon could bring some advantages. We set the value of the time horizon of Part 1 equal to 16 and the time horizon of Part 2 from the day after the closure day found in Part 1 to 32 we set a time limit of 500 seconds for each part.

In Table 5.26 we can see that there is an improvement, but not so evident, especially for the second part the gain is less evident.

PART	*closure_day*	percentage gap
1	11	5.6
2	11+15	30

Table 5.26: $SPHS\_MS$: Meta-heuristic results with Decomposition with a starting feasible cl_day

Taking into account that the number of days used for Part 1 strictly depends on the instance considered we would like to find an automatic method to choose this number from the instance in case. For this reason, we decided to add a previous step, Part 0, in order to obtain an idea of the minimum number of days required. In Part 0 in fact we run the model with half number of cases just for a few minutes, in order to obtain the value of *closure_day* to use in the next step.
In Part 1 we set the time horizon from 1 to the value of *closure_day* found in the previous step and then run the model with the first half of patients.
Sequentially in Part 2 we set the time horizon from the day after the *closure_day* found

in Part 1 to the value of the original $|T|$ and run the model with the remaining patients, same as before.

First of all, we have to keep attention that this method could give some improvements to the next two steps, but it is not certain.

In fact, the minimum time required to find a value of cl_day, that could bring real improvements, is more than 900 seconds, which is a time very close to the total computational time we want to stay inside, and this proves the ineffectiveness of this method. Actually, giving the information to the solver to keep in memory the obtained best solution found so far, part 1 is only the continuation of part 0.

The inefficiency of these last 2 approaches is also given by the fact that we do not consider all the clinical cases all at once but are divided into two partitions and so have a lower overview of the scheduling process.

Let us assume this example: consider 20 patients. The first 10 ones have everybody τ_p of 7 hours and the second 10 patients have τ_p equal to 2 hours. If we do this schedule in sequence, considering only 1 OT per day, we obtain as optimal solutions, $closure\_day$ of 10 for Part 1 and a $closure\_day$ of 3 for Part 2, completing all the surgical cases in 13 days.
But if we have considered all the 20 patients all at once we would have a completion date of 10 instead of 13.

So we could assume not think this blocks division of the time horizon, but formulating this approach in this way: first, run the model with the first partition of cases, fix them, and run the second partition attaching the results to the ones found in the previous step. We could use it also with a bin packing approach, in order to use OTs already open, maximizing their utilization. But in this case, we have to pay attention to some constraints that in the second step could be never satisfied, or satisfying, the quality of the solution is very poor. Just think for example the constraints regarding the ordering of the patients, in an increasing order. Or even at the constraints of the compatibility between patient and hospital, there could be the risk that a patient that can go only to a specific facility, could find it always already full, and cannot be allocated anywhere else.

To sum up, this method does not give improvements in terms of speed of convergence, and quality of solution.

5.4.4. Surgeons added in a second step

Looking at the relevant improvements by removing the surgeons' variables and constraints it has been decided to continue in this direction.

We widely explain why the solution presented with this relaxation cannot be the answer to the entire problem as it stands, since it is not complete. So we decided to find, in a second step, the associations of the doctors with the OTs and with the patients.

This second step was referable to solving a facility location problem. $model\_SA$ has constraints and variables removed from $SPHS$ to obtain the formulation without surgeons $SPHS\_nodoc$. The objective function now is 4.1, which aims to minimize the number of allocated doctors in the planning horizon. As Table 5.27 reports, we obtain excellent results, finding the optimal solution to this second problem in just a few seconds.

MODEL	OBJECTIVE FUNCTION	O.F. VALUE	RELMIPGAP[%]	COMPUTATIONAL TIME
$SPHS\_nodoc$	$objfun2$[2.16]	18764	0.9	1000
$model\_SA$	$objfun\_mindoc$ [4.1]	56	0	3
$SPHS$	$objfun2$[2.16]	18596	0	>6000

Table 5.27: Computational results of the two sequential problems and of the optimal

This seems so far the best model that reaches all the requirements and goals, given by the clinic.

5.4.5. Analysis

Since this formulation seemed to be the best formulation among the analyzed ones, with this specific instance, it has been decided to conduct further analysis on the found schedule, obtained by running firstly $SPHS\_nodoc$, followed by the run $model\_SA$ model, for the surgeon association.

Given the nature of our specific challenge, we made a deliberate decision to exclude the Bin-Packing approach from our future considerations. This decision was made considering the trade-off between complexity and the potential benefits derived from its inclusion. Upon analysis, it became apparent that incorporating this approach would introduce a higher level of complexity to the model, involving an increase in the number of variables and constraints. However, the resulting improvements were deemed not sufficiently substantial to warrant its inclusion.

But first, we change some parameters to have a more realistic setting. With respect to some parameter values. For example, the number of working hours per day has been changed, and for the following analysis, it assumes the value of 7 hours.

Secondly, the working hours' range has been partitioned in a thicker way, making considering the duration of each time slot 30 minutes, i.e. $slots\_per\_hour = 2$. With this modification, the set of T will change, assuming $|T| = 7 \cdot 2 + 1 = 15$. For this reason, we also adjusted the values of τ_p, given by the clinic in minutes, and $PACU_p$ adequately, according to our formulation.

Before showing the schedule proposed by the result of this formulation, let us have a general overview of the dataset.

It is composed of 140 patients, each of which has one of the five urgency factors: 45 (most urgent), 12, 6, 2, 1 (less urgent).
Figure 5.5 actually represents how many patients have a specific urgency coefficient u_p, showing practically their distribution.
While Figure 5.6 reports which is the average value of the operation time τ_p for each urgency class.

Figure 5.5: Histogram of the frequencies for each urgency factor

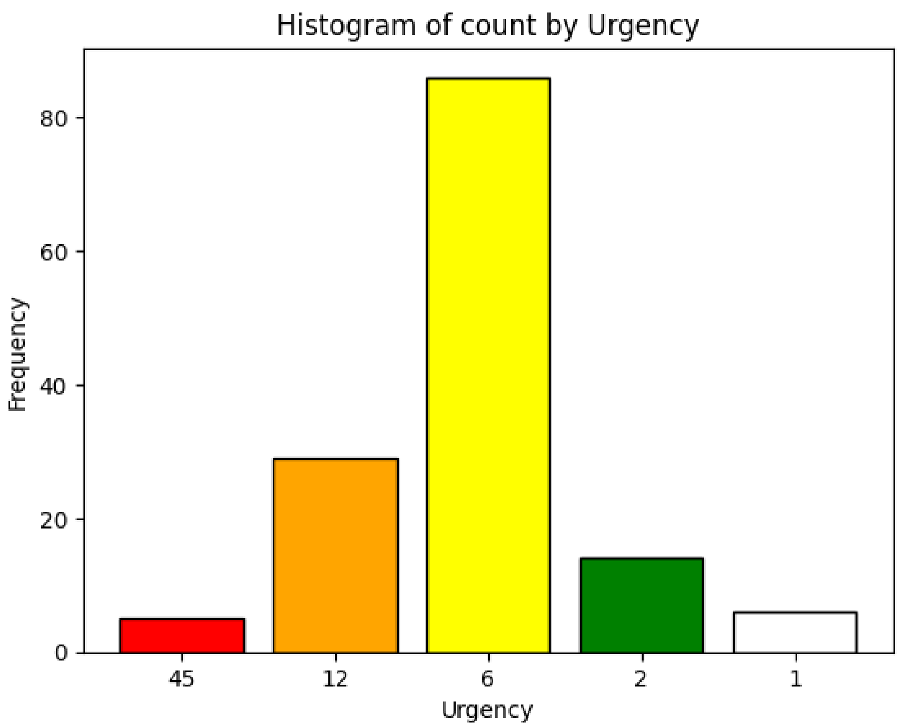
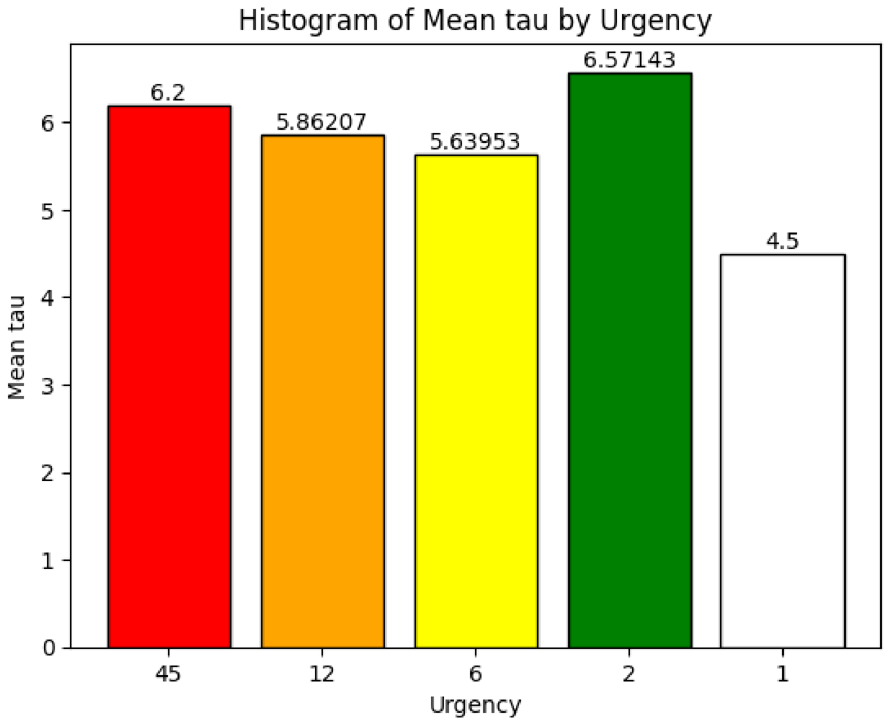


Figure 5.6: Average value of τ for each urgency factor



Schedule analysis

We have run the first model, $SPHSnodoc$, which is the most complete one, and after 1000 seconds the solver found a solution whose relmipgap is equal to 0.0092, so less than 1%, which can be considered a satisfactory solution, very close to the optimum. Subsequently, we run $model\_SA$, where the solver found the optimal schedule for the surgeons, on the basis of the patient's timetable in less than 2 seconds.

Figure 5.7 and Figure 5.8 give an overview of the different allocations of the patients.
It can be seen that, generally, in the schedule, it's preferred to operate on more patients at the begging of the time horizon to have more patients with a small operation day, and so to make it possible, patients with a smaller operating time τ_p should be operated in the first days.

Figure 5.7: Number of patients allocated for each day, and for different kinds of hospital

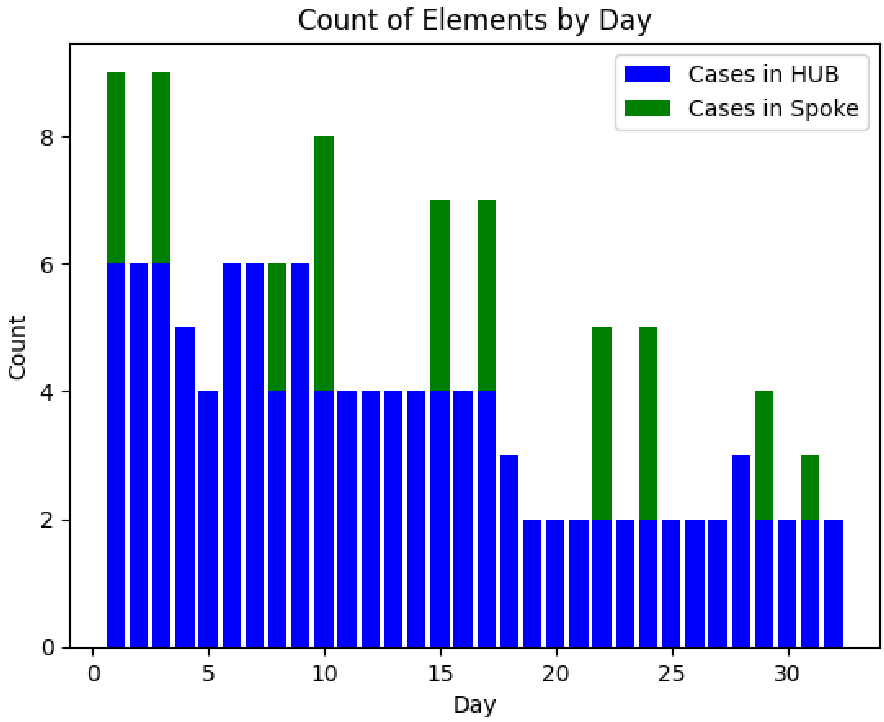
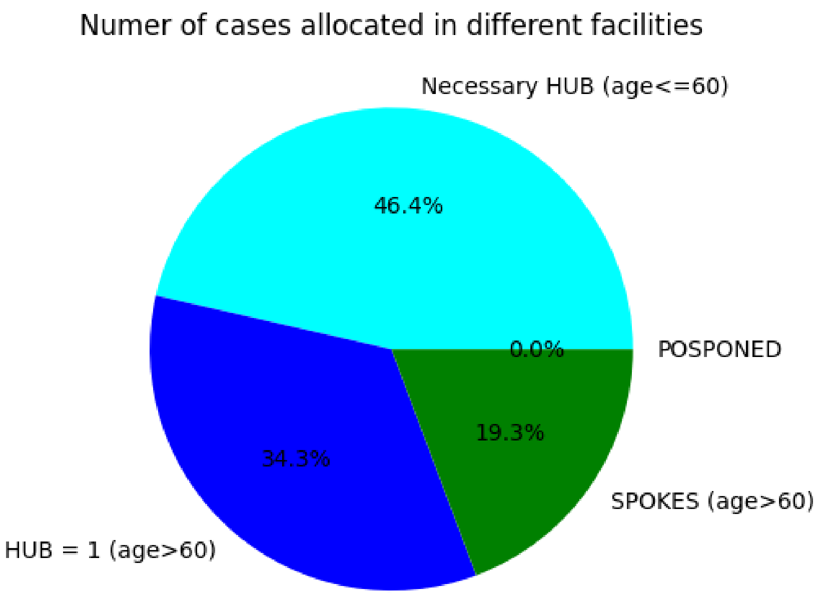


Figure 5.8: Pie diagram of patients allocation for kind of hospitals



Under time

An important KPI we mentioned, is the usage of capable resources, such as operating rooms. It could be useful to know how many hours the facility uses whenever it is open to use it. Figure 5.9 and Figure 5.10 of how many hours are used, and respectively the percentage of under-utilization, for each open bin, where we recall that the bin is composed by the OT used and the day.

Figure 5.9: Barplot of OT usage for each couple OT-day

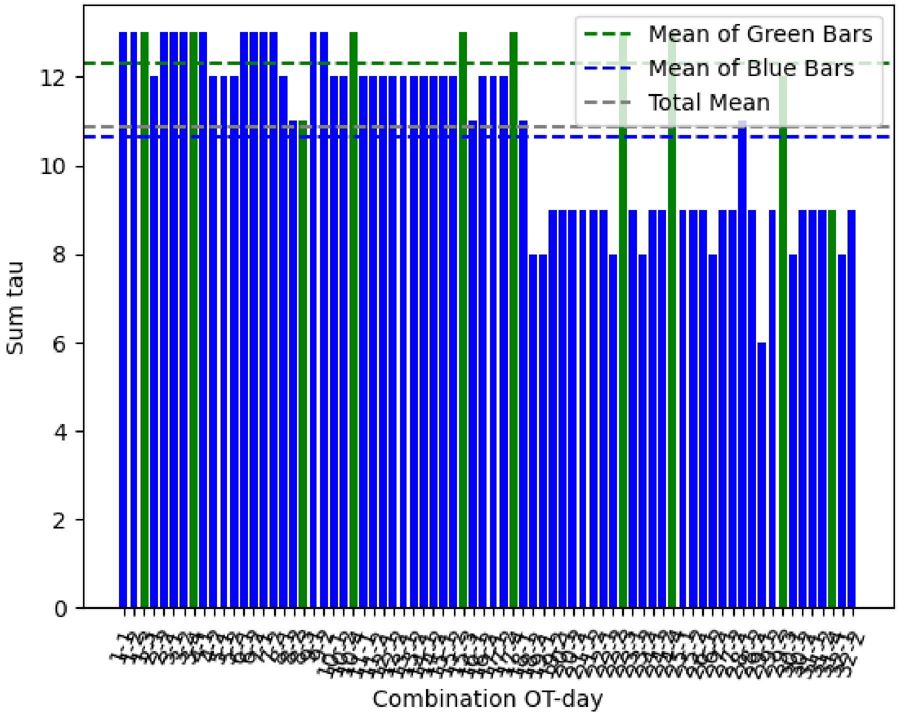
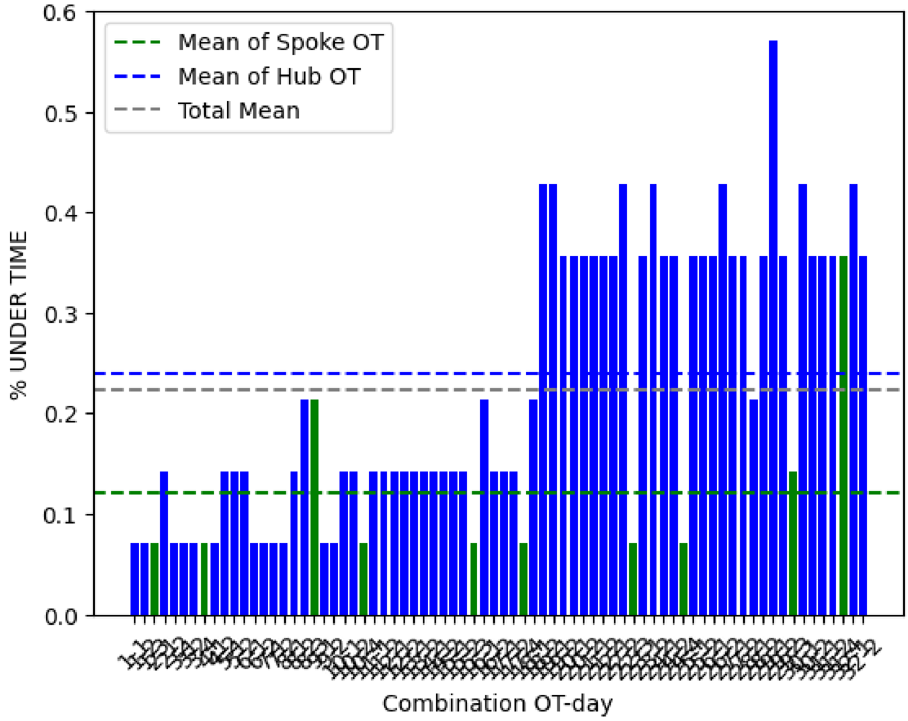
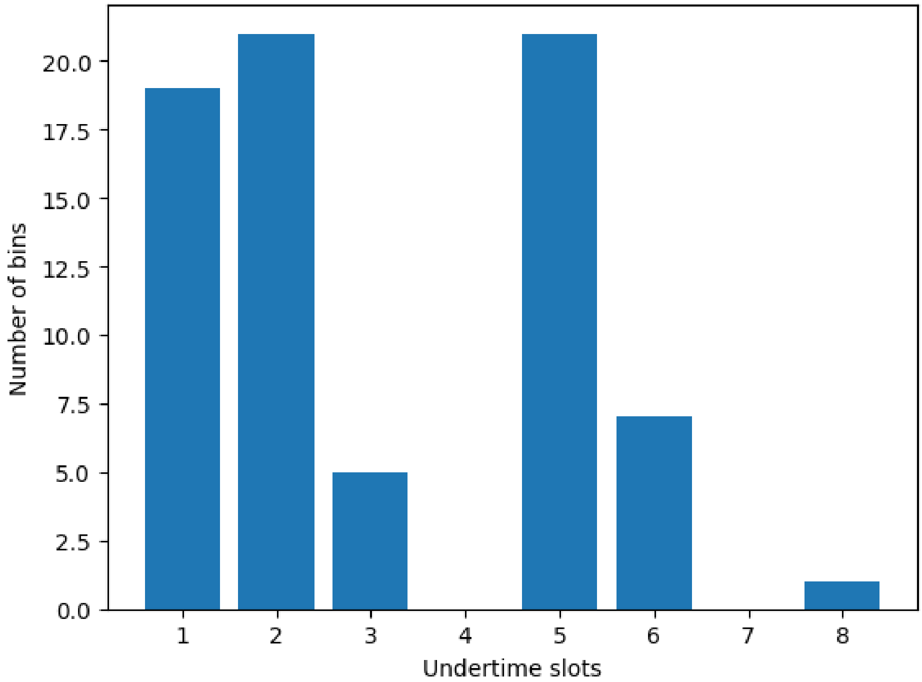


Figure 5.10: Barplot of OT undertime for each couple OT-day



To give a better overview of the previous images, Figure 5.11 shows how many used OT on different days have a specific number of undertime slots.
It is important to notice that for the kind of problem, and for the formulation we used, any patient can be allocated in the last available time slot in any operating theatre. That's because every patient has at minimum 1 slot of PACU time, so operations (surgical intervention and PACU) can not start at the last open slot. So actually 1 undertime slot is the best OT utilization we can have.

Figure 5.11: Histogram of the number of bins with specific time slots of undertime



Total number of waiting days

Figure 5.12 and 5.13 show the mean waiting time respectively before and after the schedule. It is not surprising that more or less the quantity of additional waiting days correlates directly with the duration of time a patient has been on the waiting list. In general, medical clinics tend to prioritize the assessment of a patient's urgency coefficient in determining scheduling, rather than solely considering the number of days they have been on the waiting list, which is inevitably related to the patient's urgency.

Figure 5.12: Histogram of average waiting days on waiting list

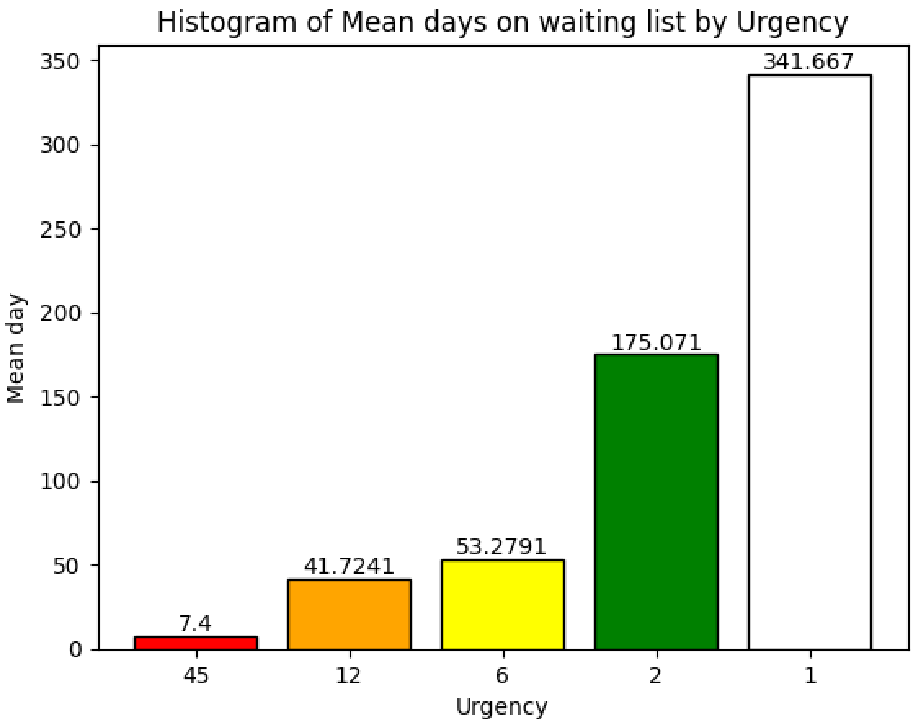
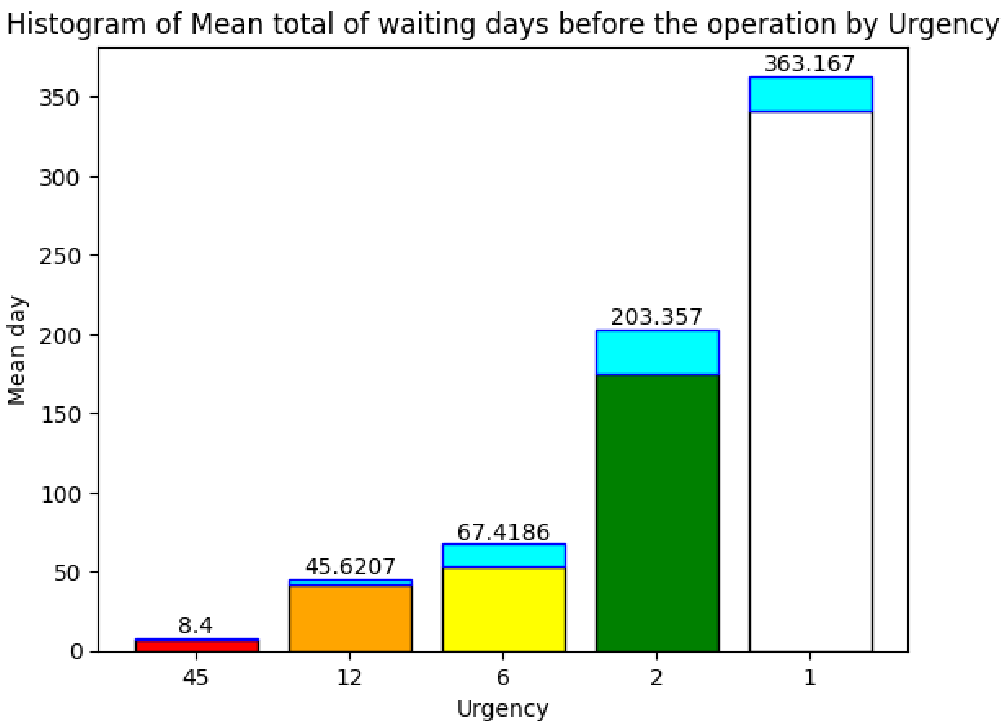


Figure 5.13: Histogram of average total waiting days



Adding extra working time slots

To follow this idea we wanted actually understand how an addition of extra working hours to the defined one could improve the quality of the solutions in terms of patient's and clinic's benefits.

The solution proposed above was basically cost-demanding since the hospital would have added 54 hours of medical staff, to the original amount of working hours. So a more realistic decision the medical center could adopt is to add an amount of extra time just once a week, to only one between the available hospitals.

So it has been decided to add a parameter $extra$ which indicates the amount of time (in terms of time slots) the hospital could give in addition to the planned hours.
It must be added a new binary variable e_{hd} which assumes value 1 if these extra time slots are used in hospital h on day d, 0 otherwise. Remember the addition of working hours to hub facilities leads to a doubling of the availability of operating theatres, since, in our case, each hub has two OTs.

This extra time could be used only once a week, in the preferred facility, which possibly, by increasing the number of working days on a specific day in a selected hospital, allows the medical network to require fewer numbers of OTs necessary to complete all the operations. For this reason, it has been added these constraints (5.1) that guarantee the usage of these extra slots at least once a week:

$$\sum_{h \in H}(ex_{hd1} + ex_{hd2}) \leq 1 \qquad \forall d1 \in D, d2 \in D \text{ s.t. } d2 > d1 \text{ and } \lfloor\frac{d1}{7}\rfloor = \lfloor\frac{d2}{7}\rfloor \tag{5.1}$$

and update the time capacity, so allow the possibility to have operations that can end before the new working hour, the ordinary ones plus those added, by replacing constraints (2.13) with new ones (5.2):

$$\begin{gathered}(t + \tau_p + PACU_p) \cdot y_{podt} \leq (|T| - 1 + e_{hd} \cdot extra) \cdot a_{hd} \\ \forall p \in P, \forall h \in Hr, \forall o \in OH_h, \forall d \in D, \forall t \in T\end{gathered} \tag{5.2}$$

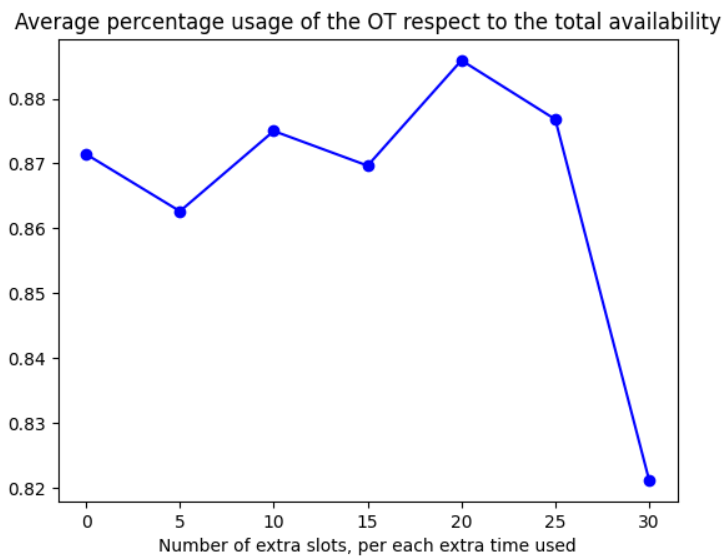
With our parameters and set, in the defined time horizon there are 5 different weeks.

We let the objective function unchanged. However, it could be added to the objective function of the formulation a factor that gives a cost to the usage of this extra time, in order to use it only if it really brings improvements to the entire schedule.

Therefore, we would like to analyze the eventual benefits of the scheduling, by letting the *extra* equal to 1,2 or more reaching 6 time slots, namely 3 hours.
Since it was difficult to analyze the benefits by comparing all the different cases, it has been decided to evaluate, as a Performance Measure, the ratio between the total number of hours where OT is used and the total available hours of OT. It is considered as available hours, all the time that an OR is open on a specific day, namely at least a patient is allocated there on that day.

It has been run the same model, with the addition of variable e_{hd} and constraints by changing the value of $extra$, from 0 to 6 time slots. Every run lasted around 1000 seconds The results presented in Figure 5.14, which represent the ratio between used hours and total open hours show that not always having more working hours per week decreases the number of OTs to open. Actually only if there is a reorganization of the schedule, by compacting the surgical operation, there could be used less Operating Rooms, and consequently, to have a lower amount of total available hours.
There is also another explanation for this "zig-zag" behavior. Since we set a time limit for each run of 1000 seconds, we did not obtain optimal solutions, but some very close to the optimum, with relmipgaps that go from 0.0092 (with $extra = 0$) to 0.0213 (with $extra = 6$). We have a relation between the increase of extra hours with the increase of the relmipgap of the run, since increasing the number of $extra$ parameters we actually admit new feasible solutions, and so there are much more nodes to explore.

Figure 5.14: Ratio between the total number of hours of OT used and the total available hours of open OT



5.5. Delays behavior

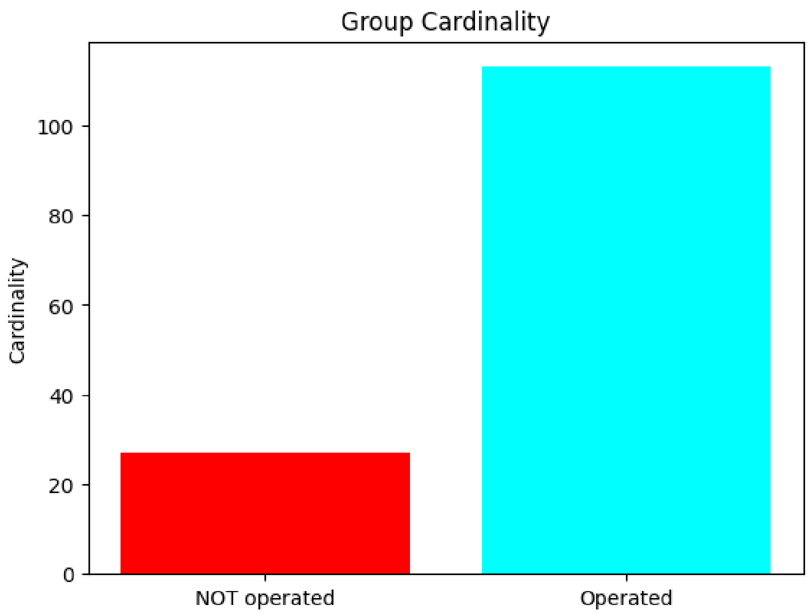
Now we want to understand how the solution treats the patients on delay.

We reduced the number of available days of the time horizon, from 32 to 20. Other sets and parameters remain unchanged. Possible extra working hours are not considered.

We run the model and find the optimal solution after 872 seconds.

In this planning horizon, the best schedule of the problem with those instances, allocates 113 patients in the time horizon, letting the remaining 27 not be scheduled, as Figure 5.15 reports.

Figure 5.15: Histogram of operations



We would want to understand the criteria of this selection, which kind of patient is more sensitive to the non-operability. We first look at the urgency factor u_p. Figure 5.16 shows the number of patients that would be operated on in the planning horizon and the number of who will have a postponed intervention.

The urgency factor gives strong weight to the decision of operability, but it is not so crucial for the decision of postponing some operations. In other words, an high value of u_p heavily brings the schedule to operate the patient p in the planning time, but a low value of u_p does not imply the schedule almost surely would postpone the patient's intervention.

Figure 5.16: Barplot of operations

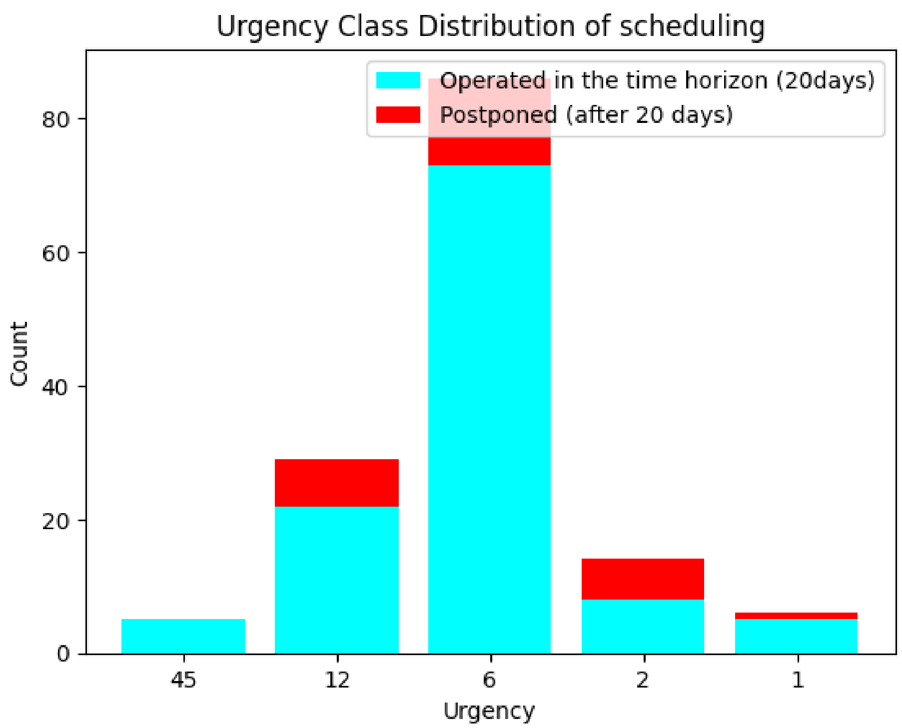


Figure 5.17: Urgency Class distribution Schedule

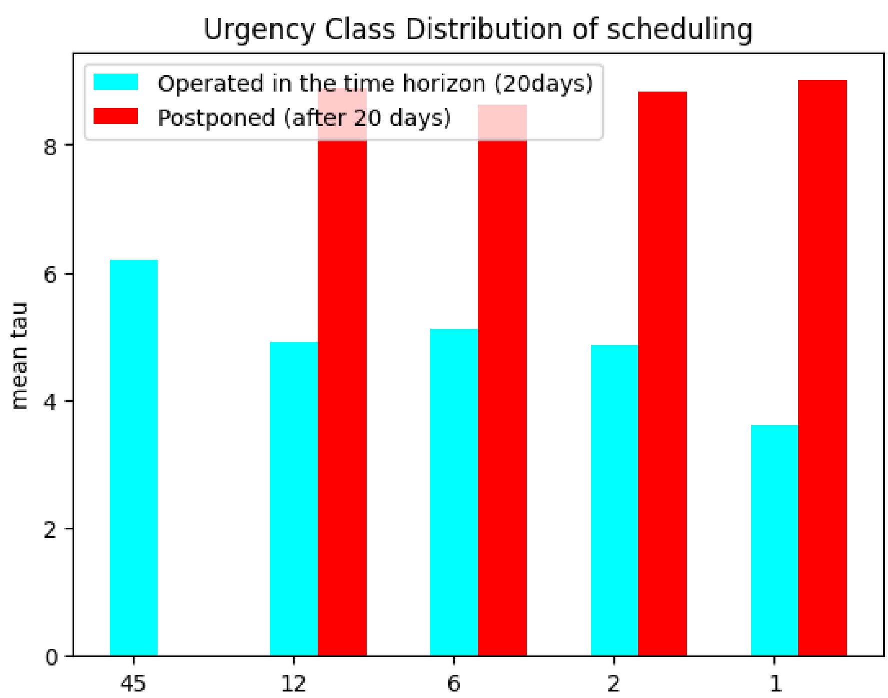
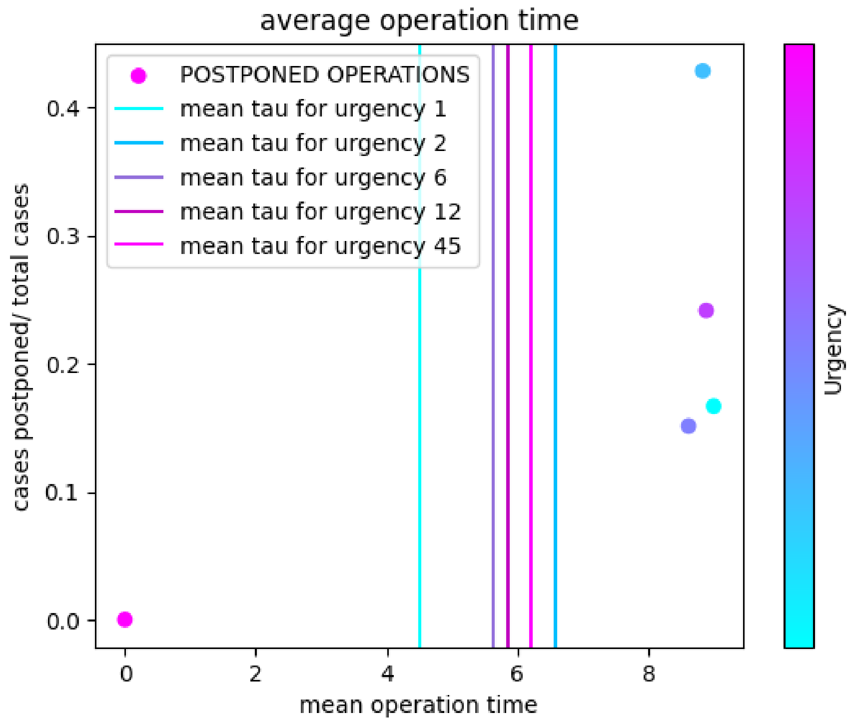


Figure 5.18: ratio of usage of OT on Urgency Class and mean operation time



So a possible explanation could be the length τ of the intervention. In fact, in 14 working slots, we could allocate a few prolonged operations or several rapid interventions.

So we want to analyze if there is an aim to operate a more number of patients, postponing the longer ones. As Figure 5.17 represents, the average operation time for not scheduled patients, except for patients with $u_p = 45$, is between 3 and 4 time slots (90 and 120 minutes).

In Table 5.18 we want to show also the ratio between postponed and total cases, considering the two factors discussed above: urgency factor and average operation time. With the kind of color, we highlight the urgency, from light blue (less urgent) to fuchsia (more urgent). On the x-axes, we show the average intervention time only of the postponed operations, while in the y-label we are representing the ratio between postponed operations over the total surgical cases. The vertical lines instead represent the mean τ for every kind of urgency. As we can notice, all the points except the fuchsia one, lay on the right part of the lines, stressing the fact the schedule set up prefers to operate on more patients, than schedule them on the basis of the urgency factor.

6 | Conclusions Future Developments

We have built a complete model, $SPHS$, capable of solving the problem of finding an optimal day-hospital pediatric surgical operation schedule, for patients and surgeons, in a specific time horizon, taking into account the minimization of patients' waiting times, and the avoidance of delays, both weighted for the urgency coefficient of each patient, by also making the most of the Spokes network.
The schedule should respect several constraints, regarding the capacity and availability of the OT for the operations and of beds for the recovery time, the compatibility of patient-hospital, and of patient-surgeon, and the ordering of interventions in a specific working day.

The problem has approached with Integer Linear Programming, by testing on real instances of surgical cases. Due to the complexity of the problem and its onerousness, this problem could not rely on an analytical solution. For this reason, it has been proposed different variants of the problem, by making enrichment of the model or by relaxing it, removing constraints or even variables, to trying some heuristic approaches.

Then, in primis, we recognized which formulation brought evident improvements in terms of convergence speed, and then analyzed primarily the feasibility of the solution with respect to the original problem, and then if it is so, or adjusted to make it feasible, examined the quality of the solution.

An important result to notice is that formulating a model for a scheduling problem, in a Hub & Spoke context, could bring relevant improvements if a Bin-Packing approach is considered in the formulation. In fact, each operating room, on each day, can be considered as a bin, that can be opened, and where we can allocate patients in it if so.

The second main approach which has been found very efficient in terms of computational speed and quality of the solution is a matheuristic approach composed of two steps: the first one allows finding the optimal schedule of the patients, by not taking into considera-

tion the surgeons, while the second step proposes the optimal schedule of the doctors by associating them to the operation schedule found in the first phase.
The first step does not include any variables and constraints related to surgeons, so this new problem $SPHS\_nodoc$, which has the same objective function as $SPHS$ is a strong relaxation with respect to the original problem.
Once the best, or almost optimal, surgical operations schedule has been found, in less than 15 minutes, it would be run a second model, which we called $model\_SA$, which recalls a Facility Location Problem. In fact, we want to satisfy all patients' requests, namely surgeons, by minimizing the number of them. So the new model $model\_SA$ is composed essentially of the variables and constraints originally removed from $SPHS$ to build a relaxed model which did not include the models and as an objective function the minimization of the sum for each day and for each facility, the total number of surgeon involved. And in just a few seconds finds the optimal solution.

Future Developments

A first direction for future developments in the study of improved formulation or heuristics is to enhance the efficiency and reduce the computational time so as to the growth of the size of the instances that can be treated. For instance, finer time slots, larger patient populations, longer planning horizons, and a more numerous net of hospitals could be tackled and solved.
To address this, innovative techniques such as heuristic, meta-heuristic, or matheuristic algorithms could be investigated.

Another important direction for future development lies in considering uncertain operating and PACU times. In healthcare settings, it is very common for surgeries or patient recovery times to deviate from the initially estimated durations. Incorporating uncertainty into scheduling models can allow for more robust and flexible scheduling decisions. Probabilistic approaches, such as stochastic programming or robust optimization, can be explored to account for the uncertainty and mitigate the potential negative impacts of schedule disruptions. By incorporating uncertainty into the scheduling process, healthcare facilities can make more informed decisions, improve resource allocation, and enhance operational efficiency in dynamic and unpredictable environments.
To optimize schedules that consider the range of possible outcomes, which take into account possible delays, can be used different methods to approach this modeling, such as a Discrete Simulation Event, Stochastic Programming, Simulation-Based Approaches,

Robust Optimization, or even Hybrid Approaches can be studied.
These are some of the methods which could provide avenues for scheduling operations with variable operation times, considering the stochastic nature of such events.

Furthermore, the extension of scheduling models to accommodate longer time horizons and more hospitals is crucial for scalability and practical applicability. In this Thesis, it has been considered a monthly planning horizon. But healthcare systems often operate on schedules spanning weeks or even months, involving coordination across multiple hospitals or healthcare facilities. Future research can focus on developing scheduling frameworks that handle reduced or extended time frames while considering interdependencies among hospitals. Such models could incorporate factors like resource sharing and coordination of specialized equipment or staff to enable efficient scheduling at a broader system level.

To address these problems new approaches that leverage data-driven methodologies can be explored. Machine learning techniques, such as predictive analytics and classification algorithms, can help analyze historical patient data to identify patterns, optimize scheduling decisions, predict delays in operation times, and provide personalized care. Additionally, simulation models can be utilized to assess the impact of different scheduling strategies on various performance metrics, enabling healthcare practitioners to make informed decisions and test scenarios in a virtual environment before implementation.

Overall, these advancements in healthcare scheduling systems will hold the potential to enhance the efficiency, flexibility, and adaptability of healthcare scheduling, ultimately leading to improved patient care, reduced waiting times, optimized resource utilization, and better overall healthcare system performance.

Bibliography

[1] 2011 milliman medical index. *Milliman: Tech. Rep.*, 2011.

[2] Z. Abdalkareem, A. Amir, Mohammed, M. Al-Betar, P. Ekhan, Abdelaziz, and A. Hammouri. Healthcare scheduling in optimization context: A review. *Health and Technology*, 04 2021. doi: 10.1007/s12553-021-00547-5.

[3] B. Addis, G. Carello, and E. Tanfani. A robust optimization approach for the operating room planning problem with uncertain surgery duration. *Springer Proceedings in Mathematics and Statistics*, 61, 05 2013. doi: 10.1007/978-3-319-01848-5_14.

[4] B. Addis, G. Carello, A. Grosso, and E. Tanfani. Operating room scheduling and rescheduling: a rolling horizon approach. *Flexible Services and Manufacturing Journal*, 28, 01 2015. doi: 10.1007/s10696-015-9213-7.

[5] A. Agnetis, A. Coppi, M. Corsini, G. Dellino, C. Meloni, and M. Pranzo. Long term evaluation of operating theater planning policies. *Operations Research for Health Care*, 1:2012, 12 2012. doi: 10.1016/j.orhc.2012.10.001.

[6] A. Agnetis, A. Coppi, M. Corsini, G. Dellino, C. Meloni, and M. Pranzo. A decomposition approach for the combined master surgical schedule and surgical case assignment problems. *Health care management science*, 06 2013. doi: 10.1007/s10729-013-9244-0.

[7] D. Antonelli and T. Taurino. Application of a patient flow model to a surgery department. pages 1 – 6, 03 2010. doi: 10.1109/WHCM.2010.5441247.

[8] V. Augusto, X. Xie, and V. Perdomo. Operating theatre scheduling with patient recovery in both operating rooms and recovery beds. *Computers Industrial Engineering / Computers and Industrial Engineering*, 58, 03 2010. doi: 10.1016/j.cie.2009.04.019.

[9] S. Azari-Rad, A. Yontef, D. Aleman, and D. Urbach. Reducing elective general surgery cancellations at a canadian hospital. *Canadian journal of surgery. Journal canadien de chirurgie*, 56:018411–18411, 02 2013. doi: 10.1503/cjs.018411.

[10] C. Banditori, P. Cappanera, and F. Visintin. A combined optimization–simulation

approach to the master surgical scheduling problem. *IMA Journal of Management Mathematics*, 24:155–187, 03 2013. doi: 10.1093/imaman/dps033.

[11] C. Banditori, P. Cappanera, and F. Visintin. Investigating the relationship between resources balancing and robustness in master surgical scheduling. *SPRINGER PROCEEDINGS IN MATHEMATICS STATISTICS*, 61:149–162, 01 2014. doi: 10.1007/978-3-319-01848-5_12.

[12] S. Batun, B. Denton, T. Huschka, and A. Schaefer. Operating room pooling and parallel surgery processing under uncertainty. *INFORMS Journal on Computing*, 23: 220–237, 05 2011. doi: 10.1287/ijoc.1100.0396.

[13] A. Baumgart, A. Zoeller, C. Denz, H.-J. Bender, A. Heinzl, and E. Badreddin. Using computer simulation in operating room management: Impacts on process engineering and performance. pages 131 – 131, 02 2007. doi: 10.1109/HICSS.2007.593.

[14] B. Berg, B. Denton, A. Erdogan, T. Rohleder, and T. Huschka. Optimal booking and scheduling in outpatient procedure centers. *Computers Operations Research*, 50, 10 2014. doi: 10.1016/j.cor.2014.04.007.

[15] M. Calichman. Creating an optimal operating room schedule. *AORN journal*, 81: 580–8, 04 2005. doi: 10.1016/S0001-2092(06)60442-0.

[16] B. Cardoen, E. Demeulemeester, and J. Beliën. Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201: 921–932, 03 2010. doi: 10.1016/j.ejor.2009.04.011.

[17] A. Chieregato, G. Veronese, F. Curto, M. Zaniboni, F. Fossi, F. Zumbo, C. Scattolini, C. Compagnone, B. Alberti, c. a. Baciu, S. Carenini, C. Cipolla, A. Formentano, A. Guidi, F. Massimo, I. Galluccio, S. Pagani, R. Paparone, and R. Fumagalli. Emergently planned exclusive hub-and-spoke system in the epicenter of the first wave of covid-19 pandemic in italy: the experience of the largest covid-19-free icu hub for time-dependent diseases. *Minerva anestesiologica*, 87, 06 2021. doi: 10.23736/S0375-9393.21.15455-0.

[18] R. Day, R. Garfinkel, and S. Thompson. Integrated block sharing: A win–win strategy for hospitals and surgeons. *Manufacturing Service Operations Management*, 14:567–583, 10 2012. doi: 10.1287/msom.1110.0372.

[19] B. Denton and D. Gupta. A sequential bounding approach for optimal appointment scheduling. *IIE Transactions*, 35:1003–1016, 11 2003. doi: 10.1080/07408170304395.

[20] B. Denton, J. Viapiano, and A. Vogl. Optimization of surgery sequencing and schedul-

ing decisions under uncertainty. *Health care management science*, 10:13–24, 03 2007. doi: 10.1007/s10729-006-9005-4.

[21] V. T. M. B. V. J. P. S. B. S. . V. d. H. J. Does, R. Reducing start time delays in operating rooms. *Journal of Quality Technology*, 41:95–109, 2009.

[22] J. Ewing. Revenue growth and cash flow margins hit all-time lows in 2013 us not-for-profit hospital medians. *Tech: Rep., Moody's*, 2014.

[23] C. Geppert, K. Berkowitz, T. Schonfeld, and A. Tarzian. Covid-19 ethics debrief: Pearls and pitfalls of a hub and spoke model. *The Journal of clinical ethics*, 33:63–68, 04 2022. doi: 10.1086/JCE2022331063.

[24] G. Guerriero. Operational research in the management of the operating theatre: a survey. *Health Care Manag Sci*, pages 89–115, 2011.

[25] S. Harris and D. Claudio. Current trends in operating room scheduling 2015 to 2020: a literature review. *Operations Research Forum*, 3, 03 2022. doi: 10.1007/s43069-022-00134-y.

[26] R. Kasivisvanathan, H. Tilney, S. Jhanji, M. O'Mahony, P. Gruber, D. Nicol, D. Morgan, E. Kipps, and S. Rasheed. The 'hub and spoke model' for the management of surgical patients during the covid-19 pandemic. *The International Journal of Health Planning and Management*, 36, 05 2021. doi: 10.1002/hpm.3243.

[27] M. O'Kelly and Y. Lao. Mode choice in a hub-and-spoke network: A zero-one linear programming approach. *Geographical Analysis*, 23:283 – 297, 09 2010. doi: 10.1111/j.1538-4632.1991.tb00240.x.

[28] T. C. . C. L. Olivares, M. Structural estimation of the newsvendor model: An application to reserving operating room time. *Management Science*, 54:41–55, 05 2008.

[29] M. Samudra, C. Riet, E. Demeulemeester, B. Cardoen, N. Vansteenkiste, and F. Rademakers. Scheduling operating rooms: achievements, challenges and pitfalls. *Journal of Scheduling*, 19, 10 2016. doi: 10.1007/s10951-016-0489-6.

[30] . L. D. L. Sieber, T. J. Operating room management and strategies in switzerland: Results of a survey. *European Journal of Anaesthesiology*, 19:415–423, 2002.

[31] P. F. . H. M. Steins, K. Increasing utilization in a hospital operating department using simulation modeling. page 463–480, 2010.

[32] R. B. C. J. P. . R. F. Tancrez, J. S. How stochasticity and emergencies disrupt the surgical schedule. *Berlin: SpringerManagement Science*, 54:221–239, 2009.

[33] H. J. L. H. W. . v. d. A. B. J. van Essen, J. T. Decision support system for the operating room rescheduling problem. *Health Care Management Science*, 15:355–372, 2012.

[34] . D. F. Wachtel, R. E. A simple method for deciding when patients should be ready on the day of surgery without procedure-specific data. *Anesthesia and Analgesia*, 105:127–140, 2007.

List of Figures

List of Tables

Acknowledgements

I wanted to thank Mum and Dad who have always been there for me even from the beginning of my academic career, not letting me lack anything, especially their support, which was indispensable to me.
Then thanks to Maria Chiara and Francesco for always being available and helpful with anything for me.
Then I want to thank all components of the family, cousins, uncles and aunts, and grandparents, even those who are not here anymore. Thanks for telling me the right words at the right moment.

I want to thank also lifelong friends, Giulio, Andrea, Giovanni, Enrico, Silvia, and Martina. I want to thank also for important friendships born recently, like Adriana, Michele, Agata e Sofia.

Then thanks to all the people and friends with whom I went through this fantastic and not-without-hardship, academic journey. Firstly with my very first colleagues Rick, Daniel, Dario, Leonardo, Edoardo, Stefano, and Martina, with whom we have helped each other since our first exams.

And then also people I've met in various circumstances. They are too many, which have entered in my life more or less overwhelmingly. Just to say a few names Jacoco, Jack, Gigi, Franci, Sara, Chiara, Pietro, Pij, Pippo, Giovanni, Riccardo, Marco.
I want to thank also all the neighborhoods, Polifemo group, the theatre company, and the choir.

I would also like to thank Prof. Pelizzo and Dr. Costanzo of the Buzzi Hospital for the definition and information on the problem analyzed. Lastly, I would like to thank Professor Carello for making me passionate about the subjects she teaches, and then for being present during these months, and for teaching me so much.